



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

**Aplicación Web para la
recopilación, tratamiento y
visualización de datos
públicos**



Presentado por Iván Arjona Alonso
en Universidad de Burgos — 23 de abril
de 2018

Tutor: Dr. José Francisco Díez Pastor
y Dr. Jesús Manuel Maudes Raedo

Índice general

Índice general	I
Índice de figuras	III
Índice de tablas	IV
Apéndice A Plan de Proyecto Software	1
A.1. Introducción	1
A.2. Planificación temporal	1
A.3. Estudio de viabilidad	5
Apéndice B Especificación de Requisitos	7
B.1. Introducción	7
B.2. Objetivos generales	7
B.3. Catalogo de requisitos	7
B.4. Especificación de requisitos	7
Apéndice C Especificación de diseño	9
C.1. Introducción	9
C.2. Diseño de datos	9
C.3. Diseño procedimental	9
C.4. Diseño arquitectónico	9
Apéndice D Documentación técnica de programación	11
D.1. Introducción	11
D.2. Estructura de directorios	11
D.3. Manual del programador	11

D.4. Compilación, instalación y ejecución del proyecto	11
D.5. Pruebas del sistema	11
Apéndice E Documentación de usuario	13
E.1. Introducción	13
E.2. Requisitos de usuarios	13
E.3. Instalación	13
E.4. Manual del usuario	13
Bibliografía	15

Índice de figuras

A.1. Burndown del sprint 0	2
A.2. Burndown del sprint 1	3
A.3. Burndown del sprint 2	3
A.4. Burndown del sprint 3	4
A.5. Burndown del sprint 4	5

Índice de tablas

Apéndice A

Plan de Proyecto Software

A.1. Introducción

A.2. Planificación temporal

El desarrollo del proyecto se ha llevado a cabo utilizando metodologías ágiles, basándose en la metodología *scrum* con algunas modificaciones (una sola persona y sin reuniones diarias).

Se aplicó una estrategia de desarrollo incremental, con iteraciones que llamaremos *sprints*.

El resultado de cada iteración es un entregable, sobre el que se discute en la reunión posterior a cada sprint.

Se realizó, en principio, una reunión a la semana con los tutores para exponer las modificaciones realizadas en el sprint anterior y planificar los cambios a realizar en la siguiente iteración.

Estas tareas están priorizadas por el tiempo estimado de su realización, se puede ver esta estimación en el enlace de cada sprint a sus tareas.

A continuación se va a realizar un breve resumen de las tareas realizadas en cada una de las iteraciones, así como la duración de cada sprint y el gráfico *burndown* correspondiente.

Sprint 0 (16/02/2018 - 02/03/2018)

Primer sprint del proyecto. En la reunión de planificación de este primer sprint se discute de forma general de lo que va a tratar el proyecto.

Las tareas realizadas durante este sprint fueran la creación y configuración del repositorio y sobre todo investigar sobre las tecnologías y herramientas que se podían utilizar.

Se investigaron posibles fuentes de datos para implementar más adelante: INE, sepe, aeat.

La duración fue de dos semanas en lugar de una para poder documentarse sobre todos los aspectos relevantes del proyecto.

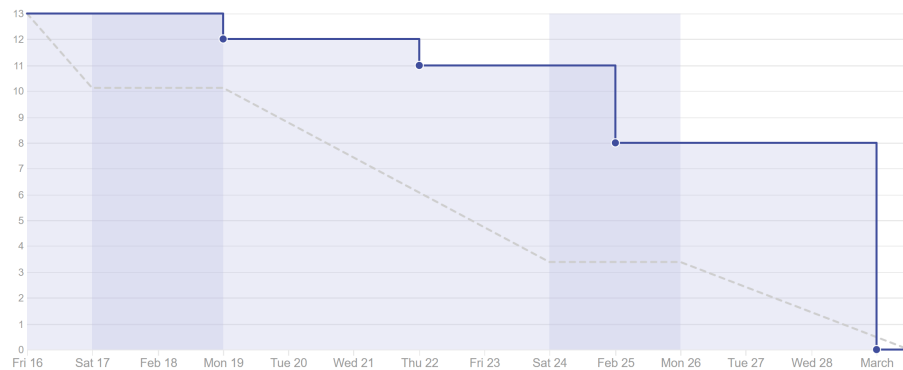


Figura A.1: Burndown del sprint 0

Tareas del sprint 0 en Github

Sprint 1 (03/03/2018 - 08/03/2018)

Un objetivo de este sprint es investigar alternativas de bases de datos no relacionadas que se podrían utilizar. Se ha elegido MongoDB.

El otro objetivo es empezar a implementar prototipos con las fuentes de datos que se habían encontrado en el sprint anterior. Se implementaron prototipos del INE, de la agencia tributaria y del sepe.

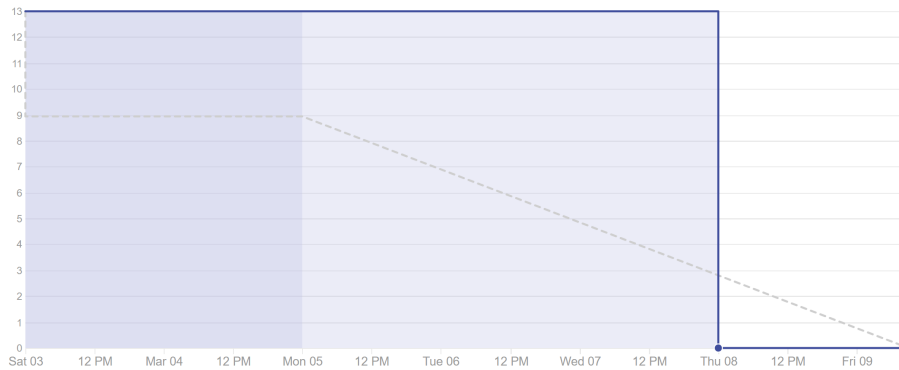


Figura A.2: Burndown del sprint 1

[Tareas del sprint 1 en Github](#)

Sprint 2 (09/03/2018 - 15/03/2018)

El primer objetivo de este sprint es crear la estructura de la página web con Flask. Utilizando un modelo vista-controlador. También se utiliza bootstrap para ahorrar trabajo en el diseño.

Se implementó un prototipo de la carga de datos hacia la base de datos y otro para la descarga de datos desde la base de datos para ser mostrados.

Se hizo una implementación de las fuentes de datos a partir de los prototipos del sprint 1 de modo que se pueda cargar todas las fuentes de manera automática.

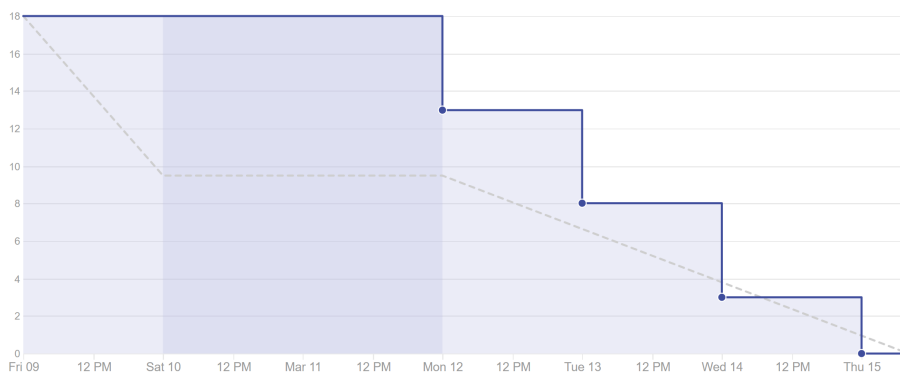


Figura A.3: Burndown del sprint 2

[Tareas del sprint 2 en Github](#)

Sprint 3 (16/03/2018 - 22/03/2018)

En este sprint se sopesaron varias plataformas para hacer el despliegue de la web. De ellas se eligió DigitalOcean y Nanobox.

Se realizó el despliegue utilizando estas plataformas. **Web desplagada.**

Se investigaron los posibles riesgos de seguridad como inyecciones NoSQL.

Se implementó un formulario para la consulta de datos en la página web. En esta primera aproximación se podía hacer una consulta comparando con una columna de una de las fuentes de datos.

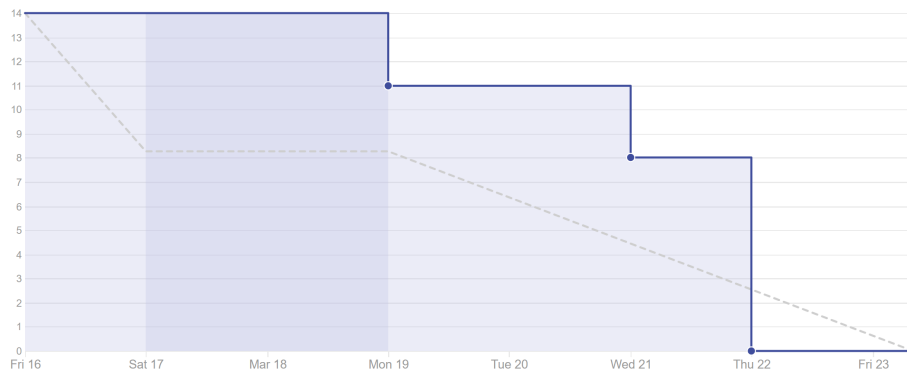


Figura A.4: Burndown del sprint 3

Tareas del sprint 3 en Github

Sprint 4 (23/03/2018 - 13/04/2018)

Este sprint coincide con semana santa, por lo que dura una semana más de lo habitual y la carga de trabajo también es mayor.

Se corrigieron errores en los tipos de las fuentes de datos al tratar con números como cadenas.

Se implementó una forma de descargar las consultas a partir del formulario.

Se mejoró interfaz gráfica y el formulario de consulta.

Se modificaron las fuentes de datos para corregir errores y añadir el código de municipio a todas ellas para más tarde poder unir las.

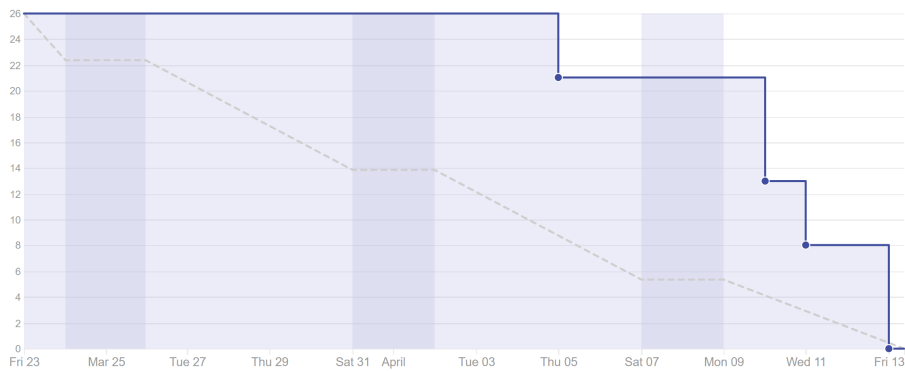


Figura A.5: Burndown del sprint 4

Tareas del sprint 4 en Github

A.3. Estudio de viabilidad

Viabilidad económica

Viabilidad legal

Apéndice B

Especificación de Requisitos

- B.1. Introducción
- B.2. Objetivos generales
- B.3. Catalogo de requisitos
- B.4. Especificación de requisitos

Apéndice C

Especificación de diseño

- C.1. Introducción
- C.2. Diseño de datos
- C.3. Diseño procedimental
- C.4. Diseño arquitectónico

Apéndice D

Documentación técnica de programación

- D.1. Introducción
- D.2. Estructura de directorios
- D.3. Manual del programador
- D.4. Compilación, instalación y ejecución
del proyecto
- D.5. Pruebas del sistema

Apéndice E

Documentación de usuario

- E.1. Introducción
- E.2. Requisitos de usuarios
- E.3. Instalación
- E.4. Manual del usuario

Bibliografía
