



# **ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**FACULTAD DE CIENCIAS**

**CARRERA ESTADÍSTICA**

## **“PRONÓSTICOS EN LOS CONSUMOS DE AGUA POTABLE SEGÚN REGISTROS HISTÓRICOS GENERADOS EN LA PARROQUIA CEBADAS DURANTE EL PERIODO 2017-2021”**

**Trabajo de Integración Curricular**

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

**INGENIERA EN ESTADÍSTICA INFORMÁTICA**

**AUTORA:**

**GISSELA MARISOL PEÑAFIEL MALAN**

Riobamba – Ecuador

2022



**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**FACULTAD DE CIENCIAS**

**CARRERA ESTADÍSTICA**

**“PRONÓSTICOS EN LOS CONSUMOS DE AGUA POTABLE  
SEGÚN REGISTROS HISTÓRICOS GENERADOS EN LA  
PARROQUIA CEBADAS DURANTE EL PERIODO 2017-2021”**

**Trabajo de Integración Curricular**

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

**INGENIERA EN ESTADÍSTICA INFORMÁTICA**

**AUTORA:** GISSELA MARISOL PEÑAFIEL MALAN

**DIRECTORA:** Ing. NANCY ELIZABETH CHARIGUAMAN MAURISACA MSc.

Riobamba – Ecuador

2022

© 2022, Gissela Marisol Peñafiel Malan

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el derecho de Autor.

Yo, GISSELA MARISOL PEÑAFIEL MALAN, declaro que el presente Trabajo de Integración Curricular es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autora asumo la responsabilidad legal y académica de los contenidos de este Trabajo de Integración Curricular; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 25 de noviembre de 2022



**Gissela Marisol Peñafiel Malan**

**180487585-2**

**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**  
**FACULTAD DE CIENCIAS**  
**CARRERA ESTADÍSTICA**

El Tribunal del Trabajo de Integración Curricular, certifica que: El Trabajo de Integración Curricular; Tipo: Proyecto de Investigación, “**PRONÓSTICOS EN LOS CONSUMOS DE AGUA POTABLE SEGÚN REGISTROS HISTÓRICOS GENERADOS EN LA PARROQUIA CEBADAS DURANTE EL PERIODO 2017-2021**”, realizado por la señorita: **GISSELA MARISOL PEÑAFIEL MALAN**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal autoriza su presentación.

	<b>FIRMA</b>	<b>FECHA</b>
Ing. Natalia Alexandra Pérez Londo MSc. <b>PRESIDENTE DEL TRIBUNAL</b>		2022-11-25
Ing. Nancy Elizabeth Chariguaman Maurisaca MSc. <b>DIRECTORA DEL TRABAJO DE INTEGRACIÓN CURRICULAR</b>		2022-11-25
Ing. Johanna Enith Aguilar Reyes MSc. <b>MIEMBRO DE TRIBUNAL</b>		2022-11-25

## **DEDICATORIA**

Este trabajo de investigación, está dedicado primero a Dios y a mi familia. A Dios, porque ha estado conmigo guiándome cada momento, cuidándome y dándome la sabiduría y el conocimiento cada día, a mis padres por el apoyo incondicional y han velado por mi bienestar y educación, a mi esposo e hijo David por su permanente apoyo y motivación en todo momento, y por ser ellos la inspiración para finalizar este trabajo.

Gissela

## **AGRADECIMIENTO**

A Dios por darme la sabiduría y fuerza para culminar mi profesión académica, también agradecerle a la familia por haber confiado y darme la oportunidad de formarme como profesional, de igual forma agradecer a mi tutora de tesis Ing. Nancy Chariguaman por su guía, comprensión, paciencia y conocimiento me orientó en el desarrollo de esta investigación y a la Ing. Johanna Aguilar que me guió de la mejor manera por varios años mi camino como estudiante y en esta investigación.

Gissela

## ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS.....	x
ÍNDICE DE ILUSTRACIONES.....	xii
ÍNDICE DE ANEXOS.....	xiv
RESUMEN.....	xv
SUMMARY.....	xvi
INTRODUCCIÓN.....	1

### CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL.....	4
1.1. Agua.....	4
1.2. Consumo de agua.....	4
1.3. Consumo de agua en la población del ecuador.....	4
1.4. El derecho humano al agua y al saneamiento.....	4
1.5. Disponibilidad de servicios básicos.....	5
1.5.1. <i>Agua para consumo humano de la parroquia cebadas</i> .....	5
1.5.2. <i>Saneamiento básico</i> .....	5
1.6. Variables cuantitativas.....	5
1.6.1. <i>Variable cuantitativa discreta</i> .....	5
1.6.2. <i>Variable cuantitativa continua</i> .....	5
1.7. Estadística.....	5
1.8. Estadística descriptiva.....	6
1.9. Población.....	6
1.10. Unidad elemental.....	6
1.11. Variable.....	6
1.12. Medidas de centralización.....	6
1.13. Media aritmética.....	7
1.14. Media.....	7
1.15. Mediana.....	8
1.16. Moda.....	8
1.17. Medidas de dispersión.....	8
1.18. Varianzas.....	8
1.19. Desviación típica.....	9
1.20. Coeficiente de variación (CV).....	9



1.21.	Medidas de forma .....	10
1.22.	Coefficiente de asimetría .....	10
1.23.	Coefficiente de curtosis .....	10
1.24.	Histograma .....	11
1.25.	Suavización exponencial (HOLT-WINTERS) .....	12
1.26.	Modelos autorregresivos integrados de promedio móvil (ARIMA).....	12

## CAPÍTULO II

2.	METODOLOGÍA DE LA INVESTIGACIÓN .....	14
2.1.	Variables.....	14
2.1.1.	<i>Variable dependiente</i> .....	14
2.1.2.	<i>Variables independientes</i> .....	14
2.2.	Operacionalización de los objetos .....	14
2.3.	Tipo de investigación .....	15
2.4.	Diseño de la investigación no experimental.....	16
2.4.1.	<i>Localización del estudio</i> .....	16
2.4.2.	<i>Población de estudio</i> .....	16
2.4.3.	<i>Tamaño de la muestra</i> .....	16
2.4.4.	<i>Técnicas de recolección de datos</i> .....	16
2.4.5.	<i>Identificaciones variables</i> .....	17
2.2.1.	<i>Modelo estadístico</i> .....	17
2.5.	Variables en estudio.....	17
2.5.1.	<i>Operacionalización de variables</i> .....	17

## CAPÍTULO III

3.	ANÁLISIS ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS .....	19
3.1.	Aplicación del análisis estadístico descriptivo.....	19
3.1.1.	<i>Genero</i> .....	19
3.1.2.	<i>Consumo de agua potable</i> .....	20
3.1.3.	<i>Precio del costo de agua potable de la Parroquia Cebadas</i> .....	24
3.2.	Modelación del consumo de agua potable .....	28
3.2.1.	<i>Descomposición de la serie de tiempo para consumo de agua en Cebadas</i> .....	28
3.2.2.	<i>Modelo ARIMA</i> .....	29
3.2.2.1.	<i>Prueba de Dickey-Fuller</i> .....	29
3.2.2.2.	<i>Función de autocorrelacion y función de autocorrelacion parcial</i> .....	32

3.2.2.3.	<i>Prueba de validez (ruido blanco) para el modelo ARIMA del consumo de agua</i>	33
3.2.3.	<b>Modelo Holt Winters</b>	35
3.2.3.1.	<i>Validación del modelo Holt-Winters de consumo de agua potable</i>	36
3.2.4.	<i>Medidas de precisión de los modelos de predicción de consumo de agua potable</i>	38
3.2.5.	<i>Pronostico del consumo de agua potable de la Parroquia Cebadas 2017-2021</i>	39
3.3.	<b>Modelación del costo de agua potable</b>	41
3.3.1.	<i>Descomposición de la serie de tiempo para el costo de agua en la Parroquia Cebadas</i>	41
3.3.2.	<b>Modelo ARIMA</b>	41
3.3.2.1.	<i>Prueba de Dickey-Fuller</i>	41
3.3.2.2.	<i>Función de autocorrelación y función de autocorrelación parcial</i>	43
3.3.2.3.	<i>Prueba de Validez (ruido blanco) para el modelo ARIMA del costo de agua potable</i>	45
3.3.3.	<b>Modelo Holt Winters</b>	46
3.3.3.1.	<i>Validación del modelo Holt-Winters sobre el costo de agua potable</i>	47
3.3.4.	<i>Medidas de Precisión de los modelos de predicción de consumo de agua potable</i>	49
3.3.5.	<i>Pronósticos del costo de agua potable en la Parroquia Cebadas</i>	50
3.4.	<b>Comprobación de hipótesis de la investigación</b>	51
3.4.1.	<i>Hipótesis del consumo de agua potable</i>	51
3.4.2.	<i>Hipótesis del costo de agua potable</i>	54
	<b>CONCLUSIONES</b>	58
	<b>RECOMENDACIONES</b>	59
	<b>BIBLIOGRAFIA</b>	
	<b>ANEXOS</b>	

## ÍNDICE DE TABLAS

<b>Tabla 1-2:</b>	Operacionalización de los objetos .....	14
<b>Tabla 2-2:</b>	Operacionalización de las variables .....	17
<b>Tabla 3-2:</b>	Consumo de agua potable en metros cúbicos Parroquia Cebadas .....	18
<b>Tabla 4-2:</b>	Costo del agua potable en dólares de Parroquia Cebadas. ....	18
<b>Tabla 1-3:</b>	Genero de los usuarios .....	19
<b>Tabla 2-3:</b>	Distribuciones de Frecuencias del consumo de agua potable.....	20
<b>Tabla 3-3:</b>	Medidas Descriptivas del consumo de agua potable. ....	21
<b>Tabla 4-3:</b>	Tendencia central del consumo de agua potable .....	22
<b>Tabla 5-3:</b>	Tendencia central del consumo de agua 2018.....	22
<b>Tabla 6-3:</b>	Tendencia central del consumo de agua potable 2019 .....	23
<b>Tabla 7-3:</b>	Tendencia central del consumo de agua potable 2020 .....	23
<b>Tabla 8-3:</b>	Tendencia central del consumo de agua potable 2021 .....	23
<b>Tabla 9-3:</b>	Medidas de dispersión del consumo de agua potable.....	23
<b>Tabla 10-3:</b>	Medida de forma del consumo de agua potable .....	24
<b>Tabla 11-3:</b>	Distribuciones de frecuencia del costo de agua potable. ....	24
<b>Tabla 12-3:</b>	Medidas estadísticas del costo de agua potable.....	25
<b>Tabla 13-3:</b>	Tendencia central del costo agua potable 2017 .....	26
<b>Tabla 14-3:</b>	Tendencia central del costo agua potable 2018.....	26
<b>Tabla 15-3:</b>	Tendencia central del costo agua potable 2019.....	27
<b>Tabla 16-3:</b>	Tendencia central del costo agua potable 2020.....	27
<b>Tabla 17-3:</b>	Tendencia central del costo agua potable 2021 .....	27
<b>Tabla 18-3:</b>	Medidas de variabilidad del costo del agua potable.....	27
<b>Tabla 19-3::</b>	Medidas de forma del costo del agua potable. ....	28
<b>Tabla 20-3:</b>	Confiabilidad de los parámetros del modelo ARIMA.....	33
<b>Tabla 21-3:</b>	Modelo Holt Winters.....	35
<b>Tabla 22-3:</b>	Medida de exactitud ARIMA vs Holt Winters.....	38
<b>Tabla 23-3:</b>	Pronósticos del consumo de agua potable. ....	39
<b>Tabla 24-3:</b>	Confiabilidad de los parámetros del modelo ARIMA.....	45
<b>Tabla 25-3:</b>	Modelo Holt Winters del costo de agua potable. ....	46
<b>Tabla 26-3:</b>	Medida de precisión ARIMA vs Holt Winters.....	49
<b>Tabla 27-3:</b>	Pronósticos del costo de agua potable en la Parroquia Cebadas. ....	50
<b>Tabla 28-3:</b>	Prueba de Kolmogorov-Smirnov para el consumo de agua potable.....	52
<b>Tabla 29-3:</b>	Consumo de agua potable reales y pronosticados. ....	52
<b>Tabla 30-3:</b>	Prueba de Kolmogorov-Smirnov del costo de agua potable. ....	55

<b>Tabla 31-3:</b>	Costo reales y pronosticados .....	55
--------------------	------------------------------------	----

## ÍNDICE DE ILUSTRACIONES

<b>Ilustración 1-1:</b>	Coeficiente de curtosis. ....	11
<b>Ilustración 1-2:</b>	Localización de la Parroquia Cebadas.....	16
<b>Ilustración 1-3:</b>	Genero de los usuarios de la Parroquia Cebadas.....	19
<b>Ilustración 2-3:</b>	Histograma del consumo de agua potable en la Parroquia Cebadas. ....	21
<b>Ilustración 3-3:</b>	Box Plot del consumo de agua potable en la Parroquia Cebadas.....	22
<b>Ilustración 4-3:</b>	Histograma del Costo de agua potable. ....	25
<b>Ilustración 5-3:</b>	Box-plot Costos de agua potable de la Parroquia Cebadas. ....	26
<b>Ilustración 6-3:</b>	Descomposición de la serie de tiempo del consumo de agua potable. ....	28
<b>Ilustración 7-3:</b>	Autocorrelograma simple y parcial del consumo de agua potable.....	29
<b>Ilustración 8-3:</b>	Prueba de Dickey-Fuller para el consumo de agua potable. ....	30
<b>Ilustración 9-3:</b>	Primera diferencia para transforma a estacionalidad.....	31
<b>Ilustración 10-3:</b>	Segunda diferencia para transformar a estacionalidad. ....	31
<b>Ilustración 11-3:</b>	Autocorrelación simple y parcial del consumo de agua potable. ....	32
<b>Ilustración 12-3:</b>	Prueba de validez del modelo.....	34
<b>Ilustración 13-3:</b>	Residuales del modelo de pronóstico del consumo de agua potable ....	35
<b>Ilustración 14-3:</b>	Consumo real y Consumo pronosticada por el método Holt-Winters.....	36
<b>Ilustración 15-3:</b>	Normalidad del modelo Holt-Winters.....	36
<b>Ilustración 16-3:</b>	Grafico de la Prueba de Ljung-Box.....	37
<b>Ilustración 17-3:</b>	Consumo real y Consumo estimado por ARIMA ....	38
<b>Ilustración 18-3:</b>	Consumo real y Consumo Pronosticada del modelo Holt-Winters.....	39
<b>Ilustración 19-3:</b>	Pronostico del consumo de agua potable. ....	40
<b>Ilustración 20-3:</b>	Descomposición de la serie de tiempo del costo de agua potable.....	41
<b>Ilustración 21-3:</b>	Prueba de Dickey-Fuller.....	42
<b>Ilustración 22-3:</b>	Primera diferencia para transformar a estacionalidad. ....	42
<b>Ilustración 23-3:</b>	Segunda diferencia para transformar a estacionalidad. ....	43
<b>Ilustración 24-3:</b>	Autocorrelación simple y autocorrelación parcial costo de agua potable. ...	44
<b>Ilustración 25-3:</b>	Prueba de validez Ljung-Box. ....	45
<b>Ilustración 26-3:</b>	Residuales del modelo de predicción del costo de agua potable.....	46
<b>Ilustración 27-3:</b>	Gasto real y Gasto pronosticado por el método Holt-Winters. ....	47
<b>Ilustración 28-3:</b>	Grafico de normalidad.....	47
<b>Ilustración 29-3:</b>	Prueba de validez de Ljung-Box. ....	48
<b>Ilustración 30-3:</b>	Costo real y costo pronosticado por ARIMA y Holt-Winters.....	49
<b>Ilustración 31-3:</b>	Pronósticos del costo con el método Holt-Winters. ....	51
<b>Ilustración 32-3:</b>	Grafico de región de rechazo ....	54

**Ilustración 33-3:** Grafico de región de rechazo. .... 57

## **ÍNDICE DE ANEXOS**

**ANEXO A:** AVAL DE LA JUNTA ADMINISTRADORA DE AGUA POTABLE

**ANEXO B:** CÓDIGO R-ESTUDIO PARA ANÁLISIS ESTADÍSTICO

**ANEXO C:** ENCUESTA SOBRE LA CALIDAD DEL AGUA POTABLE

**ANEXO D:** BASE DE DATOS TABULADA

**ANEXO E:** EVIDENCIAS FOTOGRÁFICAS

## RESUMEN

El presente trabajo de investigación tuvo por objetivo pronosticar los consumos de agua potable según los registros históricos generados en la Parroquia Cebadas durante el período 2017-2021. Para el estudio se utilizó un análisis estadístico descriptivo y modelos de pronóstico, se recolectó información real donde se utilizaron variables como género, lecturas de consumo en metros cúbicos y el costo de agua, con el fin de conocer el consumo habitual de cada hogar y proporcionar información sobre el mal uso del agua según lo recomendado por la Organización Mundial de la Salud. Como resultados se obtuvo que, en el mes de enero del 2017 hubo un consumo bajo con 536.89 m<sup>3</sup> de agua y en el mes de diciembre del 2021 llegó hasta su pico más alto en consumo con 595.60m<sup>3</sup> de agua potable, lo cual fue analizado con el modelo de pronóstico ARIMA (0,1,0) (1,0,1) [12]. Para evaluar el costo se utilizaron las medidas de precisión y el modelo Holt-Winter, que fueron validados mediante la prueba Ljung Box y se determinó que, en los próximos dos años el consumo y el costo del agua potable incrementará con un nivel de significancia del 0,05. Se concluyó que, el mejor modelo para pronosticar el consumo de agua potable en la parroquia Cebadas, fue ARIMA (0,1,0) (1,0,1) [12] con parámetros  $ma = -0.9593$  y  $sar = 0,9981$  y para pronosticar el costo del agua potable el modelo adecuado de pronóstico fue Holt-Winters, con parámetros  $\alpha = 613.97$  y  $\beta = 1424$ , donde se determinó que, tanto el consumo y el costo van a incrementarse en los próximos años de forma paulatina. Se recomienda a la junta administradora de agua potable de la parroquia Cebadas, realizar más captación de agua ya que aumentará el consumo de este líquido vital.

**Palabras clave:** <ESTADÍSTICA> <CONSUMO DE AGUA POTABLE> <MODELO DE PRONOSTICO ARIMA> <MODELO DE PRONOSTICO HOLT WINTERS> <SANAMIENTO DE AGUA>.

0159-DBRA-UTP-2023





## SUMMARY

The objective of this research work was to forecast the consumption of drinking water according to the historical records generated in the Cebadas parish during the period 2017-2021. For the study, a descriptive statistical analysis and forecast models were used, real information was collected where variables such as gender, consumption readings in cubic meters and the cost of water were employed, in order to know the habitual consumption of each home and provide information on misuse of water as recommended by the World Health Organization. As results it was obtained that, in the month of January 2017 there was a low consumption with 536.89m<sup>3</sup> of water and in the month of December 2021 it reached its highest peak in consumption with 595.60m<sup>3</sup> of drinking water, which was analyzed with the forecast model ARIMA (0,1,0) (1,0,1) [12]. To evaluate the cost, precision measures and the Holt-Winters model were used, which were validated by the Ljung-Box test, it was determined that, in the next two years, the consumption and cost of drinking water will increase with a significance level of 0.05. It was concluded that the best model to predict drinking water consumption in the Cebadas parish was ARIMA (0,1,0) (1,0,1) [12] with parameters  $ma = -0.9593$  and  $sar = 0.9981$  and to forecast the cost of drinking water, the appropriate model was Holt-Winters, with parameters  $\alpha = 613.97$  and  $\beta = 1424$ , where it was determined that both consumption and cost will gradually increase in the coming years. It is recommended to the drinking water management board of the Cebadas parish, to carry out more water collection since it will increase the consumption of this vital liquid.

**Keywords:** <STATISTICS> <DRINKING WATER CONSUMPTION> <ARIMA FORECAST MODEL> <HOLT WINTERS FORECAST MODEL> <WATER SANITATION>.

Edgar Mesías Jaramillo Moyano  
0603497397

## INTRODUCCIÓN

El consumo de agua potable es un tema relevante, ya que la demanda puede considerarse inestable debido a la constante variabilidad. La norma ecuatoriana CPE INEN 00591 recomienda que los valores utilizados en el diseño de sistemas de agua potable se establezcan en base a los estudios realizados previamente en campo, con el fin de obtener una caracterización más realista de las necesidades. En el Ecuador se caracteriza por ser una región privilegiada en relación a sus recursos naturales, uno de ellos, las reservas de agua una de las principales potencialidades, pero el acceso a los servicios de agua potable, pues según ONU-Hábitat (2012) uno de los problemas de la zona es la falta de acceso a estos servicios debido a las dificultades relacionadas con la infraestructura, evidencia y gestión del servicio, además respecto a la desigualdad en el suministro entre zonas urbana y rural. Cada ecuatoriano dispone aproximadamente de 21.000 metros cúbicos al año, según CEPAL (2014) para el consumo e higiene, el acceso al agua potable es del 92%, con una marcada diferencia entre la zona urbana, 97%, y la zona rural 82%, respectivamente.

El sector de agua, saneamiento e higiene juega un rol fundamental en el control de factores de riesgo que afectan la salud pública y que se puede generar cuando las condiciones de los sistemas de abastecimiento de agua, disposición de excretas y aguas residuales cambian por la afectación en la infraestructura que impactan el servicio que se entrega a una población. Por ello, existe la necesidad de fortalecer el sector con medidas transformadoras, tales como: la concientización de los servicios que brinda el sector público, y lograr el compromiso del Estado y la comunidad.

El trabajo conjunto con la Junta administradora de agua potable de la Parroquia de Cebadas, el Estado y comunitarios que brindan servicios de agua potable en toda la parroquia a un precio accesible. Actualmente se realiza la cobranza de agua, para cubrir los gastos de alcantarillado, infraestructuras sanitarias con esto se quiere saber cuánto hay de consumo, para ello se realizará un pronóstico con la finalidad optimizar el consumo de agua. El presente trabajo investigativo se enfocará a la base de datos de los pobladores de la parroquia Cebadas de la misma manera la gran población implica un excesivo gasto en el consumo de agua cuyo presupuesto anual es para cubrir los gastos y esto servirá para el mejoramiento de las condiciones de vida de la población.

Se realizará esta investigación con el propósito de realizar pronósticos para el consumo de agua potable según registros históricos generados en la Parroquia Cebadas, Cantón Guamote, Provincia de Chimborazo, mediante un modelo ARIMA, se evaluará dicho modelo para optimizar la toma de decisión sobre el consumo del agua.

El servicio de agua es intermitente en la mitad los centros urbanos. La presión del agua está muy por debajo de la norma, especialmente en barrios marginales. En el Ecuador, la Constitución Política garantiza el derecho de acceso al agua potable y saneamiento ambiental en el marco de una vida digna. Así, mejora la calidad de la vida de la población requiere de esfuerzos multidimensionales, y de recursos disponibles que permitan desarrollar proyectos integrales de provisión de servicios a nivel urbano y rural (Jara, 2017)

La presente investigación se enfocará a los pobladores de la Parroquia Cebadas de la misma manera la gran población implica un excesivo gasto en el consumo de agua cuyo presupuesto anual es para cubrir los gastos de alcantarillado esto servirá para el mejoramiento de las condiciones de vida de la población local. Para el Ecuador la importancia de dotar de servicios de agua potable y alcantarillado a toda la población se instrumentaliza a través del Plan Nacional de Buen Vivir (PNBV) (Planificación, 2017)

La principal medida que dispone la junta administrativa de agua potable de la Parroquia Cebadas consiste en incrementar su capacidad de captación y potabilización de agua, pero no cuenta con un proyecto a largo plazo para incrementar el caudal y así satisfacer las necesidades de cada usuario. Por esta razón, es necesario crear modelos de pronósticos para pronosticar el consumo de agua que requiere la población.

### **Antecedentes**

La provisión de agua mediante tanqueros fluviales, disposición de excretas directamente en el manglar o cuerpos hídricos y la quema de desechos sólidos, son las principales problemáticas de saneamiento ambiental actuales que se analizarán para dar una respectiva alternativa de solución. La limitada accesibilidad al poblado, restringida únicamente al transporte fluvial, la recolección de crustáceos como única alternativa económica, las constantes inundaciones a las que se enfrentan y la composición salina del suelo que limita la producción agrícola, forman parte de otros aspectos relevantes que serán analizados para caracterizar a la población, su realidad socioeconómica y de calidad de vida, que permitan tomar las decisiones más oportunas para el diseño de los sistemas de agua potable y saneamiento en relación directa con sus necesidades.

### **Planteamiento del problema**

Actualmente en la Parroquia Cebadas se realiza el cobro de los consumos de agua potable en dólares / por mes, para cubrir los gastos de alcantarillado, cunetas, entre otros. Debido a la falta

de un análisis de pronósticos de los consumos de agua potable en la Parroquia Cebadas se presenta esta investigación para pronosticar los consumos de cada hogar en los próximos años.

### **Formulación (Incógnita)**

¿Existe la necesidad de realizar pronósticos de los consumos de agua potable en la Parroquia Cebadas para los próximos años?

### **Objetivos de la investigación**

#### **Objetivo general**

Pronosticar los consumos de agua potable según los registros históricos generados en la Parroquia Cebadas durante el periodo 2017-2021, mediante un modelo ARIMA.

#### **Objetivos específicos**

- Realizar un análisis descriptivo de los consumos de agua potable generados en la Parroquia Cebadas con la ayuda del software R.
- Construir un modelo estadístico para determinar pronósticos de los consumos de agua potable durante el periodo 2017-2022.
- Elegir el mejor modelo de pronósticos de ingresos económicos del consumo del agua potable para los próximos años.

## CAPÍTULO I

### 1. MARCO TEÓRICO REFERENCIAL

#### 1.1. Agua

Se denomina agua potable o agua para el consumo humano, al que puede ser consumida sin restricción. El término se aplica al agua que cumple con las normas de calidad promulgadas por las autoridades locales e internacionales.

#### 1.2. Consumo de agua

En Ecuador se gasta 237 litros por día; sin embargo, el consumo mínimo por habitante es de 100 litros por día, según la OMS; pero el consumo mínimo al día, por habitante, establecido en la ley de Recursos Hídricos es de 200 litros, en el país. La falta de agua afecta a los más pobres, recrudescen el hambre y la desnutrición, y su mala calidad y saneamiento inadecuado repercuten en la seguridad alimentaria, los medios de subsistencia y en la oportunidad de educación para las familias más pobres; y tomando el dato de la OMS que es 237 litros al día de consumo de agua potable por habitante en Ecuador, tendríamos un requerimiento de 63.343.227 litros al día; equivalente a 63.343 m<sup>3</sup> de demanda de agua potable diario (Merchan, 2018).

#### 1.3. Consumo de agua en la población del Ecuador

Según INEC 2017, dato a diciembre 2016, tiene acceso al agua potable el 94% de la población urbana y solo el 57,50% de la población rural, dando un resultado nacional de 82,32% de acceso al líquido vital potabilizado; estaríamos por debajo de la media. Necesitamos aumentar el uso eficiente de recursos hídricos en todos los sectores y asegurar la sostenibilidad del abastecimiento de agua potable, implementar la gestión integrada de los recursos hídricos en todos los niveles, proteger y restablecer los ecosistemas relacionados con el agua, incluidos los cerros, los ríos y los acuíferos. Exigir actividades y programas para toda la comunidad, relativos al agua y saneamiento, como captación de agua, desalinización, tratamiento y reutilización (Merchan, 2018).

#### 1.4. El derecho humano al agua y al saneamiento

El 2011 fue un año importante para el mundo en materia de agua potable y saneamiento, es por ello que el derecho al agua ha sido considerado de manera tradicional como un derecho natural,

derivado de la naturaleza humana, de las condiciones históricas, de las necesidades básicas o de la propia noción de justicia.

## **1.5. Disponibilidad de servicios básicos**

### ***1.5.1. Agua para consumo humano de la parroquia cebadas***

Los pobladores de las 11 comunidades del sector 2 de la parroquia Cebadas dispone del servicio de agua potable en todos los domicilios.

### ***1.5.2. Saneamiento básico***

La red de alcantarillado se considera como un servicio básico, sin embargo, la cobertura de esta asistencia aun es limitada en las comunidades de la parroquia Cebadas. En cuanto a letrización, todas las comunidades disponen de este servicio (Chimborazo, 2017).

## **1.6. Variables cuantitativas**

Los valores que toman la variable provienen de un conteo o de una medición (instrumento de medida: metro, balanza, cronometro, etc.); se divide a su vez en:

### ***1.6.1. Variable cuantitativa discreta***

Una variable se considera cuantitativa discreta cuando los valores que toman la variable provienen de un conteo y están asociados a los números enteros.

### ***1.6.2. Variable cuantitativa continua***

Una variable se considera continua, cuando los valores que toman la variable provienen de una medición; es decir, los valores que toma se encuentran en un intervalo. También se dice que los valores que toma esta variable están asociados a los números reales (Cesar, 2020).

## **1.7. Estadística**

Es una ciencia que proporciona métodos y procedimientos para el análisis de datos, cuando estos por su naturaleza presenta incertidumbre en la toma de decisión.

Es una ciencia que se encarga de la recolección, ordenamiento, representación, análisis e interpretación de datos generados en una investigación sobre hechos, individuos o grupos de los mismos, para deducir de ello conclusiones precisas o estimaciones futuras (Cesar, 2020).

### **1.8. Estadística descriptiva**

Es parte de la estadística que consiste en la recolección, organización, presentación, análisis e interpretación de un conjunto de datos para una o más variables de interés del investigador (Cesar, 2020).

### **1.9. Población**

Es la totalidad de unidades elementales (personas, animales u objetos) sujeto a estudio y que poseen una característica común; al tamaño de la población (número de unidades elementales) se simboliza con la letra mayúscula “N”. Una población en estudio puede ser finita o infinita (Cesar, 2020).

### **1.10. Unidad elemental**

Es cada uno de los elementos que conforman la población en estudios, de la cual se requiere información. En una unidad elemental se define muchas variables de interés para el investigador (Cesar, 2020).

### **1.11. Variable**

Es una característica de interés que se desea estudiar en la unidad elemental. La variable en estudios la fija el investigador; el valor que toma la variable cambia de unidad elemental a otra. A una variable se le simboliza con las últimas letras del abecedario, tales como: X, Y, Z o también letras mayúsculas acompañadas con subíndice, tales como:  $X_1, X_2, \dots, X_k$ . (Cesar, 2020).

### **1.12. Medidas de centralización**

Son valores estadísticos cálculos con los datos de la muestra o de la población y que tienden a ubicarse en el centro de la distribución de los datos. A estas medidas se les considera como valores representativos para el conjunto de datos.

Si estos son calculados usando los datos de la muestra, se les llama valores estadísticos, estadígrafos o estimados; en cambio sí son obtenidos usando los datos de la población, se les llama parámetros (Cesar, 2020).

### 1.13. Media aritmética

Es un valor que tiende a ubicarse en el centro de la distribución de los datos, y es considerado un valor representativo para el conjunto de datos cuando los datos presenta poca variabilidad. Su valor se calcula tanto para la muestra como para la población. Cuando la media aritmética se calcula utilizando datos que no se encuentran en una tabla de distribución de frecuencias, se dice que la media aritmética se obtuvo para datos sin agrupar, en caso contrario la media aritmética se obtuvo para datos agrupados. (Cesar, 2020)

La media aritmética para la población:

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

La media aritmética para la muestra:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Donde:

$\sum$  : indica la sumatoria de los valores desde el primero hasta el n

$N$ :representa el número total de observaciones

$x$ : esta variable es sobre la que calculamos la media aritmética

$i$ : representa la posición de cada observación

$n$ : tamaño de la muestra

### 1.14. Media

Es el valor medio, es simplemente el promedio de los numero, además es un valor esperado de una variable aleatoria, simbolizada por  $\bar{x}$ , según la siguiente función de la muestra simple (Rustom J, 2012):

$$\bar{x} = \frac{\sum_{i=1}^n X_i}{n}$$

Donde:

$\sum$  : indica la sumatoria de los valores desde el primero hasta el n

$x$ : esta variable es sobre la que calculamos la media aritmética

$i$ : representa la posición de cada observación

$n$ : tamaño de la muestra



### 1.15. Mediana

Es una medida de tendencia central, que divide al conjunto de datos ordenados de la muestra, en dos partes iguales; es decir el 50% de los datos tendrá valores menores o iguales al valor de la mediana y el otro 50% de los datos con valores superiores al valor de la mediana (Cesar, 2020).

$$n = \text{par} \rightarrow me = \frac{X_n + X_{(n+1)}}{2}$$

$$n = \text{impar} \rightarrow me = x_{n+1/2}$$

Donde:

$X_i$ : valor que toma la variable X en la posición "i"

### 1.16. Moda

Sea X una variable en estudio que toma los valores  $x_1, x_2, \dots, x_n$ , la moda, es aquel valor de la variable que se representa con mayor frecuencia, se representa por ( $m_o$ ); es decir,

$$m_o = \text{valor que toma la variable para la max frecuencia } (f)_i$$

El valor de la moda se interpreta como el valor más frecuente para el conjunto de datos de una muestra o población.

El valor de la moda puede o no existir. Si un conjunto de datos tiene una sola moda a su distribución se la llama unimodal, si tiene dos modas se la llama bimodal y para más modas multimodal (Cesar, 2020).

### 1.17. Medidas de dispersión

Las medidas de dispersión vistas anteriormente reducen la información recogida de la misma a un solo valor.

### 1.18. Varianzas

Es una medida de variabilidad de los datos de la muestra, en forma similar a la varianza poblacional y calcula la tendencia de las observaciones propios a extraviar con relación a la media, por ser un estadígrafo corresponde a una función de la muestra aleatoria, según la siguiente definición (Rustom J, 2012):

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}, \text{ donde } (n - 1)g.l$$

Donde:

$s^2$ : varianza

$x_i$  : termino del conjunto de datos

$\bar{x}$ : media de la muestra

$\Sigma$  : indica la sumatoria de los valores desde el primero hasta el n

$n$ : tamaño de la muestra

### 1.19. Desviación típica

Se utiliza para comparar la variabilidad o el grado de dispersión de dos o más conjunto de datos que poseen la misma la misma unidad de medida y tienen sus promedios similares. Su valor, se define como la raíz cuadrada positiva de la varianza; es decir,

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Donde:

$s$ : desviación estándar

$x_i$  : termino del conjunto de datos

$\bar{x}$ : media de la muestra

$\Sigma$  : indica la sumatoria de los valores desde el primero hasta el n

$n$ : tamaño de la muestra

El valor de la desviación estándar para un conjunto de datos esta expresado en las mismas unidades de la variable en estudio. Su valor es mayor o igual a cero, un valor cero indica que los datos de la muestra están concentrados en un mismo punto (no hay variabilidad o dispersión), un valor mayor a cero indica la presencia de variabilidad o dispersión de los datos con respecto a la media aritmética (Cesar, 2020).

### 1.20. Coeficiente de variación (CV)

Es una medida de variación relativa no tiene unidad de medida se calcula haciendo uso de la siguiente formula:

$$CV = \frac{S}{\bar{x}} \times 100\%$$

Donde:

$CV$ : coeficiente de variación

$s$ : desviación estándar

$\bar{x}$ : media de la muestra

El CV, indica el número de veces que  $s$  (desviación estándar) y contiene la media aritmética, y cuando contiene más medias, indica mayor variabilidad en los datos.

Su valor se usa para comparar la variabilidad o dispersión entre dos o más conjuntos de datos que poseen diferentes unidades de medidas y medias aritméticas diferentes (Cesar, 2020).

### 1.21. Medidas de forma

Las medidas de las formas de una distribución se clasifican en dos: medidas de asimetría y medidas de curtosis.

### 1.22. Coeficiente de asimetría

Las medidas de asimetría indica la deformación horizontal de la curva. Cuando una distribución de frecuencias esta inclinada o alargada hacia la derecha, se llama asimetría a la derecha o positiva, y si esta alargada o inclinada hacia el lado izquierdo, se llama asimetría a la izquierda o negativa.

- Simetría: El coeficiente de asimetría  $g_1 = 0$
- Asimetría: El coeficiente de asimetría  $g_1 \neq 0$
- Asimetría positiva: Es el coeficiente de asimetría  $g_1 > 0$
- Asimetría negativa: Es el coeficiente de asimetría  $g_1 < 0$  (Amparo Olmedo, 2015)

### 1.23. Coeficiente de curtosis

La curtosis es la deformación vertical de una distribución de frecuencias. Este coeficiente determina el grado de concentración de los datos con respecto a la media aritmética; se calcula cuando la distribución de frecuencias tiene la forma simétrica. El coeficiente de curtosis se calcula por la siguiente formula (Cesar, 2020):

$$k = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{ns^4} - 3$$

Donde:

$k$ : coeficiente de curtosis

$x_i$ : valor  $i$ -esimo de las observaciones

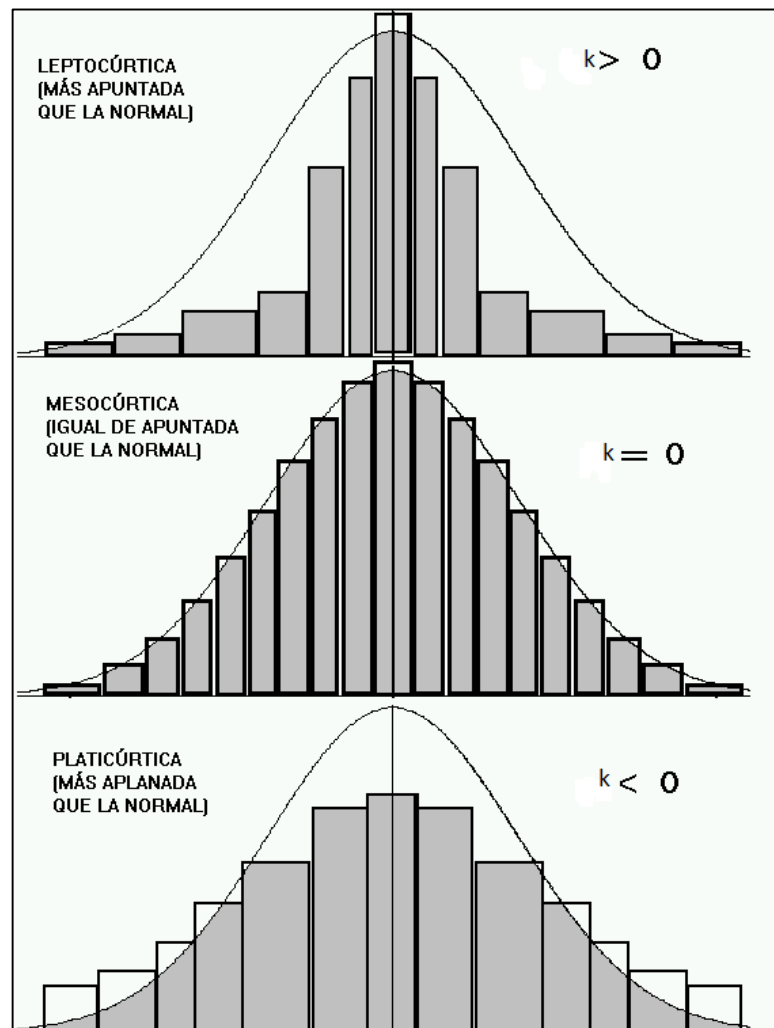
$\bar{x}$ : media aritmética de la distribución

$\sum$ : indica la sumatoria de los valores desde el primero hasta el  $n$

$n$ : tamaño de la muestra

$s$ : desviación estándar de la distribución

- Distribución leptocúrtica: cuando la parte superior de la distribución es alta, es decir, si  $k > 0$ .
- Distribución pico a pico: cuando la distribución muestra un pico moderado, es decir, su distribución no es ni alta ni baja, por ejemplo,  $k = 0$ .
- Distribución platicúrtica: cuando la distribución muestra un pico relativamente plano en su distribución, es decir, si:  $k < 0$ .



**Ilustración 1-1:** Coeficiente de curtosis.  
Fuente: curtosis (uv.es)

#### 1.24. Histograma

Sirven para representar frecuencias absolutas o relativas, cuando existen subdivisiones dentro de cada categoría.

### 1.25. Suavización exponencial (HOLT-WINTERS)

Esa una ecuación de tendencia lineal basada en el empleo de dos constantes de suavización: una para estimar el nivel actual de los valores de la serie de tiempo y otra para estimar la pendiente.

- Estimado de serie de tiempo:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1}T_{t-1})$$

- Estimado de tendencia

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta).T_{t-1}$$

- Estimado de estacionalidad

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma).S_{t-s}$$

- Pronóstico del periodo p

$$\hat{Y}_{t+p} = (L_p - pT_t).S_{t-s+p}$$

Donde:

$L_{t-1}$ : valor suavizado nuevo

$\alpha$ : constante suavizado para el nivel ( $0 < \alpha < 1$ )

$Y_t$ : Valor real de la serie en el periodo t.

$\beta$ : constante del estimado de la tendencia ( $0 < \beta < 1$ )

$T_t$ : Valor real de la serie en el periodo t.

$\gamma$ : contante de la estacionalidad ( $0 < \gamma < 1$ )

$S_t$ : estimado de la tendencia.

p: periodo para pronosticarse al futuro.

s: amplitud de la estacionalidad

$\hat{Y}_{t+p}$ : pronósticos para el periodo p

### 1.26. Modelos autorregresivos integrados de promedio móvil (ARIMA)

Son una categoría de método de pronóstico en los que valores previamente observados en la serie de tiempo se usan como variables independientes en modelo de regresión. El método de más amplio uso de esta categoría es el método de Box- Jenkins. Estos métodos hacen uso explícito de la existencia de autocorrelación (correlación de una variable rezagada uno o más periodos, consigo mismo) en las series de tiempo (Sweeney, 2018, p.1).

El modelo ARIMA (p, d, q) se puede representar como:

$$Y_t = -(\Delta^d Y_t - Y_t) + \phi_0 + \sum_{i=1}^p \phi_i \Delta^d Y_{t-i} - \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t$$

Donde:

$d$ : Corresponde a las diferencias que son necesarias para convertir la serie original en estacionaria.

$\phi_1, \dots, \phi_p$ : Son los parámetros pertenecientes a la parte autorregresivos del modelo.

$\theta_1, \dots, \theta_p$ : Los parámetros pertenecientes a la parte medias móviles del modelo.

$\phi_0$ : Es una constante.

$\varepsilon_t$ : es el término de error

## CAPÍTULO II

### 2. METODOLOGÍA DE LA INVESTIGACIÓN

#### 2.1. Variables

##### 2.1.1. Variable dependiente

Costo del agua potable. - Valor monetario de los consumos de factores que supone el ejercicio de una actividad económica destinada a la producción de un bien, servicio o actividad.

##### 2.1.2. Variables independientes

Lecturas en metros cúbicos del consumo de agua potable de cada mes. - Las lecturas del medidor de agua es un dispositivo o artefacto perteneciente a una conexión que registra con precisión la cantidad de agua que pasa a través de él.

Género. - Características fenotípicas del individuo.

#### 2.2. Operacionalización de los objetos

Tabla 1-2: Operacionalización de los objetos

OBJETIVO GENERAL	OBJETIVOS	CONCEPTO	INDICADOR	INSTRUMENTOS
Pronosticar los consumos de agua potable según los registros históricos generados en la Parroquia Cebadas durante el periodo 2017-2021, mediante un modelo ARIMA.	<ul style="list-style-type: none"><li>Realizar un análisis descriptivo de los consumos de agua potable por la Junta administrativa en la Parroquia Cebadas con la ayuda del software R.</li><li>Construir un modelo estadístico para determinar pronósticos de los consumos de agua potable durante el periodo 2017-2022.</li><li>Elegir el mejor modelo de pronósticos de</li></ul>	<b>Lecturas en metros cúbicos del consumo de agua potable de cada mes:</b> Las lecturas del medidor de agua es un dispositivo o artefacto perteneciente a una conexión que registra con	Metros cúbicos Tarifa por litro de agua potable. Sexo	Matriz de base de datos.

---

ingresos económicos del consumo del agua potable para los próximos años. precisión la cantidad de agua que pasa a través de él.

**Costo del agua potable:**

Valor monetario de los consumos de factores que supone el ejercicio de una actividad económica destinada a la producción de un bien, servicio o actividad.

**Género:**

Características fenotípicas del individuo

---

**Realizado por:** Peñafiel, Gissela, 2022.

### **2.3. Tipo de investigación**

En la presente investigación se utilizará un método de investigación cuantitativo, ya que las variables analizadas son: Costo del agua potable, lecturas en metros cúbicos del consumo de agua potable de cada mes, según el objetivo es aplicada ya que se centra en análisis de las proyecciones obtenidas en el campo del consumo agua potable en la Parroquia Cebadas, según el nivel de profundización en el objeto de estudio es exploratoria ya que nos permitirá conocer el consumo habitual de cada hogar y proporcionar información sobre el mal uso del agua con respecto a lo recomendado a la OMS, según la manipulación de variables no experimental porque la información obtenida es de una fuente secundaria, según el tipo es inductiva ya que se busca ajustar un modelo para obtener pronósticos y conocer cuál es el mejor modelo y comprobar si los modelos son o no aplicables para el consumo de agua potable, según el periodo temporal es transversal que no se sigue al objeto de estudio durante el lapso de estudio. (Hernández Sampieri et al. 2014).

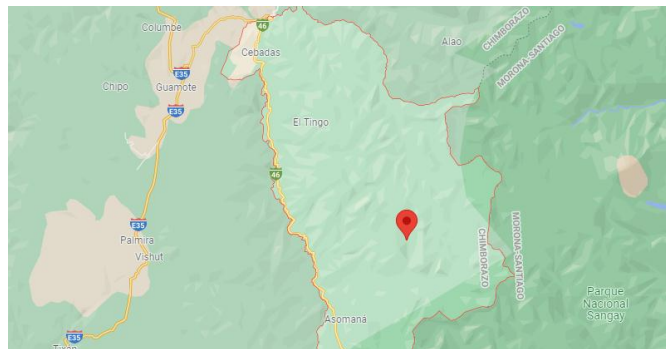


## **2.4. Diseño de la investigación no experimental**

Si utiliza un método de investigación cuantitativo y según la manipulación de variables es un diseño no experimental (Berger et al. 2018).

### **2.4.1. Localización del estudio**

El trabajo se llevó a cabo en la Junta administradora de agua potable y saneamiento de la Parroquia de Cebadas, perteneciente a la Provincia de Chimborazo.



**Ilustración 1-2:** Localización de la Parroquia Cebadas

Fuente: <https://www.google.com.ec/maps>

### **2.4.2. Población de estudio**

La población en estudio está conformada por 391 usuarios de la Parroquia Cebadas, Cantón Guamote, Provincia de Chimborazo que consume el agua potable.

### **2.4.3. Tamaño de la muestra**

No se tomará muestra ya que el estudio se lo realizará con todos los estudiantes para tener un auge más exacto.

### **2.4.4. Técnicas de recolección de datos**

En el estudio se considera una encuesta poblacional de Cebadas, además se cuenta con la base autorizada que cuenta con todos los registros de las lecturas de medidores de todos los usuarios que viven en la parroquia cebadas.

#### 2.4.5. Identificaciones variables

Generó

Costo del agua potable en dólares (\$)

Lecturas en metros cúbicos del consumo de agua potable de cada mes (m<sup>3</sup>)

#### 2.2.1. Modelo estadístico

Los modelos estadísticos a ejecutarse serán Modelo ARIMA, dada el tipo de información obtenida. Además, se usará para pronósticos del consumo de agua potable, se probará supuestos de normalidad (Jarque Bera) y independencia (Ljung-Box).

### 2.5. Variables en estudio

#### 2.5.1. Operacionalización de variables

**Tabla 2-2:** Operacionalización de las variables

Variable	Descripción	Tipo	Escala de medición	Categoría o intervalo
Genero	Sexo	Cualitativa	Nominal	
Costo del agua potable	Dólares	Cuantitativa	De razón	$[0, +\infty[$
Lecturas en metros cúbicos del consumo de agua potable de cada mes	Cantidad	Cuantitativa	De razón	$[0, +\infty[$

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 3-2:** Consumo de agua potable en metros cúbicos Parroquia Cebadas.

ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
520.965	521.965	522.965	531.965	531.965	541.965	542.965	543.195	544.195	545.195	547.195	548.195
550.764	551.747	552.201	552.204	553.385	554.395	555.398	556.399	557.4	558.402	558.403	559.404
560.686	561.687	562.793	562.81	563.812	563.826	564.002	565.994	566.104	566.736	567.289	568.566
569.104	570.126	571.522	572.705	573.148	574.813	575.398	576.422	577.709	578.729	578.093	579.881
582.104	583.126	585.522	587.705	589.848	595.913	596.398	599.422	600.678	601.429	612.093	613.011

**Realizado por:** Peñafiel, Gissela, 2022.

**Tabla 4-2:** Costo del agua potable en dólares de Parroquia Cebadas.

ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
412.32	414.32	415.32	417.32	419.32	422.32	423.25	425.48	427.48	429.48	430.48	432.17
433.20	435.72	436.63	437.10	438.28	439.23	440.00	440.00	441.23	442.23	443.23	444.72
445.02	446.90	447.84	448.78	449.78	450.79	451.01	451.18	452.19	452.94	453.07	453.17
453.40	454.41	454.32	456.17	457.55	458.47	459.76	459.30	459.13	459.95	459.00	460.37
461.37	462.70	464.37	468.90	471.61	473.62	477.34	480.46	481.46	482.19	483.35	485.84

**Realizado por:** Peñafiel, Gissela, 2022.

## CAPÍTULO III

### 3. ANÁLISIS ESTADÍSTICO E INTERPRETACIÓN DE RESULTADOS

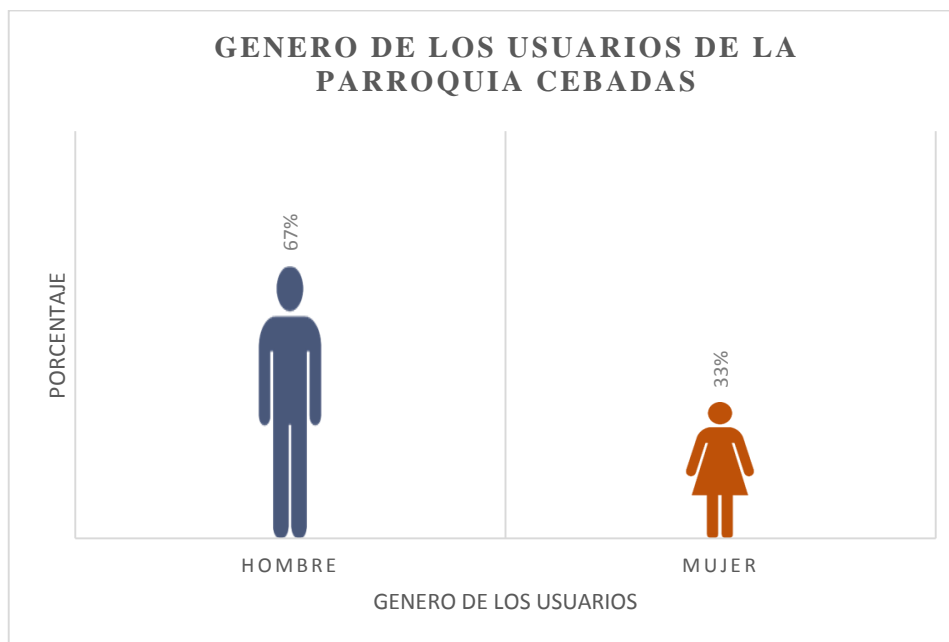
#### 3.1. Aplicación del análisis estadístico descriptivo

##### 3.1.1. Genero

**Tabla 1-3:** Genero de los usuarios

Genero		
Categorías	Frecuencia	Porcentaje
Hombre	261	67%
Mujer	130	33%
Total	391	100%

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 1-3:** Genero de los usuarios de la Parroquia Cebadas.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

Se observa que el 67% de los usuarios obtenidos de la junta administrativa que ocupan el servicio de agua potable son de sexo masculino, mientras que el 33% son de sexo femenino, con un total de 391 usuarios.

**3.1.2. Consumo de agua potable**

La distribución de frecuencia  $X_i$

$$x_1 = \{(x_{i_{min}}, x_{i_{max}}), h_i\}_{i=1}^k$$

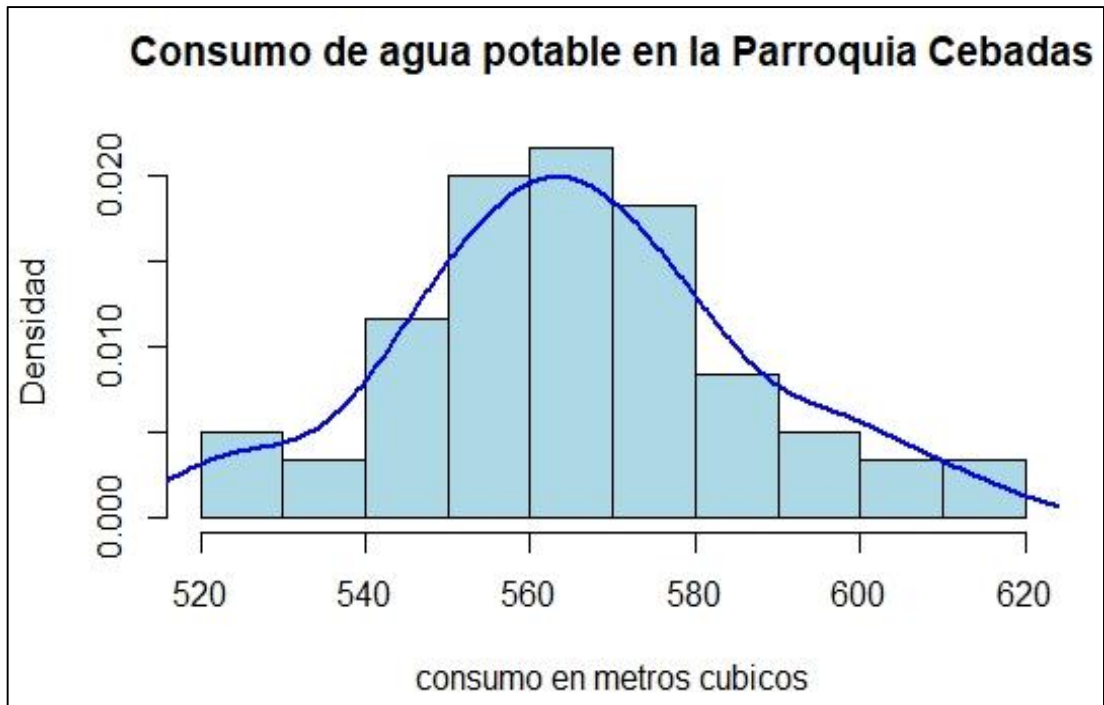
**Tabla 2-3:** Distribuciones de Frecuencias del consumo de agua potable.

i	LímInf	LímSup	Marcas de clase		Frecuencia		Frecuencia	
			$(x_1)$	$h_i$	$H_i$	absoluta	absoluta	acumulada
						$(f_i)$	$(F_i)$	
1	520	540	530	5	5	0.083	0.083	
2	540	560	550	19	24	0.32	0.40	
3	560	580	570	24	48	0.4	0.80	
4	580	600	590	8	56	0.13	0.93	
5	600	620	610	4	60	0.07	1	
				60		1		

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

El 40% de la suma de los consumos totales de los cinco años están en un intervalo de 560 a 580 metros cúbicos de agua potable.



**Ilustración 2-3:** Histograma del consumo de agua potable en la Parroquia Cebadas.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

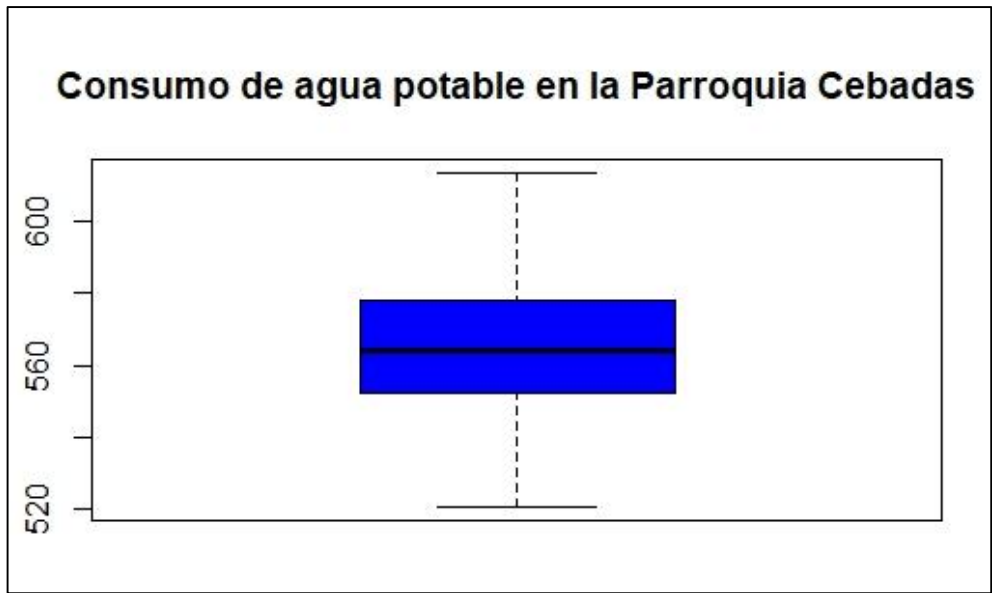
Se puede visualizar en el siguiente histograma que los datos muestran una distribución aproximadamente normal.

**Medidas descriptivas**

**Tabla 3-3:** Medidas Descriptivas del consumo de agua potable.

Mínimo	1er Cuartil	Mediana	Media	3er Cuartil	Máximo
520.965	552.203	563.914	565.367	577.805	613.001

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 3-3:** Box Plot del consumo de agua potable en la Parroquia Cebadas

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

En el gráfico de Q-Q plot se observa que, A falta de datos atípicos, el consumo mínimo es de 521 m<sup>3</sup> y el máximo de 613 m<sup>3</sup>, que en la medida de tendencia central se puede considerar como simétrico en los datos.

**Medidas de tendencia central**

**Tabla 4-3:** Tendencia central del consumo de agua potable

AÑO 2017	
Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Media	536,89
Mediana	54247
Moda	531,97

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 5-3:** Tendencia central del consumo de agua 2018

AÑO 2018	
Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Media	555.01
Mediana	554.90
Moda	#N/A

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 6-3:** Tendencia central del consumo de agua potable 2019

AÑO 2019	
Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Media	564.53
Mediana	563.91
Moda	#N/A

**Realizado por:** Peñafiel, Gissela, 2022.

**Tabla 7-3:** Tendencia central del consumo de agua potable 2020

AÑO 2020	
Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Media	574.80
Mediana	575.11
Moda	#N/A

**Realizado por:** Peñafiel, Gissela, 2022.

**Tabla 8-3:** Tendencia central del consumo de agua potable 2021

AÑO 2021	
Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Media	595.60
Mediana	596.16
Moda	#N/A

**Realizado por:** Peñafiel, Gissela, 2022.

### Interpretación:

En promedio en el año 2017 la suma de los consumos de agua e la Parroquia Cebadas era inferior al promedio del año 2020.

### Medidas de dispersión

**Tabla 9-3:** Medidas de dispersión del consumo de agua potable

Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Varianzas	435.998
desviación	20.881
Coef. de variación	0.0369

**Realizado por:** Peñafiel, Gissela, 2022.



### Interpretación:

Como se puede observar la varianza y desviación estándar, se puede concluir que el consumo de agua potable no se dispersa de la media y por lo tanto se reduce su variabilidad.

### Medida de forma

**Tabla 10-3:** Medida de forma del consumo de agua potable

Medidas Estadísticas	Consumo de agua potable m <sup>3</sup>
Asimetría	0.103984
Curtosis	0.01394

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

Se puede observar que la asimetría de la suma del consumo de agua potable de cada mes y año muestra una distribución asimétrica positiva, es decir, que tiene una concentración más alta a la izquierda de la media y el coeficiente de curtosis indica que tiene una concentración superior a la media para el consumo.

### 3.1.3. Precio del costo de agua potable de la Parroquia Cebadas

La distribución de Frecuencia  $X_i$

$$x_1 = \{(x_{i_{min}}, x_{i_{max}}), n_i\}_{i=1}^k$$

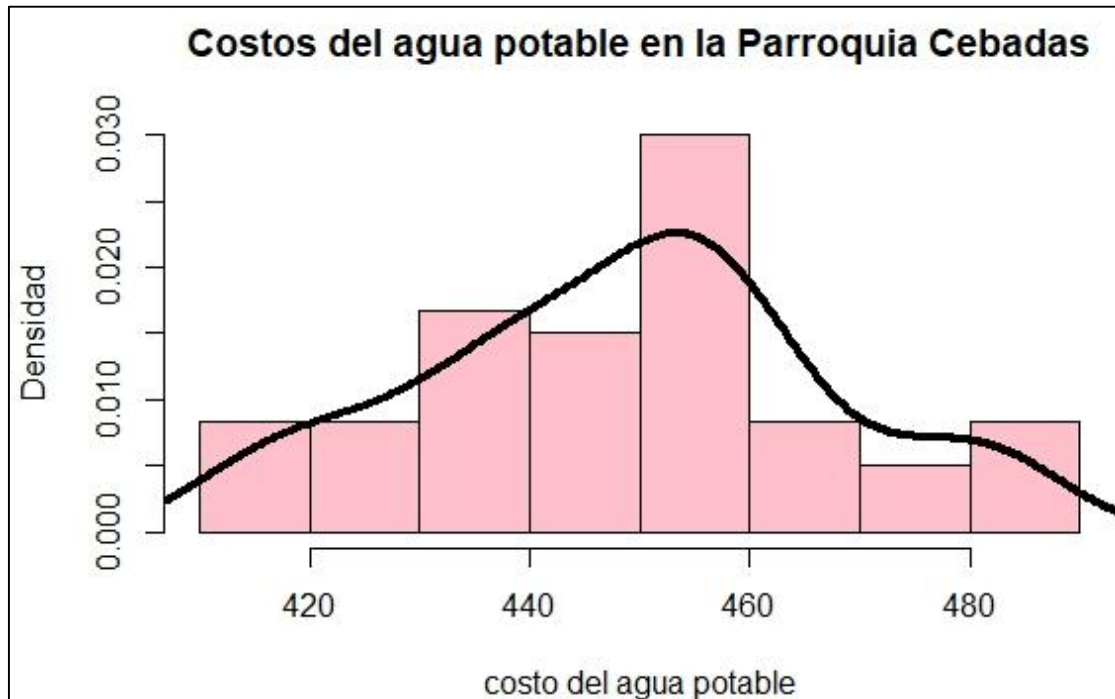
**Tabla 11-3:** Distribuciones de frecuencia del costo de agua potable.

i	LímInf	LímSup	Marcas de clase		$H_i$	Frecuencia	
			$(x_i)$	$h_i$		Frecuencia absoluta $(f_i)$	absoluta acumulada $(F_i)$
1	400	420	410	5	5	0.083	0.083
2	420	440	430	13	18	0.217	0.30
3	440	460	450	29	47	0.483	0.78
4	460	480	470	8	55	0.133	0.92
5	480	500	490	5	60	0.083	1
				60		1	

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

El 48% de la suma de los costos totales de los cinco años están en el intervalo de 440-460 dólares



**Ilustración 4-3:** Histograma del Costo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

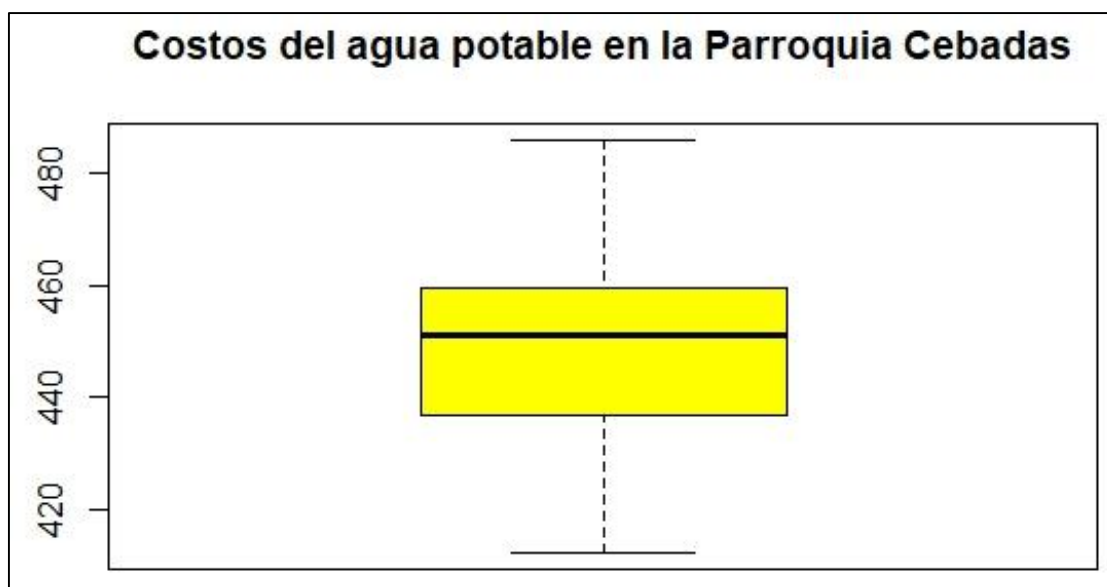
Se puede ver que el costo del agua potable más superior es de 450 a 460 , y los datos siguen una distribución normal.

**Medidas estadísticas de posición relativa**

**Tabla 12-3:** Medidas estadísticas del costo de agua potable.

Mínimo	1er Cuartil	Mediana	Media	3er Cuartil	Máximo	
412.32		436.98	450.901	446.8088	459.411	485.841

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 5-3:** Box-plot Costos de agua potable de la Parroquia Cebadas.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

La gráfica QQ no muestra datos atípicos, el costo mínimo es de 412 m3 y el costo máximo es de 486 m3, lo cual se representa en la medida de tendencia central y se puede decir que los datos muestran simetría positiva.

**Medidas de tendencia central**

**Tabla 13-3:** Tendencia central del costo agua potable 2017

AÑO 2017	
Medidas Estadísticas	Costos del agua potable (dólares)
Media	422.44
Mediana	422.79
Moda	#N/A

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 14-3:** Tendencia central del costo agua potable 2018

AÑO 2018	
Medidas Estadísticas	Costos del agua potable (dólares)
Media	439.30
Mediana	439.62
Moda	440

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 15-3:** Tendencia central del costo agua potable 2019

AÑO 2019	
Medidas Estadísticas	Costos del agua potable (dólares)
Media	450.22
Mediana	450.90
Moda	#N/A

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 16-3:** Tendencia central del costo agua potable 2020

AÑO 2020	
Medidas Estadísticas	Costos del agua potable (dólares)
Media	457.65
Mediana	458.74
Moda	#N/A

Realizado por: Peñafiel, Gissela, 2022.

**Tabla 17-3:** Tendencia central del costo agua potable 2021

AÑO 2021	
Medidas Estadísticas	Costos del agua potable (dólares)
Media	474.43
Mediana	475.48
Moda	#N/A

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

En promedio en el año 2017 la suma de los consumos de agua e la Parroquia Cebadas era inferior al promedio del año 2020.

### Medidas de variabilidad

**Tabla 18-3:** Medidas de variabilidad del costo del agua potable.

Medidas Estadísticas	Costos del agua potable (dólares)
Varianza	336.39
Desviación	18.34
Coef. de variación	0.04

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

Se puede visualizar que la varianza y la desviación no están dispersos de la media y que por tanto su variabilidad es mínima.

### Medidas de Forma

**Tabla 19-3::** Medidas de forma del costo del agua potable.

Medidas Estadísticas	Costos del agua potable (dólares)
Asimetría	-0.001967
Curtosis	-0.4247

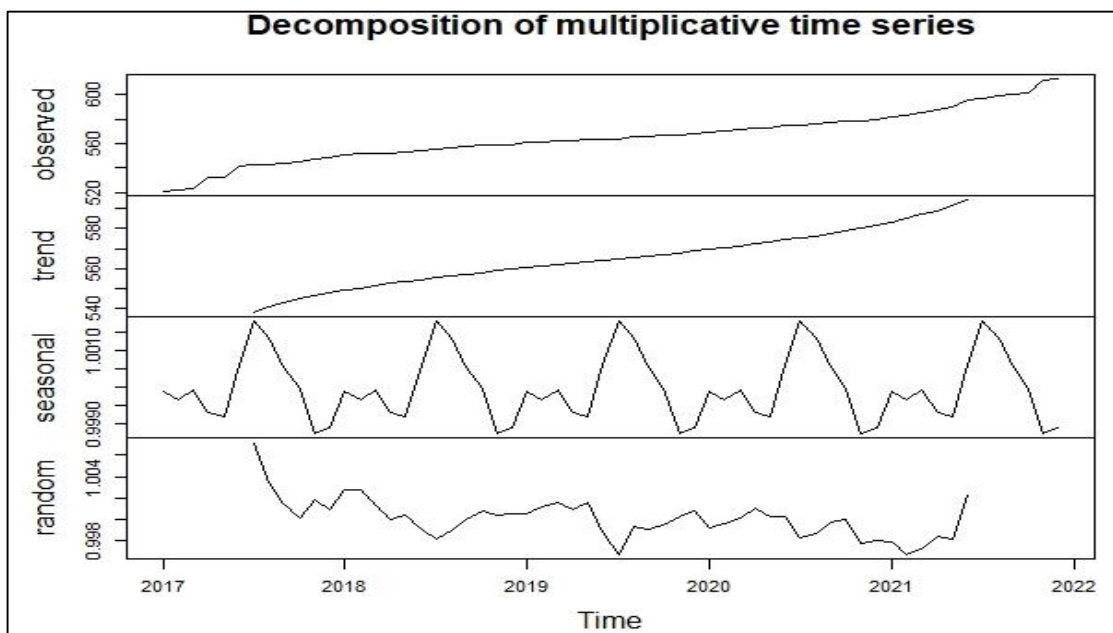
Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

Podemos visualizar que el coeficiente de asimetría de la suma del costo de agua potable por mes y año presenta una distribución asimetría positiva, es decir existe mayor concentración de los costos de agua potable al lado izquierdo de Si el valor medio, el coeficiente de curtosis indica que es platocúrtica, lo que significa que está robustamente concentrado en la media.

## 3.2. Modelación del consumo de agua potable

### 3.2.1. Descomposición de la serie de tiempo para consumo de agua en Cebadas

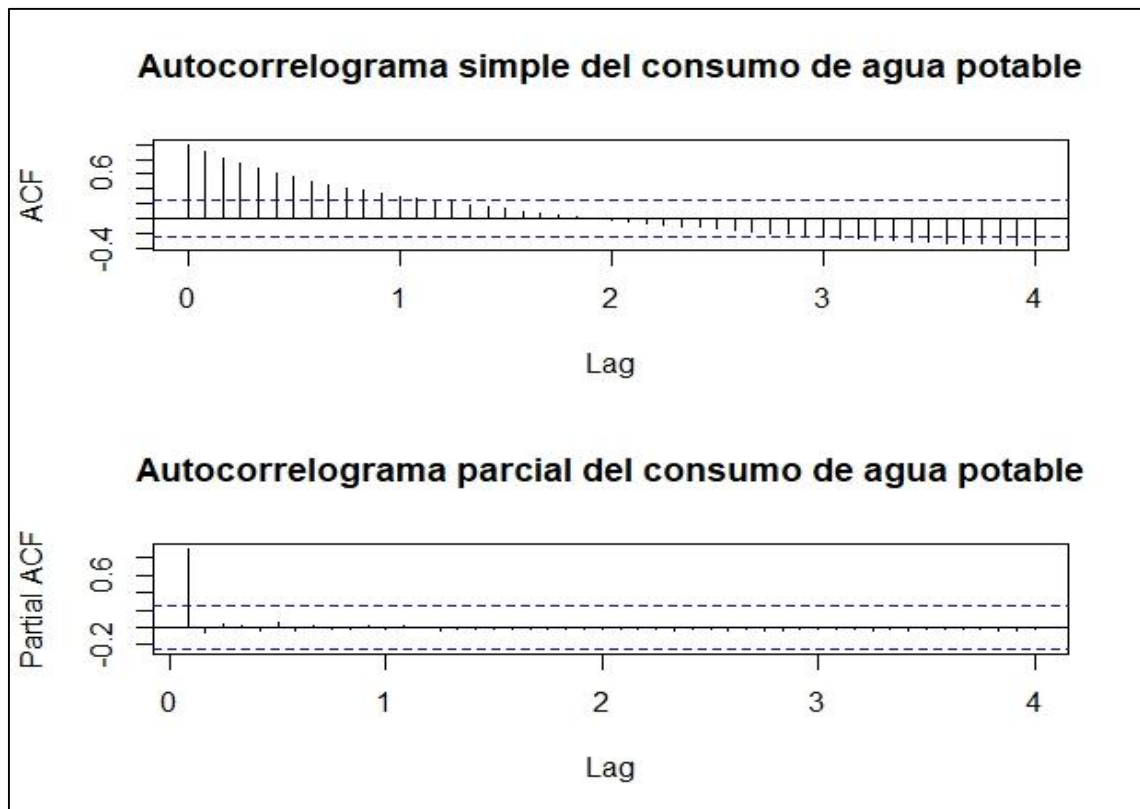


**Ilustración 6-3:** Descomposición de la serie de tiempo del consumo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

Se visualiza en este grafico que es un modelo multiplicativo ya que la varianza no es constante a lo largo del tiempo, también se observa que hay tendencia que representa el incremento del consumo de agua potable en la Parroquia Cebadas desde enero del 2017 al mes de junio del 2021.



**Ilustración 7-3:** Autocorrelograma simple y parcial del consumo de agua potable

**Realizado por:** Peñafiel, Gissela, 2022.

### Interpretación:

Se visualiza que no hay estacionalidad debemos aplicar una prueba para observar si hay estacionalidad y si no hay debemos hacerle estacionalidad para poder aplicar el modelo ARIMA la prueba más recomendada es Prueba de Dickey-Fuller.

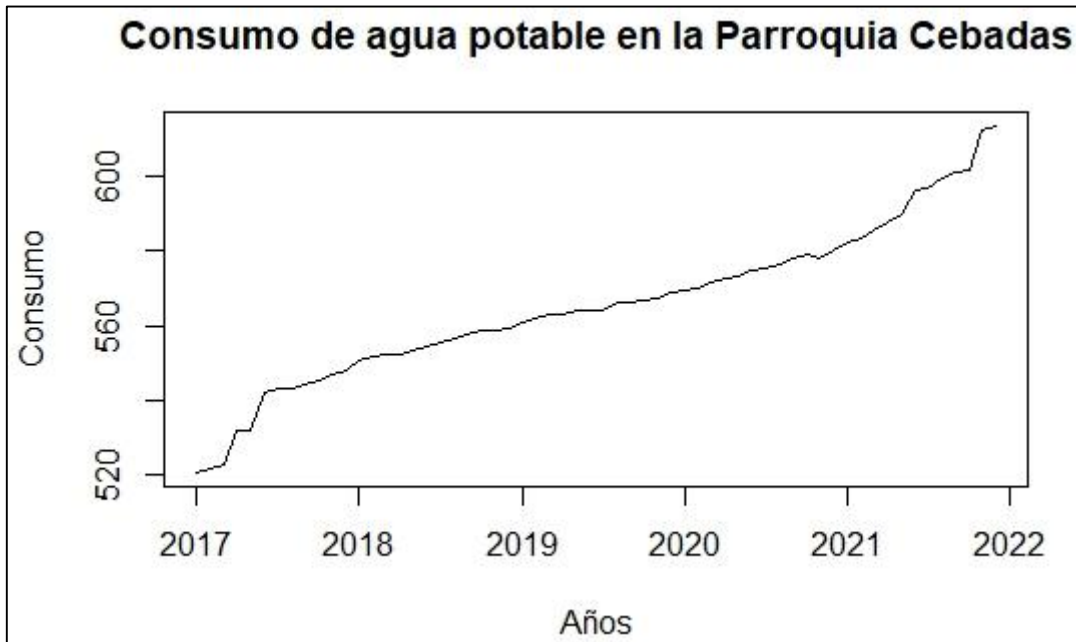
### 3.2.2. Modelo ARIMA

#### 3.2.2.1. Prueba de Dickey-Fuller

Realizamos esta prueba para observar si es estacionaria o no es estacionaria, si se aplica un modelo ARIMA O ARMA debemos saber si es estacionaria.

$H_0$ : No es estacionaria  $> 0.5$

$H_1$ : Es estacionaria  $< 0.5$



**Ilustración 8-3:** Prueba de Dickey-Fuller para el consumo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

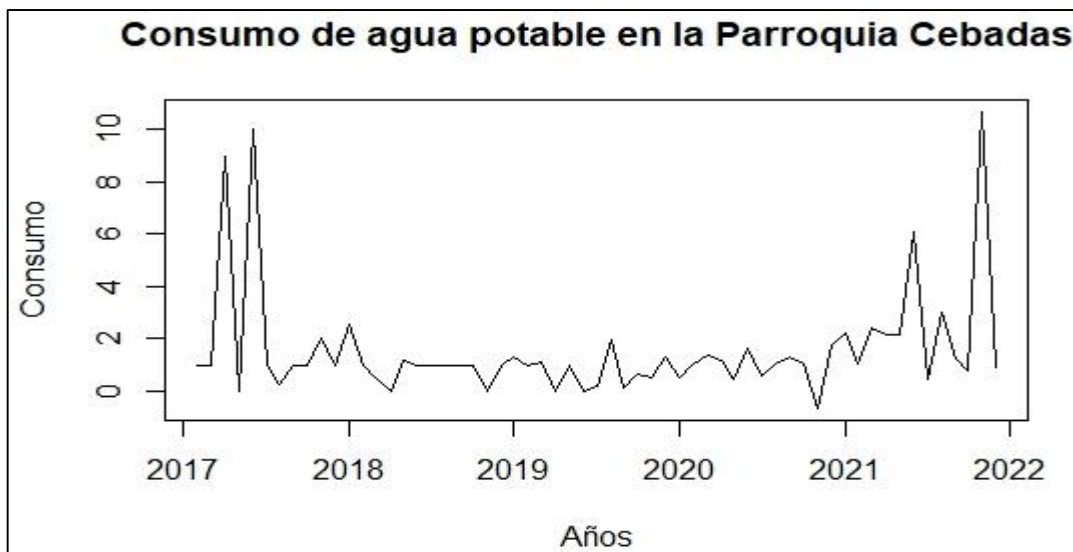
Augmented Dickey-Fuller Test

Dickey-Fuller = -0.51709, Lag order = 3, p-value = 0.9781

**Interpretación:**

Con la prueba de Dickey-Fuller obtuvimos un p-value de 0.9781 es decir que no es estacionaria, además se visualiza en el grafico que nuestra base de datos no hay estacionariedad, por lo que se realizara la primera diferencia para que sea estacionaria y así poder realizar el modelo ARIMA.

## Primera Diferencia



**Ilustración 9-3:** Primera diferencia para transformarla a estacionalidad.

Realizado por: Peñafiel, G, 2022.

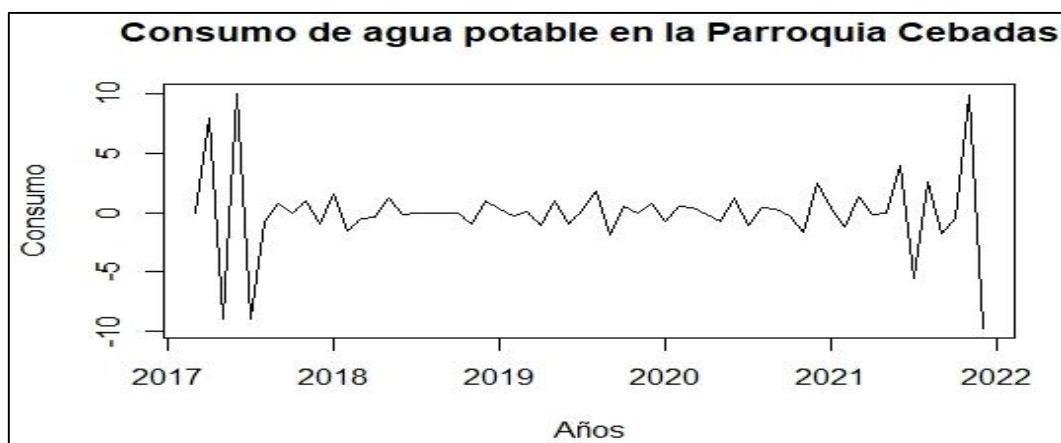
## Augmented Dickey-Fuller Test

Dickey-Fuller = -2.7158, Lag order = 3, p-value = 0.2856

### Interpretación:

Se visualiza en el gráfico que ya hay estacionariedad, pero cuando realizamos la prueba de Dickey-Fuller obtuvimos un p-value de 0.2856 es decir que no es estacionaria, por lo que se realizara la segunda diferencia para que sea estacionaria.

## Segunda Diferencia



**Ilustración 10-3:** Segunda diferencia para transformar a estacionalidad.

Realizado por: Peñafiel, Gissela, 2022.



### Augmented Dickey-Fuller Test

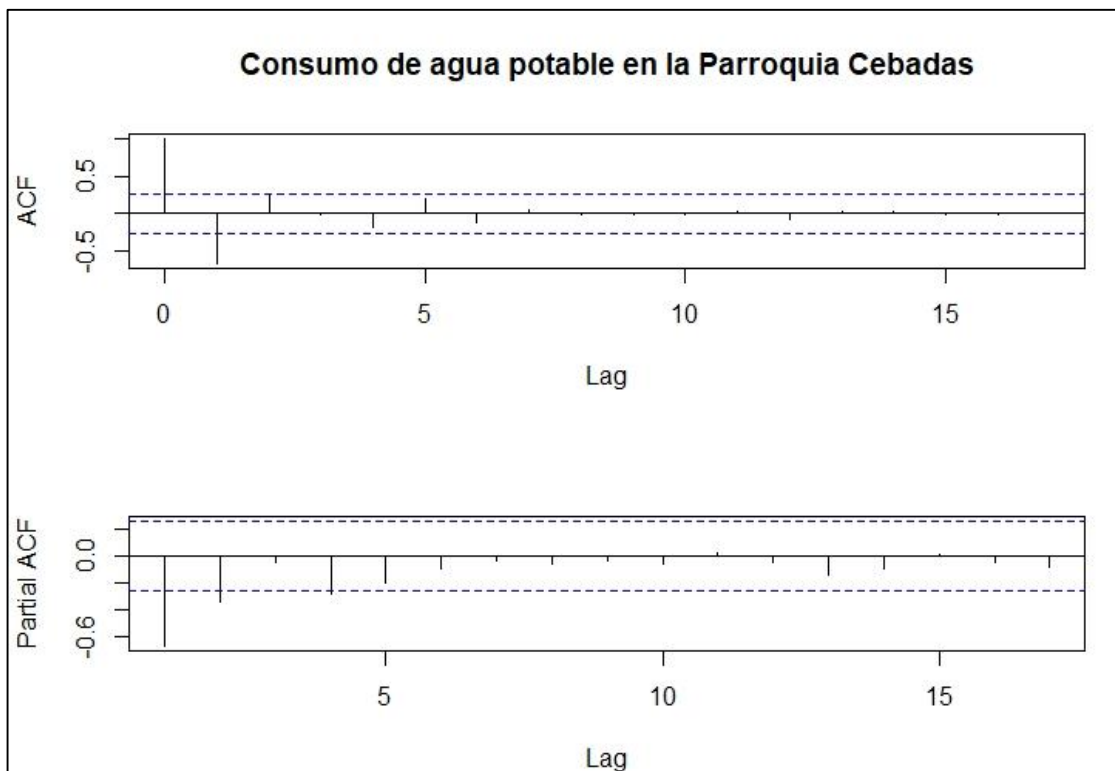
Dickey-Fuller = -7.7868, Lag order = 3, p-value = 0.01

#### **Interpretacion:**

Se visualiza en el grafico que ya hay estacionariedad ya que los datos rondan alrededor de una misma media, la media es igual a cero y todos los datos se encuentran alrededor de la media que es igual a cero y con la prueba de Dickey-Fuller obtuvimos un p-value de 0.01 es decir si hay estacionariedad, por lo que se realizara el modelo ARIMA.

#### 3.2.2.2. *Función de autocorrelacion y función de autocorrelacion parcial*

Estas dos funciones nos sirven para saber cuántas medias móviles y cuantos autoregresivos vamos a utilizar en nuestro modelo ARIMA.



**Ilustración 11-3:** Autocorrelación simple y parcial del consumo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

#### **Interpretación:**

Se visualiza en el gráfico de la función de autocorrelacion se obtuvo una media móvil y en el gráfico de autocorrelacion parcial se visualiza dos autorregresivos. Se obtuvo el siguiente modelo:

ARIMA (1,2,2)

Considerando el uso de función de autoarima, el software nos muestra que el mejor modelo viene dado por:

$$ARIMA (0,1,0) (1,0,1) [12]$$

ma1	sar1
-0.9593	0.9981

$$Y_t = 5.03 + (0.9593)Y_{t-1} - 0.9593\epsilon_{t-1}$$

La confiabilidad de los parametros de los modelos son:

**Tabla 20-3:** Confiabilidad de los parámetros del modelo ARIMA

	2.5%	97.50%
ma1	0.9709052	1.0253455
sar1	-1.2512901	-0.6673141

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

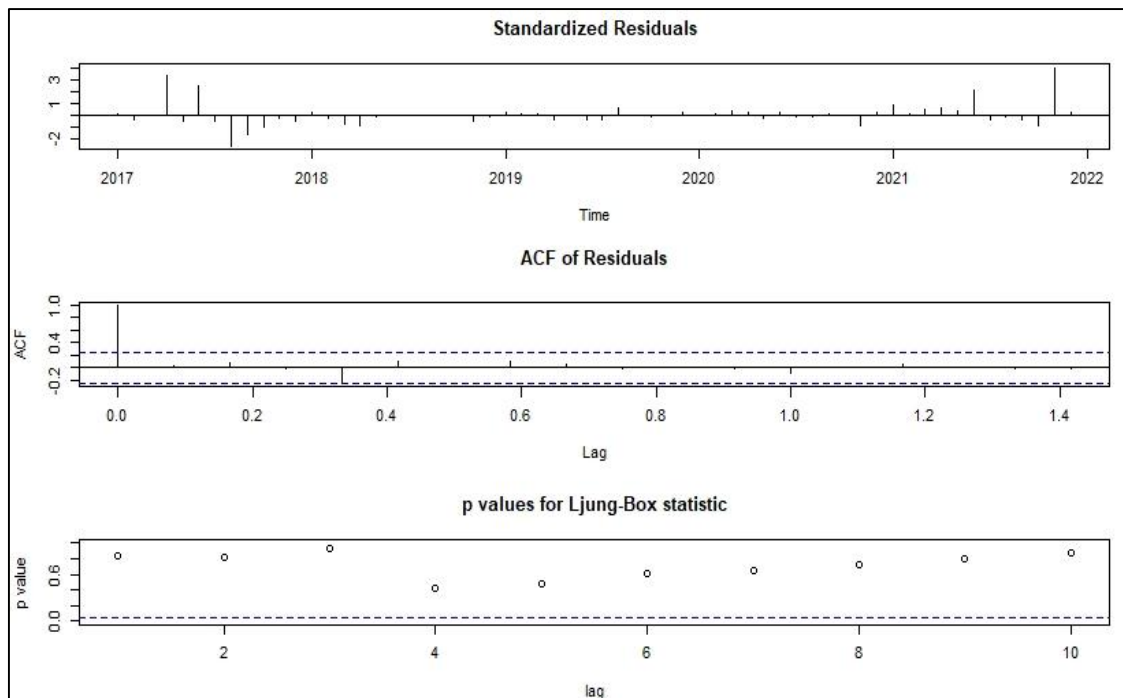
El intervalo de confianza del modelo no incluye el 0, por lo que es un modelo que se ajusta, los criterios AIC = 261,86 y el BIC = 266,01.

3.2.2.3. Prueba de validez (ruido blanco) para el modelo ARIMA del consumo de agua

Realizamos esta prueba a través de la hipótesis Ljung Box Test nos sirve para detectar si nuestro modelo es bueno y si el modelo está bien ajustado donde se comprobará que:

$$H_0: \text{Ruido blanco} > 0.5$$

$$H_1: \text{No hay ruido blanco} < 0.5$$



**Ilustración 12-3:** Prueba de validez del modelo

**Realizado por:** Peñafiel, G, 2022.

Box-Ljung test

X-squared = 0.044287, df = 1, p-value = 0.8333

**Interpretación:**

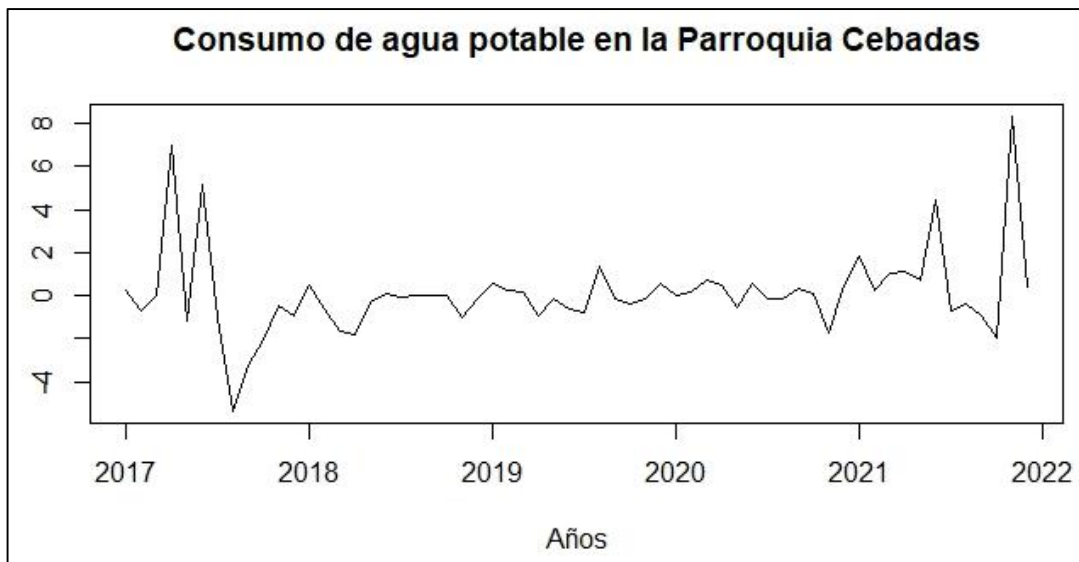
La prueba de Box-Ljung para un p-valor de 0,8333 con un nivel de confianza de 0,05 indica que el resto del modelo predictivo sobre el consumo de agua potable satisface el supuesto de independencia.

**Supuesto de heterocedasticidad**

Para probar la hipótesis de consumo heterogéneo de agua potable en las parroquias de Cebadas durante 2017-2021 según el modelo de pronóstico ARIMA, se utilizó la prueba de White para verificar:

$$H_0: \text{Existe homocedasticidad}$$

$$H_1: \text{No Existe homocedasticidad}$$



**Ilustración 13-3:** Residuales del modelo de pronóstico del consumo de agua potable .

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

Se puede evidenciar heterocedasticidad en los datos ya que la varianza de los errores no es constante en todas las observaciones realizadas.

**3.2.3. Modelo Holt Winters**

**Tabla 21-3:** Modelo Holt Winters.

Coefficientes	Valores
A	613.974
B	1.424
T	0.819

Realizado por: Peñafiel, Gissela, 2022.

Estimación de la serie con suavización exponencial Actual:

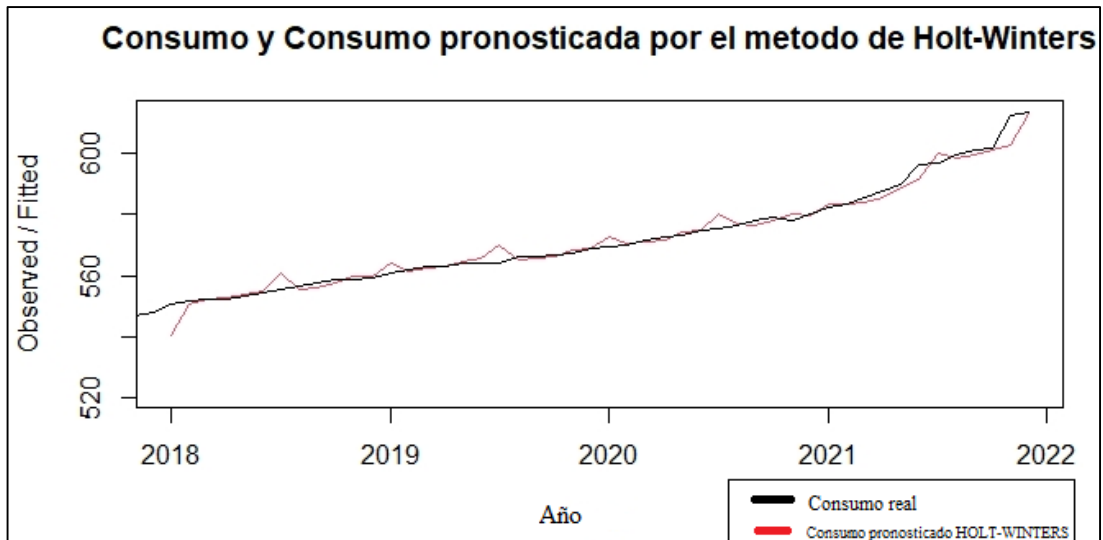
$$L_t = 613.974Y_t + (1 - 1.424)(L_t + Y_{t-1})$$

Estimación de la tendencia es:

$$T_t = 1.424(L_t + L_{t-1}) + (1 - 1.424))T_{t-1}$$

Modelo pronosticado para el periodo p para el futuro es:

$$Y_t = L_t + pT_t$$



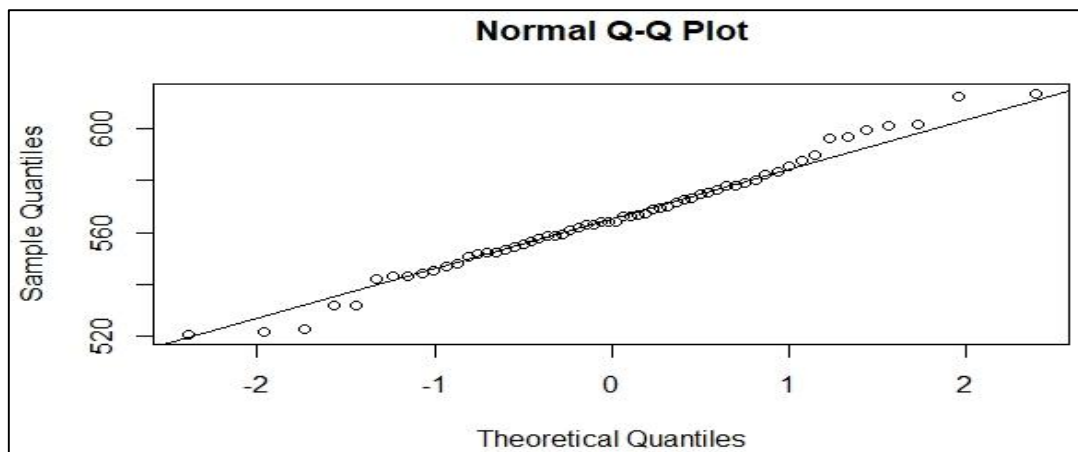
**Ilustración 14-3:** Consumo real y Consumo pronosticada por el método Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

### 3.2.3.1. Validación del modelo Holt-Winters de consumo de agua potable

#### Supuesto de Normalidad

Para comprobar la normalidad del modelo se utilizó la prueba de ajuste de Jarque-Ber, que tiene en cuenta tanto las pruebas de asimetría como las de curtosis. Para analizar el error del modelo que se verificará gráficamente con un gráfico Q-Q Plot.



**Ilustración 15-3:** Normalidad del modelo Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

#### Interpretación:

Es claro que la distribución de residuos del modelo de predicción del consumo de agua potable de la ciudad de Cebadas se asemeja a la distribución normal, como lo demuestra el valor p de la prueba de Jarque-Ber:

$H_0$ : Los datos proceden distribución normal

$H_1$ : Los datos no proceden de una distribución normal

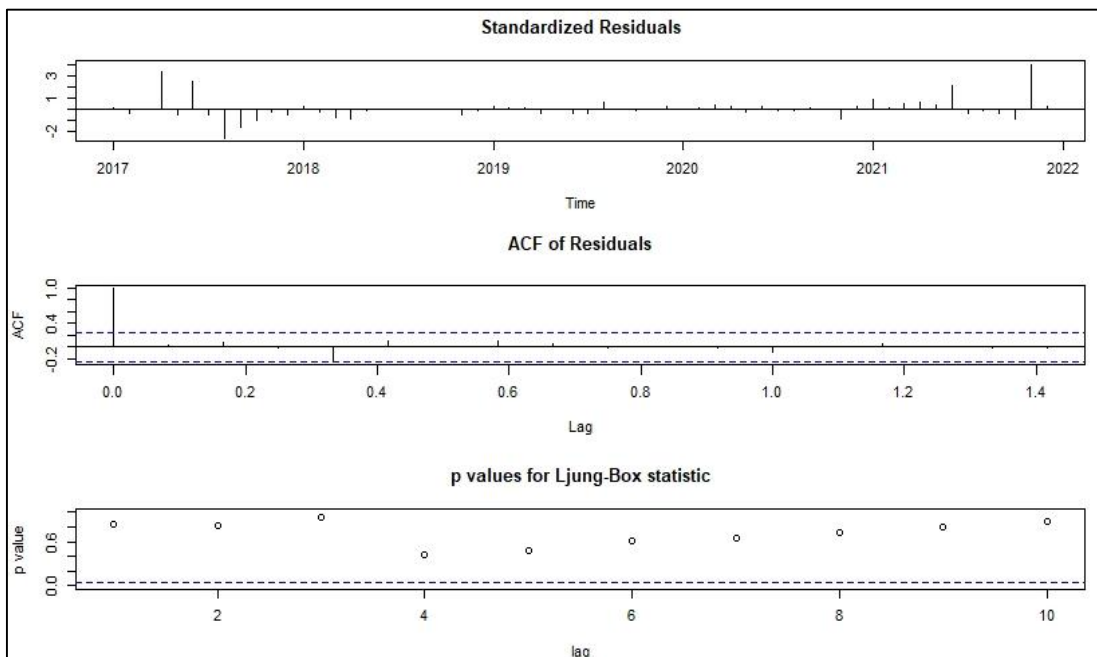
Tomando  $p = 0,9361$  con un nivel de significancia de  $0,05$ , confirmamos que el modelo cumple con el supuesto de normalidad.

### Supuesto de independencia

Para probar el supuesto de independencia del modelo de predicción ARIMA sobre el consumo de agua potable en Parroquia Cebadas durante 2017-2021, se utilizó la prueba de Box Ljung, la cual verificará:

$H_0$ : Los residuales del modelo son independientes

$H_1$ : Los residuales del modelo no son independientes



**Ilustración 16-3:** Grafico de la Prueba de Ljung-Box

**Realizado por:** Peñafiel, Gissela, 2022.

Box-Ljung test

X-squared = 0.094269, df = 1, p-value = 0.7257

### Interpretación:

En la prueba de Box-Ljung para un valor de  $p$  de  $0,725$  nos indica que el resto del modelo de predicción sobre el consumo de agua potable no cumple con el supuesto de independencia.

### 3.2.4. Medidas de precisión de los modelos de predicción de consumo de agua potable

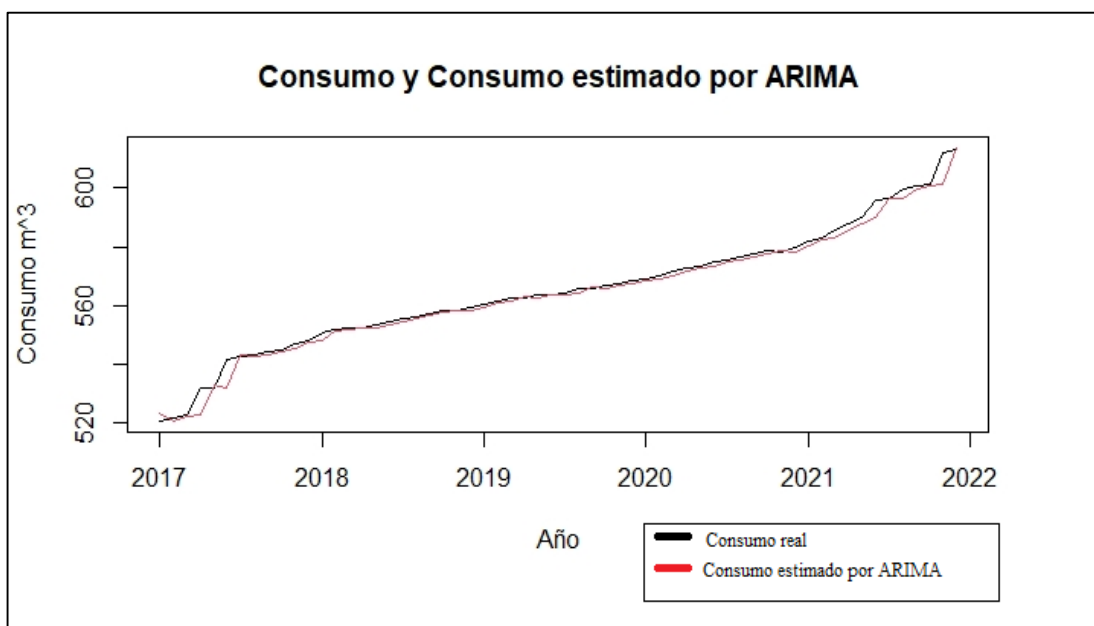
Para elegir el mejor modelo de predicción del consumo de agua potable en la parroquia Cebadas, se realizó una comparación entre las medidas de precisión del modelo ARIMA (0,1,0) (1,0,1) [12] con el modelo de Holt-Winters.

Por tanto, al cumplir con todos los supuestos de normalidad e independencia se verifica que los dos modelos tanto ARIMA y Holt.Winters son adecuados para las predicciones, por lo que procede a efectuar la comparación de las medidas de precisión.

**Tabla 22-3:** Medida de exactitud ARIMA vs Holt Winters

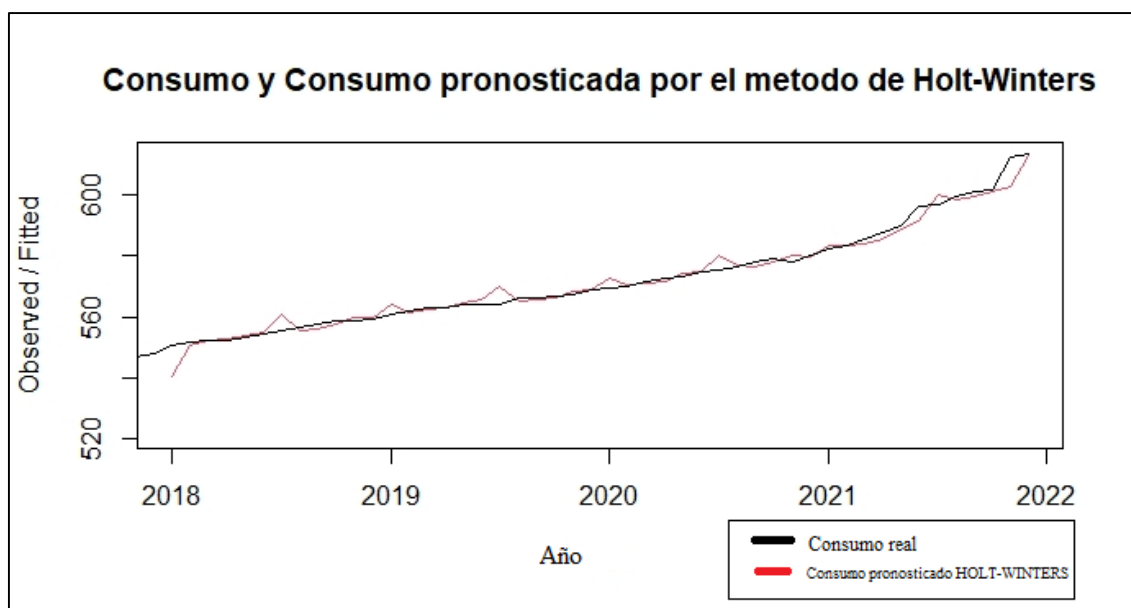
Modelo	Medidas de Precisión						
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA	0.827	2.225	1.235	0.146	0.219	0.781	-0.0602
HOLT-WINTERS	1.288	2.607	1.517	0.2259	0.268	0.959	-0.273

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 17-3:** Consumo real y Consumo estimado por ARIMA

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 18-3:** Consumo real y Consumo Pronosticada del modelo Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

Se visualizo que el mejor modelo de predicción para el consumo de agua potable en la Parroquia Cebadas es mediante el modelo ARIMA ya que este modelo gráficamente se asemeja a los datos reales.

**3.2.5. Pronostico del consumo de agua potable de la Parroquia Cebadas 2017-2021**

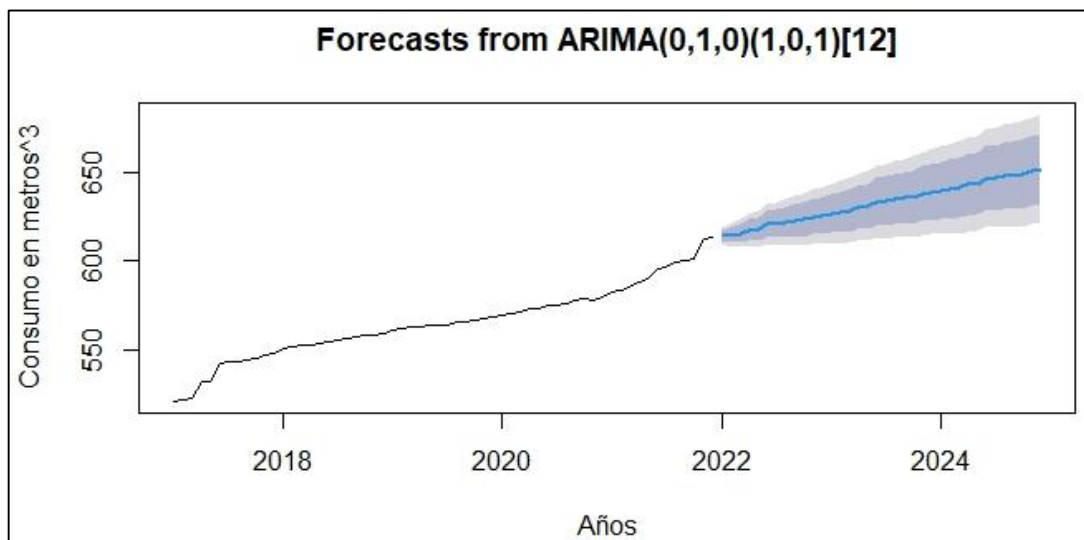
**Tabla 23-3:** Pronósticos del consumo de agua potable.

Año	Fecha	Pronóstico	L. Inferior.95	L. Superior.95
2022	enero	617.807	613.817	621.797
	febrero	620.849	615.593	626.105
	marzo	624.679	617.559	631.799
	abril	628.155	619.275	637.036
	mayo	631.791	620.956	642.626
	junio	635.354	622.493	648.215
	julio	638.950	623.945	653.956
	agosto	642.532	625.288	659.776
	septiembre	646.120	626.538	665.701
	octubre	649.705	627.693	671.716
	noviembre	653.291	628.759	677.823
	diciembre	656.877	629.737	684.016
2023	enero	660.463	630.631	690.294
	febrero	664.048	631.443	696.654
	marzo	667.634	632.174	703.094



	abril	671.220	632.828	709.612
	mayo	674.806	633.406	716.206
	junio	678.392	633.910	722.874
	julio	681.978	634.342	729.614
	agosto	685.564	634.703	736.425
	septiembre	689.150	634.995	743.305
	octubre	692.736	635.219	750.253
	noviembre	696.322	635.376	757.267
	diciembre	699.907	635.468	764.347
	enero	703.493	635.496	771.491
	febrero	707.079	635.461	778.697
	marzo	710.665	635.364	785.966
	abril	714.251	635.207	793.296
	mayo	717.837	634.989	800.685
2024	junio	721.423	634.712	808.134
	julio	725.009	634.377	815.640
	agosto	728.595	633.985	823.204
	septiembre	732.181	633.537	830.824
	octubre	735.766	633.033	838.500
	noviembre	739.352	632.474	846.231
	diciembre	742.938	631.860	854.016

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 19-3:** Pronostico del consumo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

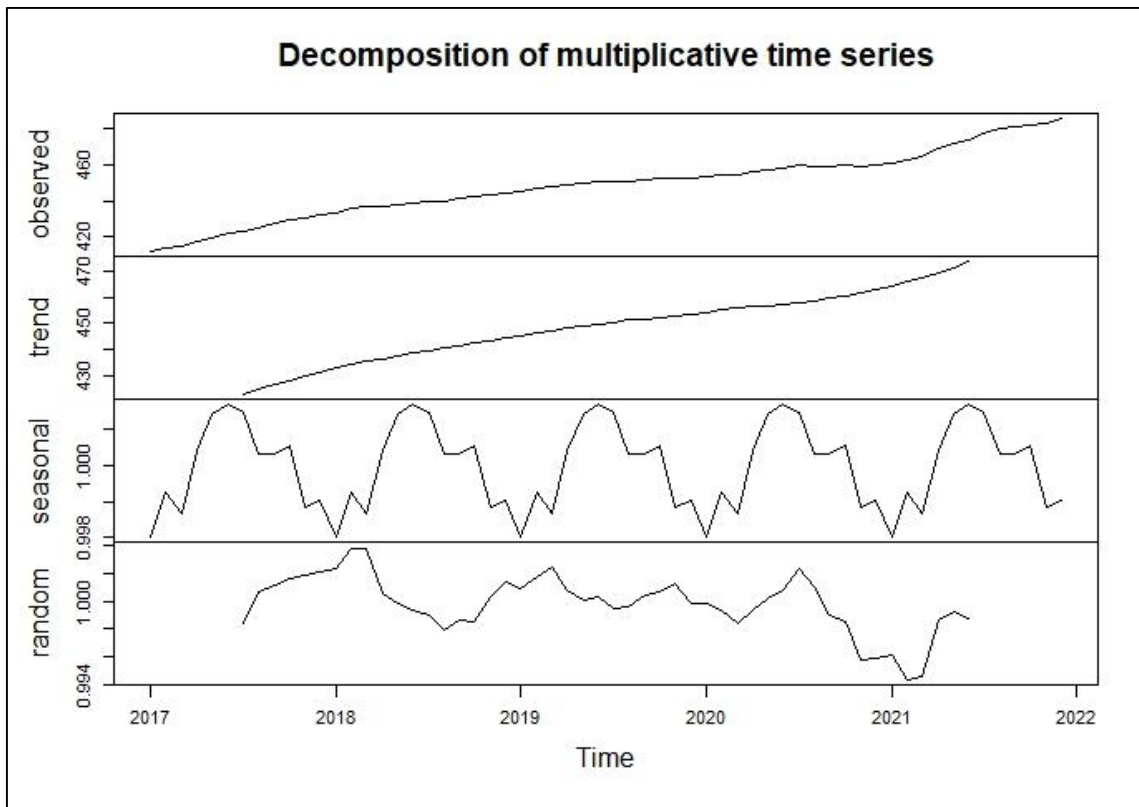
### Interpretación:

El uso del modelo descrito anteriormente puede calcular la predicción del consumo de agua potable para la cebada parroquial en el período 2017-2022. Con un intervalo confiable de 95%,

el grafico muestra que la preservación de los valores pronosticados con el modelo ARIMA (0, 1.0) (1.0.1) [12], donde puede ir a mirar una tendencia pequeña.

### 3.3. Modelación del costo de agua potable

#### 3.3.1. Descomposición de la serie de tiempo para el costo de agua en la Parroquia Cebadas.



**Ilustración 20-3:** Descomposición de la serie de tiempo del costo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

#### Interpretación:

Se visualiza en este grafico que es un modelo multiplicativo ya que la varianza no es constante a lo largo del tiempo, también se observa que hay tendencia que representa el incremento del costo de agua potable en la Parroquia Cebadas desde enero del 2017 al mes de junio del 2021.

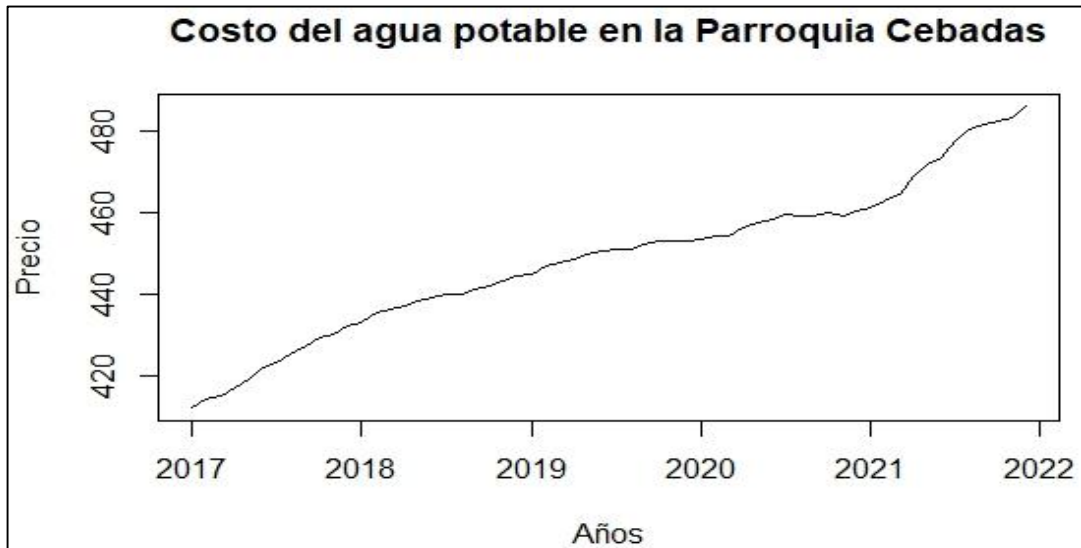
#### 3.3.2. Modelo ARIMA

##### 3.3.2.1. Prueba de Dickey-Fuller

Realizamos esta prueba para observar si es estacionaria o no es estacionaria, si se aplica un modelo ARIMA O ARMA debemos saber si es estacionaria.

$H_0$ : No es estacionaria  $> 0.5$

$H_1$ : Es estacionaria  $< 0.5$



**Ilustración 21-3:** Prueba de Dickey-Fuller

Realizado por: Peñafiel, Gissela, 2022.

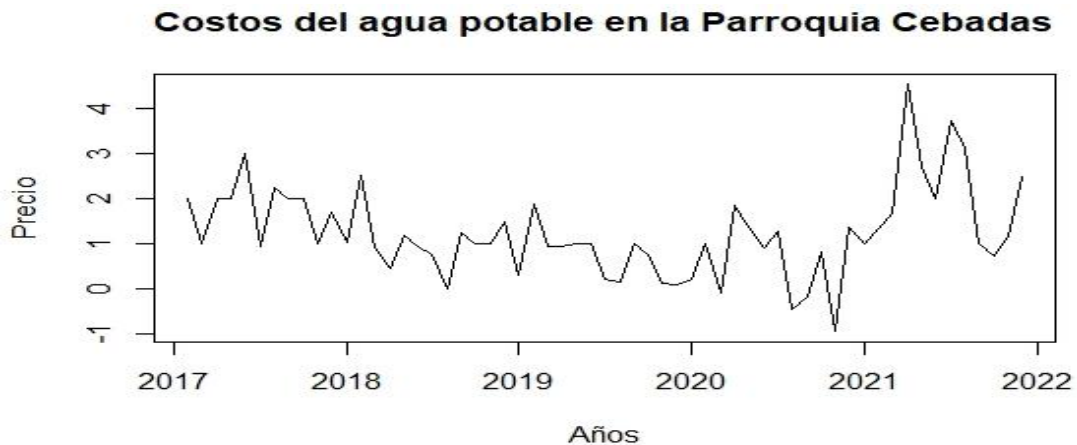
Augmented Dickey-Fuller Test

Dickey-Fuller = -2.4731, Lag order = 3, p-value = 0.3836

**Interpretación:**

Con la prueba de Dickey-Fuller obtuvimos un p-value de 0.3836 es decir que no es estacionaria, además se visualiza en el grafico que nuestra base de datos no hay estacionariedad, por lo que se realizara la primera diferencia para que sea estacionaria y así poder realizar el modelo ARIMA.

**Primera Diferencia**



**Ilustración 22-3:** Primera diferencia para transformar a estacionalidad.

Realizado por: Peñafiel, Gissela, 2022.

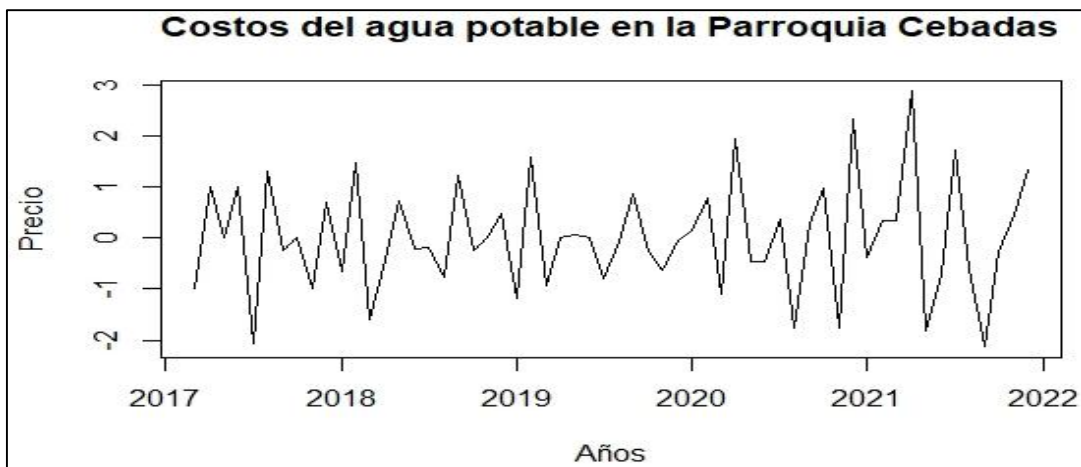
Augmented Dickey-Fuller Test

Dickey-Fuller = -2.0691, Lag order = 3, p-value = 0.5468

**Interpretación:**

Se visualiza en el grafico que ya hay estacionariedad, pero cuando realizamos la prueba de Dickey-Fuller obtuvimos un p-value de 0.5468 es decir que no es estacionaria, por lo que se realizara la segunda diferencia para que sea estacionaria.

**Segunda Diferencia**



**Ilustración 23-3:** Segunda diferencia para transformar a estacionariedad.

Realizado por: Peñafiel, Gissela, 2022.

Augmented Dickey-Fuller Test

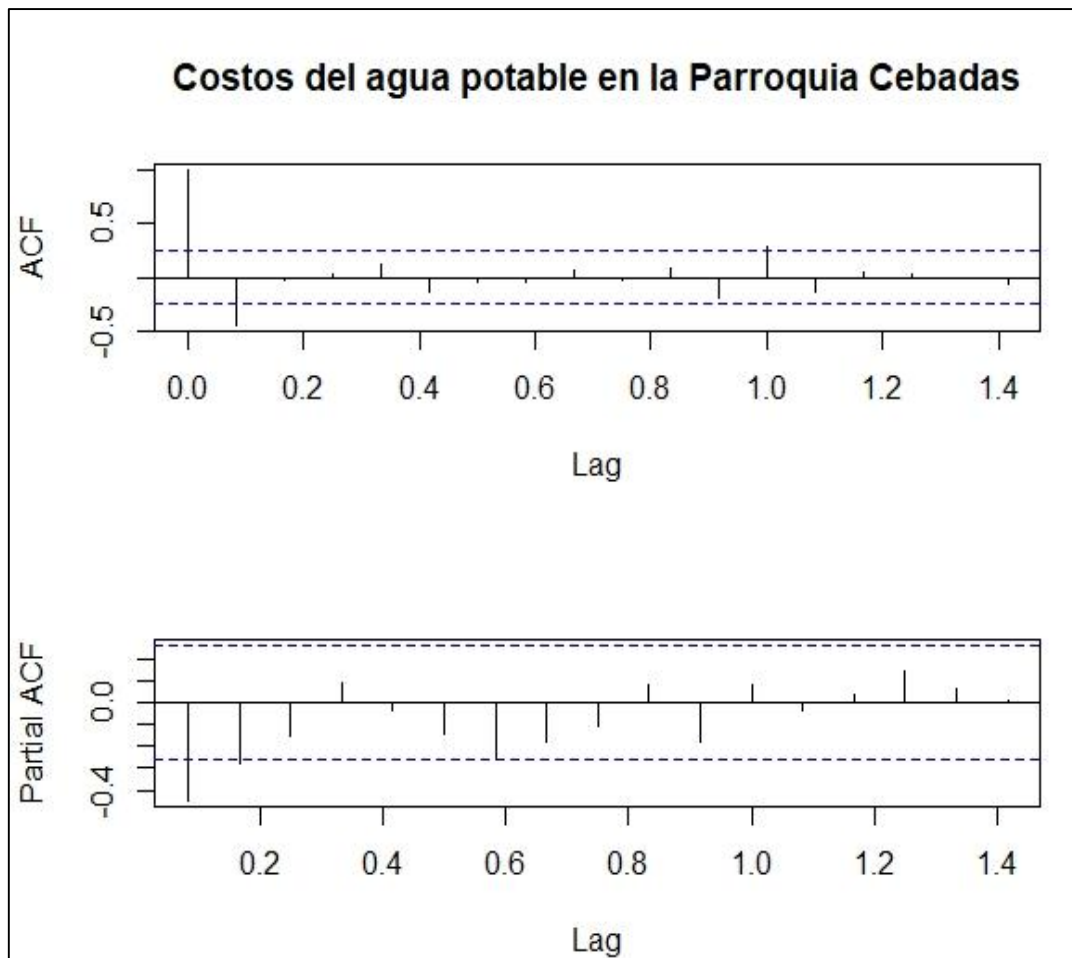
Dickey-Fuller = -4.1817, Lag order = 3, p-value = 0.01

**Interpretacion:**

Se visualiza en el grafico que ya hay estacionariedad ya que los datos rondan alrededor de una misma media, la media es igual a cero y todos los datos se encuentran alrededor de la media que es igual a cero y con la prueba de Dickey-Fuller obtuvimos un p-value de 0.01 es decir si hay estacionariedad, por lo que se realizara el modelo ARIMA.

*3.3.2.2. Función de autocorrelación y función de autocorrelación parcial*

Estas dos funciones nos sirven para saber cuántas medias móviles y cuantos autoregresivos vamos a utilizar en nuestro modelo ARIMA.



**Ilustración 24-3:** Autocorrelación simple y autocorrelación parcial del costo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

Se visualiza en el gráfico de la función de autocorrelación se obtuvo una media móvil y en el gráfico de autocorrelación parcial se visualiza dos autorregresivos. Se obtuvo el siguiente modelo:

$$ARIMA (2,1,2)$$

Considerando el uso de función de autoarima, el software nos muestra que el mejor modelo viene dado por:

$$ARIMA (1,1,1) (0,0,1) [12]$$

ar1	ma1
0.9624	-0.581

$$Y_t = 0.6582 + (1.0873)Y_{t-1} - 1.0873\epsilon_{t-1}$$

La confiabilidad de los parametros de los modelos son:

**Tabla 24-3:** Confiabilidad de los parámetros del modelo ARIMA

	2.5%	97.50%
<b>ar1</b>	0.882	1.0427
<b>ma1</b>	-0.838	-0.324
<b>sma1</b>	0.120	0.693

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

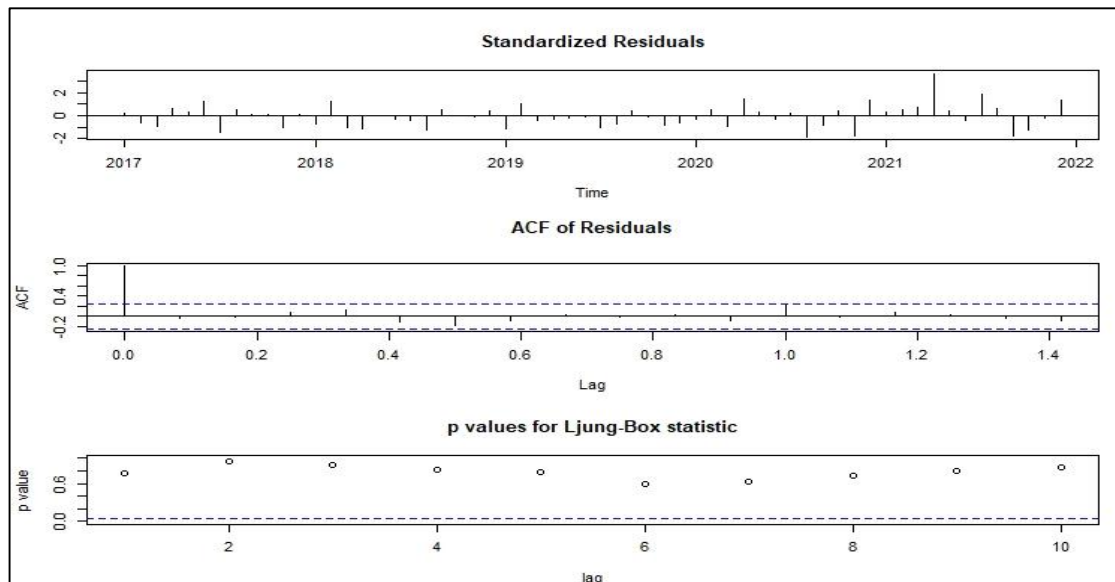
El intervalo de confianza del modelo no incluye el 0, por lo que es un modelo que se ajusta, a los criterios AIC= 154.82 y el BIC= 165.21.

3.3.2.3. Prueba de Validez (ruido blanco) para el modelo ARIMA del costo de agua potable

Realizamos esta prueba a través de la hipótesis Ljung Box Test nos sirve para detectar si nuestro modelo es bueno y si el modelo está bien ajustado donde se comprobará que:

$$H_0: \text{Ruido blanco} > 0.5$$

$$H_1: \text{No hay ruido blanco} < 0.5$$



**Ilustración 25-3:** Prueba de validez Ljung-Box.

Realizado por: Peñafiel, Gissela, 2022.

Box-Ljung test

X-squared = 0.094269, df = 1, p-value = 0.332

### Interpretación:

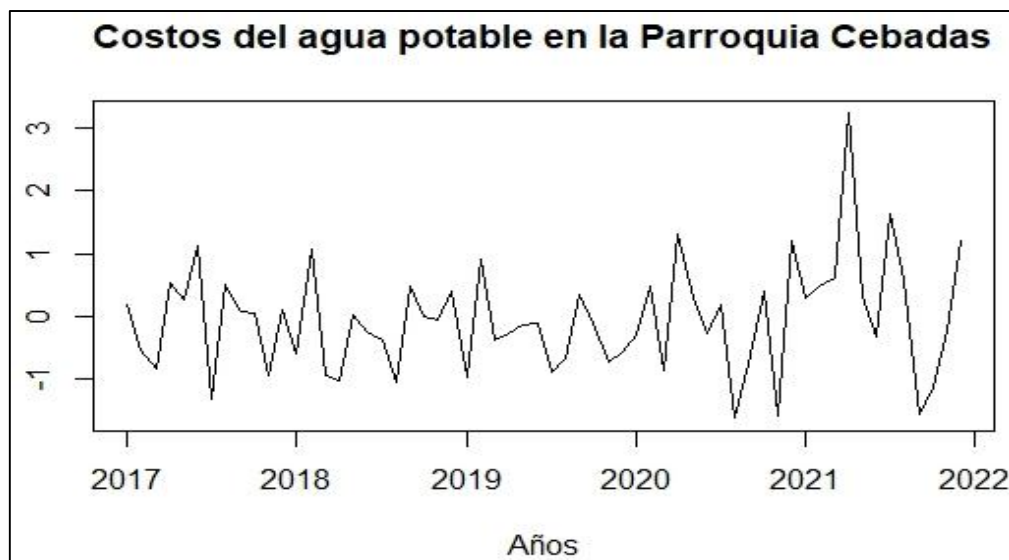
En la prueba de Box-Ljung para un p-valor de 0,332 lo que nos indica que el resto del modelo predictivo sobre el costo de agua potable no satisface el supuesto de independencia.

### Supuesto de heterocedasticidad

Para probar la hipótesis de consumo heterogéneo de agua potable en las parroquias de Cebadas durante 2017-2021 según el modelo de pronóstico ARIMA, se utilizó la prueba de White para verificar:

$H_0$ : Existe homocedasticidad

$H_1$ : No Existe homocedasticidad



**Ilustración 26-3:** Residuales del modelo de predicción del costo de agua potable.

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

Se puede evidenciar heterocedasticidad en los datos ya que la varianza de los errores no es constante en todas las observaciones realizadas.

### 3.3.3. Modelo Holt Winters

**Tabla 25-3:** Modelo Holt Winters del costo de agua potable.

Coeficientes	Valores
A	485.897
B	1.404
T	-0.427

Realizado por: Peñafiel, Gissela, 2022.

Estimación de la serie con suavización exponencial Actual:

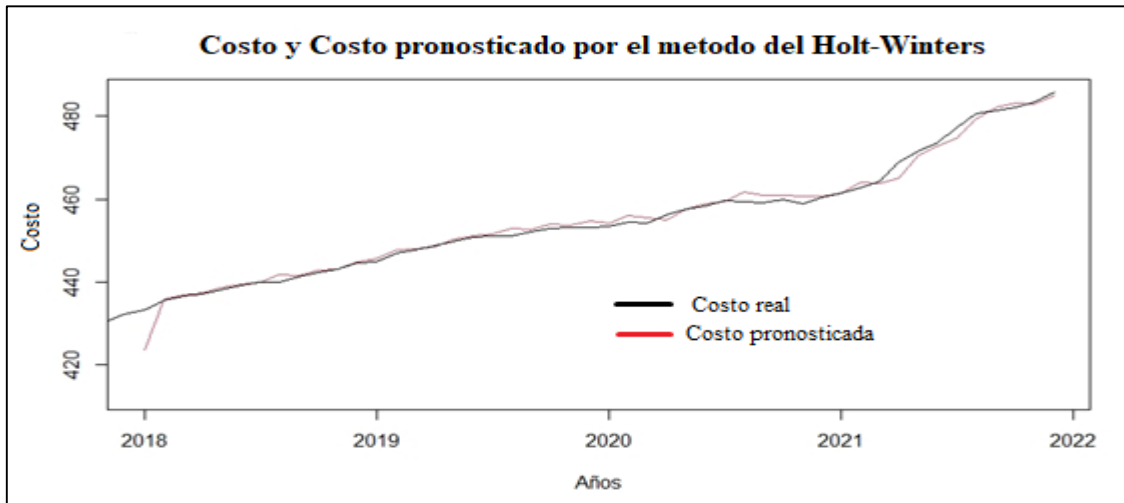
$$L_t = 485.897Y_t + (1 - 1.404)(L_t + Y_{t-1})$$

Estimación de la tendencia es:

$$T_t = 1.404(L_t + L_{t-1}) + (1 - 1.404)T_{t-1}$$

Modelo pronosticado para el periodo p para el futuro es:

$$Y_t = L_t + pT_t$$



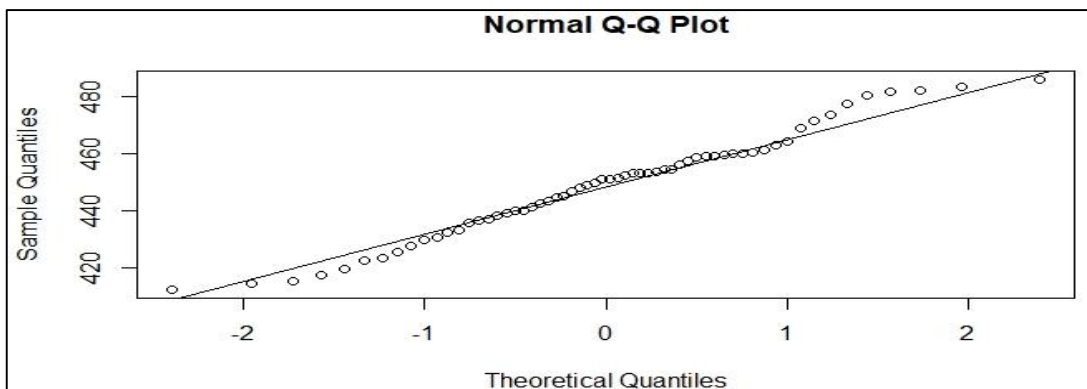
**Ilustración 27-3:** Gasto real y Gasto pronosticado por el método Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

### 3.3.3.1. Validación del modelo Holt-Winters sobre el costo de agua potable

#### Supuesto de Normalidad

Para comprobar la normalidad del modelo se utilizó la prueba de ajuste de Jarque-Ber, que tiene en cuenta tanto las pruebas de asimetría como las de curtosis. Para analizar el error del modelo que se verificará gráficamente con un gráfico Q-Q Plot.



**Ilustración 28-3:** Grafico de normalidad

Realizado por: Peñafiel, Gissela, 2022.



### Interpretación:

Es claro que la distribución de residuos del modelo de predicción del consumo de agua potable de la ciudad de Cebadas se asemeja a la distribución normal, como lo demuestra el valor p de la prueba de Jarque-Ber:

$H_0$ : Los datos proceden distribución normal

$H_1$ : Los datos no proceden de una distribución normal

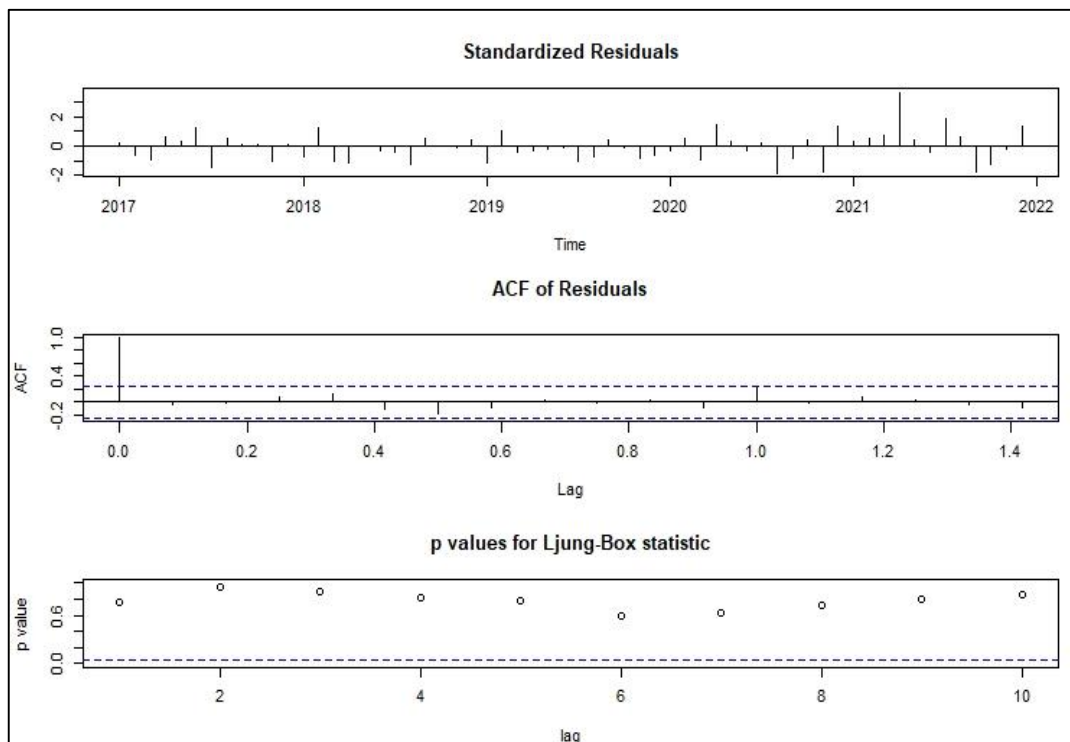
Tomando  $p = 0,7421$  con un nivel de significancia de 0,05, confirmamos que el modelo cumple con el supuesto de normalidad.

### Supuesto de independencia

Para probar el supuesto de independencia del modelo de predicción ARIMA sobre el consumo de agua potable en Parroquia Cebadas durante 2017-2021, se utilizó la prueba de Box Ljung, la cual verificará:

$H_0$ : Los residuales del modelo son independientes

$H_1$ : Los residuales del modelo no son independientes



**Ilustración 29-3:** Prueba de validez de Ljung-Box.

Realizado por: Peñafiel, Gissela, 2022.

Box-Ljung test

X-squared = 1.4991, df = 1, p-value = 0.7208

### Interpretación:

La prueba de Box-Ljung para un valor de p de 0,7208 con un nivel de confianza de 0,05 indica que el resto del modelo de predicción sobre el consumo de agua potable cumple con el supuesto de independencia.

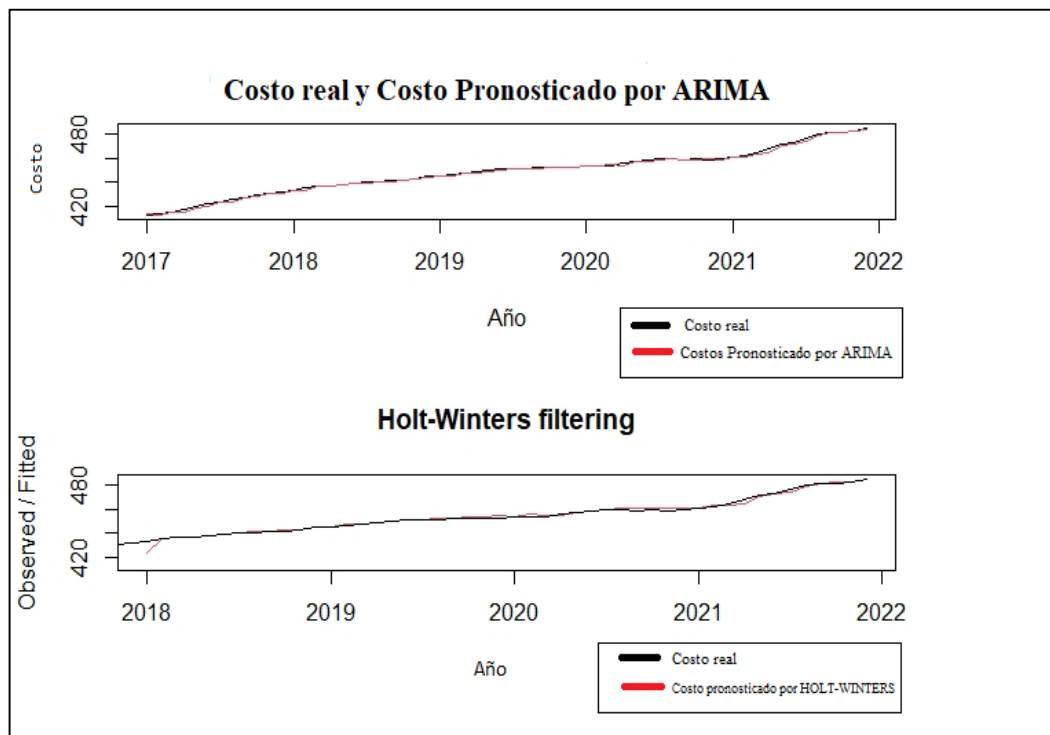
#### 3.3.4. Medidas de Precisión de los modelos de predicción de consumo de agua potable

Para elegir el mejor modelo de predicción del costo de agua potable en la parroquia Cebadas, se realizó una comparación entre las medidas de precisión del modelo ARIMA (1,1,1) (0,0,1) [12] con el modelo de Holt-Winters.

**Tabla 26-3:** Medida de precisión ARIMA vs Holt Winters.

Medida de Precisión ARIMA vs HOLT WINTERS							
Modelo	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA	0.11	0.8272	0.6566	0.0253	0.1456	0.504	-0.0450
HOLT-WINTER	-0.087	1.7853	1.0430	-0.019	0.2295	0.0802	0.1713

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 30-3:** Costo real y costo pronosticado por ARIMA y Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

### Interpretación:

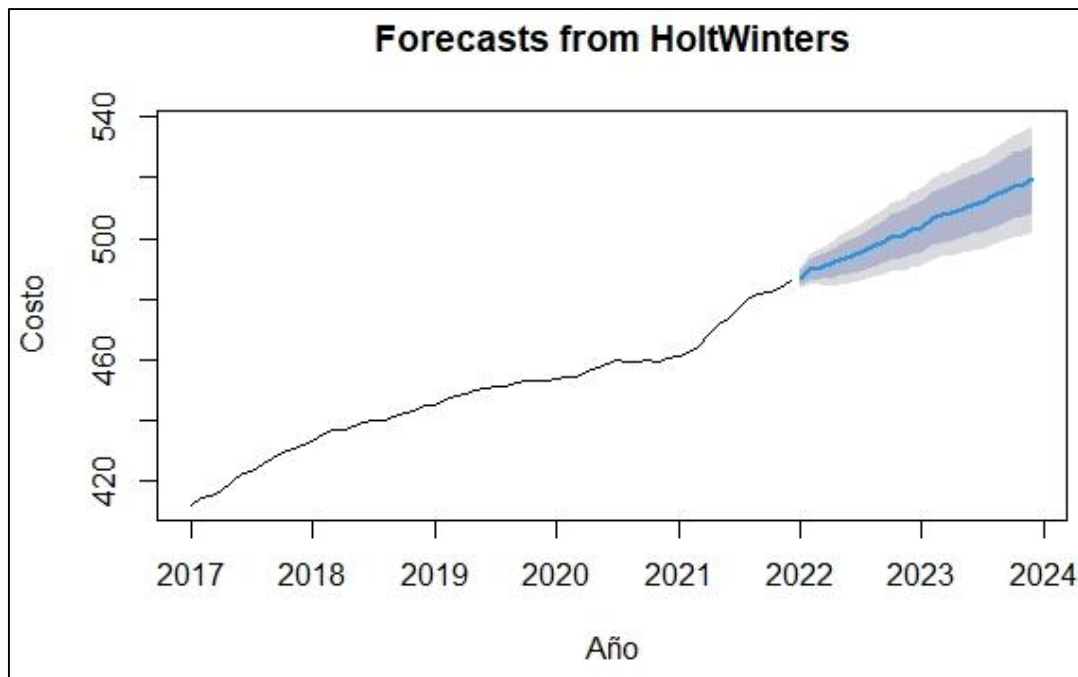
Se puede visualizar que el mejor modelo de predicción para el costo de agua potable en la Parroquia Cebadas es mediante el modelo HOLT-WINTERS ya que este modelo gráficamente se asemeja a los datos reales.

### 3.3.5. Pronósticos del costo de agua potable en la Parroquia Cebadas

**Tabla 27-3:** Pronósticos del costo de agua potable en la Parroquia Cebadas.

Año	Fecha	Pronósticos	L. Inferior.95	L. Superior.95
2022	enero	486.87	483.34	490.41
	febrero	489.49	484.49	494.48
	marzo	490.63	484.51	496.75
	abril	491.39	484.33	498.46
	mayo	492.91	485.02	500.81
	junio	494.22	485.57	502.87
	julio	495.11	485.76	504.45
	agosto	496.98	486.99	506.97
	septiembre	498.61	488.01	509.20
	octubre	500.30	489.13	511.47
	noviembre	501.09	489.37	512.80
	diciembre	502.69	490.46	514.93
2023	enero	503.72	490.99	516.46
	febrero	506.34	493.12	519.55
	marzo	507.48	493.80	521.16
	abril	508.24	494.12	522.37
	mayo	509.77	495.20	524.33
	junio	511.07	496.08	526.05
	julio	511.96	496.56	527.35
	agosto	513.83	498.04	529.63
	septiembre	515.46	499.27	531.64
	octubre	517.15	500.58	533.72
	noviembre	517.94	501.00	534.88
	diciembre	519.54	502.24	536.85
2024	enero	520.57	502.91	538.24
	febrero	523.19	505.18	541.20
	marzo	524.33	505.98	542.68
	abril	525.09	506.40	543.78
	mayo	526.62	507.59	545.64
	junio	527.92	508.57	547.27
	julio	528.81	509.14	548.48
	agosto	530.68	510.70	550.66
	septiembre	532.31	512.02	552.60
	octubre	534.00	513.40	554.59
	noviembre	534.79	513.89	555.69
	diciembre	536.39	515.20	557.59

Realizado por: Peñafiel, Gissela, 2022.



**Ilustración 31-3:** Pronósticos del costo con el método Holt-Winters.

Realizado por: Peñafiel, Gissela, 2022.

**Interpretación:**

El uso del modelo descrito anteriormente puede calcular la predicción del consumo de agua potable para la cebada parroquial en el período 2017-2022. Con un intervalo confiable de 95%, el grafico muestra que la preservación de los valores pronosticados con el modelo Holt Winters, donde puede ir a mirar una tendencia pequeña.

**3.4. Comprobación de hipótesis de la investigación**

**3.4.1. Hipótesis del consumo de agua potable**

**Formulación de la hipótesis**

$$H_0: u_{cr} \geq u_{cp}$$

$$H_0: u_{cr} < u_{cp}$$

Donde:

$u_{cr}$  = La media del consumo ( $m^3$ ) de los datos reales.

$u_{cp}$  = la media del consumo ( $m^3$ ) de los datos pronosticados.

**Nivel de significancia:**  $\alpha = 0.05$

**Estadístico de prueba:** Los datos del consumo de agua potable reales y pronosticados siguen una distribución normal según el contraste de hipótesis de Kolmogorov- Smirnov.

**Tabla 28-3:** Prueba de Kolmogorov-Smirnov para el consumo de agua potable.

Prueba de Kolmogorov-Smirnov para una muestra			
		Consumo Real	Consumo Pronosticado
N		60	36
Parámetros normales <sup>a,b</sup>	Media	565.3673	680.1922
	Desv. Desviación	20.88058	37.76735
Máximas diferencias extremas	Absoluto	.048	.068
	Positivo	.047	.068
	Negativo	-.048	-.068
Estadístico de prueba		.048	.068
Sig. asintótica(bilateral)		.200 <sup>c,d</sup>	.200 <sup>c,d</sup>
a. La distribución de prueba es normal.			
b. Se calcula a partir de datos.			

Realizado por: Peñafiel, Gissela, 2022.

Para ello, se utilizó la prueba t de Student para comparar los consumos de agua potable, teniendo en cuenta los siguientes datos:

**Tabla 29-3:** Consumo de agua potable reales y pronosticados.

Nº	Consumo Real	Consumo Pronosticado
1	520.965	617.807
2	521.965	620.849
3	522.965	624.679
4	531.965	628.155
5	531.965	631.791
6	541.965	635.354
7	542.965	638.950
8	543.195	642.532
9	544.195	646.120
10	545.195	649.705
11	547.195	653.291
12	548.195	656.877
13	550.764	660.463
14	551.747	664.048
15	552.201	667.634
16	552.204	671.220
17	553.385	674.806
18	554.395	678.392
19	555.398	681.978

20	556.399	685.564
21	557.4	689.150
22	558.402	692.736
23	558.403	696.322
24	559.404	699.907
25	560.686	703.493
26	561.687	707.079
27	562.793	710.665
28	562.81	714.251
29	563.812	717.837
30	563.826	721.423
31	564.002	725.009
32	565.994	728.595
33	566.104	732.181
34	566.736	735.766
35	567.289	739.352
36	568.566	742.938
37	569.104	
38	570.126	
39	571.522	
40	572.705	
41	573.148	
42	574.813	
43	575.398	
44	576.422	
45	577.709	
46	578.729	
47	578.093	
48	579.881	
49	582.104	
50	583.126	
51	585.522	
52	587.705	
53	589.848	
54	595.913	
55	596.398	
56	599.422	
57	600.678	
58	601.429	
59	612.093	
60	613.011	

---

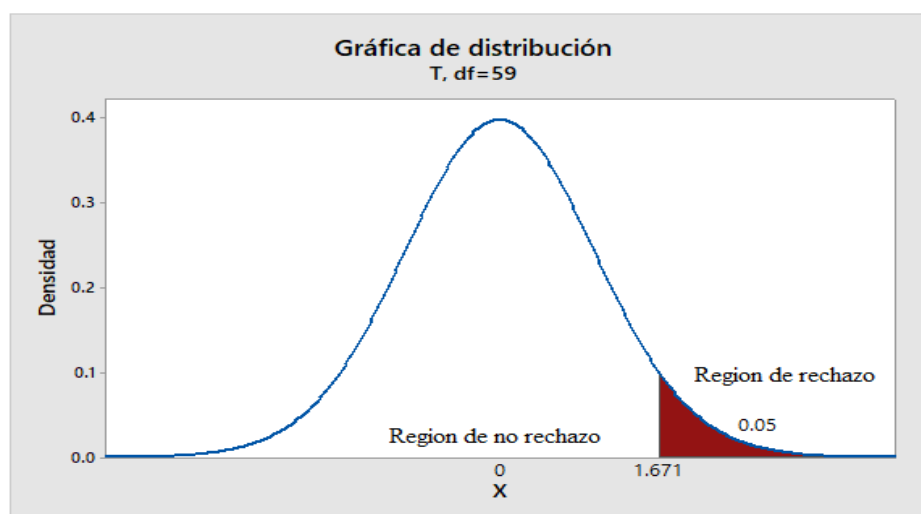
**Realizado por:** Peñafiel, Gissela, 2022.

Formula T-student

$$t' = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} \text{ g.l.}$$

**Formular una Regla de decisión**



**Ilustración 32-3:** Grafico de región de rechazo

Realizado por: Peñafiel, Gissela, 2022.

El estadístico  $t_c = 108.06$ , es decir es superior al valor calculado por lo que se rechaza la hipótesis nula.

**Decisión:** Rechazamos la hipótesis nula ( $H_0$ ), se puede decir, que el consumo de agua potable pronosticada se incrementara significativamente, con un nivel de significancia de  $\alpha = 0.05$ .

### 3.4.2. Hipótesis del costo de agua potable

**Formulación de la hipótesis**

$$H_0: u_{ctr} \geq u_{ctp}$$

$$H_0: u_{ctr} < u_{ctp}$$

Donde:

$u_{pr}$  = La media del precio del consumo ( $m^3$ ) de los datos reales.

$u_{pp}$  = la media del precio del consumo ( $m^3$ ) de los datos pronosticados.

**Nivel de significancia:**  $\alpha = 0.05$

**Estadístico de prueba:** Los datos del costo de agua potable reales y pronosticados siguen una distribución normal según el contraste de hipótesis de Kolmogorov- Smirnov.

**Tabla 30-3:** Prueba de Kolmogorov-Smirnov del costo de agua potable.

<b>Prueba de Kolmogorov-Smirnov para una muestra</b>			
		Precio Reales	Precio Pronosticado
N		60	36
Parámetros normales <sup>a,b</sup>	Media	448.8090	512.5247
	Desv. Desviación	18.34153	14.44438
Máximas diferencias extremas	Absoluto	.064	.068
	Positivo	.064	.062
	Negativo	-.060	-.068
Estadístico de prueba		.064	.068
Sig. asintótica(bilateral)		.200 <sup>c,d</sup>	.200 <sup>c,d</sup>

a. La distribución de prueba es normal.  
b. Se calcula a partir de datos.

**Realizado por:** Peñafiel, Gissela, 2022.

Para ello, se utilizó la prueba t de Student para comparar los costos de agua potable:

**Tabla 31-3:** Costo reales y pronosticados.

N°	Costo Real	Costo Pronosticado
1	412.32	486.87
2	414.32	489.49
3	415.32	490.63
4	417.32	491.39
5	419.32	492.91
6	422.32	494.22
7	423.25	495.11
8	425.48	496.98
9	427.48	498.61
10	429.48	500.30
11	430.48	501.09
12	432.17	502.69



13	433.20	503.72
14	435.72	506.34
15	436.63	507.48
16	437.10	508.24
17	438.28	509.77
18	439.23	511.07
19	440.00	511.96
20	440.00	513.83
21	441.23	515.46
22	442.23	517.15
23	443.23	517.94
24	444.72	519.54
25		520.57
26		523.19
27		524.33
28		525.09
29		526.62
30		527.92
31		528.81
32		530.68
33		532.31
34		534.00
35		534.79
36		536.39

---

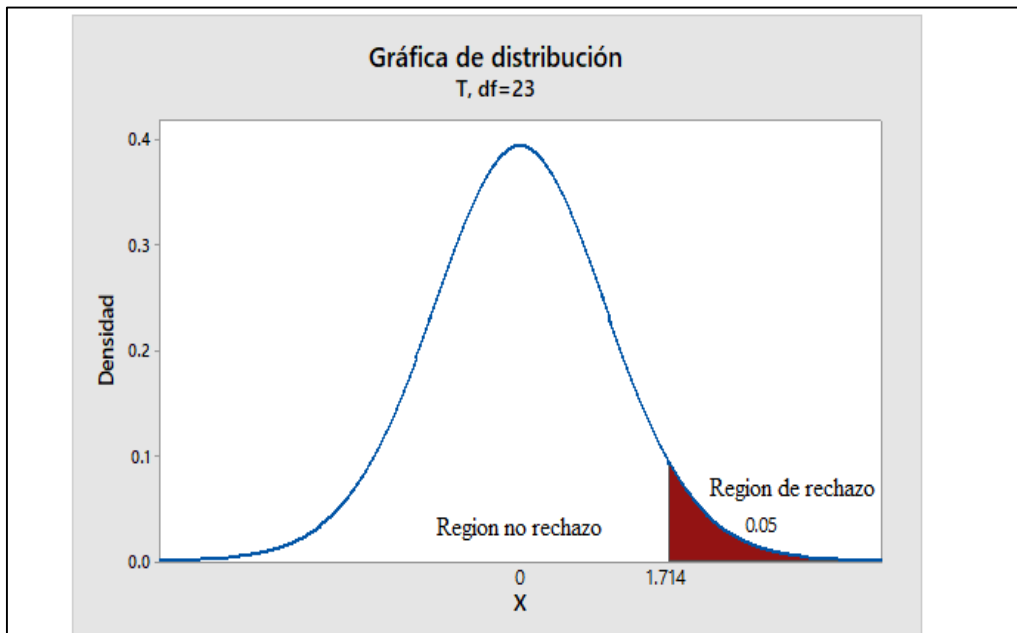
**Realizado por:** Peñafiel, Gissela, 2022.

Formula de T-student

$$t' = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} \text{ g.l.}$$

## Formular una regla de decisión



**Ilustración 33-3:** Grafico de región de rechazo.

**Realizado por:** Peñafiel, Gissela, 2022.

El estadístico  $t_c = 212.89$  es decir es superior al valor calculado por lo que se rechaza la hipótesis nula

### Decisión

Rechazamos la hipótesis nula ( $H_0$ ), se puede decir, que el costo de agua potable pronosticada se incrementara significativamente, con un nivel de significancia de  $\alpha = 0.05$ .

## CONCLUSIONES

- Se concluyó que, mediante un análisis estadístico descriptivo del consumo de agua potable, se observó en el mes de Enero del 2017 un consumo bajo con  $536.89 \text{ m}^3$ , en el mes de Diciembre del 2021 llegó hasta su pico más alto en consumo con  $595.60 \text{ m}^3$  de agua potable.
- El mejor modelo para pronosticar el consumo de agua potable en la parroquia Cebadas es *ARIMA* (0,1,0) (1,0,1) [12] con parámetros  $\mu = -0.9593$  y  $\sigma = 0,9981$ , el modelo fue confirmado por la prueba Ljung - Box test, y para pronosticar el costo del agua potable se escogió el modelo de pronóstico de Holt-Winters, con parámetros  $\alpha = 613.97$  y  $\beta = 1424$ .
- Con el uso de la prueba Ljung-Box el modelo de pronóstico *ARIMA* (0,1,0) (1,0,1) [12] sobre el consumo de agua potable en la Parroquia Cebadas se verificó que el modelo es el adecuado, de igual manera para el modelo *HOLT-WINTERS* se verificó con la prueba de Ljung-Box y normalidad, se define como los mejores modelos de pronóstico para el consumo y costo de agua potable respectivamente.
- El consumo y costo del agua potable de la Parroquia Cebadas se incrementará significativamente los próximos 2 años a un nivel de significancia de 0.05.

## **RECOMENDACIONES**

- Se sugiere a la junta administradora de agua potable realizar más captación de agua ya que aumentara el consumo en los próximos años de este líquido vital y emplear métodos estadísticos.
- Se propone igualmente realizar pronósticos periódicamente para una mejor toma de decisiones a favor de los usuarios y para la Junta administrativa en la Parroquia Cebadas.
- Se recomienda por último a la Junta administrativa de agua potable en la Parroquia Cebadas implementar y utilizar los modelos de pronósticos que se realizó en el presente trabajo sobre del consumo y costos del agua potable.

## BIBLIOGRAFIA

**ALAN, A.** *Foundations of linear and generalized linear models*. Canada: John Wiley & Sons, Inc. All rights reserved. 2015.

**AMPARO, A.** *Estadística*. [en línea] 2015. Disponible en: <http://amparooa26.blogspot.com/2015/05/en-este-seminario-avanzamos-en-la.html>

**APARICIO, J.** *Modelos lineales Aplicados en R*. (Universidad Miguel Hernández).2013.

**CESAR, M. et al.** *Estadística Descriptiva y Probabilidad para las ciencias de la información con el uso del SPSS*. Peru(Lima): Asociacion de Bibliotecologos del Peru.2020.

**HERNANDEZ, M.** *Analisis de serie de tiempo*. Mexico, 2015.

**JARA, J.** *Cierre de brechas en la cobertura de agua potable y alcantarillado sanitario*. Quito. 2017.

**JAVIER, G. et al.** *Estadística básica*. Madrid: 2011.

**LUIS, T.** *Modelos de series de tiempo*.2015.

**MERCHAN, J.** *Acceo y consumo de agua potable para todos*. [en línea] 2018. Disponible en: <http://www.eluniverso.com/opinion/2018/06/20/nota/6819557/acceso-consumo-agua-potable-todos/>

**MONROY, S.** *Estadística Descriptiva*. México: Instituto Politécnico Nacional. 2008.

**MONTESINOS, A.** *Estudio del AIC y BIC en la selección de modelos de vida con datos censurados*. [Arte] (Centro de Investigación en Matemáticas A.C.). 2011.

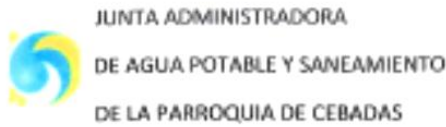
**RUSTOM J.** *Estadística descriptiva, probabilidad e inferencia*. Santiago de Chile: s.n.2012.

**VIEDMA, C.** *Estadística descriptiva e inferencial*. Madrid, España: Ediciones IDT. 2018.



## ANEXOS

### ANEXO A: AVAL DE LA JUNTA ADMINISTRADORA DE AGUA POTABLE



Cebadas, 8 de Julio del 2022

Señorita

Gissela Marisol Peñafiel Malan

ESTUDIANTE

Presente. -

De mi consideración

A nombre de la junta Administradora de Agua Potable y Saneamiento de la Parroquia de Cebadas reciba un cordial y atento saludo.

En respuesta al oficio, en la cual solicita en su parte pertinente "información de cuatro años (2017,2018,2019,2020) de los datos de consumo de agua potable según registros históricos que administra la empresa " con la finalidad de ser usado como base para los resultados que serán expuestos en la tesis de grado denominada "Pronósticos en los Consumos de Agua Potable Según Registros Históricos Generados en la Parroquia Cebadas Durante el Periodo 2017-2021"

Particular que comunico para los fines pertinentes.

Atentamente,

Sr. Marco Maldonado



**PRESIDENTE DE LA JUNTA ADMINISTRADORA DE AGUA POTABLE Y SANEAMIENTO DE LA PARROQUIA CEBADAS**  
Tel: 0997357795

Tel.: 0997367795

Dri.: Cristóbal Colón y García Moreno

Cebadas - Guamote - Ecuador

## ANEXO B: CODIGO R-ESTUDIO PARA ANALISIS ESTADÍSTICO

```
library(readxl)
Libro2 <- read_excel("Libro2.xlsx")
View(Libro2)
datos<-c(Libro2$ENERO,Libro2$EBRERO,Libro2$MARZO,Libro2$ABRIL,
         Libro2$MAYO,Libro2$JUNIO,Libro2$JULIO,Libro2$AGOSTO,
         Libro2$SEPTIEMBRE,Libro2$OCTUBRE,Libro2$NOVIEMBRE,
         Libro2$DICIEMBRE)

max(datos)
min(datos)
library(fdth)
tfr <- fdt(datos, start = 400, end = 500, h = 20)
tfr
hist(datos,main="Consumo de agua potable en la Parroquia Cebadas",ylab="Densidad",xlab =
"consumo en metros cubicos", col = "pink",freq=FALSE)
lines(density(datos),lwd = 4)
#Medidas descriptivas
#minimo
min(datos)
#cuartil 1
cuartil<-quantile(datos)
cuartil[2]
#mediana
median(datos)
#media
mean(datos)
#cuartil 3
cuartil[4]
#maximo
max(datos)

#boxplot
boxplot(datos, col="yellow", main = "Consumo de agua potable en la Parroquia Cebadas")

#Medidas de centralizacion
```

```
#mdia
```

```
mean(datos)
```

```
#mediana
```

```
median(datos)
```

```
#moda
```

```
library(modeest)
```

```
mfv(datos)
```

```
#Medidas de dispersion
```

```
#varianza
```

```
var(datos)
```

```
#desviacion tipica
```

```
sd(datos)
```

```
#coeficiente de variacion
```

```
sd(datos)/mean(datos)
```

```
#Medidas de asimetria y curtosis
```

```
#asimetria
```

```
library(EnvStats)
```

```
skewness(datos)
```

```
#curtosis
```

```
kurtosis(datos)
```

```
##### Costo del agua potable #####
```

```
library(readxl)
```

```
Libro2 <- read_excel("Libro2.xlsx")
```

```
View(Libro2)
```

```
datos<-c(Libro2$ENERO,Libro2$EBRERO,Libro2$MARZO,Libro2$ABRIL,  
         Libro2$MAYO,Libro2$JUNIO,Libro2$JULIO,Libro2$AGOSTO,  
         Libro2$SEPTIEMBRE,Libro2$OCTUBRE,Libro2$NOVIEMBRE,  
         Libro2$DICIEMBRE)
```

```
max(datos)
```

```
min(datos)
```

```
library(fdth)
```

```
tfr <- fdt(datos, start = 400, end = 500, h = 20)
```



```
tfr
hist(datos,main="Costos del agua potable en la Parroquia Cebadas",ylab="Densidad",xlab =
"costo del agua potable", col = "pink",freq=FALSE)
lines(density(datos),lwd = 4)
#Medidas descriptivas
#minimo
min(datos)
#cuartil 1
cuartil<-quantile(datos)
cuartil[2]
#mediana
median(datos)
#media
mean(datos)
#cuartil 3
cuartil[4]
#maximo
max(datos)

#boxplot
boxplot(datos, col="yellow", main = "Costos del agua potable en la Parroquia Cebadas")

#Medidas de centralizacion
#mdia
mean(datos)
#mediana
median(datos)
#moda
library(modeest)
mfv(datos)

#Medidas de dispersion
#varianza
var(datos)
#desviacion tipica
sd(datos)
#coeficiente de variacion
```

```
sd(datos)/mean(datos)
```

```
#Medidas de asimetria y curtosis
```

```
#asimetria
```

```
library(EnvStats)
```

```
skewness(datos)
```

```
#curtosis
```

```
kurtosis(datos)
```

```
##### Modelos Pronosticado#####
```

```
library(tseries)
```

```
library(urca)
```

```
library(forecast)
```

```
library(lmtest)
```

```
library(fUnitRoots)
```

```
library(readxl)
```

```
Libro3 <- read_excel("Libro3.xlsx")
```

```
View(Libro3)
```

```
datos<-c(Libro3$Consumo,Libro3$Costos)
```

```
boxplot(datos)
```

```
summary(datos)
```

```
##
```

```
par(mfrow=c(2,1),mar=c(4,4,4,1)+.1)
```

```
acf(datos, main="Autocorrelograma simple del consumo de agua potable")
```

```
#hay dos resagos pero sigue siendo estacionaria
```

```
pacf(datos, main="Autocorrelograma parcial del consumo de agua potable")
```

```
par(mfrow=c(2,1))
```

```
serie<-ts(Libro3$Consumo,frequency =12)
```

```
ar.ts=ts(Libro3$Consumo,start = c(2017,1),frequency = 12)
```

```
ar.ts
```

```
plot(ar.ts,main="Consumo de agua potable en la Parroquia Cebadas",ylab="Consumo",xlab =  
"Años")
```

```
# logaritmo
```

```
serielog=log(ar.ts)
```

```
serielog
```

```

plot(serielog)
adf.test(serielog)
#primera diferencia

dos=diff(ar.ts)
dos
plot(dos, main="Consumo de agua potable en la Parroquia Cebadas",ylab="Consumo",xlab =
"A?os")
adf.test(dos)
# segunda diferencia

tres=diff(ar.ts,differences=2)
plot(tres, main="Consumo de agua potable en la Parroquia Cebadas",ylab="Consumo",xlab =
"A?os")
adf.test(tres)
str(tres)
## la es estacionaria
## ya es modelo ARMA
##
par(mfrow=c(2,1),mar=c(4,4,4,1)+.1)
acf(tres, main="Consumo de agua potable en la Parroquia Cebadas")
#hay dos resagos pero sigue siendo estacionaria
pacf(tres, main="")
par(mfrow=c(2,1))

acf(ts(tres,frequency = 1), main="Consumo de agua potable en la Parroquia Cebadas")
pacf(ts(tres,frequency = 1), main="")

#####
auto.arima(ar.ts, trace = TRUE, test = "kpss", ic = "bic")
ajuste<- arima(ar.ts, order=c(0,1,0), seasonal=list(order=c(1,0,1)), include.mean=TRUE)
fit<-fitted(ajuste)#predichos del modelo arima
ten<-cbind(ar.ts,fit)#genera los datos predichos del ARIMA

ts.plot(ten,col=c(1,2),main="Consumo y Consumo estimado por ARIMA",ylab="Consumo en
metros^3",

```

```

      xlab="Años")#grafoca obs/pred
pron<-forecast(ajuste,h=36)
pron
plot(pron,ylab="Consumo en metros^3",xlab="Años")
summary(ajuste)
confint(ajuste)

#####
modelo1=arima(ar.ts,order = c(1,2,2))
modelo1
tsdiag(modelo1)
Box.test(residuals(modelo1),type ="Ljung-Box" )
error=residuals(modelo1)
plot(error)
pronostico<-forecast(modelo1,h=36)
pronostico
plot(pronostico)
pre<-pronostico
pre
ts.plot(pre,col=c(1,2),main="estimado por ARIMA",ylab="Consumo en m^3", xlab="Años")

## automatico
modelo<-auto.arima(ar.ts)
modelo
tsdiag(modelo)
Box.test(residuals(modelo),type ="Ljung-Box" )
error=residuals(modelo)
plot(error)
pronostico<-forecast(modelo,h=36)
pronostico
write.csv(pronostico,"pronostico.csv",sep = ",")

plot(pronostico)
pre<-pronostico
pre
ts.plot(pre,col=c(1,2),main="estimado por ARIMA",ylab="Consumo en m^3", xlab="Años")

```

```

summary(mdhwd)# medidas de precisión Holt Winters
confint(mdhwd)
summary(mdhwd)
band1<-expression("Demanda de Potencia Máxima","Estimado por Holt-Winters")#
obs/predichos demanda holt-winters
legend(2016.5,0.55,band1,lty = 1,col = c(1,2),cex=1)
plot(mdhwd, main="Consumo y Consumo pronosticada por el metodo de Holt-Winters")

fdt<-fitted(respd)#predichos del modelo arima

tdn<-cbind(ar.ts,fdt)#genera los datos predichos del ARIMA
ts.plot(tdn,col=c(1,2),main="Consumo y Consumo estimado por ARIMA",ylab="Consumo
m^3",
      xlab="Año")#grafica obs/pred
band<-expression("Demanda de Potencia Máxima","Estimado por ARMA")
legend(2016,0.84,band,lty = 1,col = c(1,2),cex=1)
summary(mdhwd)
summary(forecast(mdhwd,h = 38, level = 95))
mwhp<-forecast(mdhwd,h = 38, level = 95)

plot(forecast(mdhwd,h = 38, level = 95),ylim=c(0.2,1.2),ylim=c(0.4,1.0), main="Pronósticos de
la demanda de Potencia máxima en MW/h")

plot(mdhwd)
write.csv(mwhp,file = "pronosticos.csv",sep = ";")
# Normalidad
jarqueberaTest(Libro3$Consumo)

hist(Libro3$Consumo, freq = F, main="Residuales del modelo", ylim = c(0,8))
lines(density(Libro3$Consumo))
qqnorm(Libro3$Consumo)
qqline(Libro3$Consumo)

#estacionariedad
shapiro.test(Libro3$Consumo)

```

```
## Independencia:
```

```
# fas y fap de los residuos
```

```
par(mfrow=c(2,1))
```

```
acf(ajuste$resid, main="")
```

```
pacf(ajuste$resid, main="")
```

```
# Fas estimadas y Ljung-Box
```

```
tsdiag(respd)
```

```
Box.test(residuals(mdhwd),type ="Ljung-Box")
```

```
##### Costo DEL AGUA
```

```
Libro3 <- read_excel("Libro3.xlsx")
```

```
View(Libro3)
```

```
datos<-c(Libro3$Consumo,Libro3$Costos)
```

```
boxplot(datos)
```

```
summary(datos)
```

```
serie<-ts(Libro3$Costos,frequency =12)
```

```
ar.ts1=ts(Libro3$Costos,start = c(2017,1),frequency = 12)
```

```
ar.ts1
```

```
plot(ar.ts1,main="Costo del agua potable en la Parroquia Cebadas",ylab="Precio",xlab =  
"Años")
```

```
# logaritmo
```

```
serielog=log(ar.ts1)
```

```
serielog
```

```
plot(serielog)
```

```
adf.test(serielog)
```

```
#primera diferencia
```

```

dos=diff(ar.ts1)
dos
plot(dos, main="Costos del agua potable en la Parroquia Cebadas",ylab="Precio",xlab =
"Años")
adf.test(dos)
# segunda diferencia

tres1=diff(ar.ts1,differences=2)
plot(tres1, main="Costos del agua potable en la Parroquia Cebadas",ylab="Precio",xlab =
"Años")
adf.test(tres1)
str(tres1)
## la es estacionaria
## ya es modelo ARMA
##
par(mfrow=c(2,1),mar=c(4,4,4,1)+.1)
acf(tres1, main="Costos del agua potable en la Parroquia Cebadas")
#hay dos resagos pero sigue siendo estacionaria
pacf(tres1, main="")
par(mfrow=c(2,1))

acf(ts(tres1,frequency = 1), main="Costos del agua potable en la Parroquia Cebadas")
pacf(ts(tres1,frequency = 1), main="")

#####
auto.arima(ar.ts1, trace = TRUE, test = "kpss", ic = "bic")
ajuste1<- arima(ar.ts1, order=c(1,1,1), seasonal=list(order=c(0,0,1)), include.mean=TRUE)
fit1<-fitted(ajuste1)#predichos del modelo arima
ten1<-cbind(ar.ts1,fit1)#genera los datos predichos del ARIMA

ts.plot(ten1,col=c(1,2),main="Costo y Costo estimado por ARIMA",ylab="Costos del agua
potable",
xlab="Años")#grafoca obs/pred
lines(density(ten1),lwd = 12)
pron1<-forecast(ajuste1,h=36)
pron1
write.csv(pronostico1,"pronosticos1.csv",sep = ",")

```

```

plot(pron1)
summary(ajuste1)
confint(ajuste1)

#####
modelo2=arima(ar.ts1,order = c(1,2,2))
modelo2
tsdiag(modelo2)
Box.test(residuals(modelo2),type ="Ljung-Box" )
error=residuals(modelo2)

plot(error,main="Costos del agua potable en la Parroquia Cebadas",xlab="Años",,ylab="Error")
pronostico1<-forecast(modelo2,h=36)
pronostico1
plot(pronostico1)

pre1<-pronostico1
pre1
ts.plot(pre1,col=c(1,2),main="estimado por ARIMA",ylab="Costos del agua potable",
xlab="Años")

## automatico
modelo1<-auto.arima(ar.ts1)
modelo1
tsdiag(modelo1)
Box.test(residuals(modelo1),type ="Ljung-Box" )
error=residuals(modelo1)
plot(error)
pronostico1<-forecast(modelo1,h=36)
pronostico1
write.csv(pronostico1,"pronostico.csv",sep = ",")

plot(pronostico1)
pre1<-pronostico1
pre1
ts.plot(pre1,col=c(1,2),main="estimado por ARIMA",ylab="Costos", xlab="Años")

```



```
#### metodo holwinter
```

```
#modelo de predicción del Costos del agua potable en la Parroquia Cebadas
```

```
auto.arima(ar.ts, trace = TRUE, test = "kpss", ic = "bic")
```

```
respd<- arima(ar.ts, order=c(1,0,1), include.mean=TRUE)
```

```
respd
```

```
confint(respd)
```

```
summary(respd)#medidas de precisión ARIMA
```

```
mdhwd<- HoltWinters(ar.ts,seasonal = "additive")
```

```
summary(mdhwd)# medidas de precisión Holt Winters
```

```
confint(mdhwd)
```

```
summary(mdhwd)
```

```
plot(mdhwd)
```

```
band1<-expression("Demanda de Potencia Máxima","Estimado por Holt-Winters")#
```

```
obs/predichos demanda holt-winters
```

```
legend(2016.5,0.55,band1,lty = 1,col = c(1,2),cex=1)
```

```
plot(mdhwd, main="Costo y Costo estimado pronosticada por el metodo de Holt-Winters",  
ylab="Precio",xlab="Años")
```

```
fdt<-fitted(respd)#predichos del modelo arima
```

```
tdn<-cbind(ar.ts,fdt)#genera los datos predichos del ARIMA
```

```
ts.plot(tdn,col=c(1,2),main="Costo y Costo estimado por ARIMA",ylab="Precio",  
xlab="Año")#grafica obs/pred
```

```
band<-expression("Demanda de Potencia Máxima","Estimado por ARMA")
```

```
legend(2016,0.84,band,lty = 1,col = c(1,2),cex=1)
```

```
summary(mdhwd)
```

```
summary(forecast(mdhwd,h = 38, level = 95))
```

```
mwhp<-forecast(mdhwd,h = 38, level = 95)
```

```
plot(forecast(mdhwd),ylab="Costo",xlab="Año")
```

```
plot(mdhwd)
```

```
write.csv(mdhwd,file = "pronosticos1.csv",sep = ";")
```

```
# Normalidad
jarqueberaTest(Libro3$Costos)

hist(Libro3$Costos, freq = F, main="Residuales del modelo", ylim = c(0,8))
lines(density(Libro3$Costos))
qqnorm(Libro3$Costos)
qqline(Libro3$Costos)

#estacionariedad
shapiro.test(Libro3$Costos)

plot(respd, main="Test de raíz unitaria">#estacionariedad raiz unitaria
adf.test(Libro3$Costos, alternative ="stationary")#test de estacionariedad

## Independencia:

# fas y fap de los residuos

par(mfrow=c(2,1))
acf(ajuste$resid, main="")
pacf(ajuste$resid, main="")

# Fas estimadas y Ljung-Box
tsdiag(respd)
Box.test(residuals(mdhwd),type ="Ljung-Box")
```

**ANEXO C: ENCUESTA SOBRE LA CALIDAD DE AGUA POTABLE**

**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**CARRERA DE ESTADISTICA**

**Encuesta de opinión sobre la calidad del servicio de Agua Potable**

**Encuesta N°.....**

Objetivo. - Analizar la calidad del servicio de agua potable de los usuarios de la Junta Administradora de Agua Potable de la Parroquia de Cebadas.

Su opinión es importante, conteste las siguientes

1. Sexo
  - a) Femenino
  - b) Masculino
  
2. Edad  
.....
  
3. Estado Civil
  - a) Soltero(a)
  - b) Separado(a)
  - c) Casado(a)
  - d) Viudo(a)
  
4. Ocupación económica
  - a) Agricultor
  - b) Profesional
  - c) Empresario
  - d) Comerciante
  - e) Otro:.....
  
5. ¿Cuál es el ingreso mensual familiar estimado?  
..... dólares
  
6. ¿Cómo calificaría el servicio de agua potable?
  - a) Bueno
  - b) Regular
  - c) Malo
  
7. ¿Cuántas personas viven en su hogar?  
.....
  
8. ¿Cuántas horas del día tiene el servicio de agua potable?  
.....
  
9. ¿Cómo calificarías la presión del agua que llega a su hogar?
  - a. Bueno
  - b. Regular
  - c. Malo
  
10. ¿El agua que llega a su hogar presenta un olor desagradable?

- a. Si
- b. No

11. ¿El agua que llega a su hogar presenta un buen sabor?

- a. Si
- b. No

12. ¿El agua que llega a su hogar presenta turbidez (Que está alterado por una cosa que le quita la transparencia o la claridad)?

- a. Si
- b. No

13. En caso de interrupción ¿Cuántos días tardan en reparar el servicio?

..... días.

## ANEXO D: BASE DE DATOS TABULADA

base\_tabulada.xlsx - Excel

ARCHIVO INICIO INSERTAR DISEÑO DE PÁGINA FÓRMULAS DATOS REVISAR VISTA

Calibri 11 Fuente Ajustar texto General \$ - % 000 00 00 Formato condicional Dar formato como tabla Estilos de celda Insertar Eliminar Formato Celdas Autosuma Rellenar Borrar Ordenar y filtrar Buscar y seleccionar

Portapapeles Fuente Alineación Número Estilos Celdas Modificar

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	Numero	sexo	edad	Estado_Civil	ocupacion_economico	ingreso	calificacion	num_personas	horas_dia_d	calificacion	OLOR_AGUA	sabor_agua	transparenci	reponer_agua			
2	1	Femenino	54	Separado(a)	Comerciante	394	Bueno	7	12	Bueno	No	Si	No	1			
3	2	Femenino	59	Casado(a)	Agricultor	394	Bueno	3	12	Regular	Si	Si	No	1			
4	3	Masculino	65	Casado(a)	Agricultor	100	Regular	5	12	Bueno	No	Si	No	1			
5	4	Masculino	54	Casado(a)	Agricultor	394	Regular	2	9	Bueno	No	Si	No	1			
6	5	Femenino	55	Casado(a)	Comerciante	500	Bueno	2	10	Bueno	No	Si	No	1			
7	6	Masculino	59	Casado(a)	Agricultor	394	Bueno	7	9	Bueno	No	Si	No	1			
8	7	Masculino	72	Casado(a)	Agricultor	100	Regular	7	12	Bueno	No	Si	No	1			
9	8	Masculino	39	Casado(a)	Agricultor	394	Regular	4	12	Bueno	No	Si	No	1			
10	9	Masculino	68	Casado(a)	Agricultor	70	Regular	5	12	Bueno	No	Si	No	1			
11	10	Femenino	47	Soltero(a)	Agricultor	394	Bueno	4	10	Bueno	No	Si	No	3			
12	11	Masculino	47	Casado(a)	Profesional	700	Bueno	2	12	Bueno	No	Si	No	1			
13	12	Masculino	53	Casado(a)	Agricultor	394	Bueno	4	12	Bueno	No	Si	No	1			
14	13	Masculino	65	Casado(a)	Agricultor	500	Bueno	2	12	Bueno	Si	Si	No	1			
15	14	Masculino	57	Casado(a)	Agricultor	120	Bueno	7	12	Bueno	No	Si	No	1			
16	15	Femenino	54	Casado(a)	Agricultor	394	Bueno	2	12	Bueno	No	Si	No	1			
17	16	Femenino	30	Casado(a)	Agricultor	394	Bueno	2	10	Bueno	No	Si	No	1			
18	17	Masculino	44	Casado(a)	Agricultor	394	Bueno	4	12	Bueno	No	Si	No	1			
19	18	Femenino	46	Casado(a)	Empresario	1200	Bueno	5	12	Bueno	No	Si	No	1			
20	19	Masculino	34	Soltero(a)	Agricultor	394	Regular	3	12	Bueno	No	Si	No	2			
21	20	Femenino	59	Casado(a)	Agricultor	394	Bueno	3	12	Bueno	No	Si	No	1			
22	21	Masculino	50	Casado(a)	Agricultor	394	Regular	2	12	Bueno	No	Si	No	1			
23	22	Masculino	56	Casado(a)	Agricultor	394	Bueno	7	12	Bueno	No	Si	No	1			
24	23	Femenino	30	Casado(a)	Agricultor	394	Bueno	2	12	Bueno	No	Si	No	1			
25	24	Femenino	54	Casado(a)	Agricultor	394	Malo	7	12	Bueno	No	Si	No	1			
26	25	Femenino	64	Casado(a)	Agricultor	394	Regular	5	12	Bueno	No	Si	No	1			
27	26	Masculino	49	Separado(a)	Agricultor	394	Bueno	6	12	Regular	No	Si	No	1			

Hoja1

## ANEXO E: EVIDENCIAS FOTOGRÁFICAS

### REUNION CON EL SR. ALCANDE DELFIN QUISPE



### USUARIOS DEL AGUA POTABLE



## POSICIONAMIENTO DE NUEVOS DIRECTIVOS



## VISITA A TANQUES DE TRATAMIENTO A NIVEL DE LA PARROQUIA DE CEBADAS



DIRECTIVOS DE LA JUNTA ADMINISTRADORA DE AGUA POTABLE Y SANEAMIENTO



TANQUES DE TRATAMIENTO DE AGUAS SERVIDAS







**UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y DOCUMENTAL**

**REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA**

**Fecha de entrega:** 14 / 02 / 2023

<b>INFORMACIÓN DEL AUTOR/A (S)</b>
<b>Nombres – Apellidos:</b> PEÑAFIEL MALAN GISSELA MARISOL
<b>INFORMACIÓN INSTITUCIONAL</b>
<b>Facultad:</b> CIENCIAS
<b>Carrera:</b> ESTADÍSTICA
<b>Título a optar:</b> INGENIERA EN ESTADÍSTICA INFORMÁTICA
<b>f. Analista de Biblioteca responsable:</b> Lcdo. Holger Ramos, MSc.

0159-DBRA-UPT-2023

