



UNIVERSIDAD DE BURGOS  
ESCUELA POLITÉCNICA SUPERIOR  
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería  
Informática**

**Sistema de Información sobre  
Matriculación**



Presentado por Mario de la Parte Izquierdo  
en Universidad de Burgos — 2 de julio  
de 2019

Tutor: Carlos Pardo Aguilar





UNIVERSIDAD DE BURGOS  
ESCUELA POLITÉCNICA SUPERIOR  
Grado en Ingeniería Informática



D. Carlos Pardo Aguilar, profesor del departamento de Ingeniería Civil, área de Lenguajes y Sistemas Informáticos.

Expone:

Que el alumno D. Mario de la Parte Izquierdo, con DNI 71305494C, ha realizado el Trabajo final de Grado en Ingeniería Informática titulado *Sistema de Información sobre Matriculación*.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 2 de julio de 2019

Vº. Bº. del Tutor:

D. Carlos Pardo Aguilar



## Resumen

En la actualidad, existe una gran cantidad de información o datos, los cuales componen una parte muy importante en las grandes empresas y organizaciones de todo el mundo. Cada día se genera multitud de nueva información y es indispensable almacenarla para posteriormente poder interpretarla adecuadamente.

El desarrollo de este proyecto viene motivado por realizar un Sistema de Información, o lo que es lo mismo, un almacén electrónico sobre la matriculación de alumnos en la Universidad de Burgos.

Con la creación del **Sistema de Información sobre Matriculación** se pretende crear una aplicación que sea capaz de procesar, almacenar, administrar, organizar y visualizar correctamente información relevante a la matriculación.

De esta manera, la Universidad de Burgos contará con un sistema cuya información se podrá utilizar para la toma de decisiones.

## Descriptores

Sistema de Información, Matrícula Universitaria, Aplicación Sigma

### **Abstract**

Nowadays, there is a large amount of data and information, which is a very great value for big companies and organizations all around the world. Every single day a multitude of new information is created, and it is indispensable to keep it saved, in order to be used properly later on.

The research and development of this project is to create a new Information System to keep an electronic store, based on Burgos University students's enrolment data.

With this new **Enrolment Information System**, I pretend to create an application that is capable of processing, storing, managing, organising and correctly visualising relevant enrolment information.

In this way, Burgos University will have an Enrolment System whose information can be used to find out things or to make decisions.

### **Keywords**

Information System, University Enrollment, Sigma Application

---

# Índice general

---

<b>Índice general</b>	<b>III</b>
<b>Índice de figuras</b>	<b>V</b>
<b>Introducción</b>	<b>1</b>
1.1. Estructura de la memoria . . . . .	2
<b>Objetivos del proyecto</b>	<b>3</b>
2.1. Objetivo general . . . . .	3
2.2. Objetivos técnicos . . . . .	3
<b>Conceptos teóricos</b>	<b>5</b>
3.1. Sistema de Información . . . . .	5
3.2. Gráficos Representados . . . . .	9
<b>Técnicas y herramientas</b>	<b>13</b>
4.1. Metodologías . . . . .	13
4.2. Lenguaje de Programación . . . . .	14
4.3. Entorno de Desarrollo . . . . .	14
4.4. Control de Versiones . . . . .	15
4.5. Documentación . . . . .	15
4.6. Otras Herramientas . . . . .	16
<b>Aspectos relevantes del desarrollo del proyecto</b>	<b>19</b>
5.1. Inicio del proyecto . . . . .	19
5.2. Metodologías . . . . .	20
5.3. Toma de decisiones . . . . .	20

5.4. Librerías para el tratamiento y manipulación de datos . . . . .	24
5.5. Interfaz de usuario del proyecto . . . . .	26
5.6. Problemas encontrados . . . . .	26
<b>Trabajos relacionados</b>	<b>29</b>
<b>Conclusiones y Líneas de trabajo futuras</b>	<b>31</b>
7.1. Conclusiones . . . . .	31
7.2. Líneas de trabajo futuras . . . . .	32
<b>Bibliografía</b>	<b>33</b>



---

## Índice de figuras

---

3.1. Componentes de un Sistema de Información . . . . .	6
3.2. Información de Caja y Bigotes . . . . .	10
3.3. Diagrama de Caja y Bigotes . . . . .	10
3.4. Gráfico apilado horizontalmente de asignaturas por curso . . . . .	11
5.5. Logotipo del Sistema de Información sobre Matriculación . . . . .	20
5.6. Datos del fichero original (.xls) . . . . .	22
5.7. Datos del fichero parseado con columna añadida (.csv) . . . . .	23
5.8. Error al intentar abrir con Excel los ficheros originales . . . . .	26



---

# Introducción

---

En la actualidad, existe una gran cantidad de información o datos, los cuales componen una parte muy importante en las grandes empresas y organizaciones de todo el mundo. Cada día se genera multitud de nueva información y es indispensable almacenarla para posteriormente poder interpretarla correctamente. Es imprescindible por lo tanto, saber extraer e identificar información relevante a partir de ficheros o documentos poco legibles o difíciles de entender a priori.

En este punto es cuando toma especial interés la creación de un Sistema de Información, o lo que es lo mismo, un almacén electrónico. En dichos almacenes se protege y mantiene una gran cantidad de datos e información, de manera fiable, segura y fácil de administrar.

Además de estas funciones de almacenamiento y administración, un Sistema de Información también permite organizar, entender y utilizar los datos para la toma de decisiones. Para esta tarea, es necesario contar con cierta capacidad de análisis, ya que hay que extraer información concreta, destacada y relevante; para posteriormente poder visualizarla con ayuda de elementos visuales como gráficos.

En la realización de este proyecto se propone la creación de un Sistema de Información, para procesar, almacenar y representar visualmente la información sobre la matriculación de alumnos en la Universidad de Burgos.

De esta forma, en el proyecto se podrán diferenciar varias funcionalidades:

- Creación de la Base de Datos (BBDD).
- Preprocesamiento de los ficheros Excel (.xls) descargados de *Sigma*.

- Carga de datos en la Base de Datos (BBDD) a partir de los ficheros (.csv) generados.
- Visualización de diferentes tipos de gráficos en función de los datos de la BBDD y lo que el usuario seleccione.

## 1.1. Estructura de la memoria

La memoria se estructura de la siguiente manera:

- **Introducción:** se describe brevemente el contexto y el proyecto realizado. Posteriormente se realiza una sección donde se expone la estructura de la memoria.
- **Objetivos del proyecto:** se exponen los objetivos del proyecto, divididos en objetivo general y objetivos técnicos.
- **Conceptos teóricos:** se exponen los conceptos teóricos y básicos para comprender tanto el proyecto como el desarrollo del mismo.
- **Técnicas y herramientas:** se explican las metodologías y herramientas utilizadas durante el desarrollo del proyecto.
- **Aspectos relevantes del desarrollo del proyecto:** se exponen los aspectos más importantes que han surgido en el desarrollo del proyecto, así como la toma de decisiones, cambios y problemas encontrados.
- **Trabajos relacionados:** se exponen aplicaciones, proyectos y empresas que ofrecen soluciones en un ámbito o campo similar al estudiado.
- **Conclusiones y líneas de trabajo futuras:** se explican las conclusiones finales que se obtienen después de la realización del proyecto, así como futuras mejoras del mismo.
- **Bibliografía:** conjunto de referencias bibliográficas utilizadas en la memoria.

---

# Objetivos del proyecto

---

A continuación se definen los objetivos del proyecto realizado, divididos en dos apartados:

## 2.1. Objetivo general

- Desarrollar una aplicación para analizar datos relacionados con la matriculación de alumnos en la Universidad de Burgos (UBU).

## 2.2. Objetivos técnicos

Se han propuesto cinco objetivos técnicos a realizar:

- Extraer los datos o información relevante de ficheros Excel (.xls), utilizando librerías concretas de Python.
- Crear la Base de Datos(BBDD) para almacenar la información anteriormente extraída, con una estructura de tablas, campos y claves adecuada.
- Identificar y crear gráficos o estadísticos que resulten útiles para visualizar y comparar información.
- Desarrollar una aplicación en Python que unifique todo lo anterior, así como realizar una interfaz gráfica agradable para el usuario.
- Conseguir que la aplicación sea fiable, usable y robusta.



---

# Conceptos teóricos

---

En este apartado se van a explicar aquellos conceptos teóricos básicos que son necesarios para comprender el proyecto.

## 3.1. Sistema de Información

### Definición

Para comenzar, hay que explicar que no existe una definición de consenso en la propia definición de Sistema de Información. De hecho, existen multitud de definiciones diferentes sobre cómo se define un Sistema de Información.

Los autores Laudon y Laudon definen un Sistema de Información como un conjunto de módulos relacionados entre sí que son capaces de obtener (o reutilizar), procesar, almacenar y distribuir cierta información para que sirva de apoyo para la toma de decisiones [2]. A parte de suministrar apoyo en decisiones importantes, también pueden ayudar a detectar problemas o carencias difíciles de ver sin la ayuda de estos sistemas.

### Componentes de un Sistema de Información

Aunque existen numerosas definiciones y no existe una definición general o global, la mayoría de Sistemas de Información pueden representarse a través del diagrama de la figura 3.1.

Se pueden apreciar cinco elementos principales. En primer lugar los elementos de entrada, que en nuestro proyecto serían los ficheros (.xls) originales descargados de *Sigma*.

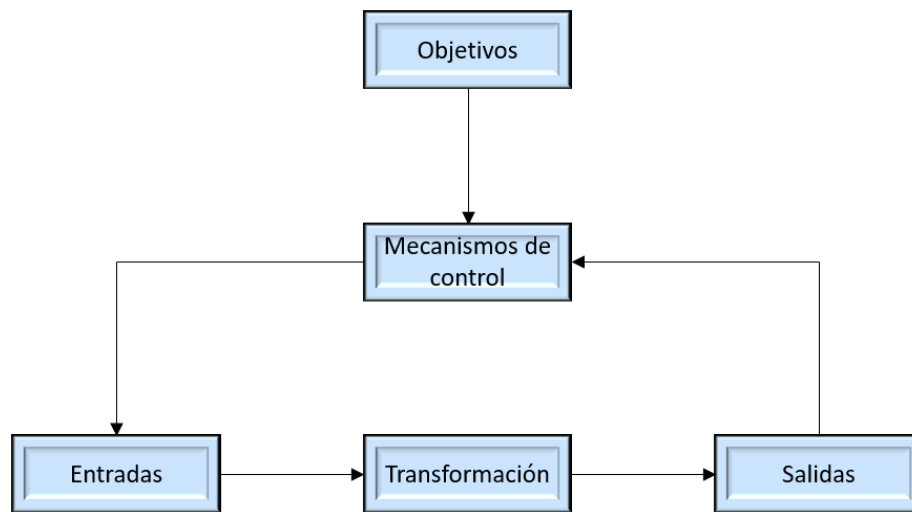


Figura 3.1: Componentes de un Sistema de Información

A continuación estaría un elemento de modificación o transformación, que en nuestro caso sería el preprocesado de los ficheros originales anteriores en ficheros (.csv) reordenados, modificados y sin ningún tipo de error. También se podría incluir la carga de datos a la Base de Datos creada con anterioridad.

Seguidamente estaría el sistema de salida, donde se visualizan los resultados obtenidos. En este proyecto, el sistema de salida serían los diferentes tipos de gráficos que se pueden obtener a partir de la información que seleccione el usuario y los datos existentes o disponibles en la BBDD.

Además de estas 3 secciones, se aprecian otras dos secciones más. Una de ellas es el mecanismo de control, que es el proceso encargado de lograr los objetivos, que sería el quinto y último elemento. En nuestro proyecto se podrían identificar numerosos mecanismos de control, como por ejemplo que los ficheros que se puedan seleccionar en los botones de *Preprocesar* y *Cargar Archivos* sean únicamente (.xls) y (.csv) respectivamente. Otros mecanismos de control serían la no introducción de datos repetidos en la BBDD o la selección de opciones de datos que realmente se encuentran en la BBDD, entre otros.

En cuanto a los objetivos de nuestro sistema de información, hay que destacar que se definen en el apartado anterior denominado *Objetivos del proyecto*.



## Características de un Sistema de Información

Algunas de las características más comunes que comparten casi todos los sistemas de información, son las siguientes:

- **Relevancia.** El sistema debe ser capaz de generar información importante y necesaria para la empresa u organización. Adicionalmente la información generada debe ser fiable[3].
- **Apoyo en la toma de decisiones.** Estos sistemas suelen ser repetitivos y capaces de soportar decisiones no estructuradas que no suelen repetirse.
- **Flujo independiente.** Hay que destacar que en este tipo de sistemas, existe un flujo de procesamiento de datos (tanto de manera interna como externa), pero también existe un flujo independiente de los propios sistemas de información. Estos flujos independientes suelen estar integrados a sistemas ya existentes (en el caso del proyecto, la aplicación *Sigma*).
- **Integración.** Esta característica se refiere al nexo de unión que debe existir entre el propio sistema de información y la empresa u organización. Con esta característica es más sencillo coordinar los diferentes departamentos o divisiones, así como agilizar la toma de decisiones[1].
- **Control.** Esta característica no es común u obligatoria en todos los tipos de sistemas de información, pero algunos pueden contener funciones de control interno. La finalidad de este control es asegurar que la información que se genera es fiable y que los datos que se obtienen son protegidos y controlados adecuadamente.

## Tipos de Sistemas de Información

A continuación se van a exponer los diferentes tipos de sistemas de información, desde un punto de vista empresarial.

- **Sistemas de procesamiento de transacciones.** Más conocido como **TPS**<sup>1</sup> por sus siglas en inglés. Estos tipos de sistemas se categorizan como básicos, ya que son útiles a nivel operacional dentro de la organización. Es decir, se encargan de realizar transacciones diarias necesarias para el correcto funcionamiento de la empresa[5].

---

<sup>1</sup>Transaction Processing System

- **Sistemas de control de procesos de negocio.** Más conocido como **BPM**<sup>2</sup> por sus siglas en inglés. Estos sistemas se encargan de monitorizar y controlar procesos industriales (con la ayuda de sensores y otros dispositivos) y realizar ajustes en tiempo real que controlan los mismos.
- **Sistemas de colaboración empresarial.** Más conocido como **ERP**<sup>3</sup> por sus siglas en inglés. Este es uno de los tipos de sistemas más utilizados, ya que prestan ayuda (como por ejemplo controlar el flujo de información) a los directivos de una empresa.
- **Sistemas de Información Ejecutiva.** Más conocido como **EIS**<sup>4</sup> por sus siglas en inglés. Estos sistemas son herramientas orientadas a usuarios de nivel gerencial, ya que generalmente permiten monitorizar estados de variables de un área determinada de la empresa a partir de información interna y externa de la misma[11].
- **Sistemas de Información de Gestión o Gerencial.** Más conocido como **MIS**<sup>5</sup> por sus siglas en inglés. Este tipo de sistemas se encargan de recopilar y procesar información de diferentes fuentes para ayudar en el proceso de toma de decisiones dentro de una empresa u organización. Estos sistemas a su vez, están orientados a solucionar problemas empresariales en general.

De los diferentes tipos de sistemas de información que se han mencionado, nuestro proyecto se asemejaría al último (*Sistema de Información de Gerencial (MIS)*). Como característica curiosa, comentar que las siglas de nuestro proyecto, así como su logotipo, son las siglas de *Management Information System* al revés, como se aprecia en la figura 5.5.

## Ventajas de un Sistema de Información

Las ventajas más destacadas de un sistema de información son las siguientes:

- **Detección de problemas.** Gracias a un sistema de información se pueden detectar problemas para su posterior resolución.

---

<sup>2</sup>Business Process Management

<sup>3</sup>Enterprise Resource Planning

<sup>4</sup>Executive Information System

<sup>5</sup>Management Information System

- **Disminución de costes.** Se disminuye el costo de mano de obra y se optimizan tiempos y tareas[9].
- **Administración de activos.** Los activos pueden ser tangibles o intangibles. Sean del tipo que sean, los sistemas de información actuales se han convertido en una herramienta crucial para la administración de activos.
- **Ventaja competitiva.** Los sistemas de información se han convertido en una de las principales vías para obtener ventaja competitiva en el ámbito empresarial.

## 3.2. Gráficos Representados

### Diagrama de Caja y Bigotes

Los diagramas de cajas y bigotes (o diagramas de cuartiles) son un tipo de gráficas que representan una gran cantidad de información de manera muy visual y esquematizada. A su vez, se pueden apreciar características estadísticas relevantes como la simetría y la dispersión de un conjunto de datos.

Toda esta valiosa información se representa mediante unas pequeñas cajas muy intuitivas, como se aprecia en la figura 3.2.

Se diferencian cinco partes fundamentales:

- **Tres Cuartiles (Q1, Q2 y Q3).** Hay que destacar que el segundo cuartil(Q2) coincide con la mediana y representa la relación entre el primer y tercer cuartil. El primer cuartil identifica el valor por debajo del cual queda un 25 % de todos los datos de la muestra ordenada. Del mismo modo, el tercer cuartil es el valor por debajo del cual quedan el 75 % de los datos de la muestra.
- **Máximo.** Representa el valor máximo de los datos.
- **Mínimo.** Representa el valor mínimo de los datos.

A continuación, en la figura 3.3 se puede apreciar un diagrama de caja y bigotes que se genera mediante nuestra aplicación, donde se pueden observar los conceptos teóricos anteriormente mencionados.

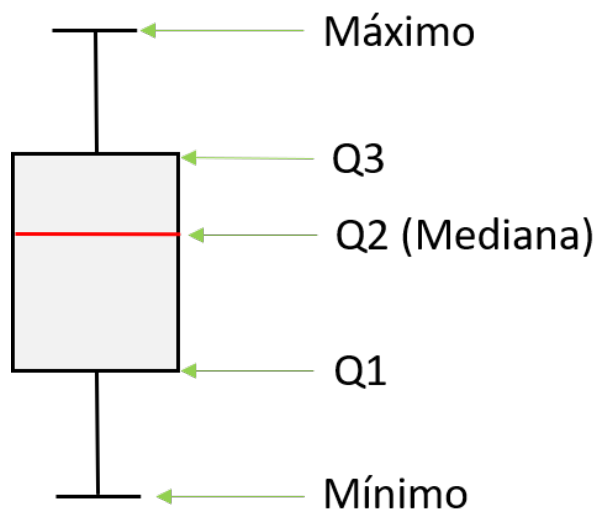


Figura 3.2: Información de Caja y Bigotes

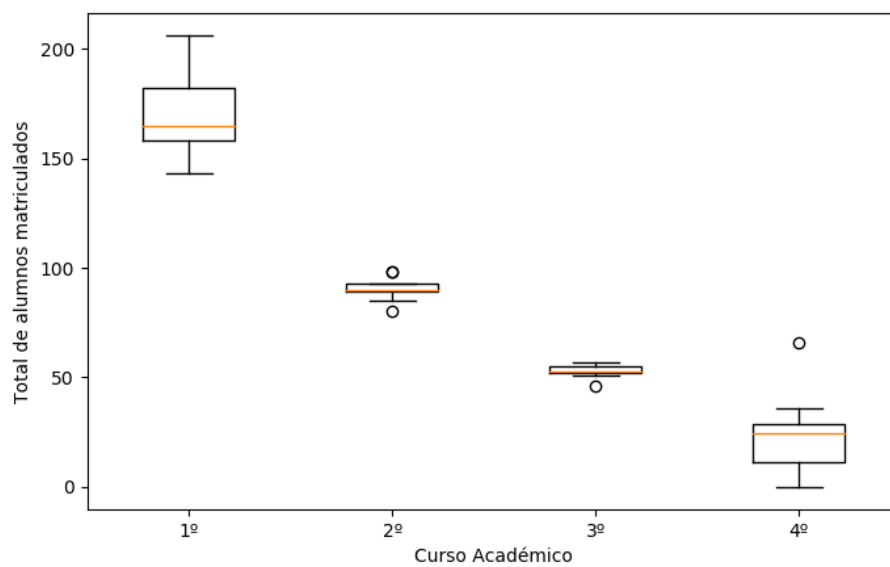


Figura 3.3: Diagrama de Caja y Bigotes

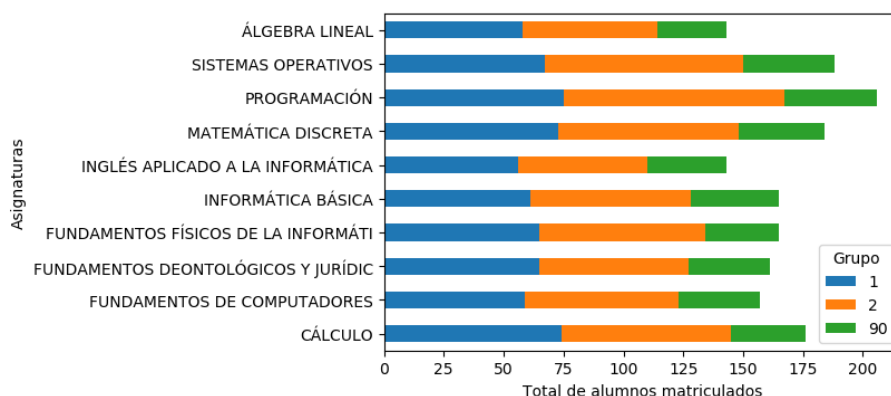


Figura 3.4: Gráfico apilado horizontalmente de asignaturas por curso

## Gráfico de Barras

Los gráficos de barras son un tipo de representación gráfica realizada en un eje cartesiano de las frecuencias de una variable cualitativa o discreta. [4]

En uno de los ejes se posicionan las diferentes categorías de la variable cualitativa (en el caso de nuestro proyecto, las asignaturas de un determinado curso) y en el otro eje se posiciona la frecuencia de cada asignatura (en nuestro proyecto sería la cantidad de alumnos matriculados).

La orientación del gráfico puede ser horizontal o vertical. Hay que destacar que en nuestro proyecto se utiliza un **gráfico apilado y horizontal**. Es decir, las diferentes asignaturas del curso seleccionado se sitúan en el eje vertical y las barras apiladas de los diferentes grupos (teóricos o prácticos) aumentan horizontalmente.

Este tipo de distribución horizontal se suele utilizar cuando existen numerosas categorías o el nombre de las mismas es demasiado extenso, como es el caso (generalmente hay diez asignaturas por curso).

Hay que destacar que en los gráficos apilados, las diferentes barras se dividen en segmentos de diversos colores (para poder apreciarlos y diferenciarlos del resto) y cada uno de estos colores, representa una serie. En nuestro caso, estas series se corresponden con los diferentes grupos (teóricos o prácticos) que componen las asignaturas. A continuación, en la figura 3.4 se pueden apreciar todos los conceptos en un gráfico obtenido por nuestra aplicación.



---

# Técnicas y herramientas

---

En este apartado se van a exponer las técnicas metodológicas y herramientas de desarrollo que se han utilizado para la realización del proyecto. Se detallarán las razones principales por los que se ha usado esa herramienta y no otra.

## 4.1. Metodologías

### Scrum

Para realizar la planificación correcta del proyecto, se ha utilizado *Scrum*, que es una metodología ágil de desarrollo.

- Se ha utilizado una estrategia orientada a un desarrollo incremental y basada en *sprints*.
- La duración media de cada *sprint* era aproximadamente de una semana.
- Al inicio de cada *sprint* se definían las tareas o *issues* a realizar, las cuales tenían que ser realizadas en un cierto intervalo de tiempo.
- Cada *sprint* se planificaba cuando se finalizaban las tareas o *issues* del anterior *sprint*.
- Al final de cada *sprint* se revisan todas las tareas realizadas, así como ver si se han logrado los objetivos fijados y solucionado los problemas encontrados.

Con la utilización de esta metodología se ha logrado evitar la realización de una planificación y ejecución completa desde el inicio del proyecto.

## 4.2. Lenguaje de Programación

### Python

El lenguaje de programación utilizado ha sido Python<sup>6</sup> en la versión 3.7.1.

Las razones por las que se ha decidido utilizar Python son las siguientes:

- Es uno de los lenguajes de programación más sencillos de aprender ya que su sintaxis es muy entendible.
- Es un lenguaje gratuito, multiplataforma y de código abierto.
- Gracias a las dos anteriores razones, se ha convertido en un lenguaje tan popular y utilizado, que ha dado lugar a que se desarrollen multitud de librerías, módulos y programas de software libre. Gran parte de estas librerías destacan en el ámbito de manejo de ficheros, tratamiento y visualización de datos.
- Del mismo modo, al ser un lenguaje utilizado por tantas personas, hace que existan numerosos foros, blogs y páginas en las que apoyarse cuando surgen dudas o se necesita ayuda.

## 4.3. Entorno de Desarrollo

### Jupyter Notebook

Como entorno de desarrollo principal se ha utilizado Jupyter Notebook<sup>7</sup> en la versión 5.7.4.

Se trata de una aplicación web de código abierto que permite tanto el desarrollo como la ejecución del código. Esta aplicación se puede lanzar directamente desde un navegador (sin instalar nada) o se puede instalar con *Anaconda Navigator*. Para el desarrollo del proyecto se utilizó la segunda opción.

Otra de las grandes ventajas de esta aplicación es la agilidad en el desarrollo, ya que al tratarse de una aplicación cuya ejecución es en vivo, se pueden realizar pruebas de manera rápida e intuitiva.

---

<sup>6</sup>[www.python.org](http://www.python.org)

<sup>7</sup>[www.jupyter.org](http://www.jupyter.org)



## SQLite

SQLite<sup>8</sup> se trata de un sistema de gestión de bases de datos (BDD) relacionales de pequeño tamaño. Una de las características de este sistema de gestión es que no necesita un servidor para poder utilizarse, ya que los datos se almacenan en un único fichero en el sistema host. Hay que destacar que Python incluye soporte para SQLite desde la versión 2.5 incorporado en la Biblioteca Estándar como el módulo *sqlite3*, que es el módulo que se ha utilizado en el desarrollo del proyecto [12].

## 4.4. Control de Versiones

### GitHub

GitHub<sup>9</sup> se trata de una plataforma cuya función principal es la de hospedar repositorios y permitir el desarrollo colaborativo. Es una plataforma de las más usadas y por esta razón es la que se ha utilizado a lo largo del grado y en particular en la realización de este proyecto. Hay que destacar que gracias a formar parte de la Universidad de Burgos y ser estudiante, se ha obtenido la versión *PRO* (licencia de estudiantes). Aun así hay que destacar que se trata de una herramienta gratuita. Por último comentar que se ha utilizado tanto *GitHub Desktop* (aplicación de escritorio) como la plataforma web.

## 4.5. Documentación

### Texmaker

Para la realización de la documentación con L<sup>A</sup>T<sub>E</sub>X<sup>10</sup> se ha utilizado el editor Texmaker<sup>11</sup>. Se trata de un editor gratuito, el cual contiene las herramientas y características necesarias para desarrollar y editar documentos con L<sup>A</sup>T<sub>E</sub>X.

Hay que señalar que también incluye corrección ortográfica, auto-completado, plegado de código y un visor incorporado en pdf con soporte de *synctex* y un modo de visualización continua [10].

---

<sup>8</sup>[www.sqlite.org](http://www.sqlite.org)

<sup>9</sup>[www.github.com](http://www.github.com)

<sup>10</sup>[www.latex-project.org](http://www.latex-project.org)

<sup>11</sup>[www.xmlmath.net/texmaker](http://www.xmlmath.net/texmaker)

## 4.6. Otras Herramientas

En este apartado se van a explicar otras herramientas destacadas que se han utilizado a lo largo del proyecto.

### DB Browser

DB Browser<sup>12</sup> es una herramienta gratuita y de código abierto cuyo principal objetivo es la administración de Bases de Datos que utilizan SQLite como motor de las mismas. Esta herramienta cuenta con numerosas funcionalidades, entre las que se encuentran la creación de BBDD, tablas, índices, entradas, importar y exportar archivos, entre otras.

Hay que destacar que en este proyecto se ha utilizado la aplicación de escritorio para tareas de visualización de datos de la BBDD y comprobación de los mismos.

### Sublime Text 3 y Notepad++

Tanto Sublime Text 3<sup>13</sup> como Notepad++<sup>14</sup> son editores de código que pueden ser utilizados como entornos de desarrollo, ya que pueden interpretar numerosos lenguajes de programación.

En un primer lugar ambos editores se utilizaron para el desarrollo del proyecto, pero finalmente el uso de estas herramientas fue la de edición y visualización de ficheros. Hay que destacar funcionalidades como la mostrar caracteres ocultos de *Notepad++*, la visualización del texto en función de la sintaxis o lenguaje de programación que se elija de *Sublime Text 3* y las herramientas de búsqueda de ambos editores.

### Nitro Pro

Nitro Pro<sup>15</sup> es una herramienta gráfica cuya funcionalidad reside en la creación y edición de ficheros (.pdf). Hay que destacar que al tratarse de una herramienta de pago, se ha utilizado la versión de prueba de la misma, ya que contaba con las funcionalidades necesarias para el proyecto.

Esta herramienta se ha utilizado para la edición de ficheros (.pdf) como *about.pdf* y *ayuda.pdf* de la interfaz gráfica.

---

<sup>12</sup>[www.sqlitebrowser.org](http://www.sqlitebrowser.org)

<sup>13</sup>[www.sublimetext.com](http://www.sublimetext.com)

<sup>14</sup>[www.notepad-plus-plus.org](http://www.notepad-plus-plus.org)

<sup>15</sup>[www.gonitro.com](http://www.gonitro.com)

## Excel

Excel<sup>16</sup> pertenece a la categoría de programas conocidos como hojas de cálculo [6].

De hecho es una de las herramientas más utilizadas y potentes en el análisis de datos, ya que cuenta con una gran cantidad de funcionalidades relacionadas con este área.

En la realización del proyecto, se ha utilizado principalmente para modificar (pruebas) y visualizar tanto los ficheros originales (.xls), como los generados (.csv).

## Photoshop

Photoshop<sup>17</sup> es un editor de imágenes dedicado principalmente para el retoque de fotografías y creación de gráficos.

Esta herramienta se ha utilizado para pequeñas tareas de diseño gráfico, como la realización de los botones de tipos de gráfico, el logotipo de la aplicación, la edición de las imágenes de la documentación...etc.

---

<sup>16</sup>[www.support.office.com/es-es/excel](http://www.support.office.com/es-es/excel)

<sup>17</sup>[www.photoshop.com](http://www.photoshop.com)



---

# Aspectos relevantes del desarrollo del proyecto

---

En este apartado se van a recoger los aspectos más importantes que han surgido en el desarrollo del proyecto. Se incluirán la toma de decisiones, los posibles cambios, la aparición de problemas y las soluciones establecidas.

## 5.1. Inicio del proyecto

Al principio se propuso la idea de la creación de un Sistema de Información (*Data Ware-House* o almacén de datos), para almacenar información relevante para la matriculación de alumnos en la Universidad de Burgos (UBU) y su futura utilización para la toma de decisiones.

Por esta razón, la idea transmitida por el tutor fue la creación de un sistema para poder almacenar y visualizar cualquier grado o máster de la UBU.

El tutor comentó que los archivos que se iban a utilizar de partida, eran un tipo de ficheros descargado desde una aplicación denominada *Sigma* y tenían un error de formato.

Con este conjunto de ideas y proposiciones, se realizaron los *Objetivos del proyecto* y se enviaron al tutor para la supervisión de los mismos.

Tras la aprobación o visto bueno, se empieza a realizar el proyecto, con la creación del repositorio en *GitHub* el día 26 de febrero de 2019.



Figura 5.5: Logotipo del Sistema de Información sobre Matriculación

## 5.2. Metodologías

A lo largo del desarrollo del proyecto se ha usado la *metodología Scrum*. Se trata de una metodología ágil basada en *sprints*, en este caso, de desarrollo incremental con revisiones semanales.

Por lo tanto, la duración estimada de cada *sprint* es de una semana, si bien ha habido varios *sprints* que han tenido una duración superior. Al finalizar cada *sprint*, se planificaba el siguiente, creando sus *issues* o tareas a realizar en dicho *sprint*. Cuando estas tareas se realizaban, se cambiaba el estado del *issue* correspondiente a *Closed* o cerrado. Para ver todos los *sprints* realizados, así como cada *issue* detallado, se puede acceder al repositorio de *GitHub* desde el siguiente enlace<sup>18</sup>.

## 5.3. Toma de decisiones

### Interfaz de usuario

Para la realización de la interfaz de usuario se ha utilizado la librería *tkinter* de *Python*, ya que cuenta con numerosas funcionalidades que se explicarán en la siguiente sección.

El objetivo principal era la creación de una interfaz de usuario sencilla y funcional, y se ha conseguido gracias a la librería anteriormente mencionada.

---

<sup>18</sup>[www.github.com/mdi0007/Sistema-Informacion-sobre-Matriculacion](https://www.github.com/mdi0007/Sistema-Informacion-sobre-Matriculacion)

También hay que exponer que se optó por la realización de una aplicación de escritorio, ya que era la mejor opción para el uso que se espera dar a dicha aplicación, no teniendo en principio mucho sentido la creación de una aplicación web.

## Creación de una BBDD

Hay que destacar que una vez preprocesados los ficheros originales de *Sigma*, los gráficos que se fueron creando de prueba, se realizaban a partir de los ficheros generados (.csv). Más adelante se consideró oportuno la creación de una base de datos específica para el proyecto, con una estructura de tablas y restricciones concretas. Para esta labor se utilizó el sistema de gestión *SQLite*, ya que cumplía con creces nuestros requisitos (características principales de *SQLite* explicadas en el apartado anterior de *Técnicas y Herramientas*). De esta forma se diseñó una base de datos con 3 tablas principales (explicado en detalle en el apartado *Especificación de diseño* de los anexos), y las gráficas se empezaron a generar mediante consultas *SQL*, aportando simplicidad y eficiencia.

## Modificación de archivos (.csv) generados

Relacionado con lo anterior, se decidió modificar los archivos que nuestra aplicación generaba a partir de los ficheros *Sigma* originales.

La decisión de esta modificación, fue la de realizar la carga de datos en la BBDD de una forma más eficiente y en bloque, evitando tener que leer fila por fila el documento y comprobando si los datos de dicha fila se debían añadir o no.

Es decir, se optó por eliminar las cabeceras repetidas en el caso de tener más de una Titulación o Plan en un mismo fichero. Como se aprecia en la figura 5.6 si se eliminaban las cabeceras repetidas (recuadro rojo), se perdía la información del nombre de la titulación o plan (recuadro verde).

Como esta información no se debía perder, se añadió una nueva columna (llamada *Plan*) como se aprecia en la figura 5.7(en color verde) para almacenar esa información de las cabeceras repetidas, de las cuales se iba a prescindir.

De esta manera se obtuvo un fichero (.csv) final mejor estructurado (con una única cabecera general) y preparado para hacer más eficiente la carga de datos a la BBDD.

Figura 5.6: Datos del fichero original (.xls)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN						
1	Control de Grupos de Matrícula																																													
2	Año : 2018 / 1º Fei ## Cent : 1 / C13 Plan ( 26 Pla 26 Tri Todas																																													
3	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN ENTORNOS SEGUROS																																													
4	Asignatura																																													
5	CódigDescripcu Gr. Tipo L A Tri Vj Turno Máx in Matr Propios																																													
6	Re No Re Nc Má Ex As Dis Mé Asi Disp.																																													
7	8093	INFRAE	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	4564	1014645	ANDUJAS	N	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
8	8094	MODEL	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	1125	53453	Arnaiz GS	S	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
9	8094	MODEL	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	1335	55897	Díez Pas	S	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
10	8095	ARQUI	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	4745	1016217	Martínez	S	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
11	8096	ALMAC	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	4584	1015089	Sánchez	N	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
12	8096	ALMAC	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	4746	1016218	BREGÓN	S	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
13	8097	CONCE	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	7	20	0	0	0	7	1	Manual	4588	1015153	Villafañ	N	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
14	8097	CONCE	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	7	20	0	0	0	7	1	Manual	4747	1016219	TEJERIN	S	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
15	8098	PROCE	1	90	Teori	S	1	Mixto	0	27	0	0	27	0	8	19	0	0	0	8	1	Manual	4585	1015091	Álvarez	N	S	Plan 263 - MÁSTER UNIVERSITARIO EN INTELIGENCIA DE NEGOCIO Y BIG DATA EN E																		
16	8112	ARQUI	1	90	Teori	S	1	Mixto	0	10	0	0	10	0	8	2	0	0	0	8	1	Manual	1403	1000235	Cámara	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
17	8113	INTELI	1	90	Teori	S	1	Mixto	0	15	0	0	15	0	11	4	0	0	0	11	1	Manual	1125	53453	Arnaiz GS	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
18	8113	INTELI	1	90	Teori	S	1	Mixto	0	15	0	0	15	0	11	4	0	0	0	11	1	Manual	1642	39809	Herrero	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
19	8114	ARQUI	1	90	Teori	S	1	Mixto	0	10	0	0	10	0	7	3	0	0	0	7	1	Manual	3137	35796	Redond	N	N	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
20	8114	ARQUI	1	90	Teori	S	1	Mixto	0	10	0	0	10	0	7	3	0	0	0	7	1	Manual	4544	1013983	Rodríguez	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
21	8115	SISTEM	1	90	Teori	S	1	Mixto	0	15	0	0	15	0	12	3	0	0	0	12	1	Manual	1595	90225	González	S	N	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
22	8115	SISTEM	1	90	Teori	S	1	Mixto	0	15	0	0	15	0	12	3	0	0	0	12	1	Manual	3259	48887	Vaquero	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		
23	8116	DIRECC	1	90	Teori	S	2	Mixto	0	15	0	0	15	0	12	3	0	0	0	12	1	Manual	2912	63707	Arco	S	S	Plan 264 - MÁSTER UNIVERSITARIO EN INGENIERÍA INFORMÁTICA (PLAN 2018)																		

Figura 5.7: Datos del fichero parseado con columna añadida (.csv)

## 5.4. Librerías para el tratamiento y manipulación de datos

En este apartado se van a explicar las diferentes librerías o bibliotecas que se han utilizado en el desarrollo del proyecto y su función principal. Todas las librerías explicadas a continuación son de *Python*.

### **re**

La primera librería que se ha utilizado en el proyecto ha sido `re`<sup>19</sup>, ya que dicha librería contiene las funciones necesarias para trabajar con expresiones regulares.

Las expresiones regulares se han utilizado sobretodo en la parte inicial de parsear los ficheros descargados de Sigma. De esta manera se ha podido separar y extraer información por filas, celdas y contenido de las mismas.

### **pandas**

`Pandas`<sup>20</sup> es una librería que ofrece numerosas estructuras de datos de gran rendimiento y herramientas de análisis de datos.

Esta librería se ha utilizado principalmente para abrir y crear archivos con *Python* y para crear *dataframes* o estructuras auxiliares donde guardar datos.

### **sqlite3**

Como ya se ha comentado en un apartado anterior, la librería `sqlite3`<sup>21</sup> proporciona una base de datos relacional de pequeño tamaño, ya que no necesita un servidor para poder utilizarse, porque los datos se almacenan en un único fichero en el sistema host.

Esta librería se ha utilizado para la creación de la base de datos (BBDD), así como para la carga de datos y los procesos de consultas a la misma.

---

<sup>19</sup>[www.docs.python.org/3/library/re](http://www.docs.python.org/3/library/re)

<sup>20</sup>[www.pandas.pydata.org](http://www.pandas.pydata.org)

<sup>21</sup>[www.docs.python.org/2/library/sqlite3.html](http://www.docs.python.org/2/library/sqlite3.html)

## Tkinter

Tkinter<sup>22</sup> es una librería con numerosas funciones para hacer posible la creación de una interfaz gráfica en *Python*.

Se trata de una librería orientada a objetos y gracias a su facilidad de uso y rapidez para realizar una GUI, es una de las librerías más destacadas de este lenguaje de programación.

Por lo tanto, se ha utilizado principalmente en la creación de la interfaz gráfica (botones, desplegados, ventanas...etc).

## math

La librería o componente math<sup>23</sup> incluye principalmente funciones matemáticas, para realizar operaciones aritméticas.

## matplotlib

La librería matplotlib<sup>24</sup> cuenta con multitud de funciones y características para generar gráficos. Hay que destacar que se pueden generar una gran variedad de gráficos (de sector, de barras...), así como personalizar los mismos (ejes, etiquetas, fuente, leyenda...)

Esta librería se ha utilizado para la realización y personalización de los gráficos de la interfaz gráfica.

## os

La librería os<sup>25</sup> de Python permite poder usar funcionalidades relacionadas con el Sistema Operativo.

Las funciones más destacadas de esta librería y las que se han utilizado son las que informan sobre el entorno del Sistema Operativo y las que permiten navegar por la estructura de directorios, ya sea para leer o modificar archivos.

---

<sup>22</sup>[www.docs.python.org/2/library/tkinter.html](http://www.docs.python.org/2/library/tkinter.html)

<sup>23</sup>[www.docs.python.org/3/library/math.html](http://www.docs.python.org/3/library/math.html)

<sup>24</sup>[www.matplotlib.org](http://www.matplotlib.org)

<sup>25</sup>[www.docs.python.org/3/library/os.html](http://www.docs.python.org/3/library/os.html)

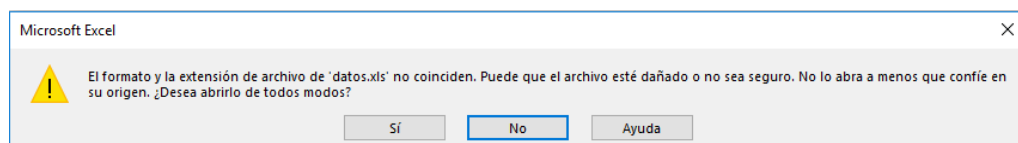


Figura 5.8: Error al intentar abrir con Excel los ficheros originales

## 5.5. Interfaz de usuario del proyecto

Para la realización de este proyecto se ha realizado una GUI (Graphical User Interface) o interfaz de usuario para hacer más sencillo la comunicación entre el usuario y el sistema.

Se ha optado por la realización de una aplicación de escritorio, ya que era la mejor opción para el uso que se espera que se le dé en un futuro a la aplicación.

## 5.6. Problemas encontrados

### Error al abrir los Excel(.xls) bajados de Sigma con Python

En un primer lugar no se esperaba que estos ficheros fueran a generar muchos problemas, ya que al tratarse de una hoja de cálculo de Excel con extensión (.xls), al abrir los ficheros con dicho programa, se mostraba una pantalla de error (5.8), indicando el siguiente tipo de error, el cual el propio programa era capaz de solucionar, pudiendo visualizar todos los datos contenidos en el archivo:

Como se conseguía ver el contenido de los ficheros, así como su formato(color de celdas, celdas unificadas...); se pensó que podría ser un pequeño problema. Realmente se apreció la dimensión del problema cuando se intentó importar y abrir los datos con librerías específicas de *Python* como *Pandas*.

Al comprobar que no se podía cargar u obtener la información de de ninguna manera, se optó por crear un analizador sintáctico o parser para realizar un parseo de los datos. De esta manera se abría el fichero (.xls) como si fuera un fichero de texto, con lo cual obteníamos un fichero (.xml) para analizar. Tras analizar y parsear este último fichero, se obtenía toda la información del (.xls) original (ROW, CELL, MergeDown, MergeAcross, DATA...etc) y con esta información se creaba un fichero resultante (.csv).

### Problema con el uso de la librería *tkinter*

La librería *tkinter* de *Python* tiene numerosas funcionalidades como se ha explicado anteriormente. Ahora bien, el manejo entre diferentes ventanas o botones en ocasiones es bastante complicado. Uno de los principales problemas que se tuvo relacionado con esta librería, fue la obtención de los datos para poder almacenarlos cuando el usuario pulsa sobre una opción de una lista desplegable (*Combobox*). Con la función *comboX.get()* únicamente se establecía el valor por defecto de la lista desplegable o de la primera opción seleccionada, no permitiendo cambiar de opción. Este problema se solucionó mediante el uso del método *bind* (opción «*ComboboxSelected*») acompañado de funciones lambda y pasándolas como parámetro *comboX.get()*.

A parte de este problema, también hay que destacar que en los botones (*Button*), cuando se le pasa el parámetro *command = funcionX*, se procede a la llamada de la *funcionX* cuando se produce el evento de pulsar o hacer clic sobre dicho botón. Ahora bien, si añadimos parámetros a la función *funcionX*, el botón deja de funcionar, ya que se ejecuta una única vez (cuando se abre la ventana donde se sitúa dicho botón) y después de esto, el botón queda inutilizado. Para solucionar este problema, por ejemplo en el botón de *Visualizar Gráfica*, donde se necesita hacer consultas a la BBDD en función de las opciones que seleccione el usuario en los desplegables, se han utilizado variables globales, para que sean accesibles desde donde se encuentra el botón, hasta donde se encuentren las funciones que son llamadas por el propio botón.



---

## Trabajos relacionados

---

Para comenzar este apartado, hay que exponer que existen o podrían existir tantos sistemas de información, como empresas u organizaciones encontramos alrededor del mundo. Esto es así ya que cada empresa u organización posee una estructura y características muy diferenciadas que la hacen única.

Por lo tanto, hay que destacar que existen numerosos trabajos y libros sobre sistemas de información, que explican conceptos teóricos, así como la forma de implantar estos sistemas.

Se ha encontrado numerosa documentación sobre sistemas de información geográfica y para el ámbito administrativo. Un libro de carácter teórico, donde se explican características, categorías y el funcionamiento de los diferentes tipos de sistemas sería el siguiente *Introducción a la gestión de sistemas de información en la empresa*[\[8\]](#).

También existe bastante documentación y trabajos sobre principios de los sistemas de información donde se explica el uso, desarrollo e implantación de los sistemas de información orientados al negocio[\[7\]](#).

Como conclusión de este apartado, destacar que no se han encontrado trabajos de investigación muy similares al proyecto realizado, si bien es cierto que como se ha comentado, existen numerosos desarrollos de otros tipos de sistemas u orientados a otro tipo de soluciones.





---

# Conclusiones y Líneas de trabajo futuras

---

En el último apartado se van a exponer las conclusiones obtenidas tras la realización del proyecto. Adicionalmente se propondrán unas pautas o sugerencias para mejoras futuras del proyecto.

## 7.1. Conclusiones

Tras la finalización del proyecto, se extraen las siguientes conclusiones:

- En primer lugar hay que destacar que se han conseguido los objetivos (general y técnicos) marcados al inicio del proyecto. De esta manera, se ha logrado crear un sistema de información sencillo, útil y fiable, que se espera que sea valioso para su uso en el ámbito de la Universidad de Burgos.
- Gracias a la realización de este proyecto, la Universidad de Burgos posee un sistema de información idóneo para generar gráficos capaces de aportar información visual muy eficaz, pudiendo ser de gran ayuda en la toma de decisiones o detección de problemas relacionados con la matriculación de alumnos.
- Otro aspecto relevante es la facilidad de elección, personalización y generación de los diferentes tipos de gráficos, ya que únicamente con cuatro clics se obtienen los gráficos que se deseen.
- Por último, una conclusión personal. Me gustaría destacar que gracias a la realización de este proyecto, realmente se han encajado o asimilado

una gran parte de las asignaturas cursadas en la carrera, ya que un pequeño proyecto como es el caso, conlleva numerosas tareas (planificación, diseño, investigación, desarrollo, documentación, entre otras muchas) y hasta el momento no se habían cohesionado todas para lograr un mismo fin. Por último destacar que los problemas encontrados durante el desarrollo, han sido uno de los principales focos de aprendizaje, sirviendo para la búsqueda de soluciones y autoaprendizaje continuo que es tan importante en esta carrera.

## 7.2. Líneas de trabajo futuras

Hay que señalar que aunque se realice la entrega del proyecto, no quiere decir que se trate de un proyecto perfecto y cerrado, ya que existen aspectos que se pueden mejorar de cara al futuro. A continuación numero una serie de cambios que me hubiera gustado realizar y los cuales se podrían realizar más adelante:

- Modificar la aplicación para que cuente con una interfaz gráfica más óptima visualmente. Se podría realizar mediante el uso de otras herramientas, ya que como se ha explicado la librería *tkinter* de *Python* tiene ciertas limitaciones.
- Dentro de la BBDD, existen campos de las tablas, cuya información no se ha explotado o utilizado. Sería muy interesante generar otros tipos de gráficos, por ejemplo relacionados con los profesores, ya que se dispone de la estructura e información necesaria.

---

## Bibliografía

---

- [1] Características de los sistemas de información, 2019.
- [2] Vicenç Fernández Alarcón. *Desarrollo de Sistemas de Información. Una Metodología Basada En El Modelado*. Univ. Politèc. de Catalunya, 2006.
- [3] Valor J Andreu R, Ricart JE. *Estrategia y sistemas de información*. McGraw-Hill, 1991.
- [4] Instituto Nacional de Estadística. Tipos de gráficos, 2019.
- [5] Kyocera. Los 6 principales tipos de sistemas de información, 2017.
- [6] Moisés Ortiz. Excel total, 2019.
- [7] Effy Oz. *ADMINISTRACIÓN DE LOS SISTEMAS DE INFORMACIÓN*. CENGAGE Learning, 2008.
- [8] Joaquín Guiral Herrando Rafael Lapiedra Alcamí, Carlos Devece Carañana. *Introducción a la gestión de sistemas de información en la empresa*. Publicacions de la Universitat Jaume I, 2018.
- [9] Julia Máxima Uriarte. 10 características de un sistema de información, 2019.
- [10] Wikipedia. Texmaker — wikipedia, la enciclopedia libre, 2018.
- [11] Wikipedia. Sistemas de información ejecutiva — wikipedia, la enciclopedia libre, 2019.
- [12] Wikipedia. Sqlite — wikipedia, la enciclopedia libre, 2019.