



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO

Grado en Estadística

Aplicación de técnicas estadísticas para el análisis de las incidencias en las líneas de autobuses urbanos de la empresa AUVASA

Autora: Sandra Farto González

**Tutores: Jesús Alberto Tapia García
Itziar Fernández Martínez**

Año 2023

Agradecimientos

Me gustaría dedicar unas palabras de agradecimiento a todas las personas que me han ayudado a realizar este Trabajo de Fin de Grado.

A mis tutores de la Universidad de Valladolid, Jesús Alberto Tapia García e Itziar Fernández Martínez, por la paciencia e interés que han mostrado hasta el último momento en la elaboración de este trabajo.

A mi tutor de la empresa AUVASA, José Alfonso Gálvez Salinas, por su accesibilidad y cercanía en todo momento y por brindarme la oportunidad de realizar un estudio con esta empresa tan importante para la ciudad de Valladolid.

A mi familia, en especial a mi madre, por su apoyo, comprensión y paciencia.

Por último, a todas aquellas personas que me han acompañado a lo largo de estos últimos años, compañeros, amigos y profesores.

Resumen

La empresa AUVASA (Autobuses Urbanos de Valladolid, S. A.) es la sociedad municipal encargada del transporte público urbano en Valladolid. En consonancia con su proceso de crecimiento y modernización mantiene una base de datos donde registra todas las incidencias que ocurren en el servicio de autobuses que presta a la ciudad.

La empresa puso a nuestra disposición, para la realización de este trabajo, dos ficheros Excel con las incidencias registradas desde febrero de 2022 hasta el 14 de junio de 2023. El objetivo es obtener la mayor información posible sobre el servicio e identificar patrones o casuísticas repetitivas que pudieran ser analizadas en detalle por la empresa.

La parte más compleja de este estudio ha sido la depuración de la base de datos mediante RStudio y el lenguaje de programación R. Esta fase de depuración ha permitido resumir estadísticamente las incidencias registradas a lo largo de los meses de los que se dispone información, estudiar algunos periodos del año en los que la empresa modifica sus servicios y establecer una relación entre las líneas con mayor número de usuarios y las que registran mayor número de incidencias. Además se ha elaborado un tipo de calendario (*HeatMap Calendar*) que proporciona a la empresa información visual sencilla sobre los incidentes registrados anualmente.

Abstract

The company AUVASA (Autobuses Urbanos de Valladolid, S.A.) is the municipal company responsible for of urban public transportation in Valladolid. As part of its ongoing growth and modernization efforts, it maintains a comprehensive database where to record all incidents that may arise in the bus services it offers to the city.

The company has provided us, with two Excel files containing the recorded incidents from February 2022 to June 15, 2023 to conduct a statistical study. The primary objective is to gather comprehensive information about the service and identify recurring patterns or specific cases that warrant further analysis by the company.

The most challenging aspect of this study involved debugging of the database using RStudio and the R programming language. This stage enabled us to conduct statistical summaries of the incidents registered during the available months, analyze specific periods of the year where the company modified its services and establish a relationship between the lines with the highest number of users and those that registered the highest number of incidents. In addition, we developed a *HeatMap Calendar* which provides the company with simple visually intuitive representation of the annual incident registrations. This calendar offers simplified visual information for easy understanding and analysis.

Índice

1. Introducción	6
1.1. Objetivos	6
1.2. Herramientas utilizadas	7
1.3. Metodología	8
1.4. Asignaturas del Grado de Estadística de especial relevancia para este trabajo	9
1.5. Estructura de la memoria	9
2. El sistema de autobuses urbanos de Valladolid	11
2.1. Líneas de autobús urbano	11
2.2. Flota y numeración	12
2.3. Otros datos de interés	12
3. Procedimiento de Depuración	14
3.1. Base de datos inicial	14
3.2. Depuración de la base de datos	16
4. Tratamiento estadístico de la base de datos	21
4.1. Análisis inicial de las incidencias	21
4.2. Análisis de las incidencias en determinadas épocas del año	26
4.2.1. Semana Santa	26
4.2.2. Periodo estival	29
4.2.3. Fiestas de la Virgen de San Lorenzo	31
4.2.4. Navidad	34
4.3. Análisis de incidencias por línea de autobús	36
4.4. Calendario anual de incidencias	42
5. Conclusiones y trabajos futuros	44
Bibliografía	46
6. ANEXOS	47
6.1. ANEXO I. Descripción detallada de las líneas de autobús urbano	47
6.2. ANEXO II. Flota de autobuses de AUVASA (junio 2023)	55
6.3. ANEXO III. Estadísticas de viajeros por líneas	58
6.4. ANEXO IV. Código en R para la elaboración de los gráficos.	59
6.5. ANEXO V. Código en R para la elaboración del calendario de incidencias usando la librería ggplot2.	62
6.6. ANEXO VI. Calendarios de 2022 para algunos tipos de incidencias	64
6.7. ANEXO VII. Tablas de frecuencias absolutas y condicionadas por año según las líneas de autobús.	67

Índice de figuras

1.	Imagen de registros de incidencias en base de datos excel	14
2.	Ejemplo de transformación en Tipo de Incidencia	17
3.	Ejemplo de incidencia T/A	18
4.	Ejemplo de incidencia T/V	19
5.	Diagrama de flujo para el proceso de depuración de la base de datos	20
6.	Incidencias mensuales con los datos registrados	22
7.	Gráfico de barras de las frecuencias relativas condicionadas al tipo de incidencia en cada año	24
8.	Gráfico de líneas en función del Tipo de Incidencia por mes	24
9.	Representación temporal de incidencias en los meses de 2022	25
10.	Total de Incidencias diarias en el mes de abril. Años 2022 y 2023	26
11.	Semana Santa. Días de la semana. Años 2022 y 2023	27
12.	Algunos tipos de Incidencias en el mes de abril por año	28
13.	Algunas categorías de la variable <i>Grupo</i> en el mes de abril	28
14.	Total de Incidencias diarias en el periodo estival. Año 2022	29
15.	Representación en periodo estival del número de Incidencias por Tipo de Incidencia	30
16.	Representación en periodo estival del número incidencias en algunas categorías de la variable <i>Grupo</i>	31
17.	Número de incidencias en el mes de septiembre de 2022	32
18.	Número total de incidencias por tipo de incidencias en el mes de septiembre	33
19.	Número total de incidencias en diversas categorías de la variable <i>Grupo</i> en el mes de septiembre	33
20.	Total de Incidencias diarias en diciembre de 2022 y enero de 2023	34
21.	Número de incidencias según Tipos de Incidencias en diciembre de 2022 y enero de 2023	35
22.	Número de incidencias en diversas categorías de la variable <i>Grupo</i> en diciembre de 2022 y enero de 2023	35
23.	Número de incidencias por Tipo de Línea salvo las Ordinarias	37
24.	Número de incidencias en las Líneas Ordinarias por mes según año	37
25.	Gráfico de barras de frecuencias relativas de incidencias por línea de autobús y año	39
26.	Total incidencias Grupo ACC en las 10 líneas Ordinarias más demandadas	39
27.	Total incidencias Grupo COM en las 10 líneas Ordinarias más demandadas	39
28.	Total incidencias Grupo PAR en las 10 líneas Ordinarias más demandadas	39
29.	Total incidencias Grupo PER en las 10 líneas Ordinarias más demandadas	39
30.	Total incidencias Grupo VEH en las 10 líneas Ordinarias más demandadas	40
31.	Total incidencias Grupo TIC en las 10 líneas Ordinarias más demandadas	40
32.	Frecuencias relativas de incidencias y viajeros para las Líneas Ordinarias. Año 2022	40
33.	Diagrama de dispersión. Año 2022	41
34.	Calendario de Incidencias en 2022	42
35.	Calendario de Incidencias en 2023	43
36.	Estadísticas de viajeros en el año 2022.	58
37.	Calendario de Incidencias tipo ACC en 2022.	64
38.	Calendario de Incidencias tipo PB en 2022.	65
39.	Calendario de Incidencias tipo PH en 2022.	66

Índice de tablas

1.	Clasificación y codificación de las incidencias del servicio de autobuses	15
2.	Tabla de frecuencias de incidencias según el mes y el año	21
3.	Tabla de frecuencias según el Tipo de Incidencia y Año	23
4.	Categorías de la variable Grupo para su representación	27
5.	Frecuencia relativa de incidencias condicionada al tipo de línea en cada año en los tipos de Líneas de autobús.	36
6.	Incidencias por líneas. Frecuencias absolutas y condicionadas. Parte 1	67
7.	Incidencias por líneas. Frecuencias absolutas y condicionadas. Parte 2	68
8.	Incidencias por líneas Ordinarias y Búho para los grupos de incidencias en 2022 Frecuencias absolutas.	69
9.	Incidencias por líneas Ordinarias y Búho para los grupos de incidencias en 2023 Frecuencias absolutas.	70

1. Introducción

La empresa AUVASA (Autobuses Urbanos de Valladolid, S. A.) es la sociedad municipal encargada del transporte público urbano en Valladolid. Desde su constitución, en el año 1982, su actividad y modernización ha ido creciendo en respuesta a las necesidades de los ciudadanos.

En el año 2022, AUVASA celebró sus 40 años de historia, presentando una nueva imagen de la empresa acorde a su proceso de renovación, modernización y actualización ofreciendo además nuevos servicios de movilidad a través de dos nuevas marcas: MOVASA, encargada de aparcamientos y ascensores, y BIKI, que es un servicio de alquiler de bicicletas.

En este trabajo se ha tenido muy presente el nuevo diseño de la empresa y se ha utilizado su actual código de colores en las distintas tablas y figuras que se han realizado.

Dentro de este proceso de modernización y mejora, la empresa, realiza un seguimiento de las incidencias en el transporte urbano (información sobre calles cortadas, siniestros en los vehículos, incidentes dentro del propio autobús, averías, mobiliario urbano defectuoso, bajas, horas extras, . . .) que ocurren y que registra en una base de datos. AUVASA ha manifestado su interés en explotar esa información con el objetivo de identificar patrones o casuísticas repetitivas, que una vez detectadas, serán analizadas en detalle por el Departamento de Inspección de la empresa.

Con este fin, AUVASA ha proporcionado dos archivos Excel con todos los registros de las incidencias desde febrero de 2022 hasta mediados de junio de 2023 con el fin de llevar a cabo un estudio estadístico. Esta base de datos está compuesta por registros en periodos de tiempos consecutivos y variables que describían los aspectos que la empresa consideraba más importantes, como la fecha de la incidencia, la línea de autobús involucrada, si se producían o no anomalías en el vehículo o en la línea, entre otras.

1.1. Objetivos

En este trabajo se ha realizado un análisis estadístico de las principales incidencias producidas en el servicio de transporte público de la ciudad de Valladolid entre febrero de 2022 y junio de 2023, con el objetivo general de identificar patrones o casuísticas repetitivas que puedan ser posteriormente evaluadas en detalle por la empresa gestora (AUVASA) con el fin de incrementar la calidad del servicio que presta a los ciudadanos.

Como objetivos específicos podemos enumerar:

1. Integrar toda la información de las dos bases de datos proporcionadas por la empresa en un único *data.frame* de R.
2. Depurar la base de datos.
3. Resumir estadísticamente las principales incidencias registradas y analizar el comportamiento en algunos periodos del año en los que AUVASA adapta el servicio que presta a la ciudad puesto que varía la demanda de usuarios (Navidad, Semana Santa, meses de verano y Ferias y Fiestas de San Lorenzo).

4. Estudiar con más detalle las incidencias en las líneas de autobús que registran mayor número de ellas.
5. Diseñar un calendario (*HeatMap Calendar*) que registre las incidencias a lo largo del año y que permita una consulta rápida y visual a la empresa.
6. Proponer recomendaciones a AUVASA de cara a recopilar la información de forma más eficiente para facilitar su análisis estadístico.

1.2. Herramientas utilizadas

A continuación se enumera el software y las principales librerías utilizadas que se ha utilizado en este trabajo:

- **R**: es un entorno y lenguaje de programación interpretado con un claro enfoque estadístico. Ampliamente utilizado, permite la posibilidad de cargar diferentes bibliotecas o paquetes proporcionando un amplio abanico de herramientas estadísticas (modelos lineales y no lineales, tests estadísticos, análisis de series temporales, algoritmos de clasificación, etc.) y la posibilidad de realizar gráficos profesionales. Se trata de software libre que se distribuye bajo la licencia GNU (General Public License), que es una licencia de derecho de autor usada en el mundo del software libre y código abierto y garantiza a los usuarios finales la libertad de usar, estudiar, compartir (copiar) y modificar el software) y está disponible para los sistemas operativos más comunes. Además, puede integrarse con distintas bases de datos y con bibliotecas que facilitan su uso desde lenguajes de programación interpretados como Perl y Python. Véase r-project.org [1], Grolemond y Wickham [2] o Maindonald y Braun [3] para más información. Para llevar a cabo este trabajo se han utilizado las siguientes librerías:
 - *readxl*: nos permite, entre otras cosas, importar ficheros Excel a R.
 - *knitr*: utilizada para exportar tablas en \LaTeX .
 - *lubridate*: posee funciones que permiten la manipulación de fechas.
 - *ggplot2*: es una librería que permite la creación de una gran variedad de gráficos siguiendo la llamada *The Grammar of Graphics*.
- **RStudio**: es un entorno de desarrollo integrado que permite programar, depurar y ejecutar código en R enfocado a la computación estadística y gráficos. Se trata también de software libre que se distribuye bajo la licencia AGPL (Affero General Public License) y, al igual que en el caso de R, está disponible para los sistemas operativos más comunes (docs.posit.co [4]).
- **Overleaf**: es un editor online utilizado para escribir, editar y publicar documentos científicos con \LaTeX . Este último, es un procesador de textos orientado a la creación de documentos escritos con una alta calidad tipográfica. Véase overleaf.com [5] y latex-project.org [6].

1.3. Metodología

Para realizar este trabajo se han seguido varios pasos o etapas en cada uno de los cuales se ha aplicado una metodología concreta que comentaremos seguidamente.

- Etapa 1. Elección del software de trabajo.

Se valoró trabajar con uno de estos dos lenguajes: Python o R. Por un lado, Python es uno de los lenguajes de programación cada vez más utilizados para análisis estadísticos con librerías como Pandas, Numpy, PLoty, Scikit Learn, etc, de gran utilidad para este tipo de trabajos.

Por otro lado R es un entorno y lenguaje de programación claramente enfocado al análisis estadístico y utilizado desde los años 90. Este lenguaje se ha utilizado en la gran mayoría de las asignaturas de la carrera y este hecho llevó a elegir R.

- Etapa 2. Integración y depuración de la base de datos.

La fase de depuración de datos es imprescindible para realizar un buen análisis estadístico y tiene tres funciones principales: facilitar el procesado de los datos, mejorar la calidad eliminando errores y proporcionar información para medir la calidad de la base de datos.

Se pueden distinguir dos métodos para la depuración de datos (Universidad de Granada [7]) :

- El enfoque tradicional o microdepuración: consiste en la detección/corrección de los errores en todos los registros.
- La macrodepuración: consiste en hacer una selección de unidades influyentes, de valores influyentes o de ambos y una corrección de sus posibles errores. Los criterios que determinan esa influencia son la representatividad del registro en todos los datos (su peso) o la magnitud de los valores de los campos en los registros.

Además, existen diferentes técnicas y algoritmos: microdepuración, macrodepuración, depuración selectiva, depuración interactiva y depuración automática (véase Waal et al. [8], Revilla [9]).

Este trabajo se ha realizado siguiendo el método tradicional de microdepuración. Se han revisado los datos recogidos, con el fin de detectar y corregir los errores y dar un tratamiento adecuado a la falta de respuesta. Este proceso constituye una de las tareas más costosas y complejas del análisis estadístico y en este estudio ha supuesto varios meses de trabajo.

La metodología a seguir cuando se localiza un registro conflictivo es la siguiente:

- Consultar con fuente de información para obtener el valor correcto o para aclarar informaciones ambiguas.
- Ignorar instancias con valores sospechosos.
- Ajustar valores dudosos que proporcionaban conclusiones inconsistentes.
- Asignar a registros con falta de información un valor consistente (corrección manual).

- Etapa 3. El tratamiento estadístico de los datos se ha realizado mediante tablas y técnicas gráficas de estadística descriptiva. Entre los gráficos utilizados destacamos:
 - Gráfico de líneas. Es un tipo de gráfico estadístico que sirve para representar gráficamente una serie temporal.
 - Diagrama de barras. Es un tipo de gráfico que se utiliza para representar las frecuencias de cada categoría de una variable cualitativa.
 - Heatmap o mapa de calor. Es una representación gráfica que permite la visualización de datos a través de colores en dos dimensiones. La variación del color puede ser por tono o intensidad, haciendo más sencillo y atractivo el entendimiento de los datos.

En este trabajo se ha utilizado este tipo de representaciones para elaborar un calendario para la empresa con las incidencias del transporte urbano que permita trasladar a los empleados la información que considere necesaria de forma rápida e intuitiva. Así mismo, dado que es una empresa de carácter municipal, pueden usarse algunos de estos gráficos como información pública.

Dado que la visualización de datos es una parte muy importante que nos permite resumir la información y extraer conclusiones, es importante que los gráficos sean capaces de captar la atención del usuario. Uno de los aspectos que nos puede ayudar es elegir una paleta adecuada de colores. En este trabajo se ha cuidado la elección de colores de forma que usen como base los colores de la empresa y armonicen con estos.

1.4. Asignaturas del Grado de Estadística de especial relevancia para este trabajo

Para la realización de este trabajo han resultado imprescindibles muchos de los conocimientos adquiridos en los estudios de Grado, citar las asignaturas de *Estadística Descriptiva* y *Computación Estadística*. También, asignaturas de Informática como *Fundamentos de informática* y *Análisis y diseño de bases de datos*.

Otras asignaturas imprescindibles para ampliar el estudio recogido en este trabajo a medida que la empresa vaya disponiendo de más datos son *Series Temporales*, *Inferencia Estadística* y *Técnicas de Aprendizaje Automático*.

1.5. Estructura de la memoria

La memoria consta de cinco secciones:

- La presente sección (Sección 1) además de ser de carácter introductorio, presenta las herramientas que se han utilizado para el trabajo así como la metodología seguida para el mismo.
- En la Sección 2 se proporciona la información básica de la empresa relativa al transporte urbano de viajeros. Se enumeran brevemente las líneas de autobuses urbanos

(se amplía la información en el ANEXO I), se aportan datos sobre la flota de vehículos que posee la empresa (ANEXO II) y se comentan otros datos de interés que se encuentran registrados en la base de datos.

- En la Sección 3 se procede a la integración y depuración de la base de datos. Se describen las variables consideradas para el estudio y se detalla el proceso seguido para la depuración de datos.
- En la Sección 4 se realiza el tratamiento estadístico de los datos analizando, con más detalle, ciertas épocas del año en las cuales la empresa modifica su flota y horarios. También, se analizan las incidencias por la líneas de autobús.
- En la Sección 5 se presentan las conclusiones del trabajo realizado así como propuestas de trabajo futuro.

2. El sistema de autobuses urbanos de Valladolid

La empresa AUVASA (Autobuses Urbanos de Valladolid, S. A.) es una sociedad municipal, creada en 1982, y la encargada del transporte público dentro del término municipal de Valladolid. También presta sus servicios en los municipios de Simancas y La Cistérniga, vinculados por dos contratos.

La empresa lleva un registro en una base de datos con la información de las incidencias acaecidas a lo largo de cada jornada. Cada incidencia es un registro de la base de datos.

Seguidamente, veremos con más detalle cierta información sobre la empresa que será imprescindible para describir las variables que se manejan en este estudio.

2.1. Líneas de autobús urbano

Actualmente, hay 20 líneas Ordinarias de autobús (2 de ellas circulares), 9 líneas a polígonos industriales, 2 líneas lanzadera al Campus Universitario Miguel Delibes, 7 líneas de servicio especial matinal, 5 de nocturno (Búho), 6 líneas Fútbol que dan servicio al estadio José Zorrilla en los días de partido, 5 líneas especiales para diferentes ferias u otros eventos culturales al Real de la Feria, 2 líneas con otros servicios y 3 con servicios especiales asociadas a determinadas fechas y eventos. A continuación se destaca la información más relevante de cada una de estas líneas. La información completa sobre la frecuencia de las líneas puede consultarse en el Anexo I.

– **Líneas Ordinarias.**

Líneas: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 16, 17,18, 19, 24, C1, C2, H.

- **Líneas Búho:** prestan servicio nocturno viernes, sábados y vísperas de festivo (salvo en Nochebuena y Nochevieja) a partir de las 23:30 y hasta las 3:00 horas. Durante las fiestas de la ciudad de Valladolid (primera semana de septiembre) se da servicio todos los días.

Líneas: B1, B2, B3, B4, B5.

- **Líneas a polígonos industriales:** complementan las líneas Ordinarias para dar servicio a las dos grandes áreas industriales de Valladolid.

Líneas: P1, P2, P3, P6, P7, P13, PSC1, PSC3 San Pedro Regalado–Polígono Argales–La Rubia.

- **Líneas matinales.** Las líneas de la red matinal replican parte de los recorridos del servicio ordinario desde la periferia de Valladolid. Salen de la parada de origen a las 6:50 horas en dirección al centro de la ciudad de lunes a viernes no festivos. Solo tienen recorrido de ida.

Líneas: M1, M2, M3, M4, M5, M6, M7.

- **Líneas universitarias.** Se trata de 2 líneas lanzadera que recorren los ejes Sur–Norte y Suroeste–Norte a las 7:20 y 8:20 horas en dirección al centro de la ciudad y

el Campus Miguel Delibes de lunes a viernes lectivos, según el calendario académico de la Universidad de Valladolid. Estas líneas sólo tienen recorrido de ida.

Líneas : U1, U8.

- **Líneas especiales.** Hay dos tipos de líneas, Líneas F (fútbol) y Líneas R (que llegan al recinto ferial la semana de Ferias de la ciudad).

Líneas F: F1, F2, F3, F4, F5, F6.

Líneas R: R, R1, R2, R3, R4, R5.

- **Otros servicios.** Añadidos recientemente para cubrir algunos recorridos.

Líneas: 23, 26.

- **Otros servicios especiales.** El Día de Todos los Santos entran en servicio dos líneas especiales a los cementerios vallisoletanos de Las Contiendas y de El Carmen, con recorridos radiales desde el centro de la ciudad. También, en la concentración motera Pingüinos, se establece una línea lanzadera desde la zona centro con parada en varios puntos de la ciudad. La línea Cementerio El Carmen, La Línea Cementerio de Las Contiendas, La línea Concentración motera Pingüinos.

- **Bus turístico.** Valladolid cuenta con un autobús turístico que tiene una duración aproximada de 50 minutos.

2.2. Flota y numeración

A fecha de junio de 2023, AUVASA posee una flota de 155 autobuses (si tenemos en cuenta el autobús turístico) a los que se les asigna una numeración que podemos consultar en la tabla que figura en el Anexo II de esta memoria y donde también se encuentran las líneas a las que habitualmente se asignan los vehícululos y sus características técnicas.

La abreviatura RR referida a un vehículo indica que no está asignado a una línea ordinaria y su servicio es reforzar estas líneas, sustituir bajas o bien estar como vehículo de reserva en puntos estratégicos de la ciudad. El objetivo de estos vehículos es garantizar siempre la mejor oferta de servicio al ciudadano.

2.3. Otros datos de interés

De cara a nuestro estudio también debemos tener en cuenta:

- **Conductores.** Cada conductor de la empresa tiene asignado un número que es su identificador y que le representa de forma unívoca en AUVASA (aunque en este trabajo se dispone de esta información no será mostrada puesto que se trata de un dato sensible que esta protegido por el principio de confidencialidad).

- **Inspectores.** Son trabajadores que registran las incidencias de los vehículos, las rutas y los conductores de las mismas. Al igual que los conductores tienen un identificador único (que tampoco se muestra por la misma razón que para el caso de los conductores). También son los responsables del control de servicio (tanto fuera como dentro de un autobús) y de la inspección de la infraestructura (paradas y autobuses).
- **Número de parada y calle.** Como descriptores básicos de las incidencias se registran los números de parada, calle, entre otros datos. Por esta razón, la base de datos de la empresa tiene una variable denominada *número de parada*. Toda la información acerca de estos números de parada y las calles la podemos encontrar en la siguiente dirección web:

http://www.auvasa.es/auv_mapaparadas.asp.
- **Número de viajeros.** El número de usuarios en el año 2022 creció con respecto a 2021 superando los 1.400.000 viajeros (Anexo III). Evidentemente, no se dispone de todos los datos para 2023 y por esta razón no se aportan.

3. Procedimiento de Depuración

Los datos proporcionados por la empresa constituyen el pilar fundamental sobre el que se apoya este trabajo. La información utilizada a lo largo del proyecto procede del seguimiento de las incidencias en las líneas de autobús en el periodo comprendido desde febrero de 2022 hasta el 14 de junio de 2023. En esta sección se explica el proceso de depuración de datos que se ha seguido.

3.1. Base de datos inicial

Los dos ficheros Excel de la base de datos (con formato xlsx) que nos proporcionó AUVA-SA mostraban información sobre los incidentes ocurridos en las líneas de autobús urbano. En concreto, eran ficheros que contenían registros de periodos de tiempos consecutivos. El primero contenía 3614 registros desde el 2 de febrero de 2022 hasta el 11 de noviembre de 2022. El segundo constaba de 6366 instancias y abarcaba desde el 2 de febrero de 2023 hasta el 14 de junio de 2023.

Submission Date	Source App	I.C.P.	Fecha	Tipo de Incidencia	Grupo	Descripción	Conductor (nº)	Conductor (Nombre y Apellido)	Vehículo nº	Línea	Archivos adjuntos	OBSERVACIONES	Parada	Submission IP	Last Update Date	Submission ID
11/11/2022	xxx-xxxx	11/10/2022	V - Comunicaciones internas, anomalías en línea, servicio, etc.	COM	Manifestación Plaza del Carmen. Entre las horas indicadas se desarrolla la manifestación por las PZ contra.	999								80.28.123.193		5439632103918793447
11/11/2022	xxx-xxxx	11/10/2022	T - Anomalías en vehículos	VEH	Llama el conductor arriba indicado para comunicar que un turismo ha golpeado el PN Zorilla.	xxx	xxxx xxxx		222	1				80.28.123.193		5439610153917457321
11/11/2022	xxx-xxxx	11/10/2022	A - Accidentes y siniestros	ACC	Llama el conductor arriba indicado para comunicar que un turismo ha golpeado el incorporación en jornada de trabajo.	xxx	xxxx xxxx		222	1				80.28.123.193		5439609793919867043
11/10/2022	xxx-xxxx	11/09/2022	PH - Personal en servicio, prolongación jornada, horas extra, etc.	PER	El conductor indicado se incorpora a las 15:30.	xxx	xxxx xxxx				https://www.joinfor.m.com/uploads/get_drow/2204712941			80.28.123.193		5438742293917650082
11/10/2022	xxx-xxxx	11/09/2022	V - Comunicaciones internas, anomalías en línea, servicio, etc.	COM	Se detectan fallos en la plantilla. Se comunican.	999								80.28.123.193		5438741233917019752
10/31/2022	xxx-xxxx	10/29/2022	TV - Anomalías en el servicio por vehículos.	VEH	Noo Bata el conductor q xxx C/ DARSENA / VILLANUBLA.	xxx	xxxx xxxx		232	1			838 Plaza Castilla y León 1	80.28.123.193		5430296343917278472
09/26/2022	I20-Anton	09/25/2022	T - Anomalías en vehículos	VEH	Se recibe una llamada	103	Juan Carlos de Vega Paune	64		6				80.28.123.193	sept. 27, 2022	5399844753911701951

Figura 1: Imagen de registros de incidencias en base de datos excel

En la Figura 1 podemos observar una imagen de algunas de las incidencias registradas procedentes de uno de los dos ficheros de datos. En la imagen, y por motivos de confidencialidad, se han sustituido con la letra X tanto los identificadores únicos del personal como sus nombres. En esta figura vemos que la información registrada para cada incidencia es la siguiente:

- **Submission Date:** fecha en la que se ha registrado la incidencia.
- **SourceApp:** columna que se encontraba vacía.
- **I.C.P.:** inspector que ha registrado la incidencia.
- **Fecha:** se indica el mes/día/año en el que se ha producido la incidencia.

- **Tipo de Incidencia:** esta variable representa la clasificación en tipos de incidencias que se ha utilizado en la base de datos. En la Tabla 1 se recogen todos tipos de incidencias así como la codificación usada en la base de datos.

CÓDIGO	DESCRIPCIÓN
A	Accidentes y siniestros: colisiones, incidentes con pasajeros, etc
C	Anomalías en el Billetaje
D	Partes diarios Subjefes
G	Software SAE, puesto GPS: fallos en pantallas de comunicaciones, etc
I	Máquinas embarcadas, software embarcado, software de letreros, tarjetas retiradas, etc
M	Mobiliario Urbano: roturas de marquesinas, paneles de tiempo defectuosos, etc
PB	Personal en servicio, absentismos, bajas, etc.
PH	Personal en servicio: prolongación jornada, horas extra, etc.
T	Anomalías en vehículos: lunas rotas, espejos retrovisores, rampas que no descienden, etc.
Tr.I	Partes de trabajo de inspección
TV	Anomalías en el servicio por vehículos: retrasos, etc.
U	Usuarios: quejas y reclamaciones
V	Comunicaciones internas, anomalías en líneas y servicios, cortes de calles, etc
VP	Incidencias de conductores: equivocaciones al tomar calles, etc

Tabla 1: Clasificación y codificación de las incidencias del servicio de autobuses

- **Grupo:** es una variable categórica que utiliza la empresa y que en algunas de sus categorías agrupa varios tipos de incidencias, pero en otras categorías se corresponde sólo con una de las ya existentes.
 - ACC: igual que la categoría A de la variable Tipo de Incidencia.
 - COM: igual que la categoría V.
 - PAR: igual que la categoría M.
 - PER: que comprende los tipos PB, PH y VP.
 - VEH: abarca los tipos T y TV.
 - TIC: agrupa los tipos C, G e I.
- **Descripción:** el inspector correspondiente describe la incidencia que ha tenido lugar.
- **Conductor (nº):** se corresponde con el número de conductor relacionado con esa incidencia.

- **Conductor (Nombre y Apellidos):** que no se mostrará por la Ley de Protección de Datos y por cuestiones de confidencialidad.
- **Vehículo n°:** el número de vehículo involucrado en una incidencia.
- **Línea:** a la que se vincula esa incidencia.
- **Archivos adjuntos:** que contiene la URL a fotos o documentos que el inspector considera aconsejable añadir.
- **Observaciones:** son informaciones descritas a mayores si el inspector lo ha considerado oportuno.
- **Parada:** número y dirección de la parada de autobús.
- **Submisión IP:** es la dirección IP del ordenador desde el que se registró el incidente.
- **Submisión ID:** dato que proporciona automáticamente la aplicación que tiene la empresa.

3.2. Depuración de la base de datos

Para el proceso de depuración se procedió primero a importar los datos con RStudio (las hojas de cálculo de Excel no son la herramienta más adecuada para el tratamiento de datos, su depuración y su posterior visualización gráfica) y se llevó a cabo de forma exhaustiva, con el fin de garantizar la calidad del análisis, siguiendo el enfoque tradicional. Se observó que todas las variables estaban guardadas como tipo *character*. Los pasos que se dieron fueron los siguientes:

1. Se eliminaron las variables sin datos: *SourceApp* y *OBSERVACIONES*.
2. Se procedió a revisar cada una de las variables con el fin de identificar otras situaciones que pudiesen ser conflictivas: errores tipográficos, datos incoherentes (por ejemplo fechas de incidentes posteriores a sus fechas de registro), etc. Estas variables son:
 - 2.1. *Submission Date*. Se procedió a transformar esta variable de tipo carácter a la clase *Date* de R. Se comprobó que no hubiera fechas de años diferentes a 2022 o 2023.
 - 2.2. *I.C.P.* Se eliminó el nombre del inspector manteniendo únicamente el identificador de cara a un mejor manejo de la base de datos y a mantener la confidencialidad del personal.
 - 2.3. *Fecha*. Se convirtieron a la clase *Date* de R y se comprobó que los años se hubieran introducido correctamente. Se encontraron 5 registros con años incorrectos. Se corrigieron tres de ellos y se procedió a la eliminación de otros dos.
 - 2.4. *Tipo de Incidencia*. Se comprobó que los registros hiciesen referencia a las distintas categorías definidas por la empresa, eliminando la descripción que aparecía junto a ellas de cara a hacer más rápido el análisis (véase Figura 2).

Submission Date	I.C.P.	Fecha	Tipo de Incidencia	Grupo	
12-22-2022	XXX	21/12/2022	A - Accidentes y siniestros	ACC	A – 21/12/20 Comunica qu realizar los p realice el par
12-22-2022	XXX	19/12/2022	M - Mobiliario Urbano	PAR	Paseo Lorenz El bus se ha t recoger viaje compensar h

↓

Submission Date	I.C.P.	Fecha	Tipo de Incidencia	Grupo	
12-22-2022	XXX	21/12/2022	A	ACC	A – 21/12/20 Comunica qu realizar los p realice el par
12-22-2022	XXX	19/12/2022	M	PAR	Paseo Lorenz El bus se ha t recoger viaje compensar h

Figura 2: Ejemplo de transformación en Tipo de Incidencia

- 2.5. *Grupo*. Se comprobó que los formatos de la variable fueran correctos.
- 2.6. *Descripción*. Se eliminaron aquellos registros que no tenían apenas información y en los que, además, esta variable descripción estaba vacía.
- 2.7. *Conductor (n°)*. Esta variable debe constar de uno, dos o tres dígitos. Al encontrar observaciones diferentes procedimos a preguntar al personal encargado en AUVASA. Se nos informó de que esos casos representaban imposibilidad de identificar al conductor con lo cual se modificaron esos datos introduciendo el valor -1 y se cambió el tipo de variable a numérica.
- También había valores numéricos como 0, 000 y 999 que son valores que no se corresponden con ningún conductor. Tras un análisis de la descripción de estos casos se observó que algunas de esas incidencias o no estaban asociadas a un conductor o se correspondían con algún apoyo a conductores, descansos, etc y se mantuvo su valor.
- 2.8. *Conductor (Nombre y Apellidos)*. Se utilizó esta variable para comprobar la coherencia con la variable Conductor (n°) tras lo cual se eliminó la variable.
- 2.9. *Vehículo n°*. Los valores de esta variable varían entre 1 y 611. El valor 999 se emplea, de nuevo, bien para indicar ausencia de información o para otro tipo de sucesos.
- 2.10. *Línea*. Se observaron datos con valores que no pertenecían a ninguna de las líneas publicitadas por la empresa. Tras preguntar al personal encargado, se constató que algunos de esos valores hacían referencia a turnos, servicios de personal, etc. Otros valores eran errores en la recogida de datos (como $L - 1$ en lugar de 1) y se sustituyeron por el valor correcto. Además, se separaron aquellos registros en los que la variable *Línea* contenía como valor el de los turnos de trabajo que se guardaron en una nueva variable llamada *Turno*.

También se crea una nueva variable categórica, *TipoLinea*, con el fin de tener una primera aproximación a los tipos de líneas con más incidencias.

La variable *TipoLinea* puede tomar los valores: ordinaria, Búho, matinal, ferial, fútbol, otros servicios, polígono, refuerzo, servicio personal, universitaria y múltiple. Casi todos estos valores se corresponden con los tipos de líneas que oferta la empresa, salvo servicio personal y múltiple. *Servicio personal* agrupa aquellas instancias que tienen en la variable *Línea* valores del tipo EN, siendo N un número (Ejemplo: E2, E11, E22...) y que la empresa indicó que se refería a servicios de personal (se utiliza para trasladar al personal de la empresa).

Múltiple se utiliza en aquellos registros que en la variable *Línea* tienen múltiples líneas (por ejemplo contiene 17, 19, P3) y además son de distinto tipo (en el ejemplo inmediatamente anterior dos Ordinarias y una de polígonos).

- 2.11. *Archivos adjuntos*. Se eliminó esta variable ya que cuando tiene valor no nulo contiene enlaces a archivos adjuntos que suelen ser fotografías y no se tiene acceso a ellos.
- 2.12. *Parada*. Se comprobó exclusivamente que los valores estuvieran escritos correctamente.
- 2.13. *Last Update Date*. Es una variable prácticamente vacía en el primer fichero Excel y totalmente vacía en el segundo, por ello, se procedió a su eliminación.
- 2.14. *Submission IP* y *Submission ID* son variables que se eliminaron ya que no se han utilizado en nuestro análisis los datos.

Submission Date	I.C.P.	Fecha	Tipo de Incidencia	Grupo	Descripción	Conductor (nº)	Vehículo nº	Línea	Parada
2022-03-07	XXX	04/03/2022	T	VEH	C.Málaga/Pso.San Isidro. El ICP 119 comunica en el pasillo que este bus de esta línea ha tenido un accidente con un coche y se encuentra tirando refrigerante. Nos dice que está taller avisado. Se manda tras iniciar el servicio a inspección desde la central, pues desconocemos más datos	XXX	114	PSC2	
2022-03-07	XXX	04/03/2022	A	ACC	C.Málaga/Pso.San Isidro. El ICP 119 comunica en el pasillo que este bus de esta línea ha tenido un accidente con un coche y se encuentra tirando refrigerante. Nos dice que está taller avisado. Se manda tras iniciar el servicio a inspección desde la central, pues desconocemos más datos	XXX	114	PSC2	

Figura 3: Ejemplo de incidencia T/A

- 3. En este punto se inició la búsqueda de datos duplicados. Primero, se identificaron incidencias exactamente iguales, eliminando sus duplicados. Después, se detectaron casos de incidencias casi iguales, en las que una de ellas amplía la información proporcionada por la otra. En total, se eliminaron de la base de datos 87 registros.

Durante este proceso se detectaron también registros que sólo se diferenciaban en dos variables: *Tipo de Incidencia* y *Grupo*. Tras consultarlo con la empresa se constató que se trataba de incidentes que se correspondían con varios tipos de incidencias distintas y que se registraban varias veces con dos o más códigos diferentes. En la Figura 3 se pueden observar dos incidencias que sólo se diferencian en el valor que toma la variable *Tipo de Incidencia*. En este caso se produce un accidente (A) y además tiene consecuencias en la información que la empresa debe comunicar a taller (T).

Submission Date	I.C.P.	Fecha	Tipo de Incidencia	Grupo	Descripción	Conductor (nº)	Vehículo nº	Línea	Parada
2022-04-08	XXX	07/04/2022	V	COM	9:15H- 8228 -LS - CXX El conductor arriba indicado comunica que al pasar por la Avenida del Valle del Egueva dirección Belén a la altura del Paseo del Cauce oye el impacto de una piedra que han tirado al autobús que rompe la primera luna lateral del lado derecho. Cuando mira el conductor no ve a la persona que ha tirado la piedra. Se avisa a policía que acude para hablar con el conductor	XX	228	8	679 Avenida Valle Egueva 41 frente Económicas
2022-04-08	XXX	07/04/2022	T	VEH	El conductor arriba indicado comunica que al pasar por la Avenida del Valle del Egueva dirección Belén a la altura del Paseo del Cauce oye el impacto de una piedra que han tirado al autobús que rompe la primera luna lateral del lado derecho. Cuando mira el conductor no ve a la persona que ha tirado la piedra. Se avisa a policía que acude para hablar con el conductor	XX	228	8	679 Avenida Valle Egueva 41 frente Económicas
2022-04-08	XXX	07/04/2022	T	VEH	El conductor arriba indicado comunica que al pasar por la Avenida del Valle del Egueva dirección Belén a la altura del Paseo del Cauce oye el impacto de una piedra que han tirado al autobús que rompe la primera luna lateral del lado derecho. Cuando mira el conductor no ve a la persona que ha tirado la piedra. Se avisa a policía que acude para hablar con el conductor	XX	N/A	N/A	N/A

Figura 4: Ejemplo de incidencia T/V

De cara a un análisis más ágil, se asignó a todas estas incidencias una nueva categoría para la variable con el código T/A y se eliminaron los elementos que quedaban duplicados. Encontramos un total de 138 instancias de tipo T/A.

La situación anterior también se reproducía con diferentes valores de la variable *Tipo de Incidencia* por ello se permitieron los siguientes nuevos valores para esta variable: T/A, A/V, PB/V, PB/VP, PH/VP, T/VP, TV/V, TV/VP, V/VP, A/T/M, A/T/VP.

Además se encontraron varios registros para una misma incidencia con distinto Tipo (código) pese a haber una categoría que englobaba estas situaciones. Puede verse un ejemplo en la Figura 4. A mayores, algunos de esos registros sólo diferían del análogo en una nueva frase añadida en la descripción de otro. Un ejemplo de esta situación se observa en la Figura 4. Todos estos casos se revisaron uno a uno ya que no suponían un gran número de instancias (unos 30 registros).

En la Figura 5 podemos observar el diagrama de flujo que describe el proceso de depuración completo.

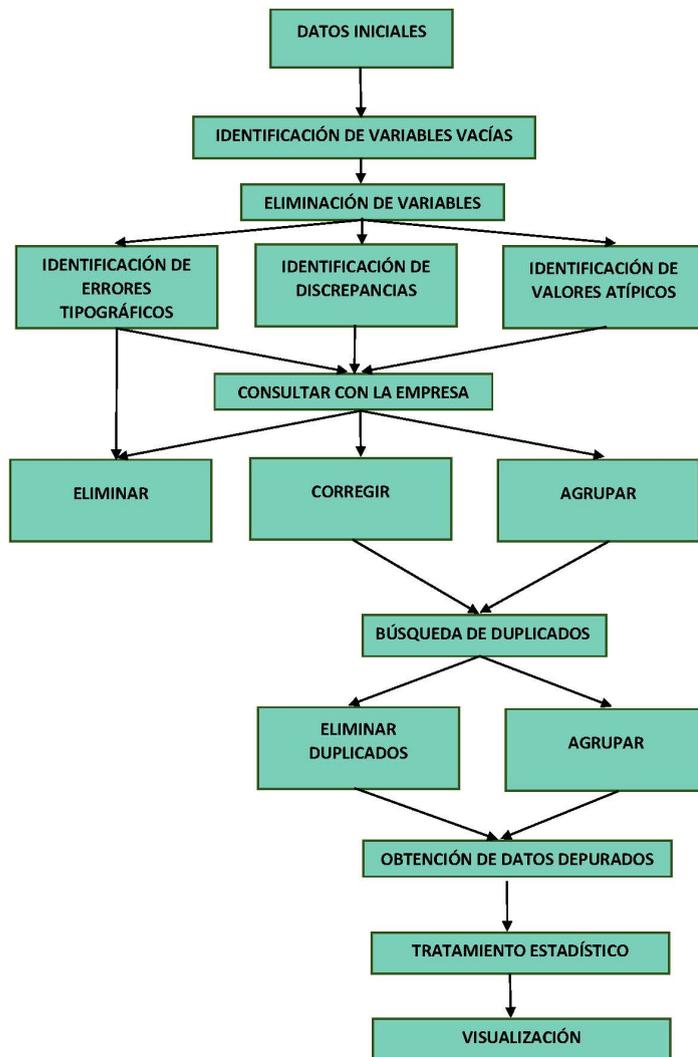


Figura 5: Diagrama de flujo para el proceso de depuración de la base de datos

4. Tratamiento estadístico de la base de datos

Una vez que disponemos de los datos ya depurados, se ha procedido a realizar un análisis estadístico mediante distintos métodos descriptivos con la finalidad de organizar, sintetizar y describir la información utilizando tablas, gráficos y otras herramientas estadísticas que faciliten la comprensión de la información con el objetivo de identificar tendencias, patrones, relaciones, correlaciones, etc, y extraer conclusiones útiles para la empresa (Fernández *et al.* [10], Fernández–Cuesta [11] y Montero [12]).

4.1. Análisis inicial de las incidencias

Se ha comenzado realizando una tabla de frecuencias (Tabla 2) donde se ha representado el número de incidencias a lo largo de cada mes registrado de los años 2022 y 2023 (no se dispone de los datos de enero de 2022 y sólo se tienen los datos de casi la primera mitad de 2023 lo que produce valores nulos en la tabla).

Los datos de esta tabla se ha representado gráficamente en la Figura 6. Puesto que en el mes de febrero de 2022 sólo disponemos de dos incidencias registradas y no se tiene de más información, se han eliminado los datos de este mes.

MES	2022	2023	Total
enero	0	374	374
febrero	2	353	355
marzo	467	367	834
abril	406	301	707
mayo	469	451	920
junio	378	242	620
julio	296	0	296
agosto	318	0	318
septiembre	489	0	489
octubre	399	0	399
noviembre	393	0	393
diciembre	354	0	354
Total	3971	2088	6059

Tabla 2: Tabla de frecuencias de incidencias según el mes y el año

En el año 2022, se alcanzan mínimos relativos en los meses de abril y julio. Analizando con más detalle el crecimiento de la gráfica y el calendario de dicho año, se comprueba que estos meses con menor número de incidencias coinciden con Semana Santa (del 8 al 17 de abril de 2022) y con el periodo estival. De hecho, en el mes de julio de 2022 las incidencias alcanzan el mínimo absoluto de ese año. Tras estos periodos comentados previamente se produce, de nuevo, un incremento de los incidentes que parece coincidir con el regreso de las vacaciones llegando a alcanzarse el máximo en septiembre de 2022, seguido de cerca por los máximos relativos de los meses de marzo y mayo.

En el año 2023, en los meses comprendidos entre enero y mayo, parece reproducirse el comportamiento del mes de abril donde encontramos un mínimo en el número de inci-

dencias (de nuevo en este mes volvió a celebrarse la Semana Santa) para después volver a crecer en el mes de mayo.

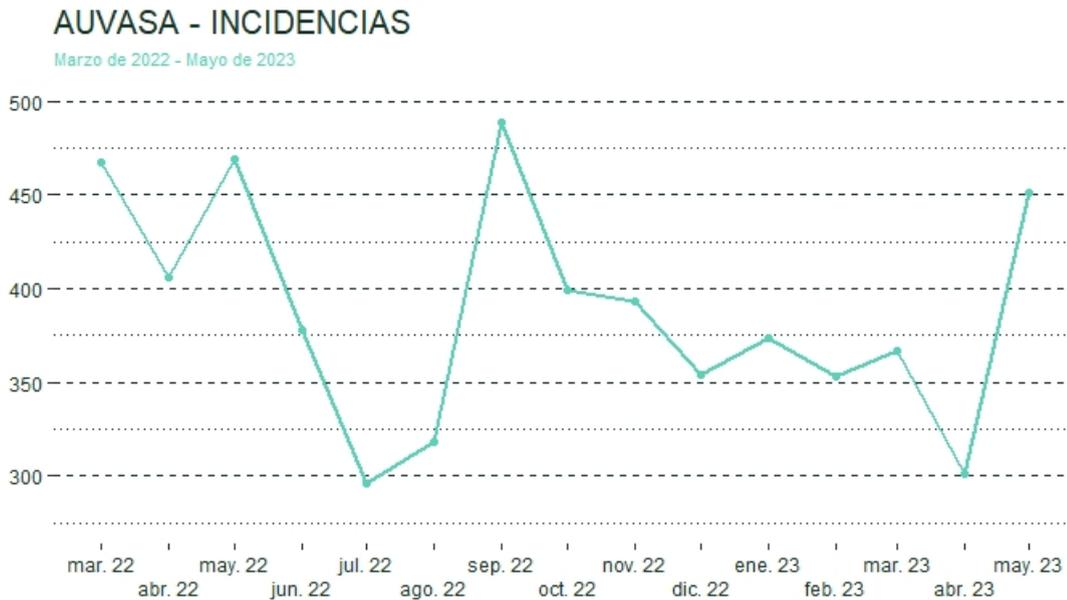


Figura 6: Incidencias mensuales con los datos registrados

Con el fin de profundizar en el análisis, se ha construido la tabla de frecuencias de todas las incidencias registradas (Tabla 3).

Es preciso tener en cuenta que se disponen de prácticamente todos los datos para 2022 (10 meses desde marzo a diciembre) y sólo de los datos de la primera mitad del año 2023 (desde enero a mitad de junio). Se observa que los tipos con más de una incidencia, salvo los correspondientes a anomalías en los vehículos y accidentes (T/A) en 2022, son prácticamente despreciables y su eliminación de cara a una mejor visualización gráfica apenas influye en el análisis.

Incidencia	Año		Total	Fr. Rel.
	2022	2023		
V	1314	784	2098	34.63
PH	883	470	1353	22.33
PB	370	221	591	9.75
A	261	226	487	8.04
TV	271	83	354	5.84
M	215	82	297	4.90
VP	202	92	294	4.85
I	116	76	192	3.17
G	118	39	157	2.59
T/A	138	0	138	2.28
T	53	14	67	1.11
D	9	0	9	0.15
C	7	1	8	0.13
A/T/VP	3	0	3	0.05
PB/VP	2	0	2	0.03
V/VP	2	0	2	0.03
A/T/M	1	0	1	0.02
A/V	1	0	1	0.02
PB/V	1	0	1	0.02
PH/VP	1	0	1	0.02
T/VP	1	0	1	0.02
TV/V	1	0	1	0.02
TV/VP	1	0	1	0.02

Tabla 3: Tabla de frecuencias según el Tipo de Incidencia y Año

En la Figura 7 se puede observar un gráfico de barras en el que se han representado las frecuencias condicionadas por años de los Tipos de Incidencias mostrados en la Tabla 3 sin tener en cuenta los menos representativos (los 10 últimos tipos que involucran más de una incidencia).

En ambos años, las incidencias se deben principalmente a comunicaciones internas (V), seguidas de aquellas que hacen referencia a la prolongación de jornadas, descansos y horas extras (PH).

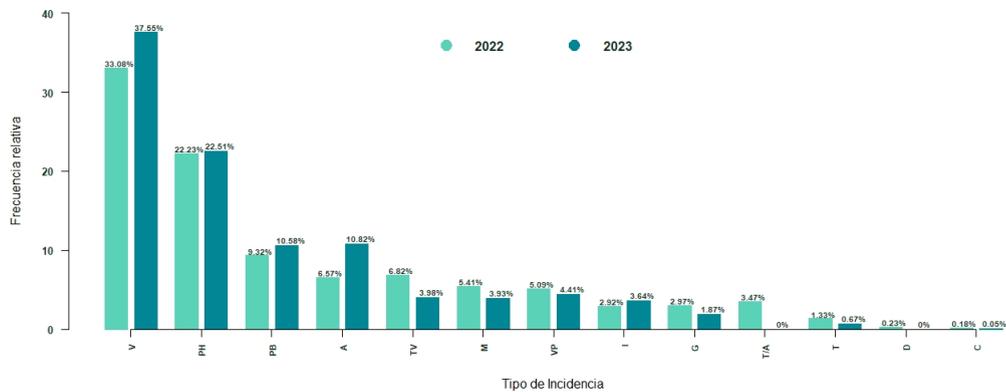


Figura 7: Gráfico de barras de las frecuencias relativas condicionadas al tipo de incidencia en cada año

Se observa, que aunque ambos años parecen seguir la misma distribución, las incidencias de los tipos V (comunicaciones internas), PH (prolongación de jornadas o descansos), PB (absentismos o bajas), A (accidentes), TV (anomalías en el servicio) e I (equipos y tarjetas) aumentaron sensiblemente su frecuencia en el presente año, 2023, respecto a 2022. Destacar, que en lo que va de año, no se han registrado incidencias de tipo T/A (accidente con consecuencias para transmitir a taller), sin embargo, han aumentado las de tipo A (accidentes y siniestros) respecto al año anterior.

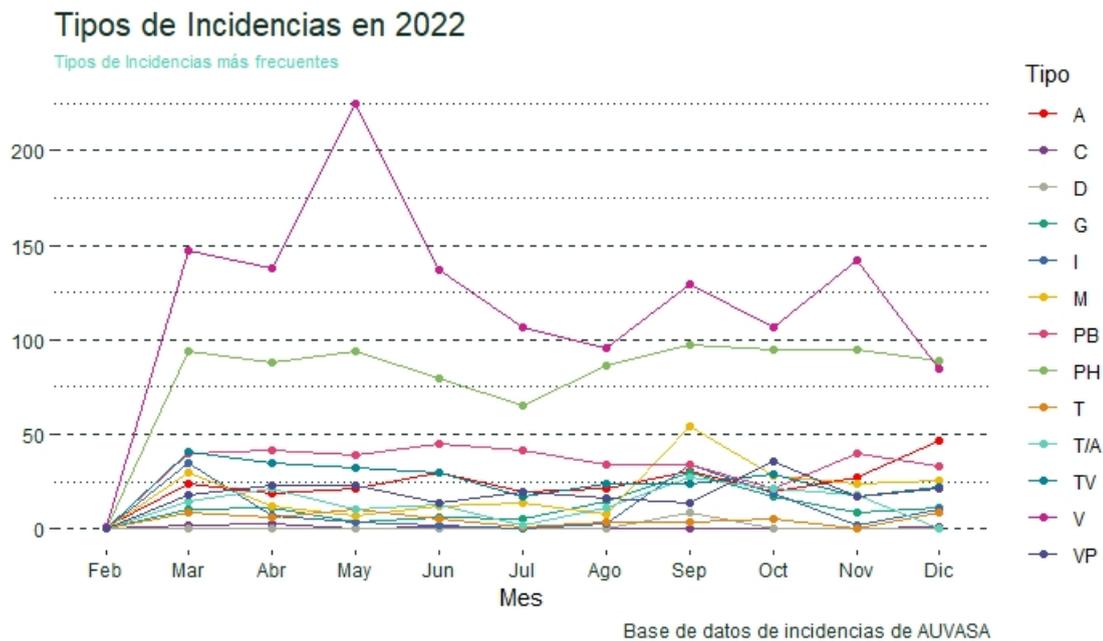


Figura 8: Gráfico de líneas en función del Tipo de Incidencia por mes

En la Figura 8 se puede ver un gráfico temporal con los tipos de incidencias a lo largo de los meses del año 2022. La incidencia más frecuente se debe a comunicaciones internas (V) alcanzando su máximo absoluto en el mes de mayo y su mínimo absoluto en el mes de diciembre. Como ya comentamos, le siguen en frecuencia las incidencias debidas a prolongación de jornadas, descansos y horas extra (PH).

La visualización de la gráfica de líneas anterior (Figura 8) es bastante compleja debido, entre otras cosas, al número de incidencias diferentes y a los valores cercanos que alcanzan algunas de ellas entre sí. Por ello, en lugar de utilizar un gráfico de líneas, se ha programado en R un tipo de representación gráfica (Figura 9) que nos permite visualizar mejor los tipos de incidencias y analizar éstas una a una (en el Anexo IV se encuentra el código programado de uno de estos gráficos). La escala utilizada es la misma para todas las incidencias para permitir realizar comparaciones entre ellas con mayor facilidad.

Se puede observar que son las comunicaciones internas (V) las más numerosas alcanzando su máximo en el mes de mayo, seguidas en importancia por las incidencias debidas a prolongación de jornadas, horas extra, etc (PH) que permanece prácticamente constante a lo largo de todo el año con un ligero descenso en los meses de verano, así como las debidas a bajas y absentismo (PB). Los siniestros y accidentes (A) parecen alcanzar su máximo absoluto en el mes de diciembre.

Las observaciones comentadas previamente nos han llevado a realizar un estudio más detallado de determinadas épocas del año en las siguientes subsecciones, bien por coincidir con periodos vacacionales o con algún evento de especial relevancia en la ciudad.

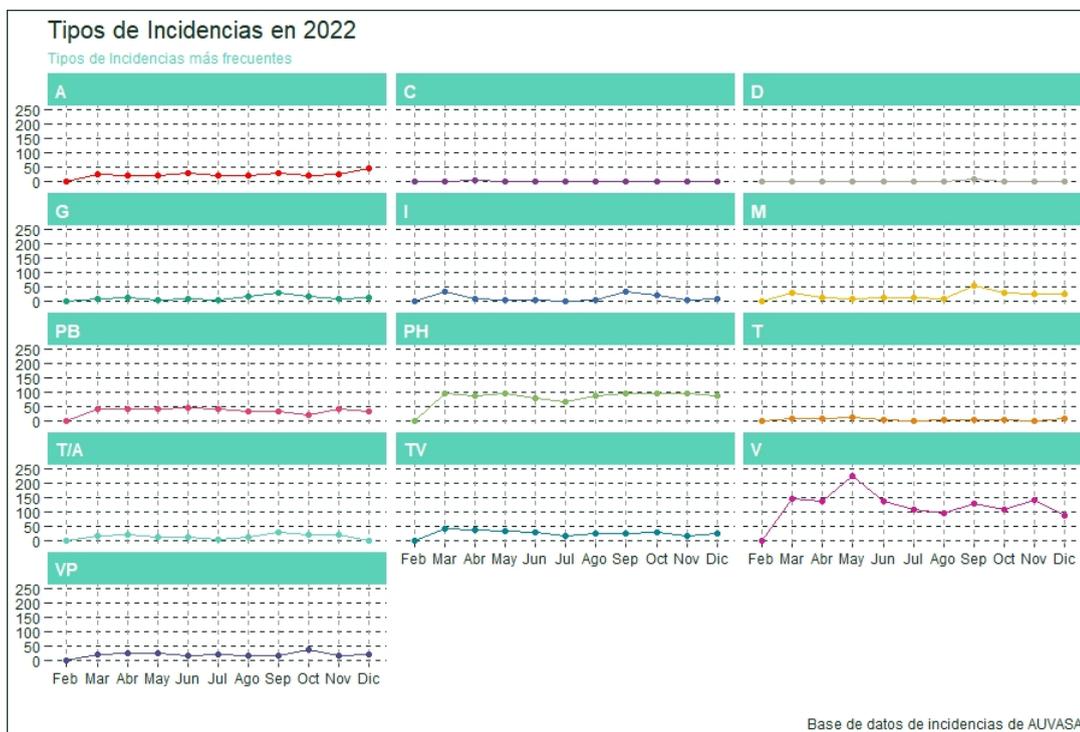


Figura 9: Representación temporal de incidencias en los meses de 2022

4.2. Análisis de las incidencias en determinadas épocas del año

Como se comentó previamente había dos meses, abril y julio, en los que el número de incidencias era mínimo y un mes, diciembre de 2022, donde las incidencias alcanzaban su máximo en este periodo.

En abril, tanto en 2022 como en 2023, se celebró la Semana Santa; julio es un mes del periodo estival y en diciembre se celebra la Navidad. Por ello, se realizó un análisis de estos periodos que se amplió a las Ferias y Fiestas de la Virgen de San Lorenzo en septiembre.



Figura 10: Total de Incidencias diarias en el mes de abril. Años 2022 y 2023

4.2.1. Semana Santa

La Semana Santa en 2022 se celebró entre los días 8 (Viernes de Dolores) y 17 de abril (Domingo de Resurrección) y entre el 31 de marzo (Viernes de Dolores) y el 9 de abril (Domingo de Resurrección) en 2023. Sin embargo, no parece observarse en esos días ningún patrón de su comportamiento. El número de incidencias sí alcanza el mínimo en esos días para el caso del año 2022, pero no así en la Semana Santa del año 2023.

En la Figura 10 se puede ver un gráfico de líneas con el total de incidencias ocurridas a lo largo de todos los días del mes de abril en los años 2022 y 2023. Para intentar ver si se producen comportamientos similares en ambos años, se han representado en la Figura 11 los días de la semana en los años 2022 y 2023 comenzando por el Viernes de Dolores hasta llegar al Domingo de Resurrección. El patrón observado cada uno de los años es diferente quizá debido a que durante el año 2022 todavía estaban vigentes algunas medidas restrictivas contra la COVID-19, lo que pudo tener influencia en la actividad y desplazamientos de los ciudadanos durante este periodo.

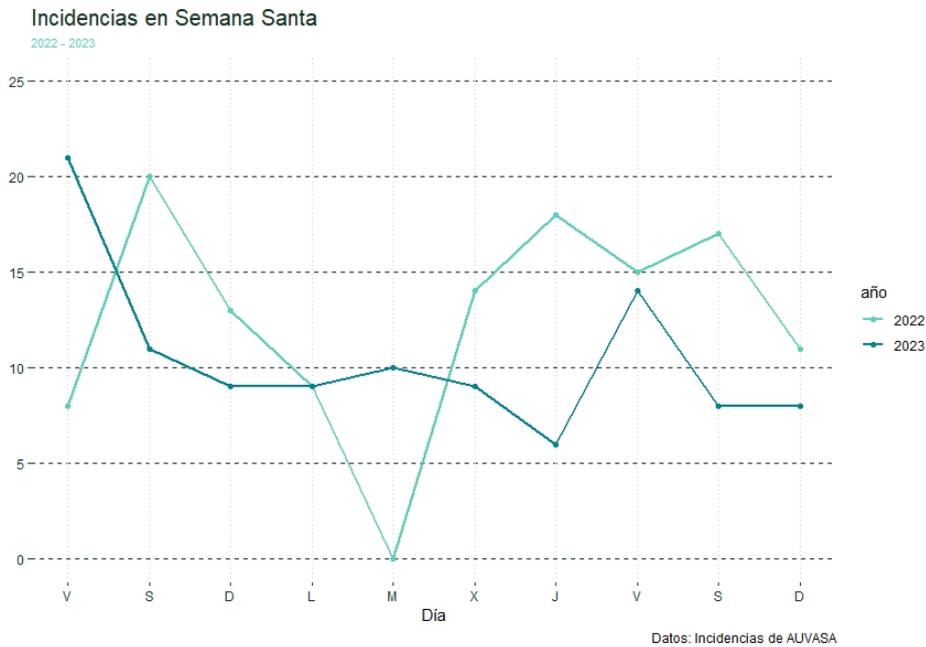


Figura 11: Semana Santa. Días de la semana. Años 2022 y 2023

En Semana Santa muchas calles se cortan al tráfico debido a que las procesiones transcurren por diferentes zonas del centro de la ciudad. Quizás este hecho explique porqué el número de incidencias máximo se alcanza fuera de estas fechas.

Para un mejor análisis de los tipos de incidencias se ha tenido en cuenta la agrupación que hace la empresa en su variable *Grupo*, en concreto, las categorías (*ACC*, *VEH* y *TIC*). Se han modificado las dos primeras ampliando los tipos de incidencias que agrupan. En la Tabla 4 podemos ver un resumen de éstas.

CÓDIGO (Grupo)	INCIDENCIAS	DESCRIPCIÓN
ACC	A, T/A, A/T/M, A/T/VP, A/V y TV	Accidentes Y siniestros con aviso a taller e incidencia de conductor
VEH	T, TV, TV/V, TV/VP	Incidentes, anomalías en vehículos y las anomalías en los servicios provocadas por vehículos
TIC	C, G, I	Incidencias de Software y GPS, problemas con los equipos, tarjetas retiradas y las anomalías en billeteaje

Tabla 4: Categorías de la variable Grupo para su representación

En las Figuras 12 y 13 observamos dos gráficos para algunos tipos de incidencias. En el primero, se observan las incidencias que no hemos agrupado y en el segundo gráfico se muestran las incidencias agrupadas. Se han trazado rectas verticales (del mismo color al utilizado para el año) abarcando los días correspondientes de la Semana Santa.

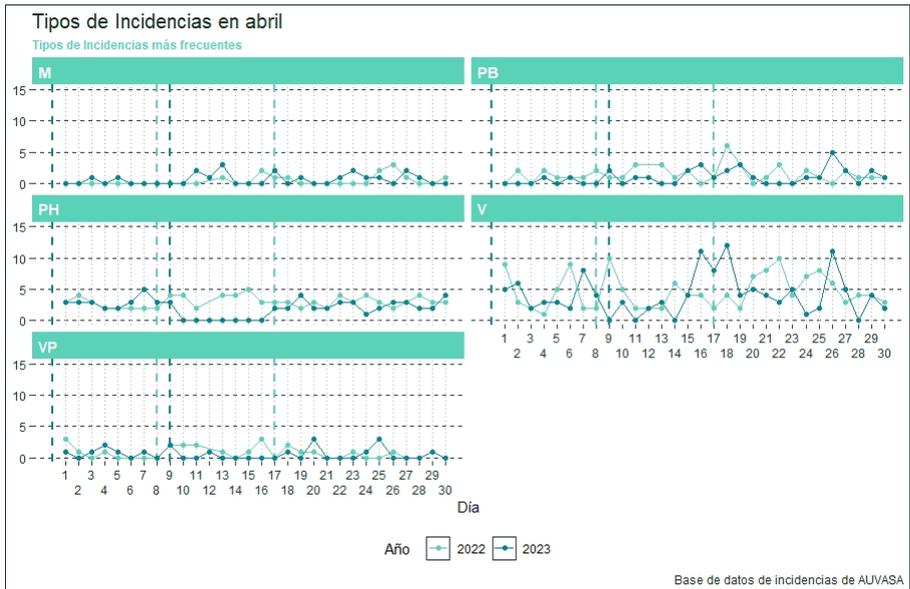


Figura 12: Algunos tipos de Incidencias en el mes de abril por año

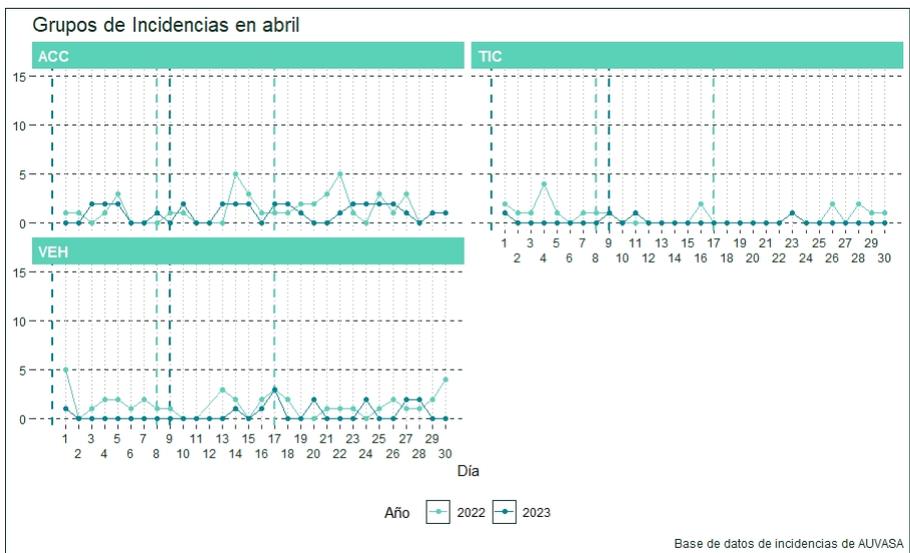


Figura 13: Algunas categorías de la variable *Grupo* en el mes de abril

Observando cada uno de los tipos y grupos de incidencias representados se aprecia que durante los días que transcurrió la Semana Santa ambos años, se produce un incremento de incidencias ocasionadas por la necesidad de cubrir horas extra y ampliaciones de jornada (PH). También en el año 2022 los accidentes y siniestros (A) con consecuencias para taller (VEH) alcanzaron un máximo el día de Jueves Santo.

Además, también se debe señalar que el día 22 de abril de 2022 se alcanzó otro máximo

en los siniestros coincidiendo con el fin de semana festivo en que se celebra el *Día de la Comunidad*.

4.2.2. Periodo estival

AUVASA, todos los años durante los meses de julio y agosto, se establecen nuevos horarios en algunas líneas de autobuses en los días laborables adaptándose al descenso de la demanda como consecuencia de las vacaciones escolares y de gran parte de la ciudadanía.

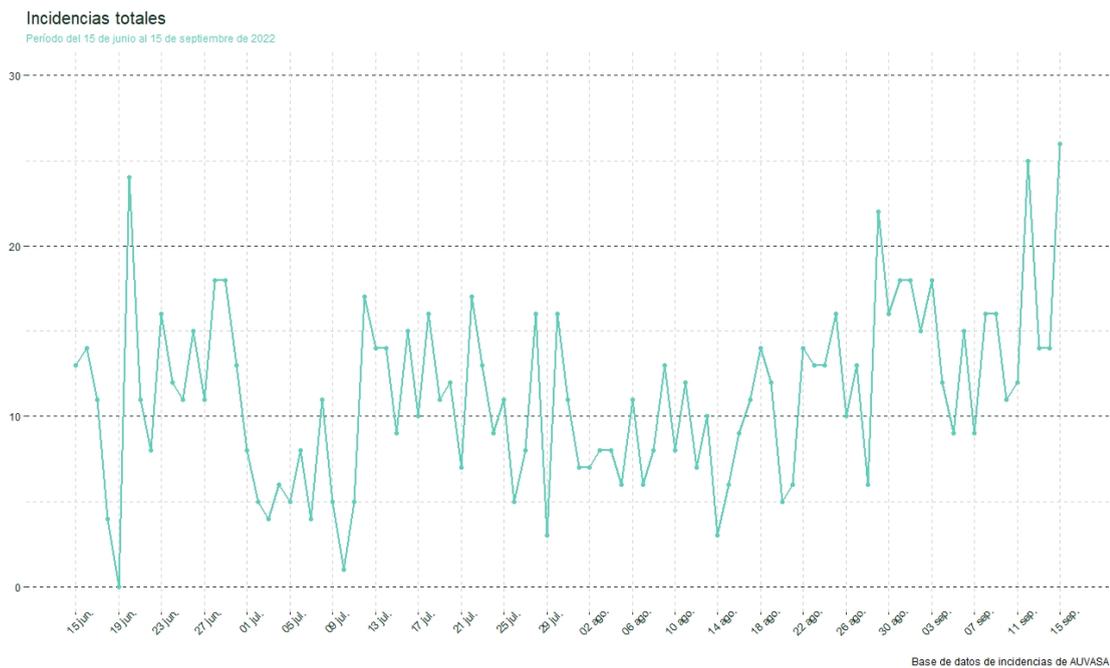


Figura 14: Total de Incidencias diarias en el periodo estival. Año 2022

En la Figura 14 se ha representado el número total de incidencias ocurridas entre el 14 de junio y el 14 de septiembre de 2022. Se puede observar que las incidencias comienzan disminuyendo en julio, oscilan después y se incrementan de forma notable en el mes de septiembre.

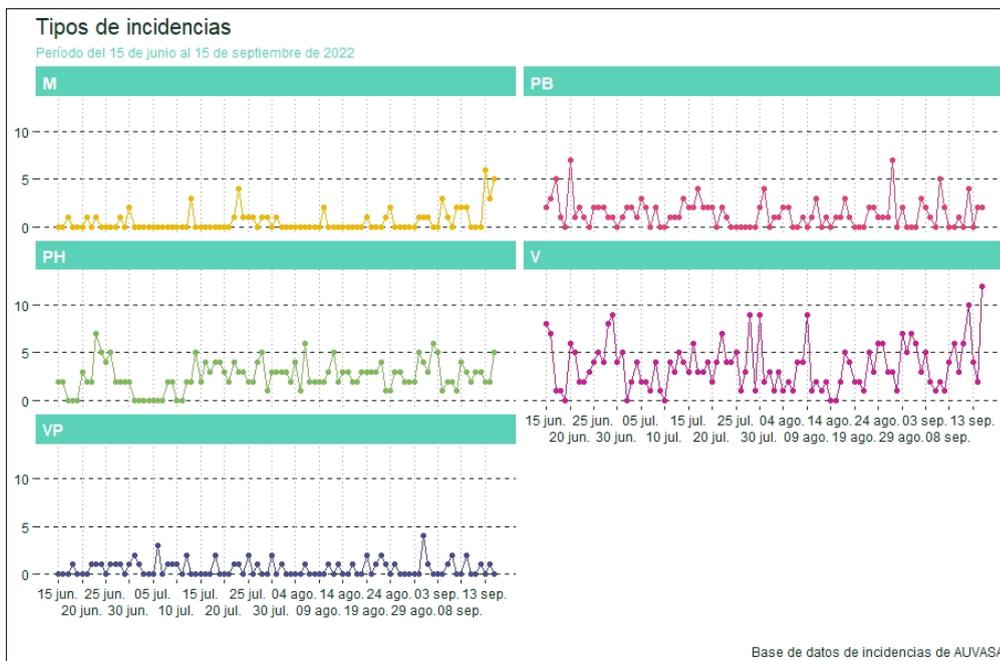


Figura 15: Representación en periodo estival del número de Incidencias por Tipo de Incidencia

En la Figura 15 se representan algunos tipos de incidencias aislados entre el 14 de junio de 2022 y el 15 de septiembre del mismo año. En la Figura 16 se corresponde con las incidencias agrupadas en el mismo periodo. De estos gráficos los únicos datos que parecen destacables son la disminución de las incidencias producidas por problemas informáticos y tecnológicos (TIC), las de mobiliario urbano (M) y las anomalías en vehículos (T).

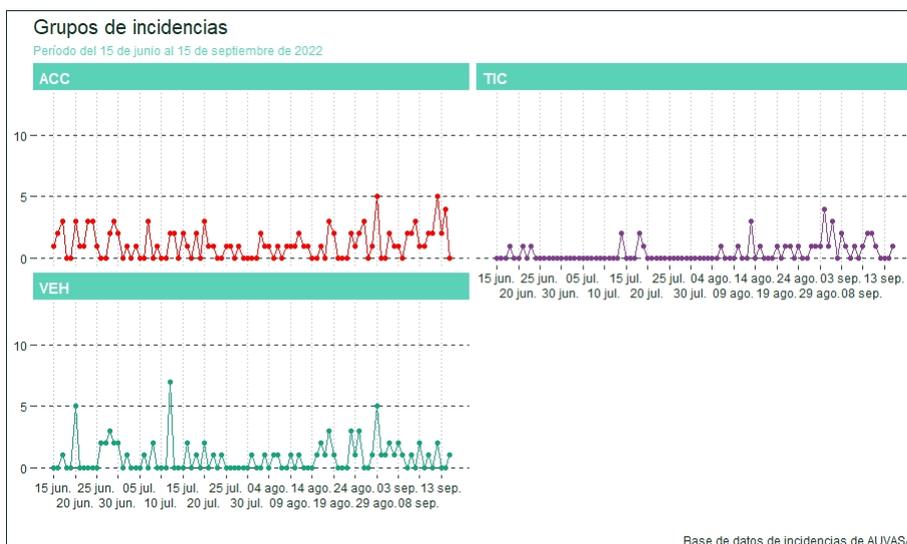


Figura 16: Representación en periodo estival del número incidencias en algunas categorías de la variable *Grupo*

4.2.3. Fiestas de la Virgen de San Lorenzo

Hemos considerado oportuno estudiar el comportamiento de las incidencias en el mes de septiembre cuando tienen lugar las Ferias y Fiestas de la Virgen de San Lorenzo que se celebraron el pasado año 2022 entre los días 2 y 11 de septiembre.

En el periodo de fiestas, para satisfacer el aumento de la demanda y el acceso al recinto ferial, AUVASA incrementa el número de autobuses y la plantilla en servicio. Así mismo, es una época con muchos actos festivos en las calles de la ciudad, que experimenta un notable incremento de visitantes. Pese a ello, y como se observa en la Figura 17, el número de incidencias se reduce en relación con los restantes días del mes.

En el gráfico se observa que en este periodo se alcanzó el mínimo absoluto en los días 5 y 7 de septiembre con un total de 9 incidencias. El número máximo de incidencias este mes se alcanzó el martes día 20 de septiembre, momento en el cual, la ciudad, ya ha recuperado su actividad habitual tras el verano y las Ferias y Fiestas.

Incidencias. Septiembre de 2022



Figura 17: Número de incidencias en el mes de septiembre de 2022

Prestando más atención al tipo de incidencias de estos días, observamos, en las Figuras 18 y 19, que los accidentes y siniestros es muy grande si se comparan con el resto del mes (no alcanzan tampoco el valor máximo), al igual que ocurre con las incidencias tipo TIC y las comunicaciones internas (V). Sin embargo, en estos días, sí se han alcanzado los máximos en bajas (PB) y necesidad de horas extra (PH)

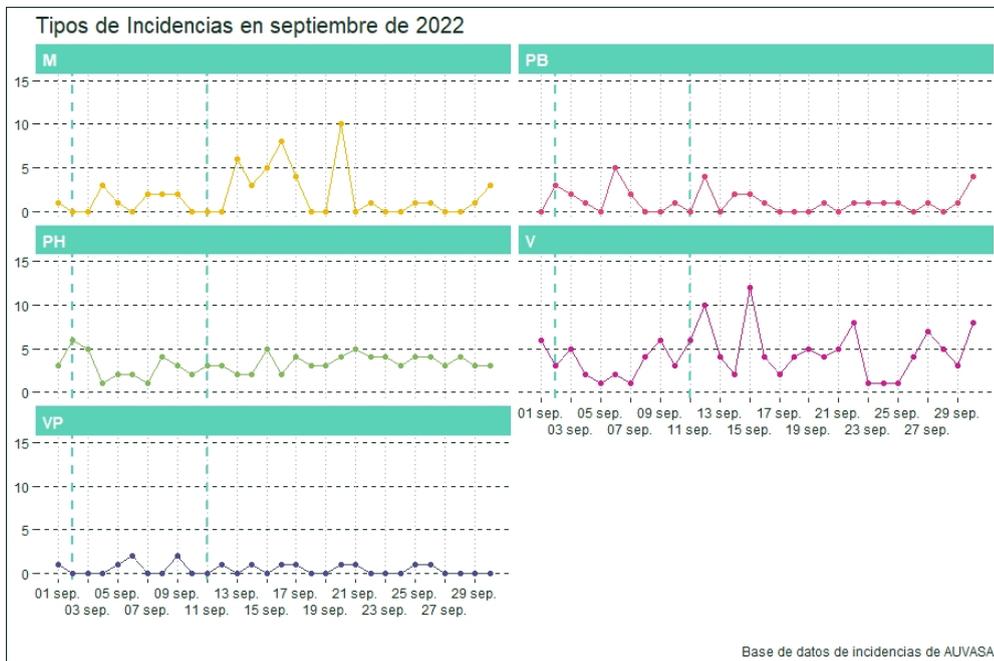


Figura 18: Número total de incidencias por tipo de incidencias en el mes de septiembre

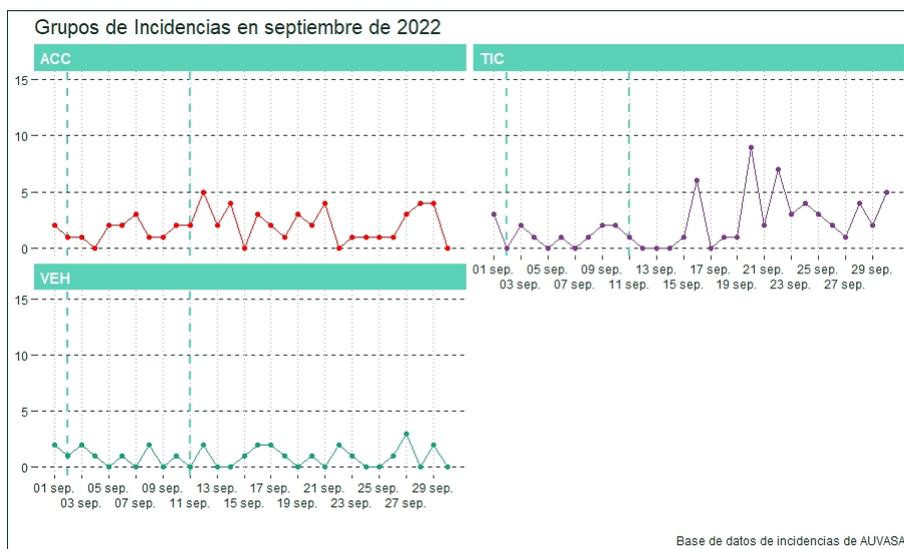


Figura 19: Número total de incidencias en diversas categorías de la variable Grupo en el mes de septiembre

4.2.4. Navidad

AUVASA reduce los horarios en algunas líneas ciertos días de Navidad, y suspende otras, de acuerdo con el descenso de la demanda de la población. En concreto, los últimos horarios de las líneas Ordinarias los días 24 y 31 de diciembre desde su parada de origen son aquellos anteriores a las 20:00 horas. Los días 25 de diciembre de 2022 y 1 de enero de 2023 se modificó la salida de los primeros servicios de las líneas Ordinarias, que comenzaron, en la mayoría de los casos, a las 10:30 horas. El servicio Búho se suspendió los días 24 y 31 de diciembre.

En la Figura 20 observamos un gráfico con el número total de incidencias en los meses de diciembre y enero. Los puntos marcados en el gráfico en color oscuro corresponden a los datos del periodo que coincide con las vacaciones escolares de Navidad.

Se observa que tres de los cuatro mínimos de este mes se producen durante estos días, en concreto, en los días 24 (Nochebuena), 25 (Navidad) y 26 (domingo en 2022). Tras un repunte de incidentes localizado en el lunes 27 de diciembre (final de este primer grupo de fiestas), el resto del periodo navideño parece registrar menos incidentes que el resto del mes de enero.

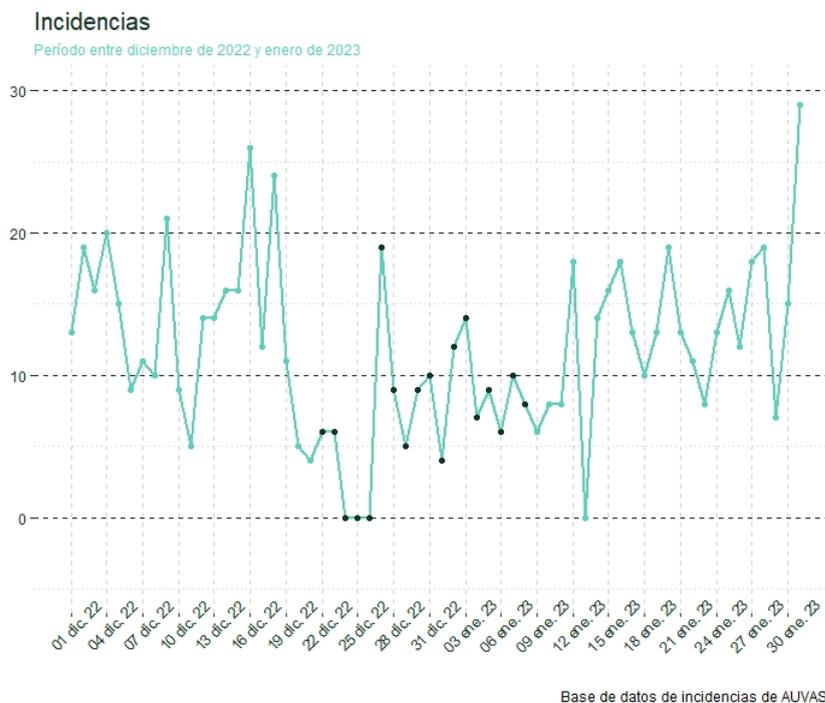


Figura 20: Total de Incidencias diarias en diciembre de 2022 y enero de 2023



Figura 21: Número de incidencias según Tipos de Incidencias en diciembre de 2022 y enero de 2023

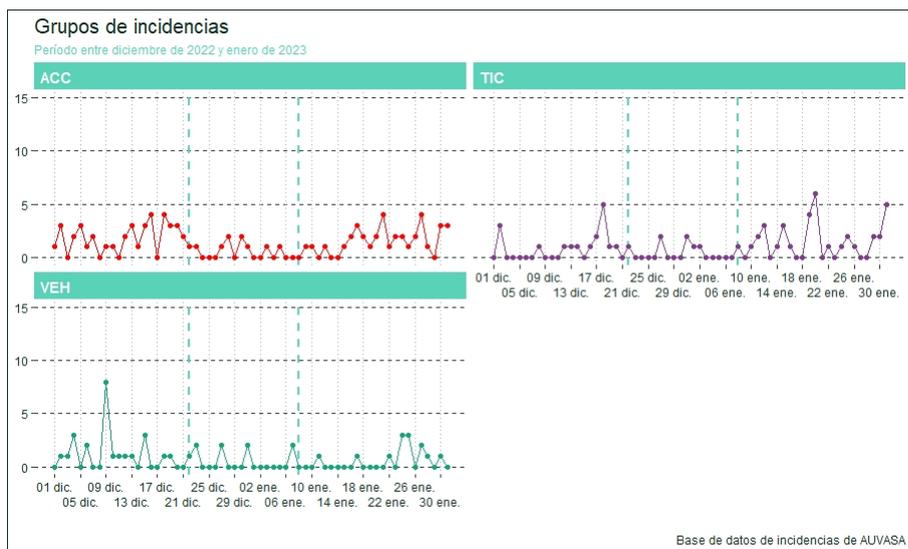


Figura 22: Número de incidencias en diversas categorías de la variable *Grupo* en diciembre de 2022 y enero de 2023

En las Figuras 21 y 22 se observa el número de incidencias de cada tipo a lo largo de los meses de diciembre y enero. Las líneas verticales discontinuas que se han señalado abarcan los días de las vacaciones escolares durante la Navidad de 2022. Se puede observar que en estos días se produce una disminución del número de accidentes y siniestros (ACC), así como de las incidencias TIC. En contraste, el 27 de diciembre se alcanza el máximo número de incidencias en el mobiliario urbano (M) de los dos meses (diciembre y enero). El número de bajas de personal (PB) y el de ampliación de jornada y horas de trabajo (PH) alcanza a lo largo de varias jornadas su valor mínimo.

4.3. Análisis de incidencias por línea de autobús

En esta subsección se realizó un estudio de las incidencias por líneas de autobús. Recordemos que se creó una nueva variable *TipoLinea*, con valores: ordinaria, Búho, matinal, ferial, fútbol, otros servicios, polígono, refuerzo, servicio personal, universitaria y múltiple.

La tabla de frecuencias condicionadas por año (expresada en porcentajes) en para el 2022 y el 2023 (Tabla 5) muestra que el mayor número de incidencias se produce en las líneas Ordinarias, que son además las que tienen mayor número de usuarios (Anexo III). A menor escala, le siguen las líneas Búho, que de nuevo son las que siguen en demanda a las Ordinarias. Este hecho, podría indicar que hay una correlación entre el número de incidencias que registran este tipo de líneas y el número de viajeros de las mismas.

Tipo de Línea	2022	2023
Ordinaria	95.134 %	96.288 %
Búho	2.679 %	1.747 %
Otros servicios	0.492 %	0.218 %
Fútbol	0.383 %	0.437 %
Matinal	0.328 %	0.546 %
Polígono	0.328 %	0.109 %
Refuerzo	0.273 %	0.109 %
Universitaria	0.164 %	0.218 %
Servicio de Personal	0.109 %	0.218 %
Ferial	0.055 %	0.000 %
Múltiple	0.055 %	0.109 %

Tabla 5: Frecuencia relativa de incidencias condicionada al tipo de línea en cada año en los tipos de Líneas de autobús.

Se han representado a través de gráficos (Figura 23 y Figura 24) el número total de incidencias a lo largo de los 11 meses de 2022 y los casi primeros seis meses de 2023 (tenemos presente que no hay datos a partir de julio de 2023 lo cual se ve reflejado en las curvas correspondientes) con el fin de intentar ver tendencias en los datos.

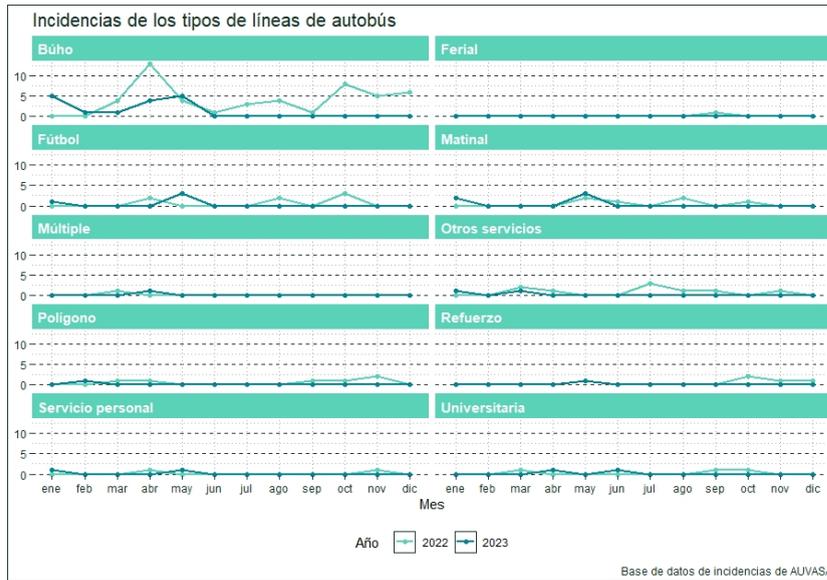


Figura 23: Número de incidências por Tipo de Línea salvo las Ordinarias

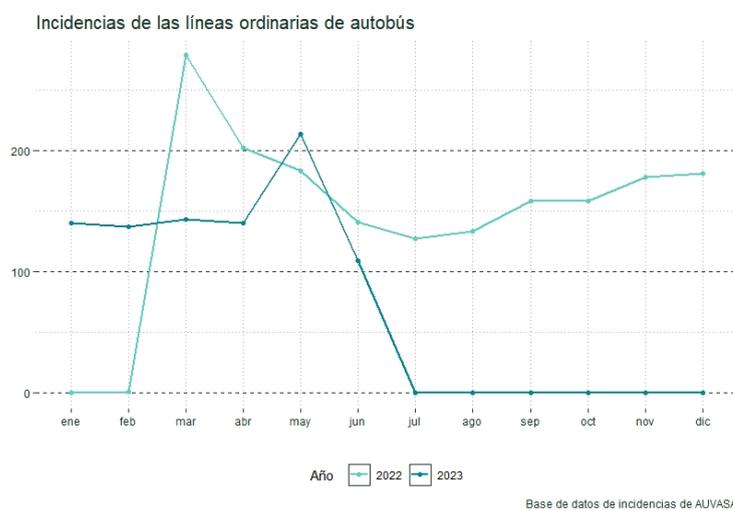


Figura 24: Número de incidências en las Líneas Ordinarias por mes según año

Lo más destacable de estas figuras es:

- Las líneas Búho alcanzan el máximo número de incidencias en abril el año 2022 y para los datos disponibles de 2023 en mayo.
- En las líneas universitarias se observan claramente los periodos lectivos y los no lectivos.
- Las líneas Ordinarias alcanzan un máximo absoluto en marzo de 2022 y un máximo relativo (por el momento) en mayo de 2023.
- Si nos fijamos en los meses de junio, julio y agosto para el año 2022 en las líneas Ordinarias, observamos que se produce un descenso de las incidencias. Esto puede ser debido a que se reduce la frecuencia de algunas líneas Ordinarias (a partir del 1 julio) y coincide con el período vacacional de una gran parte de los ciudadanos, así como de los centros escolares. Sin embargo, si nos fijamos en el período estival del 2022 referente a las líneas Búho, se observa que en junio el número de incidencias descienden para luego aumentar durante julio y agosto.

Dado que las líneas Ordinarias recogen más del 95 % de las incidencias, se han analizado en mayor profundidad. Para ello, se construyeron las tablas de frecuencias absolutas y condicionadas por año de estas líneas (Anexo VII) donde se observa que las diez líneas con más incidencias por orden para el año 2022 son:

1, 2, 6, C1, 8, C2, 7, 9, 3, 5.

Si sumamos las incidencias de las líneas circulares para poder compararlas por el número de viajeros que figura en el Anexo III (no se dispone de la información para usuarios de las líneas circulares por separado), las líneas con más incidencias son :

C, 1, 2, 6, 8, 7, 9, 3, 5.

Las líneas Ordinarias de autobús ordenadas de mayor a menor número de usuarios son:

1, C, 2, 6, 8, 7, 9, 3, 5, 19, 18, 4, H, 16, 14, 13, 17, 10.

Desde la Figura 26 hasta la Figura 31 se representan los gráficos de barras para las 10 líneas Ordinarias que registran más incidencias. Se han representado el total de incidencias atendiendo a la variable *Grupo* que hace la empresa (véase la sección Procedimiento de Depuración en su subsección Base de datos inicial).

Se puede observar que, salvo en el grupo de incidencias que involucran al mobiliario urbano (PAR) donde la línea 7 es la que tiene más registros, en los demás grupos es la línea con más viajeros (línea 1) la que más casos registra. En el caso de los demás grupos, parece observarse una correlación entre las líneas con más usuarios y el número de incidencias.

Otros grupos de incidencias que pueden tener, a priori, mayor impacto económico son: ACC (accidentes y siniestros), VEH (anomalías en vehículos) y PER (personal de baja, ampliación de jornadas y horas extra). En estos casos, se observa que las líneas 6 y 7 pese a su número de usuarios registran menos accidentes. La línea 3 adelanta en anomalías en vehículos a otras líneas más frecuentadas.



Figura 25: Gráfico de barras de frecuencias relativas de incidencias por línea de autobús y año

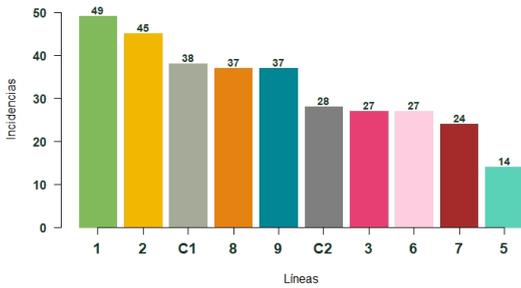


Figura 26: Total incidencias Grupo ACC en las 10 líneas Ordinarias más demandadas

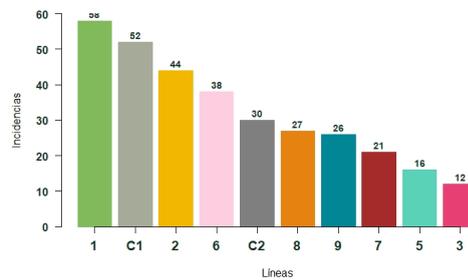


Figura 27: Total incidencias Grupo COM en las 10 líneas Ordinarias más demandadas

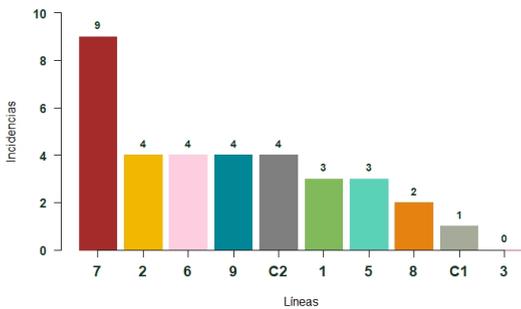


Figura 28: Total incidencias Grupo PAR en las 10 líneas Ordinarias más demandadas

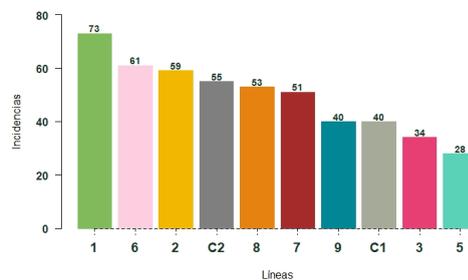


Figura 29: Total incidencias Grupo PER en las 10 líneas Ordinarias más demandadas

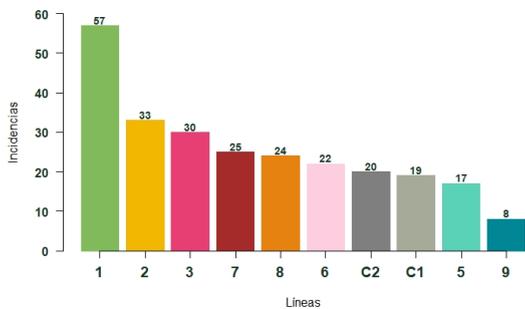


Figura 30: Total incidencias Grupo VEH en las 10 líneas Ordinarias más demandadas

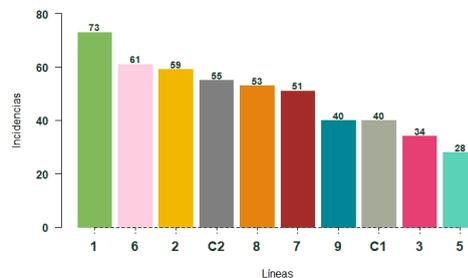


Figura 31: Total incidencias Grupo TIC en las 10 líneas Ordinarias más demandadas

En el Anexo VII se adjuntan más tablas de frecuencias para las líneas Ordinarias y Búho en función de los grupos de incidencias.

En la Figura 32 hemos representado en el año 2022 las frecuencias relativas del total de incidencias, por un lado, y del número de usuarios, por otro, para las líneas Ordinarias con el fin de observar la relación entre ambas variables.

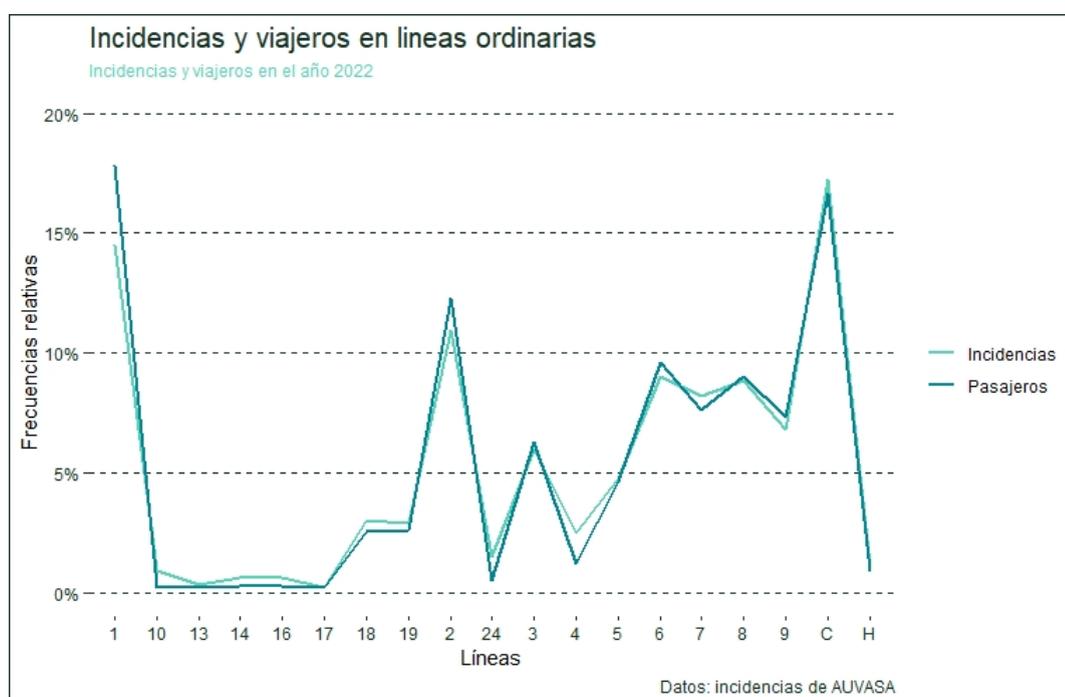


Figura 32: Frecuencias relativas de incidencias y viajeros para las Líneas Ordinarias. Año 2022

Se puede observar que ambas curvas tienen un comportamiento similar, lo que parece corroborar la correlación comentada previamente. Se ha construido también el diagrama de dispersión para todas las líneas, salvo para el autobús turístico, (Figura 33) y los datos parecen mostrar una relación muy fuerte positiva entre ambas variables.



Figura 33: Diagrama de dispersión. Año 2022

Se ha calculado el coeficiente de correlación de Pearson

$$r = \frac{S_{xy}}{S_x \cdot S_y}$$

siendo S_x y S_y las desviaciones típicas muestrales de las variables x e y , respectivamente y S_{xy} la covarianza muestral de las variables x e y . El valor obtenido es $r = 0,98$ que corrobora que la correlación lineal es positiva y alta por lo que se deduce que, en efecto, a mayor demanda de una línea, se produce mayor número de incidencias.

4.4. Calendario anual de incidencias

Se ha escrito en R, haciendo uso de la librería *ggplot2*, el código para diseñar un calendario para la empresa (ANEXO V de la memoria) con los días de la semana coloreados de acuerdo al número de incidencias ocurridas y utilizando como base los colores representativos de AUVASA. La intención es que de un simple vistazo se puedan visualizar los días con más incidencias en el calendario.

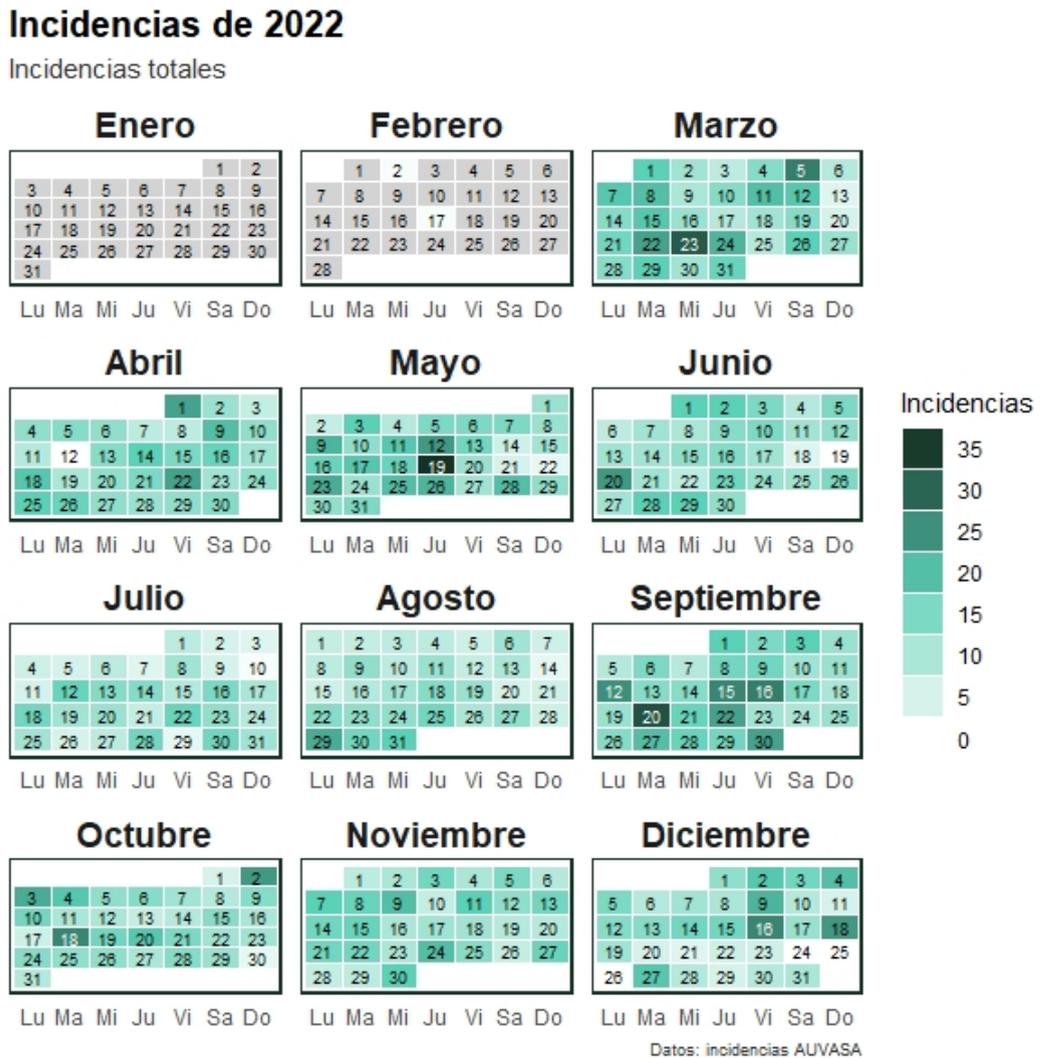


Figura 34: Calendario de Incidencias en 2022

En la Figura 34 se observa el calendario de Incidencias del año 2022. Los días coloreados en gris indican que no se dispone de datos. Aquellos coloreados de blanco corresponden a días sin incidentes. A medida que el color se oscurece el número de incidencias se incrementa.

Incidencias de 2023

Incidencias totales

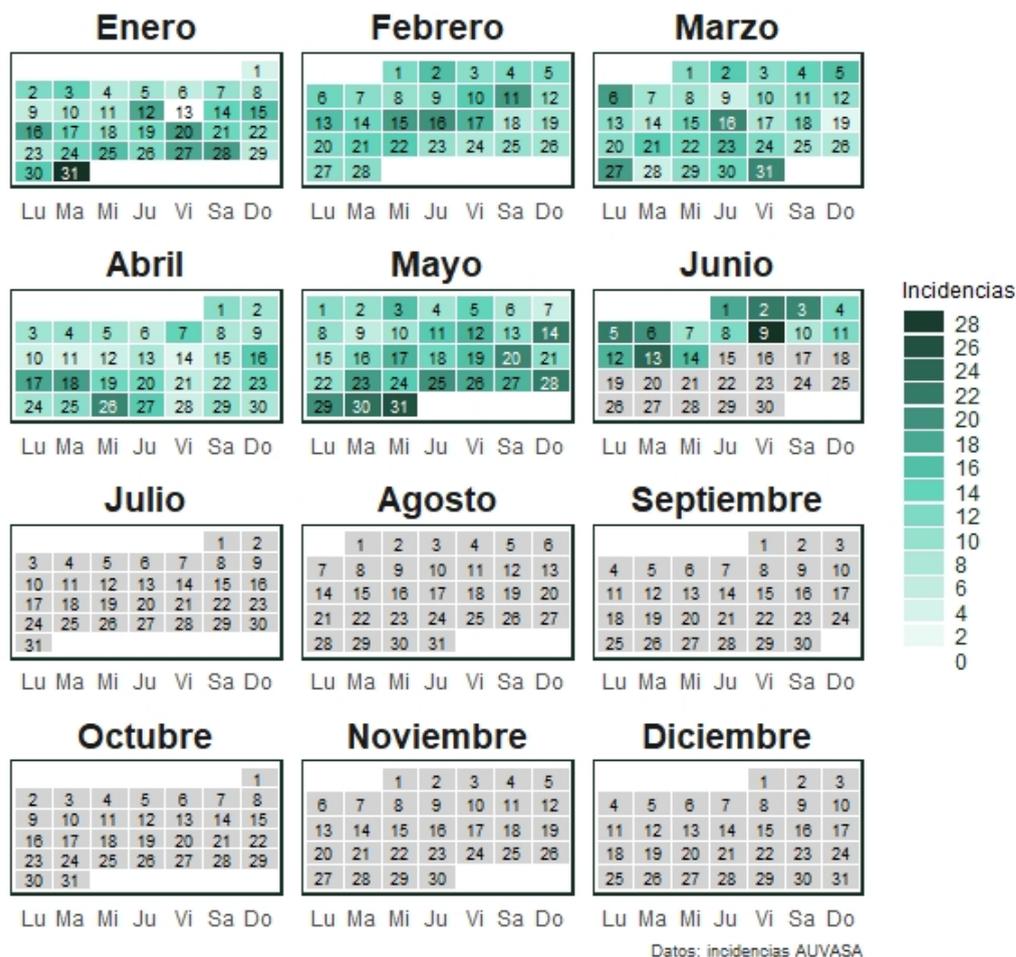


Figura 35: Calendario de Incidencias en 2023

En la Figura 35 se observa el correspondiente calendario para los datos registrados de 2023.

En el Anexo VI se incluyen calendarios del año 2022 para algunos de los tipos de incidencias de mayor interés.

5. Conclusiones y trabajos futuros

Las principales conclusiones obtenidas de este trabajo fin de grado son:

- Algunas de las variables utilizadas en la base de datos permiten una respuesta abierta que dificulta el análisis estadístico. Además, se ha observado en algunas de ellas, falta de uniformidad en los formatos, por ejemplo, en la variable *Línea* se pueden encontrar valores como 8 o L-8 para referirse al mismo dato.
- El 95 % de las incidencias están asociadas a las líneas de autobús Ordinarias, seguidas de las líneas Búho con un 2,37 %.

Las diez líneas Ordinarias con mayor número de incidencias en orden decreciente son: 1, 2, 6, C1, 8, C2, 7, 9, 3 y 5. Estas líneas son las que mayor número de usuarios registrados tuvieron en el año 2022. Se comprobó que estas dos variables (*líneas Ordinarias*) y *número de viajeros*) estaban altamente correlacionadas, indicando que un aumento en el número de usuarios implica un aumento del número de incidencias.

- Tanto en 2022 como en 2023, el número de incidencias en el mes de abril parece experimentar un descenso respecto de los meses inmediatamente anterior y siguiente. Una hipótesis era que se podía deber a la celebración de la Semana Santa, sin embargo, no parece repetirse el patrón en los dos años considerados, podemos asumir, por lo tanto, que pueda deberse al corte de muchas calles de la ciudad debido al paso de las procesiones de Semana Santa y a las vacaciones escolares en este periodo.
- En el periodo comprendido desde el 1 de julio hasta finales del mes de agosto de 2022, AUVASA reduce la frecuencia de sus líneas ya que se produce un descenso de la demanda como consecuencia de las vacaciones de una gran parte de la población. Analizando este periodo, se observa un descenso significativo del número de incidencias, que alcanzan su valor mínimo en el mes de julio de 2022. Estudiando los incidentes de este periodo por líneas se observa una disminución de incidencias en las de tipo ordinario, pero un crecimiento en las líneas de tipo Búho.
- Se han analizado también los datos correspondientes al mes de septiembre de 2022 ya que se celebraron las Ferias y Fiestas de la Virgen de San Lorenzo. Pese a que AUVASA incrementa el número de autobuses con las líneas R que llegan al recinto ferial (véase Anexo I), e incrementa los días y horarios de las líneas Búho, no se ha producido un aumento del total de las incidencias, pero sí se alcanza el máximo anual en las incidencias de mobiliario urbano. El número total de incidencias alcanza el máximo anual en el mes de septiembre.
- Los datos correspondientes a la Navidad de 2022 (las vacaciones escolares transcurrieron entre el 22 de diciembre de 2022 y el 8 de enero de 2023), muestran un descenso en las incidencias registradas en este periodo, salvo en las relativas a mobiliario urbano que experimentan un aumento alcanzando su máximo el día 27 de diciembre. En cuanto a las líneas de autobús, las únicas que experimentaron un incremento de incidencias fueron las líneas Búho.
- Para explotar la información recogida en algunas de las variables sería preciso realizar un procesamiento de lenguaje natural (PLN) que detecte e identifique frases

y/o palabras clave que faciliten la extracción de conclusiones. Las respuestas cerradas contribuyen también a minimizar el error humano durante la recogida de datos. Una alternativa es que estas variables de descripción libre se complementen con nuevas variables de respuesta cerrada (de tipo dicotómico) que reflejen parte de la información.

En consonancia con las etapas y conclusiones obtenidas se propone a la empresa AUVA-SA las siguientes recomendaciones:

- Rediseñar la base de datos presente en la empresa, de forma que contenga más variables de respuesta cerrada y que complementen o amplíen la descripción de un incidente. También, en la medida de lo posible, se recomienda automatizar la entrada de datos.
- En la variable *Tipo de Incidencia* en la categoría A (accidentes y siniestros), se propone añadir nuevas variables que indiquen si ha habido daños materiales o no, daños a personas o no, etc.
- De los dos ficheros que nos proporcionó la empresa, en el primero de ellos además de registrar la fecha de la incidencia, se registraba la hora, sin embargo, esta información no constaba en el segundo fichero. Un estudio interesante sería analizar los accidentes y siniestros en función de la hora del suceso para observar si hay relación con los momentos con mayor tráfico en la ciudad.
- Profundizar en el análisis de la base de datos. Sería interesante desarrollar una herramienta de procesamiento del lenguaje natural (PLN) de la variable *Descripción* desarrollando algoritmos híbridos de categorización de textos que partan de un clasificador que pueda entrenarse (aprendizaje automático) y de un clasificador de textos basado en reglas.
- Ampliar el estudio utilizando series temporales cuando la empresa disponga de una mayor cantidad de datos recogidos a lo largo de más años con el fin de detectar patrones, datos atípicos y desarrollar modelos predictivos.
- Desarrollar una aplicación web que permita el acceso a los calendarios (*Heatmap Calendar*) elaborados en este trabajo con posibilidad de que sean interactivos de forma que pulsando sobre un día concreto del año se muestre la información sobre las incidencias recogidas ese día.

Bibliografía

- [1] r-project.org. *The R Project for Statistical Computing*.
<https://www.r-project.org>.
Visitado el 20 de junio de 2023.
- [2] G. Grolemond y H. Wickham. *R for Data Science*. O'Reilly Media, 2017.
- [3] J. Maindonald, y J. Braun. *Data analysis and graphics using R. An example-based approach*. Cambridge University press, 2007.
- [4] RStudio IDE User Guide.
<https://docs.posit.co/ide/user/#about-this-guide>.
Visitado el 20 de junio de 2023.
- [5] overleaf.com. Overleaf documentation.
<https://www.overleaf.com/learn>.
Visitado el 20 de junio de 2023.
- [6] latex-project.org. *The latex project*. <https://www.latex-project.org/>.
Consultado en junio de 2023.
- [7] Universidad de Granada. *Introducción a la depuración e imputación de datos estadísticos en el proceso estadístico. Datos, errores, datos ausentes y controles (edits). Métodos básicos para la depuración e imputación de datos estadísticos. Estrategia de la depuración e imputación*.
<https://www.studocu.com/es/document/universidad-de-granada/estadistica/tema-15-depuracion-e-imputacion/40332990>.
Visitado el 20 de junio de 2023.
- [8] T. de Waal, J. Pannekoek, S. Scholtus. *Handbook of Statistical Data Editing and Imputation*. John Wiley & Sons, 2011.
- [9] P. Revilla. *Métodos estadísticos de depuración e imputación de datos*. Tesis doctoral. Universidad de Alcalá, 2013.
- [10] S. Fernández, J.M. Cordero, A., Córdoba. *Estadística Descriptiva*. ESIC Editorial, 2002.
- [11] C. Fernandez-Cuesta, F. Fuentes-García. *Curso de Estadística Descriptiva. Teoría y práctica*. Ariel Economía, 1995.
- [12] J.M. Montero. *Estadística Descriptiva*. Ed. Thomson, 2007.
- [13] D. Peña Sánchez de Rivera. *Análisis de series temporales*. Alianza Editorial, 2007.
- [14] E. uriel, A. Peiró. *Introducción al Análisis de Series Temporales*. Editorial AC, 2000.

6. ANEXOS

6.1. ANEXO I. Descripción detallada de las líneas de autobús urbano

Toda la información que se proporciona a continuación es pública y se puede acceder a ella en http://www.auvasa.es/auv_lineas.asp

- **Líneas Ordinarias.** Se trata de un total de 20 líneas de autobús, de las cuales dos de ellas son circulares.

- **Línea 1.** Barrio España–Covaresa.

Tiene un recorrido diametral que cruza Valladolid de norte a sur pasando por el barrio de España, La Rondilla, el centro de la ciudad y el paseo de Zorrilla en toda su extensión, utilizando su carril bus, hasta el barrio de Covaresa. Da servicio a varios puntos de interés de la capital como el Hospital Clínico Universitario, la plaza de toros, el parque del Campo Grande, salas culturales como el Teatro Calderón, el Museo Patio Herreriano o el Laboratorio de las Artes (LAVA), y centros comerciales y grandes almacenes. Asimismo, da servicio al aparcamiento disuasorio de Las Moreras.

- **Línea 2.** San Pedro Regalado–Covaresa.

Es una línea diametral que cruza Valladolid de norte a sur pasando por el centro de la ciudad y las estaciones de autobuses y ferrocarriles. También usa parte del carril bus del paseo de Zorrilla. Da servicio a varios puntos de interés de la ciudad como el Hospital Clínico Universitario, la plaza de toros, el parque del Campo Grande, salas culturales como el Teatro Calderón o el Laboratorio de las Artes (LAVA), espacios comerciales, así como al aparcamiento disuasorio de la calle La India. Tiene también paradas en varios intercambiadores del transporte metropolitano de Valladolid en la estación de autobuses (cabecera de multitud de líneas) y en la plaza de Juan de Austria (paseo de Zorrilla).

- **Línea 3.** Las Flores–Girón.

Es una línea diametral que cruza Valladolid desde el barrio de Las Flores hasta Girón y el entorno del Cementerio Las Contiendas, atravesando el centro de la ciudad. Al pasar por la plaza Circular, permite el transbordo con varias líneas del transporte metropolitano de Valladolid y también da servicio a los aparcamientos disuasorios de Plaza del Milenio y calle La Vía.

- **Línea 4.** Pilar de Jalón–Fuente Berrocal.

Une los barrios de Fuente Berrocal y Pinar de Jalón con el centro de Valladolid. A lo largo de su recorrido da servicio a lugares como la Feria de Valladolid, el Museo Patio Herreriano, el parque del Campo Grande, la Universidad Europea Miguel de Cervantes o la sede del Instituto Nacional de la Seguridad Social (INSS) de la calle Boston, así como la estación de autobuses con sus paradas en el paseo del Arco de Ladrillo. También pasa por las cercanías de los aparcamientos disuasorios de Feria de Muestras, Las Moreras y Paseo del Renacimiento.

- **Línea 5.** La Victoria–Santa Ana–Entreopinos.

Une la urbanización Entreopinos, perteneciente al municipio de Simancas, con el barrio vallisoletano de La Victoria pasando por el centro de la capital. A lo largo de su recorrido pasa algunos centros comerciales, el Laboratorio de las Artes (LAVA), la plaza de toros y el parque del Campo Grande, así como a los aparcamientos disuasorios cercanos a la Plaza del Milenio y la Feria de Valladolid.

- **Línea 6.** Delicias–La Victoria.

Tiene un recorrido diametral que cruza Valladolid de sudeste a noroeste pasando por el centro de la ciudad, el Hospital Universitario Río Hortega (que se sitúa uno de sus extremos) y los barrios de Las Delicias, Huerta del Rey, La Victoria y Puente Jardín. Da también servicio a los aparcamientos disuasorios cercanos a Las Moreras, la Plaza del Milenio y la Feria de Valladolid.

- **Línea 7.** Arturo Eyries–Los Santos Pilarica.

Con recorrido diametral, atraviesa Valladolid de suroeste a nordeste, pasando por el centro de la ciudad además de los barrios de Arturo Eyries, Belén, Pilarica, Los Santos Pilarica y el primer tramo del paseo de Zorrilla. Da servicio a diversos puntos de interés como el Polideportivo Pisuerga, la plaza de toros, el parque del Campo Grande, el palacio de Santa Cruz, la Jefatura Provincial de Tráfico y la escuela de Ingenierías Industriales de la Universidad de Valladolid, así como a los aparcamientos disuasorios de Cuatro de Marzo y calle La Vía, finalizando en el barrio de Los Santos–Pilarica aprovechando el nuevo túnel entre las calles Nochevieja y Andrómeda.

- **Línea 8.** Parquesol–Belén.

Es una línea diametral que cruza Valladolid de suroeste a nordeste pasando por el centro. Presta servicio a puntos de interés de la ciudad como el Centro Cultural Miguel Delibes, el estadio José Zorrilla, el Campus Miguel Delibes de la Universidad de Valladolid, así como el apeadero ferroviario de Valladolid–Universidad y los aparcamientos disuasorios cercanos a Las Moreras y la Plaza del Milenio. También pasa por las cercanías de lugares como el Polideportivo Huerta del Rey, el palacio de Santa Cruz, el Hospital Clínico Universitario y la facultad de Medicina y Enfermería. Las mañanas de los días festivos da servicio con una parada más cercana al mercadillo que se instala en las inmediaciones del estadio José Zorrilla.

- **Línea 9.** Parquesol–Delicias–Polígono San Cristóbal.

Su recorrido diametral cruza Valladolid de oeste a este pasando por el centro. Nace en el barrio de Parquesol y llega hasta el Hospital Río Hortega todos los días, y al polígono industrial de San Cristóbal los días laborables. También pasa por las estaciones de autobuses y de ferrocarril, la plaza de toros y el parque del Campo Grande. Da servicio al aparcamiento disuasorio de calle La India.

- **Línea 10.** Parquesol–Villa del Prado.

Con servicio en días laborables, es una línea que une el extremo sur del barrio

vallisoletano de Parquesol con Villa de Prado, a través del barrio de Arturo Eyrías, el paseo de Zorrilla y la plaza de Juan de Austria, y da servicio al aparcamiento disuasorio del barrio Cuatro de Marzo.

- **Línea 13.** Plaza España–La Cistérniga–Polígono La Mora.

Tiene servicio en días laborables y une el centro de Valladolid con el barrio de Las Delicias, La Cistérniga y el polígono industrial de La Mora.

- **Línea 14.** Plaza España–La Cistérniga–Polígono San Cristóbal.

Esta línea presta su servicio en días laborables y conecta el centro de Valladolid y el polígono industrial de San Cristóbal, pasando por el barrio de Las Delicias y dando servicio a la Universidad Europea Miguel de Cervantes.

- **Línea 16.** Pinar de Jalón–Paseo Zorrilla 71.

Une el barrio de Pinar de Jalón y el polígono industrial de Argales con el paseo de Zorrilla y el nudo de comunicaciones de la plaza de Juan de Austria. Tiene servicio todo el día de lunes a viernes, y los sábados por la mañana.

- **Línea 17.** Rondilla–Polígono San Cristóbal.

Tiene servicio en días laborables y une el barrio de La Rondilla y el polígono industrial de San Cristóbal, pasando por el centro de Valladolid y el Hospital Universitario Río Hortega.

- **Línea 18.** Puente Duero–San Isidro–La Cistérniga.

Une el barrio de Puente Duero con el pueblo de La Cistérniga pasando por el centro de Valladolid, además de la plaza de toros de Valladolid, el parque del Campo Grande, el Laboratorio de las Artes (LAVA), algunos centros comerciales y el aparcamiento disuasorio de la calle La Vía. Los domingos y festivos solo realiza el recorrido entre la Plaza España y La Cistérniga.

- **Línea 19.** Puente Duero–Hospital Río Hortega–La Cistérniga.

Une el barrio de Puente Duero con el pueblo de La Cistérniga pasando por el centro de Valladolid, además de enclaves como el Hospital Río Hortega, el Polígono Industrial de San Cristóbal, la plaza de toros de Valladolid, el Laboratorio de las Artes (LAVA), algunos centros comerciales, y el aparcamiento disuasorio de la calle La Vía. Comparte casi todo su recorrido con la línea 18, cuyas frecuencias se alternan cada media hora. Así, ambas asumen la demanda de viajeros entre Valladolid y La Cistérniga, y entre el centro y sur de la capital y con Puente Duero y el Pinar de Antequera.

- **Línea 24.** La Overuela–Plaza Poniente.

Une los barrios de La Overuela y La Victoria con el centro de Valladolid, pasando por el barrio de Huerta del Rey por un lado y por el Hospital Clínico por otro. Da servicio a Mercaolid y al polígono industrial anejo a la avenida de Burgos, así como a los aparcamientos disuasorios de Paseo del Renacimiento, Feria de Muestras y Las Moreras.

- **Líneas C1 y C2.** Parquesol–La Victoria–Parquesol.

Se trata de dos líneas circulares que arrancan en el barrio de Parquesol atravesando diversos barrios de la ciudad hasta llegar al barrio de la Victoria para, de nuevo, dirigirse hacia Parquesol.

- **Línea H.** Es una línea a horario que une el paseo de Zorrilla de Valladolid, en el nudo de comunicaciones de la plaza de Juan de Austria, con el Hospital Río Hortega a través del barrio de Las Delicias. Tiene servicio de lunes a viernes laborables y sirve de lanzadera para enlazar con gran parte de las líneas de autobús que provienen o se dirigen al sur de la ciudad.

– **Líneas Búho** que prestan servicio nocturno viernes y vísperas de festivo (salvo en Nochebuena y Nochevieja) a partir de las 23:30 y hasta las 3:00 horas. Durante las fiestas de la ciudad de Valladolid (primera semana de septiembre) se da servicio todos los días, teniendo la última salida del centro a las 4:00 en viernes y vísperas de festivo, y el resto de días a las 2:00 horas. Esta red cuenta con tres líneas diametrales (B1, B2 y B3) y dos radiales (B4 y B5). Todas salen del centro de Valladolid a las horas en punto facilitando de esta forma los transbordos. Actualmente cuenta con las siguientes cinco líneas:

- **Línea B1.** Covaresa–Fuente Dorada–San Pedro Regalado.
- **Línea B2.** Delicias–Doctrinos–La Victoria.
- **Línea B3.** Parquesol–Fuente Dorada–Las Flores.
- **Línea B4.** Plaza Zorrilla–Entrepinos–(Puente Duero).
- **Línea B5.** Duque de la Victoria–La Cistérniga.

– **Líneas a polígonos industriales** que complementan las líneas Ordinarias para dar servicio a las dos grandes áreas industriales de Valladolid, los Polígonos de Argales y de San Cristóbal con salidas a las 5:10/5:15 y 6:15 horas de lunes a viernes no festivos. Se trata de 9 líneas que unen los barrios periféricos de Valladolid, el centro de la ciudad y los polígonos industriales de Argales y San Cristóbal a primera hora de la mañana. Solo tienen recorrido de ida y durante el mes de agosto se suprimen los servicios de las 5:10/5:15 horas. Estas líneas son:

- **Línea P1.** San Pedro Regalado–Polígono Argales–La Rubia.
- **Línea P2.** Barrio España–Polígono Argales–Covaresa.
- **Línea P3.** La Victoria–Polígono San Cristóbal.
- **Línea P6.** La Victoria–Polígono Argales.
- **Línea P7.** Los Santos Pilarica–Polígono San Cristóbal.

- **Línea P13.** Covaresa–Polígono San Cristóbal.
 - **Línea PSC1.** Parquesol–Polígono Argales–Polígono San Cristóbal.
 - **Línea PSC2.** Rondilla–Pajarillos–Polígono San Cristóbal.
 - **Línea PSC3.** Las Flores–Delicias–Polígono Argales–Parquesol.
- **Líneas matinales.** Las líneas de la red Matinal replican parte de los recorridos del servicio ordinario desde la periferia de Valladolid. Salen de la parada de origen a las 6:50 en dirección al centro de la ciudad de lunes a viernes no festivos. Solo tienen recorrido de ida. Estas líneas son:
- **Línea M1.** Covaresa – Fuente Dorada.
Recorre a primera hora de la mañana el barrio de Covaresa, Parque Alameda, el paseo de Zorrilla, las estaciones de autobuses y ferrocarriles y el centro de Valladolid.
 - **Línea M2.** San Pedro Regalado – Fuente Dorada.
Recorre el barrio de San Pedro Regalado con el Hospital Clínico y el centro de Valladolid.
 - **Línea M3.** Delicias (Hosp. Río Hortega) – Fuente Dorada.
Comunica el Hospital Río Hortega y el barrio de Delicias con el centro de Valladolid.
 - **Línea M4.** La Victoria – Plaza España.
Une los barrios de La Victoria, Girón y Huerta del Rey con el centro de Valladolid.
 - **Línea M5.** Parquesol – Plaza España.
Une los barrios de Parquesol, Villa de Prado y Huerta del Rey con el centro de Valladolid.
 - **Línea M6.** Las Flores – Fuente Dorada.
Comunica los barrios de Las Flores, Pilarica y Belén y el entorno del Hospital Clínico con el centro de Valladolid.
 - **Línea M7.** La Cistérniga – Fuente Dorada.
Comunica el pueblo de La Cistérniga con el centro de Valladolid.
- **Líneas universitarias.** Se trata de 2 líneas lanzadera que recorren los ejes Sur–Norte y Suroeste–Norte a las 7:20 y 8:20 horas en dirección al centro de la ciudad y el Campus Miguel Delibes de lunes a viernes lectivos, según el calendario académico de la Universidad de Valladolid. Estas líneas sólo tienen recorrido de ida.

- **Línea U1.** Covaresa – Campus Miguel Delibes.

Recorre a primera hora de la mañana el barrio de Covaresa, Parque Alameda, el paseo de Zorrilla, las estaciones de autobuses y ferrocarriles, el centro de la ciudad y los barrios de Belén y Pilarica.

- **Línea U8.** Parquesol – Campus Miguel Delibes.

Une los barrios vallisoletanos de Parquesol, Villa de Prado y Huerta del Rey con el centro de la ciudad y la zona de facultades del Campus Miguel Delibes.

- **Líneas especiales.** Hay dos tipos de líneas, Líneas F (fútbol) y Líneas R (que llegan al recinto ferial la semana de Ferias de la ciudad).

Las Líneas F entran en servicio en los días en que el Real Valladolid juega partidos como local, en el estadio José Zorrilla. Parten de la parada de origen en los distintos barrios de la ciudad una hora antes del comienzo de los partidos y, una vez acabado el encuentro, al poco tiempo los autobuses parten del estadio para regresar a las paradas de origen. Estas son:

- **Línea F1.** Covaresa – La Rubia - Arturo Eyrías – Estadio José Zorrilla.

Sale del barrio de Covaresa de Valladolid, pasa por el Paseo de Zorrilla y los barrios de La Rubia y Arturo Eyrías, y se dirige hacia el Estadio José Zorrilla.

- **Línea F2.** Delicias (Hosp. Río Hortega) – Plaza Circular – Estaciones – Plaza Juan de Austria – Estadio José Zorrilla.

Sale del barrio vallisoletano de Las Delicias, pasa por las estaciones, de tren y autobús, y la plaza Juan de Austria para dirigirse hacia el Estadio José Zorrilla.

- **Línea F3.** Las Flores – Pajarillos – Doctrinos – Estadio José Zorrilla.

sale del barrio vallisoletano de Las Flores, pasa por el centro de la ciudad y se dirige hacia el Estadio José Zorrilla.

- **Línea F4.** Los Santos-Pilarica – Doctrinos – Estadio José Zorrilla.

Sale del barrio vallisoletano de Los Santos Pilarica, pasa por el Barrio de La Pilarica, el centro de la ciudad y se dirige hacia el Estadio José Zorrilla.

- **Línea F5.** Belén (Universidad) – Angustias – Doctrinos – Estadio José Zorrilla.

Sale del barrio vallisoletano de Belén, en las cercanías del Campus Miguel Delibes y el apeadero de Valladolid-Universidad, pasa por el entorno del Hospital Clínico y el centro de la ciudad y se dirige hacia el Estadio José Zorrilla.

- **Línea F6.** San Pedro Regalado – Rondilla – La Victoria – Estadio José Zorrilla.

Sale del barrio de San Pedro Regalado de Valladolid, pasa por los barrios de Los Viveros, Rondilla, La Victoria y Huerta del Rey y se dirige hacia el Estadio José Zorrilla.

Las líneas R circulan durante la semana de las Ferias y fiestas de la Virgen de San Lorenzo de Valladolid. Estas líneas son:

- **Línea R.** Recinto Ferial - Plaza Poniente (directo).
Sirve de lanzadera entre el centro de la ciudad y las dársenas del estadio José Zorrilla, junto al que se sitúa el recinto ferial.
- **Línea R1.** Recinto Ferial - San Pedro Regalado (nocturno).
Es un servicio nocturno que une el Recinto Ferial con los barrios de Huerta del Rey, el centro, La Rondilla y San Pedro Regalado.
- **Línea R2.** Recinto Ferial - Covaresa (nocturno).
Es un servicio nocturno que une el Recinto Ferial con los barrios de Parquesol, Arturo Eyriès, La Rubia y Covaresa.
- **Línea R3.** Recinto Ferial - La Victoria (nocturno).
Servicio nocturno que une el Recinto Ferial con el barrio de La Victoria pasando por los barrios de Huerta del Rey y Girón.
- **Línea R4.** Recinto Ferial - Las Flores (nocturno).
Servicio nocturno que comunica el Recinto Ferial con el paseo de Zorrilla y los barrios de Las Delicias, Pajarillos y Las Flores.
- **Línea R5.** Recinto Ferial - Belén (Universidad, nocturno).
Servicio nocturno que comunica el Recinto Ferial con el centro de la ciudad y los barrios de Pilarica y Belén.

– **Otros servicios.**

- **Línea 23.** Paseo de Zorrilla 130 – Cañada Real
- **Línea 26.** Plaza España – San Juan de Dios

– **Otros servicios especiales.** Cada año con motivo del Día de Todos los Santos entran en servicio dos líneas especiales a los cementerios vallisoletanos de Las Contiendas y de El Carmen, con recorridos radiales desde el centro de la ciudad. También, ante eventos alejados del núcleo urbano en la Antigua Hípica Militar, como la concentración motera Pingüinos, se establece una línea lanzadera desde la zona centro con parada en varios puntos de la ciudad.

- **La línea Cementerio El Carmen.**
- **La Línea Cementerio de Las Contiendas.**
- **La línea Concentración motera Pingüinos.**

- **Bus turístico.** Valladolid cuenta con un autobús turístico de doble piso con piso superior descubierto. Está dotado de un servicio de información sonora en cada asiento, con cinco canales para cinco idiomas (alemán, español, francés, inglés y portugués), que relata brevemente las características urbanísticas de la zona que va recorriendo, haciendo hincapié en los edificios, conventos, monumentos, museos, espacios naturales y palacios de su entorno y alrededores, con datos de interés de carácter histórico, artístico, arquitectónico y urbano, combinado con datos y fechas de interés cultural y festivo. El vehículo tiene una capacidad total de ochenta plazas, una de ellas para silla de ruedas en el piso inferior, y lleva rampa motorizada para acceso de personas con movilidad reducida en su puerta central. El recorrido tiene una duración aproximada de 50 minutos.

Los horarios que tiene este servicio son:

- Durante el invierno (1 de octubre a 30 de marzo):

Viernes: 17:00 h - 18:00 horas.

Sábados y domingos: 12:00 h - 13:00 h - 17:00 h - 18:00 horas.

- Durante el verano (1 de abril a 30 de septiembre):

Viernes: 17:00 h - 18:00 h -19:00 horas.

Sábados y domingos: 12:00 h - 13:00 h - 17:00 h - 18:00 h - 19:00 horas.

Además de estos horarios realiza servicios especiales concertados en días laborables que no tengan el servicio habitual.

El autobús realiza el siguiente trayecto:

Parada 1. Acera de Recoletos fte. n.º 6

Parada 2. Paseo Isabel la Católica esq. San Agustín

Parada 3. Calle San Quirce, 5

Parada 4. Calle Angustias, 21

Parada 5. Plaza de Portugalete

Parada 6. Calle Colón, 18

Parada 7. Calle Real de Burgos, 3

Parada 8. Plaza de la Universidad

6.2. ANEXO II. Flota de autobuses de AUVASA (junio 2023)

Nº Bus	Matricula	Marca	Modelo	Año Matr.	Combustible	Piadas	Rampa
001	9209 GZB	MERCEDES BENZ	628 02 CITARO	2010	GASÓLEO	108	M/A
002	9242 GZB	MERCEDES BENZ	628 02 CITARO	2010	GASÓLEO	108	M/A
003	4685 HCM	MERCEDES BENZ	628 02 CITARO	2011	GASÓLEO	108	M/A
004	4647 HCM	MERCEDES BENZ	628 02 CITARO	2011	GASÓLEO	108	M/A
005	3918 HCM	MERCEDES BENZ	628 02 CITARO	2011	GASÓLEO	108	M/A
006	6223 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
007	6340 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
008	6450 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
009	6481 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
010	6526 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
011	6555 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
012	8551 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
013	8753 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
014	9858 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
015	9871 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
016	9886 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
017	9895 KFK	IVECO	URBANWAY	2017	GASÓLEO	104	A
060	5719 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
061	5680 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
062	5655 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
063	5625 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
064	5603 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
065	5754 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
066	5574 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
067	5554 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
068	5541 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
069	5516 GHY	MAN	LIONS CITY 273 F/GLP (A21)	2008	GLP	94	M/A
070	4503 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
071	4538 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
072	4771 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
073	4799 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
074	4838 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
075	4870 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
076	4914 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
077	4940 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
078	4973 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
079	4988 GPN	MAN	LIONS CITY 273 F/GLP (A21)	2009	GLP	94	M/A
119	7745 BPP	MAN	NL 243 F GLP (J22)	2001	GLP	88	A
120	7746 BPP	MAN	NL 243 F GLP (J22)	2001	GLP	88	A
121	7747 BPP	MAN	NL 243 F GLP (J22)	2001	GLP	88	A
122	7748 BPP	MAN	NL 243 F GLP (J22)	2001	GLP	88	A
123	7749 BPP	MAN	NL 243 F GLP (J22)	2001	GLP	88	A
124	5808 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
125	5817 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
126	5821 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
127	5824 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
128	5833 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
129	5838 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
130	5842 CCM	MAN	NL 243 F GLP (J22)	2002	GLP	88	A
131	5005 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
132	5001 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
133	4997 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
134	4980 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
135	4953 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A

Nº Bus	Matricula	Marca	Modelo	Año Matr.	Combustible	Plazas	Rampa
136	4947 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
137	4939 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
138	4991 CHG	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
139	3414 CHP	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
140	3418 CHP	MAN	NL 243 F GLP (J22)	2003	GLP	88	A
141	6837 CHN	MAN	NL 243 F GLP (J22)	2003	GLP	91	A
142	6888 CHN	MAN	NL 243 F GLP (J22)	2003	GLP	91	A
143	3372 CRK	MAN	NL 243 F GLP (J22)	2004	GLP	91	A
144	3393 CRK	MAN	NL 243 F GLP (J22)	2004	GLP	91	A
223	4052 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
224	4166 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
225	4061 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
226	4152 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
227	4189 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
228	4114 GZM	MAN	NG 320 F ARTICULADO	2010	GASÓLEO	134	M/A
229	9789 HCD	MAN	NG 320 F ARTICULADO	2011	GASÓLEO	134	M/A
230	0442 HCF	MAN	NG 320 F ARTICULADO	2011	GASÓLEO	134	M/A
231	0307 HCF	MAN	NG 320 F ARTICULADO	2011	GASÓLEO	134	M/A
232	0516 HCF	MAN	NG 320 F ARTICULADO	2011	GASÓLEO	134	M/A
233	0590 HCF	MAN	NG 320 F ARTICULADO	2011	GASÓLEO	134	M/A
301	0365 HGS	MAN	LION S CITY HIBRIDO	2011	HIBRIDO	88	M/A
302	5207 JRN	VECTIA	VERIS.12 HYBRID	2016	HIBRIDO	93	A
303	5459 JRN	VECTIA	VERIS.12 HYBRID	2016	HIBRIDO	93	A
304	1136 JSH	VECTIA	VERIS.12 HYBRID	2016	HIBRIDO	93	A
305	1158 JSH	VECTIA	VERIS.12 HYBRID	2016	HIBRIDO	93	A
306	1192 JSH	VECTIA	VERIS.12 HYBRID	2016	HIBRIDO	93	A
307	8876 KSJ	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
308	8862 KSJ	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
309	0886 KSK	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
310	8878 KSJ	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
311	8877 KSJ	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
312	8865 KSJ	VECTIA	VERIS.12 HYBRID	2018	HIBRIDO	93	A
313	1219 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
314	1231 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
315	1249 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
316	1267 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
317	1312 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
318	1356 KZL	VECTIA	VERIS.12 HYBRID	2019	HIBRIDO	93	A
330	2194 LJW	IVECO	URBANWAY GNC	2020	GNC	101	A
331	2177 LJW	IVECO	URBANWAY GNC	2020	GNC	101	A
332	2116 LJW	IVECO	URBANWAY GNC	2020	GNC	101	A
333	5814 LMH	MERCEDES BENZ	CITARO NGT	2020	GNC	95	M/A
334	5874 LMH	MERCEDES BENZ	CITARO NGT	2020	GNC	95	M/A
335	5915 LMH	MERCEDES BENZ	CITARO NGT	2020	GNC	95	M/A
336	6005 LMH	MERCEDES BENZ	CITARO NGT	2020	GNC	95	M/A
337	7339 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
338	7436 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
339	7553 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
340	7634 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
341	7692 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
342	7841 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
343	7904 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
344	7984 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
345	5494 LXL	SOLARIS	URBINO GNC	2022	GNC	103	M/A

Nº Bus	Matricula	Marca	Modelo	Año Matr.	Combustible	Plazas	Rampa
346	8267 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
347	5530 LXL	SOLARIS	URBINO GNC	2022	GNC	103	M/A
348	8407 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
349	8479 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
350	8557 LXM	SOLARIS	URBINO GNC	2022	GNC	103	M/A
351	5586 LXL	SOLARIS	URBINO GNC	2022	GNC	103	M/A
352	0412 LYN	SOLARIS	URBINO GNC	2022	GNC	103	M/A
353	4982 LZD	SOLARIS	URBINO GNC	2022	GNC	103	M/A
354	0432 LYN	SOLARIS	URBINO GNC	2022	GNC	103	M/A
355	4370 LZG	SOLARIS	URBINO GNC	2022	GNC	103	M/A
356	0493 LYN	SOLARIS	URBINO GNC	2022	GNC	103	M/A
357	4479 LYY	SOLARIS	URBINO GNC	2022	GNC	103	M/A
358	0511 LYN	SOLARIS	URBINO GNC	2022	GNC	103	M/A
359	3313 MFH	SOLARIS	URBINO GNC	2023	GNC	103	M/A
360	3336 MFH	SOLARIS	URBINO GNC	2023	GNC	103	M/A
361	3376 MFF	SOLARIS	URBINO GNC	2023	GNC	103	M/A
362	3230 MFF	SOLARIS	URBINO GNC	2023	GNC	103	M/A
363	5426MGB	SOLARIS	URBINO GNC	2023	GNC	103	M/A
364	3355 MFH	SOLARIS	URBINO GNC	2023	GNC	103	M/A
365	9478 MFJ	SOLARIS	URBINO GNC	2023	GNC	103	M/A
401	9578 LZF	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
402	4612 LYY	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
403	4699 LYY	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
404	9648 LZF	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
405	5028 LYY	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
406	5120 LYY	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
407	2742 LZJ	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
408	4141 LZZ	SOLARIS	URBINO GNC ART.	2022	GNC	135	M/A
409	0297 MFK	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
410	0298 MFK	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
411	3220 MFH	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
412	3262 MFH	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
413	0296 MFK	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
414	3274 MFH	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
415	5532MGB	SOLARIS	URBINO GNC ART.	2023	GNC	135	M/A
601	9659 MCF	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
602	7079 MCM	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
603	8353 MDT	IRIZAR	IETRAM ART.	2023	ELÉCTRICO	105	M/A
604	7086 MCM	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
605	7084 MCM	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
606	0654 MCT	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
607	3941 MDG	IRIZAR	IETRAM ART.	2022	ELÉCTRICO	105	M/A
608	1622 MDK	IRIZAR	IETRAM ART.	2023	ELÉCTRICO	105	M/A
609	0090 MDJ	IRIZAR	IETRAM ART.	2023	ELÉCTRICO	105	M/A
610	5290 MDK	IRIZAR	IETRAM ART.	2023	ELÉCTRICO	105	M/A
611	5288 MDK	IRIZAR	IETRAM ART.	2023	ELÉCTRICO	105	M/A

6.3. ANEXO III. Estadísticas de viajeros por líneas

Líneas	2022	%	2021	%	Dif.	% Var.
Ordinarias	1.449.527	97,1%	1.060.139	97,8%	389.388	36,73%
Matinales	2.092	0,1%	1.633	0,2%	459	28,11%
Polígonos	3.892	0,3%	3.171	0,3%	721	22,74%
Búho	5.631	0,4%	0	0,0%	5.631	100,00%
Fútbol	4.784	0,3%	0	0,0%	4.784	100,00%
Otros	4.237	0,3%	829	0,1%	3.408	411,10%
C. Personal	1.645	0,1%	1.719	0,2%	-74	-4,30%
Antiguas						
Ajuste	21.464	1,4%	16.078	1,5%	5.386	33,50%
TOTALES	1.493.272	100,0%	1.083.569	100,0%	409.703	37,81%

Líneas	2022	%	2021	%	Dif.	% Var.
1	256.958	17,2%	183.049	16,9%	73.909	40,38%
2	177.324	11,9%	124.226	11,5%	53.098	42,74%
3	90.613	6,1%	67.525	6,2%	23.088	34,19%
4	17.051	1,1%	12.822	1,2%	4.229	32,98%
5	67.329	4,5%	48.502	4,5%	18.827	38,82%
6	138.956	9,3%	105.557	9,7%	33.399	31,64%
7	110.075	7,4%	79.453	7,3%	30.622	38,54%
8	129.811	8,7%	89.072	8,2%	40.739	45,74%
9	105.677	7,1%	79.278	7,3%	26.399	33,30%
10	2.830	0,2%	2.055	0,2%	775	37,71%
13	3.234	0,2%	2.837	0,3%	397	13,99%
14	4.003	0,3%	3.497	0,3%	506	14,47%
16	4.005	0,3%	3.201	0,3%	804	25,12%
17	2.835	0,2%	2.187	0,2%	648	29,63%
18	36.202	2,4%	26.973	2,5%	9.229	34,22%
19	37.509	2,5%	27.785	2,6%	9.724	35,00%
C	240.736	16,1%	183.596	16,9%	57.140	31,12%
23	0	0,0%	426	0,0%	-426	-100,00%
24	6.973	0,5%	5.722	0,5%	1.251	21,86%
26	985	0,1%	504	0,0%	481	95,44%
33	4.172	0,3%	3.043	0,3%	1.129	37,10%
H	12.249	0,8%	8.829	0,8%	3.420	38,74%
Matinal	2.092	0,1%	1.633	0,2%	459	28,11%
Polígono	3.892	0,3%	3.171	0,3%	721	22,74%
Búho	5.631	0,4%	0	0,0%	5.631	100,00%
Fútbol	4.784	0,3%	0	0,0%	4.784	100,00%
Otros	4.237	0,3%	829	0,1%	3.408	411,10%
Personal	1.645	0,1%	1.719	0,2%	-74	-4,30%
Antiguas						
Ajuste	21.464	1,4%	16.078	1,5%	5.386	33,50%
TOTALES	1.493.272	100,0%	1.083.569	100,0%	409.703	37,81%

Figura 36: Estadísticas de viajeros en el año 2022.

Estos datos son de carácter público y pueden consultarse en:

http://www.auvasa.es/auv_opendata.asp

6.4. ANEXO IV. Código en R para la elaboración de los gráficos.

ANEXO IV. Gráfico de Líneas

```
ggplot(data = incidencias, aes(x = fecha, y = cont)) +  
  geom_line(size = 0.8, color = "#5AD2B8") +  
  geom_point(color = "#5AD2B8") +  
  
  labs(title = "Incidencias totales",  
        subtitle = "Período del 15 de junio al 15 de septiembre de 2022",  
        caption = "Base de datos de incidencias de AUVASA",  
        x = NULL, y=NULL) +  
  
  theme(panel.grid = element_blank(),  
        panel.grid.major.y = element_line(size=.05,color='#143325',linetype=2),  
        panel.grid.minor.y = element_line(size=.05,color='lightgrey',linetype=2),  
        panel.grid.major.x = element_line(size=.05,color='lightgrey',linetype=2),  
        panel.background = element_rect(fill='white'),  
        panel.border = element_blank(),  
  
        axis.text = element_text(color='#143325'),  
        axis.text.x = element_text(angle =45, family = "sans"),  
  
        plot.background = element_rect(fill='white', color='white'),  
        plot.title = element_text(color="#143325",size=15),  
        plot.subtitle = element_text(color="#5AD2B8",size=9)) +  
  
  scale_y_continuous(limits = c(0,30)) +  
  scale_x_date(breaks = seq(dmy("15/06/2022"), dmy("15/09/2022"), "4 days"),  
              date_labels = "%d %b")
```

ANEXO IV. Gráfico de Barras

```
barras <- barplot(tabla,
                 beside = T,
                 xlab="Tipo de Incidencia",
                 ylab="Frecuencia relativa",
                 ylim = c(0, 40),
                 las=2,
                 col = c("#5AD2B8", "#008695"),
                 border = c("#5AD2B8", "#008695"),
                 font.axis=2,
                 col.axis="#143325",
                 cex.axis = 0.8,
                 space = c(0.3, 0.8),
                 width = 1.2,
                 cex.names = 0.7,
                 axis.lty = 1)

legend("top", legend = rownames(tabla),
      col = c("#5AD2B8", "#008695"),
      bty = "n", pch = 20,
      pt.cex = 3, cex = 1,
      box.col = "#143325",
      box.lty = 3, text.col = "#143325",
      text.font = 2, ncol = 2)

text((c(barras[1,], barras[2,] + 0.05)+0.15),
     c(tabla[1,], tabla[2,] +0.1)+0.6,
     paste0(c(tabla[1,], tabla[2,]), "%"),
     col="#143325", font = 2, cex = 0.6)
```

ANEXO IV. Gráfico para datos multivariantes usando el comando *facet_wrap*

```
ggplot(data = fechas, aes(x = day(dia), y = freq, group = año, color=año)) +
  geom_vline(mapping=aes(xintercept=8),
             color = "#5AD2B8", size = 0.8, linetype = 2) +
  geom_vline(mapping=aes(xintercept=17),
             color = "#5AD2B8",size = 0.8, linetype = 2) +
  geom_vline(mapping=aes(xintercept=9),
             color = "#008695", size = 0.8, linetype = 2) +
  geom_vline(mapping=aes(xintercept=0),
             color = "#008695", size = 0.8, linetype = 2) +
  geom_line() +
  geom_point() +

  scale_color_manual(values=c("#5AD2B8","#008695"), name = "Año") +
  labs(title = "Tipos de Incidencias en abril",
       subtitle = "Tipos de Incidencias más frecuentes",
       caption = "Base de datos de incidencias de AUVASA",
       x = "Día", y=NULL) +

  theme(panel.grid = element_blank(),
        panel.grid.major.y = element_line(size=0.05,color='#143325',linetype=2),
        panel.grid.major.x = element_line(size=0.05,color='grey65',linetype=3),
        panel.background = element_rect(fill='white'),
        panel.border = element_blank(),

        axis.text = element_text(color='#143325'),

        plot.background = element_rect(fill='white', color= "#143325", size=1),
        plot.title = element_text(color="#143325",size=15),
        plot.subtitle = element_text(color="#5AD2B8",size=9, face = "bold"),
        plot.caption = element_text(color="#143325"),

        strip.text=element_text(face='bold', size=11, hjust=0, color='white'),
        strip.background=element_rect(fill="#5AD2B8"),

        legend.key = element_rect(color = "#143325", size = 0.7, fill = "white"),
        legend.position = "bottom") +

  scale_x_continuous(breaks = c(1:30), guide = guide_axis(n.dodge = 2)) +
  scale_y_continuous(limits = c(0,15)) +
  facet_wrap(fechas$tipo, ncol=2)
```

6.5. ANEXO V. Código en R para la elaboración del calendario de incidencias usando la librería ggplot2.

5.5.1. Preparación del data.frame para el calendario

```
# Creación del data.frame con datos para el calendario

df <- data.frame(
  fecha = seq(
    dmy("01/01/2023"),
    dmy("31/12/2023"),
    "days"),
  cont = rep(NA, 365))

d <- data.frame(table(datos23$Fecha))

names(d) <- c("fecha", "cont")
d$fecha <- as.Date(d$fecha)

df <- merge(df, d, by="fecha", all= T)
df <- df[,-2]
names(df) <- c("fecha", "cont")
df$cont[13] <- 0

calendario <- data.frame(fecha=df$fecha,
  cont=df$cont,
  weekday=yday(df$fecha, week_start=1, label=T, abbr=F),
  mes=month(df$fecha, abbr = F, label = T),
  semana=isoweek(df$fecha),
  dia=day(df$fecha),
  texto = rep("black", 365))

calendario$texto[df$cont > 20] <- "white"
calendario$semana <- calendario$semana+1
calendario$semana[calendario$mes == "enero" &
  (calendario$semana == 53 | calendario$semana == 54)] <- 1
```

5.5.2. Código del Heatmap Calendar usando la libreria ggplot2

```
# Funcion para poner en mayúscula la letra inicial de los meses
capitalize <- function(string) {
  substr(string, 1, 1) <- toupper(substr(string, 1, 1))
  string
}

# Gráfico del Heatmap Calendar usando la libreria ggplot2
ggplot(calendario, aes(weekday,-semana, fill = cont)) +
  geom_tile(colour = "white") +
  geom_text(aes(label = day(df$fecha)), size = 2.5, color = calendario$texto) +
  theme(aspect.ratio = 1/2,
        axis.title.x = element_blank(),
        axis.title.y = element_blank(),
        axis.text.y = element_blank(),
        axis.ticks = element_blank(),

        plot.title = element_text(face="bold", size = 15),
        plot.subtitle = element_text(color = "grey20"),
        plot.caption = element_text(color = "grey18", size = 7),

        panel.grid = element_blank(),
        panel.background = element_blank(),
        panel.border = element_rect(colour = "#143325", fill=NA, size=1),

        strip.background = element_blank(),
        strip.text = element_text(face = "bold", size = 14)) +
  scale_fill_gradientn(colours = c('white', '#5AD2B8', '#143325'),
                      na.value = "grey90",
                      name= "Incidencias",
                      n.breaks = 15) +
  guides(fill = guide_legend(title.position = "top",
                             reverse = T, ncol = 1,
                             keywidth = 1.2, keyheight = 0.7,
                             label.theme = element_text(color='#143325', size=9),
                             title.theme = element_text(size = 9))) +
  scale_x_discrete(labels=c("Lu","Ma","Mi","Ju","Vi","Sa","Do")) +
  facet_wrap(~mes, nrow = 4, ncol = 3, scales = "free",
            labeller = labeller(mes = capitalize)) +
  labs(title = "Incidencias de 2023",
       subtitle = "Incidencias totales",
       caption = "Datos: incidencias AUVASA")
```


ANEXO VI. Incidencias de tipo PB (Bajas, absentismo, ...)

Incidencias de tipo PB de 2022

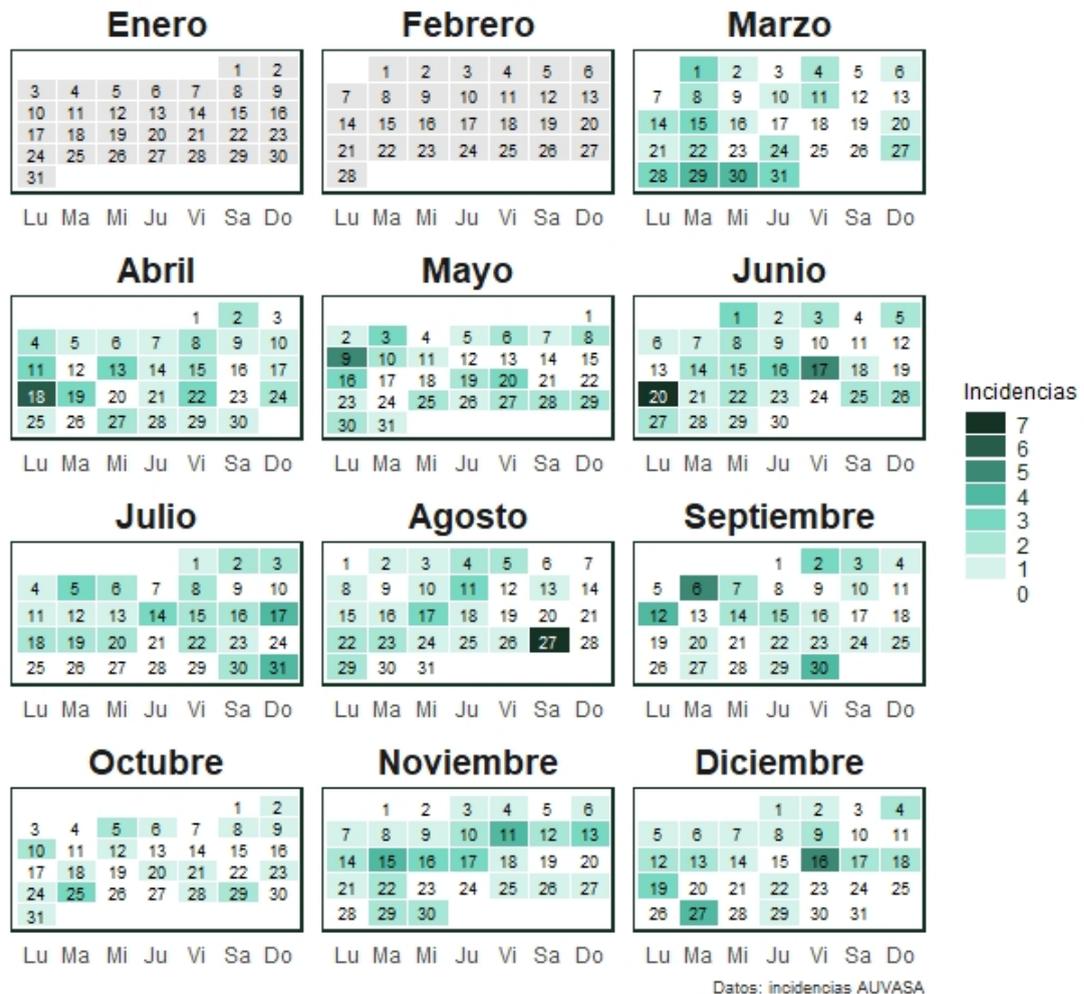


Figura 38: Calendario de Incidencias tipo PB en 2022.

ANEXO VI. Incidencias de tipo PH (Ampliación de jornada, horas extra,...)

Incidencias de tipo PH de 2022



Datos: incidencias AUVASA

Figura 39: Calendario de Incidencias tipo PH en 2022.

6.7. ANEXO VII. Tablas de frecuencias absolutas y condicionadas por año según las líneas de autobús.

Línea \ Año	2022		2023	
	Fr. Abs.	% Fr. Cond.	Fr. Abs.	% Fr. Cond.
1	250	13.6686714	142	15.5021834
2	189	10.3335156	91	9.9344978
6	156	8.5292510	68	7.4235808
C1	153	8.3652269	75	8.1877729
8	152	8.3105522	86	9.3886463
C2	144	7.8731547	52	5.6768559
7	141	7.7091307	40	4.3668122
9	117	6.3969382	51	5.5676856
3	103	5.6314926	62	6.7685590
5	82	4.4833242	86	9.3886463
18	52	2.8430837	34	3.7117904
19	50	2.7337343	18	1.9650655
4	43	2.3510115	12	1.3100437
24	26	1.4215418	8	0.8733624
H	19	1.0388190	22	2.4017467
10	16	0.8747950	7	0.7641921
B2	14	0.7654456	3	0.3275109
B1	13	0.7107709	6	0.6550218
14	11	0.6014215	7	0.7641921
16	11	0.6014215	6	0.6550218
B3	10	0.5467469	1	0.1091703
B4	10	0.5467469	6	0.6550218
26	7	0.3827228	1	0.1091703
13	6	0.3280481	6	0.6550218
17	4	0.2186987	3	0.3275109
18, 19	4	0.2186987	1	0.1091703
RR	4	0.2186987	0	0.0000000
33	3	0.1640241	0	0.0000000
F1	3	0.1640241	1	0.1091703
23	2	0.1093494	1	0.1091703
B5	2	0.1093494	0	0.0000000
F6	2	0.1093494	0	0.0000000

Tabla 6: Incidencias por líneas. Frecuencias absolutas y condicionadas. Parte 1

Línea \ Año	2022		2023	
	Fr. Abs.	% Fr. Rel.	Fr. Abs.	% Fr. Rel.
M1	2	0.1093494	2	0.2183406
P1	2	0.1093494	1	0.1091703
P3	2	0.1093494	0	0.0000000
PSC2	2	0.1093494	0	0.0000000
U8	2	0.1093494	1	0.1091703
17, 19, P3	1	0.0546747	0	0.0000000
3, 18, 19	1	0.0546747	0	0.0000000
4, H	1	0.0546747	0	0.0000000
6, H	1	0.0546747	0	0.0000000
7, 8	1	0.0546747	0	0.0000000
7, 9	1	0.0546747	0	0.0000000
C1, C2	1	0.0546747	1	0.1091703
E13	1	0.0546747	0	0.0000000
E22	1	0.0546747	0	0.0000000
F2	1	0.0546747	1	0.1091703
F5	1	0.0546747	0	0.0000000
M2	1	0.0546747	0	0.0000000
M3	1	0.0546747	0	0.0000000
M4	1	0.0546747	1	0.1091703
M7	1	0.0546747	0	0.0000000
P7	1	0.0546747	0	0.0000000
PSC1	1	0.0546747	0	0.0000000
R	1	0.0546747	0	0.0000000
Refuerzo	1	0.0546747	0	0.0000000
U1	1	0.0546747	1	0.1091703
1, 18, 19	0	0.0000000	1	0.1091703
1, 2, 7, 8	0	0.0000000	1	0.1091703
2, 9, P2, B1	0	0.0000000	1	0.1091703
3, 5, 6, 8	0	0.0000000	1	0.1091703
C1, C2, 5, 3	0	0.0000000	1	0.1091703
E11	0	0.0000000	1	0.1091703
E2	0	0.0000000	1	0.1091703
F3	0	0.0000000	1	0.1091703
F4	0	0.0000000	1	0.1091703
M6	0	0.0000000	2	0.2183406
RR-B1	0	0.0000000	1	0.1091703

Tabla 7: Incidencias por líneas. Frecuencias absolutas y condicionadas. Parte 2

Línea	ACC	COM	PAR	PER	TIC	VEH
1	49	58	3	73	10	57
2	45	44	4	59	4	33
C1	38	52	1	40	3	19
8	37	27	2	53	9	24
9	37	26	4	40	2	8
C2	28	30	4	55	7	20
3	27	12	0	34	0	30
6	27	38	4	61	4	22
7	24	21	9	51	11	25
18	17	14	0	14	1	6
5	14	16	3	28	4	17
19	10	10	2	20	0	8
24	4	2	1	16	0	3
4	4	10	4	14	4	7
13	3	1	0	1	0	1
14	2	2	0	6	0	1
B2	2	4	0	5	1	2
H	2	6	0	6	3	2
16	1	3	1	5	0	1
B1	1	11	0	1	0	0
B3	1	2	0	3	1	3
B4	1	4	1	2	1	1
B5	1	0	0	1	0	0
10	0	4	1	8	0	3
17	0	1	0	3	0	0
18, 19	0	0	0	4	0	0
3, 18, 19	0	1	0	0	0	0
33	0	2	1	0	0	0
4, H	0	1	0	0	0	0
6, H	0	1	0	0	0	0
7, 8	0	1	0	0	0	0
7, 9	0	1	0	0	0	0
C1, C2	0	1	0	0	0	0

Tabla 8: Incidencias por líneas Ordinarias y Búho para los grupos de incidencias en 2022
Frecuencias absolutas.

Línea	ACC	COM	PAR	PER	TIC	VEH
1	35	51	0	39	7	10
C1	31	18	0	21	2	3
8	26	21	0	30	1	8
C2	17	13	1	15	1	5
9	15	10	1	20	1	4
5	14	23	1	44	1	3
6	14	22	1	27	0	4
3	13	19	2	20	0	8
7	12	17	0	9	0	2
2	10	24	2	40	1	14
18	8	8	0	15	1	2
19	5	6	1	5	0	1
H	4	2	0	14	1	1
24	3	3	0	1	0	1
10	2	3	1	0	0	1
14	2	3	0	2	0	0
13	1	2	0	2	0	1
17	1	1	0	0	0	1
1, 18, 19	0	0	1	0	0	0
1, 2, 7, 8	0	1	0	0	0	0
16	0	4	0	2	0	0
18, 19	0	0	0	1	0	0
3, 5, 6, 8	0	1	0	0	0	0
4	0	3	0	7	1	1
B1	0	4	0	2	0	0
B2	0	2	0	1	0	0
B3	0	0	1	0	0	0
B4	0	5	0	1	0	0
C1, C2	0	0	1	0	0	0
C1, C2, 5, 3	0	1	0	0	0	0

Tabla 9: Incidencias por líneas Ordinarias y Búho para los grupos de incidencias en 2023
Frecuencias absolutas.