



# **ESCUELA SUPERIOR POLITECNICA DE CHIMBORAZO**

**FACULTAD DE CIENCIAS**

**ESCUELA DE FÍSICA Y MATEMÁTICA**

**CARRERA DE INGENIERIA EN ESTADÍSTICA INFORMÁTICA**

**“Análisis Estadístico Multivariante de los Principales Componentes Químicos sobre Hachas Moneda de Cobre para determinar diferentes grupos de acuerdo a su composición.”**

## **TESIS DE GRADO**

Previa la obtención del Título de

**Ingeniera en Estadística Informática**

Presentado por:

**María Angélica Villa Moyota.**

Riobamba – Ecuador

- 2012 -

## **AGRADECIMIENTO**

Quiero agradecer infinitamente a Dios por haber guiado mi camino hasta ahora, a mis Padres y hermanos por siempre haberme brindado su fuerza y apoyo incondicional para llegar hasta donde estoy ahora.

Además un infinito gracias a mis Profesores en especial a mi director de tesis quien con sus conocimientos y experiencia me apoyo para la realización de este tema final de tesis, también a la Empresa Ecuacorriente S. A ya que gracias a su apoyo y apertura a la investigación fue posible realizar este estudio y finalmente a mis amigas y compañeras de carrera por su amistad y apoyo dentro de lo académico y personal.

## **DEDICATORIA**

La concepción de este trabajo investigativo está dedicada a mis Padres, pilares fundamentales en mi vida, sin ellos jamás hubiera podido conseguir lo que hasta ahora. Su tenacidad y lucha insaciable han hecho de ellos el gran ejemplo a seguir y destacar, no solo para mí, sino para mis hermanos y familia en general.

También quiero dedicar este trabajo a dos personas que han marcado mi vida, a Ud. Papito Jorge que con su partida me dejó un inmenso vacío y aunque ya no está conmigo sus enseñanzas y el amor por la vida y la familia, son las virtudes que día a día trataré de seguir y finalmente al amor de mi vida Matías por ser el hombrecito que le trajo luz y alegría a nuestra familia.

NOMBRE	FIRMA	FECHA
Dra. Yolanda Díaz <b>DECANA FACULTAD DE CIENCIAS</b>	<hr/>	<hr/>
Dr. Richard Pachacama <b>DIRECTOR ESCUELA DE FÍSICA Y MATEMÁTICA</b>	<hr/>	<hr/>
Dr. Luis Vera <b>DIRECTOR DE TESIS</b>	<hr/>	<hr/>
Mat. Marcelo Cortez <b>MIEMBRO DEL TRIBUNAL</b>	<hr/>	<hr/>
<b>MIEMBRO DEL TRIBUNAL</b>	<hr/>	<hr/>
<b>DIRECTOR CENTRO DE DOCUMENTACIÓN</b>	<hr/>	<hr/>

Yo, María Angélica Villa Moyota soy la responsable de las ideas, doctrinas y resultados expuestos en este trabajo de investigación; y el patrimonio intelectual de la tesis de grado pertenece a la ESCUELA SUPERIOR POLITECNICA DE CHIMBORAZO.

---

María Angélica Villa Moyota

## INDICE GENERAL

AGRADECIMIENTO .....	- 2 -
DEDICATORIA .....	- 3 -
INDICE GENERAL.....	- 6 -
INDICE DE FIGURAS .....	- 12 -
INDICE DE TABLAS .....	- 13 -
INDICE DE GRÁFICOS .....	- 15 -
INDICE DE FOTOS .....	- 17 -
INTRODUCCION.....	- 18 -
ANTECEDENTES .....	- 20 -
JUSTIFICACION .....	- 24 -
OBJETIVOS.....	- 25 -
OBJETIVO GENERAL .....	- 25 -
OBJETIVOS ESPECIFICOS.....	- 25 -
HIPOTESIS.....	- 26 -

I. CAPITULO .....	- 27 -
MARCO TEORICO.....	- 27 -
1.1. ARQUEOLOGIA CUANTITATIVA: LA FUNCION DEL ANALISIS MULTIVARIANTE..	- 27 -
1.2. ANALISIS MULTIVARIADO.....	- 29 -
1.2.1. Descripción de Datos Multivariantes .....	- 30 -
1.2.1.1. Tipos de Variables. ....	- 30 -
1.2.1.2. Matriz de Datos.....	- 32 -
1.2.2. Análisis Exploratorio de Datos. ....	- 32 -
1.2.2.1. Análisis Univariante.....	- 33 -
1.2.2.2. Medidas de Centralización – Vector de Medias .....	- 35 -
1.2.2.3. Matriz de Varianzas y Covarianzas.....	- 36 -
1.2.2.4. Variables Redundantes.....	- 38 -
1.2.2.5. Medidas Globales de Variabilidad .....	- 38 -
1.2.2.6. Medidas de Dependencia Lineal.....	- 41 -

1.2.2.7.	Análisis de Datos Atípicos.....	- 44 -
1.2.3.	Componentes Principales.....	- 52 -
1.2.3.1.	Calculo del Primer Componente.....	- 53 -
1.2.3.2.	Selección del número de Componentes. ....	- 55 -
1.2.4.	Análisis de Conglomerados.....	- 56 -
1.2.4.1.	Algoritmos Jerárquicos.....	- 58 -
1.2.4.2.	Conglomerados por variables.....	- 64 -
1.2.5.	Análisis Factorial. ....	- 64 -
1.2.5.1.	El Modelo Factorial. ....	- 65 -
1.2.5.2.	Número Máximo de Factores.....	- 67 -
1.2.5.3.	El método del Factor Principal.....	- 67 -
1.2.5.4.	Estimación Máximo Verosímil. ....	- 70 -
II.	CAPITULO.....	- 72 -
	TECNICAS EXPERIMENTALES.....	- 72 -
2.1	PROCESO DE RESTAURACION.....	- 72 -



2.1.1	Deterioro de Piezas Metálicas. ....	- 73 -
2.1.2	<i>Productos de la corrosión.</i> .....	- 74 -
2.1.3	Proceso de Restauración aplicado en Hachas Moneda .....	- 75 -
2.1.3.1	Tratamiento previo de la pieza.....	- 76 -
2.1.3.2	Limpieza de Objetos.....	- 77 -
2.1.3.3	Consolidación de la pieza. ....	- 79 -
2.2	APLICACIÓN DE LA MÁQUINA THERMO NITON .....	- 80 -
2.2.1	Máquina Thermo Niton XL3 XRF.....	- 80 -
2.2.2	Análisis de las Hachas Moneda.....	- 81 -
2.3	OBTENCION DE DATOS.....	- 82 -
III.	CAPITULO .....	- 83 -
	APLICACIÓN DE LAS TECNICAS DE ANALISIS ESTADISTICO MULTIVARIADO.....	- 83 -
3.1	DATOS.....	- 83 -
3.1.1	Hachas Moneda (Individuos).....	- 83 -
3.1.2	Elementos Químicos (Variables).....	- 84 -

3.1.3	Matriz de Datos.....	- 89 -
3.2	ANALISIS MULTIVARIADO.....	- 101 -
3.2.1	Análisis Exploratorio de Datos (AED). ....	- 101 -
3.2.1.1	Normalidad. ....	- 101 -
3.2.1.2	Análisis Univariante.....	- 106 -
3.2.1.3	El vector de medias. ....	- 108 -
3.2.1.4	Matriz de Varianzas y Covarianzas.....	- 108 -
3.2.1.5	Variables Redundantes.....	- 109 -
3.2.1.6	Datos atípicos.....	- 116 -
3.2.1.7	Medidas Globales de Variabilidad. ....	- 121 -
3.2.1.8	Medidas de Dependencia Lineal.....	- 122 -
3.2.2	Componentes Principales. ....	- 129 -
3.2.3	Análisis de Conglomerados.....	- 139 -
3.2.3.1	Conglomerados por observaciones. ....	- 140 -
3.2.3.2	Conglomerados por Variables. ....	- 145 -

3.2.4	Análisis Factorial. ....	- 147 -
IV.	CAPITULO .....	- 156 -
	CONCLUSIONES Y RECOMENDACIONES .....	- 156 -
4.1	CONCLUSIONES .....	- 156 -
4.2	RECOMENDACIONES .....	- 158 -
	RESUMEN .....	- 160 -
	SUMARY .....	<b>¡Error! Marcador no definido.</b>
	BIBLIOGRAFIA.....	<b>¡Error! Marcador no definido.</b>
	ANEXOS.....	- 165 -
	ACUERDO DE COOPERACION ACADEMICO CIENTIFICO .....	- 166 -
	RESOLUCION PARA LA SUBSCRIPCION DEL ACUERDO DE COOPERACION.....	- 173 -
	FOTOS DE LA INVESTIFACION. ....	- 175 -
	INSTALACIONES DE ECSA. ....	- 175 -
	PROCESO DE RESTAURACION Y TOMA DE DATOS. ....	- 178 -

## INDICE DE FIGURAS

Figura 1: Mapa de Ecuador con sitios arqueológicos.....	21
Figura 2: Mapa de América destacando las zonas de comercialización entre Ecuador y México.....	23
Figura 3: Ejemplo de Matriz de dispersión.....	45
Figura 4: Ejemplo de Diagrama de Caja.....	46
Figura 5: Ejemplo de Aplicación de las Caras de Chernoff.....	47
Figura 6: Hacha moneda corroída.....	74
Figura 7: Frasco de Benzotriazol.....	79
Figura 8: Máquina Thermo Niton XL3 XRF.....	80
Figura 9: Datos Obtenidos con la Thermo Niton.....	82

## INDICE DE TABLAS

Tabla I: Matriz de Datos Inicial.....	90
Tabla II: Prueba de Kolmogorov Smirnov.....	102
Tabla III: Análisis Univariante Descriptivo.....	106
Tabla IV: Análisis Univariante Descriptivo Robusto.....	107
Tabla V: Matriz de Varianzas y Covarianzas Inicial.....	109
Tabla VI: Autovalores.....	110
Tabla VII: Matriz de Datos sin Variables Redundante.....	111
Tabla VIII: Análisis Univariante Descriptivo Final.....	120
Tabla IX: Matriz de Varianzas y Covarianzas Final.....	121
Tabla X: Medidas Globales de Variabilidad.....	121
Tabla XI: Matriz de Correlación.....	122
Tabla XII: Coeficiente de Determinación.....	124
Tabla XIII: Porcentajes para las Estimaciones.....	126
Tabla XIV: Estimaciones para los Modelos Ajustados Aceptables.....	126

Tabla XV: Intervalos de Confianza para los Coeficientes de Regresión del modelo.....	127
Tabla XVI: Varianza Total Explicada, selección de Componentes.....	130
Tabla XVII: Primeras dos Componentes Principales.....	130
Tabla XVIII: Ordenación de las Hachas Moneda de Cobre según las dos Primeras Componentes.....	133
Tabla XIX: Distribución de las observaciones en los Conglomerados.....	142
Tabla XX: Observaciones más cercanas.....	143
Tabla XXI: Distribución de las Variables en los Conglomerados.....	146
Tabla XXII: Comunalidades.....	147
Tabla XXIII: Varianza Total Explicada, selección de Factores.....	149
Tabla XXIV: Matriz de Correlaciones.....	151
Tabla XXV: Prueba KMO y Prueba de Bartlett.....	152
Tabla XXVI: Matriz de Factores.....	153

## INDICE DE GRÁFICOS

Gráfico 1: Histogramas con curva Normal de Sb, Sn, Ag, Bal, Bi, As, Se y Au.....	104
Gráfico 2: Histogramas con curva Normal de Pb, Cu, Ni, Fe, Ti, Ca y K.....	105
Gráfico 3: Diagramas de Caja.....	116
Gráfico 4: Matriz de Dispersión.....	117
Gráfico 5: Caras de Chernoff.....	118
Gráfico 6: Diagramas de Dispersión de las Correlaciones relativas altas.....	123
Gráfico 7: Diagramas de Dispersión de las Correlaciones relativas bajas.....	123
Gráfico 8: Gráfico de Sedimentación de Componentes Principales.....	131
Gráfico 9: Representación de los pesos de las dos primeras componentes.....	132
Gráfico 10: Representación de las observaciones sobre las dos Primeras Componentes.....	139
Gráfico 11: Dendrograma por Observaciones.....	140
Gráfico 12: Análisis de Conglomerados por Observaciones en Minitab.....	141
Gráfico 13: Dendrograma por Variables.....	145

Gráfico 14: Análisis de Conglomerados por Variables en Minitab.....	146
Gráfico 15: Gráfico de Sedimentación para Análisis Factorial.....	150
Gráfico 16: Representación de los Elementos Químicos componentes de Hachas Moneda de Cobre en los tres primeros factores.....	155



## INDICE DE FOTOS

Foto 1: Limpieza con bisturí de materias adherida por la corrosión.....	77
Foto 2: Aplicación de HCL al 10%.....	78
Foto 3: Muestra Pulida para retirar el oxido.....	78
Foto 4: Hachas con aplicación de Benzotriazol.....	79
Foto 5: Análisis con la Thermo Niton.....	81
Foto 6: Instalaciones de ECSA (1).....	174
Foto 7: Instalaciones de ECSA (2).....	175
Foto 8: Instalaciones de ECSA (3).....	175
Foto 9: Personal del Campamento El Mirador.....	176
Foto 10: Tesista.....	176
Foto 11: Primer contacto con las Hachas Moneda.....	177
Foto 12: Máquina Thermo Niton.....	178

## INTRODUCCION

Dentro de la historia pre-colonial del Ecuador ha existido el hallazgo de diferentes piezas arqueológicas que de alguna manera nos han mostrado el desarrollo de las culturas que habitaron y se desarrollaron dentro del territorio ecuatoriano, dentro de estos hallazgos se encuentran las **“hachas moneda”**, artefactos que usualmente servían de intercambio comercial entre varias culturas de la costa norte del Perú, Ecuador y México desde el período del Horizonte Medio hasta la llegada de los españoles.

Dentro de esta investigación la Estadística y la Arqueología avanzaran juntas ya que en esta ocasión el estudio arqueológico de la composición química de las hachas moneda de cobre necesita de un gran e importante aporte de la aplicación de herramientas estadísticas específicamente dentro del campo del Análisis Multivariado que nos permitirá identificar los diferentes grupos de hachas halladas en un lugar específico de la costa ecuatoriana, esto nos conducirá de alguna manera a inferir la procedencia de estos artefactos arqueológicos.

De acuerdo a indagaciones previas a la realización de este trabajo investigativo no se han encontrado estudios arqueológicos - estadísticos que realicen un estudio similar al que se aplica dentro de este trabajo, por lo que lo convierte en pionero dentro de este ámbito, además es importante señalar que la Estadística especialmente el Análisis Multivariado, dentro de la arqueología se ha convertido y se sigue desarrollando como

una potente herramienta para determinar el desarrollo de la historia no solo nacional sino que a demás mundial.

Para entender y envolvernos en este tema el presente proyecto de tesis estará conformado por una breve introducción a los antecedentes de la Hachas moneda, su origen, fabricación y posible distribución dentro de ciertas culturas de la costa ecuatoriana.

El primer capítulo presenta el estudio formal del Análisis Estadístico Multivariado, describe los supuestos que deben cumplirse para la aplicación de las técnicas multivariantes además se describen teóricamente dichas técnicas, que son aplicables para el desarrollo de nuestro estudio.

El capítulo dos mostrara las técnicas experimentales que fueron utilizadas para la toma de datos de los componentes químicos de las hachas moneda con la máquina Thermo Niton y el proceso de restauración aplicado a las hachas previo a su análisis.

El capítulo tres presenta la aplicación de las técnicas multivariantes a la base de datos obtenida dentro de la fase experimental.

Finalmente el capítulo cuatro trata de las conclusiones obtenidas posterior al estudio investigativo, que tendría que ver con la evaluación de los resultados obtenidos, además las debidas recomendaciones de acuerdo a los objetivos planteados y pretendidos al iniciar este estudio investigativo.

## **ANTECEDENTES**

La historia del Ecuador es un conjunto de sucesos en el tiempo, en el cual los territorios que actualmente pertenecen a la República del Ecuador cambian en su aspecto físico y forma de gobierno. La historia de la República puede dividirse en cuatro etapas: Etapa Prehispánica, Etapa Hispánica (Conquista, colonización y Colonia), Independencia y República.

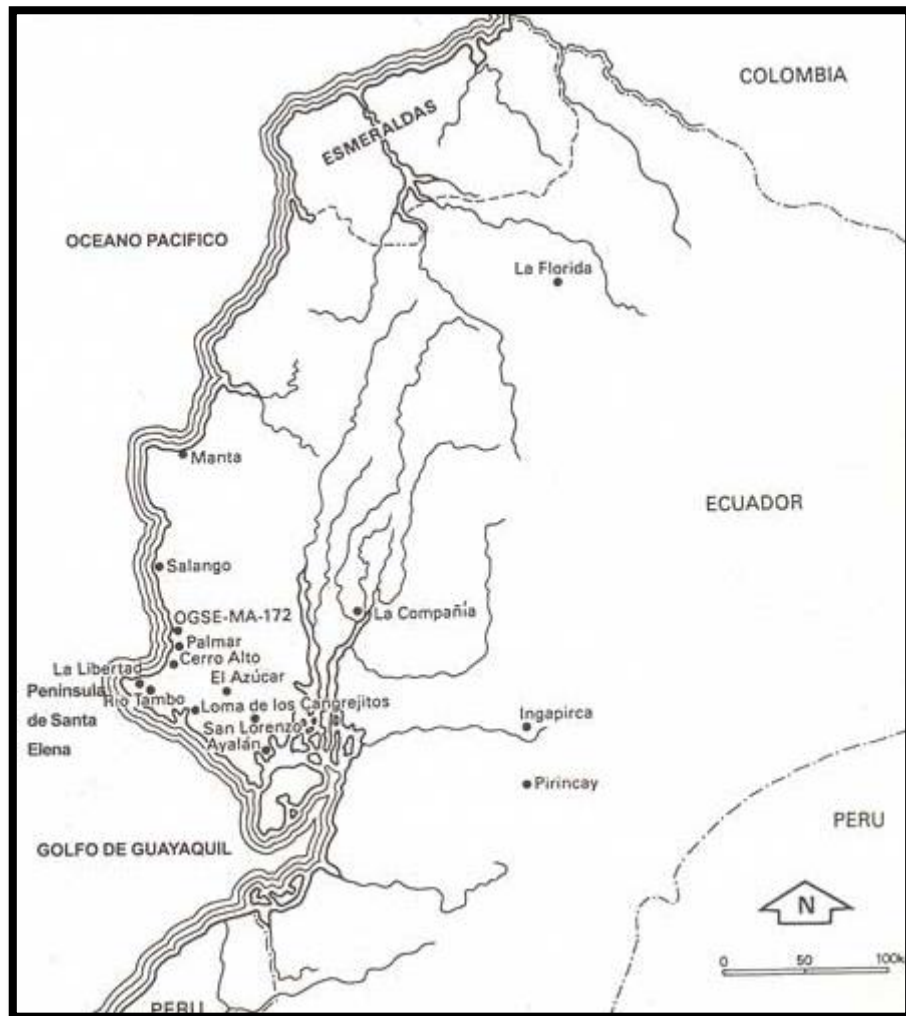
La historia en el Ecuador durante la etapa Prehispánica, puede dividirse en cuatro períodos antes de la invasión del Inca, estas son:

- Período Pre cerámico
- Período Formativo Tardío
- Período de Desarrollo Regional
- Período de Integración

Siendo de esta manera el periodo formativo tardío el periodo durante el cual según estudios arqueológicos se demuestra la aparición, fabricación y utilidad de las Hachas moneda de cobre.

En Ecuador los objetos metálicos más antiguos aparecen en la costa, en el sitio de Salango cerca de la ciudad de Manta (Fig. 1). Pertenecen a la fase Chorrera y datan de alrededor del año 1500 antes de nuestra era. Son de oro, plata y cobre, y martillados.

Más tarde, y antes del año 500 de nuestra era, los artesanos ecuatorianos, al igual que los mesoamericanos, comenzaron a fabricar agujas, pinzas, hachas moneda y argollas.



*Figura 1: Mapa de Ecuador mostrando los sitios arqueológicos mencionados en el texto*

Las hachas moneda son objetos en forma de T, martillados y sumamente delgados, hechos de aleaciones de cobre - arsénico. Las aleaciones de cobre - plata son comunes en el norte del Perú, en Ecuador y en el estado de Michoacán. Las culturas en las que

de acuerdo a excavaciones se han hallado hachas moneda son las culturas de la costa ecuatoriana Manteño – Huancavilca y Milagro- Quevedo que estaban establecidas en las actuales provincias de el Oro, Manabí, Guayas (Santa Elena, Anyulla, Cangrejillo) y Los Ríos (las Palmas).

Las hachas moneda han sido halladas en fardas funerarias, en paquetes de 15 o 20 hachas, que ha sido un número repetitivo en estos hallazgos, sus pesos varían de 15 a 105gr, estos paquetes se encontraban amarrados al cuerpo del difunto con hilos de algodón

Las hachas moneda se caracterizan por un peso más o menos constante, y al parecer eran raspadas eventualmente para reconocer su “autenticidad”, ya que muestran en sus lados una serie de estrías que no se justifican en el momento de su elaboración. Estas “reglas de juego” o normas convencionales, no solamente fueron reconocidas a nivel regional, sino que las compartieron con otras sociedades del Occidente de México y del norte de Perú.

Una relación del siglo XVI, la de Salomón Jerez, describe el encuentro de europeos con una balsa Manteña en las aguas de lo que es ahora Ecuador, y nos da una idea más concreta de cómo sucedió. Los Manteños dominaban el territorio en la costa norte - central del Ecuador. Eran grandes navegantes y comerciantes marítimos. Sabemos que viajaban a lo largo de las costas del Pacífico de Sudamérica. La relación describe ciertos objetos de metal que llevaban los navegantes ecuatorianos en sus balsas, y entre ellos

había cintas, pinzas, cascabeles y hachas moneda de metal. El documento relata que los marinos ecuatorianos llevaban tales objetos para cambiarlos por la concha *Spondylus*, que habita las aguas tibias ecuatoriales desde el norte del golfo de Guayaquil hasta el golfo de California. Probablemente fueron estos navegantes ecuatorianos quienes llevaron las hachas moneda y otros objetos a México, tal vez para intercambiarlos por la concha *Spondylus* que procesaban los pobladores de las costas de Guerrero, Michoacán, Coloma y Nayarit (Fig. 2).

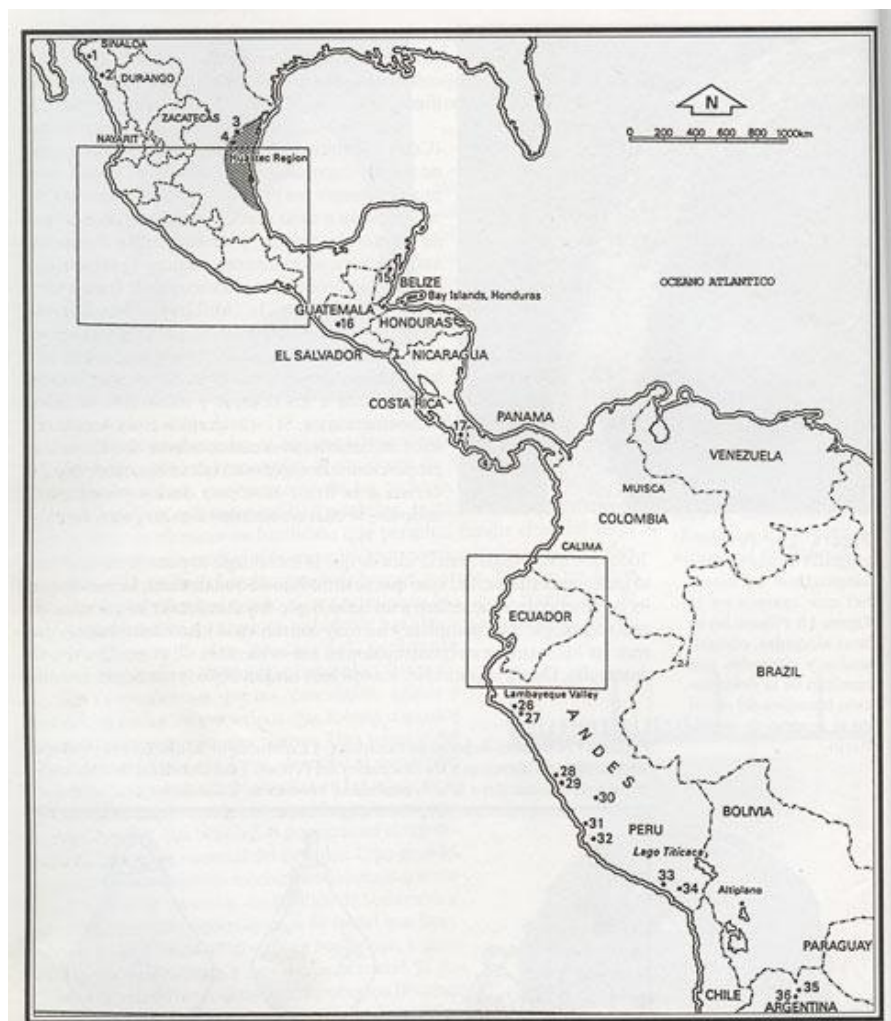


Figura 2: Mapa de las Américas donde se destaca la zona de posible comercialización entre Ecuador y México

## **JUSTIFICACION**

El presente trabajo de investigación, a través de la aplicación de técnicas estadísticas multivariantes, constituirá un aporte a investigaciones ya efectuadas dentro del ámbito minero en el Ecuador. Específicamente, mediante el estudio estadístico de hachas moneda pre-hispánicas se podrá determinar diferentes grupos de acuerdo a su composición química.

Es importante mencionar que el Cobre es uno de los minerales más utilizados en la fabricación de diversos objetos indispensables en el desarrollo de la humanidad y al determinar yacimientos del mismo se podrá contribuir con el desarrollo económico de nuestro país.

De los resultados de esta investigación se beneficiara el sector minero ecuatoriano y en particular la empresa minera Ecuacorriente, que además colaborará en el desarrollo del presente trabajo investigativo.



## **OBJETIVOS**

### **OBJETIVO GENERAL**

Aplicar técnicas estadísticas multivariantes en el estudio de los principales componentes químicos encontrados sobre hachas moneda de Cobre para determinar diferentes grupos de acuerdo a su composición química.

### **OBJETIVOS ESPECIFICOS**

- Recolectar la información necesaria para el análisis a través de la medición de varios componentes químicos sobre hachas moneda de Cobre.
- Determinar posibles variables redundantes y analizar la posible presencia de datos atípicos.
- Encontrar, si existen, grupos de hachas moneda con características químicas similares.
- Determinar el origen y el tipo de Cobre utilizado en la fabricación de hachas moneda.

- Determinar la presencia de factores externos que pudieron haber intervenido dentro del hallazgo y fabricación de estos instrumentos pre-hispánicos.

## **HIPOTESIS**

La aplicación de técnicas estadísticas multivariantes en el estudio de los principales componentes químicos encontrados sobre hachas moneda de Cobre permitirá determinar grupos existentes formados de acuerdo a su composición química.

## **I. CAPITULO**

### **MARCO TEORICO**

#### **1.1.ARQUEOLOGIA CUANTITATIVA: LA FUNCION DEL ANALISIS MULTIVARIANTE.**

Si intentamos hacer una valoración histórica del uso de las técnicas estadísticas podemos ver cómo con anterioridad a la II Guerra Mundial, la preponderancia de la arqueología tradicional hacía inexistente la utilización de las matemáticas y con ello, de los principios de cuantificación. La aproximación teórica del arqueólogo se veía restringida a una "identificación de culturas espacio-temporales a partir de los vestigios materiales que la caracterizarán". Será tras finalizar la guerra, cuando empiecen a proliferar estas técnicas.

En Francia, ligados a la arqueología prehistórica de larga tradición en este país, empiezan a aparecer los primeros trabajos en los que se introducen técnicas gráficas y de estadística elemental. Muestra de ello son los trabajos de Bordes en los años 50, que utiliza diagramas acumulativos para identificar culturas prehistóricas, los de

Spaulding en el año 1953, con los primeros cálculos de correlación, para evidenciar los tipos de objetos, o las primeras seriaciones para establecer cronologías relativas de Brainerd y Robinson en 1951, todos ellos expuestos en el trabajo de Djindjian 1984-85.

Si estos son los primeros intentos con estadísticas sencillas, será en la década de los 60, cuando los métodos cuantitativos en arqueología empezarán a adquirir importancia.

Ese camino hacia la cuantificación de los datos, está motivado por el deseo de eliminar la carga subjetiva del investigador y dotar a la Arqueología de un aparente carácter científico. Por ello, cuando Shenan 1992 habla de "Revolución cuantitativa", indica que ésta se llevó a cabo para dar un aspecto científico a la Arqueología y desligarla de su imagen tradicional.

Es necesario destacar que en su mayor parte las técnicas aplicadas dentro del estudio de Arqueología cuantitativa proceden del área de la Estadística conocida como Análisis Multivariante, difiere de otras técnicas usada ya que el análisis Multivariante tiene una base teórica bien fundamentada en las Matemáticas y la Estadística.

A continuación se describirán las Técnicas Multivariantes aplicadas dentro de esta investigación.

## **1.2.ANALISIS MULTIVARIADO**

El Análisis Multivariante es un conjunto de métodos estadísticos y matemáticos, destinados a describir e interpretar los datos que provienen de la observación de varias variables estadísticas, estudiadas conjuntamente.

Además pretende los siguientes objetivos.

- Resumir el conjunto de variables en unas pocas nuevas variables, construidas como transformaciones de las originales, con la mínima pérdida de información.
- Encontrar grupos en los datos si existen.
- Clasificar nuevas observaciones en grupos definidos.
- Relacionar dos conjuntos de variables.

Las técnicas de análisis Multivariante tienen aplicaciones en todos los campos científicos y comenzaron desarrollándose para resolver problemas de clasificación en Biología, se extendieron para encontrar variables indicadoras y factores en Psicometría, Marketing y las Ciencias sociales y han alcanzado una gran aplicación en Ingeniería y Ciencias de la computación como herramientas para resumir la

información y diseñar sistemas de clasificación automática y de reconocimiento de patrones.<sup>1</sup>

## **1.2.1. Descripción de Datos Multivariantes**

### ***1.2.1.1. Tipos de Variables.***

Una variable es una característica que al ser medida en diferentes individuos es susceptible de adoptar diferentes valores.

Existen diferentes tipos de variables:

- **Variables cualitativas**

Son las variables que expresan distintas cualidades, características o modalidad. Cada modalidad que se presenta se denomina atributo o categoría y la medición consiste en una clasificación de dichos atributos. Las variables cualitativas pueden ser dicotómicas cuando sólo pueden tomar dos valores posibles como sí y no, hombre y mujer o son politómicas cuando pueden adquirir tres o más valores. Dentro de ellas podemos distinguir:

---

<sup>1</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

- **Variable cualitativa ordinal o variable cuasicuantitativa:** La variable puede tomar distintos valores ordenados siguiendo una escala establecida, aunque no es necesario que el intervalo entre mediciones sea uniforme, por ejemplo: leve, moderado, grave.
- **Variable cualitativa nominal:** En esta variable los valores no pueden ser sometidos a un criterio de orden como por ejemplo los colores o el lugar de residencia.
- **Variabes cuantitativas**

Son las variables que se expresan mediante cantidades numéricas. Las variables cuantitativas además pueden ser:

- **Variable discreta:** Es la variable que presenta separaciones o interrupciones en la escala de valores que puede tomar. Estas separaciones o interrupciones indican la ausencia de valores entre los distintos valores específicos que la variable pueda asumir. Ejemplo: El número de hijos (1, 2, 3, 4, 5).
- **Variable continua:** Es la variable que puede adquirir cualquier valor dentro de un intervalo especificado de valores. Por ejemplo la masa (2,3 kg, 2,4 kg, 2,5 kg,...) o la altura (1,64 m, 1,65 m, 1,66 m,...), que solamente está limitado por la precisión del aparato medidor, en teoría permiten que siempre exista un valor entre dos variables, también puede ser el dinero o un salario dado.

### **1.2.1.2. Matriz de Datos**

Sobre cada uno de los  $n$  individuos que conforman la muestra de estudio, son observados  $p$  características, dando origen a la matriz de datos  $X$  de orden  $(n \times p)$ , cada una de estas  $p$  variables se denomina una variable escalar o univariante y el conjunto de las  $p$  variables forman una variable vectorial o multivariante.

Denotaremos por  $x_{ij}$  al elemento genérico de esta matriz, que representa el valor de la variable escalar  $j$  sobre el individuo  $i$ . Es decir:

Datos  $x_{ij}$       donde  $i = 1, \dots, n$  representa el individuo;

$j = 1, \dots, p$  representa la variable

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

### **1.2.2. Análisis Exploratorio de Datos.**

La finalidad del Análisis Exploratorio de Datos (AED) es examinar los datos previamente a la aplicación de cualquier técnica estadística. De esta forma el analista consigue un entendimiento básico de sus datos y de las relaciones existentes entre las variables analizadas.



El AED proporciona métodos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de datos, tratamiento y evaluación de datos ausentes, identificación de casos atípicos y comprobación de los supuestos subyacentes en la mayor parte de las técnicas multivariantes.

### ***1.2.2.1. Análisis Univariante<sup>2</sup>***

Describir datos multivariantes supone estudiar cada variable aisladamente y además las relaciones entre ellas. El estudio univariante de la variable escalar  $x_j$  implica calcular su media:

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$$

Se calcula una medida de variabilidad con relación a la media, promediando las desviaciones entre los datos y su media. Si definimos las desviaciones mediante  $d_{ij} = (x_{ij} - \bar{x}_j)^2$ , donde el cuadrado se toma para prescindir del signo, se define la desviación típica por:

---

<sup>2</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

$$S_j = \sqrt{\frac{\sum_{i=1}^n d_{ij}}{n}} = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n}}$$

Y su cuadrado es la varianza,  $S_j^2 = \sum_{i=1}^n d_{ij}/n$ . Para comparar la variabilidad de distintas variables conviene construir medidas de variabilidad relativa que no dependan de las unidades de medida. Una de estas medidas es el coeficiente de variación.

$$CV_j = \sqrt{\frac{S_j^2}{\bar{x}_j^2}}$$

Donde de nuevo se toman los cuadrados para prescindir del signo y suponemos que  $\bar{x}_j$  es distinto de cero. En tercer lugar, conviene calcular los coeficientes de asimetría, que miden la simetría de los datos respecto a su centro, y que se calculan como:

$$A_j = \frac{1}{n} \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^3}{S_j^3}$$

Este coeficiente es cero para una variable simétrica. Cuando el valor absoluto del coeficiente es aproximadamente mayor que uno podemos concluir que los datos tienen una distribución claramente asimétrica.

Una característica importante de un conjunto de datos es su homogeneidad. Si las desviaciones  $d_{ij}$  son muy distintas, esto sugiere que hay datos que se separan mucho

de la media y por lo tanto alta heterogeneidad, el coeficiente de homogeneidad puede escribirse de la siguiente forma.

$$H_j = \frac{\frac{1}{n} \sum_{i=1}^n (d_{ij} - S_j^2)^2}{S_j^4}$$

Este coeficiente es siempre mayor o igual a cero. Desarrollando el cuadrado del numerador este coeficiente puede escribirse también como:

$$H_j = \frac{1}{n} \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^4}{S_j^4} - 1 = K_j - 1$$

El primer miembro de esta expresión,  $K_j$ , es una forma alternativa de medir la homogeneidad y se conoce como coeficiente de kurtosis. Como  $H_j \geq 0$ , el coeficiente de kurtosis será igual o mayor que uno. Ambos coeficientes miden la relación entre la variabilidad de las desviaciones y la desviación media.

### **1.2.2.2. Medidas de Centralización - Vector de Medias<sup>3</sup>**

La medida de centralización más utilizada para describir datos multivariantes es el vector de medias, que es un vector de dimensión  $p$  cuyos componentes son las medias de cada una de las  $p$  variables.

---

<sup>3</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_p \end{bmatrix}$$

Su expresión a partir de la matriz de datos es:

$$\bar{x} = \frac{1}{n} X' \mathbf{1}$$

Donde:

$\mathbf{1} \rightarrow$  representa siempre el vector de unos de la dimensión adecuada.

$X' \rightarrow$  son los vectores columna que contienen los valores de las p variables en cada elemento de la muestra.

$$\bar{x} = \frac{1}{n} [x_1 \quad \dots \quad x_p] \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

### **1.2.2.3. Matriz de Varianzas y Covarianzas<sup>4</sup>**

Para variables escalares la variabilidad respecto a la media se mide habitualmente por la varianza, o su raíz cuadrada, la desviación típica. La relación lineal entre dos variables se mide por la covarianza. Se define esta matriz como:

---

<sup>4</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

$$S = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})'$$

Es una matriz cuadrada y simétrica de orden p. su diagonal principal está formada por las varianzas de las variables y fuera de la diagonal se encuentra las covarianzas entre los pares de variables.

La matriz de varianzas y covarianzas, que llamaremos para simplificar matriz de covarianzas, es la matriz simétrica de orden p con forma:

$$S = \begin{bmatrix} S_1^2 & S_{12} & \dots & S_{1p} \\ S_{21} & S_2^2 & \dots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \dots & S_p^2 \end{bmatrix}$$

**Propiedades:**

Así como la varianza es siempre un número no negativo, la matriz de varianzas y covarianzas tiene una propiedad similar: es semidefinida positiva. Esta propiedad asegura que si y es cualquier vector,  $y'Sy \geq 0$ . También la traza, el determinante y los valores propios de esta matriz son no negativos.

#### **1.2.2.4. Variables Redundantes<sup>5</sup>**

Variabes redundantes quiere decir que existe una alta relación lineal entre las variables.

Es decir que la matriz  $S$  tiene una raíz característica o autovalor igual a cero y  $w$  es el vector característico asociado a la raíz característica cero. Multiplicado por  $w'$  se obtiene,  $(\tilde{X} w)'(\tilde{X} w) = 0$ , que implica  $\tilde{X} w = 0$ , y concluimos que una variable es una combinación lineal exacta de las otras. En consecuencia, es posible reducir la dimensionalidad del sistema eliminando esta variable.

#### **1.2.2.5. Medidas Globales de Variabilidad**

Cuando se tiene un conjunto de variables aleatorias, se puede considerar la variabilidad de cada una de ellas, así como las covarianzas que pueden hallarse entre los distintos pares de variables, de modo que se tiene una gran cantidad de información que en muchas ocasiones es deseable, por no decir necesaria, resumir en un solo número. Aunque resumir una gran cantidad de información en un solo número puede parecer algo muy bueno, al hacerlo se debe pagar el precio de perder mucha información y, lo que es peor aún, en muchos casos, se obtienen medidas resumen

---

<sup>5</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

que no corresponden con lo que sucede en el conjunto de variables. Son varias las medidas resumen o medidas globales de variabilidad que existen. Entre ellas podemos citar las siguientes:

#### **1.2.2.5.1. Varianza Generalizada<sup>6</sup>**

Una mejor medida de variabilidad global es la varianza generalizada, concepto debido a Wilks (1932). Se define como el determinante de la matriz de varianzas y covarianzas, es decir,

$$VG = |S|$$

Su raíz cuadrada se denomina desviación típica generalizada, y tiene las propiedades siguientes:

Está bien definida, ya que el determinante de la matriz de varianzas y covarianzas es siempre no negativo.

Es una medida del área (para  $p = 2$ ), volumen (para  $p = 3$ ) o hipervolumen

(Para  $p > 3$ ) ocupado por el conjunto de datos.

---

<sup>6</sup> Tomado de MORALES Víctor Hugo y Vargas José Alberto, *Revista Colombiana de Estadística*, Junio 2008.

Un inconveniente de la varianza generalizada es que no sirve para comparar conjuntos de datos con distinto número de variables, ya que tiene la dimensión del producto de las variables.

#### **1.2.2.5.2. Varianza Efectiva<sup>7</sup>**

Para evitar estos inconvenientes, Peña & Rodríguez (2003) han propuesto como medida global de variabilidad la varianza efectiva, dada por:

$$VE = |S|^{1/p}$$

Cuya ventaja es que cuando todas las variables son de igual dimensión, tiene las unidades de la varianza. Para matrices diagonales, esta medida es simplemente la media geométrica de las varianzas. Se observa que, como el determinante es el producto de los valores propios, la varianza efectiva es la media geométrica de los valores propios de la matriz S que, por ser definida positiva, serán siempre no negativos. La varianza efectiva tiene en cuenta la dependencia conjunta, ya que si una variable fuese combinación de las demás, al existir un valor propio nulo, será cero.

---

<sup>7</sup> Tomado de MORALES Víctor Hugo y Vargas José Alberto, *Revista Colombiana de Estadística*, Junio 2008.



### **1.2.2.6. Medidas de Dependencia Lineal<sup>8</sup>**

Un objetivo fundamental de la descripción de los datos multivariantes es comprender la estructura de dependencias entre las variables. Estas dependencias pueden estudiarse:

- entre pares de variables;
- entre una variable y todas las demás;
- entre el conjunto de todas las variables.

A continuación se estudiara estos tres aspectos.

#### **1.2.2.6.1. Dependencia por pares. Matriz de Correlación.**

Este tipo de dependencia es la más utilizada en la estadística descriptiva, y su estudio se inicia con el cálculo y la interpretación (que suponemos conocida) del coeficiente de correlación lineal de Pearson. Supongamos que  $x_j$  y  $x_k$  son dos variables en  $\mathbb{R}^p$ , entonces este coeficiente es:

$$r_{jk} = \frac{s_{jk}}{s_j s_k}$$

---

<sup>8</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

Y tiene las siguientes propiedades:

$$0 \leq r_{jk} \leq 1;$$

Si existe una relación exacta entre las variables,  $x_{ij} = a + bx_{ik}, i = 1 \rightarrow p$ , entonces

$$|r_{jk}| = 1;$$

$r_{jk}$  es invariante ante transformaciones lineales de las variables.

La dependencia por pares entre las variables, en el caso multivariante entre las columnas de la matriz X, se mide simultáneamente por la matriz de correlación R definida como:

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix}$$

Es una matriz cuadrada, simétrica, con unos sobre la diagonal principal, y evidentemente semidefinida positiva.

#### **1.2.2.6.2. Dependencia de cada variable y el resto. Regresión Múltiple.**

Sin llegar a una situación extrema puede ser, que hayan variables que sean muy dependientes de las demás y lo que interesa es medir su grado de dependencia.

Supongamos que por alguna razón estamos interesados en la variable  $x_j$ , que para simplificar la notación denotaremos por  $y$ , y llamaremos variable predictora (observe que  $x_j$ , es una columna de nuestra matriz  $X$ ). Supongamos además que, a priori, queremos expresar la variable  $y$  como una combinación lineal de las variables restantes  $x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_p$  que llamaremos variables explicativas o regresores. Entonces, nuestra intención es encontrar los "mejores" valores de  $\hat{\beta}_k$  para encontrar:

$$\hat{y}_i = \bar{y} + \hat{\beta}_1(x_{i1} - \bar{x}_1) + \dots + \hat{\beta}_p(x_{ip} - \bar{x}_p) \quad i = 1, \dots, n$$

Una medida descriptiva de la capacidad predictiva del modelo es el cociente entre la variabilidad explicada por la regresión y la variabilidad total, y tal medida se llama coeficiente de determinación o coeficiente de correlación múltiple al cuadrado, y se denota por:

$$R_{j,1,\dots,p}^2 = 1 - \frac{s_r^2(j)}{s_j^2}$$

### 1.2.2.6.3. Coeficiente de Dependencia.

Para obtener una medida conjunta de la dependencia entre las variables podemos utilizar el determinante de la matriz de correlación, que mide el alejamiento del conjunto de variables de la situación de perfecta dependencia lineal.

$$0 \leq |R| \leq 1$$

### **1.2.2.7. Análisis de Datos Atípicos**

Uno de los problemas en el análisis de datos es manejar los valores atípicos dentro de un grupo de datos. Estas observaciones son muy distintas al resto (por errores de medida, cambio de instrumentos o heterogeneidad de los elementos).

Es frecuente entre 1-3% en experimentos controlados y alrededor de un 5% en experimentos incontrolados.

Las consecuencias de un dato atípico pueden ser graves: distorsionan medias y desviaciones típicas enmascaran las relaciones existentes entre ellas.

Para identificar datos atípicos es importante destacar en primera instancia las representaciones y graficas y finalmente el análisis analítico para identificar o descartar la presencia de datos atípicos.

#### **1.2.2.7.1. Representaciones Gráficas.**

El primer paso de cualquier análisis multivariante es representar gráficamente las variables individualmente, mediante un histograma o un diagrama de caja. Estas representaciones son muy útiles para detectar asimetrías, heterogeneidad, datos atípicos etc. En segundo lugar conviene construir los diagramas de dispersión de las variables por pares, y esta posibilidad se incluye ya en muchos programas de ordenador. Con  $p$  variables existen  $p(p-1)/2$  gráficos posibles que pueden disponerse

en forma de matriz y son muy útiles para entender el tipo de relación existente entre pares de variables, e identificar puntos atípicos en la relación bivariante. En particular, estos gráficos son importantes para apreciar si existen relaciones no lineales, en cuyo caso la matriz de covarianzas puede no ser un buen resumen de la dependencia entre las variables.<sup>9</sup>

Para identificar posibles datos atípicos mediante gráficos estadísticos son muy útiles plasmar los datos mediante diagramas de dispersión y diagramas de caja.

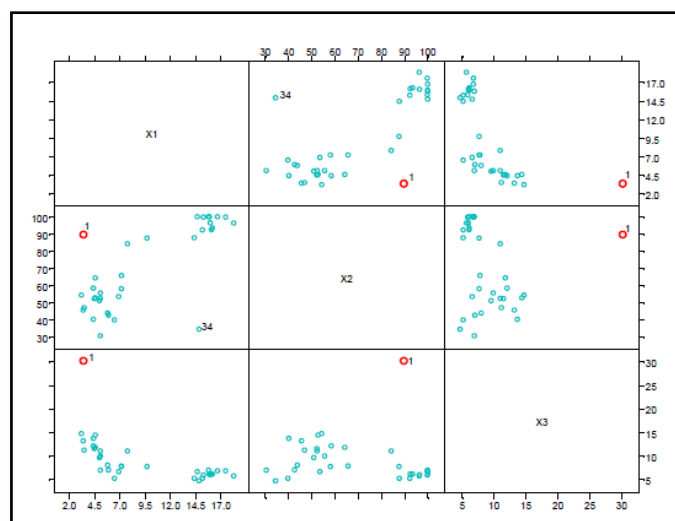


Fig. 3: Ejemplo de una matriz de dispersión con datos sospechosos.

---

<sup>9</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

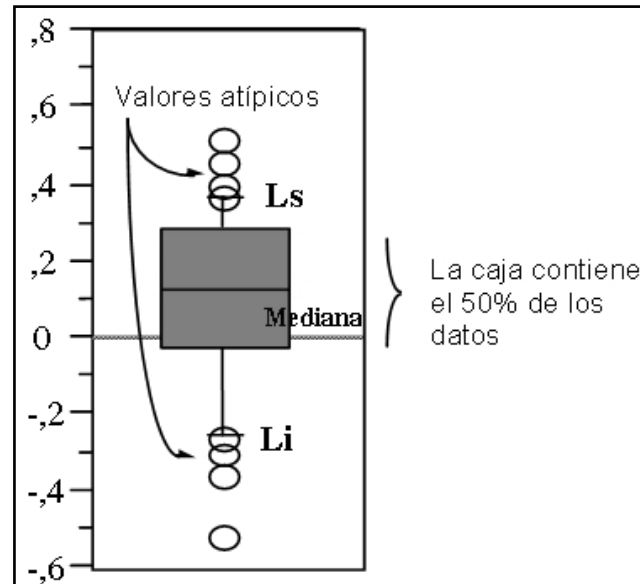


Fig. 4: Ejemplo de diagrama de caja con datos atípicos.

Como se muestra en las Fig. 4.1 y Fig. 4.2 la presencia de datos potencialmente sospechosos ya que claramente son diferentes al resto de los datos que posiblemente son homogéneos.

#### 1.2.2.7.2. Representación mediante figuras.

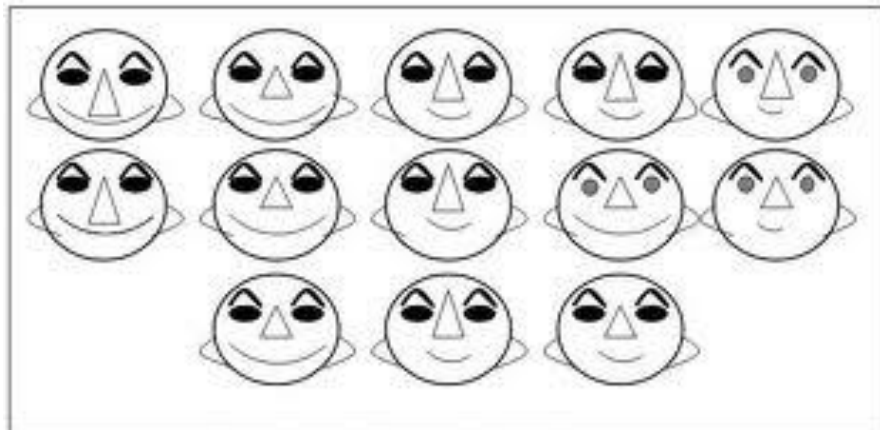
Para más de tres variables se utilizan principalmente dos tipos de métodos gráficos. El primero, es mostrar los datos mediante figuras planas, asociando cada variable a una característica del gráfico. El segundo, es buscar conjuntos de proyecciones en una y dos dimensiones que revelen aspectos característicos de los datos.

Existen muchas alternativas posibles para representar los datos mediante figuras. Chernoff ha propuesto la utilización de caras, que tienen la ventaja de nuestra facilidad para reconocer patrones en este formato y el inconveniente de que la representación

es muy dependiente de las variables escogidas para representar cada rasgo. Por ejemplo, la boca y la forma de la cabeza son rasgos más llamativos que las orejas o la longitud de la nariz, y el mismo conjunto de datos puede sugerir distintos patrones de similitud entre las observaciones según la asociación elegida entre rasgos y variables.

Si asociamos cada variable a un rasgo de una figura plana, podemos representar cada elemento en la muestra por una figura geométrica. En estas representaciones las similitudes entre figuras indican las similitudes entre los elementos, y los valores atípicos aparecerán como figuras discordantes con el resto como se muestra en la Fig.

4.3.



*Fig. 5: Aplicación de Caras de Chernoff.*

### **1.2.2.7.3. Identificación de grupos de atípicos<sup>10</sup>.**

Hay dos filosofías para tratar con la heterogeneidad. La primera es utilizar estimadores robustos, que son estimadores diseñados para verse poco afectados por cierta contaminación de atípicos. La segunda es detectar los atípicos, y aplicar el cálculo de los estimadores a las muestras limpias de atípicos. Ambos enfoques son complementarios, pero ha sido de gran utilidad la aplicación del segundo método.

El procedimiento para detectar grupos de atípicos es eliminar de la muestra todos los puntos sospechosos, de manera que evitemos el enmascaramiento y podamos calcular el vector de medias y la matriz de covarianzas sin distorsiones. A continuación identificaremos con estos estimadores la distancia de cada punto sospechoso respecto al centro de los datos, y consideraremos atípicos a los muy alejados. El primer paso para identificar las observaciones sospechosas es detectar aquellas que lo sean claramente respecto a una variable.

Observemos que cualquier observación atípica multivariante debe aparecer como atípica al menos en una dirección de proyección: la definida por la recta que une el centro de los datos con el dato atípico. En base a esta idea, Stahel (1981) y Donoho (1982) propusieron generar muchas direcciones al azar, proyectar los puntos sobre

---

<sup>10</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*



estas direcciones y marcar como datos atípicos a aquellas observaciones que aparecen como extremas en estas proyecciones. Para generar direcciones al azar pueden tomarse muestras al azar de  $p$  puntos, calcular el plano que las contiene y tomar como dirección el vector ortogonal al plano.

Este método funciona bien con pocas variables, pero al aumentar la dimensión del problema el número de direcciones que necesitamos generar para cubrir razonablemente el espacio y tener garantías de éxito aumenta exponencialmente. Una solución propuesta por Peña y Prieto (2001), es proyectar los datos sobre ciertas direcciones específicas, escogidas de manera que tengan alta probabilidad de mostrar los atípicos cuando existan. Hemos comentado que en muestras univariantes una pequeña proporción de atípicos hace aumentar el coeficiente de kurtosis, lo que sugiere investigar las direcciones donde los puntos proyectados tengan máxima kurtosis univariante. Por otro lado, un grupo grande de atípicos puede producir bimodalidad y baja kurtosis, por lo que conviene también explorar las direcciones donde los puntos proyectados tengan mínima kurtosis. La idea del procedimiento es buscar  $p$  direcciones ortogonales de máxima kurtosis y  $p$  direcciones ortogonales de mínima kurtosis, eliminar provisionalmente los datos extremos en estas direcciones, calcular la media y la matriz de covarianzas con los datos no sospechosos y después identificar los datos atípicos como aquellos que son extremos con la distancia de

Mahalanobis calculada con las estimaciones no contaminadas. Dada la muestra multivariante  $(x_1, \dots, x_n)$ , el proceso se realiza como sigue:

1. Sean  $\bar{x}$  y  $S$  el vector de medias y la matriz de covarianzas de los datos. Estandarizar los datos de forma multivariante y sean  $z_i = S_x^{-1/2}(x_i - \bar{x})$  los datos estandarizados con media cero y matriz de covarianzas identidad. Tomar  $j = 1$  y  $z_i^{(1)} = z_i$ .
2. Calcular la dirección  $d_j$  con norma unidad que maximiza el coeficiente de kurtosis univariante de los datos proyectados. Llamando  $y_i^{(j)} = d_j' z_i^j$ , a los datos proyectado sobre la dirección  $d_j$ , esta dirección se obtiene como solución de:

$$\max \sum (y_i^j - \bar{y}^j)^4 + \lambda(d'd - 1)$$

3. Proyectar los datos sobre un espacio de dimensión  $p - j$  definido como el espacio ortogonal a la dirección  $d_j$ . Para ello tomar  $z^{(j+1)} = (I - d_j d_j') z^j$ . Hacer  $j = j + 1$ .
4. Repetir (2) y (3) hasta obtener las  $p$  direcciones,  $d_1, \dots, d_p$ .
5. Repetir (2) y (3) pero ahora minimizando la kurtosis en lugar de maximizarla para obtener otras  $p$  direcciones,  $d_{p+1}, \dots, d_{2p}$ .
6. Considerar como sospechosos aquellos puntos que en alguna de estas  $2p$  direcciones están claramente alejados del resto, es decir, verifican

$$\frac{|y_i^{(j)} - \text{med}(y^{(j)})|}{\text{MEDA}(y^{(j)})} > 5$$

A continuación se eliminan todos los valores sospechosos detectados y se vuelve a 2 para analizar los datos restantes. La estandarización multivariante ahora se realizará con la nueva media y matriz de covarianzas de los datos restantes. Los pasos 2 a 6 se repiten hasta que no se detecten más datos atípicos o se haya eliminado una proporción de datos prefijada, por ejemplo un máximo del 40% de los datos.

Una vez que la muestra no contenga más valores sospechosos con el criterio anterior se calcula el vector de medias,  $\bar{x}_R$ , y la matriz de covarianzas,  $S_R$ , de los datos no sospechosos, y las distancias de Mahalanobis para los sospechosos como:

$$\mathbf{d}_R^2(\mathbf{x}_i, \bar{\mathbf{x}}_R) = (\mathbf{x}_i - \bar{\mathbf{x}}_R) S_R^{-1} (\mathbf{x}_i - \bar{\mathbf{x}}_R)'$$

Finalmente se consideran valores atípicos a aquellos valores mayores que  $p + 3\sqrt{2p}$  (recordemos que el valor promedio de la distancia de Mahalanobis es  $p$ ). Algunos puntos del conjunto de sospechosos serán atípicos y otros no. Los atípicos son desechados, y los buenos incorporados al conjunto de puntos. Finalmente, se calculará un vector de medias,  $\bar{x}_f$ , y una matriz de covarianzas,  $S_f$ , con los puntos no atípicos, que serán las estimaciones finales obtenidas de los datos.

### **1.2.3. Componentes Principales.**

Estas técnicas fueron inicialmente desarrolladas por Pearson a finales del siglo XIX y posteriormente fueron estudiadas por Hotelling en los años 30 del siglo XX. Sin embargo, hasta la aparición de los ordenadores no se empezaron a popularizar.

Para estudiar las relaciones que se presentan entre  $p$  variables correlacionadas (que miden información común) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tenga repetición o redundancia en la información) llamado conjunto de componentes principales.

Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra.

De modo ideal, se buscan  $r < p$  variables que sean combinaciones lineales de las  $p$  originales y que estén incorreladas, recogiendo la mayor parte de la información o variabilidad de los datos.

Si las variables originales están incorreladas de partida, entonces no tiene sentido realizar un análisis de componentes principales.

El análisis de componentes principales es una técnica matemática que no requiere la suposición de normalidad multivariante de los datos, aunque si esto último se cumple se puede dar una interpretación más profunda de dichos componentes.

Su utilidad es doble<sup>11</sup>:

1. Permite representar óptimamente en un espacio de dimensión pequeña, observaciones de un espacio general  $p$ -dimensional. En este sentido componentes principales es el primer paso para identificar posibles variables latentes o no observadas, que están generando la variabilidad de los datos.
2. Permite transformar las variables originales, en general correladas, en nuevas variables incorreladas, facilitando la interpretación de los datos.

### ***1.2.3.1. Cálculo del Primer Componente<sup>12</sup>.***

El primer componente principal será la combinación lineal de las variables originales que tenga varianza máxima. Los valores de este primer componente en los  $n$  individuos se representarán por un vector  $z_1$ , dado por

$$z_1 = \mathbf{X}a_1$$

---

<sup>11</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

<sup>12</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

Como las variables originales tienen media cero también  $z_1$  tendrá media nula. Su varianza será:

$$\mathbf{Var}(z_1) = \frac{1}{n} z_1' z_1 = \frac{1}{n} a_1' X' X a_1 = a_1' S a_1$$

Donde  $S$  es la matriz de varianzas y covarianzas de las observaciones. Es obvio que podemos maximizar la varianza sin límite aumentando el módulo del vector  $a_1$ . Para que la maximización de  $Var(z_1)$  tenga solución debemos imponer una restricción al módulo del vector  $a_1$ , y, sin pérdida de generalidad, impondremos que  $a_1' a_1 = 1$ . Introduciremos esta restricción mediante el multiplicador de Lagrange:

$$M = a_1' S a_1 - \lambda(a_1' a_1 - 1)$$

Y maximizaremos esta expresión de la forma habitual derivando respecto a los componentes de  $a_1$  e igualando a cero. Entonces:

$$\frac{\partial M}{\partial a_1} = 2S a_1 - 2\lambda a_1 = 0$$

$$S a_1 = \lambda a_1$$

Que implica que  $a_1$  es un vector propio de la matriz  $S$ , y  $\lambda$  su correspondiente valor propio. Para determinar qué valor propio de  $S$  es la solución de la ecuación  $S a_1 = \lambda a_1$  tendremos en cuenta que, multiplicando por la izquierda por  $a_1'$  esta ecuación,

$$a'_1 S a_1 = \lambda a'_1 a_1 = \lambda$$

Por consiguiente  $\lambda$  es la varianza de  $z_1$ . Como esta es la cantidad que queremos maximizar,  $\lambda$  será el mayor valor propio de la matriz  $S$ . Su vector asociado,  $a_1$ , define los coeficientes de cada variable en el primer componente principal.

De esta manera se podrá calcular de manera similar la segunda, y las componentes necesarias para el estudio.

### ***1.2.3.2. Selección del número de Componentes.***

Se han sugerido distintas reglas para seleccionar el número de componentes a mantener:

1. Realizar un gráfico de  $\lambda_1$  frente  $a_1$ . Comenzar seleccionando componentes hasta que los restantes tengan aproximadamente el mismo valor de  $\lambda_1$ . La idea es buscar un codo en el gráfico, es decir, un punto a partir del cual los valores propios son aproximadamente iguales. El criterio es quedarse con un número de componentes que excluya los asociados a valores pequeños y aproximadamente del mismo tamaño.
2. Seleccionar componentes hasta cubrir una proporción determinada de varianza, como el 80% o el 90%. Esta regla es arbitraria y debe aplicarse con cierto cuidado. Por ejemplo, es posible que un único componente de tamaño recoja el 90% de la

variabilidad y sin embargo pueden existir otros componentes que sean muy adecuados para explicar la forma de las variables.

3. Desechar aquellos componentes asociados a valores propios inferiores a una cota, que suele fijarse como la varianza media,  $\sum \lambda_1/p$ . En particular, cuando se trabaja con la matriz de correlación, el valor medio de los componentes es 1, y esta regla lleva a seleccionar los valores propios mayores que la unidad. De nuevo esta regla es arbitraria: una variable que sea independiente del resto suele llevarse un componente principal y puede tener un valor propio mayor que la unidad. Sin embargo, si esta incorrelada con el resto puede ser una variable poco relevante para el análisis, y no aportar mucho a la comprensión del fenómeno global.

#### **1.2.4. Análisis de Conglomerados.**

Conocido también como Análisis Clusters, Análisis de Conjuntos o Análisis de Conglomerados, constituye una importante técnica de agrupación de datos. En el caso de Arqueología, la necesidad de clasificar los ítems arqueológicos y contribuir tipologías precisas conduce a una utilización, cada vez mas amplia de este tipo de técnicas estadísticas, solventando determinados problemas que aparecen en la arqueología tradicional como son:



- Manejo de grandes cantidades de datos de tipo multivariante que, debido a su dimensionalidad, son difíciles de estudiar a menos que puedan clasificarse en grupos manejables con la mínima pérdida de información.
- Necesidad de disponer de un método de agrupación útil que introduzca un grado de objetividad no obtenible por observación directa.
- Utilización simultánea de varias características a lo largo del proceso, para evitar soluciones descriptivas basadas, en una única característica diferenciadora.

El análisis de conglomerados estudia tres tipos de problemas:

- **Partición de los datos.** Disponemos de datos que sospechamos son heterogéneos y se desea dividirlos en un número de grupos prefijado, de manera que:
  1. Cada elemento pertenezca a uno y solo uno de los grupos;
  2. Todo elemento quede clasificado;
  3. Cada grupo sea internamente homogéneo.
- **Construcción de jerarquías.** Deseamos estructurar los elementos de un conjunto de forma jerárquica por su similitud. Por ejemplo, tenemos una encuesta de atributos de distintas profesiones y queremos ordenarlas por similitud. Una clasificación jerárquica implica que los datos se ordenan en niveles, de manera que los niveles superiores contienen a los inferiores.

Estrictamente, estos métodos no definen grupos, sino la estructura de asociación en cadena que pueda existir entre los elementos. Sin embargo, como veremos, la jerarquía construida permite obtener también una partición de los datos en grupos.

- **Clasificación de variables.** En problemas con muchas variables es interesante hacer un estudio exploratorio inicial para dividir las variables en grupos. Este estudio puede orientarnos para plantear los modelos formales para reducir la dimensión que estudiaremos más adelante. Las variables pueden clasificarse en grupos o estructurarse en una jerarquía.

Los métodos de partición utilizan la matriz de datos, pero los algoritmos jerárquicos utilizan la matriz de distancias o similitudes entre elementos. Para agrupar variables se parte de la matriz de relación entre variables.

#### ***1.2.4.1. Algoritmos Jerárquicos<sup>13</sup>***

Dada una matriz de distancias o de similitudes se desea clasificar los elementos en una jerarquía. Los algoritmos existentes funcionan de manera que los elementos son sucesivamente asignados a los grupos, pero la asignación es irrevocable, es decir, una vez hecha, no se cuestiona nunca más. Los algoritmos son de dos tipos:

---

<sup>13</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

1. **De aglomeración.** Parten de los elementos individuales y los van agregando en grupos.
2. **De división.** Parten del conjunto de elementos y lo van dividiendo sucesivamente hasta llegar a los elementos individuales.

Los algoritmos de aglomeración requieren menos tiempo de cálculo y son los más utilizados.

#### **1.2.4.1.1. Métodos Aglomerativos**

Los algoritmos aglomerativos que se utilizan tienen siempre la misma estructura y sólo se diferencian en la forma de calcular las distancias entre grupos. Su estructura es:

1. Comenzar con tantas clases como elementos,  $n$ . Las distancias entre clases son las distancias entre elementos originales.
2. Seleccionar los dos elementos más próximos en la matriz de distancias y formar con ellos una clase.
3. Sustituir los dos elementos utilizados en (2) para definir la clase por un nuevo elemento que represente la clase construida. Las distancias entre este nuevo elemento y los anteriores se calculan con uno de los criterios que comentamos a continuación.
4. Volver a (2) y repetir (2) y (3) hasta que tengamos todos los elementos agrupados en una clase única.

#### 1.2.4.1.2. Criterios para definir distancias entre grupos

Supongamos que tenemos un grupo A con  $n_a$  elementos, y un grupo B con  $n_b$  elementos, y que ambos se fusionan para crear un grupo (AB) con  $n_a + n_b$  elementos. La distancia del nuevo grupo, (AB), a otro grupo C con  $n_c$  elementos, se calcula habitualmente por alguna de las cinco reglas siguientes:

1. **Encadenamiento simple o vecino más próximo.** La distancia entre los dos nuevos grupos es la menor de las distancias entre grupos antes de la fusión. Es decir:

$$d(\mathbf{C}; \mathbf{AB}) = \min (d_{CA}, d_{CB})$$

Como este criterio sólo depende del orden de las distancias será invariante ante transformaciones monótonas: obtendremos la misma jerarquía aunque las distancias sean numéricamente distintas. Se ha comprobado que este criterio tiende a producir grupos alargados, que pueden incluir elementos muy distintos en los extremos.

2. **Encadenamiento completo o vecino más alejado.** La distancia entre los dos nuevos grupos es la mayor de las distancias entre grupos antes de la fusión. Es decir:

$$d(\mathbf{C}; \mathbf{AB}) = \max (d_{CA}, d_{CB})$$

Este criterio será también invariante ante transformaciones monótonas de las distancias al depender, como el anterior, del orden de las distancias. Tiende a producir grupos esféricos.

3. **Media de grupos.** La distancia entre los dos nuevos grupos es la media ponderada entre las distancias entre grupos antes de la fusión. Es decir:

$$d(\mathbf{C}; \mathbf{AB}) = \frac{n_a}{n_a + n_b} d_{CA} + \frac{n_b}{n_a + n_b} d_{CB}$$

Como se ponderan los valores de las distancias, este criterio no es invariante ante transformaciones monótonas de las distancias.

4. **Método del centroide.** Se aplica generalmente sólo con variables continuas.

#### 1.2.4.1.3. El método de Ward

Un proceso algo diferente de construir el agrupamiento jerárquico ha sido propuesto por Ward y Wishart. La diferencia con los métodos anteriores es que ahora se parte de los elementos directamente, en lugar de utilizar la matriz de distancias, y se define una medida global de la heterogeneidad de una agrupación de observaciones en grupos. Esta medida es  $W$ , la suma de las distancias euclídeas al cuadrado entre cada elemento y la media de su grupo:

$$\mathbf{W} = \sum_g \sum_{i \in g} (\bar{x}_{ig} - \bar{x}_g)' (\bar{x}_{ig} - \bar{x}_g)$$

Donde  $\bar{x}_g$  es la media del grupo  $g$ . El criterio comienza suponiendo que cada dato forma un grupo,  $g = n$  y por tanto  $\mathbf{W}$  es cero. A continuación se unen los elementos que produzcan el incremento mínimo de  $\mathbf{W}$ . Obviamente esto implica tomar los más próximos con la distancia euclídea. En la siguiente etapa tenemos  $n - 1$  grupos,  $n - 2$  de un elemento y uno de dos elementos. Decidimos de nuevo unir dos grupos para que  $\mathbf{W}$  crezca lo menos posible, con lo que pasamos a  $n - 2$  grupos y así sucesivamente hasta tener un único grupo. Los valores de  $\mathbf{W}$  van indicando el crecimiento del criterio al formar grupos y pueden utilizarse para decidir cuántos grupos naturales contienen nuestros datos.

Puede demostrarse que, en cada etapa, los grupos que deben unirse para minimizar  $\mathbf{W}$  son aquellos tales que:

$$\min \frac{n_a n_b}{n_a + n_b} (\bar{x}_a - \bar{x}_b)' (\bar{x}_a - \bar{x}_b)$$

#### 1.2.4.1.4. El Dendrograma

El dendrograma, o árbol jerárquico, es una representación gráfica del resultado del proceso de agrupamiento en forma de árbol. Los criterios para definir distancias que hemos presentado tienen la propiedad de que, si consideramos tres grupos, A, B, C, se verifica que

$$d(A, C) \leq \max\{d(A, B), d(B, C)\}$$

Y una medida de distancia que tiene esta propiedad se denomina ultramétrica. Esta propiedad es más fuerte que la propiedad triangular, ya que una ultramétrica es siempre una distancia.

En efecto si  $d^2(A, C)$  es menor o igual que el máximo de  $d^2(A, B)$ ,  $d^2(B, C)$  forzosamente será menor o igual que la suma  $d^2(A, B) + d^2(B, C)$ . El dendrograma es la representación de una ultramétrica, y se construye como sigue:

1. En la parte inferior del gráfico se disponen los  $n$  elementos iniciales.
2. Las uniones entre elementos se representan por tres líneas rectas. Dos dirigidas a los elementos que se unen y que son perpendiculares al eje de los elementos y una paralela a este eje que se sitúa al nivel en que se unen.
3. El proceso se repite hasta que todos los elementos están conectados por líneas rectas.

Si cortamos el dendrograma a un nivel de distancia dado, obtenemos una clasificación del número de grupos existentes a ese nivel y los elementos que los forman.

El dendrograma es útil cuando los puntos tienen claramente una estructura jerárquica, pero puede ser engañoso cuando se aplica ciegamente, ya que dos puntos pueden

parecer próximos cuando no lo están, y pueden aparecer alejados cuando están próximos.

#### ***1.2.4.2. Conglomerados por variables.***

El análisis de conglomerados de variables es un procedimiento exploratorio que puede sugerir procedimientos de reducción de la dimensión, como el análisis factorial o los métodos de correlación canónica. La idea es construir una matriz de distancias o similitudes entre variables y aplicar a esta matriz un algoritmo jerárquico de clasificación.

#### **1.2.5. Análisis Factorial.**

El Análisis Factorial es el nombre genérico que se da a una clase de métodos estadísticos multivariantes cuyo propósito principal es sacar a la luz la estructura subyacente en una matriz de datos. Analiza la estructura de las interrelaciones entre un gran número de variables no exigiendo ninguna distinción entre variables dependientes e independientes. Utilizando esta información calcula un conjunto de dimensiones latentes, conocidas como FACTORES, que buscan explicar dichas interrelaciones. Es, por lo tanto, una técnica de reducción de datos dado que si se cumplen sus hipótesis, la información contenida en la matriz de datos puede expresarse, sin mucha distorsión, en un número menor de dimensiones representadas por dichos FACTORES.



El análisis factorial está relacionado con los componentes principales, pero existen ciertas diferencias. En primer lugar, los componentes principales se construyen para explicar las varianzas, mientras que los factores se construyen para explicar las covarianzas o correlaciones entre las variables. En segundo lugar, componentes principales es una herramienta descriptiva, mientras que el análisis factorial presupone un modelo estadístico formal de generación de la muestra dada.

#### ***1.2.5.1. El Modelo Factorial<sup>14</sup>.***

##### **Hipótesis básicas:**

Supondremos que observamos un vector de variables  $\mathbf{x}$ , de dimensiones  $(p \times 1)$ , en elementos de una población. El modelo de análisis factorial establece que este vector de datos observados se genera mediante la relación:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u}$$

Donde:

---

<sup>14</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

$\mathbf{f}$   $\longrightarrow$  es un vector  $(m \times 1)$  de variables latentes o factores no observadas. Supondremos que sigue una distribución  $N_m(0, I)$ , es decir los factores son variables de media cero e independientes entre sí y con distribución normal.

$\Lambda$   $\longrightarrow$  es una matriz  $(p \times m)$  de constantes desconocidas  $(m < p)$ . Contiene los coeficientes que describen como los factores,  $\mathbf{f}$ , afectan a las variables observadas,  $\mathbf{x}$ , y se denomina matriz de carga.

$\mathbf{u}$   $\longrightarrow$  es un vector  $(p \times 1)$  de perturbaciones no observadas. Recoge el efecto de todas las variables distintas de los factores que influyen sobre  $\mathbf{x}$ . Supondremos que  $\mathbf{u}$  tiene distribución  $N_p(0, \psi)$  donde  $\psi$  es diagonal, y que las perturbaciones están incorreladas con los factores  $\mathbf{f}$ .

Con estas tres hipótesis deducimos que:

- a)  $\boldsymbol{\mu}$  es la media de las variables  $\mathbf{x}$ , ya que tanto los factores como las perturbaciones tienen media cero;
- b)  $\mathbf{x}$  tiene distribución normal, al ser suma de variables normales, y llamando  $\mathbf{V}$  a su matriz de covarianzas.

$$\mathbf{x} \sim N_p(\boldsymbol{\mu}, \mathbf{V})$$

### **1.2.5.2. Número Máximo de Factores**

Una regla muy sencilla de acuerdo al número máximo de factores podemos establecer que, cuando  $p$  no es muy grande (menor de 10) aproximadamente el número máximo de factores debe ser menor que la mitad del número de variables menos uno. Por ejemplo, el número máximo de factores con 7 variables es 3.

### **1.2.5.3. El método del Factor Principal<sup>15</sup>**

El método del factor principal es un método para estimar la matriz de carga basado en componentes principales. Evita tener que resolver las ecuaciones de máxima verosimilitud, que son más complejas. Tiene la ventaja de que la dimensión del sistema puede identificarse de forma aproximada. Se utiliza en muchos programas de ordenador por su simplicidad. Su base es la siguiente: supongamos que podemos obtener una estimación inicial de la matriz de varianzas de las perturbaciones  $\hat{\psi}$ . Entonces, se puede escribir que:

$$S - \hat{\psi} = \Lambda\Lambda'$$

Y como  $S - \hat{\psi}$  es simétrica, siempre puede descomponerse como:

---

<sup>15</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

$$\mathbf{S} - \hat{\boldsymbol{\psi}} = \mathbf{H}\mathbf{G}\mathbf{H}' = (\mathbf{H}\mathbf{G}^{1/2})(\mathbf{H}\mathbf{G}^{1/2})'$$

Donde  $\mathbf{H}$  es cuadrada de orden  $p$  y ortogonal,  $\mathbf{G}$  es también de orden  $p$ , diagonal y contiene las raíces características de  $\mathbf{S} - \hat{\boldsymbol{\psi}}$ . El modelo factorial establece que  $\mathbf{G}$  debe ser diagonal del tipo:

$$\begin{bmatrix} \mathbf{G}_{1 \times m} & \mathbf{O}_{m \times (p-m)} \\ \mathbf{O}_{(p-m) \times m} & \mathbf{O}_{(p-m) \times (p-m)} \end{bmatrix}$$

Ya que  $\mathbf{S} - \hat{\boldsymbol{\psi}}$  tiene rango  $m$ . Por tanto, si llamamos  $\mathbf{H}_1$  a la matriz  $p \times m$  que contiene los vectores propios asociados a los valores propios no nulos de  $\mathbf{G}_1$  podemos tomar como estimador de  $\boldsymbol{\Lambda}$  la matriz  $p \times m$ :

$$\hat{\boldsymbol{\Lambda}} = \mathbf{H}_1 \mathbf{G}_1^{1/2}$$

Con lo que resolvemos el problema. Observemos que la normalización resultante es:

$$\hat{\boldsymbol{\Lambda}}' \hat{\boldsymbol{\Lambda}} = \mathbf{G}_1^{1/2} \mathbf{H}_1' \mathbf{H}_1 \mathbf{G}_1^{1/2} = \mathbf{G}_1 = \text{Diagonal}$$

Ya que los vectores propios de matrices simétricas son ortogonales, por lo que  $\mathbf{H}_1' \mathbf{H}_1 = \mathbf{I}_m$ .

Por tanto, con este método se obtienen estimadores de la matriz  $\hat{\boldsymbol{\Lambda}}$  con columnas ortogonales entre sí.

En la práctica la estimación se lleva a cabo de forma iterativa como sigue:

1. Partir de una estimación inicial de  $\widehat{\Lambda}_i$  o de  $\widehat{\psi}_i$  mediante  $\widehat{\psi} = \text{diag}(S - \widehat{\Lambda}'\widehat{\Lambda})$ ,
2. Calcular la matriz cuadrada y simétrica  $Q_i = S - \widehat{\psi}_i$ .
3. Obtener la descomposición espectral de  $Q_i$  de forma

$$Q_i = H_{1i}G_{1i}H'_{1i} + H_{2i}G_{2i}H'_{2i}$$

Donde  $G_{1i}$  contiene los  $m$  mayores valores propios de  $Q_i$  y  $H_{1i}$  sus valores propios.

Elegiremos  $m$  de manera que los restantes vectores propios contenidos en  $G_{2i}$  sean todos pequeños y de tamaño similar. La matriz  $Q_i$  puede no ser definida positiva y algunos de sus valores propios pueden ser negativos. Esto no es un problema grave si estos valores propios son muy pequeños y podemos suponerlos próximos a cero.

4. Tomar  $\widehat{\Lambda}_{i+1} = H_{1i}G_{1i}^{1/2}$  y volver a (1). Iterar hasta convergencia, es decir hasta que

$$\|\Lambda_{n+1} - \Lambda_n\| < \epsilon.$$

Los estimadores obtenidos serán consistentes pero no eficientes, como en el caso de Máxima verosimilitud. Tampoco son invariantes ante transformaciones lineales, como los MV, es decir, no se obtiene necesariamente el mismo resultado con la matriz de covarianzas y con la de correlaciones.

#### 1.2.5.4. Estimación Máximo Verosímil<sup>16</sup>.

La estimación de los parámetros mediante el método de máximo verosimilitud puede obtenerse mediante el siguiente algoritmo iterativo.

1. Partir de una estimación inicial. Si tenemos una estimación  $\hat{\Lambda}_i$ , ( $i = 1$  la primera vez), por ejemplo por el método del factor principal, se calcula la matriz  $\hat{\psi}_i$  mediante  $\hat{\psi}_i = \text{diag}(S - \hat{\Lambda}_i \hat{\Lambda}'_i)$ . Alternativamente, podemos estimar la matriz  $\hat{\psi}_i$  directamente por el método del factor principal.
2. Se calcula la matriz cuadrada simétrica  $\mathbf{A}_i = \hat{\psi}_i^{-1/2}(S - \hat{\psi}_i)\hat{\psi}_i^{-1/2} = \hat{\psi}_i^{-1/2}S\hat{\psi}_i^{-1/2} - \mathbf{I}$ . Esta matriz pondera los términos de  $S$  por su importancia en términos de los componentes específicos.
3. Se obtiene la descomposición espectral de  $A_i$  de forma que

$$\mathbf{A}_i = \mathbf{H}_{1i}\mathbf{G}_{1i}\mathbf{H}'_{1i} + \mathbf{H}_{2i}\mathbf{G}_{2i}\mathbf{H}'_{2i}$$

Donde los  $m$  mayores valores propios de  $A_i$  están en la matriz diagonal ( $m \times m$ ),  $\mathbf{G}_{1i}$  y los  $p - m$  menores de la  $\mathbf{G}_{2i}$  y  $\mathbf{H}_{1i}$  y  $\mathbf{H}_{2i}$  contienen los correspondientes vectores propios.

---

<sup>16</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

4. Se toma  $\hat{\Lambda}_{i+1} = \hat{\psi}_i^{1/2} H_{1i} \hat{\psi}_i^{1/2}$  y se sustituye en la función de verosimilitud, que se maximiza respecto a  $\psi$ . Esta parte es fácil de hacer con un algoritmo de optimización no lineal. Con el resultado obtenido se vuelve a (2), iterando hasta la convergencia.

Puede ocurrir que este algoritmo converja a un máximo local donde algunos de los términos de la matriz  $\psi$  sean negativos. Esta solución impropia se denomina a veces una solución de Heywood. Los programas existentes cambian entonces esos valores por números positivos e intentan encontrar otro máximo local, aunque no siempre el algoritmo converge.

## **II. CAPITULO**

### **TECNICAS EXPERIMENTALES**

#### **2.1 PROCESO DE RESTAURACION**

La restauración de piezas arqueológicas es una etapa, a veces imprescindible, en el estudio de los objetos encontrados en los yacimientos arqueológicos. En el caso particular de la restauración de objetos metálicos, monedas, joyas, etc., el o los métodos normalmente empleados suelen ser de tipo químico y caracterizado por ser abrasivos y de larga duración. Además, y por la propia naturaleza del procedimiento, sólo suele ser posible el tratamiento pieza a pieza. En esencia, los métodos químicos se basan en la eliminación de la capa de productos oxidados que recubre la pieza. Sin embargo, existe la posibilidad de invertir el proceso de corrosión, causante del deterioro, por métodos electroquímicos.

Es importante mencionar que todo el proceso para la obtención de los datos requeridos para la investigación se lo realizó en las instalaciones del campamento El Mirador de la Empresa Ecuacorriente S. A, ubicado en la Parroquia Tundayme, Cantón El Pangui, Provincia de Zamora Chinchipe.



### **2.1.1 Deterioro de Piezas Metálicas.**

Es de sobra conocido que los objetos metálicos de interés arqueológico suelen encontrarse muy raramente en su estado original, ya que la acción oxidante del medio ambiente provoca que las piezas halladas presenten un estado variable de deterioro.

La presencia de oxígeno y agua en el medio ambiente es causa fundamental de este deterioro provocado por el proceso de corrosión del metal. A esta causa fundamental se añade normalmente la presencia de sales (gran concentración de iones cloruro en las proximidades del mar), la acción de la luz, la presencia de ciertas sustancias oxidantes, de bacterias, etc. Es por ello que el medio ambiente, específico de cada zona, influye extraordinariamente en la velocidad de corrosión del metal o lo que es lo mismo, en el estado de deterioro que presenta la pieza metálica. Como ejemplo claro de lo antedicho, la velocidad de corrosión, expresada en masa perdida como miligramos por decímetro cuadrado por día (mdd), de una pieza de cobre, varía entre 0,1 mdd en atmósfera rural, a 8 mdd en una pieza sumergida en el mar, pasando por el valor de 0,3 mdd en atmósfera marina. (EVANS, 1960, 1968, 1975).



*Fig.6. hacha Moneda corroída*

### **2.1.2 Productos de la corrosión.**

Puesto que el proceso de restauración se basa en la inversión del proceso: metal original -> productos de corrosión, es lógico que el método a aplicar dependa de la composición de estos productos y sea por tanto interesante describir muy brevemente algunos de ellos. Suelen ser:

- **PLATA:** Pueden existir sulfuras de plata como productos de corrosión debido a la presencia de ácido sulfhídrico en la atmósfera como contaminante. Ello provoca un deslustrado de la pieza que adquiere un color oscuro. La restauración de la pieza suele ser sencilla en estos casos. Más difícil se presenta este proceso cuando existen cloruros en el medio, debido a que el mayor tamaño del cristal de cloruro de plata provoca grandes deformaciones en los objetos.

- **COBRE Y SUS ALEACIONES:** Las piezas construidas con este metal o con sus aleaciones, sufren una severa corrosión en presencia de cloruros y nitratos, con formación de sales solubles que producen un picado en la pieza (ANGELUCCI, 1978) cuando ésta se introduce en una disolución acuosa que puede llegar a la destrucción de la misma. Estos aniones, en presencia de  $O_2$  y  $CO_2$ , se transforman lentamente en oxiclорuros y oxicarbonatos de cobre también solubles. Si la pieza es de bronce, los productos que se forman son, aparte de los anteriores, diferentes formas de óxido de estaño. En cualquier caso, la capa verde azulada que recubre la pieza envuelve a otra capa de óxido de cobre que a su vez descansa sobre metal cobre. En casos de corrosión muy severa puede llegar a no existir núcleo metálico.
- **HIERRO:** El hierro se corroe fácilmente conformación de óxidos e hidróxidos de hierro que por su volumen causan deformación en las piezas. La presencia de aniones forma sales de hierro ferroso y férrico. La corrosión de piezas construidas con este metal es tan severa que suele ser muy difícil la restauración.

### **2.1.3 Proceso de Restauración aplicado en Hachas Moneda**

Al momento de empezar con la aplicación de un tratamiento y limpieza de piezas arqueológicas es importante indicar que la limpieza de cualquier objeto conlleva la realización de tres etapas:

- a) Tratamiento previo de la pieza.
- b) Limpieza del objeto.
- c) Consolidación de la pieza.

### ***2.1.3.1 Tratamiento previo de la pieza.***

Antes de empezar con el tratamiento previo se procedió a numerar las hachas, la numeración se dio aleatoriamente sin un orden específico; el formato de numeración es: 1002-1, 1002-2, ..., 1002-194.

En la primera fase se trata de librar a la pieza de los productos y adherencias ajenas a su propia naturaleza. Entre los procedimientos más usuales se encuentra el tratamiento por frotación con cepillo. En el caso específico de limpieza para las hachas en estudio, las materias adheridas fueron eliminadas mediante el uso de bisturíes y punzones de acero. El tratamiento se llevó a cabo bajo la supervisión de expertos, ya que las piezas arqueológicas son sumamente delicadas y siempre se corre el peligro de fractura de la pieza.



*Foto. 1: Limpieza con bisturí de materias adheridas por la corrosión*

### **2.1.3.2 Limpieza de Objetos.**

La segunda fase conlleva la limpieza de los productos de corrosión. Existen distintos procedimientos de limpieza que se emplean según la naturaleza del objeto a restaurar y según la naturaleza de los productos de corrosión.

Mediante la evaluación de el experto en restauración se acordó en utilizar ácido clorhídrico al (HCL) al 10%. Este ácido se utiliza generalmente en metalurgia a veces se utiliza para disolver la capa de óxido que recubre un metal, previo a procesos como galvanizado, extrusión, u otras técnicas y en este caso se lo utilizo con el propósito de remover el óxido que cubría a las hachas.



*Foto 2: Aplicación de HCL al 10%.*

Dejando actuar el ácido por un periodo de 10 a 15 minutos para a continuación implementar finalmente la pulidora que ayudara a retirar en su mayor parte el óxido.



*Foto 3: Muestra pulida para retirar el óxido*

### **2.1.3.3 Consolidación de la pieza.**

Esta fase es la de consolidación de las piezas. Se trata de preservar a éstas de una posterior corrosión y de intentar remediar la fragilidad de la pieza debido al enérgico tratamiento realizado en la fase previa. Una primera etapa fue el secado térmico a fin de eliminar el agua tras el lavado de la pieza. Posteriormente se trata de proteger a la pieza de una ulterior corrosión, sobre todo porque las hachas moneda estudiadas habían pasado por una fuerte corrosión de cloruros. Tras la eliminación de los cloruros la pieza fue protegida mediante el tratamiento con benzotriazol.



**Fig. 7:** Frasco de Benzotriazol



**Foto 4:** Hachas con aplicación de Benzotriazol

## **2.2 APLICACIÓN DE LA MÁQUINA THERMO NITON**

### **2.2.1 Máquina Thermo Niton XL3 XRF.**

Este modelo monta un tubo de rayos X de 50 kV (máximo) y 100 microamperios (máximo) con la opción de incorporar un ánodo de plata para análisis de elementos ligeros. Peso inferior a los 1,3 kg y dimensiones de 240 x 230 x 95 mm. Altas prestaciones y gran velocidad en la detección y medición de concentraciones de elementos contaminantes. CPU interna de 533 MHz. En ocasiones los límites de detección son extremadamente bajos y demandan instrumentación analítica especialmente diseñada para estas aplicaciones.

Este instrumento está sujeto a lo indicado en el Real Decreto sobre Instalaciones Nucleares y Radiactivas.



*Fig. 8: Máquina Thermo Niton XL3 XRF*



### **2.2.2 Análisis de las Hachas Moneda.**

Luego de someter a las piezas arqueológicas al proceso de restauración se procedió a analizarlas con la ayuda de la Máquina Thermo Niton, que como ya se dijo esta máquina mide el porcentaje cada elemento que químico que forma parte de la estructura y composición de las hachas moneda.

La medición es sumamente fácil el tiempo de aplicación en cada pieza es de 90 segundos, el tipo de análisis aplicado para este caso fue el de Mining (minería), la radiación que emite este aparato no es nocivo y por esta razón no fue necesario utilizar vestimenta para radiación.



*Foto 5: Análisis con la Thermo Niton*

### 2.3 OBTENCION DE DATOS

Finalmente luego del análisis con la Thermo Niton a las 194 piezas arqueológicas se obtuvo un total de 29 elementos químicos hallados en la composición de las hachas moneda de cobre, el software incluido en dicha máquina posee la aplicación para importar directamente al Excel la tabla presentada en la fig. 2.4.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194
Reading No	Time	Type	Duration	Units	Sigma Value	Sequence	Flags	SAMPLE	LOCATION	INSPECTOR	MISC	NOTE	Ba	Ba Error	Sb	Sb Error	Sn	Sn Error	Cd	Cd Error	Pd	Pd Error	A																																																																																																																																																																										
44	20/09/2011 13:04	Mining	90 %		4	Final		Hacha pulida 1002-1	Mirador	Angelica V		<LOD	0,067 <LOD	0,022 <LOD	0,024 <LOD	0,009 <LOD	0,005																																																																																																																																																																																
49	21/09/2011 8:40	Mining	90 %		4	Final		muestra pulida 1002-2	Mirador	Angelica V		<LOD	0,072 <LOD	0,022 <LOD	0,023 <LOD	0,01 <LOD	0,005																																																																																																																																																																																
50	21/09/2011 8:57	Mining	90 %		4	Final		muestra pulida 1002-3	Mirador	Angelica V		<LOD	0,067 <LOD	0,02 <LOD	0,021 <LOD	0,009 <LOD	0,005																																																																																																																																																																																
51	21/09/2011 9:01	Mining	90 %		4	Final		muestra pulida 1002-4	Mirador	Angelica V		<LOD	0,083 <LOD	0,028 <LOD	0,026 <LOD	0,01 <LOD	0,006																																																																																																																																																																																
52	21/09/2011 9:14	Mining	90 %		4	Final		muestra pulida 1002-5	Mirador	Angelica V		<LOD	0,072 <LOD	0,022 <LOD	0,028 <LOD	0,014 <LOD	0,01 <LOD	0,006																																																																																																																																																																															
53	21/09/2011 9:24	Mining	90 %		4	Final		muestra pulida 1002-6	Mirador	Angelica V		<LOD	0,073 <LOD	0,032 <LOD	0,013 <LOD	0,025 <LOD	0,009 <LOD	0,006																																																																																																																																																																															
54	21/09/2011 9:47	Mining	90 %		4	Final		muestra pulida 1002-7	Mirador	Angelica V		<LOD	0,07 <LOD	0,085 <LOD	0,017 <LOD	0,022 <LOD	0,009 <LOD	0,005																																																																																																																																																																															
56	21/09/2011 10:42	Mining	90 %		4	Final		hacha pulida 1002-8	Mirador	Angelica V		<LOD	0,068 <LOD	0,02 <LOD	0,022 <LOD	0,009 <LOD	0,005																																																																																																																																																																																
57	21/09/2011 10:46	Mining	90 %		4	Final		hacha pulida 1002-9	Mirador	Angelica V		<LOD	0,072 <LOD	0,022 <LOD	0,025 <LOD	0,009 <LOD	0,004																																																																																																																																																																																
58	21/09/2011 10:52	Mining	90 %		4	Final		hacha pulida 1002-10	Mirador	Angelica V		<LOD	0,067 <LOD	0,02 <LOD	0,021 <LOD	0,009 <LOD	0,005																																																																																																																																																																																
59	21/09/2011 10:55	Mining	90 %		4	Final		hacha pulida 1002-11	Mirador	Angelica V		<LOD	0,07 <LOD	0,022 <LOD	0,022 <LOD	0,01 <LOD	0,005																																																																																																																																																																																
60	21/09/2011 11:10	Mining	90 %		4	Final		hacha pulida 1002-12	Mirador	Angelica V		<LOD	0,071 <LOD	0,021 <LOD	0,045 <LOD	0,015 <LOD	0,01 <LOD	0,006																																																																																																																																																																															
61	21/09/2011 11:22	Mining	90 %		4	Final		muestra pulida 1002-13	Mirador	Angelica V		<LOD	0,073 <LOD	0,024 <LOD	0,012 <LOD	0,024 <LOD	0,01 <LOD	0,006																																																																																																																																																																															
62	21/09/2011 11:25	Mining	90 %		4	Final		muestra pulida 1002-14	Mirador	Angelica V		<LOD	0,073 <LOD	0,04 <LOD	0,014 <LOD	0,026 <LOD	0,011 <LOD	0,007																																																																																																																																																																															
63	21/09/2011 11:28	Mining	90 %		4	Final		muestra pulida 1002-15	Mirador	Angelica V		<LOD	0,072 <LOD	0,022 <LOD	0,031 <LOD	0,013 <LOD	0,009 <LOD	0,006																																																																																																																																																																															
64	21/09/2011 11:32	Mining	90 %		4	Final		muestra pulida 1002-16	Mirador	Angelica V		<LOD	0,069 <LOD	0,023 <LOD	0,022 <LOD	0,01 <LOD	0,005																																																																																																																																																																																
65	21/09/2011 11:38	Mining	90 %		4	Final		muestra pulida 1002-17	Mirador	Angelica V		<LOD	0,066 <LOD	0,02 <LOD	0,022 <LOD	0,009 <LOD	0,005																																																																																																																																																																																
66	21/09/2011 11:42	Mining	90 %		4	Final		muestra pulida 1002-18	Mirador	Angelica V		<LOD	0,071 <LOD	0,021 <LOD	0,023 <LOD	0,01 <LOD	0,005																																																																																																																																																																																
67	21/09/2011 12:06	Mining	90 %		4	Final		muestra pulida 1002-19	Mirador	Angelica V		<LOD	0,074 <LOD	0,021 <LOD	0,024 <LOD	0,01 <LOD	0,005																																																																																																																																																																																
68	21/09/2011 12:15	Mining	90 %		4	Final		muestra pulida 1002-20	Mirador	Angelica V		<LOD	0,074 <LOD	0,023 <LOD	0,026 <LOD	0,009 <LOD	0,006																																																																																																																																																																																
69	21/09/2011 12:25	Mining	90 %		4	Final		muestra pulida 1002-21	Mirador	Angelica V		<LOD	0,068 <LOD	0,021 <LOD	0,021 <LOD	0,011 <LOD	0,005																																																																																																																																																																																

Fig. 9: Datos obtenidos con la Thermo Niton

Luego de consultar con un Geólogo sobre los elementos químicos que generalmente se hallan en la extracción del cobre se recomienda trabajar con 15 elementos y finalmente se obtiene la matriz con la que se trabajara en adelante para el estudio estadístico respectivo.

### **III. CAPITULO**

## **APLICACIÓN DE LAS TECNICAS DE ANALISIS ESTADISTICO MULTIVARIADO.**

### **3.1 DATOS.**

Como se dijo en el capitulo anterior, la base de datos con la que se trabajara en adelante son los datos tomados con la máquina Thermo Niton la misma que detecto un total de 29 elementos químicos, pero luego de consultar con experto geólogo se recomienda trabajar únicamente con 15 de estos ya que son los elementos que generalmente están asociados con la extracción del cobre, los elementos que fueron descartados se debió a que estos elementos químicos pudieron ser adheridos a las piezas por su exposición al medio ambiente.

#### **3.1.1 Hachas Moneda (Individuos).**

Objetos arqueológicos ecuatorianos realizados en lámina de cobre con un perfil en forma de hacha, pero con los bordes levantados. Por sus características utilitarias, las hachas monedas fueron consideradas asimilables a una moneda. Sus particularidades hacen ver que fueron objetos mercantilistas intercambiables pues

reúnen los requisitos de: útil, incrementable, portátil, indestructible, homogénea, divisible, estable y cognoscible.

No se manufacturaron solamente en el contexto del actual territorio ecuatoriano; evidencias encontradas en México hacen suponer de una posible universalidad en el mundo americano de este objeto-mercancía, con anterioridad a la llegada de los conquistadores españoles. El número de hachas que se analizara es 194.

En el Ecuador, se han realizado hallazgos de hachas monedas en las Provincias del Oro, Manabí, Guayas (Santa Elena, Anyulla, Cangrejillo), Los Ríos (Las Palmas), donde se asentaron las Culturas Manteño Huancavilca y Milagro Quevedo (400 a 1500 d.C.). Encontradas en fardos funerarios, en paquetes de 20 hachas miniaturas, número repetitivo en estos hallazgos. Variaban de 15gr a 105gr, dependiendo a la cultura a la que pertenecían su tamaño aproximado es de 7cm de ancho por 12,6cm de alto y cuyo espesor oscila entre 3mm y 5mm.

### **3.1.2 Elementos Químicos (Variables).**

A continuación se detalla cada una de las 15 variables a ser analizadas.

**Antimonio (Sb):** El antimonio no es un elemento abundante en la naturaleza; raras veces se encuentra en forma natural, a menudo como una mezcla isomorfa con arsénico. El antimonio se encuentra principalmente en la naturaleza como (estibnita, antimonita). Forma parte por lo general de los minerales de cobre, plata y plomo. El

antimonio se puede encontrar en los suelos, agua y aire en cantidades muy pequeñas. El antimonio contamina principalmente los suelos. Puede viajar grandes distancias con las aguas subterráneas hacia otros lugares y aguas superficiales.

**Estaño (Sn):** Se funde a baja temperatura; tiene gran fluidez cuando se funde y posee un punto de ebullición alto. Es suave, flexible y resistente a la corrosión en muchos medios. Una aplicación importante es el recubrimiento de envases de acero para conservar alimentos y bebidas. Otros empleos importantes son: aleaciones para soldar, bronces, pletres y aleaciones industriales diversas. Los productos químicos de estaño, tanto inorgánicos como orgánicos, se utilizan mucho en las industrias de galvanoplastia, cerámica y plásticos, y en la agricultura.

**Plata (Ag):** En la mayor parte de sus aplicaciones, la plata se alea con uno o más metales. La plata, que posee las más altas conductividades térmica y eléctrica de todos los metales, se utiliza en puntos de contacto eléctrico y electrónico. También se emplea mucho en joyería y piezas diversas.

La plata es un elemento bastante escaso. Algunas veces se encuentra en la naturaleza como elemento libre (plata nativa) o mezclada con otros metales. Sin embargo, la mayor parte de las veces se encuentra en minerales que contienen compuestos de plata. Aproximadamente tres cuartas partes de la plata producida son un subproducto de la extracción de otros minerales, sobre todo de cobre y de plomo.

**Arsénico (As):** Al arsénico se le encuentra natural como mineral de cobalto, aunque por lo general está en la superficie de las rocas combinado con azufre o metales como Mn, Fe, Co, Ni, Ag o Sn. El óxido, arsenolita, se encuentra como producto de la alteración debida a los agentes atmosféricos de otros minerales de arsénico, y también se recupera de los polvos colectados de los conductos durante la extracción de Ni, Cu y Sn.

**Selenio (Se):** La abundancia de este elemento, ampliamente distribuido en la corteza terrestre, se estima aproximadamente en  $7 \times 10^{-5}\%$  por peso, encontrándose en forma de seleniuros de elementos pesados y, en menor cantidad, como elemento libre en asociación con azufre elemental. Sus minerales no se encuentran en suficiente cantidad para tener utilidad, como fuente comercial del elemento, y por ello los minerales de sulfuro de cobre seleníferos son los que representan la fuente primaria.

**Oro (Au):** El oro se clasifica como metal pesado y noble; en el comercio es el más común de los metales preciosos. El cobre, la plata y el oro están en el mismo grupo en la tabla periódica. La fuente del símbolo químico, Au, es su nombre en latín aurum (amanecer radiante).

El oro metálico, o natural, y varios minerales telúricos son las únicas formas de oro presentes en la Tierra. El oro natural existe en las rocas y minerales de otros metales, especialmente en el cuarzo y la pirita, o puede estar disperso en arenas y gravas (oro de aluvión).

**Plomo (Pb):** Es flexible, inelástico, se funde con facilidad, es relativamente resistente al ataque de los ácidos sulfúrico y clorhídrico. Pero se disuelve con lentitud en ácido nítrico. El plomo es anfótero, ya que forma sales de plomo de los ácidos, así como sales metálicas del ácido plúmbico. El plomo forma aleaciones con muchos metales y, en general, se emplea en esta forma en la mayor parte de sus aplicaciones. Todas las aleaciones formadas con estaño, cobre, arsénico, antimonio, bismuto, cadmio y sodio tienen importancia industrial.

**Cobre (Cu):** Uno de los metales de transición e importante metal no ferroso. Su utilidad se debe a la combinación de sus propiedades químicas, físicas y mecánicas, así como a sus propiedades eléctricas y su abundancia. El cobre fue uno de los primeros metales usados por los humanos.

Es uno de los metales que puede tenerse en estado más puro, es moderadamente duro, es tenaz en extremo y resistente al desgaste.

**Níquel (Ni):** metal duro, blanco plateado, dúctil y maleable. La mayor parte del níquel comercial se emplea en el acero inoxidable y otras aleaciones resistentes a la corrosión. También es importante en monedas como sustituto de la plata. El níquel finamente dividido se emplea como catalizador de hidrogenación.

El níquel es un elemento bastante abundante, constituye cerca de 0.008% de la corteza terrestre y 0.01% de las rocas ígneas. En algunos tipos de meteoritos hay

cantidades apreciables de níquel, y se piensa que existen grandes cantidades en el núcleo terrestre. Está presente en pequeñas cantidades en el agua de mar, el petróleo y en la mayor parte del carbón.

**Hierro (Fe):** El hierro es el cuarto elemento más abundante en la corteza terrestre (5%). Es un metal maleable, tenaz, de color gris plateado y magnético. El hierro se encuentra en muchos otros minerales y está presente en las aguas freáticas y en la hemoglobina roja de la sangre.

El uso más extenso del hierro (fierro) es para la obtención de aceros estructurales; también se producen grandes cantidades de hierro fundido y de hierro forjado.

**Titanio (Ti):** Mientras que su comportamiento químico muestra muchas semejanzas con el del silicio y el zirconio, como un elemento del primer grupo de transición, la química de la solución acuosa, especialmente de los estados de oxidación más bajos, tiene algunas semejanzas con la del cromo y el vanadio.

**Bismuto (Bi):** Se estima que la corteza terrestre contiene cerca de 0.00002% de bismuto. Existe en la naturaleza como metal libre y en minerales. Los principales depósitos están en Sudamérica, pero en Estados Unidos se obtiene principalmente como subproducto del refinado de los minerales de cobre y plomo.

**Calcio (Ca):** Es el tercer metal más abundante en la corteza terrestre. Los compuestos de calcio constituyen 3.64% de la corteza terrestre. El metal es trimorfo, más duro que



el sodio, pero más blando que el aluminio. La distribución del calcio es muy amplia; se encuentra en casi todas las áreas terrestres del mundo. Este elemento es esencial para la vida de las plantas y animales, ya que está presente en el esqueleto de los animales, en los dientes, en la cáscara de los huevos, en el coral y en muchos suelos.

**Potasio (K):** Este metal reactivo es ligero y blando. Se parece mucho al sodio en su comportamiento en forma metálica.

El potasio es un elemento muy abundante y es el séptimo entre todos los elementos de la corteza terrestre; el 2.59% de ella corresponde a potasio en forma combinada. El agua de mar contiene 380 ppm, lo cual significa que el potasio es el sexto más abundante en solución.

**Bal:** Esta variable es el conjunto de elementos químicos detectados por la máquina Thermo Niton cuyo número atómico es inferior a 14, estos son el Hidrogeno, Helio, Litio, Berilio, Boro, Carbono, Nitrógeno, Oxígeno, Flúor, Neón, Sodio, Magnesio, Aluminio, Sílice; estos elementos se encuentran presentes dentro de la composición química de las monedas pero sus porcentajes son mínimos, por esta razón la Niton los agrupa en una sola variable.

### **3.1.3 Matriz de Datos.**

A continuación se muestra la matriz de datos inicial.

TABLA I: MATRIZ DE DATOS INICIAL															
Hacha	Sb	Sn	Ag	Bal	Bi	As	Se	Au	Pb	Cu	Ni	Fe	Ti	Ca	K
1	0	0	0,024	18,323	0	0,664	0	0	0,077	80,794	0	0	0	0	0
2	0	0	0,019	17,216	0	0,377	0	0	0,151	82,058	0	0	0	0	0
3	0	0	0,022	25,117	0	0,772	0	0	0,217	73,156	0	0,425	0,023	0,077	0,089
4	0	0	0,039	11,746	0,029	0,722	0	0	0,146	87,073	0	0	0,016	0,074	0,071
5	0	0,028	0,028	15,891	0	0,349	0	0	0,033	83,547	0	0	0	0	0
6	0,032	0	0,029	17,585	0	0,373	0	0	0	81,742	0	0,072	0	0	0
7	0,085	0	0,027	18,429	0,021	0,337	0	0	0	80,909	0,053	0	0	0	0
8	0	0	0,035	19,398	0	0,319	0	0	0,074	79,927	0	0,118	0	0	0
9	0	0	0,019	18,009	0	0,187	0	0	1,183	80,459	0	0,064	0	0	0
10	0	0	0,032	18,542	0	0,256	0	0	0,107	80,725	0	0,183	0	0	0
11	0	0	0,033	17,931	0	0,186	0	0	0,234	81,522	0	0,035	0	0	0
12	0	0,045	0,036	17,441	0,026	0,251	0	0	0,038	82,006	0	0,043	0	0	0
13	0,024	0	0,03	16,641	0	0,301	0	0	0,075	82,651	0	0,092	0	0	0
14	0,04	0	0,03	16,13	0	0,416	0	0	1,18	82,079	0	0	0	0	0
15	0	0,031	0,017	17,303	0	0,087	0	0	0,247	82,219	0	0	0	0	0
16	0	0	0,027	17,909	0	0,257	0	0	0,777	80,885	0	0	0	0	0
17	0	0	0,03	18,986	0	0,196	0	0	0,135	80,497	0	0,038	0	0	0
18	0	0	0,037	20,352	0	0,18	0	0	0,055	79,24	0	0	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
19	0	0	0,023	16,139	0	0,578	0	0	1,014	82,113	0	0	0	0	0
20	0	0	0,02	16,585	0	0,461	0	0	0,037	82,716	0	0	0	0	0
21	0	0	0,031	17,771	0	0,253	0	0	0,292	81,545	0	0	0	0	0
22	0,033	0	0,029	14,619	0	0,357	0	0	0,351	84,534	0	0	0	0	0
23	0,038	0	0,029	16,891	0	0,413	0	0	0,069	82,429	0,047	0	0	0	0
24	0	0	0,039	17,758	0	0,455	0	0	0	81,562	0	0,031	0,013	0	0
25	0	0	0,023	18,453	0	0,245	0	0	0,235	80,916	0	0	0	0	0
26	0,027	0	0,028	17,486	0	0,402	0	0	0,84	81,082	0	0,038	0	0	0
27	0	0	0,024	16,374	0	0,298	0	0	0,628	82,535	0	0	0	0	0
28	0,072	0	0,018	18,141	0	0,26	0	0	0,855	80,453	0	0,037	0	0,077	0
29	0	0	0,017	19,215	0	0,255	0	0	0,058	80,303	0	0,042	0	0	0
30	0	0	0,028	20,181	0	0,115	0	0	0,145	79,357	0	0,037	0	0	0
31	0,024	0,031	0,026	16,286	0	0,239	0	0	0,676	82,615	0	0	0	0	0
32	0,03	0	0,028	19,394	0	0,328	0	0	0,616	79,502	0	0	0,014	0	0
33	0,027	0	0,046	20,187	0,049	0,229	0	0	0,034	79,275	0	0,066	0	0	0
34	0,03	0	0,024	18,635	0	0,679	0	0	0,258	79,746	0	0,491	0	0	0
35	0	0	0,052	18,928	0	0,28	0,015	0	0,65	79,907	0	0,035	0	0	0
36	0	0	0,014	22,497	0	0,346	0	0	0,228	75,812	0	0,785	0,013	0,143	0
37	0	0	0,027	19,337	0	0,365	0	0	0,339	79,778	0	0,049	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
38	0,038	0	0,027	19,536	0	0,266	0	0	0,105	79,663	0	0,173	0	0,067	0
39	0	0	0,017	18,057	0,025	0,163	0	0	0	81,56	0	0,038	0	0	0
40	0,069	0	0,028	15,624	0	0,593	0	0	0,041	83,475	0,06	0	0	0	0
41	0	0	0,032	16,889	0	0,49	0	0	0,412	82,034	0	0	0	0	0
42	0,027	0	0,019	18,879	0	0,417	0	0	0,285	80,257	0	0	0	0	0
43	0,028	0	0,025	15,399	0	0,27	0	0	0,723	83,427	0	0	0,016	0	0
44	0	0	0,042	14,782	0	0,497	0	0	0,302	84,191	0	0,088	0	0	0
45	0	0	0,029	16,561	0	0,399	0	0	0,155	82,641	0,05	0,103	0	0	0
46	0	0	0,032	18,702	0	0,194	0	0	0,053	80,849	0	0,04	0	0	0
47	0	0	0,023	22,411	0	0,529	0	0	0,056	76,502	0,044	0,311	0	0	0
48	0	0	0,04	19,406	0,02	0,191	0	0	0	80,118	0	0,095	0	0	0
49	0	0,029	0,024	19,312	0	0,211	0	0	0,793	79,538	0	0	0	0	0
50	0,026	0	0,037	18,965	0	0,181	0	0	0,557	79,998	0	0,094	0	0	0
51	0	0	0,035	17,583	0	0,275	0	0	0,883	81,072	0	0,061	0	0	0
52	0	0	0,031	20,633	0,023	0,382	0	0	0,114	78,705	0	0,035	0	0	0
53	0,031	0	0,032	19,785	0	0,176	0	0	0,082	79,741	0,045	0,039	0	0	0
54	0	0,035	0,019	20,531	0	0,319	0	0	0,102	78,624	0	0,15	0,012	0	0,059
55	0,046	0	0,024	19,034	0	0,333	0	0	0,294	80,144	0	0,033	0	0	0
56	0,047	0	0,017	14,742	0	0,696	0	0	0,569	83,796	0	0	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
57	0,032	0	0,03	20,144	0	0,35	0	0	0	79,224	0,053	0,075	0	0	0
58	0	0	0,024	19,473	0	0,346	0	0	0,364	79,618	0	0,036	0	0	0
59	0	0	0,033	19,98	0	0,202	0	0	0,74	78,864	0	0,089	0	0	0
60	0	0	0,026	19,901	0	0,463	0	0	0,14	79,258	0	0,049	0	0	0
61	0	0	0,033	21,311	0	0,213	0	0	0,031	78,172	0	0,125	0	0	0
62	0	0	0,046	19,25	0	0,466	0	0	0,405	78,008	0	1,332	0,028	0,22	0,081
63	0	0	0,042	20,551	0	0,263	0	0	0,07	78,923	0,041	0,033	0	0	0
64	0	0	0,057	18,92	0	0,311	0	0	0,158	80,41	0	0	0	0	0
65	0,03	0	0,021	19,198	0	0,143	0	0	0,668	79,766	0	0,068	0	0	0
66	0	0	0,017	22,349	0	0,244	0	0	0,158	76,73	0	0,278	0	0	0,071
67	0	0	0,021	18,312	0	0,232	0	0	0,188	81,103	0	0	0	0	0
68	0,038	0	0,035	16,023	0	1,308	0	0	0	82,33	0,112	0,043	0	0	0
69	0	0,029	0,027	18,928	0	0,264	0	0	0,559	79,99	0	0,084	0	0	0
70	0,029	0	0,031	19,97	0	0,265	0	0	0,403	79,174	0	0,042	0	0	0
71	0,025	0	0	17,625	0	0,165	0	0	0,32	81,717	0	0,048	0,017	0	0
72	0,024	0	0,028	18,112	0	0,492	0	0	0,034	81,132	0,049	0,06	0	0	0
73	0	0	0,043	17,02	0	0,312	0	0	1,258	81,259	0	0,064	0	0	0
74	0	0	0,016	19,356	0	0,238	0	0	0,888	79,094	0	0,26	0	0	0
75	0,029	0	0,015	19,757	0	0,233	0	0	0	79,521	0	0,265	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
76	0,023	0	0,023	21,025	0	0,541	0	0	0,226	77,771	0	0,242	0	0	0
77	0	0	0,053	18,482	0	0,275	0	0	0,784	80,215	0	0,066	0	0	0
78	0,033	0	0,031	17,569	0	0,435	0	0	0	81,747	0	0,075	0	0	0
79	0	0	0,018	20,859	0	0,398	0	0	0,107	78,314	0	0,109	0	0	0
80	0	0	0,033	17,613	0	0,279	0	0	0	81,911	0	0,047	0	0	0
81	0,037	0	0,014	20,237	0	0,28	0	0	0,066	79,042	0	0,193	0,013	0	0
82	0	0	0,027	18,851	0	0,165	0	0	0,038	80,722	0	0,065	0	0	0
83	0	0	0,028	20,277	0	0,233	0	0	0,188	79,122	0	0,047	0	0	0
84	0	0	0,034	20,733	0	0,246	0	0	0	78,812	0,04	0,038	0	0	0
85	0,026	0,07	0,022	20,297	0	0,18	0	0	0	79,167	0	0,117	0	0	0
86	0,025	0	0,025	21,879	0	0,27	0	0	0	77,399	0	0,233	0	0	0
87	0	0	0,033	18,028	0	0,252	0	0	0,163	81,319	0	0,083	0	0	0
88	0	0	0,033	18,702	0	0,364	0	0	1,015	79,723	0	0,04	0	0	0
89	0,031	0	0,011	15,764	0	0,318	0	0	0,34	83,326	0,074	0,075	0	0	0
90	0	0	0,02	20,398	0	0,217	0	0	0,026	78,599	0,044	0,427	0,019	0,095	0
91	0	0	0,027	16,359	0	0,303	0	0	0,119	82,866	0	0,19	0	0	0
92	0	0	0,018	21,679	0	0,279	0	0	0,037	77,752	0	0,072	0	0	0
93	0	0	0,029	20,899	0,049	0,368	0	0	0,055	78,434	0	0	0	0	0
94	0,025	0	0,044	18,924	0	0,314	0	0	0,259	80,247	0	0,088	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
95	0	0	0,023	18,13	0	0,213	0	0	0,072	81,334	0	0,067	0	0	0
96	0	0,034	0,027	17,584	0	0,328	0	0	1,082	79,099	0	1,494	0,016	0,162	0
97	0	0	0,019	23,694	0	0,3	0	0	0	75,778	0	0,073	0	0	0
98	0	0	0,033	18,786	0	0,646	0	0	0,108	80,062	0	0,174	0	0	0
99	0	0	0,048	21,225	0	0,241	0	0	0	78,23	0	0,133	0	0	0
100	0	0	0,023	21,015	0	0,367	0	0	0,141	78,065	0	0,175	0,019	0	0
101	0	0	0,035	18,592	0	0,131	0	0	0	81,027	0	0,083	0	0	0
102	0	0	0,018	24,501	0	0,126	0	0	0,409	74,283	0	0,3	0	0,23	0
103	0,101	0,046	0	18,937	0	0,354	0	0	0,376	80,061	0	0,033	0	0	0
104	0,026	0	0,024	21,491	0	0,196	0	0	0,286	77,797	0	0,086	0	0	0
105	0,044	0	0,022	21,892	0	0,46	0	0	0,669	76,491	0	0,258	0	0	0
106	0,096	0	0,033	21,533	0	0,2	0	0	0,262	77,638	0	0,108	0	0	0
107	0	0	0,021	20,652	0	0,27	0	0	0,591	78,225	0	0,103	0	0	0
108	0	0	0,021	20,151	0	0,289	0	0	0,054	78,254	0	0,783	0,019	0,211	0
109	0	0	0,024	17,578	0	0,408	0	0	0,217	81,55	0	0,118	0	0	0
110	0,024	0	0,035	18,316	0	0,362	0	0	0	81,099	0	0,078	0	0	0
111	0	0	0,031	18,969	0	0,116	0	0	0,043	80,658	0	0,069	0	0	0
112	0,034	0	0,037	17,637	0	0,33	0	0	0,09	81,681	0	0,04	0	0	0
113	0	0	0,022	16,608	0	0,328	0	0	0,039	82,737	0	0,147	0	0	0

<b>TABLA I: MATRIZ DE DATOS INICIAL (Continuación)</b>															
<b>114</b>	0	0	0,031	18,387	0,036	0,231	0	0	0,146	80,921	0	0,116	0	0	0
<b>115</b>	0	0	0,051	16,408	0	0,232	0	0	0,035	83,1	0	0,067	0	0	0
<b>116</b>	0,026	0,031	0,025	16,866	0	0,301	0	0	0,733	81,887	0	0,034	0	0	0
<b>117</b>	0	0	0,031	18,001	0	0,26	0	0	0,537	80,939	0	0,116	0	0	0
<b>118</b>	0	0	0,026	19,685	0	0,464	0	0	0,137	79,427	0	0,078	0	0	0
<b>119</b>	0,033	0	0,031	16,401	0	0,291	0	0	0,105	82,974	0	0,068	0	0	0
<b>120</b>	0,06	0	0,019	17,819	0	0,407	0	0	0,166	81,353	0	0,055	0	0	0
<b>121</b>	0,027	0,029	0,062	25,995	0	0,24	0	0	0,243	73,086	0	0,114	0,024	0	0
<b>122</b>	0	0	0,029	21,534	0	0,21	0	0	0,154	77,875	0	0,061	0	0	0
<b>123</b>	0	0	0,025	17,375	0	0,341	0	0	0,242	81,765	0	0,124	0	0	0
<b>124</b>	0	0	0,037	20,467	0	0,297	0	0	0,503	78,518	0	0,073	0	0	0
<b>125</b>	0	0	0,045	19,566	0,042	0,316	0	0	0	79,833	0	0,062	0	0	0
<b>126</b>	0	0	0,02	21,106	0	0,309	0	0	0,405	77,963	0	0,066	0	0	0
<b>127</b>	0,033	0	0,024	16,279	0	0,266	0	0	0,032	83,102	0	0,111	0	0	0
<b>128</b>	0,043	0	0,015	17,661	0	0,34	0	0	0,36	81,373	0,046	0,116	0	0	0
<b>129</b>	0	0	0,032	19,581	0	0,269	0	0	0,245	79,697	0	0,068	0	0	0
<b>130</b>	0,03	0	0,02	19,011	0	0,283	0	0	0,158	80,31	0	0,104	0	0	0
<b>131</b>	0	0	0,028	20,197	0	0,444	0	0	0,311	78,861	0	0	0	0	0
<b>132</b>	0	0	0,045	18,841	0	0,259	0	0	1,048	79,336	0	0,27	0,012	0	0



TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
133	0,024	0	0,02	22,345	0	0,28	0	0	0,247	76,974	0	0,032	0	0	0
134	0	0	0,032	17,351	0	0,338	0	0	0,28	81,812	0	0,073	0	0	0
135	0,026	0	0,03	23,563	0	0,356	0	0	0	75,892	0	0	0	0	0
136	0,043	0	0,079	23,085	0,042	0,343	0	0	0	76,218	0,045	0,065	0	0	0
137	0	0	0,028	19,49	0	0,113	0	0	0,071	79,886	0	0,2	0,013	0	0
138	0	0	0,029	23,641	0	0,28	0	0	0,252	75,273	0	0,318	0	0,077	0
139	0,025	0	0,029	15,447	0	0,381	0	0	0,039	83,941	0,045	0	0	0	0
140	0	0	0,036	20,215	0	0,507	0	0	0,311	78,596	0	0,12	0,015	0	0
141	0,025	0,026	0,02	19,324	0	0,144	0	0	0,025	80,259	0	0,082	0	0	0
142	0	0	0,033	20,496	0	0,189	0	0	0,782	78,35	0	0,056	0	0	0
143	0	0	0,039	17,817	0	0,244	0	0	0	81,787	0	0	0	0	0
144	0	0	0,039	15,762	0	0,238	0	0	0,473	83,311	0	0,07	0	0	0
145	0,051	0	0,037	17,654	0	0,376	0	0	0,683	81,117	0	0	0	0	0
146	0	0	0,017	19,024	0	0,151	0	0	0,374	80,312	0	0	0	0	0
147	0	0	0,03	18,925	0,039	0,312	0	0	0,172	80,383	0	0	0	0	0
148	0	0	0,022	19,977	0	0,309	0	0	0,21	79,172	0	0,121	0,016	0	0
149	0	0,03	0,038	24,702	0	0,108	0	0	0,058	74,901	0,043	0	0	0	0
150	0	0	0,026	21,718	0	0,217	0	0	0,068	77,818	0	0	0	0	0
151	0	0	0,05	16,894	0,152	1,026	0	0	0,138	81,492	0,044	0,054	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL (Continuación)															
152	0,038	0	0,051	21,518	0,042	0,345	0	0	0,238	77,603	0	0,033	0	0	0
153	0	0	0,026	17,906	0	0,19	0	0	0,277	81,43	0	0,032	0,016	0	0
154	0	0	0,043	22,397	0	0,306	0	0	0,333	76,676	0	0,046	0	0,084	0
155	0	0	0,016	21,454	0	0,257	0	0	0,079	77,956	0	0,089	0	0	0
156	0,039	0	0,051	21,902	0,045	0,228	0	0	0,074	77,499	0	0,066	0	0	0
157	0	0	0,029	19,157	0	0,177	0	0	0,18	80,297	0	0,046	0	0	0
158	0	0	0,036	21,816	0,027	0,347	0	0	0	77,505	0	0,078	0	0	0
159	0	0	0,016	17,95	0	0,181	0	0	0,377	80,939	0	0,297	0,015	0,083	0
160	0,036	0	0,032	19,327	0,022	0,222	0	0	0,062	80,158	0	0,039	0	0	0
161	0	0	0,022	19,701	0	0,23	0	0	0,752	79,197	0	0	0	0	0
162	0	0	0,04	17,093	0	0,356	0	0	0,111	82,243	0	0,031	0	0	0
163	0,034	0	0,032	17,239	0	0,355	0	0	0,065	82,072	0	0	0	0	0
164	0	0	0,024	17,162	0	0,426	0	0	0,405	81,634	0	0,109	0,013	0,092	0
165	0	0	0,031	18,431	0	0,401	0	0	0,304	80,62	0	0,102	0	0	0
166	0	0	0,037	18,247	0	0,245	0	0	0,696	80,649	0	0	0	0	0
167	0	0	0,06	14,691	0,037	0,878	0	0	0,057	84,082	0	0,055	0	0	0
168	0,028	0	0,029	18,518	0	0,266	0	0	0,612	80,434	0	0,032	0	0	0
169	0	0	0,031	17,68	0	0,303	0	0	0,129	81,557	0	0,148	0	0	0
170	0	0	0,018	16,362	0	0,23	0	0	0,331	82,902	0	0,04	0	0	0

TABLA I: MATRIZ DE DATOS INICIAL(Continuación)															
171	0,028	0	0,023	19,213	0	0,453	0	0	0	80,127	0	0	0	0	0
172	0	0	0,029	18,094	0	0,422	0	0	0	81,232	0	0,067	0	0	0
173	0	0	0,026	20,749	0	0,4	0	0	0,522	77,22	0	0,707	0,027	0,182	0,071
174	0	0	0,035	20,962	0,018	0,316	0	0	0,04	78,43	0	0,082	0	0	0
175	0	0	0,021	14,984	0	0,392	0	0	0	84,401	0	0,068	0	0	0
176	0	0	0,035	19,53	0	0,327	0	0	0,082	79,858	0	0,07	0	0	0
177	0	0	0,025	20,309	0	0,386	0	0	0,437	78,282	0	0,374	0,014	0	0
178	0	0	0,045	18,598	0	0,418	0	0	0,472	80,244	0	0,079	0	0	0
179	0,043	0	0,026	22,898	0	0,279	0	0	0	76,487	0	0,106	0	0	0
180	0	0	0,033	18,38	0	0,447	0	0	0	80,902	0	0,111	0	0	0
181	0	0	0,043	17,644	0	0,182	0	0	0	81,856	0	0,135	0	0	0
182	0	0	0,032	18,629	0	0,16	0	0	0,063	80,824	0	0,117	0	0	0
183	0,062	0	0,027	19,997	0	0,147	0	0	0,064	79,419	0	0,126	0	0	0
184	0	0	0,038	18,9	0	0,506	0	0	0	80,302	0	0,083	0	0	0
185	0	0,048	0,027	19,192	0	0,295	0	0	0,085	79,911	0	0,245	0	0	0
186	0	0	0,077	18,247	0,064	0,435	0	0	0,471	80,519	0	0,069	0	0	0
187	0	0	0,028	18,307	0	0,217	0	0	0	81,235	0	0,054	0	0	0
188	0	0	0,029	20,851	0	0,182	0	0	0,03	78,679	0	0,115	0	0	0
189	0	0	0,016	33,906	0	0,233	0	0	0,048	64,713	0	0,853	0	0,088	0

<b>TABLA I: MATRIZ DE DATOS INICIAL(Continuación)</b>															
<b>190</b>	0	0	0,027	23,155	0	0,408	0	0	0,088	76,159	0	0,06	0	0	0
<b>191</b>	0	0	0,027	17,951	0	0,195	0	0	0	81,553	0	0,138	0	0	0
<b>192</b>	0	0,034	0,033	18,997	0	0,212	0	0	0,122	80,364	0	0,115	0	0	0
<b>193</b>	0,058	0	0,017	19,176	0	0,213	0	0	0	80,364	0	0,04	0	0	0
<b>194</b>	0,029	0	0,025	18,056	0	0,242	0	0	0,184	81,292	0	0,072	0	0	0

Fuente: Datos obtenidos con la máquina Thermo Niton XL3 XRF

## **3.2 ANALISIS MULTIVARIADO.**

### **3.2.1 Análisis Exploratorio de Datos (AED).**

Como se dijo en el Capítulo II, es importante iniciar con un AED dentro del Análisis Multivariado de esta forma se detectara previamente las posibles anomalías dentro de la base de datos que en análisis posteriores pudieran alterar los resultados.

No existe un algoritmo o un orden específico para el análisis exploratorio de datos es así que este análisis empieza a partir de la eliminación de variables redundantes.

#### **3.2.1.1 Normalidad.**

La prueba a utilizarse es la de Kolmogorov-Smirnov, mediante el paquete estadístico SPSS.

Las hipótesis a probar son:

$H_0 =$  *El conjunto de datos sigue una distribución Normal.*

$H_1 =$  *El conjunto de datos no sigue una distribución Normal.*

La Tabla II muestra la aplicación del test.

Tabla II: Prueba de Kolmogorov-Smirnov																
		Sb	Sn	Ag	Bal	Bi	As	Se	Au	Pb	Cu	Ni	Fe	Ti	Ca	K
N		194	194	194	194	194	194	194	194	194	194	194	194	194	194	194
Parámetros	Media	,01210	,00297	,02956	19,07520	,00416	,32136	,00008	,00000	,25443	80,04378	,00482	,11054	,00208	,01011	,00228
normales <sup>a,b</sup>	Desviación típica	,019963	,010416	,010894	2,399369	,015302	,154389	,001077	,000000 <sup>c</sup>	,284660	2,416634	,015929	,187421	,005762	,037591	,012892
Diferencias más extremas	Absoluta	,403	,530	,129	,081	,504	,125	,523		,186	,079	,526	,287	,517	,524	,539
	Positiva	,403	,530	,129	,081	,504	,125	,523		,169	,057	,526	,287	,517	,524	,539
	Negativa	-,272	-,388	-,070	-,059	-,393	-,096	-,471		-,186	-,079	-,381	-,278	-,359	-,394	-,430
Z de Kolmogorov-Smirnov		5,613	7,378	1,793	1,135	7,022	1,744	7,291		2,587	1,102	7,328	3,993	7,202	7,293	7,510
Valor p		,000	,000	,003	,152	,000	,005	,000		,000	,176	,000	,000	,000	,000	,000

a. La distribución de contraste es la Normal.  
 b. Se han calculado a partir de los datos.  
 c. La distribución no tiene varianza para esta variable. No es posible realizar la prueba de Kolmogorov-Smirnov para una muestra.

Fuente: Matriz de datos Hachas Moneda vs. Elementos Químicos analizados con SPSS

Como se observa en la Tabla II según la obtención del **valor p** se aceptan únicamente dos hipótesis nulas es decir que la variable Bal y Cobre (Cu) están distribuidas normalmente, de acuerdo estos resultados probablemente las 13 variables restantes no provengan de una población normal pero la falta de normalidad se puede deber a que los datos de estas variables poseen muy pocos valores diferentes de cero, en el caso específico del Oro (Au) todos sus valores son igual a cero por esta razón como se muestra también en la Tabla II no se puede realizar el test de Kolmogorov-Smirnov.

Luego de este análisis de normalidad se continuara de todas formas con el estudio ya que en muchos de los siguientes análisis la normalidad no es indispensable.

A continuación los Gráficos 1 y 2 muestran los histogramas de las 15 variables y sobre estos la curva aproximada de la distribución normal.

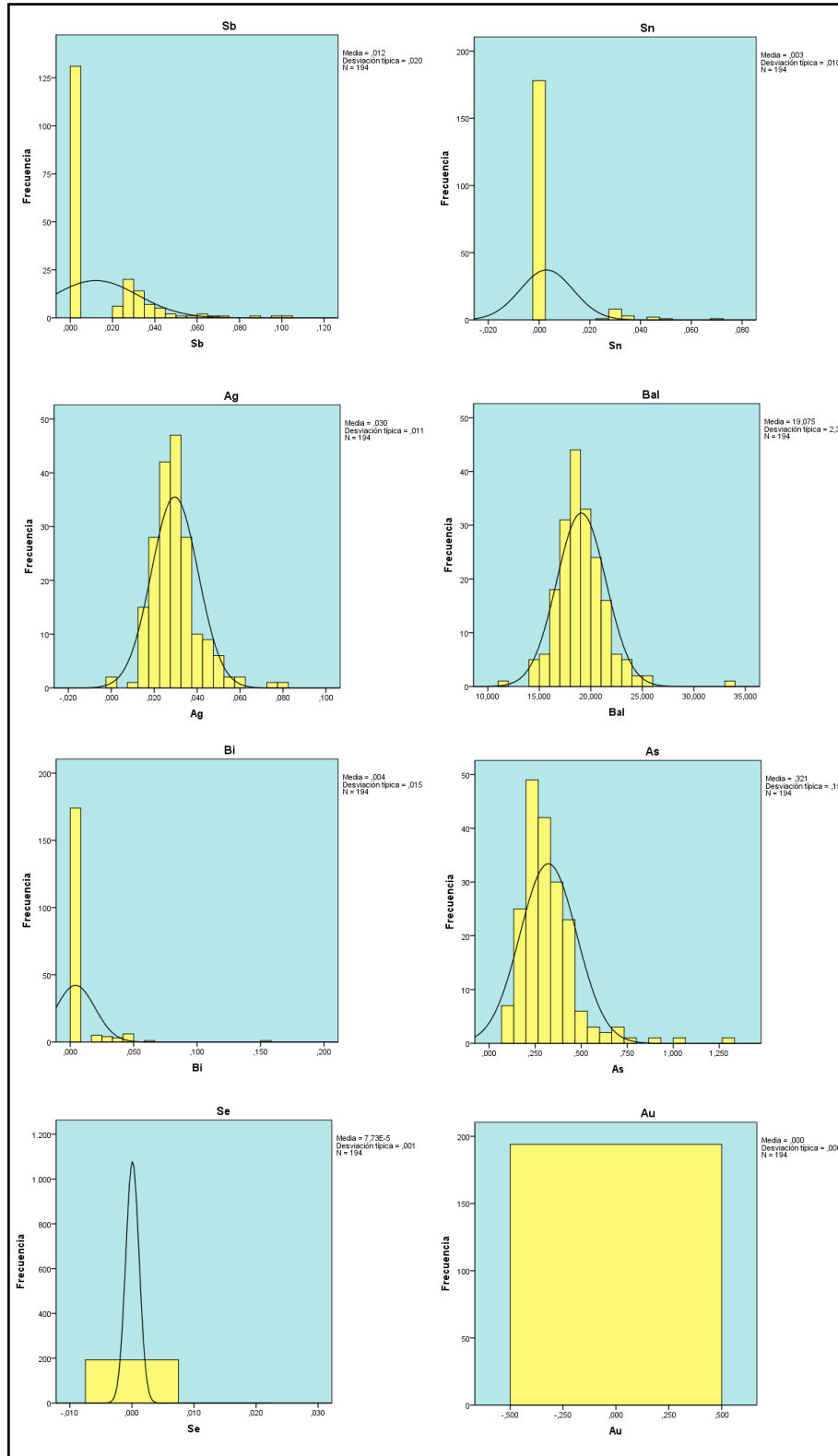


Gráfico 1: Histogramas con curva normal de Sb, Sn, Ag, Bal, Bi, As, Se, Au



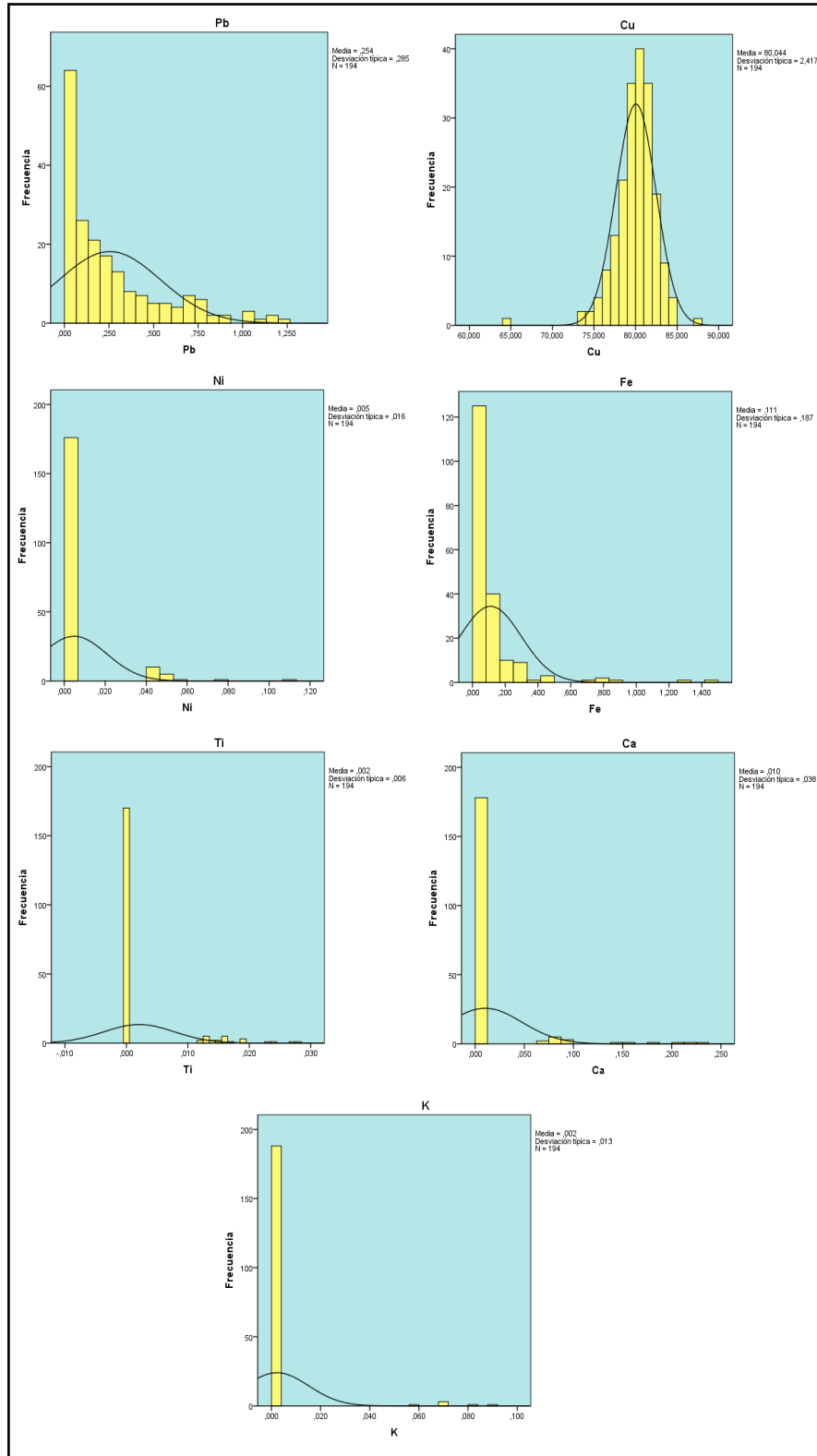


Gráfico 2: Histogramas con curva normal de Pb, Cu, Ni, Fe, Ti, Ca, K.

### 3.2.1.2 Análisis Univariante.

El análisis univariante es necesario ya que a partir de este análisis se puede caracterizar individualmente variable por variable. Este análisis se encuentra resumido en la Tabla III, a continuación.

TABLA III: ANALISIS UNVARIANTE DESCRIPTIVO INICIAL							
	Sb	Sn	Ag	Bal	Bi	As	Pb
<b>Media</b>	0,012	0,003	0,030	19,075	0,004	0,321	0,254
<b>Desviación Típica</b>	0,020	0,010	0,011	2,393	0,015	0,154	0,284
<b>Varianza</b>	0,000	0,000	0,000	5,727	0,000	0,024	0,081
<b>Coefficiente de Variación</b>	1,646	3,499	0,368	0,125	3,664	0,479	1,116
<b>Coefficiente de Asimetría</b>	1,819	3,711	1,210	1,378	5,880	2,499	1,451
<b>Coefficiente de Homogeneidad</b>	2,620	13,305	2,823	6,238	45,937	9,782	0,550
<b>Coefficiente de Kurtosis</b>	3,620	14,305	3,823	7,238	46,937	10,782	1,550

Fuente: Análisis Univariante realizado en EXCEL

TABLA III: ANALISIS UNVARIANTE DESCRIPTIVO INICIAL (Continuación)						
	Cu	Ni	Fe	Ti	Ca	K
<b>Media</b>	80,044	0,005	0,111	0,002	0,010	0,002
<b>Desviación Típica</b>	2,410	0,016	0,187	0,006	0,037	0,013
<b>Varianza</b>	5,810	0,000	0,035	0,000	0,001	0,000
<b>Coefficiente de Variación</b>	0,030	3,297	1,691	2,767	3,707	5,644
<b>Coefficiente de Asimetría</b>	-1,537	3,600	4,633	2,698	4,188	5,604
<b>Coefficiente de Homogeneidad</b>	7,054	13,328	25,383	5,327	17,111	29,275
<b>Coefficiente de Kurtosis</b>	8,054	14,328	26,383	6,327	18,111	30,275

Fuente: Análisis Univariante realizado en EXCEL

Como se observa en la Tabla III los porcentajes de la variable Cobre son mucho más constantes que las demás variables con una variabilidad relativa del 3%, mientras que el coeficiente de variación para el Calcio es el más alto esto sugiere que los porcentajes presentes de este elemento en la composición de las hachas varía mucho, esto se puede deber a la exposición de las piezas al medio ambiente.

Con respecto al coeficiente de asimetría para todas las variables es mayor que 1, lo que indica claramente que los datos presentan una alta distribución de asimetría.

De acuerdo a los coeficientes de homogeneidad y kurtosis es importante mencionar que la variabilidad de las desviaciones es alta y por tal motivo existe la presencia de datos atípicos univariantes.

En cuanto al coeficiente de kurtosis en las variables Sn, Bi, As, Se, Ni, Fe, Ca, y K es sumamente alto, mayor a 10, es decir que en estas variables la heterogeneidad se debe exclusivamente a la presencia de datos atípicos presentes en las variables nombradas anteriormente. En el caso de la variable Pb el coeficiente de kurtosis indica que existe una mezcla de dos poblaciones ya que este coeficiente es menor a 2, por esta razón los datos de esta variable son de igual manera heterogéneos.

TABLA IV: ANALISIS UNIVARIANTE DESCRIPTIVO ROBUSTO													
	Sb	Sn	Ag	Bal	Bi	As	Pb	Cu	Ni	Fe	Ti	Ca	K
<b>Mediana</b>	0	0	0,03	18,9	0	0,29	0,15	80,3	0	0,07	0	0	0
<b>MEDA</b>	0	0	0,01	1,3	0	0,07	0,13	1,3	0	0,04	0	0	0
<b>MEDA/Mediana</b>			0,18	0,1		0,24	0,91	0,0		0,54			

*Fuente: Análisis Univariante Robusto realizado en EXCEL*

De acuerdo a la mediana y la Meda se puede confirmar lo anteriormente dicho, ya que los valores de la media y las desviaciones típicas difieren de los valores obtenidos en la Mediana y MEDA se concluye que existe la presencia de valores extremos.

### 3.2.1.3 *El vector de medias.*

Como es conocido la media es la medida de centralización más utilizada pero en la matriz de datos se la denomina vector de medias y como se enuncia en el Capítulo I este vector contiene las medias de cada variable como se muestra a continuación.

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_{15} \end{bmatrix} = \begin{bmatrix} 0,012 \\ 0,003 \\ 0,030 \\ 19,075 \\ 0,004 \\ 0,321 \\ 0 \\ 0 \\ 0,254 \\ 80,044 \\ 0,005 \\ 0,111 \\ 0,002 \\ 0,010 \\ 0,002 \end{bmatrix}$$

### 3.2.1.4 *Matriz de Varianzas y Covarianzas.*

Como se ha comentado en el Capítulo I la variabilidad de las variables se mide por su varianza o su raíz cuadrada la desviación típica y la relación lineal entre dos variables se mide por la covarianza. En el caso multivariante toda esta información se puede resumir en la matriz de varianzas y covarianzas.

A continuación se indica la matriz de varianzas y covarianzas para las 15 variables en estudio.

TABLA V: MATRIZ DE VARIANZAS Y COVARIANZAS INICIAL (S)													
	Sb	Sn	Ag	Bal	Bi	As	Pb	Cu	Ni	Fe	Ti	Ca	K
Sb	0,000												
Sn	0,000	0,000											
Ag	0,000	0,000	0,000										
Bal	-0,003	0,000	-0,001	5,727									
Bi	0,000	0,000	0,000	-0,001	0,000								
As	0,000	0,000	0,000	-0,089	0,001	0,024							
Pb	0,000	0,000	0,000	-0,115	-0,001	-0,001	0,081						
Cu	0,003	-0,001	0,001	-5,684	0,001	0,063	0,034	5,810					
Ni	0,000	0,000	0,000	-0,003	0,000	0,001	-0,001	0,003	0,000				
Fe	0,000	0,000	0,000	0,131	0,000	0,001	0,005	-0,180	0,000	0,035			
Ti	0,000	0,000	0,000	0,001	0,000	0,000	0,000	-0,002	0,000	0,001	0,000		
Ca	0,000	0,000	0,000	0,019	0,000	0,000	0,001	-0,028	0,000	0,005	0,000	0,001	
K	0,000	0,000	0,000	0,002	0,000	0,000	0,000	-0,004	0,000	0,001	0,000	0,000	0,000

Fuente: Análisis realizado en EXCEL

En la diagonal de esta matriz consta las varianzas de cada variable, como una de las propiedades de esta matriz es la simetría se muestra únicamente los elementos de la diagonal inferior ya que los elementos de la diagonal superior serán iguales.

Tanto la matriz de varianzas y covarianzas como el vector de medias serán fundamentales para los análisis que a continuación se realizaran.

### 3.2.1.5 Variables Redundantes.

El principal propósito de encontrar variables redundantes es disminuir la dimensionalidad de la matriz de análisis, esto es posible solo si alguna de las variables es combinación de las demás. Es decir que posiblemente las 15 variables no son independientes ya que se podría despejar una cualquiera en función de las demás.

<b>TABLA VI: AUTOVALORES</b>
11,5187
0,1498
0,0469
0,0222
0,0007
0,0004
0,0002
0,0001
0,0003
0,0002
0,0001
0,0001
0,0001
0,0001
0,0001
0,0000

*Fuente: Análisis  
realizado en MINITAB*

De acuerdo a la Tabla VI se observa que existen autovalores muy pequeños cerca a cero y esto a su vez indica que existen variables redundantes.

Y de hecho de acuerdo a lo calculado existen 7 variables redundantes ya que Au, Se, Ti, K, Sn, Bi y Ni; resultaron ser combinaciones casi exactas de las demás. Y por esta razón en adelante serán eliminadas de análisis posteriores.

La matriz de datos con la que en adelante se trabajara es la siguiente:

TABLA VII: MATRIZ DE DATOS SIN VARIABLES REDUNDANTES								
Hacha	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
1	0	0,024	18,323	0,664	0,077	80,794	0	0
2	0	0,019	17,216	0,377	0,151	82,058	0	0
3	0	0,022	25,117	0,772	0,217	73,156	0,425	0,077
4	0	0,039	11,746	0,722	0,146	87,073	0	0,074
5	0	0,028	15,891	0,349	0,033	83,547	0	0
6	0,032	0,029	17,585	0,373	0	81,742	0,072	0
7	0,085	0,027	18,429	0,337	0	80,909	0	0
8	0	0,035	19,398	0,319	0,074	79,927	0,118	0
9	0	0,019	18,009	0,187	1,183	80,459	0,064	0
10	0	0,032	18,542	0,256	0,107	80,725	0,183	0
11	0	0,033	17,931	0,186	0,234	81,522	0,035	0
12	0	0,036	17,441	0,251	0,038	82,006	0,043	0
13	0,024	0,03	16,641	0,301	0,075	82,651	0,092	0
14	0,04	0,03	16,13	0,416	1,18	82,079	0	0
15	0	0,017	17,303	0,087	0,247	82,219	0	0
16	0	0,027	17,909	0,257	0,777	80,885	0	0
17	0	0,03	18,986	0,196	0,135	80,497	0,038	0
18	0	0,037	20,352	0,18	0,055	79,24	0	0
19	0	0,023	16,139	0,578	1,014	82,113	0	0
20	0	0,02	16,585	0,461	0,037	82,716	0	0
21	0	0,031	17,771	0,253	0,292	81,545	0	0
22	0,033	0,029	14,619	0,357	0,351	84,534	0	0
23	0,038	0,029	16,891	0,413	0,069	82,429	0	0
24	0	0,039	17,758	0,455	0	81,562	0,031	0
25	0	0,023	18,453	0,245	0,235	80,916	0	0
26	0,027	0,028	17,486	0,402	0,84	81,082	0,038	0
27	0	0,024	16,374	0,298	0,628	82,535	0	0
28	0,072	0,018	18,141	0,26	0,855	80,453	0,037	0,077
29	0	0,017	19,215	0,255	0,058	80,303	0,042	0
30	0	0,028	20,181	0,115	0,145	79,357	0,037	0
31	0,024	0,026	16,286	0,239	0,676	82,615	0	0
32	0,03	0,028	19,394	0,328	0,616	79,502	0	0
33	0,027	0,046	20,187	0,229	0,034	79,275	0,066	0
34	0,03	0,024	18,635	0,679	0,258	79,746	0,491	0
35	0	0,052	18,928	0,28	0,65	79,907	0,035	0
36	0	0,014	22,497	0,346	0,228	75,812	0,785	0,143
37	0	0,027	19,337	0,365	0,339	79,778	0,049	0
38	0,038	0,027	19,536	0,266	0,105	79,663	0,173	0,067

TABLA VII: MATRIZ DE DATOS SIN VARIABLES REDUNDANTES (Continuación)								
39	0	0,017	18,057	0,163	0	81,56	0,038	0
40	0,069	0,028	15,624	0,593	0,041	83,475	0	0
41	0	0,032	16,889	0,49	0,412	82,034	0	0
42	0,027	0,019	18,879	0,417	0,285	80,257	0	0
43	0,028	0,025	15,399	0,27	0,723	83,427	0	0
44	0	0,042	14,782	0,497	0,302	84,191	0,088	0
45	0	0,029	16,561	0,399	0,155	82,641	0,103	0
46	0	0,032	18,702	0,194	0,053	80,849	0,04	0
47	0	0,023	22,411	0,529	0,056	76,502	0,311	0
48	0	0,04	19,406	0,191	0	80,118	0,095	0
49	0	0,024	19,312	0,211	0,793	79,538	0	0
50	0,026	0,037	18,965	0,181	0,557	79,998	0,094	0
51	0	0,035	17,583	0,275	0,883	81,072	0,061	0
52	0	0,031	20,633	0,382	0,114	78,705	0,035	0
53	0,031	0,032	19,785	0,176	0,082	79,741	0,039	0
54	0	0,019	20,531	0,319	0,102	78,624	0,15	0
55	0,046	0,024	19,034	0,333	0,294	80,144	0,033	0
56	0,047	0,017	14,742	0,696	0,569	83,796	0	0
57	0,032	0,03	20,144	0,35	0	79,224	0,075	0
58	0	0,024	19,473	0,346	0,364	79,618	0,036	0
59	0	0,033	19,98	0,202	0,74	78,864	0,089	0
60	0	0,026	19,901	0,463	0,14	79,258	0,049	0
61	0	0,033	21,311	0,213	0,031	78,172	0,125	0
62	0	0,046	19,25	0,466	0,405	78,008	1,332	0,22
63	0	0,042	20,551	0,263	0,07	78,923	0,033	0
64	0	0,057	18,92	0,311	0,158	80,41	0	0
65	0,03	0,021	19,198	0,143	0,668	79,766	0,068	0
66	0	0,017	22,349	0,244	0,158	76,73	0,278	0
67	0	0,021	18,312	0,232	0,188	81,103	0	0
68	0,038	0,035	16,023	1,308	0	82,33	0,043	0
69	0	0,027	18,928	0,264	0,559	79,99	0,084	0
70	0,029	0,031	19,97	0,265	0,403	79,174	0,042	0
71	0,025	0	17,625	0,165	0,32	81,717	0,048	0
72	0,024	0,028	18,112	0,492	0,034	81,132	0,06	0
73	0	0,043	17,02	0,312	1,258	81,259	0,064	0
74	0	0,016	19,356	0,238	0,888	79,094	0,26	0
75	0,029	0,015	19,757	0,233	0	79,521	0,265	0
76	0,023	0,023	21,025	0,541	0,226	77,771	0,242	0
77	0	0,053	18,482	0,275	0,784	80,215	0,066	0



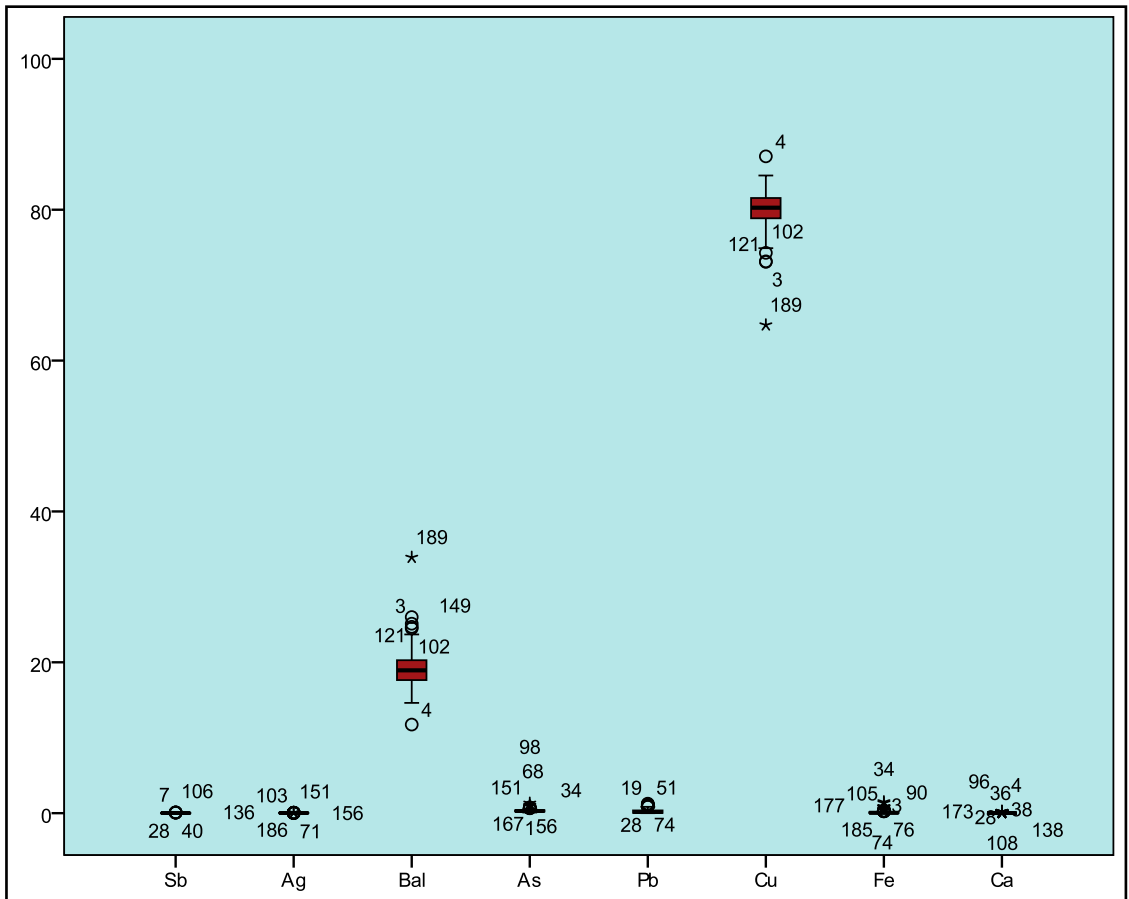
TABLA VII: MATRIZ DE DATOS SIN VARIABLES REDUNDANTES (Continuación)								
78	0,033	0,031	17,569	0,435	0	81,747	0,075	0
79	0	0,018	20,859	0,398	0,107	78,314	0,109	0
80	0	0,033	17,613	0,279	0	81,911	0,047	0
81	0,037	0,014	20,237	0,28	0,066	79,042	0,193	0
82	0	0,027	18,851	0,165	0,038	80,722	0,065	0
83	0	0,028	20,277	0,233	0,188	79,122	0,047	0
84	0	0,034	20,733	0,246	0	78,812	0,038	0
85	0,026	0,022	20,297	0,18	0	79,167	0,117	0
86	0,025	0,025	21,879	0,27	0	77,399	0,233	0
87	0	0,033	18,028	0,252	0,163	81,319	0,083	0
88	0	0,033	18,702	0,364	1,015	79,723	0,04	0
89	0,031	0,011	15,764	0,318	0,34	83,326	0,075	0
90	0	0,02	20,398	0,217	0,026	78,599	0,427	0,095
91	0	0,027	16,359	0,303	0,119	82,866	0,19	0
92	0	0,018	21,679	0,279	0,037	77,752	0,072	0
93	0	0,029	20,899	0,368	0,055	78,434	0	0
94	0,025	0,044	18,924	0,314	0,259	80,247	0,088	0
95	0	0,023	18,13	0,213	0,072	81,334	0,067	0
96	0	0,027	17,584	0,328	1,082	79,099	1,494	0,162
97	0	0,019	23,694	0,3	0	75,778	0,073	0
98	0	0,033	18,786	0,646	0,108	80,062	0,174	0
99	0	0,048	21,225	0,241	0	78,23	0,133	0
100	0	0,023	21,015	0,367	0,141	78,065	0,175	0
101	0	0,035	18,592	0,131	0	81,027	0,083	0
102	0	0,018	24,501	0,126	0,409	74,283	0,3	0,23
103	0,101	0	18,937	0,354	0,376	80,061	0,033	0
104	0,026	0,024	21,491	0,196	0,286	77,797	0,086	0
105	0,044	0,022	21,892	0,46	0,669	76,491	0,258	0
106	0,096	0,033	21,533	0,2	0,262	77,638	0,108	0
107	0	0,021	20,652	0,27	0,591	78,225	0,103	0
108	0	0,021	20,151	0,289	0,054	78,254	0,783	0,211
109	0	0,024	17,578	0,408	0,217	81,55	0,118	0
110	0,024	0,035	18,316	0,362	0	81,099	0,078	0
111	0	0,031	18,969	0,116	0,043	80,658	0,069	0
112	0,034	0,037	17,637	0,33	0,09	81,681	0,04	0
113	0	0,022	16,608	0,328	0,039	82,737	0,147	0
114	0	0,031	18,387	0,231	0,146	80,921	0,116	0
115	0	0,051	16,408	0,232	0,035	83,1	0,067	0
116	0,026	0,025	16,866	0,301	0,733	81,887	0,034	0

TABLA VII: MATRIZ DE DATOS SIN VARIABLES REDUNDANTES (Continuación)								
117	0	0,031	18,001	0,26	0,537	80,939	0,116	0
118	0	0,026	19,685	0,464	0,137	79,427	0,078	0
119	0,033	0,031	16,401	0,291	0,105	82,974	0,068	0
120	0,06	0,019	17,819	0,407	0,166	81,353	0,055	0
121	0,027	0,062	25,995	0,24	0,243	73,086	0,114	0
122	0	0,029	21,534	0,21	0,154	77,875	0,061	0
123	0	0,025	17,375	0,341	0,242	81,765	0,124	0
124	0	0,037	20,467	0,297	0,503	78,518	0,073	0
125	0	0,045	19,566	0,316	0	79,833	0,062	0
126	0	0,02	21,106	0,309	0,405	77,963	0,066	0
127	0,033	0,024	16,279	0,266	0,032	83,102	0,111	0
128	0,043	0,015	17,661	0,34	0,36	81,373	0,116	0
129	0	0,032	19,581	0,269	0,245	79,697	0,068	0
130	0,03	0,02	19,011	0,283	0,158	80,31	0,104	0
131	0	0,028	20,197	0,444	0,311	78,861	0	0
132	0	0,045	18,841	0,259	1,048	79,336	0,27	0
133	0,024	0,02	22,345	0,28	0,247	76,974	0,032	0
134	0	0,032	17,351	0,338	0,28	81,812	0,073	0
135	0,026	0,03	23,563	0,356	0	75,892	0	0
136	0,043	0,079	23,085	0,343	0	76,218	0,065	0
137	0	0,028	19,49	0,113	0,071	79,886	0,2	0
138	0	0,029	23,641	0,28	0,252	75,273	0,318	0,077
139	0,025	0,029	15,447	0,381	0,039	83,941	0	0
140	0	0,036	20,215	0,507	0,311	78,596	0,12	0
141	0,025	0,02	19,324	0,144	0,025	80,259	0,082	0
142	0	0,033	20,496	0,189	0,782	78,35	0,056	0
143	0	0,039	17,817	0,244	0	81,787	0	0
144	0	0,039	15,762	0,238	0,473	83,311	0,07	0
145	0,051	0,037	17,654	0,376	0,683	81,117	0	0
146	0	0,017	19,024	0,151	0,374	80,312	0	0
147	0	0,03	18,925	0,312	0,172	80,383	0	0
148	0	0,022	19,977	0,309	0,21	79,172	0,121	0
149	0	0,038	24,702	0,108	0,058	74,901	0	0
150	0	0,026	21,718	0,217	0,068	77,818	0	0
151	0	0,05	16,894	1,026	0,138	81,492	0,054	0
152	0,038	0,051	21,518	0,345	0,238	77,603	0,033	0
153	0	0,026	17,906	0,19	0,277	81,43	0,032	0
154	0	0,043	22,397	0,306	0,333	76,676	0,046	0,084
155	0	0,016	21,454	0,257	0,079	77,956	0,089	0

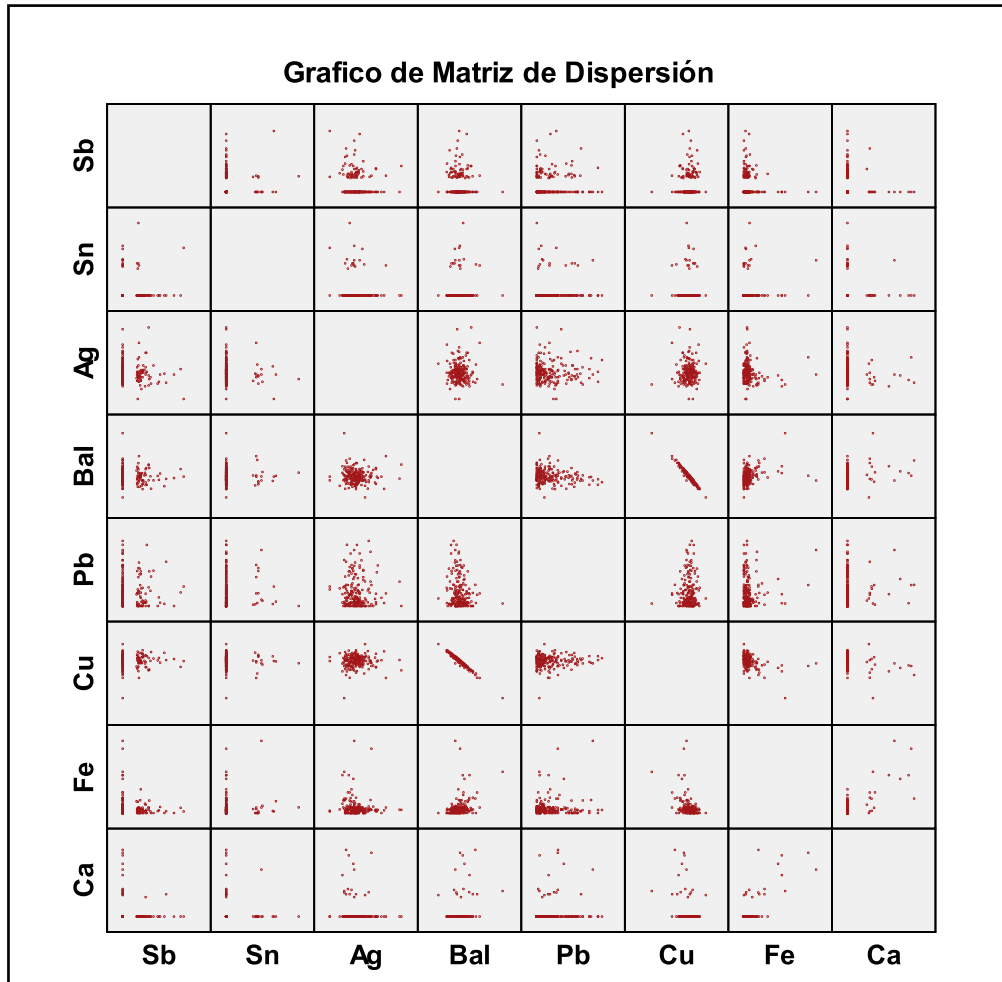
TABLA VII: MATRIZ DE DATOS SIN VARIABLES REDUNDANTES (Continuación)								
156	0,039	0,051	21,902	0,228	0,074	77,499	0,066	0
157	0	0,029	19,157	0,177	0,18	80,297	0,046	0
158	0	0,036	21,816	0,347	0	77,505	0,078	0
159	0	0,016	17,95	0,181	0,377	80,939	0,297	0,083
160	0,036	0,032	19,327	0,222	0,062	80,158	0,039	0
161	0	0,022	19,701	0,23	0,752	79,197	0	0
162	0	0,04	17,093	0,356	0,111	82,243	0,031	0
163	0,034	0,032	17,239	0,355	0,065	82,072	0	0
164	0	0,024	17,162	0,426	0,405	81,634	0,109	0,092
165	0	0,031	18,431	0,401	0,304	80,62	0,102	0
166	0	0,037	18,247	0,245	0,696	80,649	0	0
167	0	0,06	14,691	0,878	0,057	84,082	0,055	0
168	0,028	0,029	18,518	0,266	0,612	80,434	0,032	0
169	0	0,031	17,68	0,303	0,129	81,557	0,148	0
170	0	0,018	16,362	0,23	0,331	82,902	0,04	0
171	0,028	0,023	19,213	0,453	0	80,127	0	0
172	0	0,029	18,094	0,422	0	81,232	0,067	0
173	0	0,026	20,749	0,4	0,522	77,22	0,707	0,182
174	0	0,035	20,962	0,316	0,04	78,43	0,082	0
175	0	0,021	14,984	0,392	0	84,401	0,068	0
176	0	0,035	19,53	0,327	0,082	79,858	0,07	0
177	0	0,025	20,309	0,386	0,437	78,282	0,374	0
178	0	0,045	18,598	0,418	0,472	80,244	0,079	0
179	0,043	0,026	22,898	0,279	0	76,487	0,106	0
180	0	0,033	18,38	0,447	0	80,902	0,111	0
181	0	0,043	17,644	0,182	0	81,856	0,135	0
182	0	0,032	18,629	0,16	0,063	80,824	0,117	0
183	0,062	0,027	19,997	0,147	0,064	79,419	0,126	0
184	0	0,038	18,9	0,506	0	80,302	0,083	0
185	0	0,027	19,192	0,295	0,085	79,911	0,245	0
186	0	0,077	18,247	0,435	0,471	80,519	0,069	0
187	0	0,028	18,307	0,217	0	81,235	0,054	0
188	0	0,029	20,851	0,182	0,03	78,679	0,115	0
189	0	0,016	33,906	0,233	0,048	64,713	0,853	0,088
190	0	0,027	23,155	0,408	0,088	76,159	0,06	0
191	0	0,027	17,951	0,195	0	81,553	0,138	0
192	0	0,033	18,997	0,212	0,122	80,364	0,115	0
193	0,058	0,017	19,176	0,213	0	80,364	0,04	0
194	0,029	0,025	18,056	0,242	0,184	81,292	0,072	0

Fuente: Datos obtenidos con la máquina Thermo Niton XL3 XRF

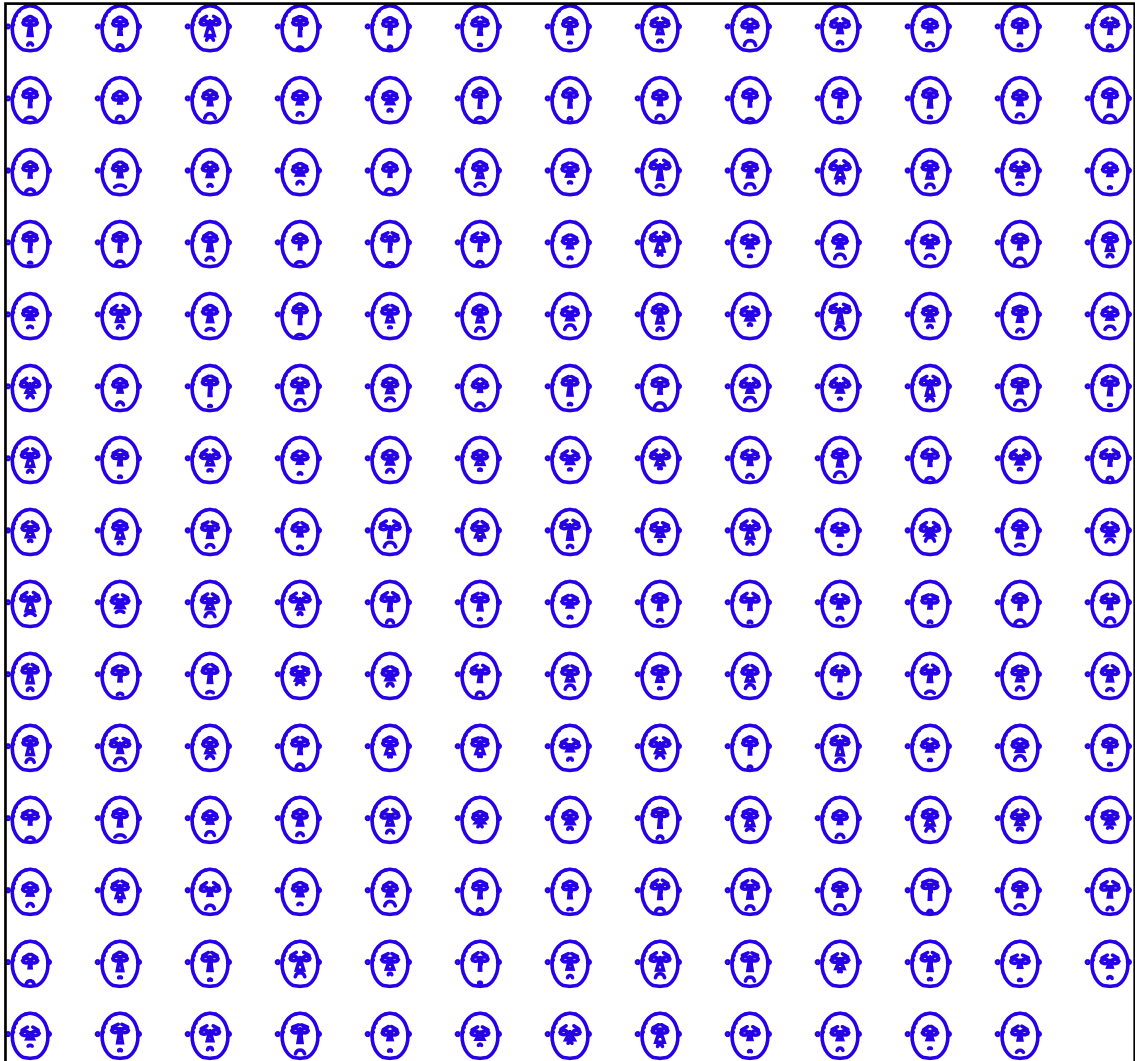
**3.2.1.6 Datos atípicos.**



**Gráfico 3:** Representación Grafica de Datos Atípicos - Diagramas de caja.



*Gráfico 4: Representación Gráfica de Datos Atípicos – Matriz de Dispersión.*



*Gráfico 5: Representación Gráfica de Datos Atípicos – Caritas de Chernoff.*

De acuerdo al análisis univariante y también la representación gráfica de datos atípicos, claramente se observa la presencia de estos, pero es necesario realizar un

análisis analítico multivariante ya que datos atípicos univariantes no siempre son datos atípicos multivariantes y para analizar datos atípicos multivariantes se utilizó el algoritmo de máxima y mínima kurtosis<sup>17</sup>.

Este algoritmo se iteró tres veces y finalmente los datos que en un principio resultaron sospechosos no fueron determinados como datos atípicos por esta razón fueron reintegrados a la matriz de datos.

Los datos detectados como potencialmente atípicos deben ser estudiados con detalle para determinar las causas de la heterogeneidad. Si estos datos no tienen un error detectable, conviene, cuando sea posible, investigar las causas de su tamaño anómalo ya que puede llevar a importantes descubrimientos. Si no hay un error en el dato y, sin embargo, es muy distinto de los demás, hay que sospechar que sobre esa observación ha actuado alguna causa que no ha estado activa en el resto de las observaciones.

Es importante mencionar que después de la aplicación del algoritmo mencionado anteriormente en el caso de que hubiera existido datos atípicos era necesario retirarlos de la matriz de datos y volver a realizar los análisis anteriores, pero en este caso con la no existencia de atípicos el nuevo análisis cambiara únicamente por la

---

<sup>17</sup> Tomado de PEÑA D., *Análisis de Datos Multivariantes*

eliminación de variables redundantes, a continuación se muestran los resultados finales.

TABLA VIII: ANALISIS UNIVARIANTE DESCRIPTIVO FINAL								
	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
<b>Media</b>	0,012	0,030	19,075	0,321	0,254	80,044	0,111	0,010
<b>Desviación Típica</b>	0,020	0,011	2,393	0,154	0,284	2,410	0,187	0,037
<b>Varianza</b>	0,000	0,000	5,727	0,024	0,081	5,810	0,035	0,001
<b>Coefficiente de Variación</b>	1,646	0,368	0,125	0,479	1,116	0,030	1,691	3,707
<b>Coefficiente de Asimetría</b>	1,804	1,201	1,367	2,479	1,440	-1,525	4,597	4,156
<b>Coefficiente de Homogeneidad</b>	2,620	2,823	6,238	9,782	0,550	7,054	25,383	17,111
<b>Coefficiente de Kurtosis</b>	3,620	3,823	7,238	10,782	1,550	8,054	26,383	18,111

Fuente: Análisis Univariante Final realizado en EXCEL

Al realizar una breve comparación entre las Tablas III y VIII, el análisis univariante final contiene menos heterogeneidad ya que los coeficientes de kurtosis mayores a 10 pertenecen a las variables Fe y Ca la asimetría se sigue conservando en todas las variables.

El vector final de medias es:

$$\bar{x}_f = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_8 \end{bmatrix} = \begin{bmatrix} 0,012 \\ 0,030 \\ 19,075 \\ 0,321 \\ 0,254 \\ 80,044 \\ 0,111 \\ 0,010 \end{bmatrix}$$

La matriz de varianzas y covarianzas final es:



TABLA IX: MATRIZ DE VARIANZAS Y COVARIANZAS FINAL ( $S_f$ )								
	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
Sb	0,0004							
Ag	0,0000	0,0001						
Bal	-0,0028	-0,0007	5,7273					
As	0,0002	0,0002	-0,0893	0,0237				
Pb	0,0000	-0,0001	-0,1153	-0,0015	0,0806			
Cu	0,0028	0,0006	-5,6842	0,0629	0,0335	5,8100		
Fe	-0,0005	-0,0002	0,1314	0,0014	0,0045	-0,1798	0,0349	
Ca	-0,0001	0,0000	0,0193	0,0002	0,0011	-0,0275	0,0053	0,0014

Fuente: Análisis realizado en EXCEL

### 3.2.1.7 Medidas Globales de Variabilidad.

En el caso de las variables estudiadas que son adimensionales porque son porcentajes de elementos químicos hallados en hachas moneda de cobre conviene estudiar medidas globales de variabilidad para comparar los distintos conjuntos de variables, para esto la VARIANZA GENERALIZADA y la VARIANZA EFECTIVA ayudaran en el análisis.

TABLA X: MEDIDAS GLOBALES DE VARIABILIDAD			
GENERALIZADA		EFECTIVA	
Varianza (VG)	9,1463E-18	Varianza (VE)	0,00741576
Desviación Típica (DG)	3,0243E-09	Desviación Típica (DE)	0,08611481

Fuente: Análisis realizado en EXCEL

Como la varianza generalizada mide el grado de dispersión en el espacio, notamos que esta no es muy alta al contrario esta es prácticamente cero.

En cuanto a la varianza efectiva como los valores son pequeños se dice que existe una gran dependencia entre las variables.

### 3.2.1.8 Medidas de Dependencia Lineal.

Uno de los objetivos fundamentales del análisis Multivariado es describir la relación de dependencia lineal entre las variables.

#### 3.2.1.8.1 Dependencia por pares (Matriz de Correlación)

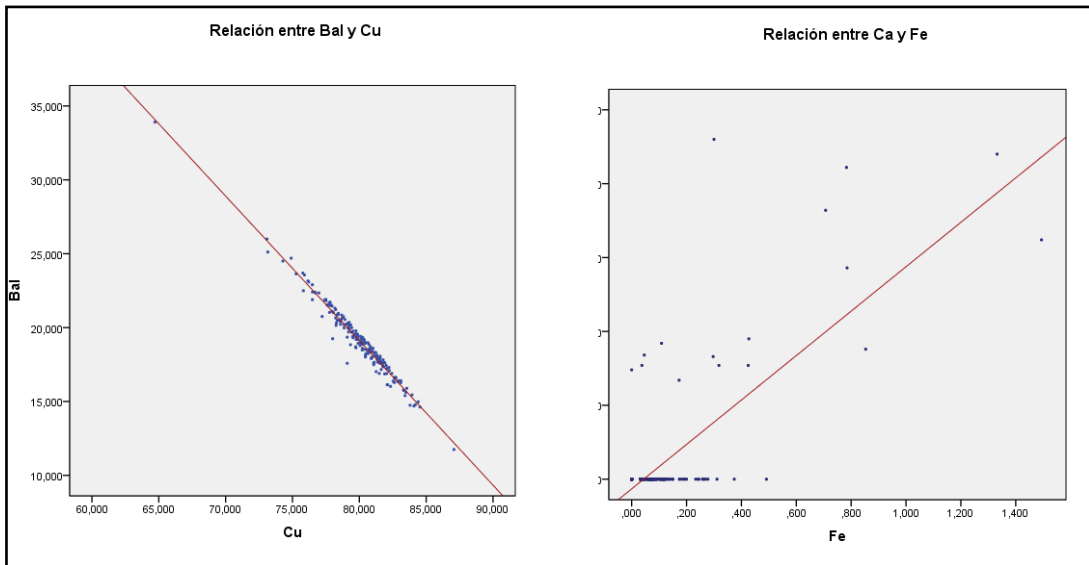
El estudio de las posibles relaciones lineales entre pares de variables se efectúa utilizando la matriz de correlación.

TABLA XI: MATRIZ DE CORRELACION (R)								
	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
Sb	1	-0,137	-0,058	0,076	0,000	0,059	-0,134	-0,108
Ag	-0,137	1	-0,026	0,138	-0,026	0,024	-0,095	-0,098
Bal	-0,058	-0,026	1	-0,242	-0,170	-0,985	0,294	0,215
As	0,076	0,138	-0,242	1	-0,034	0,169	0,048	0,027
Pb	0,000	-0,026	-0,170	-0,034	1	0,049	0,086	0,107
Cu	0,059	0,024	-0,985	0,169	0,049	1	-0,399	-0,305
Fe	-0,134	-0,095	0,294	0,048	0,086	-0,399	1	0,750
Ca	-0,108	-0,098	0,215	0,027	0,107	-0,305	0,750	1

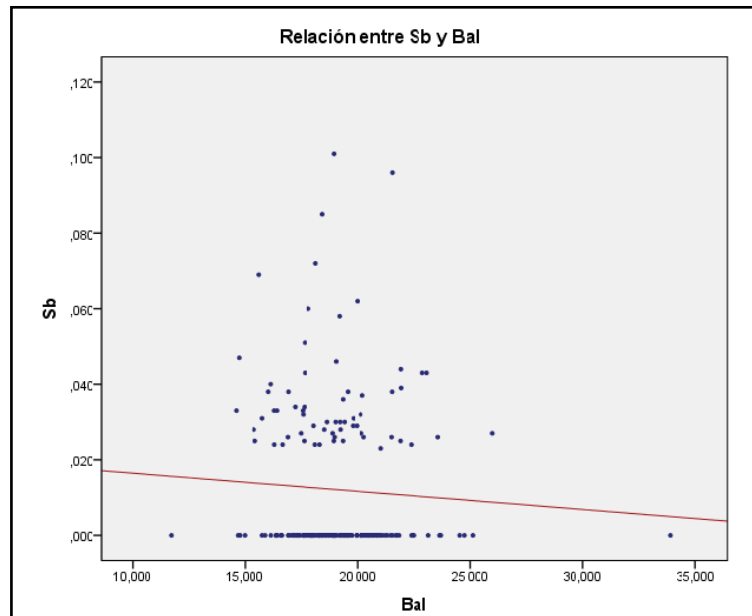
Fuente: Análisis realizado en EXCEL

En la Tabla XI se observa que la mayoría de los pares de variables están prácticamente incorreladas, es decir, no existen relaciones lineales entre los pares de variables pero podrían existir otros tipos de relaciones. Se nota además que entre las variables Cu y Bal existe correlación negativa y una dependencia inversa fuerte, o sea, a niveles altos de cobre tienden a corresponder niveles bajos de Bal. Entre las variables Ca y Fe existe en cambio correlación positiva y una dependencia directa relativamente fuerte, es decir, que a niveles altos de Ca tienden a corresponder niveles altos de Fe. Los

siguientes diagramas de dispersión corroboran visualmente lo mencionado anteriormente.



**Gráfico 6:** Diagramas de dispersión – Correlaciones relativamente altas.



**Gráfico 7:** Diagrama de dispersión – Correlación más baja.

### 3.2.1.8.2 Dependencia de cada variable y el resto.

En esta parte se pretende estudiar la posible dependencia lineal entre una cualquiera de las variables y las restantes. La siguiente tabla muestra los coeficientes de determinación al cuadrado de las 8 posibles regresiones.

TABLA XII: COEFICIENTE DE DETERMINACION	
$R_{Sb,Ag,Bal,As,Pb,Cu,Fe,Ca}^2$	0,23190336
$R_{Ag,Sb,Bal,As,Pb,Cu,Fe,Ca}^2$	0,19543015
$R_{Bal,Sb,Ag,As,Pb,Cu,Fe,Ca}^2$	0,99980145
$R_{As,Sb,Ag,Bal,Pb,Cu,Fe,Ca}^2$	0,96103828
$R_{Pb,Sb,Ag,Bal,As,Cu,Fe,Ca}^2$	0,98494184
$R_{Cu,Sb,Ag,Bal,As,Pb,Fe,Ca}^2$	0,9998038
$R_{Fe,Sb,Ag,Bal,As,Pb,Cu,Ca}^2$	0,97394967
$R_{Ca,Sb,Ag,Bal,As,Pb,Cu,Fe}^2$	0,71681873

Fuente: Análisis realizado en EXCEL

De acuerdo con la Tabla XII vemos que los modelos ajustados aceptables son los siguientes:

- La ecuación para prever Bal es:

$$Bal = 99,8 - 0,841 Sb - 1,31 Ag - 1,09 As - 0,960 Pb - 0,999 Cu - 1,08 Fe - 0,999 Ca$$

Observando los coeficientes de regresión estimados podemos afirmar que las variables de mayor peso para explicar la variable Bal son Ag y As.

- La ecuación para prever As es:

$$As = 88,0 - 0,706 Sb - 1,06 Ag - 0,882 Bal - 0,848 Pb - 0,880 Cu - 0,947 Fe - 0,881 Ca$$

Observando los coeficientes de regresión estimados podemos afirmar que las variables de mayor peso para explicar la variable As son Ag y Fe.

- **La ecuación para prever Pb es:**

$$Pb = 102 - 0,858 Sb - 1,35 Ag - 1,03 Bal - 1,11 As - 1,02 Cu - 1,11 Fe - 1,01 Ca$$

Observando los coeficientes de regresión estimados podemos afirmar que las variables de mayor peso para explicar la variable Pb son Ag, As y Fe.

- **La ecuación para prever Cu es:**

$$Cu = 99,9 - 0,842 Sb - 1,32 Ag - 1,00 Bal - 1,09 As - 0,961 Pb - 1,08 Fe - 1,00 Ca$$

Observando los coeficientes de regresión estimados podemos afirmar que las variables de mayor peso para explicar la variable Cu son Ag, As y Fe.

- **La ecuación para prever Fe es:**

$$Fe = 86,4 - 0,764 Sb - 1,18 Ag - 0,865 Bal - 0,933 As - 0,830 Pb - 0,865 Cu - 0,645 Ca$$

Observando los coeficientes de regresión estimados podemos afirmar que las variables de mayor peso para explicar la variable Fe son Ag y As.

Una vez descritas las ecuaciones para prever las variables de los modelos ajustados aceptables continuaremos con el cálculo de las estimaciones de dichas variables para las condiciones donde los porcentajes de las variables serán los siguientes:

TABLA XIII: EJEMPLO DE COMPOSICION QUIMICA DE UNA HACHA MONEDA (%)							
Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
0,05	0,04	22,80	0,70	0,63	75,90	0,75	0,12

Fuente: Análisis realizado en EXCEL

Luego de reemplazar los valores de la Tabla XIII en las ecuaciones de los modelos ajustados aceptables se obtuvieron las siguientes estimaciones:

TABLA XIV: ESTIMACIONES PARA LOS MODELOS AJUSTADOS ACEPTABLES		
Variable Respuesta	Modelos Ajustados Aceptables	Estimación (%)
Bal	$Bal = 99,8 - 0,841 Sb - 1,31 Ag - 1,09 As - 0,960 Pb - 0,999 Cu - 1,08 Fe - 0,999 Ca$	21,63
As	$As = 88,0 - 0,706 Sb - 1,06 Ag - 0,882 Bal - 0,848 Pb - 0,880 Cu - 0,947 Fe - 0,881 Ca$	-0,38
Pb	$Pb = 102 - 0,858 Sb - 1,35 Ag - 1,03 Bal - 1,11 As - 1,02 Cu - 1,11 Fe - 1,01 Ca$	-0,62
Cu	$Cu = 99,9 - 0,842 Sb - 1,32 Ag - 1,00 Bal - 1,09 As - 0,961 Pb - 1,08 Fe - 1,00 Ca$	74,68
Fe	$Fe = 86,4 - 0,764 Sb - 1,18 Ag - 0,865 Bal - 0,933 As - 0,830 Pb - 0,865 Cu - 0,645 Ca$	-0,31

Fuente: Análisis realizado en EXCEL

TABLA XV: INTERVALOS DE CONFIANZA PARA LOS COEFICIENTES DE REGRESION DEL MODELO																
Variable Respuesta	Coeficientes para los modelos ajustados aceptables															
	$\beta_0$		$\beta_1$		$\beta_2$		$\beta_3$		$\beta_4$		$\beta_5$		$\beta_6$		$\beta_7$	
	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior	Límite Superior	Límite Inferior
Bal	99,618	99,982	-1,092	-0,589	-1,775	-0,849	-1,119	-1,053	-0,978	-0,943	-1,001	-0,996	-1,122	-1,039	-1,197	-0,802
As	85,266	90,686	-0,936	-0,476	-1,483	-0,635	-0,909	-0,855	-0,878	-0,819	-0,908	-0,853	-0,997	-0,897	-1,061	-0,702
Pb	100,410	104,162	-1,119	-0,598	-1,825	-0,867	-1,044	-1,007	-1,153	-1,076	-1,043	-1,005	-1,154	-1,058	-1,219	-0,807
Cu	99,863	99,962	-1,094	-0,590	-1,779	-0,852	-1,003	-0,999	-1,120	-1,052	-0,979	-0,943	-1,124	-1,042	-1,198	-0,804
Fe	83,128	89,652	-0,989	-0,540	-1,597	-0,769	-0,898	-0,832	-0,982	-0,884	-0,866	-0,794	-0,897	-0,832	-0,842	-0,447

Fuente: Análisis realizado en EXCEL

Antes de referirnos a los modelos de regresión es importante mencionar que los intervalos de confianza mostrados en la Tabla XV indican que con una confianza del 95% cada uno de los coeficientes de los modelos de regresión considerados como aceptables están dentro de los intervalos mencionados.

Con relación a los modelos de regresión se puede decir que:

El modelo de regresión de Bal sobre las restantes variables dice que en hachas moneda donde el porcentaje de Sb es de 0,05%, de Ag es de 0,04%, de As es de 0,70%, de Pb es de 0,63%, de Cu es de 75,90%, de Fe es de 0,75% y de Ca es de 0,12% se estima la existencia de un porcentaje de Bal de alrededor del 22%.

El modelo de regresión de As sobre las restantes variables dice que en hachas moneda donde el porcentaje de Sb es de 0,05%, de Ag es de 0,04%, de Bal es de 22,8%, de Pb es de 0,63%, de Cu es de 75,90%, de Fe es de 0,75% y de Ca es de 0,12% se estima la existencia de un porcentaje de Arsénico de alrededor del -0,38%.

El modelo de regresión de Pb sobre las restantes variables dice que en hachas moneda donde el porcentaje de Sb es de 0,05%, de Ag es de 0,04%, de Bal es de 22,8%, de As es de 0,70%, de Cu es de 75,90%, de Fe es de 0,75% y de Ca es de 0,12% se estima la existencia de un porcentaje de Plomo de alrededor del -0,62%.

El modelo de regresión de Cu sobre las restantes variables dice que en hachas moneda donde el porcentaje de Sb es de 0,05%, de Ag es de 0,04%, de Bal es de 22,8%, de As es



de 0,70%, de Pb es de 0,63%, de Fe es de 0,75% y de Ca es de 0,12% se estima la existencia de un alto porcentaje de Cobre de alrededor del 75%.

El modelo de regresión de Fe sobre las restantes variables dice que en hachas moneda donde el porcentaje de Sb es de 0,05%, de Ag es de 0,04%, de Bal es de 22,8%, de As es de 0,70%, de Pb es de 0,63%, de Cu es de 75,90% y de Ca es de 0,12% se estima la existencia de un porcentaje de Hierro de alrededor del -0,31%.

### **3.2.1.8.3 Coeficiente de Dependencia.**

Para obtener una medida de dependencia global entre las variables se obtuvo el coeficiente de Dependencia, este es:

$$D = 0,75$$

De esta manera podemos concluir que globalmente que la dependencia lineal explica el 75% de la variabilidad de este conjunto de datos.

### **3.2.2 Componentes Principales.**

El análisis de “Componentes Principales”, es el punto de partida de este estudio. Una explicación resumida indica que esta técnica es útil para describir la relación de covarianza entre las variables estudiadas. Si un conjunto de variables posee algún factor común subyacente, se deduce que los valores de estas variables estarán

correlacionados entre sí, condensándose la información en términos de una pocas pero sobresalientes variables y con una pérdida mínima de información.

Para obtener el número adecuado de componentes es necesario tomar en cuenta los componentes contengan en lo posible un porcentaje alto de la varianza total explicada del total de la información inicial.

TABLA XVI: VARIANZA TOTAL EXPLICADA, SELECCIÓN DE COMPONENTES						
Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	11,519	<b>98,123</b>	98,123	11,519	98,123	98,123
2	,150	<b>1,276</b>	<b>99,399</b>	,150	1,276	99,399
3	,047	,399	99,799			
4	,022	,189	99,988			
5	,001	,006	99,994			
6	,000	,004	99,998			
7	,000	,002	99,999			
8	8,003E-5	,001	100,000			

Fuente: Análisis realizado en SPSS – Método de Extracción de Componentes Principales.

Para la selección del número de componentes (Tabla XVI), a utilizarse se tomo en cuenta, aquellos factores cuya varianza total explicada sea del 99,4% de la información inicial. De esta manera se obtuvo un nuevo resultado que las nuevas variables representan el 99,4% del problema original. Así, se ha reducido de 8 variables a 2 componentes que representan la composición química de las hachas moneda de cobre.

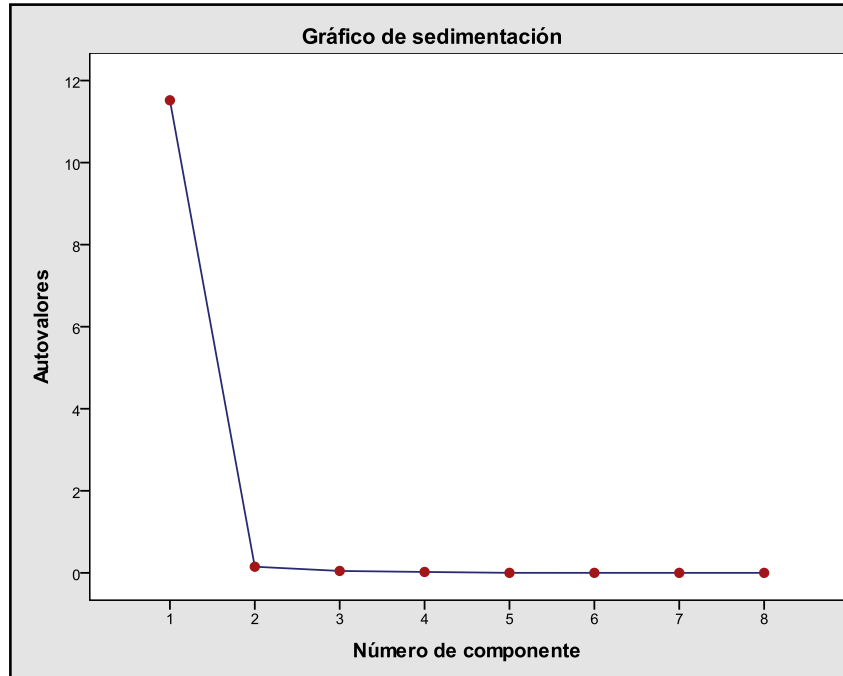


Gráfico 8: Gráfico de Sedimentación para la varianza acumulada.

El gráfico de sedimentación es explícito en cuanto a la selección adecuada del número de componentes adecuado.

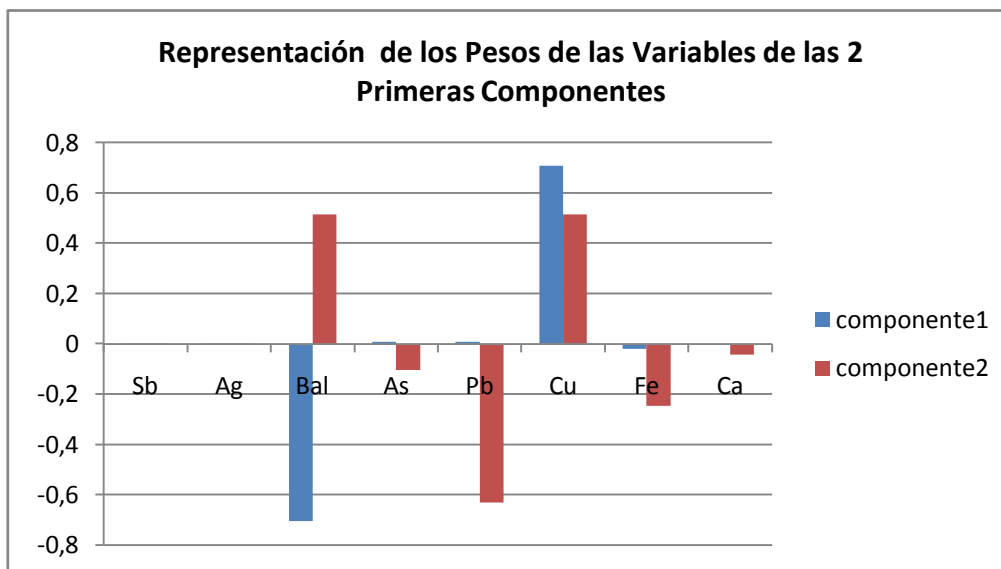
TABLA XVII: DOS PRIMERAS COMPONENTES		
	Componentes	
Variable	1	2
Sb	0,00035	0,00093
Ag	0,00008	0,00031
Bal	-0,70434	0,51518
As	0,00940	-0,10420
Pb	0,00922	-0,63034
Cu	0,70948	0,51419
Fe	-0,01926	-0,24540
Ca	-0,00290	-0,04236

Fuente: Análisis realizado en MINITAB Componentes Principales.

De acuerdo a la primera componente se puede decir que no existe en primer lugar una alta correlación positiva entre las variables ya que no todas sus coordenadas son

positivas, además a esta componente se la puede interpretar como un factor global de tamaño; por otra parte la variable con mayor peso positivo es Cu, mientras que con peso negativo destaca la variable Bal y al mismo tiempo estas dos variables se contraponen.

Es típico que a partir de la segunda componente las coordenadas tengan signos tanto positivos como negativos en definitiva el segundo componente se lo define como un factor de forma y además define las medias ponderadas de dos grupos de variables con distinto signo que se contraponen es así, que se definen el primer grupo de variables con Cu y Bal con coordenadas positivas frente al segundo grupo formado por Pb y Fe con coordenadas negativas.



**Gráfico 9:** Representación de los pesos de las dos componentes.

También es importante analizar los valores de las dos componentes para cada observación, en la Tabla XVIII se muestra el orden de las hachas moneda de cobre de acuerdo a los dos componentes.

<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES</b>			
<b>Hacha</b>	<b>Componente 1</b>	<b>Hacha</b>	<b>Componente 2</b>
4	53,5108815	39	51,213401
22	49,6846456	143	51,2073689
175	49,328974	101	51,2072605
44	49,3256469	149	51,191336
167	49,3143676	111	51,1896429
56	49,0796175	18	51,1757277
139	48,6780821	141	51,1726312
43	48,352482	84	51,1704067
40	48,2248242	193	51,1692788
5	48,0854581	187	51,1655369
89	48,0192179	82	51,1607853
144	48,0105464	80	51,150848
127	47,4935424	48	51,1500299
115	47,4018103	46	51,1429842
119	47,3185368	150	51,1362432
170	47,2969605	188	51,1316732
91	47,2693758	191	51,1273215
31	47,1509145	181	51,1270542
68	47,1369456	135	51,12484
27	47,0322782	97	51,1216054
20	47,008162	115	51,1193519
113	47,0027987	29	51,1165246
45	46,9703773	85	51,115834
13	46,9196872	53	51,1150803
19	46,9045873	99	51,1018337
14	46,8868265	61	51,1016523
23	46,588873	160	51,10136
41	46,3139108	33	51,1006822
162	46,3138765	12	51,0910313
116	46,226354	63	51,0890523
15	46,1482594	30	51,0887511

<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES (Continuación)</b>			
<b>2</b>	46,0971514	<b>5</b>	51,0883467
<b>163</b>	46,0898977	<b>125</b>	51,080991
<b>151</b>	45,9273255	<b>110</b>	51,0792706
<b>12</b>	45,8987304	<b>92</b>	51,0776239
<b>164</b>	45,8348038	<b>95</b>	51,0770945
<b>134</b>	45,8269523	<b>182</b>	51,070865
<b>123</b>	45,7753978	<b>179</b>	51,0702095
<b>80</b>	45,7100197	<b>7</b>	51,0615913
<b>73</b>	45,6766921	<b>175</b>	51,0598213
<b>181</b>	45,6465584	<b>57</b>	51,0589204
<b>78</b>	45,6255942	<b>17</b>	51,0568969
<b>6</b>	45,6102522	<b>139</b>	51,0551016
<b>71</b>	45,5657943	<b>183</b>	51,0517773
<b>112</b>	45,5313918	<b>171</b>	51,0512604
<b>109</b>	45,4803946	<b>174</b>	51,0486602
<b>143</b>	45,4789367	<b>156</b>	51,0459316
<b>169</b>	45,4111436	<b>127</b>	51,0414948
<b>24</b>	45,3622461	<b>155</b>	51,0382473
<b>21</b>	45,3424184	<b>158</b>	51,0359704
<b>128</b>	45,2970892	<b>6</b>	51,0336245
<b>26</b>	45,2203955	<b>24</b>	51,031688
<b>191</b>	45,2154187	<b>136</b>	51,0316419
<b>11</b>	45,2115683	<b>172</b>	51,0297096
<b>120</b>	45,1716369	<b>15</b>	51,0253515
<b>153</b>	45,1643962	<b>93</b>	51,0235074
<b>39</b>	45,1473499	<b>78</b>	51,0207576
<b>51</b>	45,1437312	<b>176</b>	51,0204989
<b>145</b>	45,125948	<b>157</b>	51,0138072
<b>87</b>	44,9982652	<b>137</b>	51,0116174
<b>194</b>	44,9597088	<b>20</b>	51,0044015
<b>95</b>	44,9361671	<b>163</b>	51,0036271
<b>172</b>	44,8904574	<b>113</b>	51,003574
<b>72</b>	44,8079458	<b>122</b>	51,0023101
<b>16</b>	44,7814734	<b>119</b>	51,000466
<b>159</b>	44,7805443	<b>23</b>	50,9993455
<b>117</b>	44,7505681	<b>180</b>	50,9939697
<b>187</b>	44,740886	<b>67</b>	50,9934228
<b>67</b>	44,6466268	<b>112</b>	50,9846559

<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES (Continuación)</b>			
110	44,6389673	86	50,9839302
114	44,4620469	8	50,9820708
180	44,4542775	192	50,9818099
7	44,4258648	11	50,9797942
1	44,4226873	162	50,9795998
25	44,4151988	52	50,978576
9	44,4106539	75	50,9778161
101	44,3912128	190	50,9762956
166	44,3751125	83	50,975521
28	44,3114091	13	50,9699979
186	44,2812663	64	50,9610134
182	44,2213268	72	50,9605787
165	44,2208271	184	50,9540785
10	44,2124023	87	50,9514852
46	44,1893574	81	50,9499852
168	44,0305154	194	50,9425562
82	43,993415	147	50,9407689
77	43,9015037	25	50,9389311
111	43,8644057	114	50,9365954
178	43,8385922	130	50,9338895
17	43,740403	2	50,928042
64	43,7271916	10	50,9212055
96	43,7174918	133	50,8980633
147	43,7046503	169	50,8947435
184	43,6634307	153	50,8928222
42	43,6496924	38	50,887101
192	43,6368427	40	50,8833684
94	43,6080014	91	50,8832507
130	43,5898904	79	50,8785427
146	43,5848961	21	50,8742384
98	43,5740087	54	50,8702825
193	43,5111146	129	50,8677775
157	43,4781482	71	50,8672162
103	43,4694189	1	50,8651609
55	43,4590144	45	50,8602807
34	43,4517657	60	50,8575403
29	43,441238	104	50,8521926
69	43,4252111	120	50,850175
50	43,4038453	146	50,8447085

<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES (Continuación)</b>			
<b>88</b>	43,4009519	<b>185</b>	50,8321152
<b>35</b>	43,3682596	<b>118</b>	50,8278283
<b>141</b>	43,3311396	<b>170</b>	50,8140882
<b>171</b>	43,3199239	<b>148</b>	50,8067275
<b>160</b>	43,2592765	<b>106</b>	50,8014556
<b>185</b>	43,1760388	<b>100</b>	50,7964832
<b>48</b>	43,1733016	<b>152</b>	50,7941067
<b>65</b>	43,0763032	<b>94</b>	50,7937096
<b>8</b>	43,0448677	<b>55</b>	50,7868871
<b>132</b>	43,0233704	<b>109</b>	50,7795445
<b>37</b>	42,9863236	<b>134</b>	50,7759424
<b>137</b>	42,947437	<b>90</b>	50,7754437
<b>176</b>	42,9040148	<b>123</b>	50,775263
<b>125</b>	42,8602175	<b>66</b>	50,7741233
<b>49</b>	42,837339	<b>42</b>	50,7701357
<b>58</b>	42,7773196	<b>121</b>	50,765912
<b>38</b>	42,7589143	<b>22</b>	50,7392844
<b>129</b>	42,7548637	<b>37</b>	50,7191278
<b>32</b>	42,7535208	<b>47</b>	50,7153319
<b>53</b>	42,6405889	<b>131</b>	50,7121198
<b>75</b>	42,4996503	<b>154</b>	50,7076978
<b>118</b>	42,4906978	<b>70</b>	50,7064952
<b>74</b>	42,4874616	<b>89</b>	50,7006185
<b>161</b>	42,3212209	<b>58</b>	50,6963335
<b>183</b>	42,2607143	<b>165</b>	50,6906218
<b>60</b>	42,2192363	<b>98</b>	50,666946
<b>70</b>	42,11175	<b>167</b>	50,6614794
<b>148</b>	42,1025029	<b>126</b>	50,6572993
<b>30</b>	42,0892675	<b>4</b>	50,6527035
<b>33</b>	42,0263647	<b>128</b>	50,6487691
<b>57</b>	42,0211186	<b>44</b>	50,6415707
<b>59</b>	41,8863682	<b>103</b>	50,6404048
<b>18</b>	41,8863117	<b>144</b>	50,6175831
<b>85</b>	41,8705058	<b>138</b>	50,6144948
<b>83</b>	41,8562367	<b>41</b>	50,5709477
<b>81</b>	41,8241688	<b>76</b>	50,562335
<b>62</b>	41,7680379	<b>124</b>	50,5512349
<b>131</b>	41,7314321	<b>140</b>	50,5491243
<b>140</b>	41,5290242	<b>159</b>	50,5323798



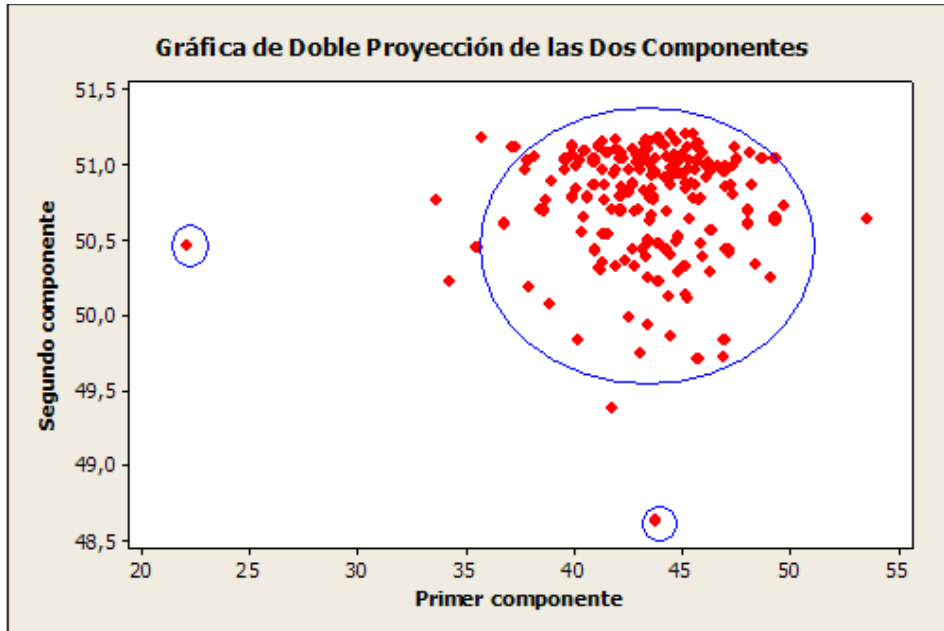
<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES (Continuación)</b>			
<b>63</b>	41,5215276	<b>50</b>	50,5113438
<b>90</b>	41,3907174	<b>117</b>	50,4975058
<b>54</b>	41,3220469	<b>164</b>	50,4863515
<b>84</b>	41,3136845	<b>178</b>	50,4813038
<b>108</b>	41,3136611	<b>69</b>	50,4806858
<b>52</b>	41,3105905	<b>168</b>	50,4769301
<b>124</b>	41,2968939	<b>189</b>	50,4747826
<b>177</b>	41,2351722	<b>102</b>	50,4635243
<b>142</b>	41,1591601	<b>65</b>	50,4524615
<b>188</b>	41,134404	<b>32</b>	50,4478823
<b>107</b>	40,9586933	<b>27</b>	50,4470858
<b>93</b>	40,9309675	<b>186</b>	50,4431983
<b>174</b>	40,8815504	<b>68</b>	50,4409466
<b>79</b>	40,8726646	<b>107</b>	50,4358641
<b>100</b>	40,5848798	<b>31</b>	50,4187991
<b>99</b>	40,5523598	<b>166</b>	50,404934
<b>126</b>	40,4524061	<b>151</b>	50,3984511
<b>61</b>	40,4508127	<b>35</b>	50,3910126
<b>76</b>	40,3703863	<b>161</b>	50,3736726
<b>155</b>	40,1983936	<b>108</b>	50,3533728
<b>173</b>	40,1657741	<b>43</b>	50,3464974
<b>122</b>	40,0853691	<b>59</b>	50,3348298
<b>104</b>	40,0609286	<b>145</b>	50,334661
<b>106</b>	39,9179574	<b>49</b>	50,32474
<b>150</b>	39,9157787	<b>142</b>	50,3193515
<b>152</b>	39,9062581	<b>177</b>	50,3069488
<b>92</b>	39,8953318	<b>16</b>	50,2998388
<b>158</b>	39,6237809	<b>116</b>	50,2925468
<b>156</b>	39,5587607	<b>34</b>	50,2509161
<b>86</b>	39,5005019	<b>56</b>	50,2504548
<b>133</b>	38,8769947	<b>3</b>	50,2309108
<b>105</b>	38,8546178	<b>77</b>	50,2280586
<b>66</b>	38,6951599	<b>36</b>	50,1931214
<b>154</b>	38,6294629	<b>51</b>	50,1443836
<b>47</b>	38,4908331	<b>28</b>	50,135477
<b>179</b>	38,1382771	<b>26</b>	50,119091
<b>36</b>	37,9310426	<b>105</b>	50,0761345
<b>136</b>	37,8171124	<b>74</b>	49,9926087
<b>190</b>	37,7274487	<b>88</b>	49,9399117

<b>TABLA XVIII: NUEVO ORDEN DE LAS HACHAS DE ACUERDO A LAS DOS PRIMERAS COMPONENTES (Continuación)</b>			
<b>135</b>	37,2505142	<b>9</b>	49,8679827
<b>97</b>	37,0754236	<b>173</b>	49,8431001
<b>138</b>	36,7516616	<b>19</b>	49,8365445
<b>149</b>	35,7433793	<b>132</b>	49,7462348
<b>102</b>	35,443453	<b>14</b>	49,7267092
<b>3</b>	34,212338	<b>73</b>	49,709522
<b>121</b>	33,5457333	<b>62</b>	49,3879149
<b>189</b>	22,016913	<b>96</b>	48,6409331

*Fuente: Análisis realizado en MINITAB*

El nuevo orden de las hachas de acuerdo a la primera componente obviamente se encuentra en los primeros lugares las hachas que en su composición contienen un porcentaje alto de Cobre, ya que esta variable fue la más destacada dentro de los pesos de la primera componente.

Mientras que en las observaciones relacionadas con la segunda componente se encuentran de igual manera en los primeros lugares las hachas que en su composición tiene un porcentaje alto de Cu y Bal conjuntamente.



*Gráfico 10: Representación de las observaciones sobre las dos primeras componentes.*

De acuerdo al Gráfico 10 se observa que posiblemente existen tres grupos notorios de hachas moneda aunque en uno de ellos hay una gran concentración de datos.

### **3.2.3 Análisis de Conglomerados.**

Este análisis es fundamental dentro del estudio ya que permitirá determinar grupos homogéneos de hachas en relación a su composición química.

Aunque este estudio es una parte importante dentro del desarrollo del trabajo investigativo hay que señalar que el análisis de conglomerados es una técnica de Análisis Exploratorio de Datos (AED), para resolver problemas de clasificación.

El análisis consiste en clasificar datos multidimensionales en un grupos homogéneos, para ello utiliza funciones de similitud o similaridad entre ellos, este análisis también se lo utilizará para agrupar variables.

### 3.2.3.1 Conglomerados por observaciones.

Con ayuda del dendrograma es posible asignar el número de conglomerados necesarios para el análisis.

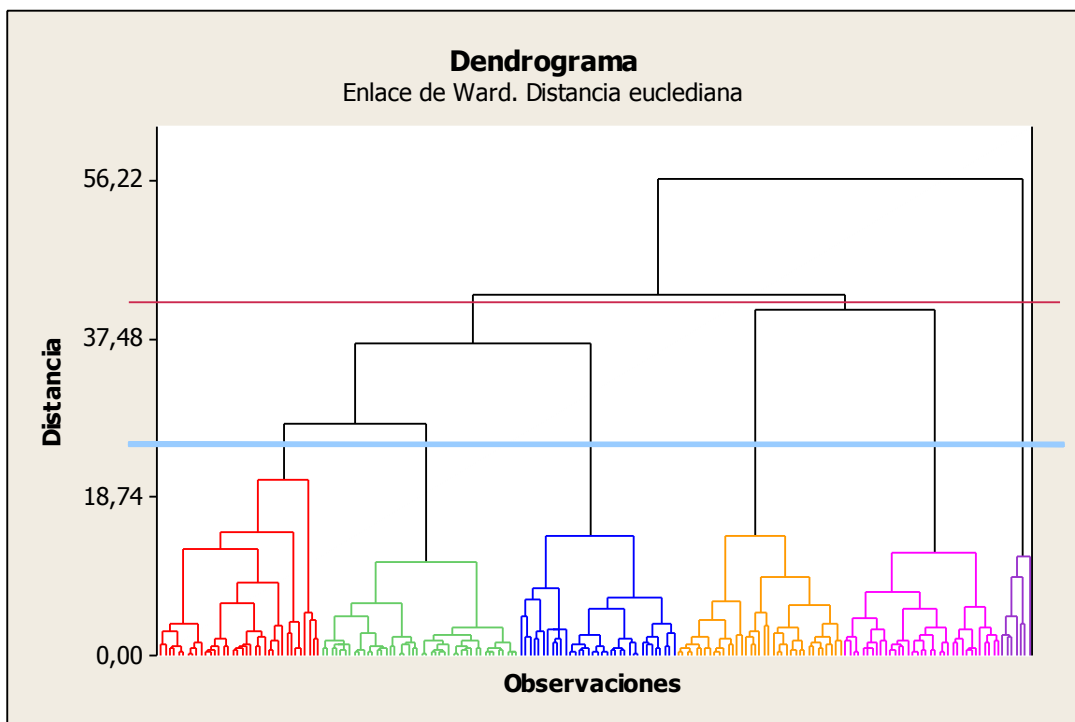


Gráfico 11: Dendrograma por observaciones.

En el dendrograma queda reflejada la formación de los conglomerados, así como las distancias entre ellos.

La decisión sobre el número óptimo de clusters es subjetiva, especialmente cuando se incrementa el número de objetos ya que si se seleccionan muy pocos, los clusters resultantes son heterogéneos y artificiales, mientras que si se seleccionan demasiados, la interpretación de los mismos suele ser complicada.

Como ayuda a la decisión sobre el número de clusters se suelen representar los distintos pasos del algoritmo y la distancia a la que se produce la fusión. En los primeros pasos el salto en las distancias será pequeño, mientras que en los últimos el salto entre pasos será mayor. El punto de corte será aquel en el que comienzan a producirse saltos bruscos, y estos saltos empiezan a partir de la línea azul y bajo esta línea están definidos claramente 6 conglomerados.

Análisis de observaciones de conglomerado: Sb. Ag. Bal. As. Pb. Cu. Fe. Ca							
Variables estandarizadas, Distancia euclidiana, Enlace de Ward							
Pasos de amalgamación							
Paso	Número de grupos	Nivel de semejanza	Nivel de distancia	Grupos incorporados	Nuevo grupo	Número de obs. en el grupo nuevo	
1	193	98,659	0,1926	60	118	60	2
2	192	98,262	0,2496	17	157	17	2
3	191	98,120	0,2700	8	176	8	2
4	190	98,062	0,2784	25	67	25	2
5	189	97,800	0,3161	92	155	92	2
6	188	97,680	0,3333	37	58	37	2
7	187	97,631	0,3403	49	161	49	2
8	186	97,455	0,3656	12	80	12	2
9	185	97,430	0,3692	87	114	87	2
10	184	97,353	0,3803	52	93	52	2
11	183	97,292	0,3890	59	142	59	2
12	182	97,006	0,4301	111	182	111	2
13	181	96,904	0,4447	6	78	6	2
14	180	96,741	0,4682	101	111	101	3
15	179	96,731	0,4696	109	123	109	2
16	178	96,715	0,4719	53	160	53	2
17	177	96,593	0,4895	82	187	82	2
18	176	96,526	0,4990	172	180	172	2
19	175	96,459	0,5087	61	188	61	2
20	174	96,391	0,5185	46	82	46	3
21	173	96,375	0,5208	13	119	13	2
22	172	96,306	0,5306	122	150	122	2
23	171	96,301	0,5314	10	192	10	2

Gráfico 12: Análisis de Conglomerados por Observaciones en MINITAB.

A partir del el análisis por conglomerados realizado en Minitab se obtuvo la Tabla XIX en donde se resumen los conglomerados y las observaciones en cada uno de ellos.

TABLA XIX: DISTRIBUCION DE LAS OBSERVACIONES PARA CADA CONGLOMERADO					
Conglomerado 1	Conglomerado 2	Conglomerado 3	Conglomerado 4	Conglomerado 5	Conglomerado 6
1	2	3	6	9	36
98	20	138	78	74	108
24	5	154	23	49	173
184	45	90	163	161	102
172	91	159	110	16	62
180	113	164	112	51	96
60	175	34	72	166	189
118	11	47	171	88	
131	21	76	13	73	
140	153	177	119	132	
12	109	105	127	32	
80	123	52	22	168	
143	134	93	139	50	
162	169	84	40	70	
115	165	174	56	65	
181	15	158	7	37	
18	170	61	183	58	
63	25	188	193	69	
48	67	122	106	117	
125	146	150	28	107	
99	29	54	103	126	
33	39	148	38	59	
94	8	79	53	142	
64	176	100	160	124	
35	10	66	57	14	
77	192	92	75	26	
178	87	155	81	145	
186	114	86	85	31	
121	185	135	141	43	
136	17	179	42	116	
152	157	104	55	19	
156	83	133	130	27	
4	129	97	194	144	

TABLA XIX: DISTRIBUCION DE LAS OBSERVACIONES PARA CADA CONGLOMERADO (Continuación)					
68	147	190	120	41	
151	30	149	128	44	
167	137		71		
	46		89		
	82				
	187				
	95				
	191				
	101				
	111				
	182				
<b>36</b>	<b>44</b>	<b>35</b>	<b>37</b>	<b>35</b>	<b>7</b>

Fuente: Distribución obtenida del análisis realizado en MINITAB

Tanto el dendrograma como el análisis analítico indican que la observación más distante al resto es la del hacha número 189 y esta forma parte del conglomerado 6, es necesario decir que el alejamiento de esta observación se debe principalmente a la influencia de la variable Bal ya que esta hacha en su composición posee el 33.9% de Bal y este porcentaje es el máximo localizado dentro de esta variable.

Por el contrario, las observaciones más cercanas entre sí son la 60 y la 118, que forman el primer grupo con un nivel de distancia de 0,19, esta distancia es la más próxima a 0.

TABLA XX: OBSERVACIONES MAS CERCANAS								
Hacha	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca
<b>60</b>	0	0,026	19,901	0,463	0,14	79,258	0,049	0
<b>118</b>	0	0,026	19,685	0,464	0,137	79,427	0,078	0

Fuente: Datos obtenidos con la máquina Thermo Niton XL3 XRF

Como se observa en la Tabla XX los porcentajes obtenidos en las observaciones 60 y 118 son muy similares además la distancia entre estas dos observaciones prácticamente es cero.

Cada uno de los conglomerados está distribuido homogéneamente es decir que las hachas que forman parte de estos conglomerados tienen una composición química similar.

En definitiva ya que aproximadamente el 60% de las observaciones queda a una distancia inferior a 2 los datos son aproximadamente homogéneos.

Obviamente ya que la formación de los clusters está dada por el porcentaje de elementos químicos de cada hacha, es muy complicado definir la procedencia de la materia prima para la elaboración de las hachas, antes se debería recoger muestras de las menas que se presumen fueron la fuente para la fabricación de las mismas, es decir se debería extraer muestras de las menas de cobre principalmente de México, Perú y Ecuador y que están cerca de los asentamientos de las culturas prehispánicas, ya que el comercio e intercambio de hachas entre las culturas ancestrales ubicadas en estos países era frecuente.

Como se mencionó anteriormente ya que aproximadamente el 60% de los datos son homogéneos se puede suponer que este porcentaje de hachas provienen de un solo lugar y probablemente pertenezcan a las culturas ecuatorianas.



### 3.2.3.2 Conglomerados por Variables.

Aunque no es muy frecuente realizar este tipo de conglomerados, se lo realiza a continuación para determinar los grupos de acuerdo a las variables.

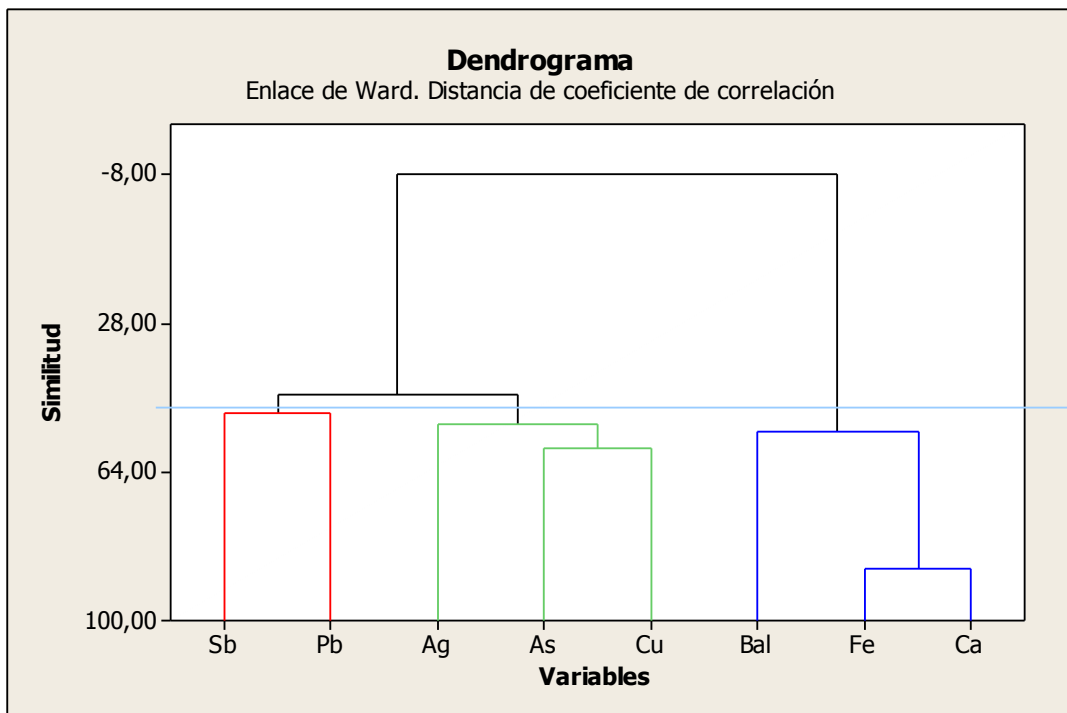


Gráfico 13: Dendrograma por variables.

Al igual que el estudio de conglomerados por observaciones se basó en el dendrograma, en esta ocasión también y de esta manera se obtienen 3 grupos claramente definidos. El análisis de Minitab nos ayuda a determinar los grupos analíticamente.

**Análisis de variables de conglomerados: Sb. Ag. Bal. As. Pb. Cu. Fe. Ca**

Distancia de coeficiente de correlación, Enlace de Ward  
Pasos de amalgamación

Paso	Número de grupos	Nivel de semejanza	Nivel de distancia	Grupos incorporados	Nuevo grupo	Número de obs. en el grupo nuevo
1	7	87,4760	0,25048	7	8	2
2	6	58,4705	0,83059	4	6	2
3	5	54,4573	0,91085	3	7	3
4	4	52,5808	0,94838	2	4	3
5	3	49,9756	1,00049	1	5	2
6	2	45,3405	1,09319	1	2	5
7	1	-7,9957	2,15991	1	3	8

Gráfico 14: Análisis de Conglomerados por Variables en MINITAB.

A partir de este análisis se obtiene la Tabla XXI donde se muestra la distribución de las variables en los tres conglomerados.

TABLA XXI: DISTRIBUCION DE LAS VARIABLES PARA CADA CONGLOMERADO		
Conglomerado 1	Conglomerado 2	Conglomerado 3
Sb	Ag	Bal
Pb	As	Fe
	Cu	Ca
<b>2</b>	<b>3</b>	<b>3</b>

Fuente: Distribución obtenida del análisis realizado en MINITAB

La interpretación en cuanto a los conglomerados por variables es simple ya que básicamente están agrupados los elementos químicos asociados entre sí, por ejemplo el Cobre y Plata en estado natural siempre están asociados, en otras palabras donde existe la presencia de Cobre siempre habrá Plata.

### 3.2.4 Análisis Factorial.

Uno de los objetivos fundamentales en este estudio es determinar las posibles variables latentes para esta investigación, para esto el análisis factorial es en sí el análisis más importante ya que permitirá inferir dichas variables o factores.

De acuerdo a lo expuesto en el Capítulo I en el análisis factorial de los componentes químicos hallados en hachas moneda de cobre se aplicó los dos métodos principales dentro de este análisis es decir el método de factor principal o método de componentes principales y el método de estimación máximo verosímil, pero el segundo método resultó no ser el adecuado para realizar el análisis.

A continuación se describirá y desarrollará el análisis factorial por el método de componentes principales con rotación equamax.

TABLA XXII: COMUNALIDADES		
	Inicial	Extracción
<b>Sb</b>	1,000	,720
<b>Ag</b>	1,000	,674
<b>Bal</b>	1,000	,961
<b>As</b>	1,000	,767
<b>Pb</b>	1,000	,508
<b>Cu</b>	1,000	,923
<b>Fe</b>	1,000	,841
<b>Ca</b>	1,000	,816

Método de extracción: Análisis de Componentes principales.

*Fuente: Análisis realizado en SPSS.*

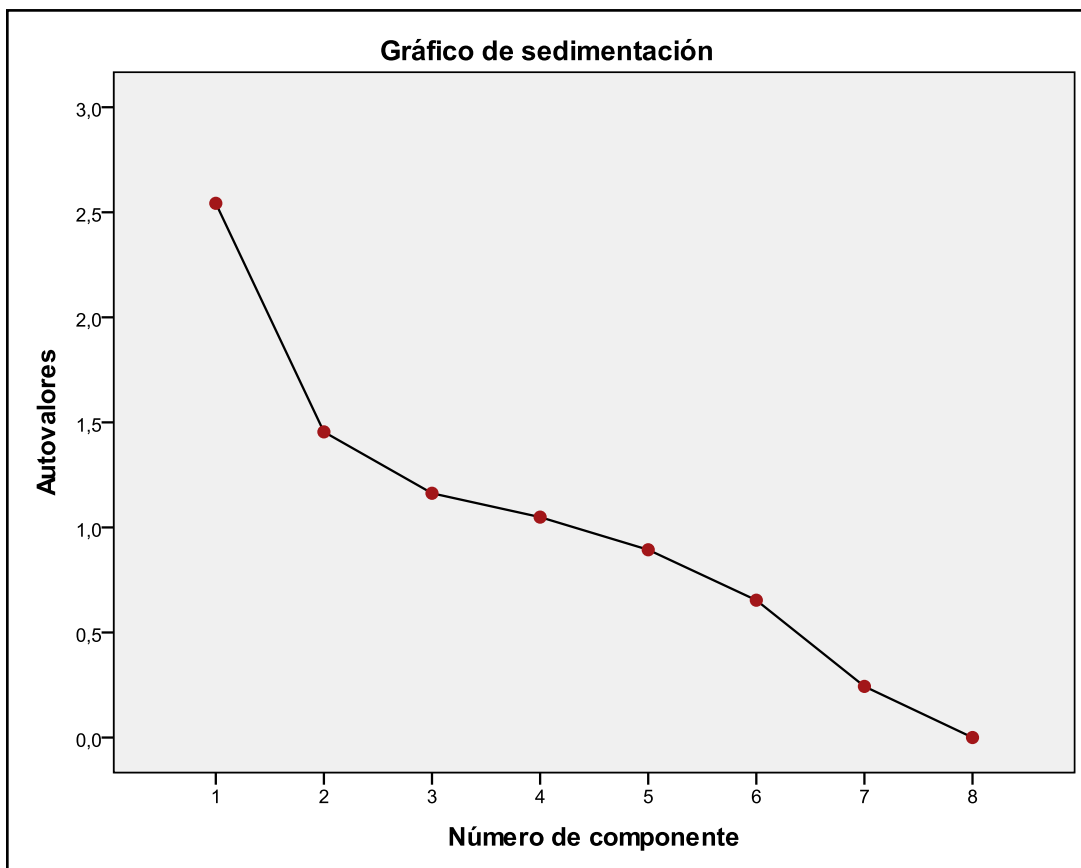
Las comunalidades de la Tabla XXII indican la proporción de variabilidad que puede ser explicada por el modelo factorial obtenido. Estudiando las comunalidades de la extracción se puede valorar cuales de las variables son las peor explicadas por el modelo. En el caso de las variables en estudio las variables Pb y Ag son las peor explicadas por el modelo, ya que sólo es capaz de reproducir el 51% y 67% respectivamente, de su variabilidad original.

TABLA XXIII: VARIANZA TOTAL EXPLICADA									
Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,543	<b>31,784</b>	31,784	2,543	31,784	31,784	2,071	25,890	25,890
2	1,455	<b>18,185</b>	49,969	1,455	18,185	49,969	1,885	23,561	49,451
3	1,163	<b>14,532</b>	64,501	1,163	14,532	64,501	1,135	14,187	63,638
4	1,049	<b>13,113</b>	<b>77,614</b>	1,049	13,113	77,614	1,118	13,976	77,614
5	,894	11,171	88,785						
6	,654	8,171	96,956						
7	,243	3,043	99,999						
8	9,746E-5	,001	100,000						

Método de extracción: Análisis de Componentes principales.

Fuente: Análisis realizado en SPSS.

La Tabla XXIII ofrece un listado del porcentaje de la varianza total explicado por cada uno de los componentes principales; por consiguiente el análisis extrae cuatro factores ya que cuatro de los autovalores extraídos son mayores que uno y dicha selección consigue explicar el 78% aproximadamente de la variabilidad total y además esta extracción de cuatro factores lo sugiere también la grafica de sedimentación.



*Gráfico 15: Grafica de sedimentación para Análisis Factorial realizado en SPSS.*

TABLA XXIV: MATRIZ DE CORRELACIONES <sup>a</sup>									
	Sb	Ag	Bal	As	Pb	Cu	Fe	Ca	
<b>Correlación</b>	<b>Sb</b>	1,000	-,137	-,058	,076	,000	,059	-,134	-,108
	<b>Ag</b>	-,137	1,000	-,026	,138	-,026	,024	-,095	-,098
	<b>Bal</b>	-,058	-,026	1,000	-,242	-,170	-,985	,294	,215
	<b>As</b>	,076	,138	-,242	1,000	-,034	,169	,048	,027
	<b>Pb</b>	,000	-,026	-,170	-,034	1,000	,049	,086	,107
	<b>Cu</b>	,059	,024	-,985	,169	,049	1,000	-,399	-,305
	<b>Fe</b>	-,134	-,095	,294	,048	,086	-,399	1,000	,750
	<b>Ca</b>	-,108	-,098	,215	,027	,107	-,305	,750	1,000
<b>Sig. (Unilateral)</b>	<b>Sb</b>		,028	,212	,145	,497	,207	,032	,067
	<b>Ag</b>	,028		,359	,027	,360	,371	,093	,087
	<b>Bal</b>	,212	,359		,000	,009	,000	,000	,001
	<b>As</b>	,145	,027	,000		,320	,009	,255	,352
	<b>Pb</b>	,497	,360	,009	,320		,249	,117	,069
	<b>Cu</b>	,207	,371	,000	,009	,249		,000	,000
	<b>Fe</b>	,032	,093	,000	,255	,117	,000		,000
	<b>Ca</b>	,067	,087	,001	,352	,069	,000	,000	
a. Determinante = 6,25E-005									

Fuente: Análisis realizado en SPSS.

La matriz de correlación ilustrada en la Tabla XXIV es la matriz de la cual parte el análisis; además en la parte inferior de esta tabla se muestra el nivel crítico unilateral asociado a cada coeficiente de correlación. Un nivel crítico menor que 0,05 indica que la correlación poblacional entre el correspondiente par de variables puede ser considerada significativamente distinta de cero, lo deseable por tanto es hallar muchos niveles críticos pequeños, como se muestra en la tabla existe una gran proporción de niveles de significancia próximos a cero, además el determinante de la matriz de correlaciones es aproximadamente cero (0,0000625), lo que indica que las variables de

la matriz de datos están linealmente relacionadas, lo cual es buen síntoma de la idoneidad del análisis. Para confirmar dicha idoneidad es necesario valorar la bondad de ajuste o adecuación de los datos analizados a un modelo factorial.

La tabla XXV muestra la medida de adecuación muestral KMO y la prueba de esfericidad de Bartlett, lo cual confirma que el modelo factorial aplicado es el adecuado para explicar los datos ya que KMO es diferente de cero y además la prueba de esfericidad de Bartlett contrasta la hipótesis nula de que la matriz de correlaciones observada es en realidad una matriz identidad. Asumiendo que los datos provienen de una distribución normal multivariante el estadístico de Bartlett asegura que el modelo factorial aplicado es el adecuado para explicar los datos ya que su nivel de significancia es prácticamente cero y a su vez es menor que el nivel de significancia de 0,05 se rechaza la hipótesis nula de esfericidad y se concluye que la muestra proviene de una distribución normal multivariante.

TABLA XXV: PRUEBA KMO Y PRUEBA DE BARTLETT		
Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,143
Prueba de esfericidad de	Chi-cuadrado aproximado	1834,334
Bartlett	Gl	28
	Sig.	,000

*Fuente: Análisis realizado en SPSS.*

A continuación se muestra la solución factorial propiamente dicha.



TABLA XXVI: MATRIZ DE FACTORES Y MATRIZ DE FACTORES ROTADOS								
MATRIZ DE FACTORIAL					MATRIZ DE FACTORES ROTADOS			
	FACTORES				FACTORES			
	1	2	3	4	1	2	3	4
<b>Sb</b>	-0,174	-0,111	-0,535	0,625	0,01	-0,148	0,791	0,268
<b>Ag</b>	-0,118	-0,067	0,808	-0,056	-0,014	-0,157	-0,712	0,377
<b>Bal</b>	0,841	-0,495	0,046	0,075	-0,964	0,148	-0,017	-0,092
<b>As</b>	-0,214	0,42	0,38	0,632	0,29	0,206	-0,005	0,8
<b>Pb</b>	-0,028	0,446	-0,273	-0,484	0,401	0,315	-0,031	-0,498
<b>Cu</b>	-0,887	0,357	-0,044	-0,075	0,913	-0,284	0,022	0,087
<b>Fe</b>	0,724	0,557	0,018	0,081	-0,217	0,89	-0,023	0,021
<b>Ca</b>	0,658	0,616	-0,023	0,056	-0,123	0,895	-0,001	-0,005

Fuente: Análisis realizado en SPSS.

Comparando las saturaciones relativas tanto de la matriz de factores como de la matriz de factores rotados en cada uno de los cuatro factores se puede apreciar que el primer factor está constituido por las variables Bal y Cu, estas variables son las que mayor peso tienen ya que los porcentajes de contenido de estos elementos químicos hallados en las hachas moneda son los más relevantes dentro del estudio, estas variables saturan en un único factor por lo que constituyen un grupo diferenciado de variables dentro de la matriz de correlaciones pero es importante mencionar que la relación existente entre estas dos variables es inversamente proporcional. Este factor parece reflejar “Composición Principal” de las hachas moneda de cobre.

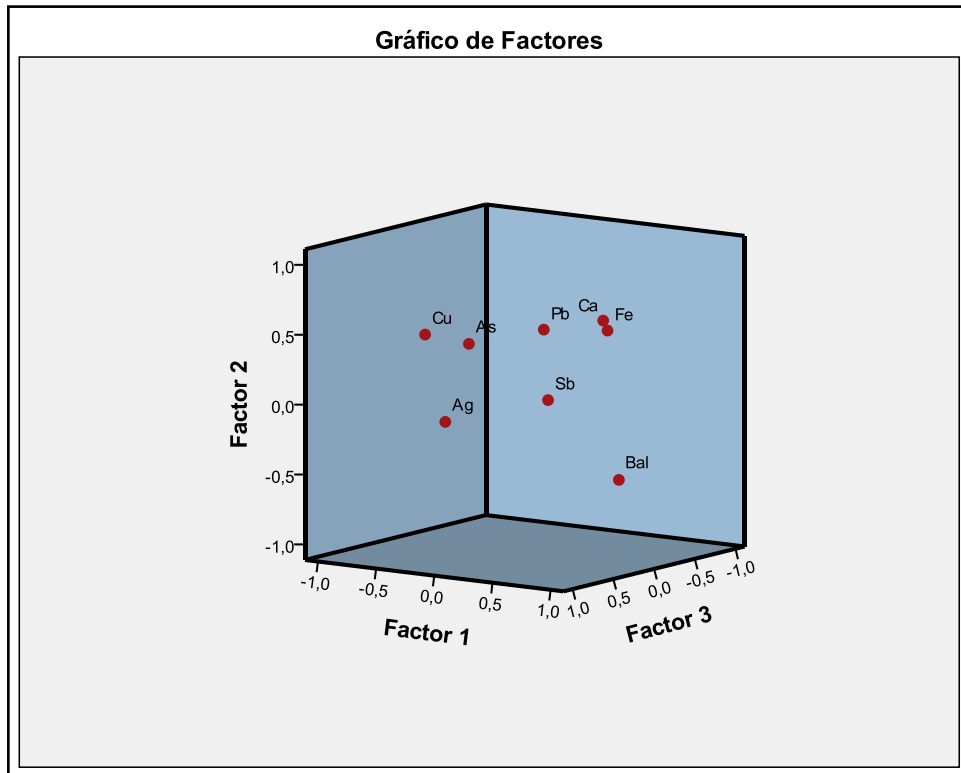
El segundo factor recoge las variables Ca y Fe, estos elementos son el cuarto y quinto respectivamente, más abundantes sobre la corteza terrestre y obviamente es muy fácil que estos elementos se hayan hallado dentro de las menas de cobre de donde se

extrajo la materia prima para la elaboración de las hachas moneda, químicamente se los podría representar por “Metales” ya que ambos pertenecen a este grupo.

El tercer factor está saturado únicamente por Ag así que este factor llevaría este único nombre.

Finalmente el cuarto factor está saturado por As y Sb, el arsénico se encuentra distribuido ampliamente sobre la corteza terrestre mientras que el antimonio no, pero a menudo se lo halla como una mezcla isomorfa con arsénico y al igual que el factor dos químicamente se los podría llamar “Semi-Conductores”.

Tanto las comunalidades (Tabla XXII), como la matriz de factores (Tabla XXVI) indican que Pb no está dentro del análisis factorial ya que el modelo no explica adecuadamente a esta variable.



**Gráfico 16:** Grafica de sedimentación para Análisis Factorial realizado en SPSS.

Finalmente el Gráfico 16 confirma lo antes dicho y analizado en la Tabla XXIII.

## **IV. CAPITULO**

### **CONCLUSIONES Y RECOMENDACIONES**

#### **4.1 CONCLUSIONES**

- El estudio de posibles variables redundantes detectó la presencia como tales de las siguientes variables: Au, Se, Ti, K, Sn, Bi y Ni. Por tal motivo se redujo la dimensionalidad, eliminando estas variables, de 15 a 8.
- El análisis descriptivo de la posible presencia de datos atípicos mediante la técnica de las direcciones de máxima y mínima kurtosis descartó la presencia de tales valores, garantizando de este modo conclusiones validas y objetivas a partir de los resultados de la investigación.
- En el estudio de la regresión lineal múltiple de cada variable sobre las restantes se escogieron cinco modelos con el propósito de hacer predicciones en cuanto sus coeficientes de determinación al cuadrado están entre el 96% y 99%. Con estos modelos se pueden prever valores de las variables Ba, As, Pb, Cu y Fe.

- Con la ayuda del análisis de conglomerados se logró localizar seis grupos homogéneos de hachas, dichos grupos están constituidos por hachas que poseen características químicas similares. Además es importante mencionar que el 60% de las hachas se separan entre ellas una distancia inferior a 2 y esto indica que la distribución de las observaciones es aproximadamente homogénea. A partir de este porcentaje se puede establecer que el 60% de las hachas analizadas pertenecen a una sola cultura y el 40% restante posiblemente proviene del comercio e intercambio con otras culturas de la costa del Pacífico.
- De acuerdo a los factores o variables latentes halladas se concluye que la aleación más sobresaliente dentro de los elementos químicos hallados en la composición de hachas moneda es la Enargita (Cobre – Arsénico). Este puede ser uno de los indicios más sobresalientes del estudio ya que de esta manera se puede empezar a localizar yacimientos de cobre en las regiones donde se asentaron las culturas que elaboraron y realizaron procesos de intercambio comercial con hachas moneda de cobre en las que la enargita tenga una gran concentración. De esta manera se podrá ubicar la procedencia de dichas piezas arqueológicas.

- Finalmente, no se pudo cumplir con el objetivo de la localización y determinación del tipo de cobre utilizado para la fabricación de las hachas moneda, ya que durante el desarrollo de la investigación se observó que para cumplir con el objetivo planteado es necesario tomar muestras de Cobre de los yacimientos cercanos a donde se asentaron las culturas que hicieron uso de este artefacto y además ampliar la muestra de hachas moneda provenientes de las distintas culturas que habitaban las costas del Pacífico desde México hasta Perú, para de esta manera poder comparar los grupos con respecto al tipo de cobre presente en la composición química de las hachas moneda y posteriormente hacer inferencias respecto al origen y al tipo de cobre.

## **4.2 RECOMENDACIONES**

- Dar mayor énfasis en la aplicación de Técnicas Estadísticas Multivariantes dentro del estudio de temas que anteriormente eran considerados netamente parte de las ciencias sociales pero que como se ha demostrado en esta investigación la aplicación de tales técnicas a la Arqueología a arrojado resultados interesantes.

- Continuar con esta investigación en las regiones de los países donde han existido hallazgos de hachas moneda ya que de esta manera se podrá definir con exactitud la procedencia y ubicación de estos artefactos de acuerdo a su composición química.
- Socializar los resultados obtenidos dentro de esta investigación con Ecuacorriente S. A que fue la empresa impulsadora e interesada en el estudio de las hachas moneda de cobre, como con entidades gubernamentales relacionadas con la arqueología para que de esta manera se pueda impulsar este tipo de estudios que no sólo promueven la aplicación de la cultura estadística, sino que además fortalecen y contribuyen a la historia e identidad del Ecuador.

## RESUMEN

El objetivo fundamental de esta investigación fue realizar un análisis estadístico multivariante de los principales componentes químicos encontrados sobre hachas moneda de Cobre para determinar diferentes grupos de acuerdo a su composición química. La aplicación de tales técnicas se han hecho por tanto dentro de un contexto arqueológico, las hachas analizadas pertenecen a una colección privada de Ecuacorriente S. A.

Esta investigación se realizó bajo los parámetros de una investigación cuantitativa experimental.

Para iniciar el estudio fue necesario realizar un proceso de restauración a las piezas arqueológicas previo a la medición de la composición química de las hachas moneda de cobre con la máquina Thermo Niton, todo este procedimiento fue realizado en el laboratorio arqueológico de ECSA ubicado en la Parroquia Tundayme, Cantón El Pangui, Provincia de Zamora Chinchipe.

En el estudio de la regresión lineal múltiple de cada variable sobre las restantes se escogieron cinco modelos con el propósito de hacer predicciones en cuanto sus coeficientes de determinación al cuadrado están entre el 96% y 99%.

Al aplicar el análisis de conglomerados se detectó la presencia de 6 grupos cada uno de los cuales está constituido por hachas moneda de características químicas similares.

Con la aplicación del análisis factorial se detectó la presencia de variables latentes las mismas que sugieren que el 60% de las hachas moneda están homogéneamente distribuidas y por ende están formadas fundamentalmente por la aleación de Cobre y Arsénico.

Por tanto se concluye que las técnicas estadísticas multivariantes aplicadas en la investigación fueron de gran ayuda para determinar los grupos de hachas formados de acuerdo a su composición química.

Finalmente se recomienda impulsar este tipo de estudios que no sólo promueven la aplicación de la cultura estadística, sino que además fortalecen y contribuyen a la historia e identidad del Ecuador.



## **ABSTRACT**

The topic is Multivariate Statistical Analysis of the Main Chemical Components of Copper Coin Axes for determining different groups according to their composition.

This research work was developed to implement a statistical culture within an archaeological context focuses on the analysis and the study of the chemical composition of Coin Axes.

The main goal was to carry out a multivariate statistical analysis of the principal components found about copper coin axes with the purpose to determine different groups according to their chemical composition. According to the application of those techniques, it was determined that axes analyzed belong to a private collection of Ecuacorriente S. A.

This study was carried out under the parameters of quantitative experimental research.

For beginning, this work was necessary to restore archaeological pieces before measuring the chemical composition of copper coin axes with Thermo Niton machine, this procedure was developed in the ECSA archaeological laboratory, which is located in Tundayme parish, El Pangui canton, Zamora Chinchipe province.

In the multiple linear regression study of each variable on the remaining, five models were chosen in order to make predictions about their coefficients to determine squared are between 96% and 99%.

By applying cluster analysis was showed the presence of 6 groups, each one from which consists of coin axes of similar chemical characteristics.

With the application of factor analysis was detected the presence of latent variables the same as suggesting that 60% of the axes are homogeneously distributed and therefore are composed mainly by the alloy of copper and arsenic.

In conclusion, the multivariate statistical techniques applied in the research were of great help to the identify groups of axes formed according to their chemical composition.

Finally it is recommended to promote this kind of study not only promote the application of statistical culture, but also reinforce and contribute to history and identity of Ecuador.

## **BIBLIOGRAFIA**

1. **BARCELÓ, J.**, *Arqueología Lógica y Estadística.*, Barcelona-España., Servei de Publicacions de l'Universitat Autònoma (ISBN 84-7488-702-X)., 1989., p., 46-49, 84-78, 88-102
2. **CUADRAS, M.**, *Nuevos Métodos de Análisis Multivariante.*, Barcelona-España., CMC Editions., 2012., p., 91-112.
3. **DOWNIE, N. y HEART, R.**, *Métodos Estadísticos Aplicados.*, México DF-Mexico., Harla., 1986., p., 117-156.
4. **PEÑA, D.**, *Análisis de Datos Multivariantes.*, Madrid-España., McGrawHill., 2002., p., 67-171, 227-253, 335-361
5. **SHENAN, S.**, *Arqueología Cuantitativa.*, Barcelona-España., Editorial Critica., 1992., p., 242-294

## **BIBLIOGRAFIA INTERNET**

### **6. ANALISIS FACTORIAL**

<http://ciberconta.unizar.es/LECCION/factorial/FACTORIALEC.pdf>

2012-04-21

### **7. ANALISIS MULTIVARIADO**

[http://www.ucm.es/info/socivmyt/paginas/materiales/analisis\\_datosy  
Multivariable/SPSS.pdf](http://www.ucm.es/info/socivmyt/paginas/materiales/analisis_datosy_Multivariable/SPSS.pdf)

2012-03-19

### **8. INFORMACION DE CULTURAS PRE-HISPANICAS**

<http://www.banrepcultural.org/book/export/html/26461>

2012-03-25

### **9. INFORMACION DE HACHAS MONEDA**

<http://www.encyclopediadelecuador.com/temasOpt>

2012-03-25

### **10. MUSEOS DEL ECUADOR**

<http://www.museos-ecuador.gob.ec>

2012-02-13

## **11. PROCESOS DE RESTAURACION**

<http://es.scribd.com/doc/78430861/Espana-T-et-al-Limpieza-y-restauracion-electroquimica-objetos>

2012-04-21

## **12. TABLA PERIODICA QUIMICA**

<http://www.lenntech.es/periodica/elementos/ba.htm>

2012-01-27

# **ANEXOS**

**ACUERDO DE  
COOPERACION  
ACADEMICO  
CIENTIFICO**

**ACUERDO DE COOPERACIÓN  
ACADÉMICO-CIENTÍFICO  
No. ECSA.A.2011.004**

**ENTRE**



Ecuacorriente S.A.

**ECUACORRIENTE S.A.**

**ESCUELA SUPERIOR POLITECNICA DE  
CHIMBORAZO**

**Y**

**MARIA ANGELICA VILLA MOYOTA**

46

### ACUERDO DE COOPERACIÓN ACADÉMICO-CIENTÍFICA

Comparecen a la suscripción del presente instrumento por una parte, la compañía **ECUACORRIENTE S.A.**, a través de su Apoderado Especial el Sr. Warren McLean, en adelante ECSA; la **ESCUELA SUPERIOR POLITECNICA DE CHIMBORAZO** (en adelante la ESPOCH) representada por el Ms. Romeo Rodríguez, en su calidad de Rector, y por otra parte la Srta. **MARÍA ANGÉLICA VILLA MOYOTA** por sus propios y personales derechos, en adelante TESISTA.

Quienes suscriben el presente instrumento declaran ser hábiles para contratar y obligarse, de manera libre y voluntaria en las calidades que comparecen, y convienen en la celebración del presente Acuerdo de Cooperación Académico-Científica, contenido en los siguientes términos y condiciones:

#### PRIMERA.- ANTECEDENTES.-

- 1.1. ECUACORRIENTE S.A., es una compañía constituida y existente bajo las leyes de la República del Ecuador, cuyo objeto social es la ejecución de actividades en todas las fases del área minera.
- 1.2. ECUACORRIENTE S.A. posee concesiones mineras legalmente otorgadas por el Estado Ecuatoriano, ubicada en la jurisdicción de la Provincia de Zamora Chinchipe.
- 1.3. Con fecha 16 de Julio de 2008, ECUACORRIENTE S.A. y la ESPOCH suscribieron un Convenio General de Cooperación Interinstitucional. Uno de los objetivos del referido Convenio fue promover y desarrollar la formación académica, científica y profesional de los estudiantes de la ESPOCH.
- 1.4. La Tesista es egresada de la carrera de Ingeniería en Estadística Informática, Facultad de Ciencias, Escuela Física y Matemática de la ESPOCH y ha manifestado su interés de desarrollar su proyecto de tesis titulada "**ANÁLISIS ESTADÍSTICO MULTIVARIANTE DE LOS PRINCIPALES COMPONENTES QUÍMICOS SOBRE HACHAS MONEDA DE COBRE PARA DETERMINAR DIFERENTES GRUPOS DE ACUERDO A SU COMPOSICIÓN**", previo a la obtención del título de Ingeniería en Estadística Informática.
- 1.5. Mediante Oficio No. 124.EFM.FC.11 de fecha 25 de Julio de 2011, LA TESISTA, a través del Director de Escuela, Dr. Richard Pachacama Mg.C., solicitó a ECSA le proporcione las facilidades para que desarrolle su tesis.
- 1.6. ECSA ha expresado su voluntad de dar las facilidades adecuadas a la TESISTA, en los términos y condiciones que han sido aceptadas en este mismo instrumento por las partes comparecientes, para que la TESISTA pueda desarrollar el contenido de su tesis para la obtención del título de grado.

#### SEGUNDA.- OBJETO.-

Con los antecedentes expuestos, por medio del presente instrumento, las partes se comprometen a:

#### ECSA:

Prestar las facilidades descritas en el presente instrumento para que la TESISTA pueda llevar a cabo las actividades académico-científicas para desarrollar el contenido de su tesis para obtener su título de grado.

MA



**TESISTA:**

Desarrollar el contenido de su tesis de grado, de acuerdo a los términos y condiciones del presente instrumento, y en fiel respeto de las normas y políticas corporativas de ECSA y de los reglamentos y normativa de la ESPOCH.

ECSA y la TESISTA podrían desarrollar un plan de visita a museos y potencial investigación de materiales similares a las hachas moneda. Para el efecto, las Partes establecerán y acordarán un cronograma y presupuesto antes de emprenderlo.

**TERCERA.- OBLIGACIONES DIRECTAS DE LAS PARTES.-**

Sin perjuicio de las demás obligaciones que se deriven del presente Acuerdo, son obligaciones de cada una de las partes especialmente las siguientes:

**3.1 Son obligaciones de ECSA:**

3.1.1 Proporcionar a la TESISTA alojamiento y alimentación en el Campamento de ECSA, o en el lugar que ECSA considere de acuerdo a las actividades de la TESISTA hasta por un máximo de 2 semanas.

3.1.2. Proporcionar a la TESISTA transporte Quito – Mirador – Quito, dependiendo de las necesidades de las actividades a desarrollarse debidamente justificadas.

3.1.2 Permitir a la TESISTA el uso del equipo XRF, siempre bajo coordinación del tutor o personal de ECSA.

3.1.3 Proporcionar a la TESISTA en sitio las hachas moneda.

3.1.4 Permitir a la TESISTA el uso de las instalaciones del campamento Mirador.

3.1.5 Asignar en calidad de Tutor a un profesional de la empresa relacionado con el título de la tesis, quien orientará el proceso de investigación.

**3.2 Son obligaciones de la TESISTA:**

3.2.1 Tratar al personal, contratistas, empleados y comunidad relacionada a la operación de ECSA con el mayor respeto y consideración, dentro de las más estrictas normas de la moral y buenas costumbres.

3.2.2 Desarrollar su tesis en estricto cumplimiento de las normas internas y políticas corporativas de ECSA.

3.2.3 Desarrollar su tesis con apego a la legislación que fuera aplicable y demás normas aplicables a las actividades de ECSA; en especial, la TESISTA generará un análisis estadístico y una base de datos de las hachas moneda en beneficio de ECSA; para lo cual, estará provisto de su laptop, flash memory y software.

3.2.4 Mantener la más absoluta reserva y confidencialidad respecto a toda la información que conozca sobre el manejo o las actividades que realiza ECSA, tales como, datos, registros o documentos que conozca y pudieran revelar la situación de ECSA, relaciones o estrategias comerciales, asuntos concernientes al personal, etc.; que llegaren a su conocimiento durante el cumplimiento de sus labores; y a, publicarlos únicamente con autorización de ECSA.

MA

MA

3.2.5 Manejar y mantener con absoluta responsabilidad todos los bienes que ECSA haya puesto a su disposición y evitar toda clase de acciones u omisiones que pudiera resultar dañosas a los bienes e intereses de ECSA o de terceras personas. En el caso de pérdida, daño o deterioro de dichos bienes, que sea imputable a la TESISISTA está obligada al inmediato pago a ECSA del costo de reposición de dichos equipos.

3.2.6 Utilizar permanentemente los equipos de protección personal que fueran suministrados por ECSA, a criterio de ésta última.

3.2.7 Asumir por su propia cuenta, cualquier riesgo sobre su persona o bienes, originado o como consecuencia de la ejecución de este instrumento, deslindando a ECSA de cualquier responsabilidad directa o indirecta de tal situación.

3.2.8 Asumir por su propia cuenta, cualquier costo o gasto que no hubiere sido contemplado dentro de los compromisos asumidos por ECSA.

3.2.9 LA TESISISTA, sus representantes y demás relacionados, no podrán presentarse en ningún momento y bajo ninguna circunstancia como empleados, representantes, contratista o agentes de ECSA. LA TESISISTA no tendrá poder alguno para comprometer, expresa o tácitamente, a ECSA con terceros.

3.2.10 A la terminación de su tesis debidamente aprobada por la UNIVERSIDAD, la TESISISTA entregará a ECSA, sin costo alguno dos (2) copias de la misma.

3.2.11 Así mismo, la TESISISTA se compromete a entregar a ECSA, sin costo alguno, cualquier información que ésta última le solicite relacionada con el contenido de su tesis. Esta obligación subsistirá por el plazo de tres (3) años contados a partir de la terminación del presente Acuerdo.

### **3.3 Son obligaciones de la ESPOCH:**

3.3.1 Supervisar el desempeño de las actividades de la TESISISTA en relación a este instrumento.

### **CUARTA: PLAZO DE DURACIÓN Y TERMINACIÓN.-**

El plazo de vigencia del presente Acuerdo de Cooperación Académico-Científico es el necesario para que la TESISISTA pueda llevar a cabo las actividades académico-científicas para el desarrollo de su trabajo de investigación de su tesis, pero no más allá de dos (2) semanas contadas a partir de la fecha de suscripción de este Acuerdo.

No obstante el plazo de vigencia, ECSA podrá dar por terminado de forma unilateral este Acuerdo sin responsabilidad ni incurrir en indemnización de ninguna naturaleza, en el caso de que la TESISISTA incumpla una o todas las obligaciones a las que se han comprometido mediante el presente instrumento. Para efectos de dicha terminación, bastará una simple notificación cursada por ECSA a las otras partes en este sentido.

### **QUINTA.-RELACIÓN CONTRACTUAL CIVIL.-**

Las partes acuerdan y declaran que el vínculo que se crea mediante el presente instrumento y la relación entre las partes es de naturaleza civil, y por lo tanto se regula por las disposiciones constantes en el Código Civil. En tal razón, las partes expresamente reconocen y declaran que entre la TESISISTA Y ECSA, y sus compañías relacionadas, funcionarios, representantes o mandatarios, no existe ni existirá relación

124



laboral alguna, por lo que ECSA no asume ninguna obligación frente a la TESISTA, que no sean aquellas que expresamente constan descritas en el presente instrumento.

LA TESISTA será responsable exclusivo frente a sí mismo por accidentes, enfermedades, percances, daños y/o similares que pudiere sufrir en su integridad física o moral, que se pueden derivar de su permanencia en el lugar en donde desarrolle su tesis designado por ECSA y/o sus desplazamientos. En consecuencia, la TESISTA exonera totalmente a ECSA y a sus compañías relacionadas de cualquier responsabilidad, que se pueda derivar directa o indirectamente de tal situación.

LA TESISTA estará cubierto con una póliza de accidentes personales contratado por ECSA.

El presente Acuerdo de Cooperación Académico-Científica, por su propia naturaleza, no tiene el carácter de oneroso ni retributivo.

#### **SEXTA.- CLÁUSULA ARBITRAL.-**

Para la resolución de toda controversia o diferencia derivada de la celebración, aplicación, interpretación, ejecución o terminación del presente instrumento, que no pueda resolverse directamente por las partes, éstas renuncian a su domicilio y fuero común y acuerdan someter la resolución de dichas controversias o diferencias a arbitraje administrado por un Tribunal Arbitral de la Universidad Técnica Particular de Loja. Este Tribunal estará integrado por un (1) solo árbitro que será nombrado de acuerdo a las normas de este Tribunal. El procedimiento arbitral será en derecho y se sujetará a lo dispuesto en la Ley de Arbitraje y Mediación, y a las regulaciones del Tribunal Arbitral de la Universidad Técnica Particular de Loja.

#### **SÉPTIMA.- MISCELÁNEOS.-**

7.1 Encabezamientos: los encabezamientos de este documento están impresos sólo como referencia y no afectarán la interpretación del contrato.

7.2 Mayúsculas y plurales: las palabras o términos utilizados en mayúsculas en este instrumento están impresos solo como referencia e incluirán a las minúsculas, sin afectar la interpretación del mismo. Las palabras utilizadas en plural incluirán su singular y viceversa.

7.3 Modificación: toda modificación o reforma al presente instrumento deberá adoptarse por acuerdo escrito entre las partes.

7.4 Continuidad: el presente Acuerdo es plenamente válido, de acuerdo a la ley aplicable. En caso de que cualquier parte o cláusula de este Acuerdo fuera considerada por cualquier Corte, juez o tribunal como violatoria a la ley aplicable, o no válida o inejecutable bajo ella, este Acuerdo continuará plenamente válido y se entenderá modificado para que la parte declarada en violación de la ley o inválida o inejecutable bajo ella, carezca de fuerza o efecto, y no afecte la validez o la posibilidad de ejecutar las restantes estipulaciones del Acuerdo.

7.5 Omisión en el ejercicio de la acción de exigir el cumplimiento: el incumplimiento por cualquier de las partes para requerir el cumplimiento o para hacer cumplir cualquiera de los términos de este Acuerdo, no será considerado como una renuncia a los derechos aquí conferidos que impida la ejecución de tales estipulaciones o la recuperación de daños por su violación.

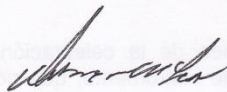
MA

7.6 Terceros Beneficiarios: el presente instrumento está celebrado para beneficio exclusivo de las Partes contratantes, por lo que nada de lo aquí expresado o mencionado será entendido como un permiso, reconocimiento, o autorización a un tercero para reclamar cualquier derecho en base a este documento.

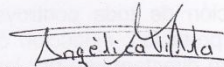
7.7 Ley Aplicable: las Partes se sujetan expresamente a lo acordado en este Acuerdo y, en todo aquello que no estuviere especificado, a las normas del Código Civil.

**OCTAVA.- RECONOCIMIENTO.-**

Las partes acuerdan que los términos y condiciones de este instrumento han sido discutidos y totalmente aceptados por cada parte compareciente.  
Para constancia de su plena aceptación a todas y cada una de las cláusulas de este documento, las partes firman este Acuerdo, en tres ejemplares de igual contenido y valor, a los 19 días del mes de Septiembre de 2011.



Sr. Warren McLean  
Apoderado Especial  
ECUACORRIENTE S.A.



Srta. María Angélica Villa Moyota  
TESISTA

124  
22



Ms. Romeo Rodríguez  
Rector  
ESCUELA SUPERIOR POLITECNICA  
DE CHIMBORAZO

**RESOLUCION PARA  
LA SUBSCRIPCION  
DEL ACUERDO DE  
COOPERACION**





epoch

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO  
SECRETARIA GENERAL

Panamericana Sur Km. 1 1/2 (03) 2605 902- 2605 917 Ext. 108 sgeneral@epoch.edu.ec

ACREDITADA

Oficio 526.SG.ESPOCH.2011  
Noviembre 30, 2011

RESOLUCIÓN 370.CP.2011  
Noviembre 29, 2011

Doctor  
Romeo Rodríguez Cárdenas  
**RECTOR DE LA ESPOCH**  
Presente

De mi consideración:

Informo a usted que el Consejo Politécnico, en sesión ordinaria realizada el martes 29 de noviembre de 2011, conoció el Oficio 674.PROCU.2011, del 23 de noviembre de 2011, suscrito por la doctora Marcela Parra C., Abogada de Procuraduría, mediante el cual informa que el ACUERDO DE COOPERACIÓN ACADÉMICO – CIENTÍFICO No.ECSA.A.2011.004 ENTRE LA ESPOCH, ECUACORRIENTE S.A Y LA SEÑORITA MARÍA ANGÉLICA VILLA MOYOTA se deduce el CONVENIO GENERAL DE COOPERACIÓN INTERINSTITUCIONAL ENTRE LA ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO, EXPLORCOBRES S.A Y ECUACORRIENTE S.A.; razón por la cual, el presente Acuerdo debe ser ratificado e incorporarse los respectivos documentos habilitantes; al respecto este Organismo por unanimidad, **RESOLVIÓ:** Autorizar al señor Rector suscriba el ACUERDO DE COOPERACIÓN ACADÉMICO – CIENTÍFICO No.ECSA.A.2011.004 ENTRE LA ESPOCH, ECUACORRIENTE S.A Y LA SEÑORITA MARÍA ANGÉLICA VILLA MOYOTA, que se deduce del CONVENIO GENERAL DE COOPERACIÓN INTERINSTITUCIONAL ENTRE LA ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO, EXPLORCOBRES S.A Y ECUACORRIENTE S.A.; cuyo objeto es: que la señorita María Angélica Villa Moyota pueda llevar a cabo las actividades académico- científicas para el desarrollo del trabajo de investigación de su tesis, por lo que, se autoriza al señor Rector suscriba el presente convenio.

Particular que comunica para los fines legales consiguientes.

Atentamente,

Dr. Julio Falconí Mejía  
**SECRETARIO GENERAL**



Copia: Sra. Vicerrectora Académica, Sr. Vicerrector Investigación y Desarrollo, Sres. Decanos, Sra. Procuradora, Sr. Secretario Académico, Sr. Dir. Dpto. Desarrollo Humano, Sr. Dir. Dpto. Financiero, Srta. Tesorera, Sr. Dir. Com. Vinculación con la Colectividad, Sres. Ecuacorriente, Srta. María Villa.

Beatriz F.

*"Saber para Ser"*

## FOTOS DE LA INVESTIFACION.

### INSTALACIONES DE ECSA.

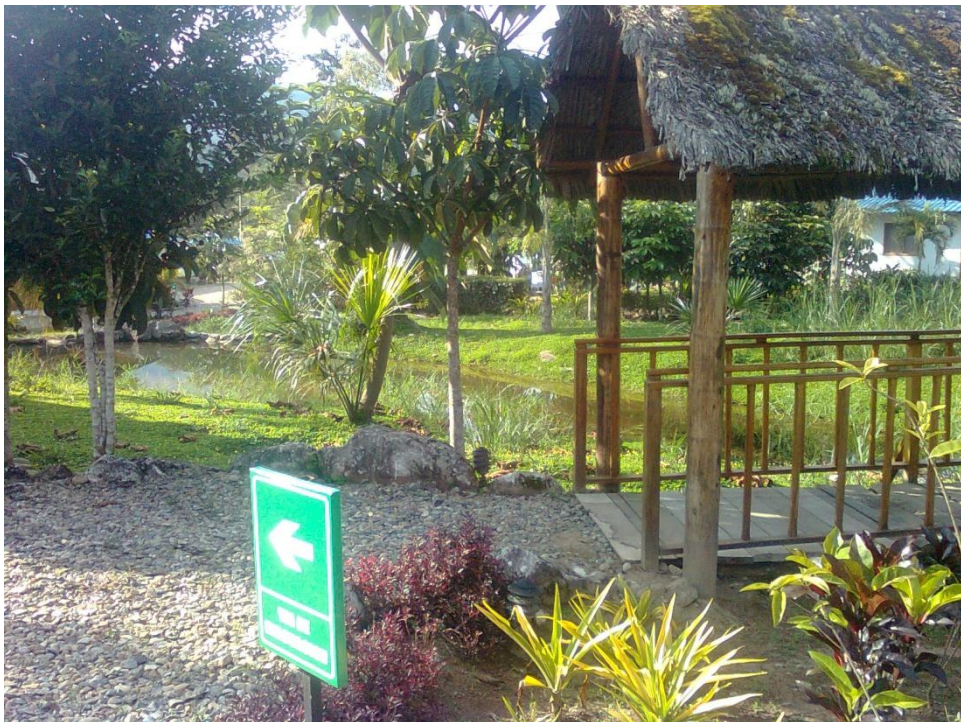


**Foto 6:** Instalaciones de ECSA (1).





**Foto 7:** Instalaciones de ECSA (2).



**Foto 8:** Instalaciones de ECSA (3).





**Foto 9:** Personal del Campamento Mirador.



**Foto 10:** Tesista.

**PROCESO DE RESTAURACION Y TOMA DE DATOS.**



*Foto 11: Primer contacto con Hachas Moneda de Cobre.*



*Foto 12: Máquina Thermo Niton*



**Foto 1:** Limpieza con bisturí de materias adheridas por la corrosión



**Foto 2:** Aplicación de HCL al 10%.



*Foto 3: Muestra pulida para retirar el oxido*



*Foto 4: Hachas con aplicación de Benzotriazol*





*Foto 5: Análisis con la Thermo Niton*