



***CUADRO DE MANDOS Y  
MODELO PREDICTIVO  
EMPRESA DE BICICLETAS***

**SOLUCIÓN DE INTELIGENCIA DE NEGOCIO**

Universidad Pablo de Olavide, Inteligencia de Negocio,  
2021

Alejandro Govantes Pola

# Índice

---

<b>1.</b>	<b>Introducción</b>	.....
<b>2.</b>	<b>Plan</b>	.....
	<b>2.1. Objetivos</b>	.....
	<b>2.2. Plan de trabajo</b>	.....
	<b>2.3. Análisis de la viabilidad</b>	.....
	<b>2.4. Riesgos</b>	.....
<b>3.</b>	<b>Análisis</b>	.....
	<b>3.1. Establecimiento de los requisitos del sistema</b>	.....
	<b>3.2. Análisis de casos de uso</b>	.....
	<b>3.3. Especificación del plan de pruebas</b>	.....
<b>4.</b>	<b>Diseño</b>	.....
	<b>4.1. Capacidad de memoria de la organización</b>	.....
	<b>4.2. Capacidad de integración de información</b>	.....
	<b>4.3. Capacidad de crear conocimiento</b>	.....
<b>5.</b>	<b>Implementación</b>	.....
	<b>5.1. Pasos para la instalación y configuración de los datos</b>	.....
	<b>5.2. Estudio de los datos</b>	.....
	<b>5.3. Cuadro de mandos</b>	.....
	<b>5.4. Modelo predictivo</b>	.....
<b>6.</b>	<b>Despliegue</b>	.....
<b>7.</b>	<b>Conclusiones</b>	.....

## 1. Introducción

Somos una empresa que tiene como objetivo analizar las bases de datos de los clientes que contactan con nuestros servicios para, a partir de dicho análisis, obtener la mayor cantidad de beneficios posibles.

En este caso vamos a trabajar con una empresa que se encarga de la venta de bicicletas. Nuestros clientes poseen una base de datos que se comprende entre 2018 y 2019, y es mediante la cual vamos a realizar los tratamientos para obtener la información que nos ayude a potenciar el rendimiento de la empresa.

La empresa ha obtenido buenos resultados en los últimos años pero se observa una falta de aprovechamiento de los datos clarísima. No se han aprovechado los meses o estaciones con mayores beneficios y se han desaprovechado ciertos parámetros que pueden resultar interesantes.

Se ha tratado la base de datos inicial para transformar ciertos parámetros y convertirlos de números a palabras para que fuese más sencillo su posterior tratamiento con Power BI.

La idea es cargar los datos para obtener gráficas y porcentajes que nos permitan sacar conclusiones, además de la realización de un modelo predictivo que ayude a saber los parámetros más importantes a la hora de obtener beneficios y conocer el futuro empresa.

Para la realización del cuadro de mandos utilizaremos la aplicación Power BI, que nos permitirá la creación de gráficas y distintos estudios del dataset. Y para el modelo predictivo se usará RStudio, mediante el que obtendremos datos relevantes para el porvenir.

## 2. Plan

### 2.1 Objetivos

Se presentan los objetivos identificados para la solución:

<b>OBJ-01</b>	<b>Creación de un cuadro de mandos</b>
<b>Versión</b>	01
<b>Autores</b>	Alberto Morales Serrano
	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Federico González Acosta
<b>Fuente</b>	Autores del proyecto
<b>Descripción</b>	Se llevará a cabo el desarrollo de un cuadro de mandos que permite consultar información relevante sobre el negocio de ventas de bicicletas.
<b>Importancia</b>	Alta
<b>Estado</b>	Finalizado
<b>Comentarios</b>	--

<b>OBJ-02</b>	<b>Análisis de datos estadísticos</b>
<b>Versión</b>	01
<b>Autores</b>	Alberto Morales Serrano
	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Federico González Acosta
<b>Fuente</b>	Autores del proyecto
<b>Descripción</b>	Análisis de las diferentes variables proporcionadas en el conjunto de datos seleccionado y cómo estas influyen en la venta de bicicletas.
<b>Importancia</b>	Alta
<b>Estado</b>	Finalizado
<b>Comentarios</b>	--

<b>OBJ-03</b>	<b>Estudio de la estacionalidad del negocio</b>
<b>Versión</b>	01
<b>Autores</b>	Alberto Morales Serrano
	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Federico González Acosta
<b>Fuente</b>	Autores del proyecto
<b>Descripción</b>	Se pretende obtener información que permita estudiar la estacionalidad del negocio. Esto se integrará en el cuadro de mandos.
<b>Importancia</b>	Alta
<b>Estado</b>	Finalizado
<b>Comentarios</b>	--

<b>OBJ-04</b>	<b>Visualización de factores relevantes</b>
<b>Versión</b>	01
<b>Autores</b>	Alberto Morales Serrano
	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Federico González Acosta
<b>Fuente</b>	Autores del proyecto
<b>Descripción</b>	Se crearán vistas en el cuadro de mandos que permitan relacionar factores relevantes, al margen de la estacionalidad, con la venta de bicicletas.
<b>Importancia</b>	Alta
<b>Estado</b>	Finalizado
<b>Comentarios</b>	--

## 2.2 Plan de trabajo

Planificación de las tareas del proyecto:

Tarea	Descripción	Duración
<b>1. Memoria del proyecto</b>	<b>Elaboración de memoria del proyecto.</b>	<b>9 días</b>
Introducción	Creación de una breve introducción al proyecto.	1 día
Plan	Definición de objetivos del proyecto, plan de trabajo, análisis de viabilidad y riesgos.	1 día
Análisis	Análisis general del proyecto, establecimiento de requisitos, casos de uso y pruebas.	1 día
Diseño	Establecimiento del proceso de diseño del proyecto	3 días
Implementación	Desarrollo de la implementación del sistema	1 día
Despliegue	Despliegue del sistema desarrollado	1 día
Conclusiones	Conclusiones al cierre del proyecto	1 día
<b>2. Búsqueda de datos</b>	<b>Proceso de búsqueda del conjunto de datos que va a utilizarse en el proyecto</b>	<b>4 días</b>
Obtención del conjunto de datos	Búsqueda de un conjunto de datos interesante.	1 día
Tratamiento de datos	Preparación de los datos del conjunto para su posterior estudio	3 días
<b>3. OBJ-01</b>	<b>Se llevará a cabo el desarrollo de un cuadro de mandos que permite consultar información relevante sobre el negocio de ventas de bicicletas.</b>	<b>4 días</b>
Implementación	Creación del cuadro de mandos con Power BI	3 días
Pruebas	Pruebas del cuadro de mandos	1 día

4. OBJ-02	Análisis de las diferentes variables proporcionadas en el conjunto de datos seleccionado y cómo estas influyen en la venta de bicicletas.	3 días
Implementación	Análisis del conjunto de datos.	2 días
Pruebas	Pruebas del análisis de datos.	1 día
5. OBJ-03	<b>Se pretende obtener información que permita estudiar la estacionalidad del negocio.</b>	2 días
Implementación	Inclusión de los datos y vistas correspondientes en el cuadro de mandos.	1 día
Pruebas	Pruebas de las vistas y datos del cuadro de mandos.	1 día
6. OBJ-04	<b>Se crearán vistas en el cuadro de mandos que permitan relacionar factores relevantes, al margen de la estacionalidad, con la venta de bicicletas.</b>	2 días
Implementación	Creación de las vistas en el cuadro de mandos.	1 día
Pruebas	Pruebas del cuadro de mandos.	1 día

Paquetes de trabajo:

Paquete de trabajo	Duración	Responsable
Memoria del proyecto	9 días	Equipo completo
Búsqueda de datos	4 días	Equipo completo
OBJ-01	4 días	Equipo completo
OBJ-02	3 días	Equipo completo
OBJ-03	2 días	Equipo completo
OBJ-04	2 días	Equipo completo
<b>Total</b>	24 días	

Paquete de trabajo	Fecha Inicio	Fecha Fin
Memoria del proyecto	01/05/21	9/05/21
Búsqueda de datos	10/05/21	13/05/21
OBJ-01	14/05/21	17/05/21
OBJ-02	18/05/21	20/05/21
OBJ-03	21/05/21	22/05/21
OBJ-04	23/05/21	24/05/21

## 2.3 Análisis de la viabilidad

El equipo al completo ha dispuesto de todos los materiales necesarios para el correcto desarrollo del proyecto sin incurrir en gastos adicionales.

Tanto el software utilizado durante el desarrollo de la memoria, Google Docs, como el utilizado en la implementación, Power BI, son herramientas gratuitas.

Los costes ligados al desarrollo de la actividad del equipo, considerando un coste de 8€/h por cada uno de los miembros del equipo, son los siguientes:

Paquete de trabajo	Total Horas	Coste Tarea
Memoria del proyecto	288 h	2.304€
Búsqueda de datos	128 h	1.024€
OBJ-01	128 h	1.024€
OBJ-02	96 h	768€
OBJ-03	64 h	512€
OBJ-04	64 h	512€
<b>Totales</b>	<b>768 h</b>	<b>6.144€</b>

## 2.4 Riesgos

Posibles riesgos durante el desarrollo del proyecto:

### 1. Error en el cálculo temporal de las tareas.

- **Posibles causas:** Falta de experiencia en la elaboración de proyectos, error al estimar de forma correcta el tiempo necesario para desarrollar una tarea.
- **Consecuencias:** Retraso en los tiempos de entrega de las distintas partes del proyecto o del proyecto completo.
- **Contingencias:** Revisión de tiempos de las tareas a realizar.

### 2. Pérdida de datos de los equipos de trabajo.

- **Posibles causas:** Fallo mecánico de un disco duro, avería de un equipo, fallo humano.
- **Consecuencias:** Pérdida de tiempo y de trabajo del equipo, retraso en la entrega del proyecto final.
- **Contingencias:** Realización de copias de seguridad del trabajo realizado hasta el momento en una carpeta de respaldo en Google Drive.

### 3. Sobrecostes del proyecto.

- **Posibles causas:** Herramientas de trabajo inadecuadas, software que no nos permita realizar el tipo de trabajo que queremos llevar a cabo y obligue a la adquisición de nuevo software con licencias.
- **Consecuencias:** Sobreprecio del proyecto, debido a la adquisición de licencias y debido a retrasos por adaptación a las nuevas herramientas.
- **Contingencias:** Estudio de la documentación de las herramientas empleadas en el desarrollo del proyecto con el fin de asegurar que no existen impedimentos técnicos para implementar el proyecto con las herramientas disponibles.

### 3. Análisis

#### 3.1 Establecimiento de los requisitos del sistema

Se presentan a continuación, todos los requisitos funcionales que deberá cumplir el usuario que aborde la aplicación:

<b>RF-01</b>	<b>Carga de datos en el cuadro de mandos</b>
<b>Autores</b>	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Alberto Morales Serrano
	Federico González Acosta
<b>Fuentes</b>	Usuarios
<b>Objetivos asociados</b>	<ul style="list-style-type: none"> <li>● OBJ-01 - Creación de un cuadro de mandos</li> <li>● OBJ-04 - Visualización de factores relevantes</li> </ul>
<b>Descripción</b>	El sistema deberá generar un cuadro de mandos, en el que se cargarán los datos de la empresa y se permitirá una visualización de estos.
<b>Comentarios</b>	--

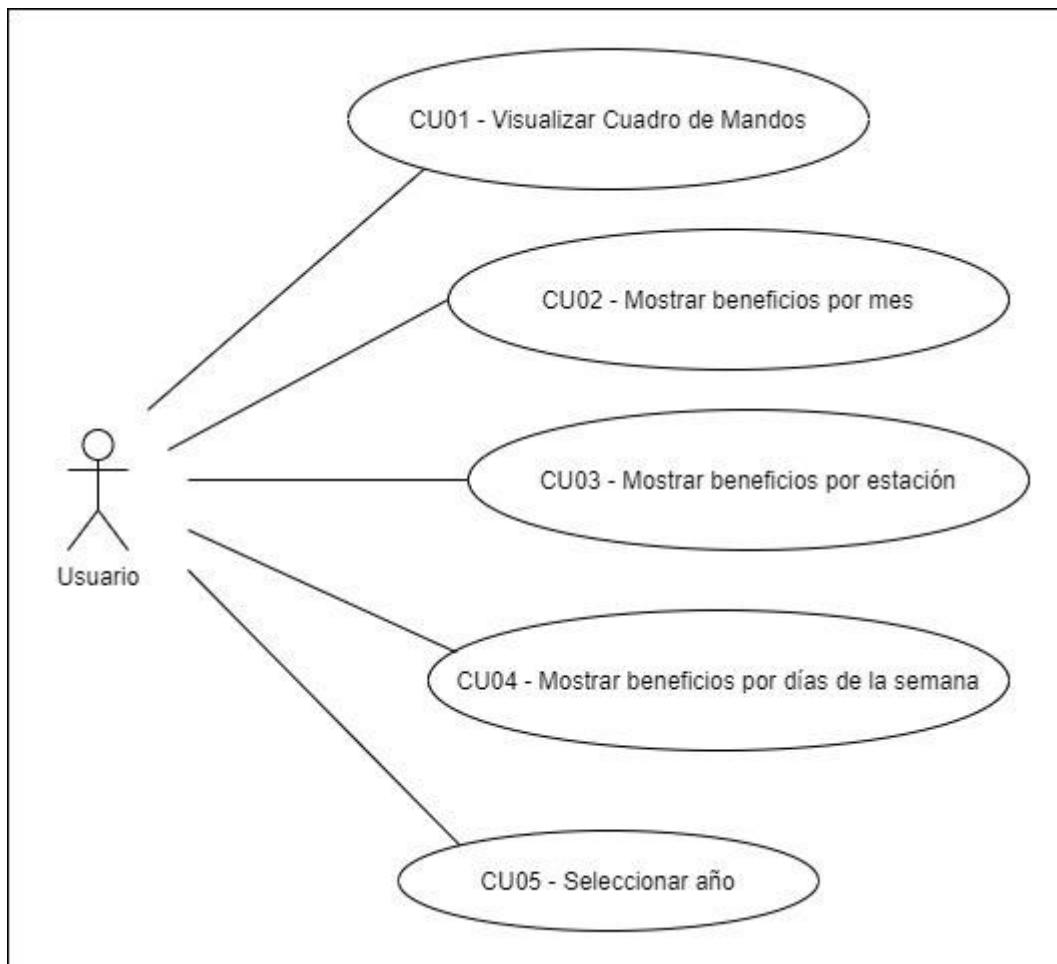
<b>RF-02</b>	<b>Mostrar beneficios en base a la estación del año</b>
<b>Autores</b>	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Alberto Morales Serrano
	Federico González Acosta
<b>Fuentes</b>	Usuarios
<b>Objetivos asociados</b>	<ul style="list-style-type: none"> <li>● OBJ-04 - Visualización de factores relevantes</li> <li>● OBJ-02 - Análisis de datos estadísticos</li> </ul>
<b>Descripción</b>	El sistema deberá permitir la visualización de los beneficios generados en cada estación del año
<b>Comentarios</b>	--

<b>RF-03</b>	<b>Mostrar beneficios mensuales</b>
<b>Autores</b>	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Alberto Morales Serrano
	Federico González Acosta
<b>Fuentes</b>	Usuarios
<b>Objetivos asociados</b>	<ul style="list-style-type: none"> <li>● OBJ-04 - Visualización de factores relevantes</li> <li>● OBJ-02 - Análisis de datos estadísticos</li> </ul>
<b>Descripción</b>	El sistema deberá permitir la visualización de los beneficios generados por cada mes
<b>Comentarios</b>	--

<b>RF-04</b>	<b>Mostrar beneficios semanales</b>
<b>Autores</b>	Alejandro Govantes Pola
	Patricia Vázquez del Cerro
	Alberto Morales Serrano
	Federico González Acosta
<b>Fuentes</b>	Usuarios
<b>Objetivos asociados</b>	<ul style="list-style-type: none"> <li>● OBJ-04 - Visualización de factores relevantes</li> <li>● OBJ-02 - Análisis de datos estadísticos</li> </ul>
<b>Descripción</b>	El sistema deberá permitir la visualización de los beneficios generados a la semana
<b>Comentarios</b>	--

### 3.2 Análisis de casos de uso

Se presenta a continuación el diagrama de casos de uso:



*Diagrama de casos de uso*

### 3.3 Especificación del plan de pruebas

Se presenta a continuación el plan de pruebas propuesto:

Pruebas de Integración e Implantación			
Pasos a seguir	Datos de entrada	Salida esperada	Observaciones
Visualizar el cuadro de mandos	DB_Bicis_PowerBI	Ninguna	--
Ejecutar RStudio	DB_Bicis_RStudio	Modelo predictivo	Obtención de datos relevantes para el porvenir

Pruebas del sistema		
Pasos a seguir	Resultados	Observaciones
Probar aplicación en distintos equipos	El sistema funciona correctamente en diferentes equipos	Los equipos deben tener instalado Power BI Desktop y RStudio
Manipulación de los datos	Se obtendrán resultados diferentes a los ya predeterminados	En este caso lo que se busca es que a pesar de modificar los datos, todo siga funcionando correctamente

## 4. Diseño

### 4.1. Capacidad de memoria de la organización

El dataset ha sido obtenido de [www.gestionexcelenciaoperacional.com](http://www.gestionexcelenciaoperacional.com). Se trata de una web encargada de realizar distintos tipos de cursos relacionados con la tecnología. Encontramos interesante esta base de datos disponible en la página y decidimos descargarla para la realización de nuestra solución de IN.

Una vez descargada la base de datos se hizo un tratamiento de la misma para que al ser cargada en Power BI se mostrara de una forma más clara y sencilla. Se cambiaron algunas columnas de forma que ciertos valores que antes eran numéricos, ahora podían asociarse directamente a lo que realmente representaban. Por ejemplo:

C	D	E	F
mes	año	estacion	dia_de_la_semana
1	2018	1	6
1	2018	1	6
1	2018	1	6
1	2018	1	6

C	D	E	F
Mes	Año	Estación	Día de la semana
Enero	2011	Invierno	Sábado
Enero	2011	Invierno	Sábado
Enero	2011	Invierno	Sábado
Enero	2011	Invierno	Sábado

#### 4.2. Capacidad de integración de información

Más allá del cambio comentado en el apartado anterior, todo lo demás estaba configurado tal y como se necesitaba para el tipo de estudio que íbamos a realizar. A la hora de cargar la base de datos en RStudio todo se pudo realizar sin tener que modificar nada ya que los datos numéricos se ajustaban al modelo predictivo.

Para el Power BI, con el dataset ya modificado, se pudieron ir obteniendo los gráficos y datos que se deseaban para poder sacar conclusiones e ideas de negocio a partir de los mismos.

Dentro del programa existen una gran cantidad de posibilidades, decenas de gráficos (circular, de colores, tablas) que permiten reconocer la información de una forma más interactiva y sencilla.

Existe la posibilidad de la realización de estudios estadísticos a partir de esos datos que le proporcionamos, tal y como podrá observarse más adelante en este documento.

La integración de la información no fue por otro motivo que para mejorar la comprensión de los datos y de esa manera hacer más vistoso el cuadro de mandos y también el modelo predictivo realizado.

#### 4.3. Capacidad de crear conocimiento

Con el análisis realizado y mediante esta solución de inteligencia de negocio podemos conocer la situación actual en la que los ciudadanos compran y usan bicicletas. También en momentos del pasado, semanas, meses y años atrás, para poder realizar un estudio de cara al futuro y predecir el momento exacto en el que las bicicletas estén más en demanda y plantear la posibilidad de poder sacar partido de esto.

Nombre del objetivo	Datos resultantes
¿Cuándo se venden más bicicletas?	Mediante los distintos gráficos realizados y estadísticas obtenidas podremos obtener los meses, estaciones o días de la semana que se venden más bicicletas.
¿Qué factores potencian la compra de bicicletas?	Mediante los distintos gráficos realizados y estadísticas obtenidas obtendremos los factores que potencian la compra.
¿Qué factores están asociados con un descenso de las ventas?	Mediante los distintos gráficos realizados y estadísticas obtenidas obtendremos los factores que afectan a la disminución de la venta.

## 5. Implementación

### 5.1. Pasos para la instalación y configuración de los datos.

Para el estudio de los datos y posteriormente la realización del cuadro de mando, se ha realizado un tratamiento previo de los datos de entrada, siendo estos los ficheros:

- DB\_Bicis\_RStudio.csv
- DB\_Bicis\_PowerBI.xlsx

El primero de ellos utilizado para transformar los valores numéricos a strings utilizando la aplicación RStudio además de asegurarnos que los datos numéricos fueran reconocidos como numéricos, ya que en algunos momentos eran detectados como strings.

Hora	Día	Mes	Año	Estación	Día de la semana	Descanso	Temperatura	Sensación térmica	Humedad	Velocidad viento	Rentas no regulares	Rentas registradas	Rentas totales
0	1	Enero	2011	Invierno	Sábado	0	9,84	14	81	0	3	13	16
1	1	Enero	2011	Invierno	Sábado	0	9,02	13	80	0	8	32	40
2	1	Enero	2011	Invierno	Sábado	0	9,02	13	80	0	5	27	32
3	1	Enero	2011	Invierno	Sábado	0	9,84	14	75	0	3	10	13
4	1	Enero	2011	Invierno	Sábado	0	9,84	14	75	0	0	1	1
5	1	Enero	2011	Invierno	Sábado	0	9,84	12	75	60	0	1	1
6	1	Enero	2011	Invierno	Sábado	0	9,02	13	80	0	2	0	2
7	1	Enero	2011	Invierno	Sábado	0	8,20	12	86	0	1	2	3
8	1	Enero	2011	Invierno	Sábado	0	9,84	14	75	0	1	7	8
9	1	Enero	2011	Invierno	Sábado	0	13,12	17	76	0	8	6	14

*Al comienzo, muchos de los datos numéricos no eran reconocidos como tal. Por lo que fue necesario configurar el dataset de forma que cada valor estuviera asociado con el tipo de dato al que representaba.*

hora	dia	mes	año	estacion	dia_de_la_semana	descanso	temperatura	sensacion_tematica	humedad	velocidad_viento	beneficios_d	beneficios_c	beneficios_cb	beneficios_totales
0	1	1	2018	1	6	0	9.84	14.395	81	0	3	13	16	
1	1	1	2018	1	6	0	9.02	13.635	80	0	8	32	40	
2	1	1	2018	1	6	0	9.02	13.635	80	0	5	27	32	
3	1	1	2018	1	6	0	9.84	14.395	75	0	3	10	13	
4	1	1	2018	1	6	0	9.84	14.395	75	0	0	1	1	
5	1	1	2018	1	6	0	9.84	12.88	75	6.0032	0	1	1	
6	1	1	2018	1	6	0	9.02	13.635	80	0	2	0	2	
7	1	1	2018	1	6	0	8.2	12.88	86	0	1	2	3	
8	1	1	2018	1	6	0	9.84	14.395	75	0	1	7	8	
9	1	1	2018	1	6	0	13.12	17.425	76	0	8	6	14	
10	1	1	2018	1	6	0	15.58	19.695	76	16.9979	12	24	36	

Para el segundo se ha utilizado la herramienta de PowerBI. Para el tratamiento de los datos se han realizado las siguientes transformaciones:

Para realizar el agrupamiento de datos nos hemos basado en la columna “Año”. El software Power Bi va a utilizar ese año para representar y diferenciar todos los demás datos de la base de datos. Tal y como se observará posteriormente, una de las características del cuadro de mandos es que se puede seleccionar el año que se desea representar y a partir del cual se desean obtener gráficos y estadísticas.

Por lo tanto, era fundamental que la columna Año estuviese bien registrada, sin ningún valor incorrecto y que el software reconociese dicha columna como un valor referente a un tipo de dato que hacía referencia a un año, es decir, que no se trataba de un dato como otro cualquiera.

Una de las características de Power BI es que permitía dentro del programa modificar los distintos parámetros, lo que permitió decirle que esa columna se refería a un año y no a un número cualquiera.

	t <sup>2</sup> <sub>3</sub> Hora	t <sup>2</sup> <sub>3</sub> Día	A <sup>B</sup> <sub>C</sub> Mes	Año	A <sup>B</sup> <sub>C</sub> Estación	Día de la semana	t <sup>2</sup> <sub>3</sub> Descanso	1,2 Temperatura	t <sup>2</sup> <sub>3</sub> Sensación térmica	t <sup>2</sup> <sub>3</sub> Humedad	t <sup>2</sup> <sub>3</sub> Velocidad viento
1	0	1	Enero	1,2 Número decimal	Sábado	0	9,84	14	81	82	0
2	1	1	Enero	\$ Número decimal fijo	Sábado	0	9,02	13	80	80	0
3	2	1	Enero	I <sup>3</sup> Número entero	Sábado	0	9,02	15	80	80	0
4	3	1	Enero	% Porcentaje	Sábado	0	9,84	14	75	75	0
5	4	1	Enero	Fecha/Hora	Sábado	0	9,84	14	75	75	0
6	5	1	Enero	Fecha	Sábado	0	9,84	12	75	75	60
7	6	1	Enero	Hora	Sábado	0	9,02	13	80	80	0
8	7	1	Enero	Fecha/Hora/Zona horaria	Sábado	0	8,2	12	86	86	0
9	8	1	Enero	Duración	Sábado	0	9,84	14	75	75	0
10	9	1	Enero	A <sup>B</sup> <sub>C</sub> Texto	Sábado	0	13,12	17	76	76	0
11	10	1	Enero	Verdadero/Falso	Sábado	0	15,58	19	76	76	16
12	11	1	Enero	Binario	Sábado	0	14,76	16	81	81	19
13	12	1	Enero	Usar configuración regional...	Sábado	0	17,22	21	77	77	19
14	13	1	Enero		Sábado	0	18,86	22	72	72	19
15	14	1	Enero		Sábado	0	18,86	22	72	72	19
16	15	1	Enero		Sábado	0	18,04	21	77	77	19

## 5.2 Estudio de los datos

Se ha realizado el estudio de los datos mediante técnicas de minería de datos, empleando la herramienta Power BI y RStudio, para obtener información de los datos usados. Este estudio ha sido realizado para predecir en qué meses del año, estación, días de la semana o en qué condiciones se produce un aumento de la cantidad de bicicletas vendidas.

En el caso de RStudio se ha realizado una regresión lineal para conocer las variables más relevantes a la hora de conocer por qué se han producido ciertos aumentos o descensos, en este caso, las variables más destacadas, resultaron ser las siguientes:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-1.668e+05	5.515e+03	-30.251	< 2e-16	***
hora	7.734e+00	2.070e-01	37.364	< 2e-16	***
mes	7.574e+00	4.207e-01	18.002	< 2e-16	***
año	8.266e+01	2.732e+00	30.258	< 2e-16	***
sensacion_termica	6.172e+00	1.689e-01	36.539	< 2e-16	***
humedad	-2.121e+00	7.858e-02	-26.988	< 2e-16	***
velocidad_del_viento	6.208e-01	1.771e-01	3.506	0.000457	***

En el caso de Power BI, gracias a las distintas herramientas que ofrece el software, se han obtenido gráficos y estadísticas detalladas referentes a los beneficios obtenidos por parte de la empresa.

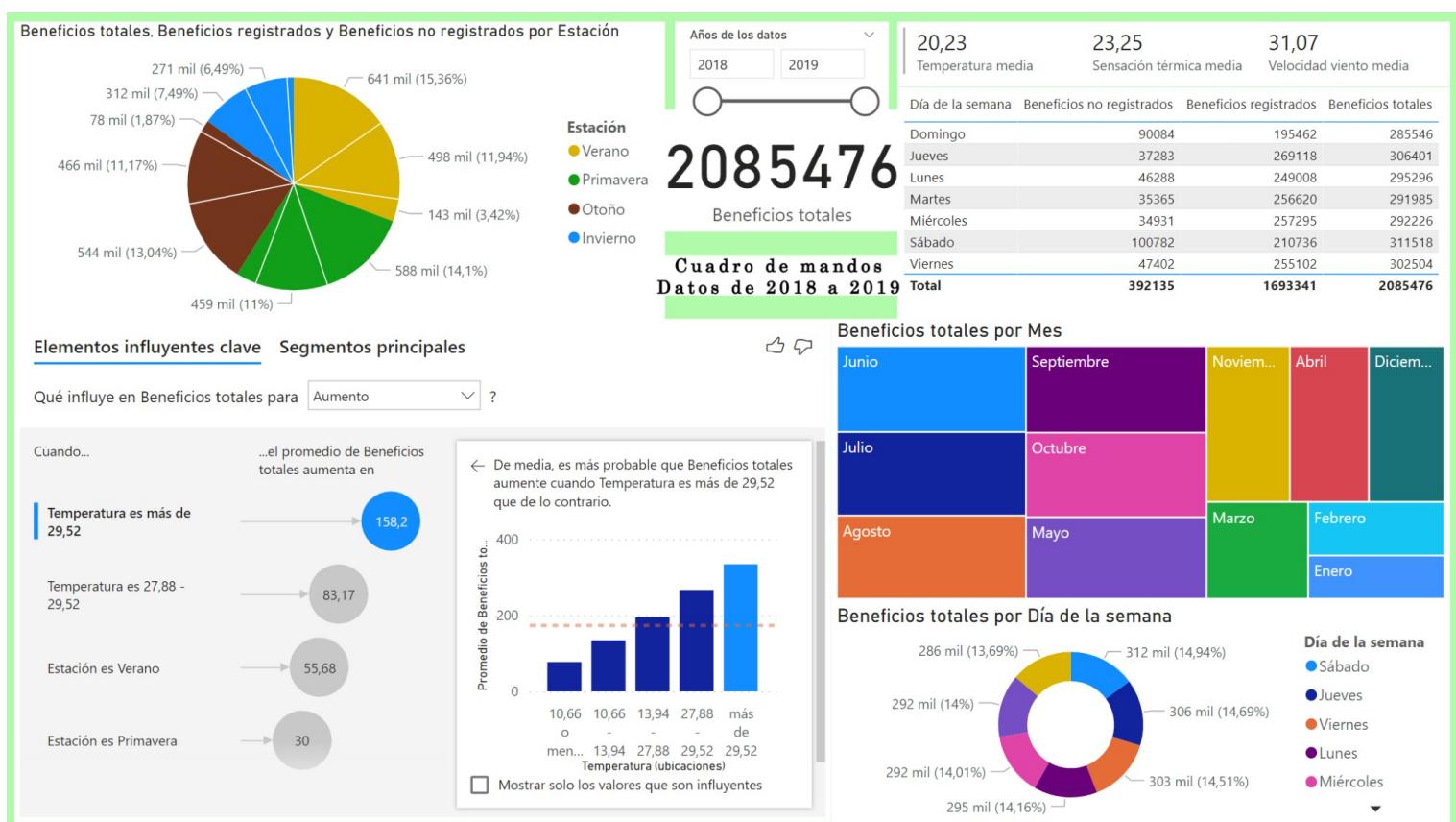
Esos gráficos y distintos apartados se mostrarán en el siguiente apartado (**5.3 Cuadro de mandos**).

## 5.3 Cuadro de mandos

En dicho cuadro de mando, se realiza el estudio de los datos obtenidos en los puntos anteriores de la implementación.

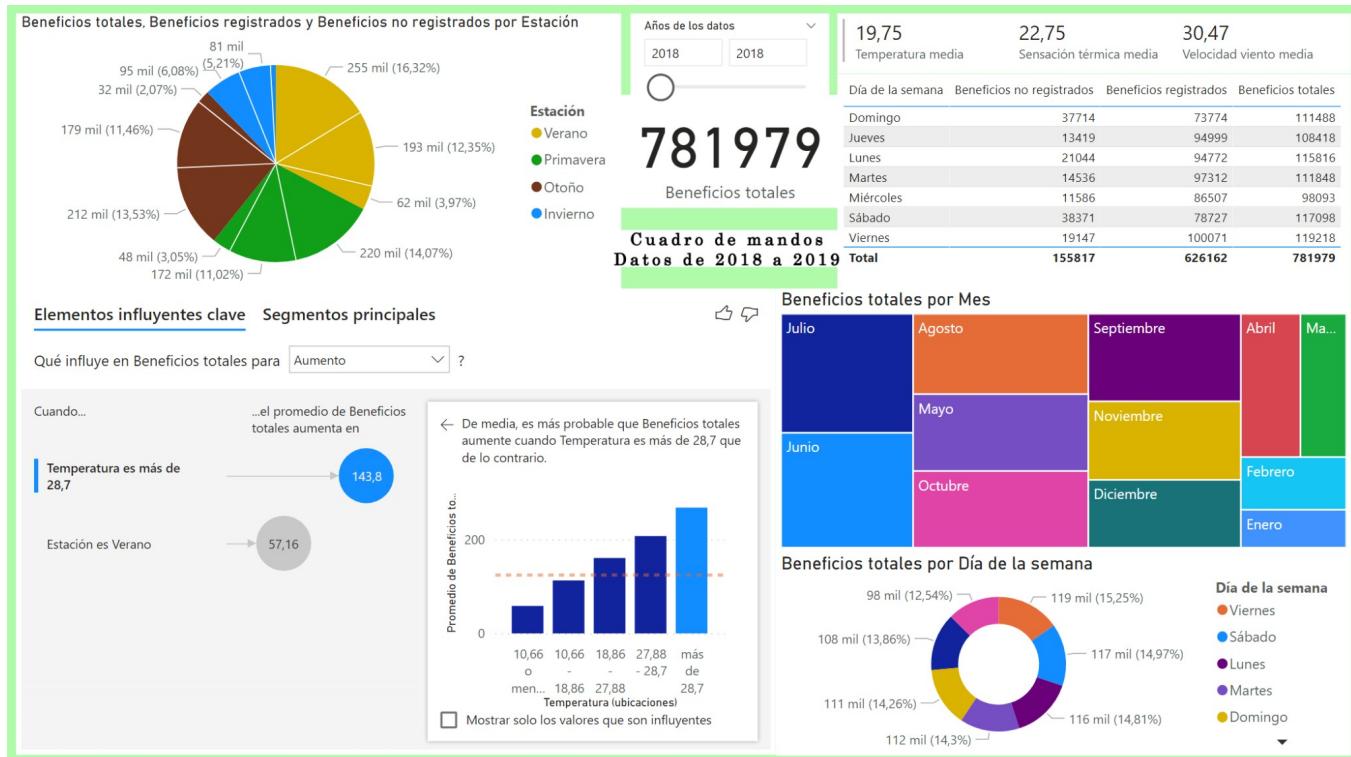
Donde mostraremos lo siguiente:

- Cuadro de mando con gráficos.

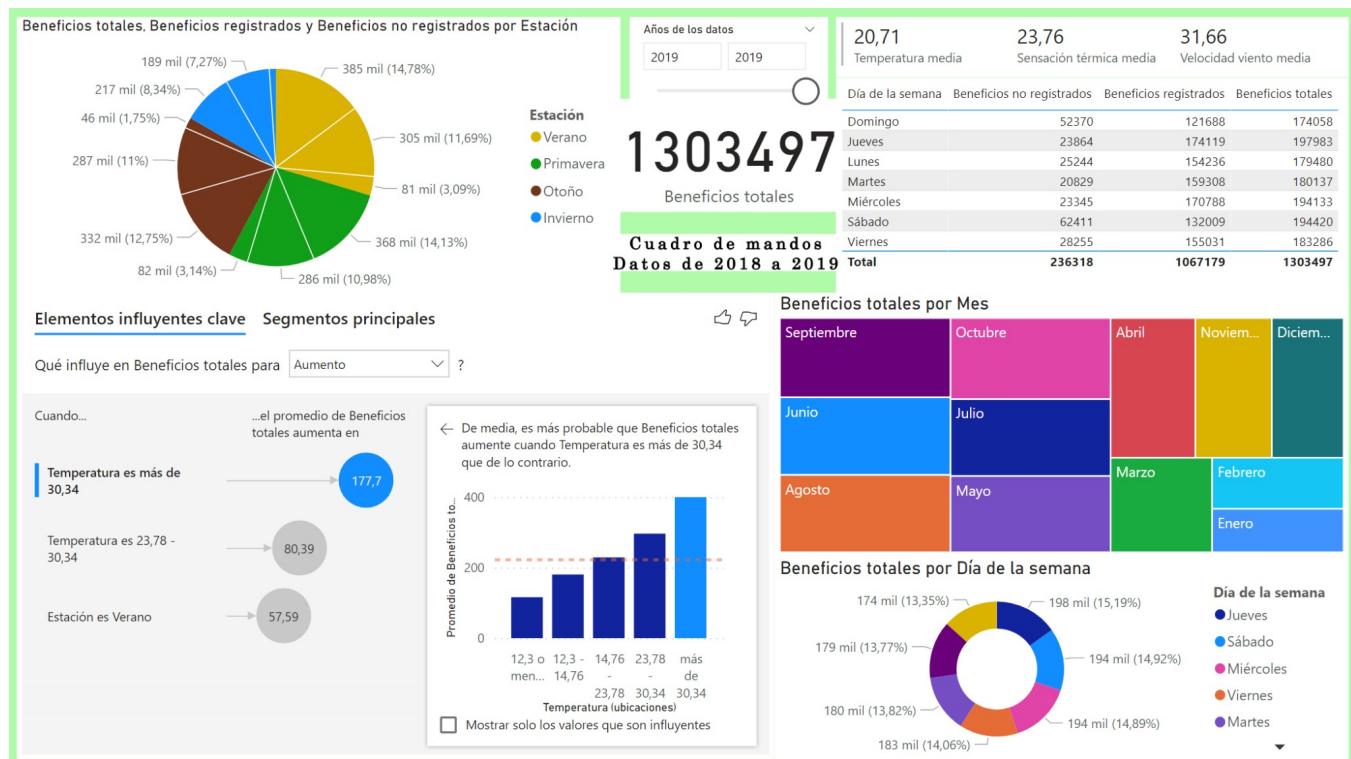


Visión general del cuadro de mandos realizado, se pueden observar una gran cantidad de gráficas que serán explicadas y mostradas en más detalle en las próximas páginas. También se cuenta con estadísticas en la parte superior derecha y en la parte inferior izquierda. Todo esto permite sacar conclusiones que nos ayudan a obtener más beneficios y al mismo tiempo ahorrar recursos en los casos en los que se estén malgastando por falta de interés del público.

- Estadísticas detalladas según la opción elegida.

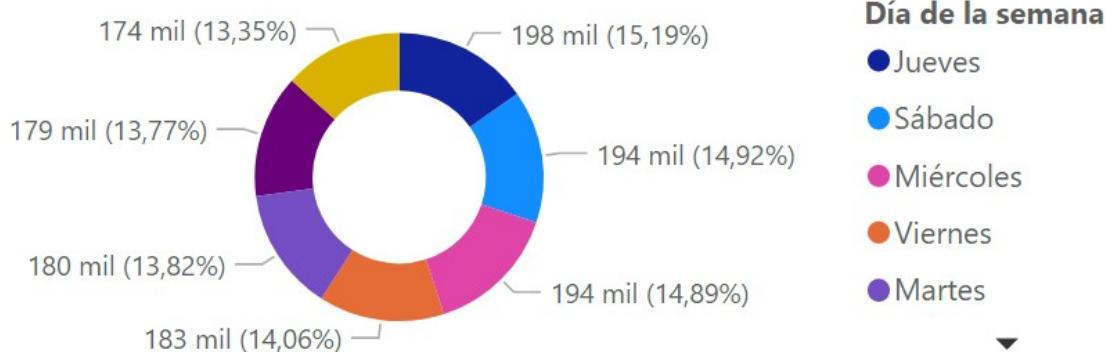


Al seleccionar el año 2018 se ajustarán los datos al cambio

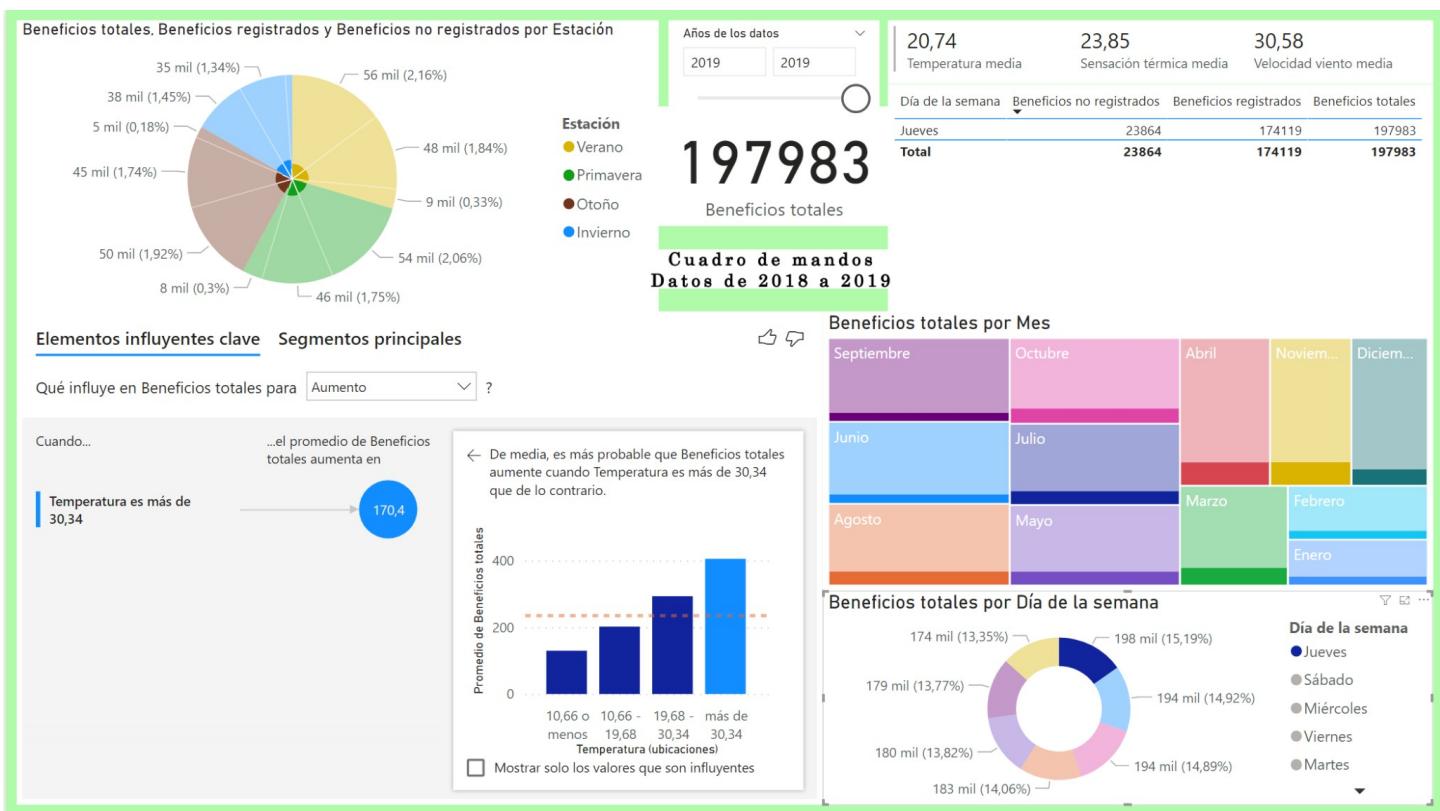


Al seleccionar el año 2019 se ajustarán los datos al cambio

## Beneficios totales por Día de la semana



*Gráfico circular que hace referencia a los días de la semana y los beneficios correspondientes a cada día*

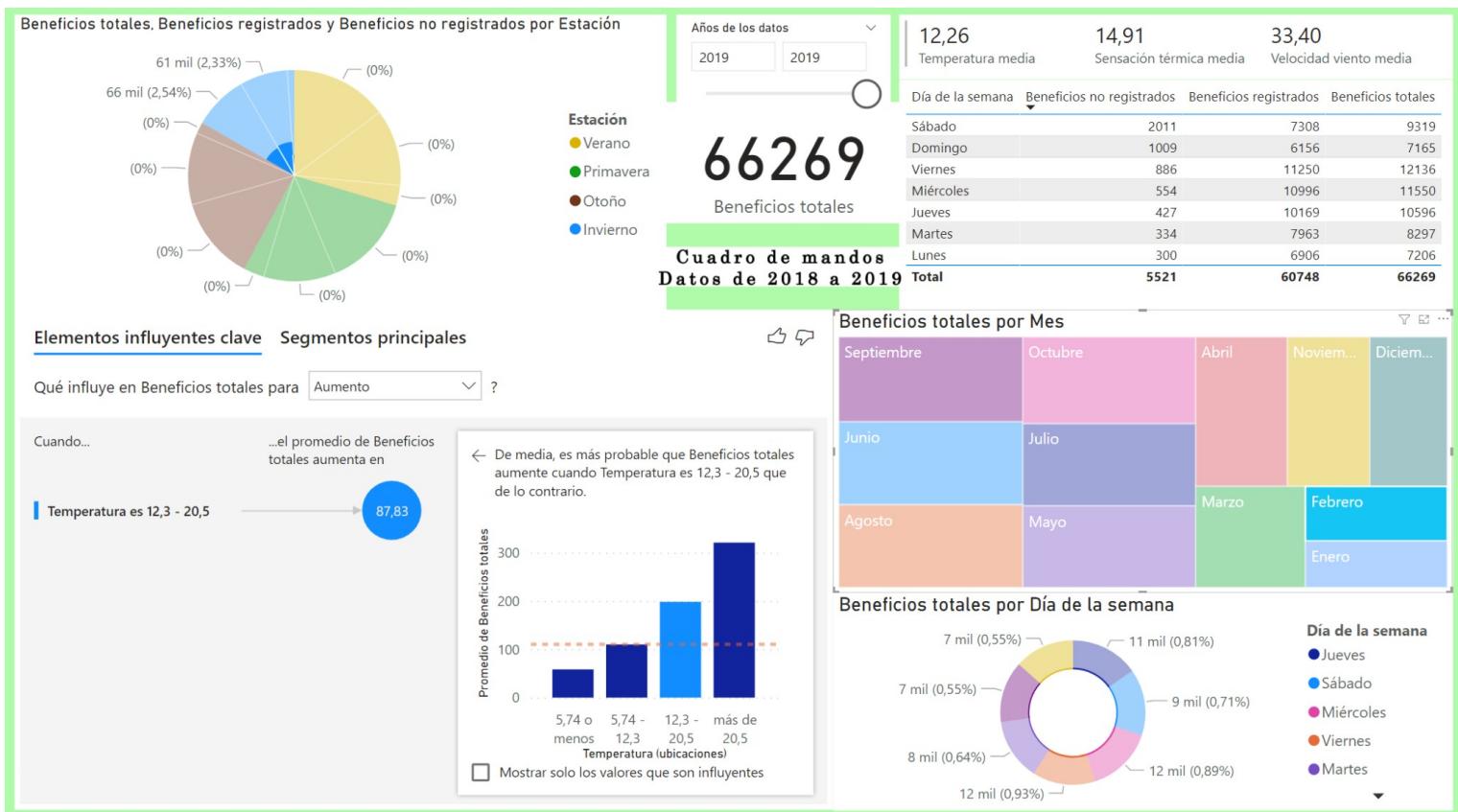


*Una vez seleccionado un día concreto, se ajustan los demás gráficos y estadísticas a dicho día escogido*

## Beneficios totales por Mes



*Gráfico de colores y rectángulos que representa por tamaños la cantidad de beneficios asociados a cada mes del año*



*Datos actualizados al seleccionar en dicho gráfico el mes de Febrero*

Elementos influyentes clave Segmentos principales



Qué influye en Beneficios totales para Aumento ▾ ?

Cuando...

...el promedio de Beneficios totales aumenta en

Temperatura es más de 30,34

177,7

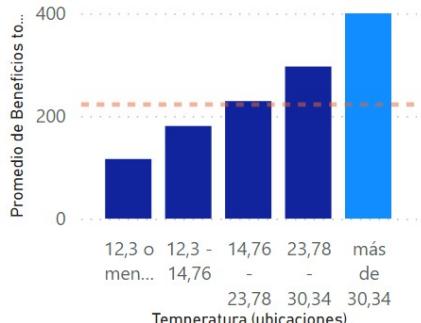
Temperatura es 23,78 - 30,34

80,39

Estación es Verano

57,59

← De media, es más probable que Beneficios totales aumente cuando Temperatura es más de 30,34 que de lo contrario.



*Estadísticas generadas a partir de los datos del dataset que representan las condiciones más relevantes a la hora de saber por qué se ha podido producir un determinado aumento de beneficios*

Elementos influyentes clave Segmentos principales



Qué influye en Beneficios totales para Disminución ▾ ?

Cuando...

...el promedio de Beneficios totales disminuye en

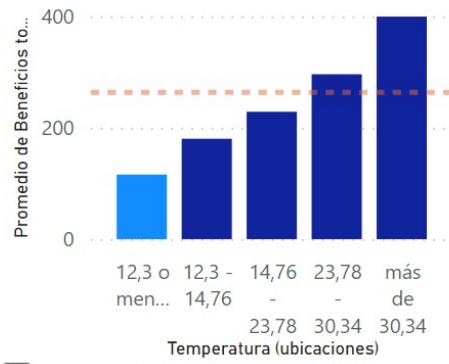
Temperatura es 12,3 o menos

147,7

Temperatura es 12,3 - 14,76

65,95

← De media, es más probable que Beneficios totales disminuya cuando Temperatura es 12,3 o menos que de lo contrario.



*Estadísticas generadas a partir de los datos del dataset que representan las condiciones más relevantes a la hora de saber por qué se ha podido producir una disminución de los beneficios*



## Elementos influyentes clave Segmentos principales

Qué influye en Beneficios totales para Aumento ▾ ?

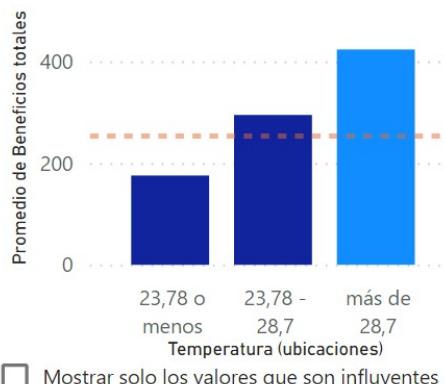
Cuando...

...el promedio de Beneficios totales aumenta en

Temperatura es más de 28,7

170,2

← De media, es más probable que Beneficios totales aumente cuando Temperatura es más de 28,7 que de lo contrario.



## Estadísticas obtenidas en cuanto a los aumentos obtenidos cuando marcamos en el gráfico de meses el mes de Septiembre

### Elementos influyentes clave Segmentos principales



Qué influye en Beneficios totales para Disminución ▾ ?

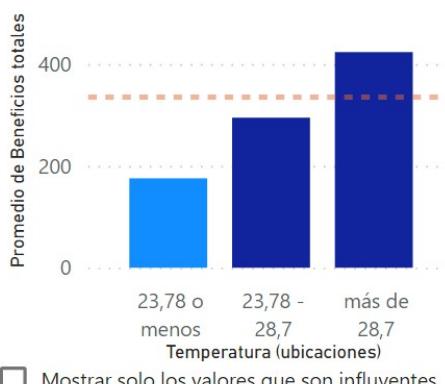
Cuando...

...el promedio de Beneficios totales disminuye en

Temperatura es 23,78 o menos

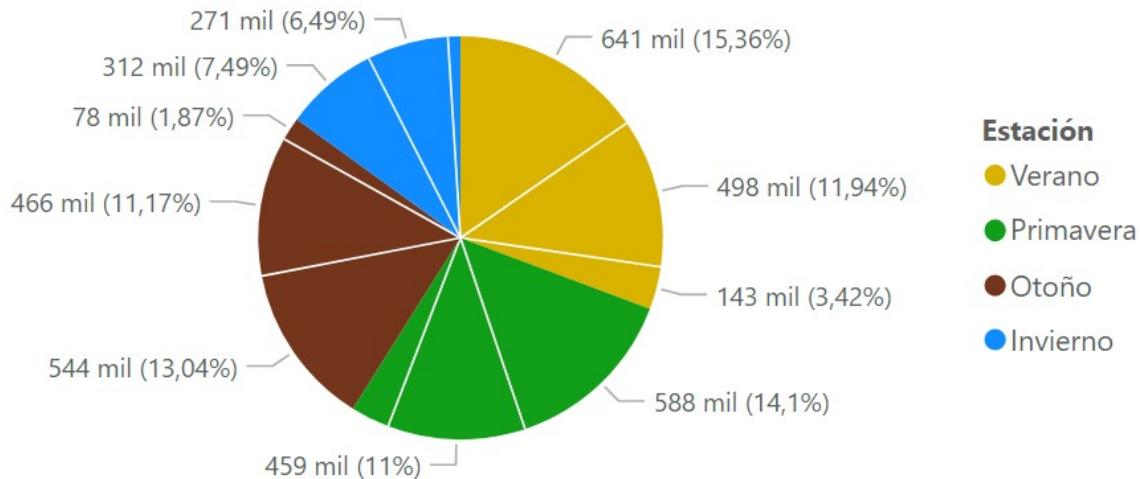
159,6

← De media, es más probable que Beneficios totales disminuya cuando Temperatura es 23,78 o menos que de lo contrario.



## Estadísticas obtenidas en cuanto a los descensos obtenidos cuando marcamos en el gráfico de meses el mes de Septiembre

### Beneficios totales, Beneficios registrados y Beneficios no registrados por Estación



*Gráfico circular que representa la cantidad de beneficios obtenidos según la estación del año*

Beneficios totales, Beneficios registrados y Beneficios no registrados por Estación

Elementos influyentes clave Segmentos principales

Qué influye en Beneficios totales para Aumento

Cuando...

- Temperatura es más de 24,6 → 272,1
- Temperatura es 19,68 - 24,6 → 125,5
- Temperatura es 13,12 - 19,68 → 40,02

De media, es más probable que Beneficios totales aumente cuando Temperatura es más de 24,6 que de lo contrario.

Promedio de Beneficiosto. vs Temperatura (ubicaciones)

Mostrar solo los valores que son influyentes

Años de los datos

2018 2019

312498 Beneficios totales

Cuadro de mandos Datos de 2018 a 2019

Día de la semana	Beneficios no registrados	Beneficios registrados	Beneficios totales
Domingo	9436	29046	38482
Jueves	3868	47580	51448
Lunes	4099	35729	39828
Martes	2859	38263	41122
Miércoles	3550	43024	46574
Sábado	12822	34161	46983
Viernes	4971	43090	48061
Total	41605	270893	312498

Beneficios totales por Mes

Junio	Septiembre	Noviem...	Abril	Diciem...
Julio	Octubre			
Agosto	Mayo	Marzo	Febrero	Enero

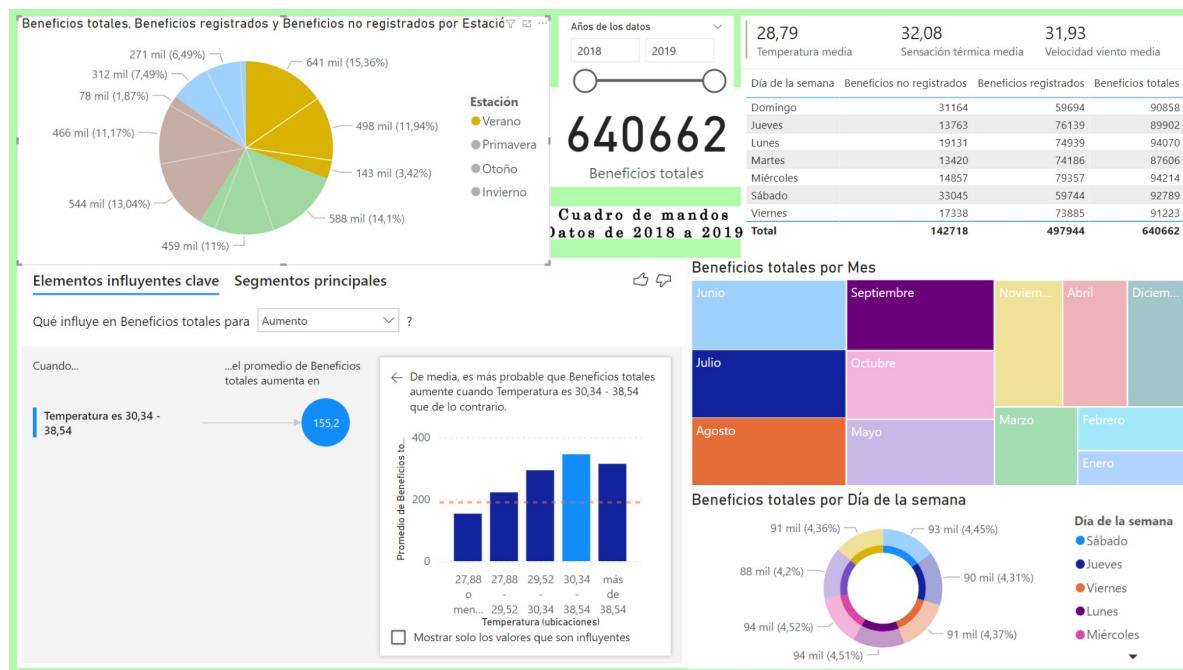
Beneficios totales por Día de la semana

Día de la semana

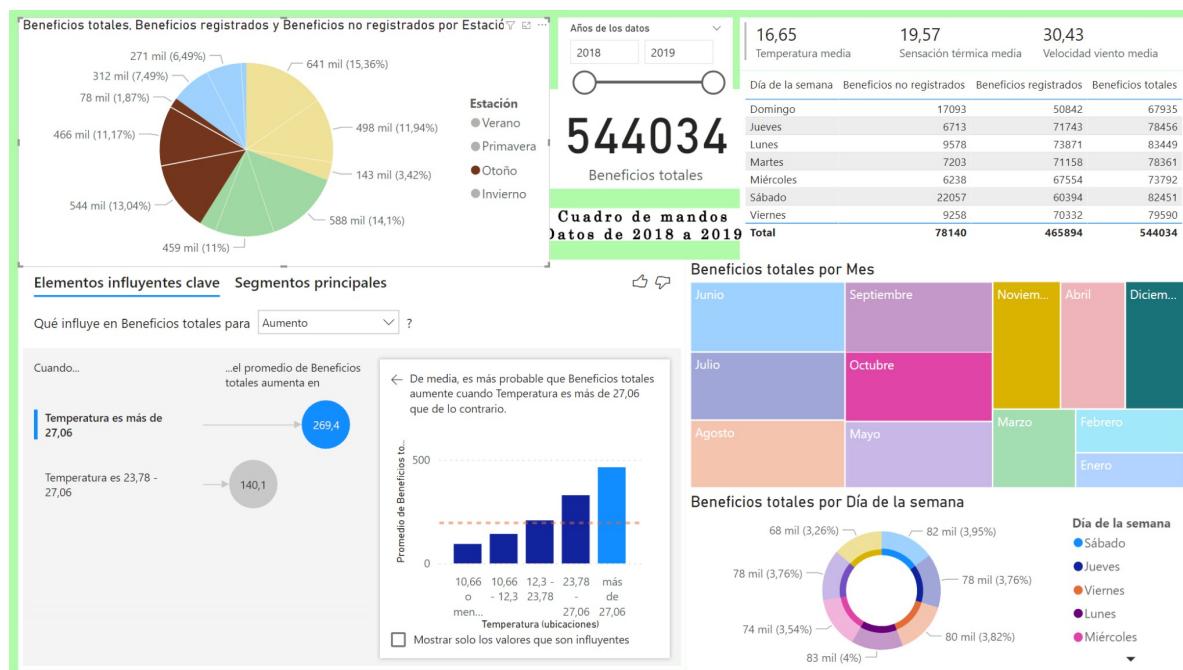
- Sábado
- Jueves
- Viernes
- Lunes
- Miércoles

Así se mostraría el cuadro de mandos una vez seleccionada la estación de invierno

## Inteligencia de negocio - 2020/2021

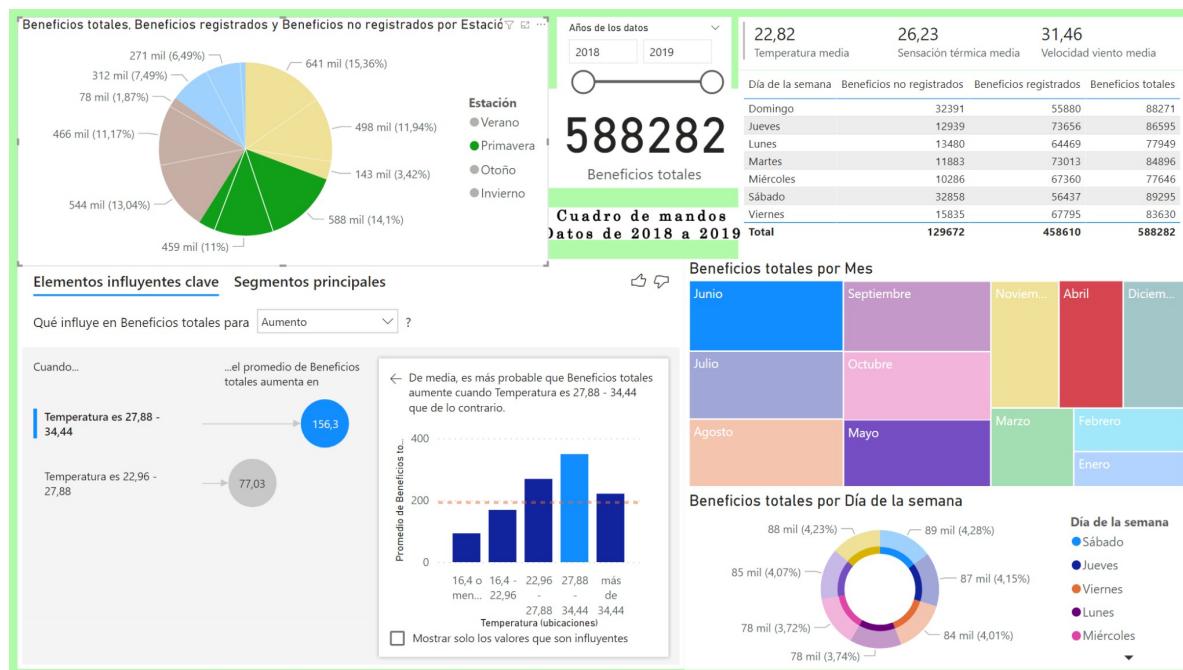


Así se mostraría el cuadro de mandos una vez seleccionada la estación de primavera



Así se mostraría el cuadro de mandos una vez seleccionada la estación de otoño

## Inteligencia de negocio - 2020/2021



Así se mostraría el cuadro de mandos una vez seleccionada la estación de primavera

20,71

Temperatura media

23,76

Sensación térmica media

31,66

Velocidad viento media

Día de la semana	Beneficios no registrados	Beneficios registrados	Beneficios totales
Domingo	52370	121688	174058
Jueves	23864	174119	197983
Lunes	25244	154236	179480
Martes	20829	159308	180137
Miércoles	23345	170788	194133
Sábado	62411	132009	194420
Viernes	28255	155031	183286
<b>Total</b>	<b>236318</b>	<b>1067179</b>	<b>1303497</b>

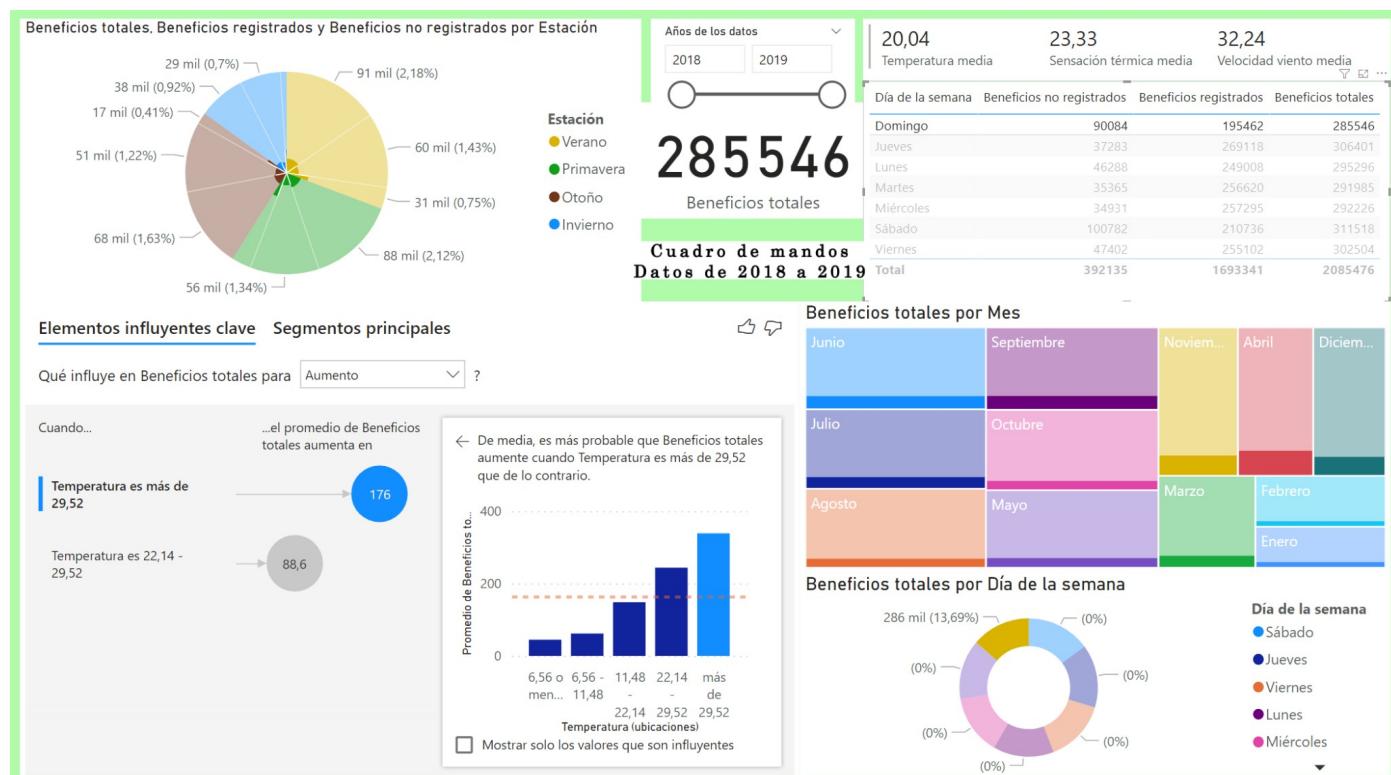
Tabla que representa distintos datos según el día de la semana, además de medidas asociadas también a los días, en este caso, al no estar seleccionado ningún día, las medias que aparecen arriba son las globales

20,31	23,57	33,61	
Temperatura media	Sensación térmica media	Velocidad viento media	

▼ □ ...

Día de la semana	Beneficios no registrados	Beneficios registrados	Beneficios totales
Domingo	52370	121688	174058
Jueves	23864	174119	197983
Lunes	25244	154236	179480
Martes	20829	159308	180137
Miércoles	23345	170788	194133
Sábado	62411	132009	194420
Viernes	28255	155031	183286
<b>Total</b>	<b>236318</b>	<b>1067179</b>	<b>1303497</b>

*Una vez seleccionado un día, en este caso el Domingo, se actualizarán las medias que aparecen arriba de la tabla y se ajustarán los demás apartados del cuadro de mandos de la siguiente forma:*



*De esta forma quedaría el dashboard tras marcar el Domingo en la tabla superior derecha*

## 5.4 Modelo predictivo

Se realizará mediante el software RStudio tal y como se ha especificado anteriormente. La idea es que mediante la introducción del dataset en el software, se pueda realizar un estudio de los datos que permita conocer de una forma lo más precisa posible, los beneficios que se obtendrán en los meses de los próximos años.

A continuación, se mostrarán imágenes que facilitarán la comprensión del proceso realizado:

**PASO 1. Importar el Dataset colocando la ubicación del .csv en el read csv**

```
base_de_datos <- read.csv("C:\\\\Users\\\\Usuario\\\\Desktop\\\\IN 20-21\\\\Trabajo\\\\Dataset\\\\DB_Bicis_RStudio.csv")
```

**PASO 2. Comprobar que se ha importado correctamente la base de datos**

```
muestra <- summary(base_de_datos)
```

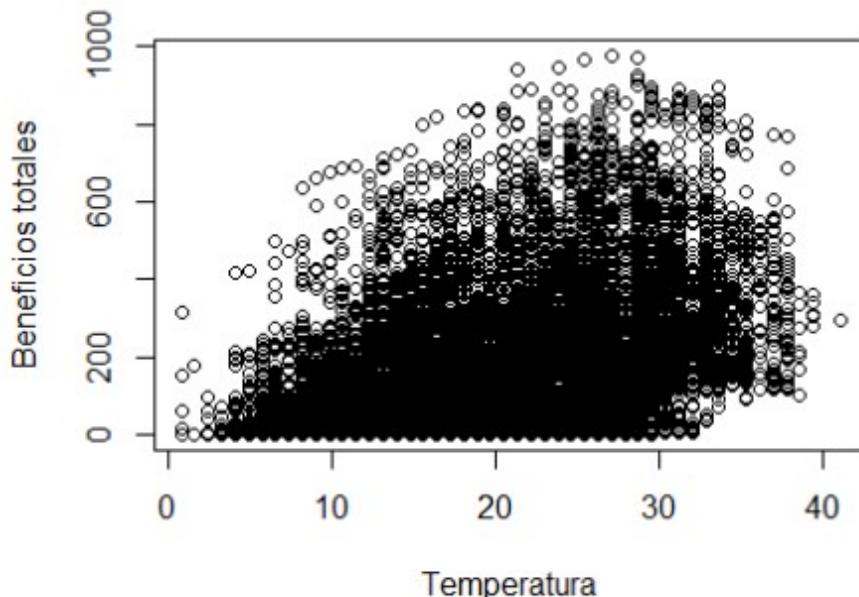
```
muestra
```

```
##      hora          dia          mes          año
##  Min.   : 0.00   Min.   : 1.000   Min.   : 1.000   Min.   :2018
##  1st Qu.: 6.00   1st Qu.: 5.000   1st Qu.: 4.000   1st Qu.:2018
##  Median :12.00   Median :10.000   Median : 7.000   Median :2019
##  Mean   :11.54   Mean   : 9.993   Mean   : 6.521   Mean   :2019
##  3rd Qu.:18.00   3rd Qu.:15.000   3rd Qu.:10.000   3rd Qu.:2019
##  Max.   :23.00   Max.   :19.000   Max.   :12.000   Max.   :2019
##      estacion      dia_de_la_semana      descanso      temperatura
##  Min.   :1.000   Min.   :1.000   Min.   :0.00000   Min.   : 0.82
##  1st Qu.:2.000   1st Qu.:2.000   1st Qu.:0.00000   1st Qu.:13.94
##  Median :3.000   Median :4.000   Median :0.00000   Median :20.50
##  Mean   :2.507   Mean   :4.014   Mean   :0.02857   Mean   :20.23
##  3rd Qu.:4.000   3rd Qu.:6.000   3rd Qu.:0.00000   3rd Qu.:26.24
##  Max.   :4.000   Max.   :7.000   Max.   :1.00000   Max.   :41.00
##      sensacion_termica      humedad      velocidad_del_viento
##  Min.   : 0.76   Min.   :  0.00   Min.   : 0.000
##  1st Qu.:16.66   1st Qu.: 47.00   1st Qu.: 7.002
##  Median :24.24   Median : 62.00   Median :12.998
##  Mean   :23.66   Mean   : 61.89   Mean   :12.799
##  3rd Qu.:31.06   3rd Qu.: 77.00   3rd Qu.:16.998
##  Max.   :45.45   Max.   :100.00   Max.   :56.997
##      beneficios_de_no_registrados      beneficios_de_registrados
##  Min.   :  0.00           Min.   :  0.0           Min.   : 1.0
##  1st Qu.:  4.00           1st Qu.: 36.0           1st Qu.: 42.0
##  Median :17.00           Median :118.0           Median :145.0
##  Mean   :36.02           Mean   :155.6           Mean   :191.6
##  3rd Qu.:49.00           3rd Qu.:222.0           3rd Qu.:284.0
##  Max.   :367.00           Max.   :886.0           Max.   :977.0
```

**PASO 3. Mostrar una gráfica que permita entender la relación entre la temperatura y los beneficios totales**

```
plot(base_de_datos$temperatura, base_de_datos$beneficios_totales, main =  
"Relación entre la temperatura y los beneficios totales", xlab =  
"Temperatura", ylab = "Beneficios totales")
```

**Relación entre la temperatura y los beneficios total**



#### PASO 4. Generar la regresión lineal

```

regresion <- lm(beneficios_totales ~ hora + dia + mes + año + estacion +
dia_de_la_semana + descanso + temperatura + sensacion_termica + humedad +
velocidad_del_viento, data = base_de_datos)
summary(regresion)

##
## Call:
## lm(formula = beneficios_totales ~ hora + dia + mes + año + estacion +
##     dia_de_la_semana + descanso + temperatura + sensacion_termica +
##     humedad + velocidad_del_viento, data = base_de_datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -305.52  -93.64  -27.70   61.85  649.10 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -1.667e+05  5.516e+03 -30.217 < 2e-16 ***
## hora          7.735e+00  2.070e-01  37.368 < 2e-16 ***
## dia           3.844e-01  2.482e-01   1.549  0.12150  
## mes           9.996e+00  1.682e+00   5.943  2.89e-09 ***
## año            8.258e+01  2.732e+00   30.225 < 2e-16 ***
## estacion      -7.774e+00  5.177e+00  -1.502  0.13324  
## dia_de_la_semana 4.393e-01  6.918e-01   0.635  0.52545  
## descanso      -4.864e+00  8.365e+00  -0.582  0.56089  
## temperatura   1.582e+00  1.038e+00   1.524  0.12752  
## sensacion_termica 4.748e+00  9.552e-01   4.971  6.76e-07 ***
## humedad        -2.115e+00  7.884e-02  -26.827 < 2e-16 ***
## velocidad_del_viento 5.582e-01  1.809e-01   3.086  0.00203 ** 
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
##
## Residual standard error: 141.7 on 10874 degrees of freedom
## Multiple R-squared:  0.3891, Adjusted R-squared:  0.3885 
## F-statistic: 629.6 on 11 and 10874 DF,  p-value: < 2.2e-16

```

### PASO 5. Ajustar la regresión lineal

```
regresion <- lm(beneficios_totales ~ hora + mes + año + sensacion_termica
+ humedad + velocidad_del_viento, data = base_de_datos)
summary(regresion)

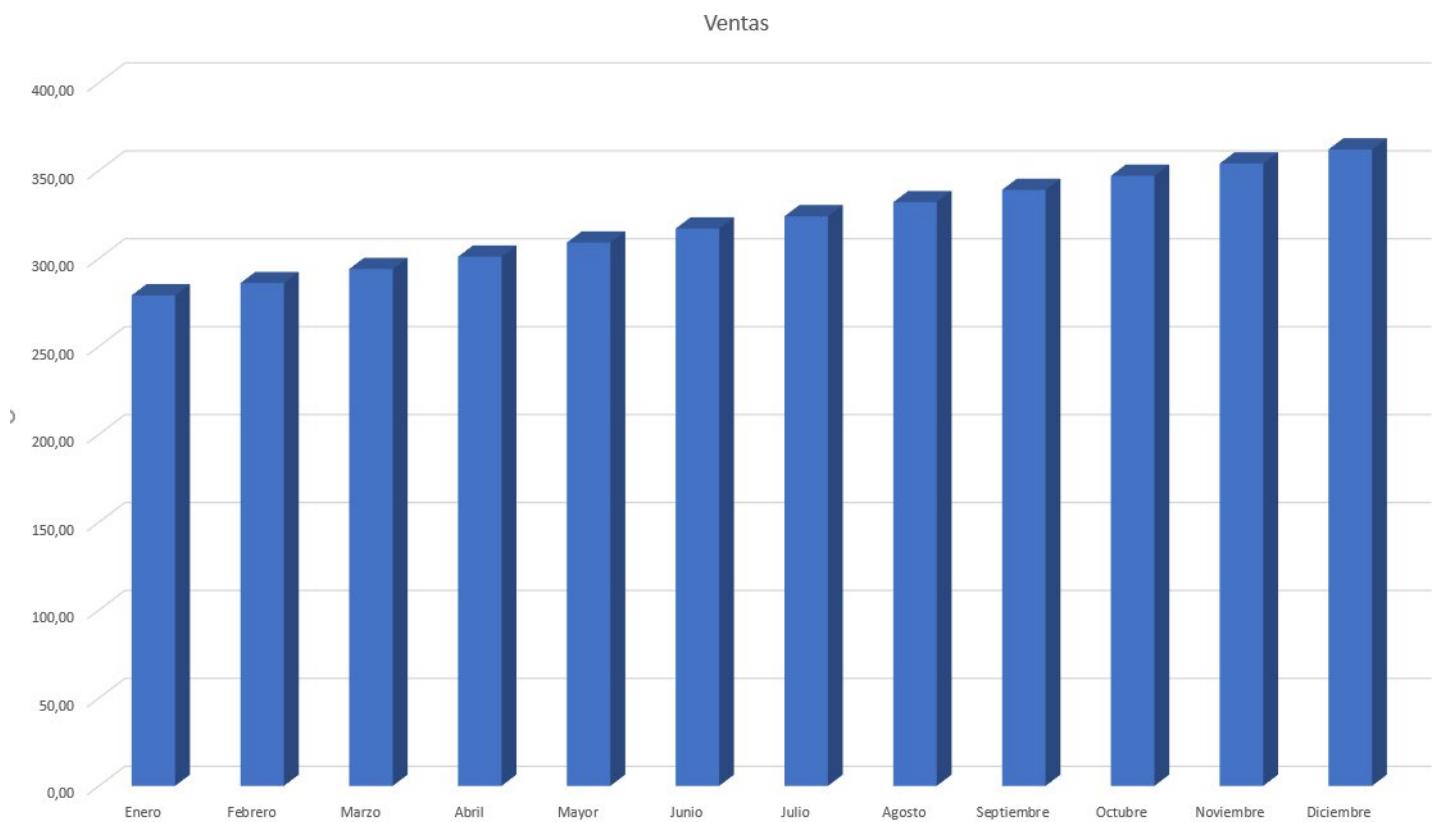
##
## Call:
## lm(formula = beneficios_totales ~ hora + mes + año + sensacion_termica
+
##     humedad + velocidad_del_viento, data = base_de_datos)
##
## Residuals:
##     Min      1Q  Median      3Q     Max 
## -308.60  -93.85  -28.34   61.05  648.09 
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)           -1.668e+05  5.515e+03 -30.251 < 2e-16 ***
## hora                  7.734e+00  2.070e-01   37.364 < 2e-16 ***
## mes                   7.574e+00  4.207e-01   18.002 < 2e-16 ***
## año                   8.266e+01  2.732e+00   30.258 < 2e-16 ***
## sensacion_termica    6.172e+00  1.689e-01   36.539 < 2e-16 ***
## humedad                -2.121e+00  7.858e-02  -26.988 < 2e-16 ***
## velocidad_del_viento  6.208e-01  1.771e-01    3.506 0.000457 *** 
## ---                
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 141.7 on 10879 degrees of freedom
## Multiple R-squared:  0.3886, Adjusted R-squared:  0.3883 
## F-statistic: 1153 on 6 and 10879 DF,  p-value: < 2.2e-16
```

### PASO 6. Realizar modelo predictivo

```
datos <- data.frame(hora = 12, mes = 1:12, año = 2020, sensacion_termica
= 24, humedad = 62, velocidad_del_viento = 13)
predict(regresion, datos)

##      1       2       3       4       5       6       7
## 279.1478 286.7215 294.2952 301.8690 309.4427 317.0164 324.5901
## 332.1638
##      9      10      11      12
## 339.7375 347.3112 354.8849 362.4587
```

A partir de dicha predicción, se puede realizar un gráfico que represente el crecimiento de las ventas el siguiente año:



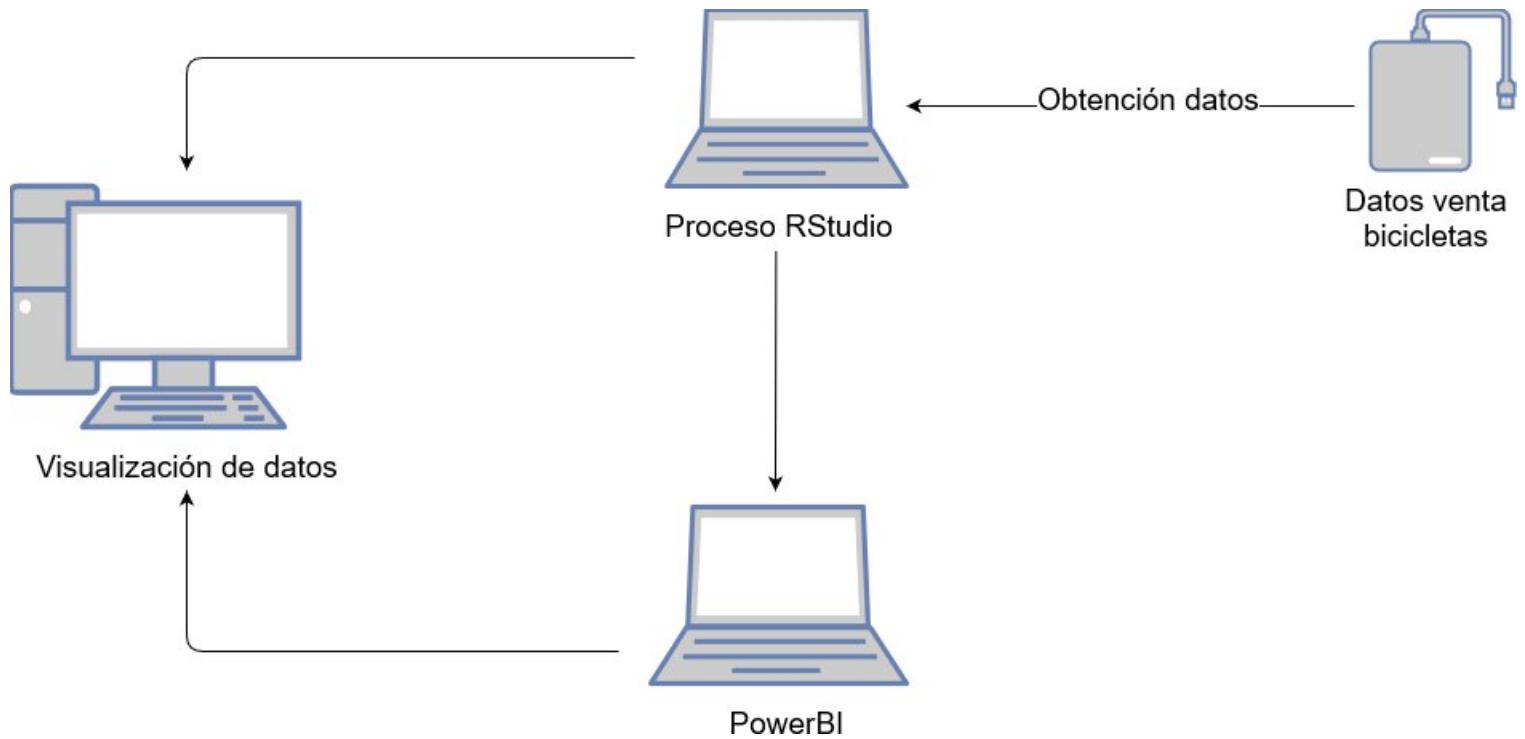
*Predicción de ventas para el año 2020 según el modelo predictivo*

El modelo predictivo también permite saber la estimación de ventas de otros años, como por ejemplo del 2021:

1	2	3	4	5	6	7	8	9	10	11	12
361.8102	369.3839	376.9576	384.5313	392.1050	399.6787	407.2525	414.8262	422.3999	429.9736	437.5473	445.1210

## 6. Despliegue

El proceso de despliegue comienza cuando nos dirigimos a la página web donde descargaremos los datos de las bicicletas, para el tratamiento de los datos en RStudio. Tras realizar los ajustes, estos se emplean en el software PowerBI. Con este software obtenemos el cuadro de mandos y podemos realizar las pruebas correspondientes.



Los casos de usos durante el proceso son los siguientes:

ID	Caso de uso
CU-01	Visualizar cuadro de mandos
CU-02	Mostrar beneficios por mes
CU-03	Mostrar beneficios por estación
CU-04	Mostrar beneficios por días de la semana
CU-05	Seleccionar año
CU-06	Mostrar datos a trás de la tabla
CU-07	Observar las medias obtenidas a partir de los demás datos
CU-08	Estudiar las estadísticas obtenidas mediante el apartado estadístico generado dentro del cuadro de mandos
CU-09	Investigar seleccionando distintos elementos del dashboard para observar cómo se va modificando todo según lo seleccionado
CU-10	Comprobar que todo funciona correctamente

## 7. Conclusiones

La inteligencia de negocio nos ofrece en este caso la posibilidad de mejorar nuestro nivel competitivo en el ámbito de las ventas de bicicletas. Con este estudio de inteligencia de negocio podemos mejorar los resultados y nuestra gestión de recursos, lo cual redundará en una mejora de la viabilidad de nuestro negocio.

Gracias al cuadro de mando elaborado en este proyecto facilitamos la toma de decisiones a la vez que proporcionamos una forma sencilla y rápida de visualizar datos de gran importancia, lo cual resulta vital para el correcto desarrollo de cualquier actividad económica.

Destacamos el uso de Power BI, que ha resultado una herramienta de fácil uso, destacando por su agilidad en términos de implementación. Ha resultado sorprendentemente sencilla la representación de los datos en el cuadro de mando, esto se ve reflejado en la claridad con la que se representa la información, con una división clara de las secciones y distintos tipos de gráficos representados.

Finalmente creemos que este tipo de trabajos de inteligencia de negocio son imprescindibles para cualquier empresa que busque una mejora. Mediante el empleo de la inteligencia de negocio podemos agilizar la obtención de información valiosa y realmente importante, lo cual es vital de cara a poder competir en los negocios. A la finalización de este proyecto hemos obtenido una información muy útil que ayudará a impulsar la mejora en la venta de bicicletas.

## Bibliografía

<https://powerbi.microsoft.com/es-es/> - Software Power BI

<https://www.rstudio.com/> - Software RStudio

<https://www.gestionexcelenciaoperacional.com/>

- Web dataset

<https://www.microsoft.com/es-es/microsoft-365/excel>

- Web de Excel

Todos las herramientas utilizadas son gratuitas o bien se ha contado con la licencia de la universidad Pablo de Olavide para su obtención.