

Bianca Estefanía Kennedy Giménez

*Estudio y análisis de carga de  
trabajo en el Centro Coordinador de  
Emergencias y Seguridad CECOES  
1-1-2*

Study and workload analysis at the Emergency  
and Security Coordination Center (CECOES  
1-1-2)

Trabajo Fin de Grado  
Grado en Matemáticas  
La Laguna, Julio de 2023

DIRIGIDO POR  
*Carlos Pérez González*

*Carlos Pérez González*

*Departamento de Matemáticas,  
Estadística e Investigación  
Operativa*

*Universidad de La Laguna  
38200 La Laguna, Tenerife*

---

## Agradecimientos

Gracias a mi familia, por apoyarme incondicionalmente en la distancia.  
Gracias a mis amigos, por confiar en mí y ser siempre una guía.  
Gracias a mi tutor, Carlos, por su paciencia y comprensión.  
Gracias a Emma, por todo.

Bianca Estefanía Kennedy Giménez  
La Laguna, 10 de julio de 2023



---

## Resumen · Abstract

### *Resumen*

---

*Este estudio analiza la carga de trabajo en el Centro Coordinador de Emergencias y Seguridad (CECOES) 1-1-2 del Gobierno de Canarias. El CECOES es un servicio público que responde a todas las llamadas de emergencia que se producen en el archipiélago canario, brindando asistencia inmediata y coordinando a los organismos de seguridad y emergencias. El objetivo de este estudio es analizar la eficiencia en la gestión de los servicios de emergencia proporcionados por el CECOES a través de indicadores cuantitativos de carga de trabajo. Se utilizó la estadística descriptiva, utilizando datos proporcionados por el CECOES, como el número de llamadas recibidas por cada operador y duración de las llamadas; también se usó modelos de análisis de series temporales ARIMA. Se evidencian patrones temporales en la carga de trabajo, con aumentos significativos en marzo y agosto, posiblemente relacionados con eventos como el estado de alarma y la incidencia de COVID-19.*

**Palabras clave:** *Carga de trabajo – Centro Coordinador de Emergencias y Seguridad (CECOES) – Estadística descriptiva – ARIMA – COVID*

## ***Abstract***

---

*This study examines the workload at the Emergency and Security Coordination Center (CECOES) 1-1-2 of the Canary Islands Government. CECOES is a public service that responds to all emergency calls occurring in the Canary archipelago, providing immediate assistance and coordinating with security and emergency agencies. The objective of this study is to analyze the efficiency in managing emergency services provided by CECOES using quantitative indicators of workload. Descriptive statistics were used, utilizing data provided by CECOES, such as the number of calls received by each operator and the duration of the calls; ARIMA time series analysis models were also used. Temporal patterns in the workload are evident, with significant increases in March and August, possibly related to events like the state of alarm and the incidence of COVID-19.*

**Keywords:** *Workload – Centro Coordinador de Emergencias y Seguridad (CECOES) – Descriptive statistics – ARIMA – COVID*

---

# Contenido

|   |     |
|---|-----|
| <b>Agradecimientos</b> .....                                | III |
| <b>Resumen/Abstract</b> .....                               | V   |
| <b>Introducción</b> .....                                   | IX  |
| <b>1. Esquema de trabajo aplicado en el estudio</b> .....   | 1   |
| 1.1. Funcionamiento del CECOES .....                        | 1   |
| 1.2. Herramientas y aplicaciones informáticas .....         | 3   |
| 1.3. Procedimientos y técnicas estadísticas .....           | 3   |
| 1.3.1. Procesamiento de datos .....                         | 4   |
| <b>2. Análisis de datos</b> .....                           | 7   |
| 2.1. Preprocesamiento de datos .....                        | 7   |
| 2.1.1. Extracción de datos .....                            | 7   |
| 2.1.2. Transformación de datos .....                        | 8   |
| 2.1.3. Carga de datos .....                                 | 9   |
| 2.2. Análisis de datos .....                                | 10  |
| 2.2.1. Análisis de datos relacionados con el COVID-19 ..... | 18  |
| 2.3. Visualización de datos .....                           | 21  |
| <b>3. Predicción de modelos de series temporales</b> .....  | 25  |
| 3.1. Series temporales y modelo ARIMA .....                 | 25  |
| 3.2. Aplicación método ARIMA .....                          | 29  |
| 3.2.1. Pruebas diagnósticas para modelos ARIMA .....        | 32  |
| 3.2.2. Predicción modelo ARIMA(3,1,0) .....                 | 35  |
| <b>A. Apéndice</b> .....                                    | 43  |
| A.1. Código en el lenguaje de programación R .....          | 43  |
| <b>Bibliografía</b> .....                                   | 45  |

**Poster** ..... 47



---

## Introducción

El objetivo principal de este Trabajo de Tin de Grado es realizar un estudio de las cargas de trabajo en el Centro Coordinador de Emergencias y Seguridad (CECOES) 1-1-2 del Gobierno de Canarias, con el propósito de analizar y mejorar la eficiencia en la gestión de los servicios de emergencias.

El CECOES juega un papel fundamental en la atención de emergencias en las islas, ya que proporciona asistencia inmediata y coordina la respuesta de los diferentes organismos de seguridad y emergencias. Por lo tanto, es de gran importancia conocer en detalle la carga de trabajo de los trabajadores del CECOES para identificar posibles problemas y mejorar la calidad de su servicio.

La Figura 1 muestra una representación gráfica de los factores de la carga de trabajo a los que están expuestas los trabajadores de CECOES, que se relaciona con subfactores como exigencias, estrés, dificultad del trabajo, horas de trabajo, malas posturas corporales y volumen de trabajo, los cuales pueden ser explorados de forma relativamente independiente.



**Figura 1.** Esquema factores que afectan a las cargas de trabajo

Además, cabe destacar que el servicio de emergencia se ha visto implicado, de una forma directa, en la gestión de la pandemia por la COVID-19 que ha azotado al planeta entero desde principios del año 2020. El CECOES se ha ocupado de la gestión de las llamadas telefónicas en las islas relacionadas con las consultas de los ciudadanos sobre el coronavirus. Por tanto, uno de los objetivos principales de este trabajo consistirá en analizar el efecto de la pandemia sobre las cargas de trabajo del CECOES.

Ante esta situación, el CECOES del Gobierno de Canarias propone que se realice un estudio de cargas de trabajo del servicio, a través de un TFG. Para sentar las bases y los objetivos, así como obtener los datos para este estudio, hemos trabajado con el responsable del Departamento de Tecnología de la Información.

Entre los objetivos finales para llevar a cabo este estudio podemos mencionar los siguientes:

- Mejorar la gestión eficiente del servicio. Comprender en detalle la carga de trabajo de los trabajadores de CECOES, planificación de turnos, de vacaciones, bajas laborales, etc. se podrán implementar medidas para optimizar los recursos, reducir tiempos de respuesta y mejorar la eficiencia general
- Identificar posibles problemas: Al analizar las cargas de trabajo, se podrán identificar posibles problemas o puntos de estrés entre los trabajadores encargados del sector de comunicaciones en el 1-1-2.. Esto incluye la detección de sobrecargas, desequilibrios en la distribución de tareas, deficiencias en la coordinación de los diferentes organismos de seguridad y emergencias, entre otros aspectos.
- Optimizar la asignación de recursos: El estudio de las cargas de trabajo ayudará a determinar de manera más precisa la cantidad y distribución de personal necesario en el CECOES. Esto permitirá una asignación óptima de recursos humanos y técnicos, evitando tanto la falta de personal como la subutilización de recursos.
- Garantizar la calidad del servicio: El CECOES desempeña un papel crucial en la atención de emergencias en las Islas Canarias. El estudio ayudará a establecer parámetros y estándares de trabajo que contribuirán a una atención de emergencias más efectiva

Como un primer paso para lograr estos objetivos, el trabajo constará de los siguientes capítulos:

En el primer capítulo, se presenta una explicación detallada del funcionamiento del CECOES, así como de las herramientas y aplicaciones informáticas utilizadas en el estudio. Además, se describen los procedimientos y técnicas estadísticas empleadas para llevar a cabo el análisis de datos.

El segundo capítulo, aplicaremos los denominados procesos ETL: extracción, transformación y carga de los datos. Se exploran características clave de los datos y se realiza un resumen detallado de las principales tendencias y patrones identificados.

El tercer capítulo, se dedica al estudio de predicción de datos mediante el uso de modelos de series temporales. Se aplican técnicas y métodos estadísticos avanzados para desarrollar modelos predictivos que permitan estimar valores futuros basados en patrones y tendencias históricas.

Por último, se presentan las conclusiones obtenidas a partir de los estudios estadísticos realizados. Se resumen los resultados significativos, se destacan las principales contribuciones y se discuten las implicaciones de los hallazgos en el contexto del análisis de datos y la toma de decisiones.

En resumen, se llevará a cabo un análisis estadístico de las cargas de trabajo durante diferentes turnos y días. El análisis proporcionará información relevante sobre el rendimiento de los trabajadores y ayudará a detectar posibles causas de fatiga y errores en la atención de emergencias. Al conocer los patrones de trabajo y los posibles problemas, se podrán implementar estrategias para mejorar la eficiencia y la calidad del servicio de atención de emergencias en el CECOES.

Nos centraremos en presentar resultados y estadísticos sobre las cargas de trabajo que sean fáciles de interpretar. En consecuencia, la información será expuesta en forma tabular y se representarán gráficos de resumen (barras, líneas, ...). Además, se reunirán o agruparán para formar cuadros de mando. El objetivo final es que la entidad que busca hacer un estudio de esta índole sea capaz de interpretar los resultados de la carga de trabajo para mejorar su funcionalidad.



## Esquema de trabajo aplicado en el estudio

Este capítulo introductorio consta de tres secciones, en la primera explicaremos el funcionamiento del CECOES; en la segunda, introduciremos las herramientas y aplicaciones utilizadas; por último, comentaremos brevemente los procedimientos ETL y la función *auto.arima()* que aplicaremos posteriormente.

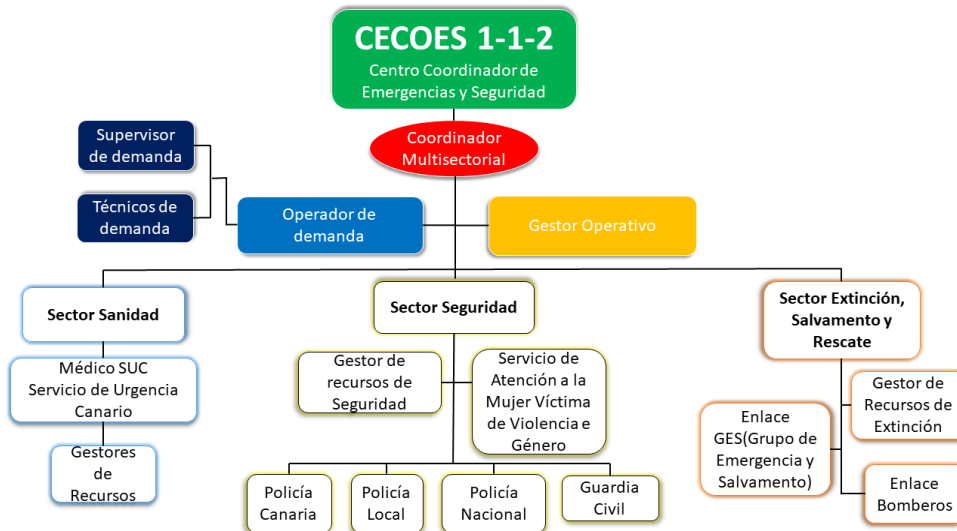
### 1.1. Funcionamiento del CECOES

A continuación, realizamos una breve descripción de este centro para poder comprender de dónde se extraen los datos en los que estamos interesados para la realización de este proyecto. El Centro Coordinador de Emergencias y Seguridad (CECOES) 1-1-2 del Gobierno de Canarias se encarga de proporcionar los servicios de emergencia necesarios para dar una respuesta eficaz a los incidentes que suceden a diario en todo el Archipiélago tanto en tierra como en aguas circundantes. Se trata de un servicio que se encuentra disponible las 24 horas del día durante los 365 del año, es accesible desde cualquier punto de la red telefónica terrestre y móvil sin necesidad de marcar prefijo. Para cubrir una media de 5000 llamadas diarias cuentan con un equipo de casi 200 profesionales distribuidos en sectores (Operación de demandas, Gestión de incidentes, Gestión de seguridad, Extinción y salvamento, Bomberos y Sanidad) supervisados por un Coordinador Multisectorial, repartidos entre las dos oficinas de las dos provincias, Santa Cruz de Tenerife y Las Palmas de Gran Canaria. Además, tienen numerosos recursos a su disposición para poder cubrir los incidentes: patrullas de policía, parques de bomberos, ambulancias, avión sanitario, helicópteros, etc.

En el esquema [1.1](#) podemos observar la organización en diferentes sectores de las oficinas del CECOES.

Veamos cómo funciona cada sector con más profundidad.

- El sector de Operación de demandas se encuentra en una sala aislada, formada por un supervisor y varios operarios que recogen las llamadas y las derivan al sector correspondiente dependiendo del tipo de incidente. Los incidentes



**Figura 1.1.** Esquema distribución CECOES

son clasificados por los operarios de esta sala primando la localización (isla, municipio, calle) para poder enviar el recurso en caso de que sea necesario y el tipo de incidente (médico, policial, incendio, etc.). Si están implicados varios sectores se le pasa al más importante, una vez cubierta su parte y solicitado el recurso, este se encargará de comunicarse con los otros sectores para que el incidente quede cubierto y resuelto en su totalidad. Los datos que se recogen en esta parte del proceso son de suma importancia para la eficacia y optimización de los diferentes recursos que se puedan precisar.

- El Coordinador/a Multisectorial es el máximo responsable en cuanto a la organización y cumplimiento de las normas de la Sala Operativa, que se encarga de la gestión de los incidentes y está formada por el resto de sectores. Su labor consiste en asegurarse de que los recursos cumplan sus cometidos y corregir los errores, asumir los roles de los técnicos que no se encuentren disponibles en la sala, atender los correos y los incidentes que se generan vía SMS y dar los avisos a la población de los organismos de protección civil. Por lo tanto, se ocupa de varias vías de comunicación, además de las llamadas que tenga que gestionar.
- El Gestor Operativo se encarga de gestionar los aspectos técnicos (informática y electrónica) de la sala.

## 1.2. Herramientas y aplicaciones informáticas

A lo largo del trabajo hemos ido utilizando diferentes herramientas para el análisis de nuestros datos y para la visualización de esto, tanto en bruto como para poder interpretar los resultados obtenidos. Las herramientas empleadas y sus funciones son las siguientes:

- Microsoft Excel [4]: Microsoft Excel (Figura 1.2) es el programa de software de hojas de cálculo líder en el sector y una herramienta avanzada de análisis y visualización de datos. En este lenguaje el CECOES nos ha dado los datos que teníamos que estudiar.



**Figura 1.2.** Logo de Microsoft Excel



**Figura 1.3.** Logo Rstudio

- RStudio [7]: RStudio (Figura 1.3) es un entorno de desarrollo integrado para el lenguaje de programación R, se trata de un software gratuito dedicado a la computación estadística y gráficos. En este programa hemos realizado el análisis de nuestros datos.
- RStudio Cloud [3]: RStudio Cloud (Figura 1.4) es la versión en la nube del software gratuito RStudio. Esta herramienta nos ha ayudado para una mejor comunicación tutor-alumna, además de evitar las posibles pérdidas de datos.
- Microsoft PowerBI [5]: Power BI (Figura 1.5) es una herramienta de visualización de datos, que nos permite compartir paneles para que los interesados del CECOES puedan ver los resultados de nuestro análisis. También es muy útil para visualizar grandes cantidades de datos gracias a los filtros que una vez establecidos podrán modificarse sin afectar a la programación que hace el analista.

## 1.3. Procedimientos y técnicas estadísticas

El análisis de datos tiene como finalidad obtener información que ayude a la mejora de la empresa, en nuestro caso, el CECOES. Para poder alcanzar este fin,

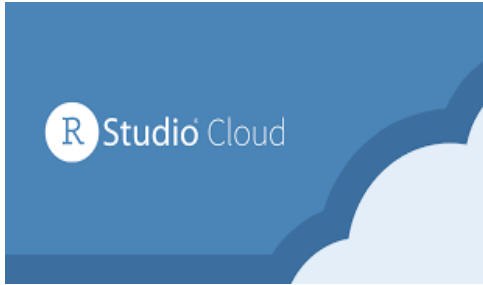


Figura 1.4. Logo de Rstudio Cloud



Figura 1.5. Logo Rstudio

hemos estudiado sus datos para saber el tiempo que emplean sus trabajadores en los diferentes sectores a lo largo del año 2020.

### 1.3.1. Procesamiento de datos

En el segundo capítulo, para realizar el análisis de los datos hemos empleado los procesos de carga de datos, ampliamente conocidos como procesos ETL y cuyo nombre viene de las siglas en inglés "Extract, Transform, Load". Para más información sobre estos procesos, se recomienda consultar, por ejemplo, las referencias [1] [6].

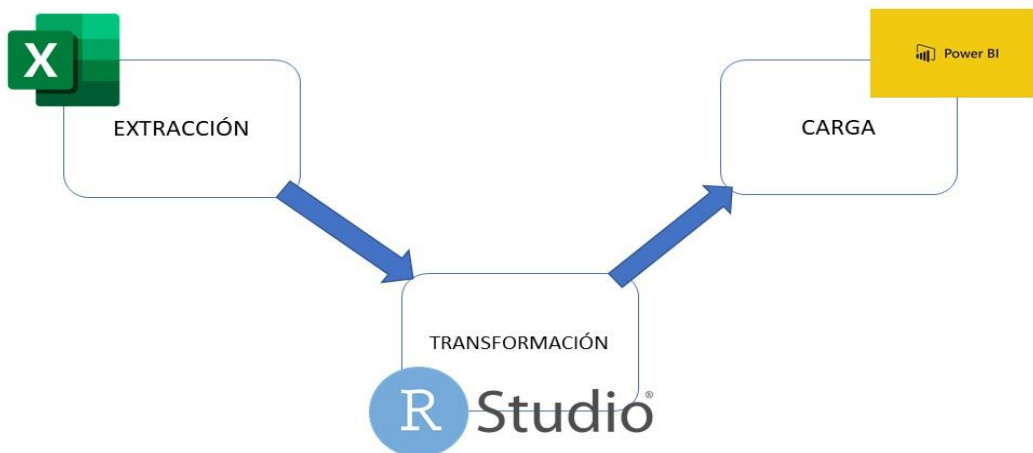


Figura 1.6. Esquema de procesos ETL

Veamos en qué consiste cada parte y qué herramientas, de las vistas en la sección anterior usaremos más adelante:



- **Extract:** la función de extracción lee datos de una base de datos fuente específica y extrae un subconjunto de datos deseado. Esto fue llevado a cabo con el software Excel.
- **Transform:** la función de transformación de los datos para convertirlos al estado deseado. Para esta sección, utilizaremos el software RStudio.
- **Load:** la función de carga se utiliza para escribir los datos resultantes (ya sea todo el subconjunto o solo los cambios) en una base de datos de destino, que puede o no existir previamente. El software que hemos utilizado es Power BI.

En el tercer capítulo, se realiza una predicción de los datos basados en el modelo de predicción de series temporales ARIMA. Para realizar este estudio, se utiliza la función *auto.arima()*, disponible en el entorno de desarrollo RStudio.

La función *auto.arima()* en R es una herramienta automatizada utilizada para ajustar modelos ARIMA a series temporales. Su objetivo principal es encontrar el modelo ARIMA óptimo para una serie temporal específica. Para lograr esto, *auto.arima()* realiza una búsqueda exhaustiva de diversas combinaciones de parámetros ARIMA, como el orden de autorregresión, el orden de integración y el orden de medias móviles.

Durante la búsqueda, la función *auto.arima()* evalúa y compara los modelos candidatos utilizando el criterio de información de Akaike (AIC) como métrica de evaluación. El AIC es un criterio que mide la calidad del ajuste de un modelo, teniendo en cuenta la complejidad del mismo. Penaliza los modelos más complejos y favorece aquellos que logran un buen ajuste con menos parámetros.

El método FPP2 es la base teórica y práctica en la que se fundamenta *auto.arima()*. Este método se centra en los principios y prácticas de la predicción de series temporales. Utilizando el enfoque FPP2, *auto.arima()* implementa la búsqueda exhaustiva y el cálculo del AIC para seleccionar el modelo ARIMA óptimo.

En resumen, *auto.arima()* ofrece una forma automatizada y eficiente de seleccionar el modelo ARIMA más adecuado para una serie temporal. Utiliza el AIC y un enfoque de búsqueda exhaustiva basado en el método FPP2 para encontrar la combinación óptima de parámetros. El modelo con el valor de AIC más bajo se considera el óptimo y se devuelve como resultado.

En los próximos capítulos, comentaremos en detalle cómo hemos aplicado este tipo de procedimientos en nuestro análisis.



# Análisis de datos

En este capítulo, analizaremos los datos proporcionados por el CECOES mediante los procesos ETL.

## 2.1. Preprocesamiento de datos

### 2.1.1. Extracción de datos

La extracción de datos se realiza mediante la entrega de tablas resumen de Excel por parte del CECOES, que es la forma en la que ellos tratan la información. Estos datos abarcan diferentes sectores y se presentan en diferentes franjas de tiempo, con el propósito de que podamos analizar la carga de trabajo de sus trabajadores.

Cabe destacar que todos los datos han sido cifrados y no contienen nombres ni ningún dato identificativo. Podemos distinguir los datos garantizando el anonimato de los trabajadores gracias a siglas que se refieren a un trabajador en concreto o gracias al número que se les asigna a las llamadas.

A continuación, se detalla el tipo de datos facilitados por el CECOES, su formato y los objetivos que se plantean obtener en este estudio.

Los datos están almacenados en un archivo Excel donde cada día del mes se representa en una hoja separada. El archivo contiene información detallada sobre las llamadas, incluyendo la marca de tiempo de cada llamada, las siglas identificativas del operador que las atiende, el número de veces que se contacta a un trabajador para un mismo incidente y la duración total de esas llamadas.

En la figura [2.1](#) podemos observar los datos de tiempos totales de los enfermeros el día 4 de enero de 2020.

De estos datos tenemos como objetivo obtener:

- El número de llamadas que cada operador obtiene a lo largo del tiempo.

- La duración de las llamadas para poder comparar el desempeño del trabajador en la empresa durante este mes.
- Las diferencias en la carga de trabajo de los trabajadores.

|    | A   | B              | C               | D            | E        | F        | G        |  |
|----|---|----------------|-----------------|--------------|----------|----------|----------|--|
| 1  | Tiempos Totales por operadores y franjas horarias |                |                 |              |          |          |          |  |
| 2  | Servicio:   | 'Enfermeros'   |                 |              |          |          |          |  |
| 3  | Fecha inicio:                                     | 04/01/2020     | Fecha fin:      | 04/01/2020   |          |          |          |  |
| 4  |   |                |                 |              |          |          |          |  |
| 5  | Fecha y hora                                      | Login          | Total contactos | Tiempo total | Locución | ACW      | BCW      |  |
| 6  |   |                |                 |              |          |          |          |  |
| 7  | 04/01/2020 00:00                                  | EAPC           | 1               | 00:00:32     | 00:00:32 | 00:00:00 | 00:00:00 |  |
| 8  | 04/01/2020 02:00                                  | EAPC           | 2               | 00:02:35     | 00:02:35 | 00:00:00 | 00:00:00 |  |
| 9  | 04/01/2020 03:00                                  | EAPC           | 1               | 00:00:23     | 00:00:23 | 00:00:00 | 00:00:00 |  |
| 10 | 04/01/2020 05:00                                  | EAPC           | 1               | 00:00:57     | 00:00:56 | 00:00:01 | 00:00:00 |  |
| 11 | 04/01/2020 06:00                                  | EAPC           | 1               | 00:00:34     | 00:00:34 | 00:00:00 | 00:00:00 |  |
| 12 | 04/01/2020 07:00                                  | EAPC           | 1               | 00:00:04     | 00:00:02 | 00:00:02 | 00:00:00 |  |
| 13 |   | TOTALES DEEAPC | 7               | 00:05:05     | 00:05:02 | 00:00:03 | 00:00:00 |  |
| 14 |   |                |                 |              |          |          |          |  |

Figura 2.1. Datos en Excel proporcionados por el CECOES

Además, analizar estos datos permite ver cómo ha afectado la pandemia actual de la COVID-19 al servicio de emergencia canario, ya que ha supuesto que el año que vamos a estudiar sea especialmente interesante por las diferentes etapas que ha sufrido de normalidad, confinamiento y nueva normalidad.

Para la extracción de los datos aplicamos un método de apilación de datos de Excel que nos permitió un manejo de los datos más eficiente en RStudio. Por ello, cargamos cada hoja del libro de excel con el siguiente código:

```
load("data112_cargas_trabajo.RData")
```

### 2.1.2. Transformación de datos

La siguiente etapa del proceso es la transformación de los datos. Éstos son extraídos del servidor de origen, carecen de formato y no se pueden utilizar en su forma original. Los datos requieren un preprocesado en varias fases; limpiados, mapeados y transformados. Para ello hemos utilizado la herramienta RStudio para trabajar con el lenguaje de programación R. Los paquetes principales que hemos utilizado son los observados en la figura 2.2.



Figura 2.2. Paquetes utilizados de RStudio.

En esta fase, hemos añadido algunas variables para que a la hora de analizar los datos el trabajo fuera más sencillo, como desglosar la fecha en año, mes y días. Esto se ha hecho con el siguiente código:

```
data112$fecha<-strptime(data112$fecha,format="%Y/%m/%d %H:%M")
data112$dia<-day(data112$fecha)
data112$mes <-month(data112$fecha,label=TRUE,abbr=FALSE)
data112$anio<-year(data112$fecha)
```

De esta forma, se añaden las columnas al final de nuestra tabla inicial, representada en la figura 2.3:

|   | prov | servicio | fecha               | login | contactos | tiempo.total | locucion | ACW      | BCW      | duracion.total | locucion.duracion | ACW.duracion | BCW.duracion | dia | anio   |
|---|------|----------|---------------------|-------|-----------|--------------|----------|----------|----------|----------------|-------------------|--------------|--------------|-----|--------|
| 1 | LPA  | DEMANDA  | 2020-01-01 22:00:00 | TAAL  | 1         | 00:00:31     | 00:00:31 | 00:00:00 | 00:00:00 | 0.5166667      | 0.5166667         | 0.0000000    | 0.0000000    | 0   | 1 2020 |
| 2 | LPA  | DEMANDA  | 2020-01-01 23:00:00 | TAAL  | 26        | 00:19:42     | 00:19:09 | 00:00:33 | 00:00:00 | 19.7000000     | 19.1500000        | 0.5500000    | 0.0000000    | 0   | 1 2020 |
| 3 | LPA  | DEMANDA  | 2020-01-01 06:00:00 | TAEP  | 1         | 00:00:09     | 00:00:09 | 00:00:00 | 00:00:00 | 0.1500000      | 0.1500000         | 0.0000000    | 0.0000000    | 0   | 1 2020 |
| 4 | LPA  | DEMANDA  | 2020-01-01 07:00:00 | TAEP  | 23        | 00:18:44     | 00:17:27 | 00:01:17 | 00:00:00 | 18.7333333     | 17.4500000        | 1.2833333    | 0.0000000    | 0   | 1 2020 |
| 5 | LPA  | DEMANDA  | 2020-01-01 08:00:00 | TAEP  | 15        | 00:20:43     | 00:19:41 | 00:01:02 | 00:00:00 | 20.7166667     | 19.6833333        | 1.0333333    | 0.0000000    | 0   | 1 2020 |
| 6 | LPA  | DEMANDA  | 2020-01-01 09:00:00 | TAEP  | 17        | 00:16:39     | 00:15:06 | 00:01:33 | 00:00:00 | 16.6500000     | 15.1000000        | 1.5500000    | 0.0000000    | 0   | 1 2020 |
| 7 | LPA  | DEMANDA  | 2020-01-01 10:00:00 | TAEP  | 21        | 00:21:44     | 00:21:25 | 00:00:19 | 00:00:00 | 21.7333333     | 21.4166667        | 0.3166667    | 0.0000000    | 0   | 1 2020 |
| 8 | LPA  | DEMANDA  | 2020-01-01 11:00:00 | TAEP  | 20        | 00:18:23     | 00:17:47 | 00:00:36 | 00:00:00 | 18.3833333     | 17.7833333        | 0.6000000    | 0.0000000    | 0   | 1 2020 |

Figura 2.3. Imagen de las columnas de la tabla final data112 de la fila 1 a la 8.

### 2.1.3. Carga de datos

El dataset que se importa en Power BI contiene información relevante para el análisis y visualización de los datos. Este dataset incluye una tabla con varias columnas que representan diferentes variables de interés. La columna "fecha" registra la fecha y hora en la que se produjo cada evento. Además, se han añadido variables adicionales mediante el desglose de la fecha en columnas como "año", "mes" y "día". Estas variables desglosadas facilitan el análisis temporal de los datos.

Adicionalmente, el dataset incluye otras columnas que proporcionan información sobre el evento en cuestión, como la hora de la llamada, las siglas del operador que la recibe y el número de veces que se contacta a un trabajador para el mismo incidente. También, se registra el tiempo total que duran esas llamadas.

Toda esta información se encuentra estructurada en filas y columnas, permitiendo su importación y utilización en Power BI para realizar análisis de datos, crear visualizaciones interactivas y obtener información significativa sobre la carga de trabajo y la gestión de los servicios de emergencias.

## 2.2. Análisis de datos

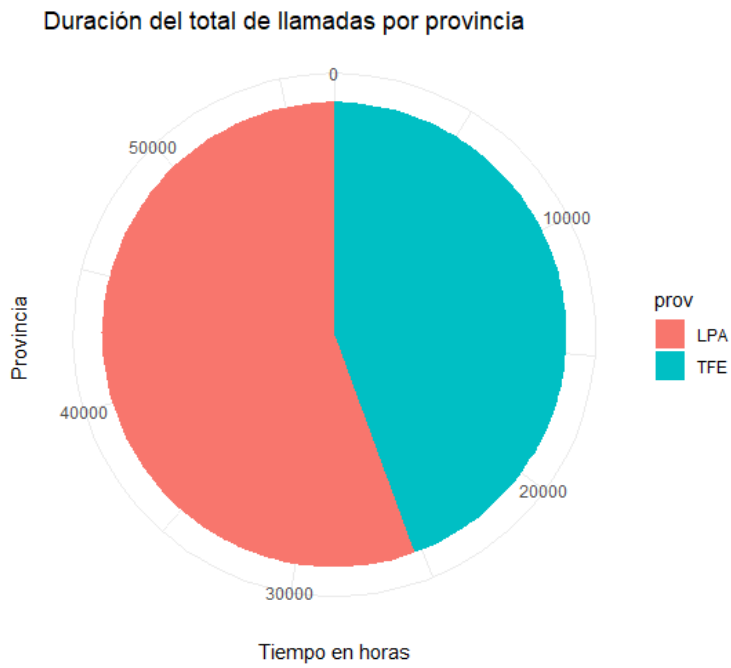
En esta sección, podremos ver un análisis de los datos de forma general, es decir, agrupándolos en meses, semanas, sectores completos y no mediante el análisis individual por trabajador.

Como hemos visto en el capítulo anterior, se obtienen los datos del tiempo de la duración de la llamada y el trabajo previo y posterior que esta conlleva, ya sea rellenando formularios, haciendo informes, etc. A lo largo del trabajo hablaremos de llamadas, refiriendonos a todas las partes del trabajo que se requiere, es decir, se han tenido en cuenta todos estos tiempos, salvo que se especifique que se han tomado una de estas tres partes para realizar nuestro análisis.

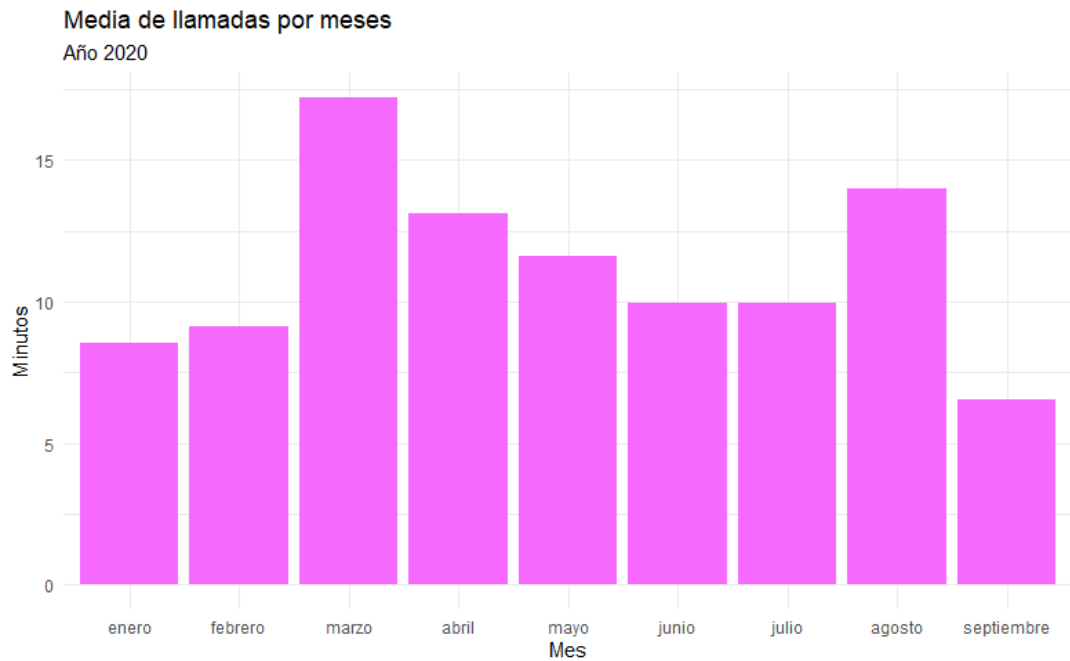
Comenzando de una manera muy general con este contenido, trabajando con la duración total, como se muestra en la figura 2.4, el tiempo empleado en las llamadas a lo largo del año 2020 por las provincias de Santa Cruz de Tenerife y Las Palmas. Podemos observar que suma es mayor en la provincias de Las Palmas que la de Santa Cruz de Tenerife. Para obtener este resultado hemos utilizado el paquete *dplyr*, sumando la duración de cada una de las llamadas sin incluir la duración de trabajo anterior y posterior.

Siguiendo con este análisis, se observa en la figura 2.5 la media de la duración de las llamadas a lo largo del año 2020, en minutos. Es destacable el aumento significativo en el mes de marzo que empieza a bajar en el mes de abril y vuelve a subir en el mes de agosto, esto coincide con las oleadas de la pandemia de la COVID-19 a lo largo de dicho año.

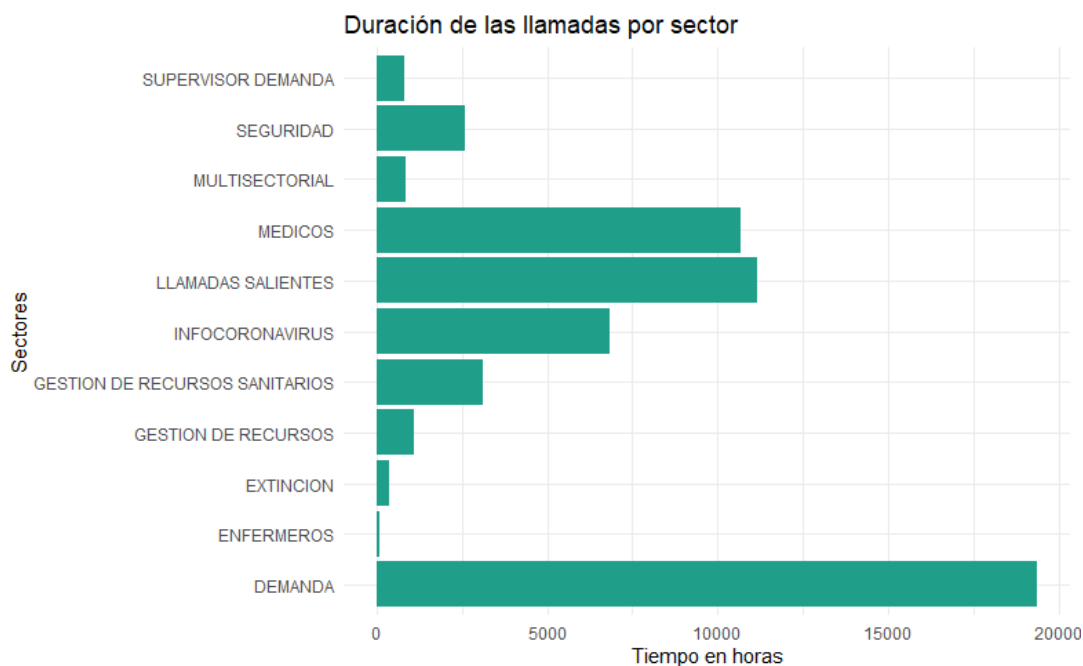
Se debe puntualizar que, evidentemente, además de las llamadas relacionadas con la pandemia, se siguieron realizando llamadas al centro de emergencias por las demás cuestiones habituales en los años precedentes a este suceso, como podemos observar en la figura 2.6. En esta gráfica, se observa que el sector de demanda es el que más llamadas recibe. Estos datos muestran que quedan registradas como demandas la recepción de llamadas y la subsiguiente derivación de la llamada al resto de los sectores.



**Figura 2.4.** Representación gráfica de la duración total de las llamadas en las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE)



**Figura 2.5.** Representación gráfica de las duración media de las llamadas en minutos.



**Figura 2.6.** Representación gráfica de la duración de las llamadas de llamadas por sector en horas.

Para entender los datos, es necesario considerar el contexto de emergencia sanitaria, y situación excepcional que caracterizó el año 2020. Específicamente, se puede destacar el mes de marzo como un período de interés, ya que fue cuando comenzó el confinamiento generalizado y se observaron reducciones significativas en sucesos comunes, como los accidentes de tráfico. Esta disminución se ilustra en la figura 2.7.

Además, si se profundiza en este análisis y se observa el trabajo medio por provincias que se muestra en la gráfica de la figura 2.8, realizando la media de la duración de todas las llamadas atendidas por los trabajadores del CECOES se observan las variaciones asociadas a acontecimientos sociales, sanitarios y políticos. En este análisis, se detalla de una manera más precisa cómo ha afectado la pandemia en cada provincia, observando que las llamadas fueron más duraderas y/o más frecuentes en la provincia de Las Palmas.

Cada contacto que tienen los trabajadores se divide en el tiempo de trabajo realizado antes de la llamada, el tiempo que dura la llamada en sí y el tiempo de trabajo posterior. Se representa en la figura 2.9, el tiempo invertido por todos los trabajadores de ambas provincias en el trabajo que realizan completando informes, mandando servicios, asegurándose que los servicios se completan, etc.

Dividiendo por provincias, se muestra en la gráfica de la figura 2.10, que la suma de la duración de este trabajo ha sido más elevada en la provincia de Las Palmas. Destacan en dicha provincia los meses de marzo, por su cambio en



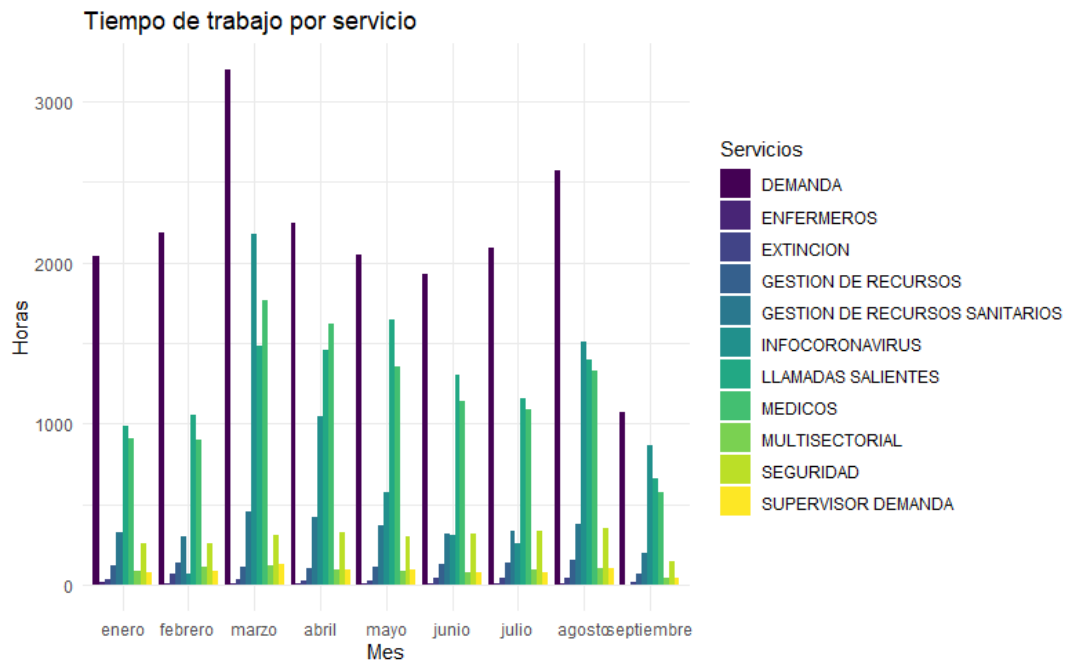


Figura 2.7. Representación gráfica del tiempo total en horas a lo largo del año 2020 por cada sector.

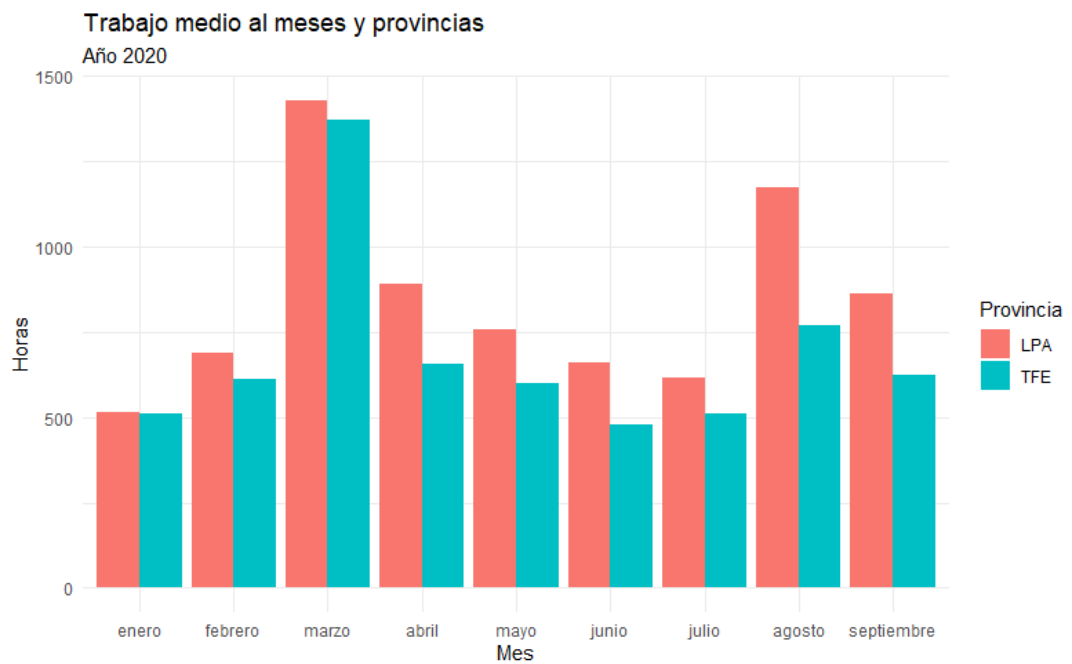
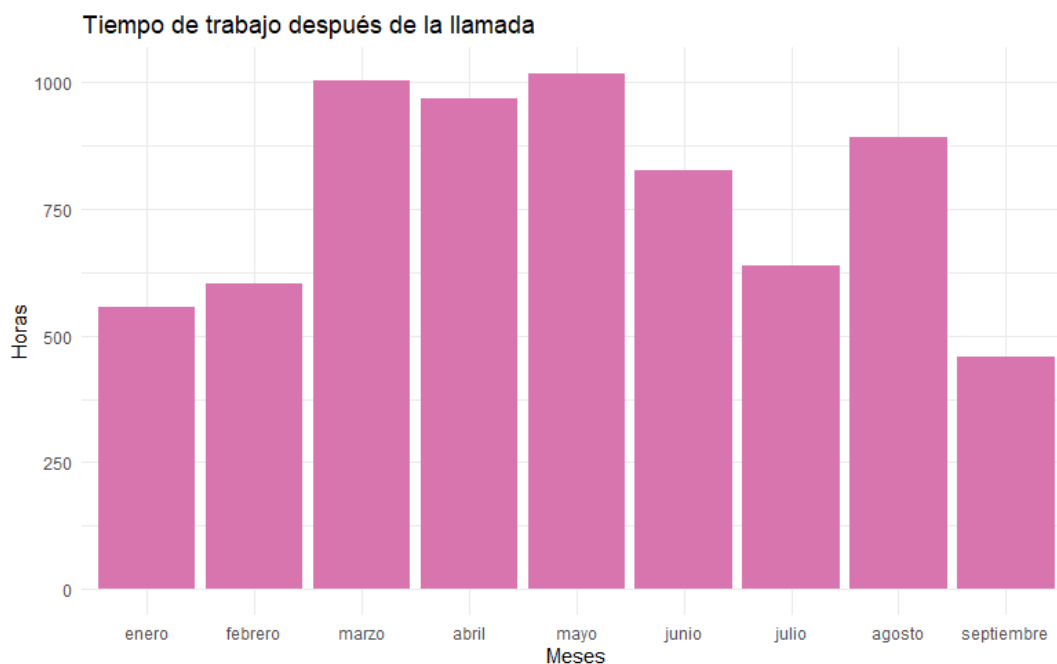


Figura 2.8. Representación gráfica de la duración de las llamadas a lo largo del año 2020 por las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE)



**Figura 2.9.** Representación gráfica de tiempo empleado por los trabajadores a lo largo del año 2020 después de las llamadas en horas

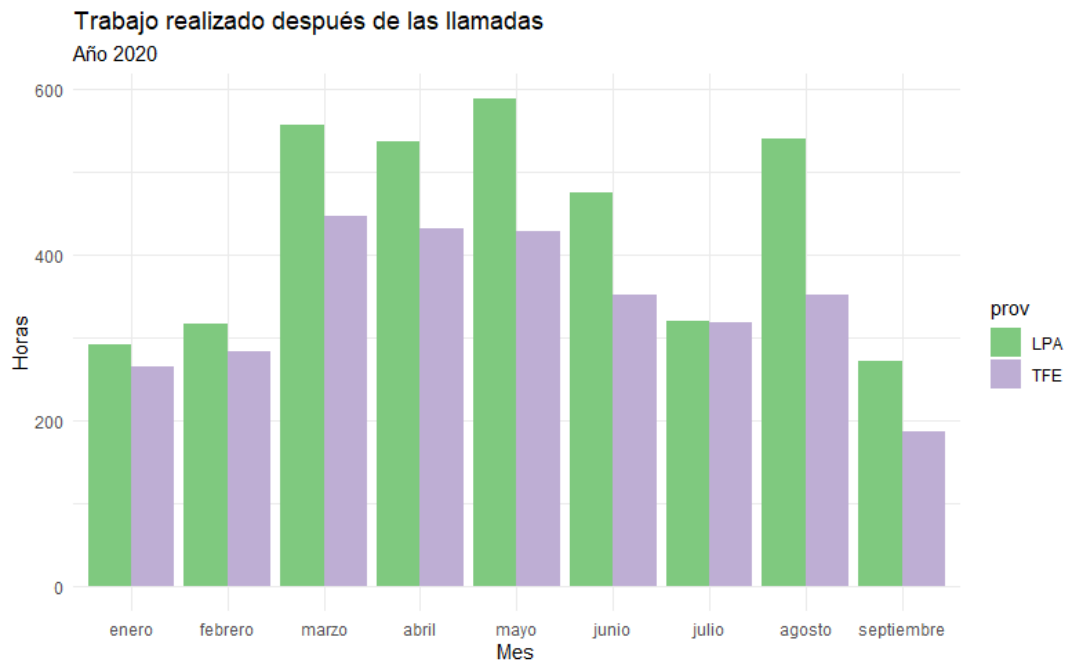
la tendencia, el mes de mayo donde alcanza su máximo y el mes de agosto ya que vuelve a haber otro pico después de un descenso de los meses anteriores.

Para garantizar la claridad y comprensión de las gráficas generadas se ha llevado a cabo un proceso de filtrado de algunos trabajadores. Esta acción fue necesaria debido a que, al no realizar dicho filtrado, la información se comprime en exceso y no permite la visualización de los datos.

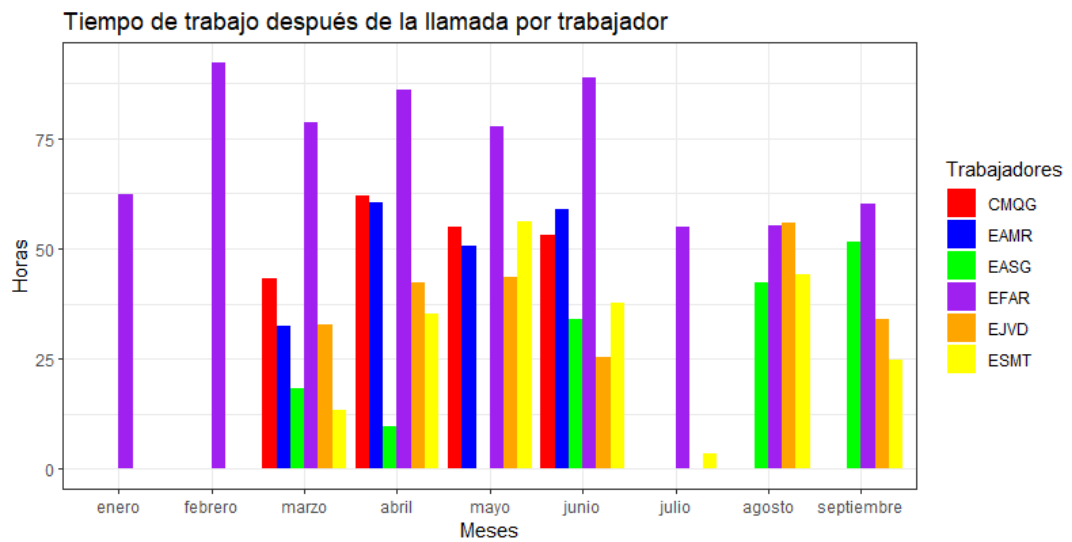
Para saber qué trabajadores escoger, se ha realizado un histograma para detectar qué trabajadores tenían los tiempos más destacados y que así, nos quedara un número pequeño de trabajadores de forma objetiva. Tras esta limpieza de datos, en el resultado final se muestran los trabajadores con más de 20 horas de llamadas a lo largo del mes.

En la figura 2.11, se observa que los trabajadores con más horas de trabajo posterior son casi exclusivamente enfermeros (identificados por la letra "E" al principio del nombre).

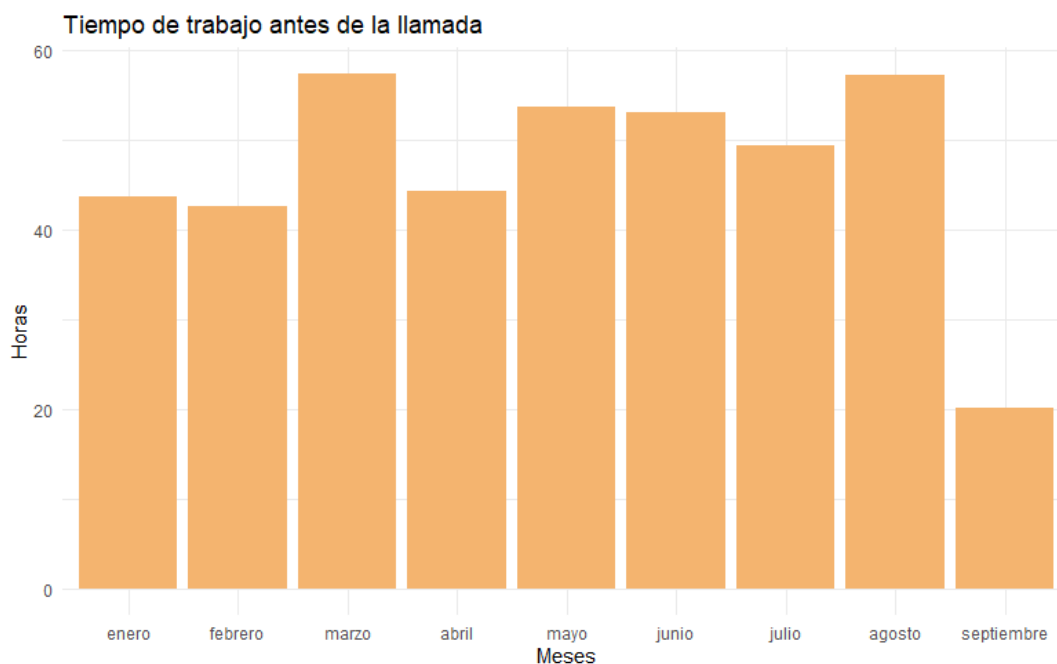
A continuación, se muestra el análisis de las horas de trabajo que se emplean antes de la llamada en la gráfica 2.12. Se observa que el número de horas empleadas empleado por los trabajadores antes de la llamadas es mucho menor que las horas empleadas después de la llamada. Además, este tipo de trabajo es consistente a lo largo del tiempo, siendo mayor de 40 horas en todos los meses. La disminución observada en el mes de septiembre se asocia a la ausencia parcial de datos para este periodo.



**Figura 2.10.** Representación gráfica de tiempo empleado por los trabajadores a lo largo del año 2020 después de las llamadas, en horas, por las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE)



**Figura 2.11.** Representación gráfica del tiempo empleado por los trabajadores seleccionados después de las llamadas durante el año 2020

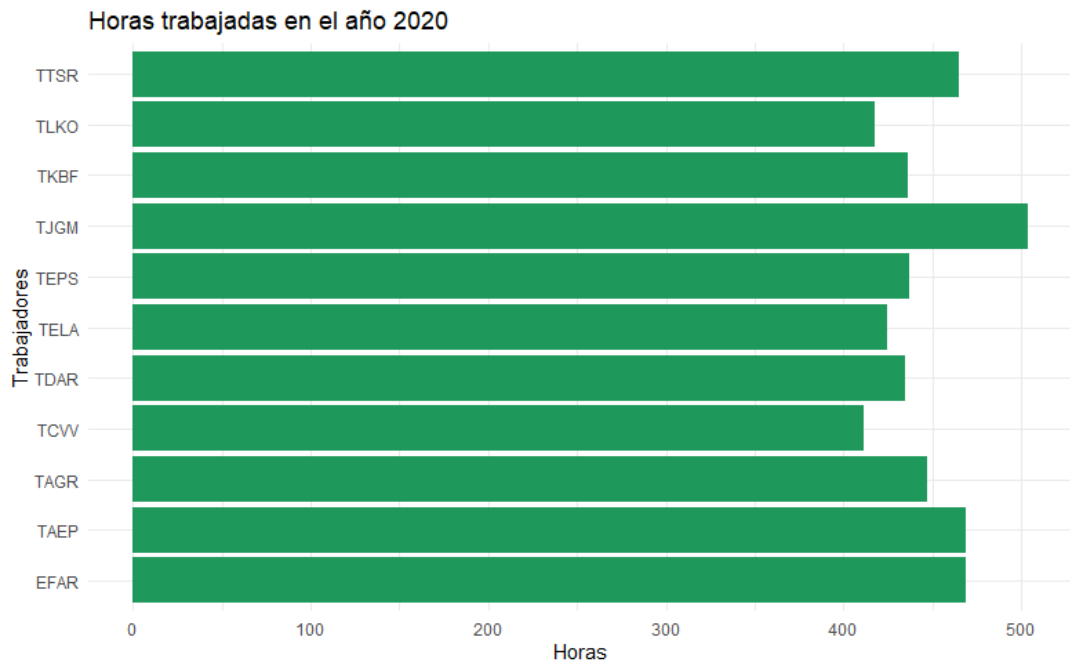


**Figura 2.12.** Representación gráfica del tiempo empleado por los trabajadores antes de la llamada por mes a lo largo del año 2020. No se dispone de datos de todo el periodo del mes de septiembre

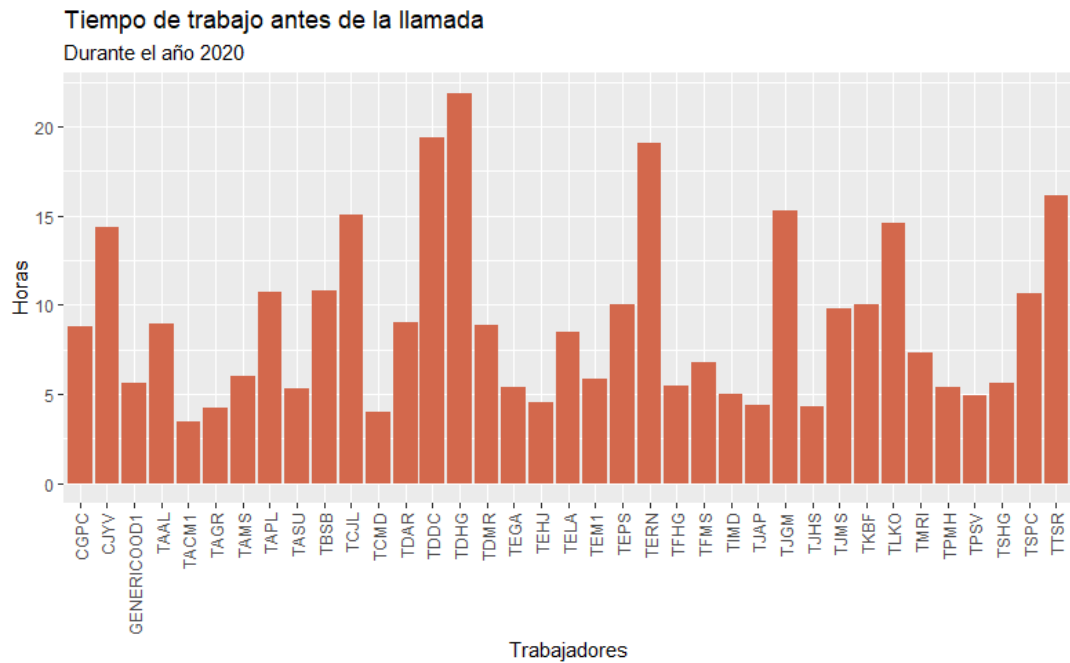
Para un análisis pormenorizado se estudia cuántas horas dedica cada trabajador al año a realizar el trabajo correspondiente antes de llamada. Para una óptima visualización de los datos en este tipo de herramientas, se ha realizado un proceso de filtrado de los trabajadores. En este análisis, se descartaron aquellos trabajadores que acumularon menos de 200 minutos en esta tarea durante el año 2020, representados en la figura 2.13. Esto se debe a que, al tener un número elevado de trabajadores, los valores extremos dificultan el análisis detallado.

Se aprecia en la figura 2.14 que los técnicos (identificados con "T" al principio de su nombre) y los médicos (identificados con "C" al principio de su nombre) son los que más trabajo realizaron antes de la llamada durante el pasado año.

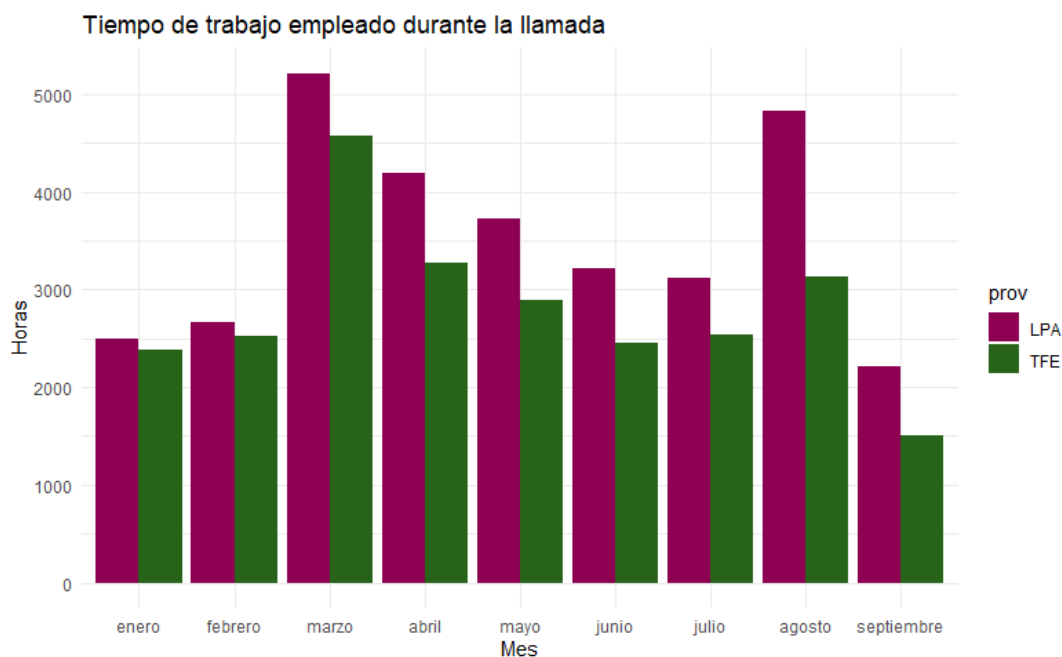
Por último, en cuanto a la división de funciones en el trabajo de las llamadas. Destacan en la figura 2.15 cifras mucho mayores y son las que marcan el trabajo total del CECOES. En el mes de marzo, se superan las 5000 horas en duración de la llamada en la provincia de Las Palmas. En este mes, la provincia de Santa Cruz de Tenerife también tiene su máximo del año. Estos resultados son convergente con análisis anteriores donde se puede apreciar el impacto durante los meses correspondientes a la pandemia.



**Figura 2.13.** Representación gráfica de las horas de trabajo empleadas por los trabajadores previamente filtrados



**Figura 2.14.** Representación gráfica del tiempo empleado por cada trabajador durante el año 2020



**Figura 2.15.** Representación gráfica del total del tiempo en horas durante la llamada a lo largo del año 2020 en las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE)

### 2.2.1. Análisis de datos relacionados con el COVID-19

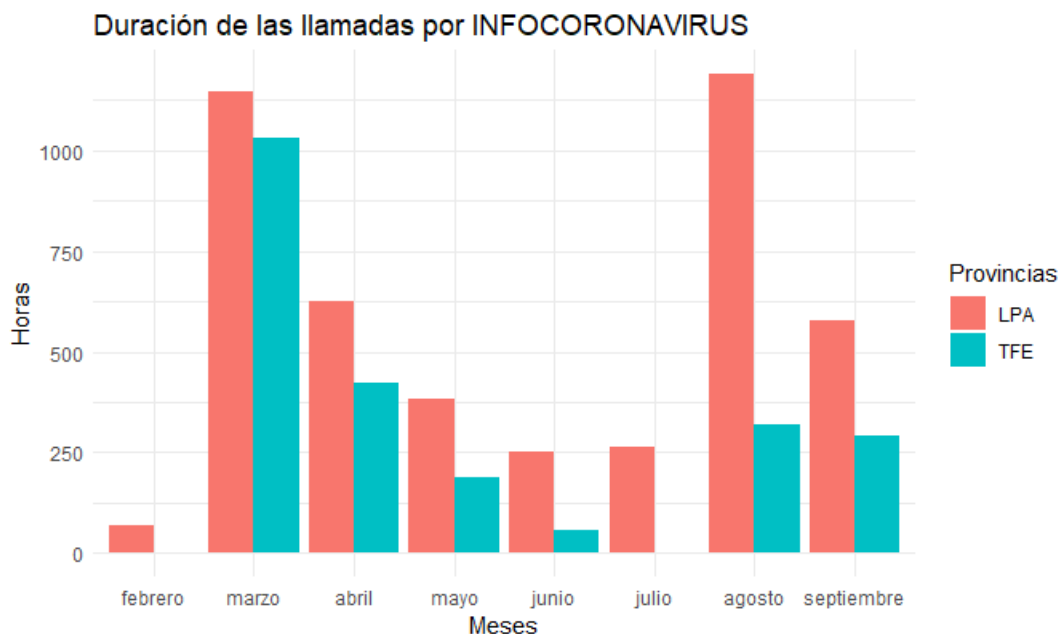
Durante el año 2020 España, junto con la mayoría de países del mundo, sufrió uno de los hechos más significativos de la historia reciente, la pandemia de Coronavirus se propagó con rapidez, provocando una emergencia sanitaria con graves implicaciones políticas, económicas y sociales. Para hacer una correcta interpretación de los datos se destacan algunas de las fechas más importantes relacionadas con la pandemia de COVID-19 en España durante el año 2020:

- 31 de enero de 2020: Se confirma el primer caso de COVID-19 en España, un turista alemán en La Gomera, Islas Canarias.
- 9 de marzo de 2020: Se declara el estado de alarma en España debido al rápido aumento de casos de COVID-19 en el país.
- 14 de marzo de 2020: Se anuncia el confinamiento nacional en España, imponiendo restricciones de movilidad y cierre de comercios no esenciales.
- 26 de marzo de 2020: España supera a China en el número de muertes por COVID-19, convirtiéndose en el país con el segundo mayor número de fallecidos en ese momento.
- 2 de abril de 2020: Se registra el pico más alto de muertes diarias por COVID-19 en España, con 950 fallecidos en 24 horas.
- 26 de abril de 2020: Se anuncia la primera fase de desconfinamiento, permitiendo ciertas actividades y flexibilizando las restricciones en algunas regiones.

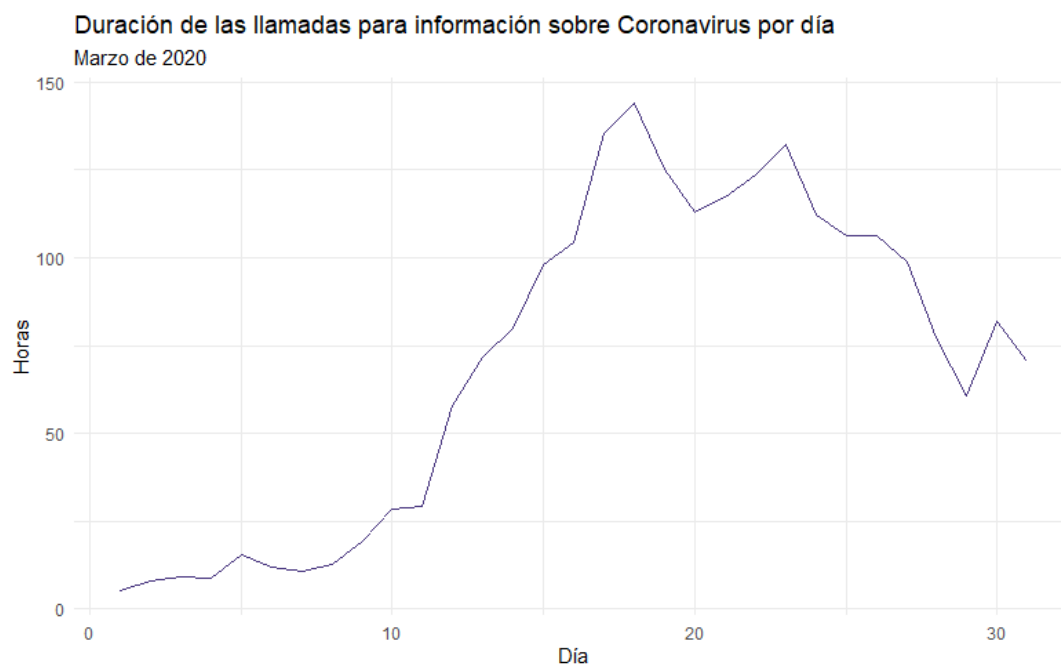
- 21 de junio de 2020: Finaliza el estado de alarma en España después de 99 días de confinamiento.
- 17 de agosto de 2020: Se establecen restricciones en Barcelona y varias localidades debido al aumento de casos de COVID-19.
- 21 de octubre de 2020: El Gobierno de España declara un nuevo estado de alarma para permitir la aplicación de medidas más estrictas y controlar la segunda ola de la pandemia.
- 27 de diciembre de 2020: Comienza la campaña de vacunación contra el COVID-19 en España, priorizando a los grupos de mayor riesgo.

En la figura 2.16 se representa la demanda de información sobre coronavirus. Los meses del año 2020 donde más demanda hubo de información sobre coronavirus, se observa que los meses más destacados en Las Palmas fueron marzo y agosto, y en Santa Cruz de Tenerife durante el mes de marzo. Estos datos coinciden con el principio del confinamiento nacional y con uno de los notables aumentos de casos durante el año.

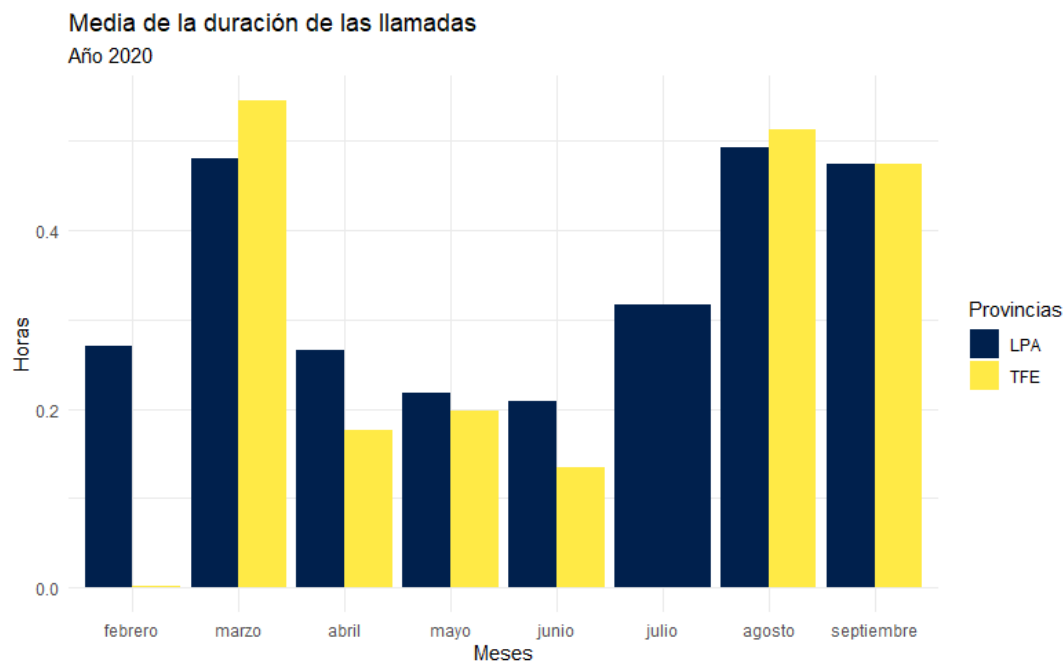
El mes de marzo, como se puede observar en la figura 2.17, resulta especialmente interesante al coincidir con el comienzo del estado de alarma. A lo largo de este mes fueron subiendo las llamadas referidas a este sector, hasta alcanzar su máximo el día 18, pocos días después del inicio del confinamiento.



**Figura 2.16.** Representación gráfica de la duración de las llamadas en horas del tiempo empleado en la demanda de información sobre Coronavirus en el año 2020 en las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE)



**Figura 2.17.** Representación gráfica de la duración de las llamadas por día para obtener información sobre el Coronavirus durante el mes de marzo del año 2020.



**Figura 2.18.** Representación gráfica de la duración media en horas de las llamadas telefónicas para las provincias de Las Palmas (LPA) y Santa Cruz de Tenerife (TFE) en los meses de febrero a septiembre. En el mes de julio no hay datos para la provincia de Santa Cruz de Tenerife.



La figura 2.17 muestra una ampliación drástica, tanto en la duración como en la frecuencia de las llamadas, en fechas cercanas al 14 de marzo. Esta fecha coincide con el anuncio y la implementación del confinamiento debido al estado de alarma decretado por el Gobierno debido al brote de Coronavirus.

En la figura 2.18, se presenta una representación gráfica detallada de la duración media en horas de las llamadas telefónicas para las provincias de Las Palmas y Santa Cruz de Tenerife, abarcando el periodo comprendido entre los meses de febrero a septiembre. Un patrón notablemente interesante se observa en los datos, revelando un aumento significativo en las medias de duración de las llamadas durante los meses de febrero, agosto y septiembre. Aunque no se ha realizado un análisis detallado del contenido de las consultas y llamadas, la duración media es un indicador indirecto de la demanda creciente de información por parte de los usuarios; podrían implicar consultas más detalladas, solicitudes de orientación o situaciones complejas que requieren un enfoque más exhaustivo. Consecuentemente esto implica un mayor trabajo y dedicación por parte del personal de emergencias del CECOES.

## 2.3. Visualización de datos

En esta sección se ha realizado un análisis de resultados más detallado dirigido al usuario final. Para este tipo de carga de datos hemos elegido el programa Power BI que nos permite una visualización más clara de los datos.

En cuanto al procedimiento de carga de datos con Excel en Power BI, se sigue un proceso que incluye los siguientes pasos:

- Preparación de los datos en Excel: antes de cargar los datos en Power BI, es fundamental realizar una preparación adecuada de los datos en Excel, del mismo modo que lo hemos hecho para la extracción de datos de datos en RStudio. Esto implica organizar y limpiar los datos, asegurándose de que estén estructurados correctamente, eliminando duplicados y resolviendo inconsistencias, entre otros pasos de limpieza y transformación necesarios.
- Importación de datos a Power BI: una vez que los datos en Excel están preparados, se pueden importar a Power BI. Esto se puede hacer utilizando la función *Obtener datos* en Power BI Desktop, que permite cargar un archivo de Excel o conectarse a una hoja de Excel existente.
- Transformación y modelado de datos: después de cargar los datos de Excel en Power BI, es posible realizar transformaciones adicionales y modelar los datos según sea necesario. Esto incluye combinar tablas, crear relaciones entre ellas, aplicar filtros y crear cálculos personalizados para preparar los datos para su análisis y visualización posterior, de una forma muy dinámica e intuitiva.

- Creación de informes y paneles: una vez que los datos están cargados y transformados en Power BI, se pueden crear informes interactivos y paneles de control utilizando las herramientas de visualización disponibles. Estos informes y paneles permiten presentar los datos de manera efectiva, aplicar filtros, explorar interactivamente y compartir información con otros usuarios.

En nuestro caso, hemos creado un filtro con los trabajadores del CECOES y a partir de este, podemos visualizar las siguientes gráficas representadas en la figura 2.19:

- Duración de la llamada por mes: esta gráfica, además de enseñar los datos esclarecedores sobre la tendencia de las llamadas a lo largo del año 2020, también permite aplicar un filtro por mes para el resto de las gráficas.
- Horas de trabajo por día: se observan los datos a lo largo del mes con los filtros que hayamos decidido aplicar.
- Horas de trabajo por trimestre: se puede observar claramente la evolución y el cambio a lo largo del tiempo. Esto permite identificar patrones, tendencias y fluctuaciones estacionales de una manera más intuitiva y comprensible.
- Duración de las llamadas por sectores: esta gráfica sirve para comparar los sectores con más demanda a lo largo de la franja temporal que necesitemos examinar. En nuestro caso, cabe destacar de las pruebas realizadas, el más demandado ha sido el sector médico.
- Duración de las llamadas por provincias: se puede ver la duración de las llamadas dependiendo de los filtros que apliquemos, como el mes, el sector o los trabajadores de mismos sectores que hayamos filtrado con nuestro filtro o las gráficas anteriormente descritas.

Destacan los siguientes beneficios de utilizar una herramienta como Power BI:

- Visualización interactiva: Power BI ofrece herramientas de visualización de datos altamente interactivas, como gráficos, tablas dinámicas y mapas. Estas visualizaciones permiten explorar los datos de manera intuitiva, facilitando la comprensión y el descubrimiento de patrones y tendencias.
- Informes y paneles dinámicos: Power BI permite crear informes y paneles dinámicos que se actualizan automáticamente a medida que los datos subyacentes cambian. Esto proporciona información en tiempo real y ayuda a tomar decisiones informadas y basadas en datos actualizados.
- Conexión con diversas fuentes de datos: Power BI es compatible con una amplia gama de fuentes de datos, incluyendo bases de datos, servicios en la nube y archivos de Excel, entre otros. Esto permite reunir datos de diferentes fuentes y combinarlos para obtener una visión completa y enriquecedora de la información.

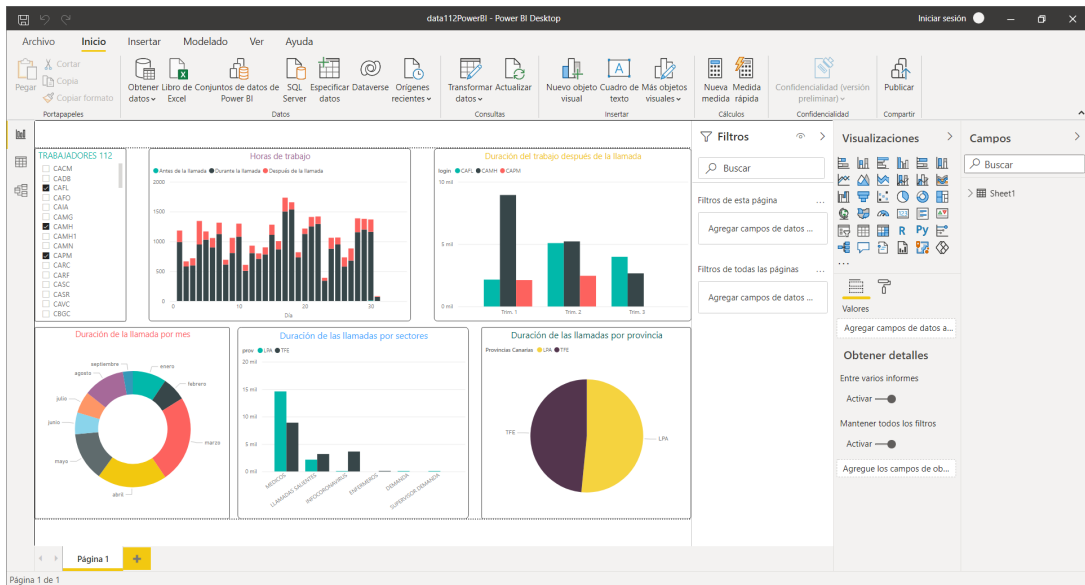


Figura 2.19. Imagen del panel de control de Power BI.

- **Análisis avanzado:** Power BI ofrece capacidades de análisis avanzado, como modelado de datos, cálculos personalizados y consultas de datos. Estas características permiten realizar análisis complejos y extraer conocimientos más profundos de los datos, lo que resulta valioso para la toma de decisiones estratégicas.



## Predicción de modelos de series temporales

En este capítulo, aplicaremos modelos de predicción de series temporales, en concreto, el modelo ARIMA para realizar una predicción de nuestro conjunto de datos, específicamente el número de llamadas recibidas por el CECOES relacionadas con la solicitud de información sobre COVID-19. Utilizaremos técnicas de análisis de series temporales para predecir valores futuros basados en patrones y tendencias observadas en los datos históricos. El modelo ARIMA (Auto-Regressive Integrated Moving Average) se utilizará para modelar y predecir el comportamiento de la serie temporal, teniendo en cuenta tanto los componentes de autoregresión como los de promedio móvil, así como la integración de la serie para lograr estacionariedad. Al aplicar el modelo ARIMA, esperamos obtener pronósticos precisos y útiles sobre el número de llamadas futuras relacionadas con la información de COVID-19 recibidas por el CECOES.

### 3.1. Series temporales y modelo ARIMA

Las series temporales son conjuntos de datos que se registran y organizan en orden cronológico, lo que permite el análisis de su evolución y comportamiento a lo largo del tiempo. Estos datos están asociados con marcas de tiempo que indican cuándo se recopilaron o registraron las observaciones.

Una característica fundamental de las series temporales es su dependencia temporal, lo que significa que las observaciones están correlacionadas y se influyen mutuamente. Esto implica que el valor actual de la serie puede depender de los valores pasados y, a su vez, puede influir en los valores futuros. Esta dependencia temporal permite analizar tendencias, patrones estacionales, ciclos y otros comportamientos a lo largo del tiempo en los datos de la serie temporal. El análisis de series temporales es utilizado para comprender y predecir el comportamiento de los datos en el futuro, aprovechando la información histórica y las relaciones temporales entre las observaciones.

En esta memoria aplicaremos el modelo ARIMA para analizar y predecir series temporales, para extender la información sobre este tema, se puede consultar [9]. Este modelo combina tres componentes:

1. Componente AR (autoregresivo): Este componente modela la dependencia de un valor actual en función de sus valores pasados. Se basa en la suposición de que los valores anteriores tienen un impacto en el valor actual. La parte "autoregresiva" del modelo ARIMA utiliza una regresión lineal de los valores pasados para predecir el valor actual.
2. Componente MA (media móvil): Este componente se basa en la dependencia de un valor actual en relación con los errores pasados. Los errores son las diferencias entre los valores reales y los valores predichos por el componente AR. El componente MA utiliza una media móvil ponderada de los errores pasados para realizar predicciones.
3. Componente I (integrado): Este componente se utiliza para lograr la estacionariedad en la serie temporal. La estacionariedad implica que la media y la varianza de la serie no cambian con el tiempo. La parte "integrada" del modelo ARIMA se refiere a la diferenciación de la serie para eliminar las tendencias y hacerla estacionaria. Esto implica tomar diferencias entre los valores observados en diferentes momentos.

El modelo ARIMA se ajusta a los datos de la serie temporal utilizando métodos estadísticos para estimar los parámetros adecuados de los componentes AR, MA e I. Una vez ajustado, el modelo puede utilizarse para realizar pronósticos futuros y analizar patrones en la serie temporal.

Los modelos ARIMA especifican explícitamente los componentes mencionados anteriormente como parámetros. Se utiliza la notación estándar ARIMA(p,d,q), donde los parámetros (p,d,q) toman valores enteros para identificar el tipo de modelo.

- Autoregresión (p): representa el número de retrasos que se incluirán en el modelo, también conocido como orden de retraso.
- Integrado (d): indica la cantidad de veces que los datos han sido diferenciados para lograr la estacionariedad, también conocido como grado de diferenciación.
- Media móvil (q): indica el número de errores pasados que se incluirán en el modelo, es decir, el tamaño de la ventana de la media móvil, también conocido como orden de la media móvil.

Por tanto, la fórmula del modelo ARIMA se compone de dos partes: el componente autoregresivo (AR) y el componente de media móvil (MA).

La parte AR está representada por el término  $(1 - \phi_1 B - \dots - \phi_p B^p)$ , donde  $\phi_1, \dots, \phi_p$  son los coeficientes del componente AR. Estos coeficientes representan las ponderaciones de los retrasos de la serie temporal en la predicción del valor actual. El operador  $B$  indica el retraso de la serie temporal en un paso de tiempo y  $B^p$  indica un retraso de  $p$  pasos. En resumen, este término AR captura las dependencias pasadas de la serie temporal.

La parte MA está representada por el término  $(1 + \theta_1 B + \dots + \theta_q B^q)\varepsilon_t$ , donde  $\theta_1, \dots, \theta_q$  son los coeficientes del componente MA y  $\varepsilon_t$  es el término de error en el tiempo  $t$ . Los coeficientes  $\theta_1, \dots, \theta_q$  representan las ponderaciones de los términos de error pasados en la predicción del valor actual. Este término MA captura las influencias de los errores pasados en el valor actual.

El término  $(1 - B)^d$  indica el componente de diferenciación integrada, donde  $d$  es el orden de diferenciación integrada. Este término se utiliza para hacer estacionaria la serie temporal, es decir, eliminar tendencias o patrones no deseados.

Además, la fórmula incluye una constante  $c$  que se suma al modelo para ajustar el nivel base de la serie temporal.

En conjunto, la fórmula del modelo ARIMA establece una relación entre los valores observados de la serie temporal, los retrasos pasados, los coeficientes del modelo, los errores pasados y el componente de diferenciación. Esta relación se utiliza para estimar y predecir los valores futuros de la serie temporal.

La fórmula del modelo ARIMA es:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \dots + \theta_q B^q)\varepsilon_t$$

donde, en resumen:

- $\phi_1, \dots, \phi_p$  son los coeficientes del componente AR.
- $B$  es el operador de retraso.
- $d$  es el orden de diferenciación integrada.
- $y_t$  es el valor observado en el tiempo  $t$  de la serie temporal.
- $c$  es una constante.
- $\theta_1, \dots, \theta_q$  son los coeficientes del componente MA.
- $q$  es el orden del componente MA.
- $\varepsilon_t$  es el término de error en el tiempo  $t$ .

Esta fórmula establece la relación entre los valores observados de la serie temporal, los retrasos pasados, los coeficientes del modelo, los errores pasados y el componente de diferenciación, lo cual permite estimar y predecir los valores futuros de la serie temporal.

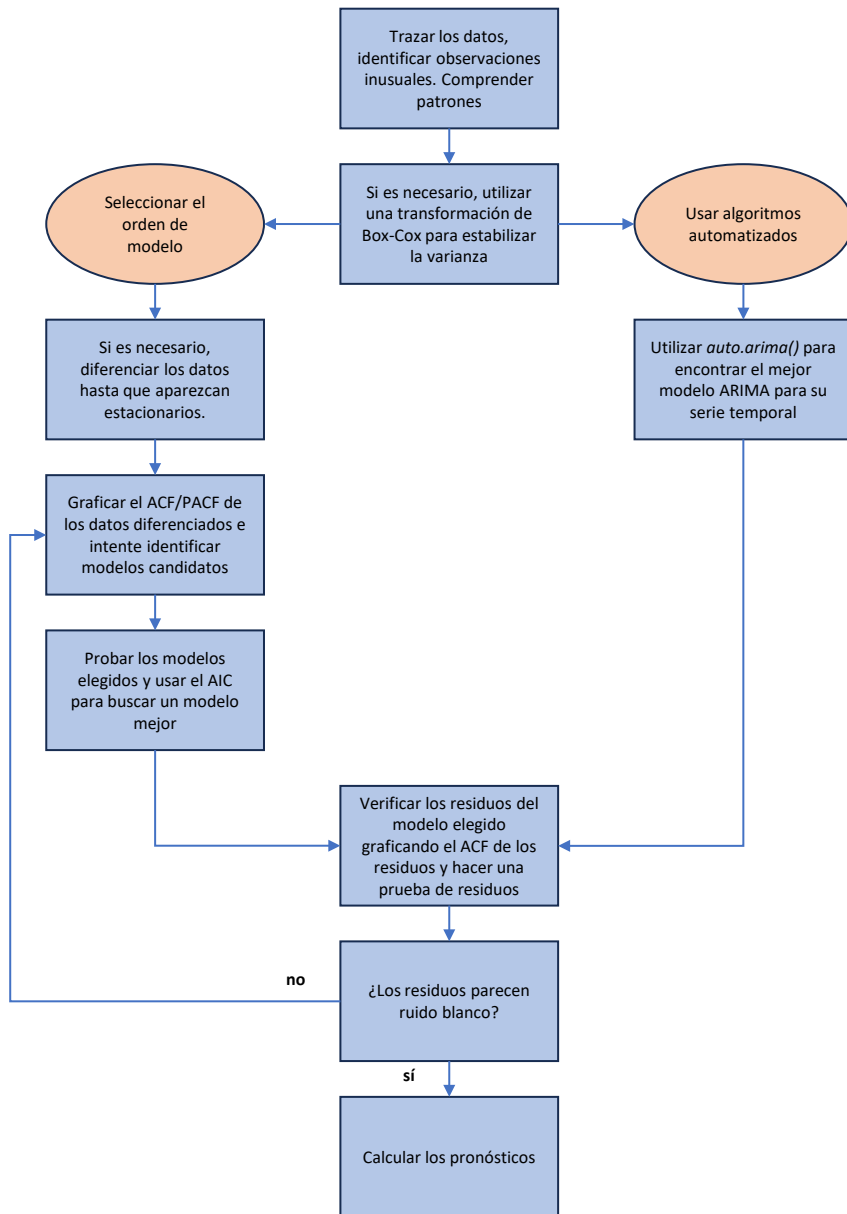


Figura 3.1. Esquema del proceso de aplicación del modelo ARIMA. Fuente adaptada: [8]



### 3.2. Aplicación método ARIMA

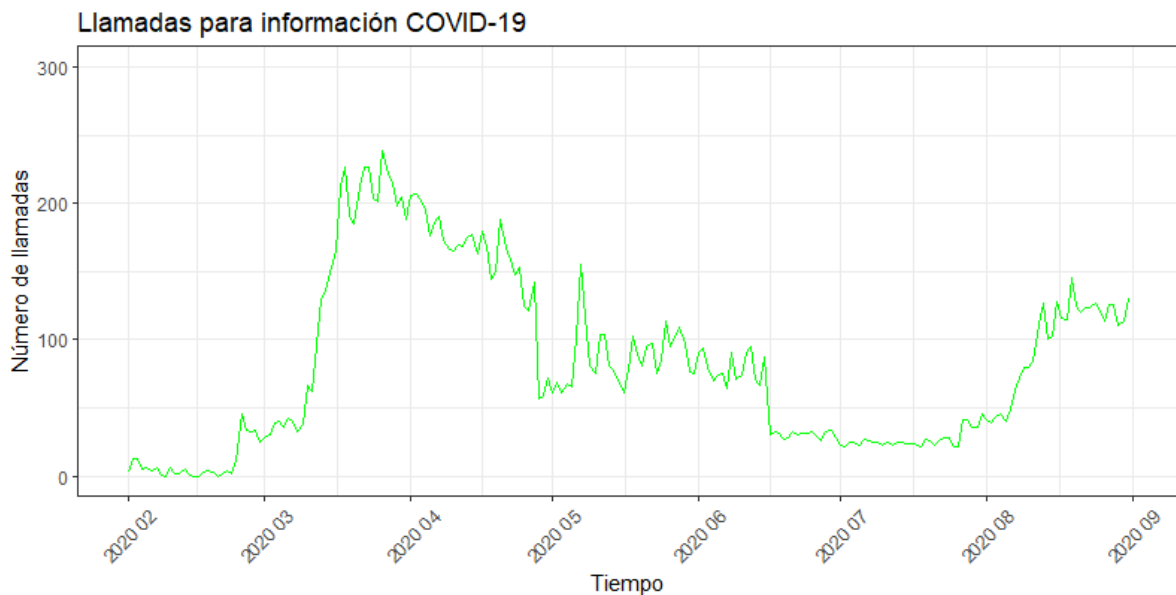
Para la aplicación de este método, hemos contado el número de llamadas que han habido por mes y por servicio a lo largo del año 2020. En concreto, nos centramos en las llamadas para información sobre COVID-19.

Para asegurarnos que nuestro análisis se podía aplicar de forma correcta, hemos tenido que comprobar que los días que no se habían recibido llamadas para el servicio que estábamos estudiando, aparecía un 0, como podemos observar en el ejemplo de la figura 3.2. Para ello, creamos una tabla con todas las combinaciones posibles entre servicios y fechas y la hemos unido con nuestra tabla original.

|    | fecha      | num_llamadas |
|----|------------|--------------|
| 46 | 2020-02-15 | 0            |
| 47 | 2020-02-16 | 0            |
| 48 | 2020-02-17 | 3            |
| 49 | 2020-02-18 | 4            |
| 50 | 2020-02-19 | 3            |
| 51 | 2020-02-20 | 0            |
| 52 | 2020-02-21 | 2            |
| 53 | 2020-02-22 | 4            |
| 54 | 2020-02-23 | 2            |
| 55 | 2020-02-24 | 11           |
| 56 | 2020-02-25 | 45           |
| 57 | 2020-02-26 | 33           |
| 58 | 2020-02-27 | 32           |
| 59 | 2020-02-28 | 34           |

**Figura 3.2.** Tabla del número de llamadas por día para pedir información sobre COVID-19 de la fila número 46 a la 59

A continuación, observamos los datos que nos ha proporcionado el CO-COES a través de una gráfica para poder identificar cualquier observación inusual. Podemos observar en la figura 3.3 que no existe un patrón estacional, encontramos cambios drásticos como el crecimiento repentino de llamadas en el mes de marzo del año 2020 y los datos no presentan ningún patrón a lo largo del tiempo, por lo que el método que apliquemos tendrá que ser para series no estacionales.



**Figura 3.3.** Representación gráfica del número de llamadas para solicitar información sobre COVID-19 a lo largo del año 2020.

Además, debemos estudiar si nuestros datos son estacionarios. Para ello, vamos a realizar un análisis de autocorrelación verificando el ACF y PACF de los datos:

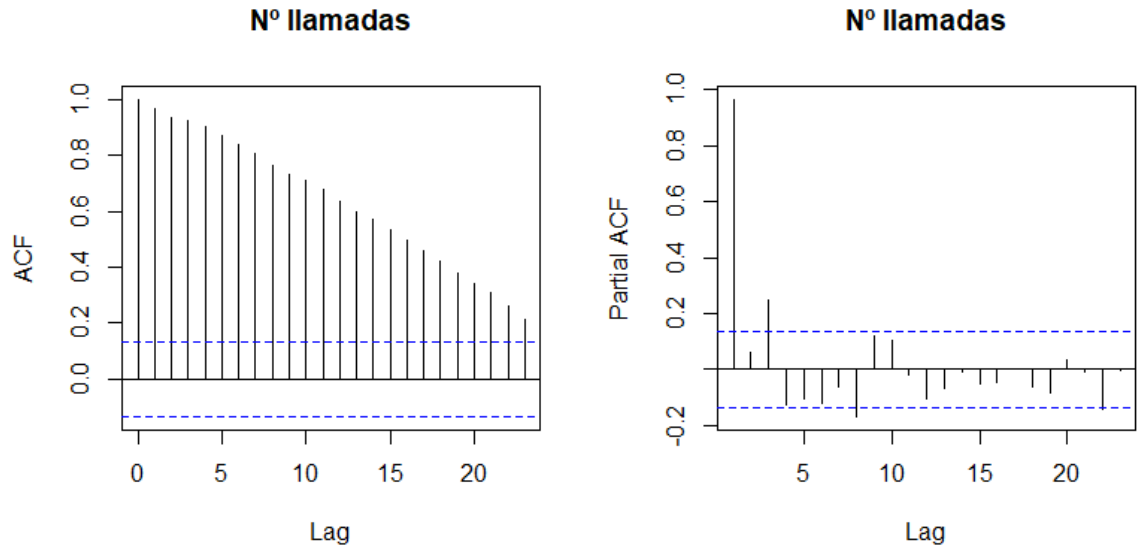
- ACF (Función de autocorrelación): Correlación entre los valores de Serie que están  $k$  separados, donde  $k$  es el orden de retraso.
- PACF (función de autocorrelación parcial): muestra la correlación parcial entre los valores de la serie que están separados por  $k$ , considerando los valores del intervalo.

A través del siguiente código:

```
par(mfrow=c(1,2))
acf(as.ts(num_llamadas_infocorona_3$num_llamadas), main="Nº llamadas")
pacf(as.ts(num_llamadas_infocorona_3$num_llamadas), main="Nº llamadas")
```

Observamos en la gráfica ACF de la figura 3.4 la autocorrelación continúa disminuyendo a medida que aumenta el retraso. En general, los procesos que tengan una ACF de este estilo (con mucha persistencia temporal), diremos que son no estacionarios.

Como hemos mencionado anteriormente, los modelos ARIMA requieren que los datos sean estacionarios. En caso de que una serie de tiempo no sea estacionaria, se puede aplicar un proceso llamado diferenciación para lograr la estacionariedad. Si la serie de tiempo no muestra patrones estacionales, se recomienda utilizar la diferenciación de primer orden para modelar los cambios en las



**Figura 3.4.** Representación gráfica ACF y PACF del número de llamadas recibidas para solicitar información sobre COVID-19 a lo largo del año 2020.

observaciones. Esto implica tomar la diferencia entre una observación y su observación anterior para obtener una nueva serie de tiempo que sea estacionaria. Esta diferenciación de retraso 1 ayuda a eliminar las tendencias y los componentes de cambio lento en los datos, lo que facilita el modelado y la predicción utilizando un modelo ARIMA. Por lo tanto, para realizar nuestra predicción aplicaremos el método de diferenciación `ffp2` de `auto.arima()` no estacional.

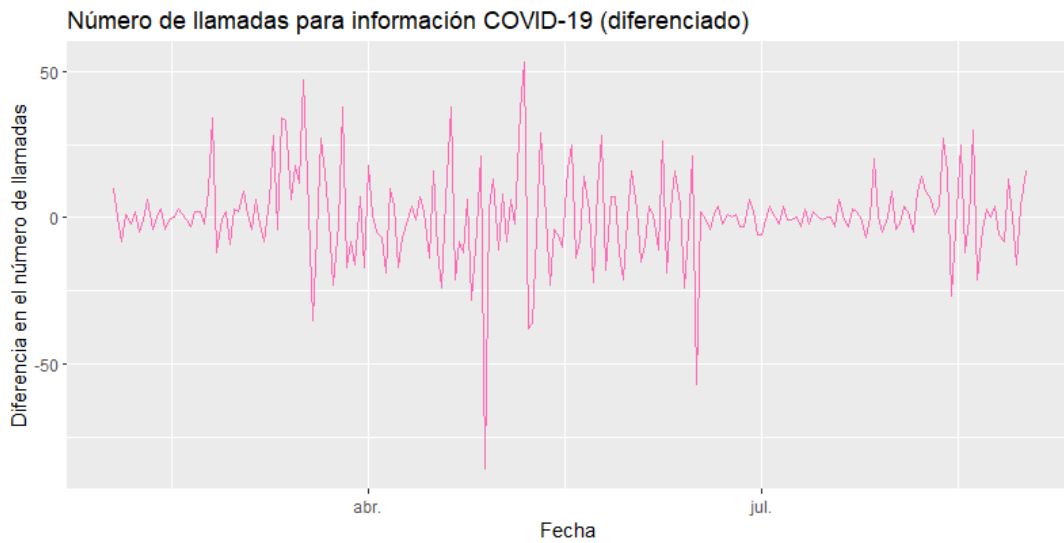
Aplicamos el método de diferenciación a nuestros datos, como observamos en la figura 3.5 y volvemos a repetir la función de autocorrelación para ver si nuestros datos son estacionarios.

A continuación, implementamos un procedimiento de ajuste automático de múltiples modelos para seleccionar el modelo de ARIMA que presente el mejor ajuste a nuestros datos con el fin de realizar nuestras predicciones con el código:

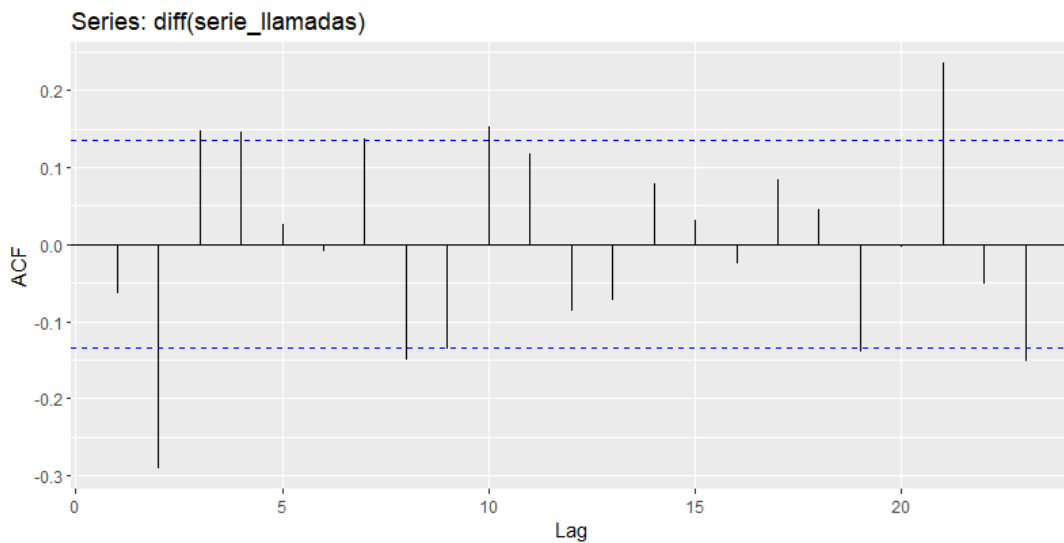
```
fit_serie_llamadas <- auto.arima(serie_llamadas, seasonal = FALSE)
summary(fit_serie_llamadas)
```

Los resultados muestran que el modelo de ARIMA que más se ajusta a nuestros datos es el ARIMA(3,1,0), como podemos observar en la figura 3.7.

Además, la distribución de los residuos es en su mayoría normal y no se observan valores atípicos. Esto es favorable, ya que implica que los intervalos de predicción serán más estrechos y más confiables.



**Figura 3.5.** Representación gráfica del número de llamadas recibidas para solicitar información sobre COVID-19 a lo largo del año 2020 diferenciada por el método ffp2.



**Figura 3.6.** Representación gráfica ACF del número de llamadas recibidas para solicitar información sobre COVID-19 a lo largo del año 2020 diferenciada por el método ffp2.

### 3.2.1. Pruebas diagnósticas para modelos ARIMA

Las pruebas diagnósticas son herramientas utilizadas para evaluar la idoneidad del modelo ARIMA en relación a los datos de una serie temporal y verificar si se cumplen los supuestos del modelo. Estas pruebas permiten determinar si existen patrones sistemáticos o estructuras residuales no capturadas por el modelo.

```

Series: serie_llamadas_2
ARIMA(3,1,0) with drift

Coefficients:
      ar1      ar2      ar3  drift
    -0.0917 -0.477  0.2793  2.0461
s.e.   0.1388  0.118  0.1384  0.9508

sigma^2 = 78.08:  log likelihood = -171.09
AIC=352.19  AICc=353.62  BIC=361.54

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set -0.00259165  8.373439  6.300671 -2.231196  10.19341  0.847149
              ACF1
Training set -0.0005346338

```

**Figura 3.7.** Verificación el resumen del informe del modelo ARIMA.

Para ello, hemos usado la función *checkresiduals()*, en R es comúnmente utilizada en el contexto del modelo ARIMA para evaluar los residuos del modelo ajustado. Después de ajustar un modelo ARIMA a una serie temporal, es importante analizar los residuos para verificar si cumplen los supuestos del modelo y si el modelo es adecuado para los datos.

Al aplicar *checkresiduals()* a un objeto de modelo ARIMA ajustado en R, se generan diferentes gráficos y pruebas diagnósticas para analizar los residuos del modelo ARIMA. Estas pruebas y gráficos ayudan a evaluar la calidad del ajuste del modelo y la adecuación de los residuos a los supuestos del modelo ARIMA.

Una prueba diagnóstica común para el modelo ARIMA es el Test de Ljung-Box, el cual se utiliza para evaluar la autocorrelación residual en los residuos del modelo ARIMA.

El Test de Ljung-Box se basa en las autocorrelaciones parciales de los residuos estandarizados. Primero, los residuos del modelo se estandarizan dividiéndolos por la desviación estándar estimada de los residuos. La fórmula para calcular los residuos estandarizados ( $\hat{\varepsilon}_t^*$ ) es:

$$\hat{\varepsilon}_t^* = \frac{\hat{\varepsilon}_t}{\sqrt{\hat{\sigma}^2}}$$

donde  $\hat{\varepsilon}_t$  son los residuos del modelo y  $\hat{\sigma}^2$  es la estimación de la varianza de los residuos.

A continuación, se calcula la estadística de prueba  $Q$  basada en las autocorrelaciones parciales de los residuos estandarizados hasta un rezago  $h$ . La fórmula para calcular la estadística de prueba  $Q$  es:

$$Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k}$$

donde  $n$  es el número de observaciones en la serie temporal y  $\hat{\rho}_k$  es la autocorrelación parcial estimada en el rezago  $k$  de los residuos estandarizados.

La estadística de prueba  $Q$  sigue una distribución chi-cuadrado con  $h$  grados de libertad. Se compara el valor de la estadística de prueba  $Q$  con el valor crítico de la distribución chi-cuadrado para un nivel de significancia dado. Si el valor de la estadística de prueba es mayor que el valor crítico, se rechaza la hipótesis nula de no autocorrelación residual significativa.

El test de Ljung-Box se utiliza como una prueba de bondad de ajuste para evaluar si el modelo ARIMA captura adecuadamente las estructuras de dependencia en los datos. Si el test indica la presencia de autocorrelación residual significativa, puede ser necesario considerar un modelo más complejo o explorar otras estructuras de dependencia.

En nuestro caso, el gráfico de residuos muestra una media alrededor de 0, la prueba de Ljung-Box da un valor  $p = 0.5$ , como podemos observar en la figura 3.8, lo cual indica que no hay suficiente evidencia para rechazar la hipótesis nula de no autocorrelación en los residuos. Esto respalda la idea de que los residuos se pueden considerar como ruido blanco.

### Ljung-Box test

```
data: Residuals from ARIMA(3,1,0) with drift
Q* = 5.8494, df = 7, p-value = 0.5574
```

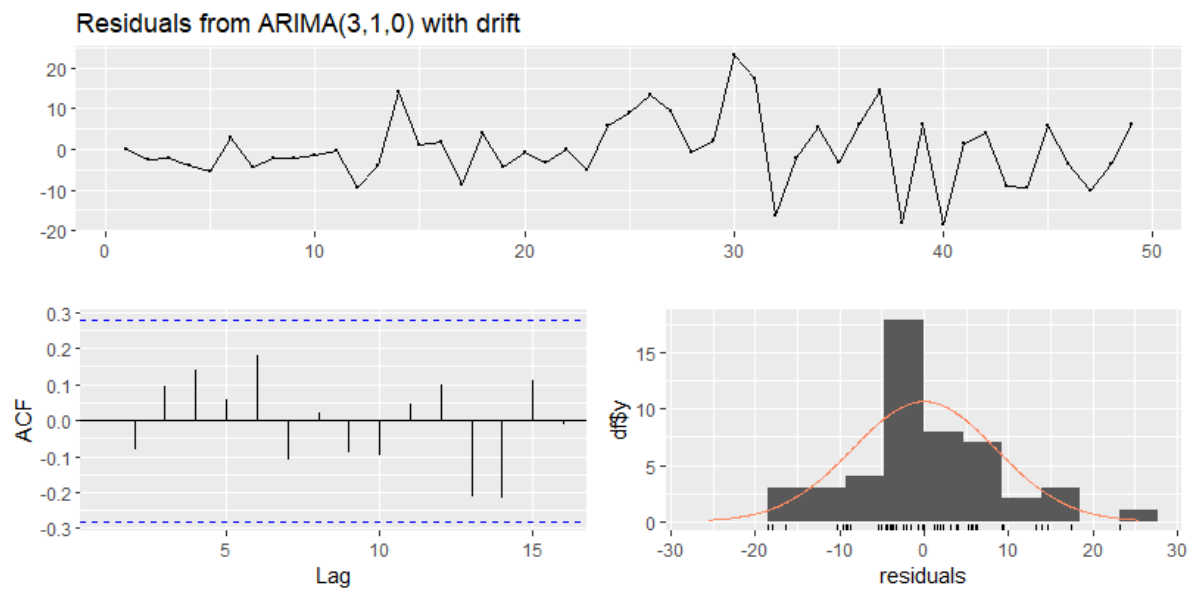
```
Model df: 3. Total lags used: 10
```

**Figura 3.8.** El código muestra la implementación del modelo ARIMA para el análisis de la serie de tiempo del Número de Llamadas.

Además, otras de las pruebas y gráficos generados por *checkresiduals()* para un modelo ARIMA incluyen:

1. Gráfico de residuos en el tiempo: Este gráfico muestra la evolución de los residuos a lo largo del tiempo. Se utiliza para detectar patrones sistemáticos, como tendencias o estacionalidad, que pueden indicar que el modelo ARIMA no captura completamente la estructura de los datos.

2. Gráfico de histograma de residuos: Este gráfico muestra la distribución de los residuos. Se utiliza para evaluar la normalidad de los residuos y detectar desviaciones significativas de la distribución normal, lo que podría indicar deficiencias en el modelo ARIMA.
3. Gráfico Q-Q de residuos: Este gráfico compara los cuantiles de los residuos con los cuantiles teóricos de una distribución de referencia, generalmente la distribución normal. Se utiliza para evaluar si los residuos siguen una distribución teórica y detectar desviaciones notables que podrían indicar problemas con el modelo ARIMA.



**Figura 3.9.** Diagnóstico de residuos del modelo ARIMA(3,1,0). Los gráficos muestran los residuos estandarizados (arriba), la autocorrelación de los residuos (abajo-izquierda) y la distribución de los residuos (abajo-derecha).

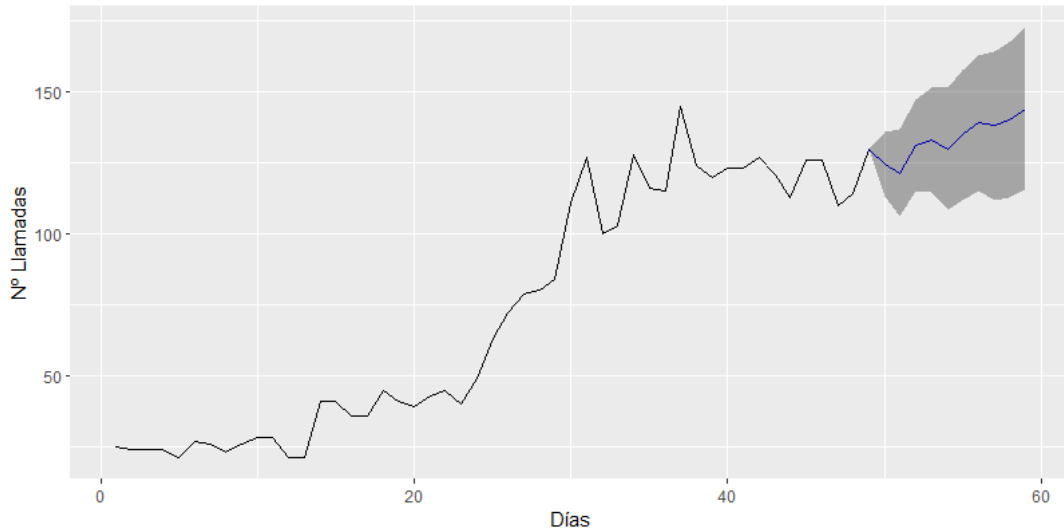
En resumen, podemos observar en la figura 3.9 los resultados muestran que los residuos se comportan como ruido blanco, lo que indica un buen ajuste del modelo ARIMA(3,1,0) a la serie de tiempo. Esto sugiere que el modelo es capaz de hacer pronósticos precisos y confiables.

### 3.2.2. Predicción modelo ARIMA(3,1,0)

El modelo ARIMA(3,1,0) seleccionado automáticamente captura la tendencia presente en los datos y genera pronósticos confiables con intervalos de predicción adecuados. 3.10.

Hemos realizado una validación de los datos al excluir los datos correspondientes al mes de septiembre de nuestra predicción. En esta validación, hemos

observado que aunque la predicción no es exacta, es confiable y se ajusta adecuadamente a los patrones y tendencias presentes en los datos.



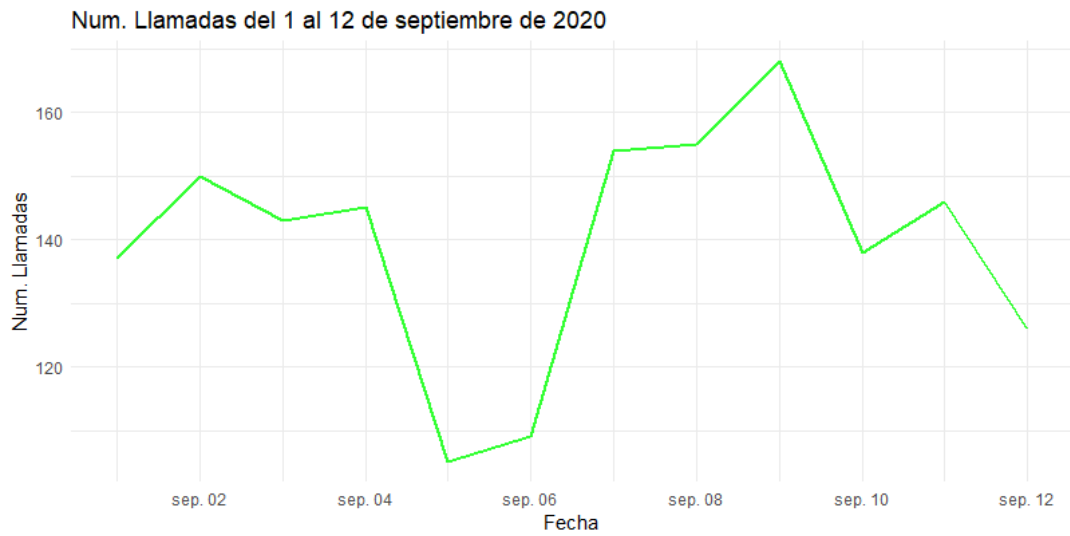
**Figura 3.10.** Representación gráfica de la predicción del número de llamadas para solicitar información sobre COVID-19 desde julio hasta agosto del año 2020.

Para evaluar la precisión de nuestra predicción, hemos realizado un procedimiento de validación cruzada dejando fuera los datos correspondientes al periodo del 1 al 12 de septiembre de 2020 [3.11](#). Durante este periodo, no se utilizó la información disponible para ajustar el modelo.

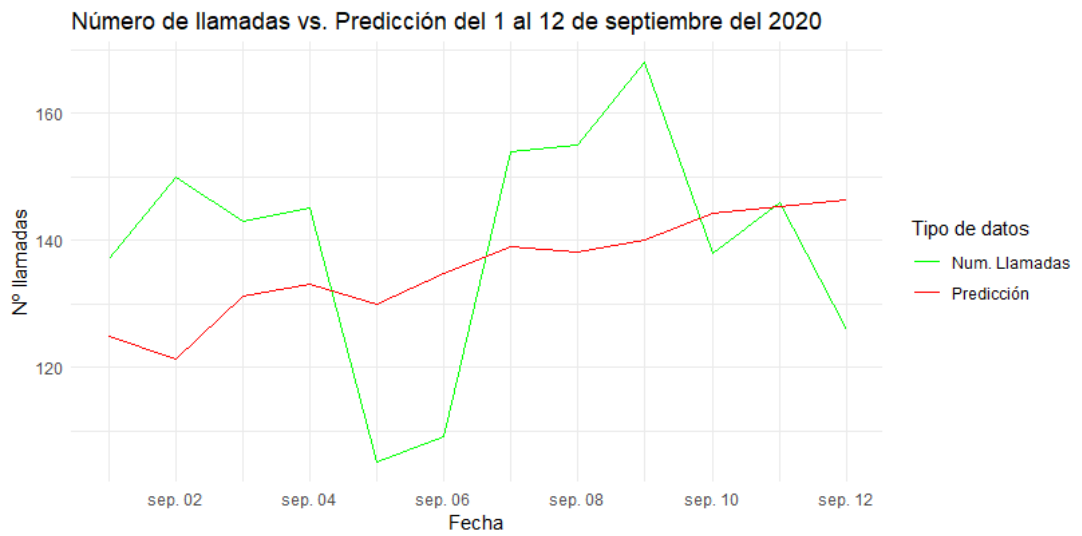
Posteriormente, comparamos nuestra predicción con los datos reales correspondientes al mismo periodo [3.12](#). Esta comparación nos permite evaluar cuán cercanas fueron nuestras predicciones a los valores reales que teníamos inicialmente.

Es importante realizar esta validación cruzada para evaluar la capacidad del modelo para generalizar y predecir con precisión más allá de los datos utilizados en su ajuste. Al comparar los resultados, podemos identificar posibles desviaciones o discrepancias entre la predicción y los datos reales, lo que nos brinda información sobre la calidad y confiabilidad de nuestro modelo.





**Figura 3.11.** Representación gráfica del número de llamadas para solicitar información sobre COVID-19 a lo largo del mes de septiembre del año 2020.



**Figura 3.12.** Representación gráfica del número de llamadas para solicitar información sobre COVID-19 a lo largo del mes de septiembre del año 2020 y la predicción ARIMA en este periodo de tiempo.



---

## Conclusiones

Sobre el estudio estadístico realizado con las herramientas principales que hemos empleado, RStudio y Power BI, podemos concluir:

1. El CECOES es responsable de coordinar la respuesta a emergencias en las Islas Canarias, proporcionando asistencia inmediata y coordinando a los organismos de seguridad y emergencias. Su función principal es asegurar la eficiencia en la gestión de los servicios de emergencia.
2. Durante el año 2020, los meses de marzo y agosto se identificaron como períodos de alta carga de trabajo para los trabajadores del COCOES. Estos meses coincidieron con la aparición de la primera y segunda ola de la COVID-19, lo que generó una demanda significativa de atención y coordinación de emergencias. La carga adicional de trabajo experimentada durante estos meses puede atribuirse directamente al impacto de la pandemia en la sociedad y a la necesidad de una respuesta rápida y eficiente por parte de los profesionales del COCOES.
3. Debido a un evento excepcional relacionado con la pandemia, los datos presentan una distribución atípica que dificulta la realización de estimaciones precisas. La anormalidad en los datos, causada por la situación extraordinaria, ha generado desafíos significativos en la tarea de obtener estimaciones confiables.
4. La importancia de la estadística descriptiva radica en su capacidad para resumir y organizar los datos de la carga de trabajo, permitiendo obtener una visión clara y concisa de los subfactores y su relación, facilitando así el análisis y la toma de decisiones informadas.
5. Este trabajo se enmarca en los años de pandemia y plantea retos específicos, como analizar y comprender la dinámica de la "nueva normalidad"
6. Los datos muestran de forma consistente un aumento de indicadores de trabajo en marzo y agosto, concretamente en Demanda de trabajo, trabajo medio,

trabajo realizado antes, durante y después de la llamada, duración de las llamadas y duración de las llamadas por información sobre Coronavirus.

7. El aumento de los indicadores de demanda de trabajo está relacionada con un efecto principal del anuncio del estado de alarma de la pandemia; y los datos de agosto están relacionados con un efecto impulsado por un pico de incidencia en la provincia de Las Palmas
8. Los modelos de predicción de series temporales son precisos y confiables para pronosticar el número de llamadas de febrero a agosto, respaldando su utilidad en la planificación y gestión de actividades relacionadas con la gestión de llamadas.
9. Este trabajo destaca la importancia de analizar la carga de trabajo en el CECOES durante la pandemia, revelando patrones significativos relacionados con el estado de alarma y el aumento de incidencias en la provincia de Las Palmas.

Una limitación de este estudio es que se centró en análisis estadísticos descriptivos básicos, lo que limitó la exploración de relaciones más complejas entre variables. A pesar de ello, se observó que los modelos de predicción de series temporales demostraron ser precisos y confiables al pronosticar el número de llamadas entre los meses de febrero y agosto. Estos hallazgos respaldan la eficacia de los modelos de series temporales en la planificación y gestión de actividades relacionadas con la gestión de llamadas, incluso en situaciones complejas.

Futuras investigaciones podrían utilizar análisis estadísticos más avanzados para obtener una comprensión más profunda y extraer información adicional de los datos proporcionados. Esto permitiría identificar patrones más sutiles y relaciones subyacentes que podrían enriquecer el análisis de la carga de trabajo en el CECOES. El uso de técnicas de Deep Learning, como redes neuronales profundas, podría permitir descubrir patrones y relaciones más complejas en los datos, así como realizar pronósticos más precisos.

Este estudio enfocó únicamente en el análisis de datos de carga de trabajo del año 2020. Esto limita la capacidad de generalizar los hallazgos y comprender las variaciones a lo largo de diferentes años. Para obtener una visión más completa y robusta de la carga de trabajo en el CECOES, sería necesario incluir datos de múltiples años y realizar análisis comparativos a lo largo del tiempo. Esto permitiría identificar tendencias a largo plazo, evaluar cambios estacionales y considerar factores externos que podrían afectar la carga de trabajo en diferentes períodos.

Durante el desarrollo del estudio, uno de los desafíos fue la dificultad de predecir cambios imprevistos debido a la naturaleza novedosa de los eventos

ocurridos durante la pandemia. Sin embargo, se observó que los modelos de predicción de series temporales demostraron ser precisos y confiables al pronosticar el número de llamadas entre los meses de febrero y agosto. Estos hallazgos respaldan la eficacia de los modelos de series temporales en la planificación y gestión de actividades relacionadas con la gestión de llamadas, incluso en situaciones complejas.

A pesar de las limitaciones mencionadas, es importante destacar la utilidad de la estadística descriptiva en el análisis de la carga de trabajo en el CECOES. La estadística descriptiva nos permite resumir y organizar los datos de manera clara y concisa, permitiendo una comprensión inicial de los datos y la identificación de patrones y tendencias importantes, a pesar de las limitaciones y el potencial de análisis más complejos.



# A

---

## Apéndice

### A.1. Código en el lenguaje de programación R

Para una mayor comprensión del código puede visitar el siguiente enlace:

<https://github.com/alu0100814231/cecoes>





---

## Bibliografía

- [1] Aggarwal, Charu C. *Data mining*. Springer International Publishing, 2015.
- [2] Gobierno de Canarias. Informes de incidencia acumulada. Disponible en: <https://www.gobiernodecanarias.org/principal/coronavirus/semaforo/>.
- [3] Grömping, Ulrike. A Data Scientist's Guide to Acquiring, Cleaning and Managing Data in R. *Journal of Statistical Software*, 86(1), 2018. 1–4.
- [4] Microsoft Corporation. Microsoft Excel (2018). Disponible en: <https://office.microsoft.com/excel>.
- [5] Microsoft Corporation. Power BI (Versión 2.118) (2023). Disponible en: <https://powerbi.microsoft.com/>.
- [6] Roiger, Richard J. *Data Mining*. Chapman and Hall/CRC, 2017.
- [7] RStudio Team. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA (2020). Disponible en: <http://www.rstudio.com/>.
- [8] Tidy Forecasting in R: ARIMA Models: <https://joonim.medium.com/tidy-forecasting-in-r-arima-models-864dc13773c1>.
- [9] Hyndman, R. J., and Khandakar Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(1), 1–22.



# Study and workload analysis at the Emergency and Security Coordination Center (CECOES 1-1-2)

 **Sección de Matemáticas**  
Universidad de La Laguna

**Bianca Estefanía Kennedy Giménez**  
Facultad de Ciencias · Sección de Matemáticas  
Universidad de La Laguna  
alu0100814231@ull.edu.es

## Abstract

*This study analyzes the workload in the Coordinating Center for Emergencies and Security (CECOES) 1-1-2 of the Canary Islands Government. CECOES coordinates and responds to emergencies, providing immediate assistance and coordinating security and emergency agencies. The goal is to improve the efficiency of emergency services provided by CECOES. Descriptive statistics are used to analyze workload, including the number of calls received by operators, call durations, and workload differences among employees. Temporal patterns in workload are observed, which increase in March and August, possibly related to the state of emergency and COVID-19 incidence.*

## 1. Introduction

CECOES, a public service in the Canary archipelago, handles emergency calls and coordinates with security agencies. This study aims to evaluate CECOES's efficiency using workload indicators, applying descriptive statistics and ARIMA time series analysis on data like call volumes and durations. We employed data loading processes known as ETL (Extract, Transform, Load).

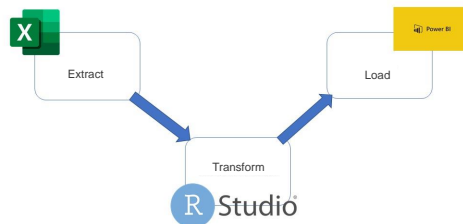


Figure 1: ETL processes diagram

## 2. Outline

The state of alert for the pandemic began in March, as evident from Figure 2, which shows an increase in the number of calls to the service. This indicates a heightened demand for assistance and information related to COVID-19 during that period.

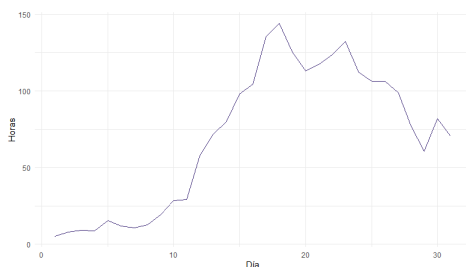


Figure 2: Duration of calls about Coronavirus per day in March 2020

By analyzing the work during calls and the total call duration, as well as considering the dynamics between different provinces, it is possible to observe indicators and trends in the work of CECOES workers. Additionally, the analysis can reveal distinctive patterns and dynamics specific to each province (see Figure 3).

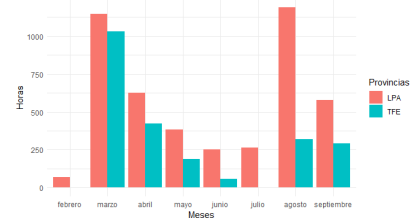


Figure 3: Duration of information calls about Coronavirus in 2020

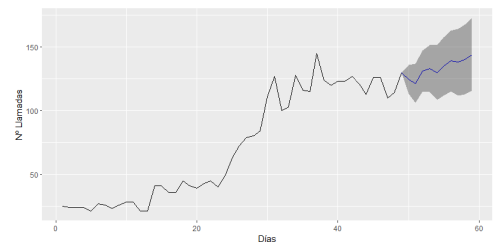


Figure 4: Prediction of the number of calls about Coronavirus.

## 3. Conclusion

This work emphasizes the importance of analyzing workload in CECOES during the pandemic, revealing significant patterns associated with the state of alarm during 2020.

## References

- [1] Roiger, Richard J. *Data Mining*. Chapman and Hall/CRC, 2017.
- [2] Informes de incidencia acumulada del Gobierno de Canarias. Available at: <https://www.gobiernodecanarias.org/principal/coronavirus/semaforo/>.