



**ESCUELA POLITÉCNICA
SUPERIOR DE CÓRDOBA**
Universidad de Córdoba



TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

ANEXO: MANUAL DE USUARIO

Análisis, Diseño e Implementación de un Dashboard
Interactivo para el Análisis de Supervivencia aplicado al
Abandono Escolar Universitario

Autor: Lucía Cabezuelo Pérez

Director: Dr. Cristóbal Romero Morales y Dr. Juan Alfonso Lara
Torralbo

Septiembre, 2025



UNIVERSIDAD DE CÓRDOBA



ÍNDICE

1. Introducción	5
2. Instalación	6
2.1 Requisitos previos	7
2.2 Descarga del código	7
2.3 Instalación de las librerías de Python requeridas	9
2.4 Desinstalación del proyecto	9
3. Interfaz	10
3.1 Home	10
3.2 Inicio	13
3.3 Ver Dataset	14
3.4 Análisis de Covariables	15
3.5 Análisis de Supervivencia	17
3.5.1 Kaplan-Meier	19
3.5.2 Regresión de Cox	22
3.5.3 Pruebas de Log Rank	26
3.5.4 Prompt IA	30

Índice de Imágenes

Ilustración 1: Descarga manual	8
--------------------------------	---

Ilustración 2: Pantalla de inicio	11
Ilustración 3: Mensaje de error	12
Ilustración 4: CSV bruto cargado	12
Ilustración 5: Dataset limpio cargado	13
Ilustración 6: Diálogo de Confirmación.....	14
Ilustración 7: Sección Ver dataset	15
Ilustración 8: Sección Análisis de Covariables	15
Ilustración 9: Gráfica variable evento	16
Ilustración 10: Análisis de covariables Género.....	17
Ilustración 11: Análisis de covariables Discapacidad	17
Ilustración 12: Sección Análisis de supervivencia.....	18
Ilustración 13: Modelos del análisis de supervivencia	19
Ilustración 14: Sección de Kaplan-Meier.....	19
Ilustración 15: Curva Supervivencia del Abandono	20
Ilustración 16: Curva de Kaplan para el Género	21
Ilustración 17: Curva de Kaplan para la Discapacidad	21
Ilustración 18: Botón "Ninguna"	22
Ilustración 19: Sección de Regresión de Cox.....	22
Ilustración 20: Interfaz de la Regresión de Cox	23
Ilustración 21: Barra de selección Cox.....	23
Ilustración 22: Opción covariable Género	24
Ilustración 23: Opción covariable Discapacidad.....	24
Ilustración 24: Opción covariables Género y Discapacidad	24
Ilustración 25: Sección Pruebas de Log Rank.....	26
Ilustración 26: Interfaz Pruebas de Log Rank	27
Ilustración 27: Opciones de selección de Log Rank	27
Ilustración 28: Opción covariable Género	28
Ilustración 29: Opción covariable Discapacidad.....	28

Ilustración 30: Opción covariables Género y Discapacidad29

Ilustración 31: Prompt IA30

Índice de Tablas

Tabla 1: Variables estudiadas Regresión de Cox26

Tabla 2: Variables estudiadas Pruebas de Log Rank.....30

1. Introducción

El propósito de este manual de usuario es proporcionar una guía detallada y accesible que facilite el uso del dashboard de análisis del abandono estudiantil en las

universidades, explicando de manera clara y comprensible sus funciones principales. Esta herramienta se ha diseñado con el objetivo de proporcionar una visión integral y dinámica de los datos relacionados con la permanencia y el abandono de los estudiantes.

En el desarrollo de dicho documento, se presentarán las instrucciones que el usuario necesita para interactuar con la herramienta, abarcando desde el inicio hasta la ejecución de las diversas opciones disponibles. Se describen los procedimientos para navegar por las diferentes secciones, filtrar y visualizar los datos, interpretar los gráficos y obtener la información significativa

Se incluyen ciertas recomendaciones de uso y buenas prácticas para garantizar que la experiencia de los usuarios sea óptima y eficiente, aprovechando todo el potencial de la herramienta.

Con un enfoque sencillo y directo, este manual está pensado como un recurso complementario tanto para los usuarios de nuevo ingreso o sin experiencia en el uso de dashboard o herramientas analíticas, como para los que usuarios con mayor experiencia o más avanzadas que deseen profundizar en las funcionalidades disponibles, brindándoles la confianza y la guía necesaria para utilizar la herramienta con éxito.

2. Instalación

Para el correcto uso por parte del usuario de la aplicación desarrollada en este proyecto, es fundamental seguir una serie de pasos de instalación que garantizan su

funcionamiento adecuado en tu equipo local. La herramienta está basada en el lenguaje de Python, y hace uso de diversas bibliotecas que facilitan el procesamiento de datos y su visualización interactiva.

En este proyecto no ha hecho falta ningún coste, ya que no se basa en ninguna inversión económica, El uso de esta herramienta es totalmente gratuito y, a parte de Python, la plataforma se ha construido mediante la biblioteca Dash.

El objetivo de esta sección del proyecto es el proporcionar al usuario una guía exhaustiva y ordenada para que el ingreso en la plataforma sea lo más exitoso posible.

2.1 Requisitos previos

Antes de ejecutar la herramienta, el usuario tiene que asegurarse de cumplir los siguientes requisitos:

1. Tener instalado Python 3.8 o superior en su equipo. En caso de no tenerlo, puedes descargarlo desde *python.org*.
2. Instalar las librerías requeridas: este paso puede llevarse a cabo simplemente ejecutando el siguiente comando en la terminal para instalar las librerías necesarias:

```
pip install -r requirements.txt
```

3. Google Drive: el archivo CSV principal, del que se realiza el análisis, al ser de un gran tamaño para ser subido a GitHub directamente. Como paso primordial en el proceso hay que descargarlo desde el enlace proporcionado en la carpeta README.md:

[Descargar CSV bruto](#)

2.2 Descarga del código

Para la descarga de y preparación del código es trascendental seguir estos pasos:

1. Descarga el proyecto como archivo comprimido ZIP.

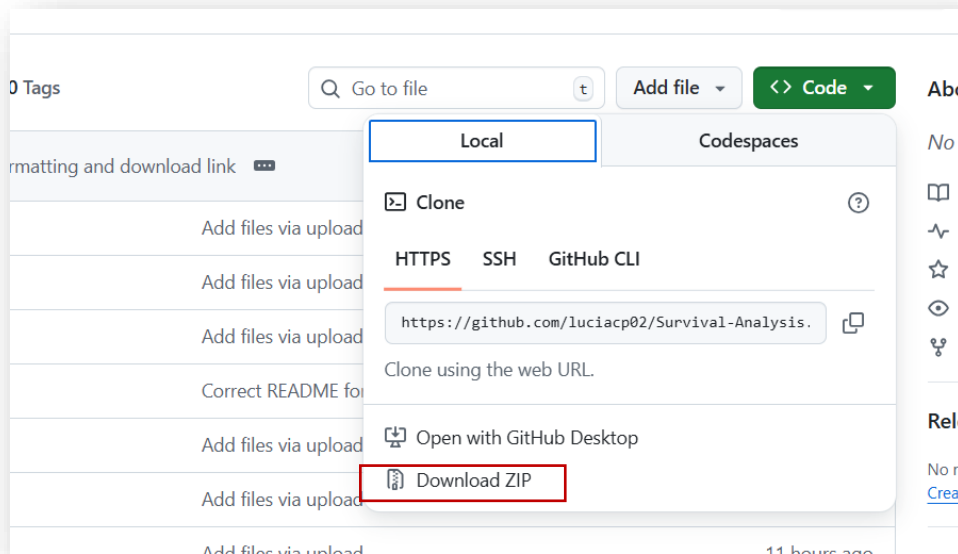


Ilustración 1: Descarga manual

2. En el caso de querer clonar el repositorio, el usuario tendrá que ejecutar en la terminal lo siguiente:

```
git clone https://github.com/luciap02/Survival-Analysis.git
```

3. **Descarga el archivo CSV:** Como se ha mencionado anteriormente. Este archivo está disponible en Google Drive y se puede acceder mediante README.md. El usuario puede hacer clic en el enlace de descarga que se proporciona en el manual o en el repositorio de GitHub.
4. Una vez instalado el archivo CSV, **sítualo en la carpeta principal**, donde se encuentran los archivos .py del proyecto para asegurar que el código funcione de manera correcta. El archivo CSV debe estar en el mismo directorio que el script *cargaDataset.py* para que la ejecución se pueda llevar a cabo con éxito.
5. **Ejecuta el código:** Una vez que hayas descargado el archivo CSV y colocado en la carpeta principal, ejecuta el script principal *cargaDataset.py* para iniciar el análisis:

```
python cargaDataset.py
```

6. Tras esperar unos segundos de la ejecución, o en el caso de error, volver a ejecutar,

el navegador abrirá la interfaz web del dashboard para comenzar con el análisis de los datos.

2.3 Instalación de las librerías de Python requeridas

1. En primer lugar, el usuario tendrá que asegurarse de tener instalado *pip*. Para ello, se tendrá que ejecutar en la terminal:

```
pip --version
```

En caso de estar instalado, debería aparecer la versión correspondiente, en caso de no tenerla instalado, puedes acceder a través de tu navegador a su instalación.

2. Seguidamente, hay que descargar las dependencias a través del archivo *requirements.txt*. Este archivo contiene todas las librerías necesarias y facilita su instalación de una sola vez ejecutando el comando en la carpeta del proyecto:

```
pip install -r requirements.txt
```

3. Una vez realizado este proceso, el usuario debe asegurarse de su instalación con la siguiente ejecución que mostrará las librerías descargadas:

```
pip list
```

2.4 Desinstalación del proyecto

Si el usuario desea de dejar de usar la herramienta, es posible proceder con su desinstalación de manera sencilla.

1. Eliminar repositorio.

Este proceso se puede realizar a mano, navegando a la carpeta del proyecto, seleccionándola y suprimiéndola.

2. Eliminar las librerías instaladas

Al haber instalado las librerías de forma global, sin entorno virtual, se puede realizar el proceso de desinstalación mediante *pip* con los siguientes comandos:

Esta ejecución listará las librerías:

```
pip freeze
```

Mediante estos comandos se desinstalarán todas ellas:

```
pip freeze | xargs pip uninstall -y
```

3. Eliminar Python

Como último paso opcional, el usuario puede desinstalar Python, si no desea mantenerlo en el sistema, de la siguiente manera:

- 1) Abrir el Panel de control
- 2) Hacer clic en “Programas”—“Desinstalar un programa”
- 3) Busca **Python** en la lista de programas instalados
- 4) Seleccionar en “Desinstalar”

3. Interfaz

Como ya se ha mencionado y detallado antes, la interfaz ha sido diseñada con el objetivo de ser clara, intuitiva y fácil de utilizar, teniendo en cuenta que muchos usuarios pueden acceder sin conocimientos técnicos avanzados.

3.1 Home

El acceso a la plataforma se considera bastante sencillo, debido a que al ejecutar el código del archivo principal de la aplicación (cargaDataset.py) y acceder a la dirección IP que aparece, la máquina donde se esté realizando el proceso muestra por pantalla la página inicial del Dashboard, abriendo una nueva pestaña en el navegador de forma

instantánea.

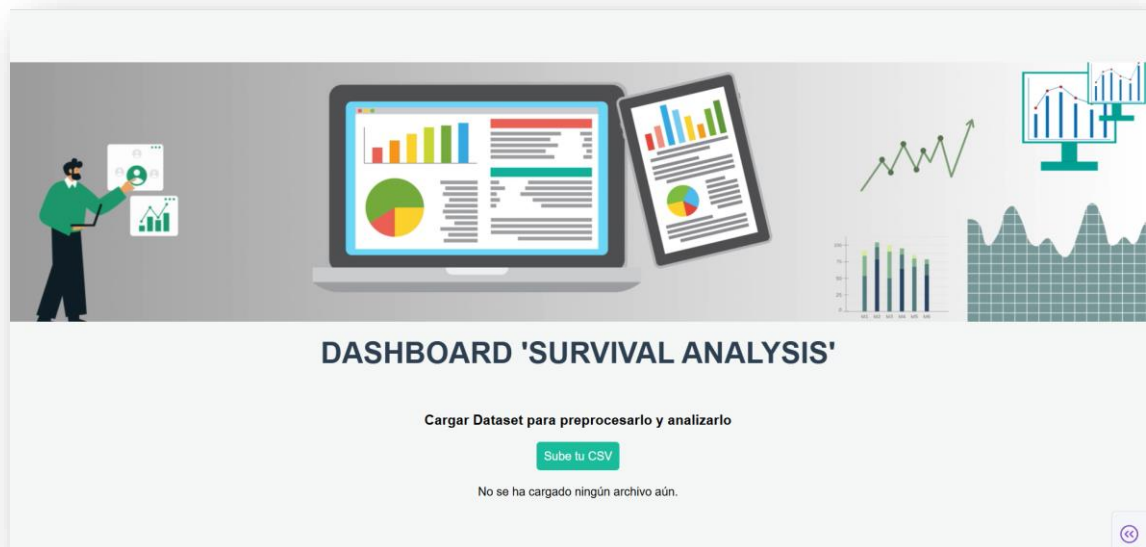


Ilustración 2: Pantalla de inicio

En esta pantalla inicial encontramos los siguientes elementos, con los cuáles dan la bienvenida al usuario y muestran unas breves instrucciones al usuario con los pasos a seguir para pasar a la siguiente etapa del análisis:

- Banner de bienvenida: logo interactivo de la plataforma dando a entender al usuario.
- Botón para la carga del dataset **“Sube el CSV”**: se muestra como única opción en esta sección, mediante el cual el usuario sube el dataset, y este, es mostrado por pantalla.
- Mensaje de error por defecto, hasta que el usuario cargue un dataset de forma correcta y este aparezca por pantalla. *“No se ha cargado ningún archivo aún”*

Cuando el usuario selecciona la opción para cargar los datos, se muestra por pantalla un recuadro interactivo para subir el CSV requerido de tus archivos. En este proyecto, solo se permite subir un CSV predeterminado para realizar el análisis, en caso contrario, se mostrará por pantalla un mensaje de error donde indicará el dataset esperado



Ilustración 3: Mensaje de error

Una vez el usuario sube el archivo que se indica en formato CSV, la interfaz muestra ese archivo visualmente en forma de tabla, donde se recogen todos los elementos y atributos del CSV bruto.

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

ANEXO: MANUAL DE USUARIO

Análisis, Diseño e Implementación de un Dashboard Interactivo para el Análisis de Supervivencia aplicado al Abandono Escolar Universitario

Autor: Lucía Cabezuelo Pérez

Director: Dr. Cristóbal Romero Morales y Dr. Juan Alfonso Lara
Torralbo

Septiembre, 2025

ÍNDICE

1. Introducción	5
2. Instalación	6
2.1 Requisitos previos	7
2.2 Descarga del código	7
2.3 Instalación de las librerías de Python requeridas	9
2.4 Desinstalación del proyecto	9
3. Interfaz	10
3.1 Home	10
3.2 Inicio	13
3.3 Ver Dataset	14
3.4 Análisis de Covariables	15
3.5 Análisis de Supervivencia	17
3.5.1 Kaplan-Meier	19
3.5.2 Regresión de Cox	22

3.5.3	Pruebas de Log Rank	26
3.5.4	Prompt IA	30

Índice de Imágenes

Ilustración 1: Descarga manual	8
Ilustración 2: Pantalla de inicio	11
Ilustración 3: Mensaje de error	12
Ilustración 4: CSV bruto cargado	12
Ilustración 5: Dataset limpio cargado	13
Ilustración 6: Diálogo de Confirmación.....	14
Ilustración 7: Sección Ver dataset	15
Ilustración 8: Sección Análisis de Covariables	15
Ilustración 9: Gráfica variable evento	16
Ilustración 10: Análisis de covariables Género.....	17
Ilustración 11: Análisis de covariables Discapacidad	17
Ilustración 12: Sección Análisis de supervivencia.....	18
Ilustración 13: Modelos del análisis de supervivencia	19
Ilustración 14: Sección de Kaplan-Meier.....	19
Ilustración 15: Curva Supervivencia del Abandono	20
Ilustración 16: Curva de Kaplan para el Género	21
Ilustración 17: Curva de Kaplan para la Discapacidad	21
Ilustración 18: Botón "Ninguna"	22
Ilustración 19: Sección de Regresión de Cox.....	22

Ilustración 20: Interfaz de la Regresión de Cox	23
Ilustración 21: Barra de selección Cox.....	23
Ilustración 22: Opción covariable Género	24
Ilustración 23: Opción covariable Discapacidad.....	24
Ilustración 24: Opción covariables Género y Discapacidad	24
Ilustración 25: Sección Pruebas de Log Rank.....	26
Ilustración 26: Interfaz Pruebas de Log Rank	27
Ilustración 27: Opciones de selección de Log Rank	27
Ilustración 28: Opción covariable Género	28
Ilustración 29: Opción covariable Discapacidad.....	28
Ilustración 30: Opción covariables Género y Discapacidad	29
Ilustración 31: Prompt IA	30

Índice de Tablas

Tabla 1: Variables estudiadas Regresión de Cox	26
Tabla 2: Variables estudiadas Pruebas de Log Rank.....	30

4. Introducción

El propósito de este manual de usuario es proporcionar una guía detallada y accesible que facilite el uso del dashboard de análisis del abandono estudiantil en las universidades, explicando de manera clara y comprensible sus funciones principales. Esta herramienta se ha diseñado con el objetivo de proporcionar una visión integral y dinámica de los datos relacionados con la permanencia y el abandono de los estudiantes.

En el desarrollo de dicho documento, se presentarán las instrucciones que el usuario necesita para interactuar con la herramienta, abarcando desde el inicio hasta la ejecución de las diversas opciones disponibles. Se describen los procedimientos para navegar por las diferentes secciones, filtrar y visualizar los datos, interpretar los gráficos y obtener la información significativa

Se incluyen ciertas recomendaciones de uso y buenas prácticas para garantizar que la experiencia de los usuarios sea óptima y eficiente, aprovechando todo el potencial de la herramienta.

Con un enfoque sencillo y directo, este manual está pensado como un recurso complementario tanto para los usuarios de nuevo ingreso o sin experiencia en el uso de dashboard o herramientas analíticas, como para los que usuarios con mayor experiencia o más avanzadas que deseen profundizar en las funcionalidades disponibles, brindándoles la confianza y la guía necesaria para utilizar la herramienta con éxito.

5. Instalación

Para el correcto uso por parte del usuario de la aplicación desarrollada en este proyecto, es fundamental seguir una serie de pasos de instalación que garantizan su funcionamiento adecuado en tu equipo local. La herramienta está basada en el lenguaje de Python, y hace uso de diversas bibliotecas que facilitan el procesamiento de datos y su visualización interactiva.

En este proyecto no ha hecho falta ningún coste, ya que no se basa en ninguna inversión económica, El uso de esta herramienta es totalmente gratuito y, a parte de Python, la plataforma se ha construido mediante la biblioteca Dash.

El objetivo de esta sección del proyecto es el proporcionar al usuario una guía exhaustiva y ordenada para que el ingreso en la plataforma sea lo más exitoso posible.

5.1 Requisitos previos

Antes de ejecutar la herramienta, el usuario tiene que asegurarse de cumplir los siguientes requisitos:

4. Tener instalado Python 3.8 o superior en su equipo. En caso de no tenerlo, puedes descargarlo desde *python.org*.
5. Instalar las librerías requeridas: este paso puede llevarse a cabo simplemente ejecutando el siguiente comando en la terminal para instalar las librerías necesarias:

```
pip install -r requirements.txt
```

6. Google Drive: el archivo CSV principal, del que se realiza el análisis, al ser de un gran tamaño para ser subido a GitHub directamente. Como paso primordial en el proceso hay que descargarlo desde el enlace proporcionado en la carpeta README.md:

Descargar CSV bruto

5.2 Descarga del código

Para la descarga de y preparación del código es trascendental seguir estos pasos:

7. Descarga el proyecto como archivo comprimido ZIP.

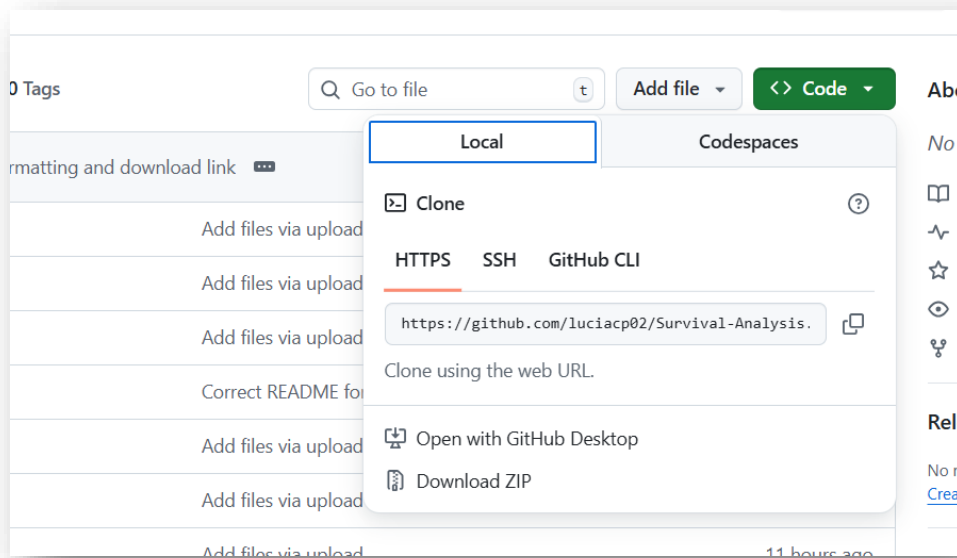


Ilustración 1: Descarga manual

8. En el caso de querer clonar el repositorio, el usuario tendrá que ejecutar en la terminal lo siguiente:

```
git clone https://github.com/luciACP02/Survival-Analysis.git
```

9. **Descarga el archivo CSV:** Como se ha mencionado anteriormente. Este archivo está disponible en Google Drive y se puede acceder mediante README.md. El usuario puede hacer clic en el enlace de descarga que se proporciona en el manual o en el repositorio de GitHub.

10. Una vez instalado el archivo CSV, **sítúalo en la carpeta principal**, donde se encuentran los archivos .py del proyecto para asegurar que el código funcione de manera correcta. El archivo CSV debe estar en el mismo directorio que el script *cargaDataset.py* para que la ejecución se pueda llevar a cabo con éxito.
11. **Ejecuta el código:** Una vez que hayas descargado el archivo CSV y colocado en la carpeta principal, ejecuta el script principal *cargaDataset.py* para iniciar el análisis:

```
python cargaDataset.py
```

12. Tras esperar unos segundos de la ejecución, o en el caso de error, volver a ejecutar, el navegador abrirá la interfaz web del dashboard para comenzar con el análisis de los datos.

5.3 Instalación de las librerías de Python requeridas

4. En primer lugar, el usuario tendrá que asegurarse de tener instalado *pip*. Para ello, se tendrá que ejecutar en la terminal:

```
pip --version
```

En caso de estar instalado, debería aparecer la versión correspondiente, en caso de no tenerla instalado, puedes acceder a través de tu navegador a su instalación.

5. Seguidamente, hay que descargar las dependencias a través del archivo *requirements.txt*. Este archivo contiene todas las librerías necesarias y facilita su instalación de una sola vez ejecutando el comando en la carpeta del proyecto:

```
pip install -r requirements.txt
```

6. Una vez realizado este proceso, el usuario debe asegurarse de su instalación con la siguiente ejecución que mostrará las librerías descargadas:

```
pip list
```

5.4 Desinstalación del proyecto

Si el usuario desea de dejar de usar la herramienta, es posible proceder con su desinstalación de manera sencilla.

4. Eliminar repositorio.

Este proceso se puede realizar a mano, navegando a la carpeta del proyecto, seleccionándola y suprimiéndola.

5. Eliminar las librerías instaladas

Al haber instalado las librerías de forma global, sin entorno virtual, se puede realizar el proceso de desinstalación mediante *pip* con los siguientes comandos:

Esta ejecución listará las librerías:

```
pip freeze
```

Mediante estos comandos se desinstalarán todas ellas:

```
pip freeze | xargs pip uninstall -y
```

6. Eliminar Python

Como último paso opcional, el usuario puede desinstalar Python, si no desea mantenerlo en el sistema, de la siguiente manera:

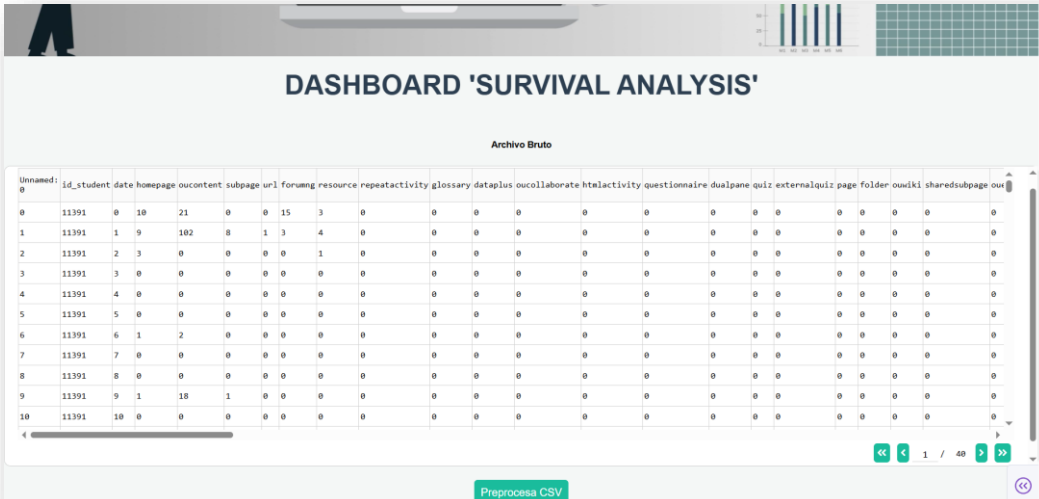
- 5) Abrir el Panel de control
- 6) Hacer clic en “Programas”—“Desinstalar un programa”
- 7) Busca **Python** en la lista de programas instalados
- 8) Seleccionar en “Desinstalar”

6. Interfaz

Como ya se ha mencionado y detallado antes, la interfaz ha sido diseñada con el objetivo de ser clara, intuitiva y fácil de utilizar, teniendo en cuenta que muchos usuarios pueden acceder sin conocimientos técnicos avanzados.

6.1 Home

El acceso a la plataforma se considera bastante sencillo, debido a que al ejecutar el código del archivo principal de la aplicación (cargaDataset.py) y acceder a la dirección IP que aparece, la máquina donde se esté realizando el proceso muestra por pantalla la página inicial del Dashboard, abriendo una nueva pestaña en el navegador de forma instantánea.



The screenshot shows a web dashboard titled "DASHBOARD 'SURVIVAL ANALYSIS'". Below the title, there is a section labeled "Archivo Bruto" (Raw File). This section contains a table with 11 columns and 11 rows of data. The columns are labeled: id_student, date, homepage, oucontent, subpage, url, forumg, resource, repeatactivity, glossary, dataplus, oucollaborate, htactivity, questionnaire, dualpane, quiz, externalquiz, page, folder, ouwiki, sharedsubpage, and ou. The data is organized into a grid where each cell contains a numerical value, mostly 0 or 1, representing the presence or absence of a feature for each student. At the bottom of the table, there is a green button labeled "Preprocesa CSV" (Preprocess CSV). The interface also includes a sidebar on the left with a logo and a top navigation bar with a bar chart and a grid icon.

Unnamed: 0	id_student	date	homepage	oucontent	subpage	url	forumg	resource	repeatactivity	glossary	dataplus	oucollaborate	htactivity	questionnaire	dualpane	quiz	externalquiz	page	folder	ouwiki	sharedsubpage	ou
0	11391	0	10	21	0	0	15	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	11391	1	9	102	8	1	3	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	11391	2	3	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	11391	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	11391	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	11391	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	11391	6	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	11391	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	11391	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	11391	9	1	18	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	11391	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Ilustración 4: CSV bruto cargado

En el momento que el conjunto de datos se ha cargado y mostrado en el dashboard, el usuario, cuenta únicamente con la opción de **“Preprocesar CSV”**, mediante el botón proporcionado seguidamente del dataset limpio. En este proceso el sistema cuenta con la funcionalidad clave para procesar y depurar los datos del conjunto bruto, y posteriormente poderlos utilizar en el análisis de supervivencia.

Gracias a este flujo de interacción, de la interfaz con el sistema de preprocesamiento, el sistema carga, en una tabla interactiva, el CSV refinado, mostrando de igual forma los datos en una tabla, pero de manera organizada y preparada para poder realizar el proceso de análisis.

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

ANEXO: MANUAL DE USUARIO

Análisis, Diseño e Implementación de un Dashboard
Interactivo para el Análisis de Supervivencia aplicado al
Abandono Escolar Universitario

Autor: Lucía Cabezuelo Pérez

Director: Dr. Cristóbal Romero Morales y Dr. Juan Alfonso Lara
Torralbo

Septiembre, 2025

ÍNDICE

1. Introducción	5
-----------------------	---

2. Instalación	6
2.1 Requisitos previos	7
2.2 Descarga del código	7
2.3 Instalación de las librerías de Python requeridas	9
2.4 Desinstalación del proyecto	9
3. Interfaz	10
3.1 Home	10
3.2 Inicio	13
3.3 Ver Dataset	14
3.4 Análisis de Covariables	15
3.5 Análisis de Supervivencia	17
3.5.1 Kaplan-Meier	19
3.5.2 Regresión de Cox	22
3.5.3 Pruebas de Log Rank	26
3.5.4 Prompt IA	30

Índice de Imágenes

Ilustración 1: Descarga manual	8
Ilustración 2: Pantalla de inicio	11
Ilustración 3: Mensaje de error	12
Ilustración 4: CSV bruto cargado	12

Ilustración 5: Dataset limpio cargado	13
Ilustración 6: Diálogo de Confirmación.....	14
Ilustración 7: Sección Ver dataset	15
Ilustración 8: Sección Análisis de Covariables	15
Ilustración 9: Gráfica variable evento	16
Ilustración 10: Análisis de covariables Género.....	17
Ilustración 11: Análisis de covariables Discapacidad	17
Ilustración 12: Sección Análisis de supervivencia.....	18
Ilustración 13: Modelos del análisis de supervivencia	19
Ilustración 14: Sección de Kaplan-Meier.....	19
Ilustración 15: Curva Supervivencia del Abandono	20
Ilustración 16: Curva de Kaplan para el Género	21
Ilustración 17: Curva de Kaplan para la Discapacidad	21
Ilustración 18: Botón "Ninguna"	22
Ilustración 19: Sección de Regresión de Cox.....	22
Ilustración 20: Interfaz de la Regresión de Cox	23
Ilustración 21: Barra de selección Cox.....	23
Ilustración 22: Opción covariable Género	24
Ilustración 23: Opción covariable Discapacidad.....	24
Ilustración 24: Opción covariables Género y Discapacidad	24
Ilustración 25: Sección Pruebas de Log Rank.....	26
Ilustración 26: Interfaz Pruebas de Log Rank	27
Ilustración 27: Opciones de selección de Log Rank	27
Ilustración 28: Opción covariable Género	28
Ilustración 29: Opción covariable Discapacidad.....	28
Ilustración 30: Opción covariables Género y Discapacidad	29
Ilustración 31: Prompt IA	30

Índice de Tablas

Tabla 1: Variables estudiadas Regresión de Cox	26
Tabla 2: Variables estudiadas Pruebas de Log Rank.....	30

7. Introducción

El propósito de este manual de usuario es proporcionar una guía detallada y accesible que facilite el uso del dashboard de análisis del abandono estudiantil en las universidades, explicando de manera clara y comprensible sus funciones principales. Esta herramienta se ha diseñado con el objetivo de proporcionar una visión integral y dinámica de los datos relacionados con la permanencia y el abandono de los estudiantes.

En el desarrollo de dicho documento, se presentarán las instrucciones que el usuario necesita para interactuar con la herramienta, abarcando desde el inicio hasta la ejecución de las diversas opciones disponibles. Se describen los procedimientos para navegar por las diferentes secciones, filtrar y visualizar los datos, interpretar los gráficos y obtener la información significativa

Se incluyen ciertas recomendaciones de uso y buenas prácticas para garantizar que la experiencia de los usuarios sea óptima y eficiente, aprovechando todo el potencial de la herramienta.

Con un enfoque sencillo y directo, este manual está pensado como un recurso complementario tanto para los usuarios de nuevo ingreso o sin experiencia en el uso de dashboard o herramientas analíticas, como para los que usuarios con mayor experiencia o más avanzadas que deseen profundizar en las funcionalidades disponibles, brindándoles la confianza y la guía necesaria para utilizar la herramienta con éxito.

8. Instalación

Para el correcto uso por parte del usuario de la aplicación desarrollada en este proyecto, es fundamental seguir una serie de pasos de instalación que garantizan su funcionamiento adecuado en tu equipo local. La herramienta está basada en el lenguaje de Python, y hace uso de diversas bibliotecas que facilitan el procesamiento de datos y su visualización interactiva.

En este proyecto no ha hecho falta ningún coste, ya que no se basa en ninguna inversión económica, El uso de esta herramienta es totalmente gratuito y, a parte de Python, la plataforma se ha construido mediante la biblioteca Dash.

El objetivo de esta sección del proyecto es el proporcionar al usuario una guía exhaustiva y ordenada para que el ingreso en la plataforma sea lo más exitoso posible.

8.1 Requisitos previos

Antes de ejecutar la herramienta, el usuario tiene que asegurarse de cumplir los siguientes requisitos:

7. Tener instalado Python 3.8 o superior en su equipo. En caso de no tenerlo, puedes descargarlo desde *python.org*.
8. Instalar las librerías requeridas: este paso puede llevarse a cabo simplemente ejecutando el siguiente comando en la terminal para instalar las librerías necesarias:

```
pip install -r requirements.txt
```

9. Google Drive: el archivo CSV principal, del que se realiza el análisis, al ser de un gran tamaño para ser subido a GitHub directamente. Como paso primordial en el proceso hay que descargarlo desde el enlace proporcionado en la carpeta README.md:

Descargar CSV bruto

8.2 Descarga del código

Para la descarga de y preparación del código es trascendental seguir estos pasos:

13. Descarga el proyecto como archivo comprimido ZIP.

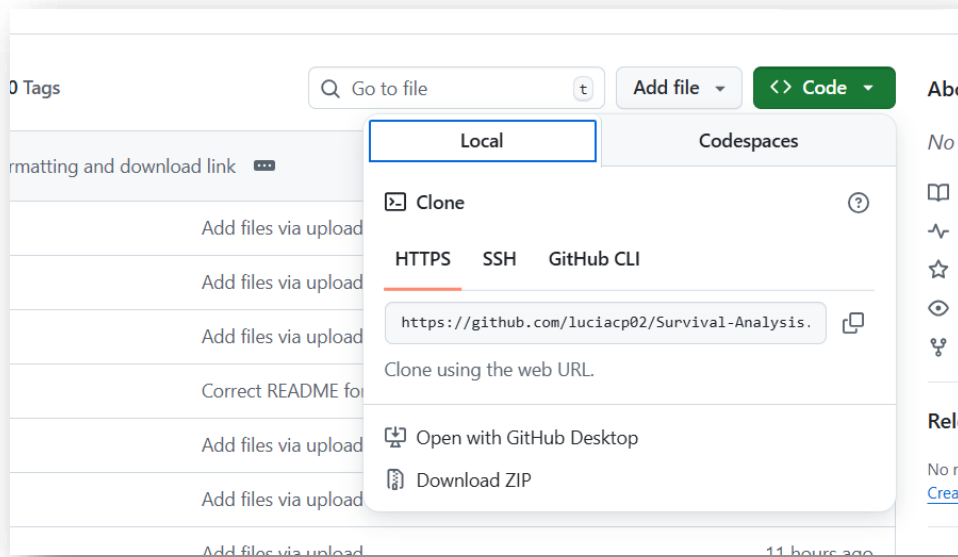


Ilustración 1: Descarga manual

14. En el caso de querer clonar el repositorio, el usuario tendrá que ejecutar en la terminal lo siguiente:

```
git clone https://github.com/luciap02/Survival-Analysis.git
```

15. **Descarga el archivo CSV:** Como se ha mencionado anteriormente. Este archivo está disponible en Google Drive y se puede acceder mediante README.md. El usuario puede hacer clic en el enlace de descarga que se proporciona en el manual o en el repositorio de GitHub.

16. Una vez instalado el archivo CSV, **sítualo en la carpeta principal**, donde se encuentran los archivos .py del proyecto para asegurar que el código funcione de manera correcta. El archivo CSV debe estar en el mismo directorio que el script *cargaDataset.py* para que la ejecución se pueda llevar a cabo con éxito.

17. **Ejecuta el código:** Una vez que hayas descargado el archivo CSV y colocado en la carpeta principal, ejecuta el script principal *cargaDataset.py* para iniciar el análisis:

```
python cargaDataset.py
```

18. Tras esperar unos segundos de la ejecución, o en el caso de error, volver a ejecutar,

el navegador abrirá la interfaz web del dashboard para comenzar con el análisis de los datos.

8.3 Instalación de las librerías de Python requeridas

7. En primer lugar, el usuario tendrá que asegurarse de tener instalado *pip*. Para ello, se tendrá que ejecutar en la terminal:

```
pip --version
```

En caso de estar instalado, debería aparecer la versión correspondiente, en caso de no tenerla instalado, puedes acceder a través de tu navegador a su instalación.

8. Seguidamente, hay que descargar las dependencias a través del archivo *requirements.txt*. Este archivo contiene todas las librerías necesarias y facilita su instalación de una sola vez ejecutando el comando en la carpeta del proyecto:

```
pip install -r requirements.txt
```

9. Una vez realizado este proceso, el usuario debe asegurarse de su instalación con la siguiente ejecución que mostrará las librerías descargadas:

```
pip list
```

8.4 Desinstalación del proyecto

Si el usuario desea de dejar de usar la herramienta, es posible proceder con su desinstalación de manera sencilla.

7. Eliminar repositorio.

Este proceso se puede realizar a mano, navegando a la carpeta del proyecto, seleccionándola y suprimiéndola.

8. Eliminar las librerías instaladas

Al haber instalado las librerías de forma global, sin entorno virtual, se puede realizar el proceso de desinstalación mediante *pip* con los siguientes comandos:

Esta ejecución listará las librerías:

```
pip freeze
```

Mediante estos comandos se desinstalarán todas ellas:

```
pip freeze | xargs pip uninstall -y
```

9. Eliminar Python

Como último paso opcional, el usuario puede desinstalar Python, si no desea mantenerlo en el sistema, de la siguiente manera:

- 9) Abrir el Panel de control
- 10) Hacer clic en “Programas”—“Desinstalar un programa”
- 11) Busca **Python** en la lista de programas instalados
- 12) Seleccionar en “Desinstalar”

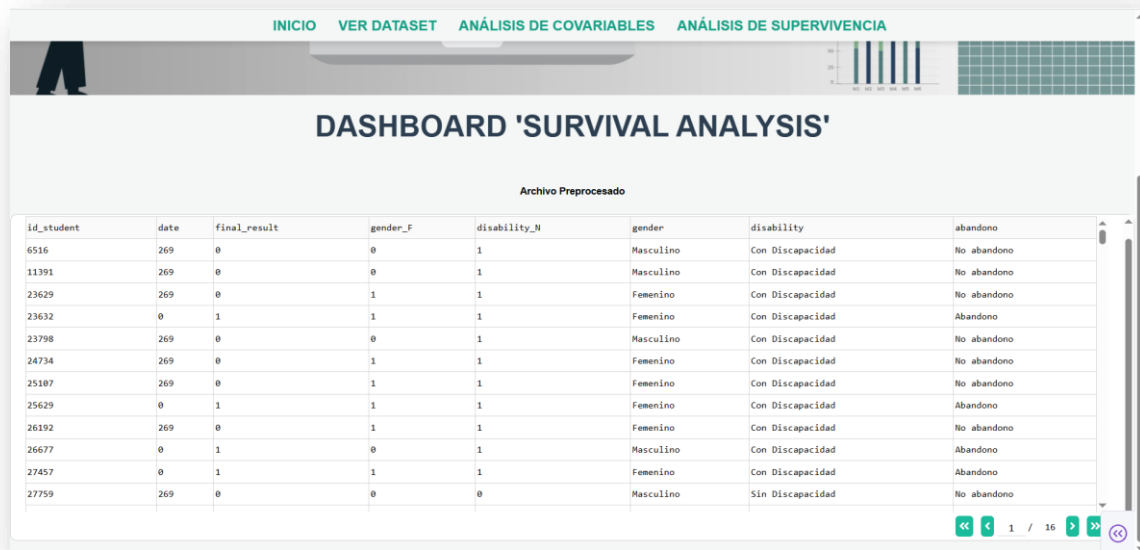
9. Interfaz

Como ya se ha mencionado y detallado antes, la interfaz ha sido diseñada con el objetivo de ser clara, intuitiva y fácil de utilizar, teniendo en cuenta que muchos usuarios pueden acceder sin conocimientos técnicos avanzados.

9.1 Home

El acceso a la plataforma se considera bastante sencillo, debido a que al ejecutar el código del archivo principal de la aplicación (cargaDataset.py) y acceder a la dirección IP que aparece, la máquina donde se esté realizando el proceso muestra por pantalla la página inicial del Dashboard, abriendo una nueva pestaña en el navegador de forma

instantánea.



id_student	date	final_result	gender_F	disability_N	gender	disability	abandono
6516	269	0	0	1	Masculino	Con Discapacidad	No abandono
11391	269	0	0	1	Masculino	Con Discapacidad	No abandono
23629	269	0	1	1	Femenino	Con Discapacidad	No abandono
23632	0	1	1	1	Femenino	Con Discapacidad	Abandono
23798	269	0	0	1	Masculino	Con Discapacidad	No abandono
24734	269	0	1	1	Femenino	Con Discapacidad	No abandono
25187	269	0	1	1	Femenino	Con Discapacidad	No abandono
25629	0	1	1	1	Femenino	Con Discapacidad	Abandono
26192	269	0	1	1	Femenino	Con Discapacidad	No abandono
26677	0	1	0	1	Masculino	Con Discapacidad	Abandono
27457	0	1	1	1	Femenino	Con Discapacidad	Abandono
27759	269	0	0	0	Masculino	Sin Discapacidad	No abandono

Ilustración 5: Dataset limpio cargado

En este momento, se aprecia en la parte superior cómo el usuario ya puede visualizar la barra de navegación y acceder a cualquiera de las secciones planteadas, seleccionando la pantalla que desee y con sus funcionalidades respectivas.

9.2 Inicio

Esta sección redirige al usuario a la pantalla inicial, reiniciando el dashboard y, por tanto, restablecer el proceso. Esto requiere la eliminación de los datos cargados y cualquier acción llevada a cabo desde el inicio.

Cuando el usuario selecciona esta opción, el sistema muestra un diálogo de confirmación, con el propósito de proteger los datos y resultados generados durante la sesión actual. Este mecanismo está diseñado para evitar la posible pérdida accidental de información.

- Para confirmar la acción y reiniciar la sesión, haga clic en “Aceptar”.
- Para cancelar y permanecer en la vista actual, haga clic en “Cancelar”.

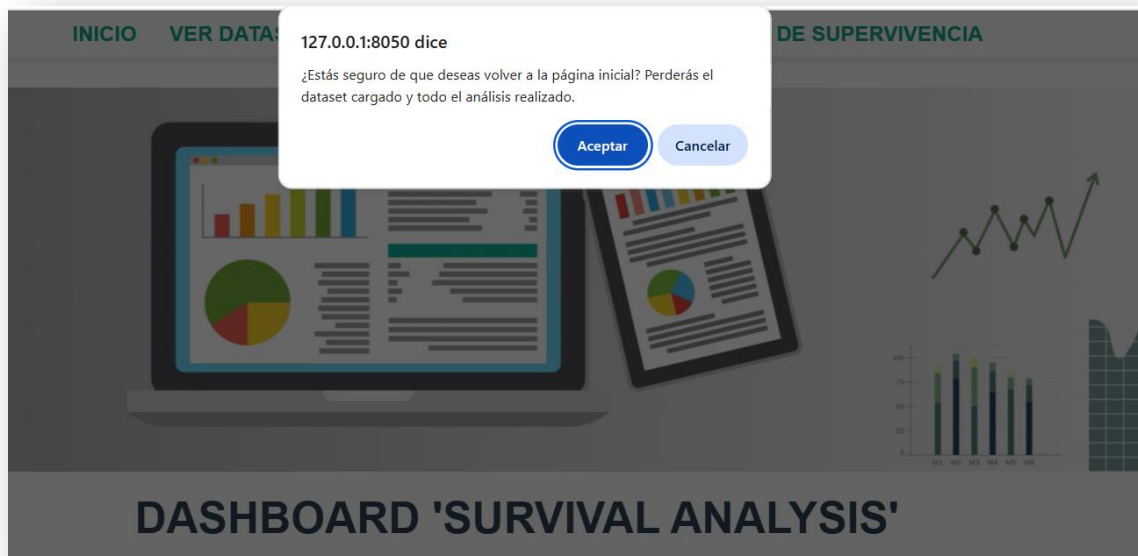


Ilustración 6: Diálogo de Confirmación

9.3 Ver Dataset

Esta sección del sistema se ofrece para que el usuario, en cualquier momento del análisis, pueda hacer clic en la opción “Ver Dataset” desde la barra de navegación superior para consultar el conjunto de datos limpio que se está utilizando en el análisis que se está llevando a cabo.

Esta funcionalidad permite revisar las variables y observaciones incluidas tras la fase de preprocesamiento.

Dataset Limpio							
id_student	date	final_result	gender_F	disability_N	abandono	gender	disability
6516	269	0	0	1	0	Masculino	Con Discapacidad
11391	269	0	0	1	0	Masculino	Con Discapacidad
23629	269	0	1	1	0	Femenino	Con Discapacidad
23632	0	1	1	1	0	Femenino	Con Discapacidad
23798	269	0	0	1	0	Masculino	Con Discapacidad
24734	269	0	1	1	0	Femenino	Con Discapacidad
25107	269	0	1	1	0	Femenino	Con Discapacidad
25629	0	1	1	1	0	Femenino	Con Discapacidad
26192	269	0	1	1	0	Femenino	Con Discapacidad
26677	0	1	0	1	0	Masculino	Con Discapacidad
27457	0	1	1	1	0	Femenino	Con Discapacidad
27759	269	0	0	0	0	Masculino	Sin Discapacidad

Ilustración 7: Sección Ver dataset

9.4 Análisis de Covariables

Esta sección se presenta como un previo análisis, donde el dashboard permite al usuario explorar de forma visual y descriptiva las covariables utilizadas en el análisis de supervivencia, con el objetivo de comprender su distribución y su posible relación con el evento de interés: el abandono universitario

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

ANEXO: MANUAL DE USUARIO

Análisis, Diseño e Implementación de un Dashboard
Interactivo para el Análisis de Supervivencia aplicado al
Abandono Escolar Universitario

Autor: Lucía Cabezuelo Pérez

Director: Dr. Cristóbal Romero Morales y Dr. Juan Alfonso Lara
Torralbo

Septiembre, 2025

ÍNDICE

1. Introducción	5
2. Instalación	6
2.1 Requisitos previos	7
2.2 Descarga del código	7
2.3 Instalación de las librerías de Python requeridas	9
2.4 Desinstalación del proyecto	9
3. Interfaz	10
3.1 Home	10
3.2 Inicio	13
3.3 Ver Dataset	14
3.4 Análisis de Covariables	15
3.5 Análisis de Supervivencia	17
3.5.1 Kaplan-Meier	19
3.5.2 Regresión de Cox	22
3.5.3 Pruebas de Log Rank	26
3.5.4 Prompt IA	30

Índice de Imágenes

Ilustración 1: Descarga manual	8
Ilustración 2: Pantalla de inicio	11
Ilustración 3: Mensaje de error	12
Ilustración 4: CSV bruto cargado	12
Ilustración 5: Dataset limpio cargado	13
Ilustración 6: Diálogo de Confirmación.....	14
Ilustración 7: Sección Ver dataset	15
Ilustración 8: Sección Análisis de Covariables	15
Ilustración 9: Gráfica variable evento	16
Ilustración 10: Análisis de covariables Género.....	17
Ilustración 11: Análisis de covariables Discapacidad	17
Ilustración 12: Sección Análisis de supervivencia.....	18
Ilustración 13: Modelos del análisis de supervivencia	19
Ilustración 14: Sección de Kaplan-Meier.....	19
Ilustración 15: Curva Supervivencia del Abandono	20
Ilustración 16: Curva de Kaplan para el Género	21
Ilustración 17: Curva de Kaplan para la Discapacidad	21
Ilustración 18: Botón "Ninguna"	22
Ilustración 19: Sección de Regresión de Cox.....	22
Ilustración 20: Interfaz de la Regresión de Cox	23
Ilustración 21: Barra de selección Cox.....	23
Ilustración 22: Opción covariable Género	24
Ilustración 23: Opción covariable Discapacidad.....	24
Ilustración 24: Opción covariables Género y Discapacidad	24
Ilustración 25: Sección Pruebas de Log Rank.....	26
Ilustración 26: Interfaz Pruebas de Log Rank	27

Ilustración 27: Opciones de selección de Log Rank	27
Ilustración 28: Opción covariable Género	28
Ilustración 29: Opción covariable Discapacidad.....	28
Ilustración 30: Opción covariables Género y Discapacidad	29
Ilustración 31: Prompt IA	30

Índice de Tablas

Tabla 1: Variables estudiadas Regresión de Cox26

Tabla 2: Variables estudiadas Pruebas de Log Rank.....30

10. Introducción

El propósito de este manual de usuario es proporcionar una guía detallada y accesible que facilite el uso del dashboard de análisis del abandono estudiantil en las universidades, explicando de manera clara y comprensible sus funciones principales. Esta herramienta se ha diseñado con el objetivo de proporcionar una visión integral y dinámica de los datos relacionados con la permanencia y el abandono de los estudiantes.

En el desarrollo de dicho documento, se presentarán las instrucciones que el usuario necesita para interactuar con la herramienta, abarcando desde el inicio hasta la ejecución de las diversas opciones disponibles. Se describen los procedimientos para navegar por las diferentes secciones, filtrar y visualizar los datos, interpretar los gráficos y obtener la información significativa

Se incluyen ciertas recomendaciones de uso y buenas prácticas para garantizar que la experiencia de los usuarios sea óptima y eficiente, aprovechando todo el potencial de la herramienta.

Con un enfoque sencillo y directo, este manual está pensado como un recurso complementario tanto para los usuarios de nuevo ingreso o sin experiencia en el uso de dashboard o herramientas analíticas, como para los que usuarios con mayor experiencia o más avanzadas que deseen profundizar en las funcionalidades disponibles, brindándoles la confianza y la guía necesaria para utilizar la herramienta con éxito.

11. Instalación

Para el correcto uso por parte del usuario de la aplicación desarrollada en este proyecto, es fundamental seguir una serie de pasos de instalación que garantizan su funcionamiento adecuado en tu equipo local. La herramienta está basada en el lenguaje de Python, y hace uso de diversas bibliotecas que facilitan el procesamiento de datos y su visualización interactiva.

En este proyecto no ha hecho falta ningún coste, ya que no se basa en ninguna inversión económica, El uso de esta herramienta es totalmente gratuito y, a parte de Python, la plataforma se ha construido mediante la biblioteca Dash.

El objetivo de esta sección del proyecto es el proporcionar al usuario una guía exhaustiva y ordenada para que el ingreso en la plataforma sea lo más exitoso posible.

11.1 Requisitos previos

Antes de ejecutar la herramienta, el usuario tiene que asegurarse de cumplir los siguientes requisitos:

10. Tener instalado Python 3.8 o superior en su equipo. En caso de no tenerlo, puedes descargarlo desde *python.org*.
11. Instalar las librerías requeridas: este paso puede llevarse a cabo simplemente ejecutando el siguiente comando en la terminal para instalar las librerías necesarias:

```
pip install -r requirements.txt
```

12. Google Drive: el archivo CSV principal, del que se realiza el análisis, al ser de un gran tamaño para ser subido a GitHub directamente. Como paso primordial en el proceso hay que descargarlo desde el enlace proporcionado en la carpeta README.md:

Descargar CSV bruto

11.2 Descarga del código

Para la descarga de y preparación del código es trascendental seguir estos pasos:

19. Descarga el proyecto como archivo comprimido ZIP.

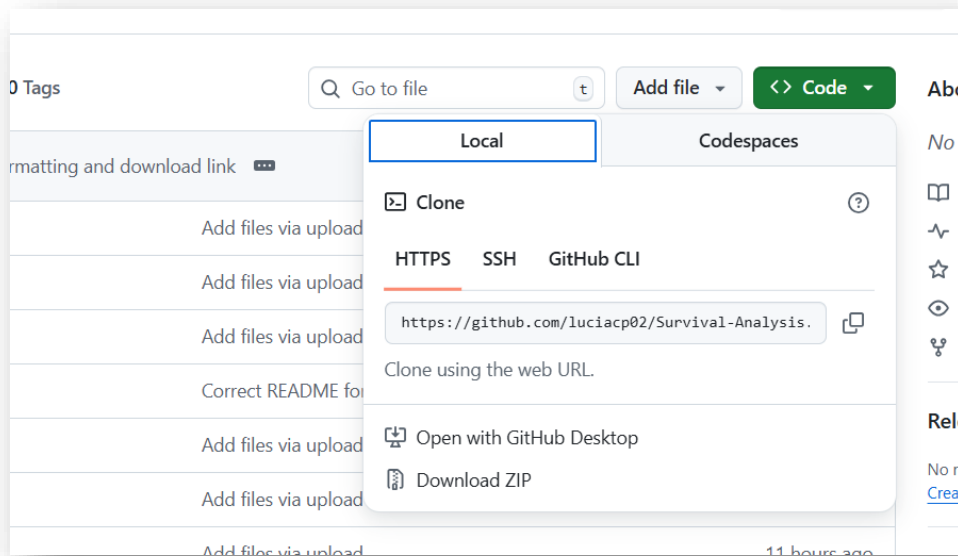


Ilustración 1: Descarga manual

20. En el caso de querer clonar el repositorio, el usuario tendrá que ejecutar en la terminal lo siguiente:

```
git clone https://github.com/luciap02/Survival-Analysis.git
```

21. **Descarga el archivo CSV:** Como se ha mencionado anteriormente. Este archivo está disponible en Google Drive y se puede acceder mediante README.md. El usuario puede hacer clic en el enlace de descarga que se proporciona en el manual o en el repositorio de GitHub.

22. Una vez instalado el archivo CSV, **sítualo en la carpeta principal**, donde se encuentran los archivos .py del proyecto para asegurar que el código funcione de manera correcta. El archivo CSV debe estar en el mismo directorio que el script *cargaDataset.py* para que la ejecución se pueda llevar a cabo con éxito.

23. **Ejecuta el código:** Una vez que hayas descargado el archivo CSV y colocado en la carpeta principal, ejecuta el script principal *cargaDataset.py* para iniciar el análisis:

```
python cargaDataset.py
```

24. Tras esperar unos segundos de la ejecución, o en el caso de error, volver a ejecutar, el navegador abrirá la interfaz web del dashboard para comenzar con el análisis de los datos.

11.3 Instalación de las librerías de Python requeridas

10. En primer lugar, el usuario tendrá que asegurarse de tener instalado *pip*. Para ello, se tendrá que ejecutar en la terminal:

```
pip --version
```

En caso de estar instalado, debería aparecer la versión correspondiente, en caso de no tenerla instalado, puedes acceder a través de tu navegador a su instalación.

11. Seguidamente, hay que descargar las dependencias a través del archivo *requirements.txt*. Este archivo contiene todas las librerías necesarias y facilita su instalación de una sola vez ejecutando el comando en la carpeta del proyecto:

```
pip install -r requirements.txt
```

12. Una vez realizado este proceso, el usuario debe asegurarse de su instalación con la siguiente ejecución que mostrará las librerías descargadas:

```
pip list
```

11.4 Desinstalación del proyecto

Si el usuario desea de dejar de usar la herramienta, es posible proceder con su desinstalación de manera sencilla.

10. Eliminar repositorio.

Este proceso se puede realizar a mano, navegando a la carpeta del proyecto, seleccionándola y suprimiéndola.

11. Eliminar las librerías instaladas

Al haber instalado las librerías de forma global, sin entorno virtual, se puede realizar el

proceso de desinstalación mediante *pip* con los siguientes comandos:

Esta ejecución listará las librerías:

```
pip freeze
```

Mediante estos comandos se desinstalarán todas ellas:

```
pip freeze | xargs pip uninstall -y
```

12. Eliminar Python

Como último paso opcional, el usuario puede desinstalar Python, si no desea mantenerlo en el sistema, de la siguiente manera:

- 13) Abrir el Panel de control
- 14) Hacer clic en “Programas”—“Desinstalar un programa”
- 15) Busca **Python** en la lista de programas instalados
- 16) Seleccionar en “Desinstalar”

12. Interfaz

Como ya se ha mencionado y detallado antes, la interfaz ha sido diseñada con el objetivo de ser clara, intuitiva y fácil de utilizar, teniendo en cuenta que muchos usuarios pueden acceder sin conocimientos técnicos avanzados.

12.1 Home

El acceso a la plataforma se considera bastante sencillo, debido a que al ejecutar el código del archivo principal de la aplicación (*cargaDataset.py*) y acceder a la dirección IP que aparece, la máquina donde se esté realizando el proceso muestra por pantalla

la página inicial del Dashboard, abriendo una nueva pestaña en el navegador de forma instantánea.

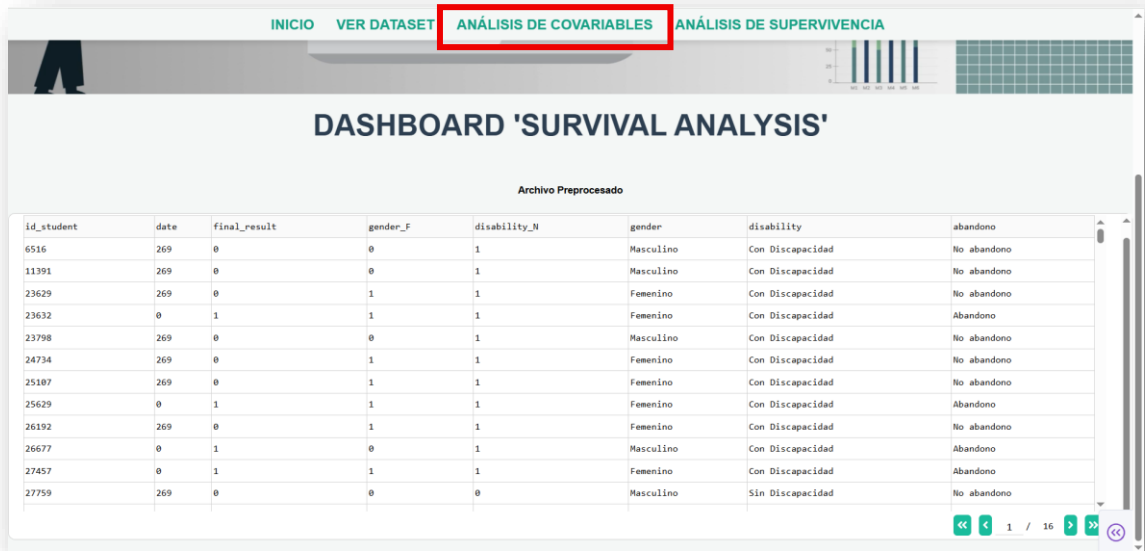


Ilustración 8: Sección Análisis de Covariables

En dicha vista se muestran gráficas de barras simples que muestran el recuento absoluto de estudiantes según cada categoría de las variables seleccionadas. Cada gráfico está diseñado para facilitar la interpretación rápida, utilizando diferentes colores para representar el estado del evento y mostrando los valores numéricos directamente sobre las barras. Estas gráficas se ven acompañadas de una breve explicación en el lateral, para ayudar a interpretar los resultados mostrados y entender las diferencias notables en determinados recuentos.

En el momento que el usuario accede a esta página mediante la barra de navegación, en el dashboard se muestra por defecto la gráfica principal, la cual expone la variable evento, Abandono frente al recuento del No abandono.

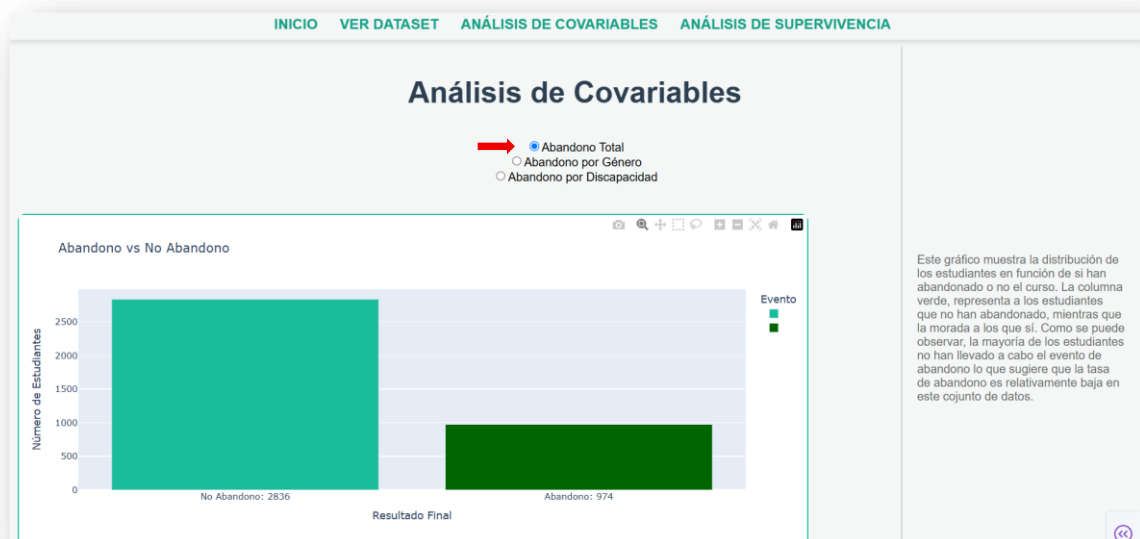


Ilustración 9: Gráfica variable evento

La vista ofrece tres opciones, mediante una selección interactiva que muestra cada gráfica correspondiente junto con su respectiva explicación en el lateral.

- **Abandono total:**

Muestra la distribución general entre estudiantes que completaron vs los que realizaron el evento y abandonaron. Esta gráfica es la que se aprecia en la imagen anterior ya que es la que el sistema muestra, primeramente.

- **Abandono por Género:**

Realiza una comparación de la tasa de abandono entre hombres y mujeres, frente al total de abandono/No abandono en cada caso, identificando posibles diferencias de género.

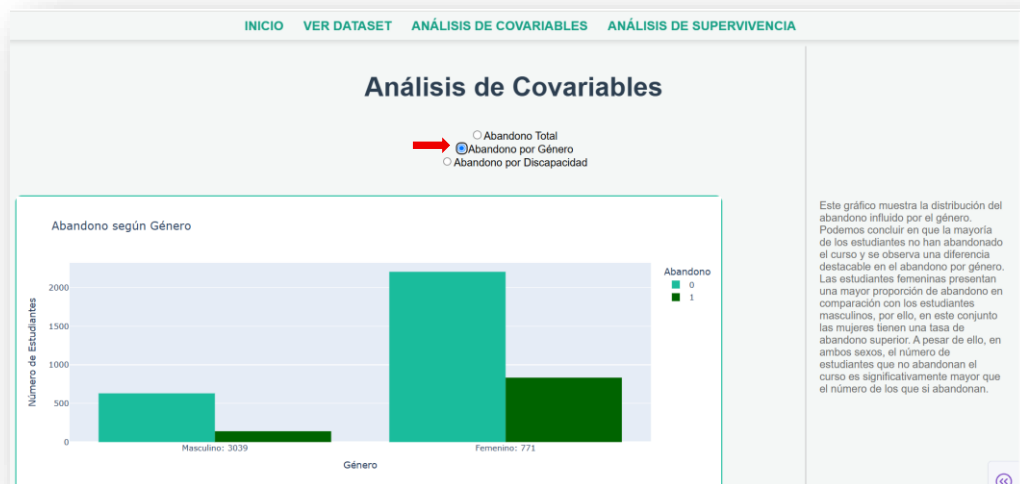


Ilustración 10: Análisis de covariables Género

- **Abandono por Discapacidad:**

Se realiza, de igual manera, un análisis que confronta la tasa Con discapacidad/Sin discapacidad, frente a la variable evento, estudiando las posibles diferencias más destacadas.

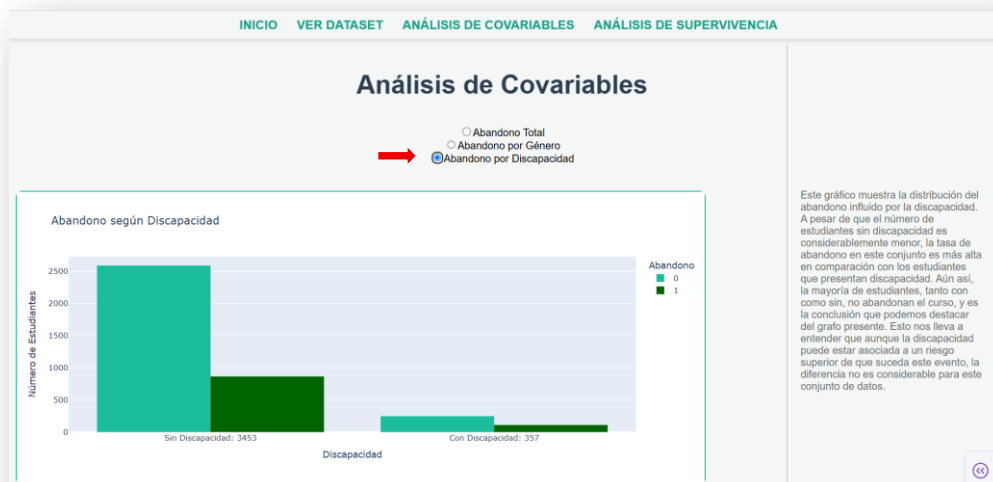


Ilustración 11: Análisis de covariables Discapacidad

12.2 Análisis de Supervivencia

Esta sección representa el corazón del análisis estadístico del dashboard. En esta parte del sistema se aplican los modelos estadísticos avanzados sobre el dataset previamente preparado y estructurado, con el principal objetivo de estimar la probabilidad de que un estudiante complete sus estudios en un tiempo determinado, o bien, el riesgo de

abandono en función de diversas covariables.

TRABAJO FIN DE GRADO

Grado en Ingeniería Informática

ANEXO: MANUAL DE USUARIO

Análisis, Diseño e Implementación de un Dashboard
Interactivo para el Análisis de Supervivencia aplicado al
Abandono Escolar Universitario

Autor: Lucía Cabezuelo Pérez

Director: Dr. Cristóbal Romero Morales y Dr. Juan Alfonso Lara
Torralbo

Septiembre, 2025

ÍNDICE

1. Introducción	5
2. Instalación	6
2.1 Requisitos previos	7
2.2 Descarga del código	7
2.3 Instalación de las librerías de Python requeridas	9
2.4 Desinstalación del proyecto	9
3. Interfaz	10
3.1 Home	10
3.2 Inicio	13
3.3 Ver Dataset	14
3.4 Análisis de Covariables	15
3.5 Análisis de Supervivencia	17
3.5.1 Kaplan-Meier	19
3.5.2 Regresión de Cox	22
3.5.3 Pruebas de Log Rank	26
3.5.4 Prompt IA	30

Índice de Imágenes

Ilustración 1: Descarga manual	8
--------------------------------	---

Ilustración 2: Pantalla de inicio	11
Ilustración 3: Mensaje de error	12
Ilustración 4: CSV bruto cargado	12
Ilustración 5: Dataset limpio cargado	13
Ilustración 6: Diálogo de Confirmación.....	14
Ilustración 7: Sección Ver dataset	15
Ilustración 8: Sección Análisis de Covariables	15
Ilustración 9: Gráfica variable evento	16
Ilustración 10: Análisis de covariables Género.....	17
Ilustración 11: Análisis de covariables Discapacidad	17
Ilustración 12: Sección Análisis de supervivencia.....	18
Ilustración 13: Modelos del análisis de supervivencia	19
Ilustración 14: Sección de Kaplan-Meier.....	19
Ilustración 15: Curva Supervivencia del Abandono	20
Ilustración 16: Curva de Kaplan para el Género	21
Ilustración 17: Curva de Kaplan para la Discapacidad	21
Ilustración 18: Botón "Ninguna"	22
Ilustración 19: Sección de Regresión de Cox.....	22
Ilustración 20: Interfaz de la Regresión de Cox	23
Ilustración 21: Barra de selección Cox.....	23
Ilustración 22: Opción covariable Género	24
Ilustración 23: Opción covariable Discapacidad.....	24
Ilustración 24: Opción covariables Género y Discapacidad	24
Ilustración 25: Sección Pruebas de Log Rank.....	26
Ilustración 26: Interfaz Pruebas de Log Rank	27
Ilustración 27: Opciones de selección de Log Rank	27
Ilustración 28: Opción covariable Género	28
Ilustración 29: Opción covariable Discapacidad.....	28

Ilustración 30: Opción covariables Género y Discapacidad29

Ilustración 31: Prompt IA30

Índice de Tablas

Tabla 1: Variables estudiadas Regresión de Cox26

Tabla 2: Variables estudiadas Pruebas de Log Rank.....30

13. Introducción

El propósito de este manual de usuario es proporcionar una guía detallada y accesible que facilite el uso del dashboard de análisis del abandono estudiantil en las

universidades, explicando de manera clara y comprensible sus funciones principales. Esta herramienta se ha diseñado con el objetivo de proporcionar una visión integral y dinámica de los datos relacionados con la permanencia y el abandono de los estudiantes.

En el desarrollo de dicho documento, se presentarán las instrucciones que el usuario necesita para interactuar con la herramienta, abarcando desde el inicio hasta la ejecución de las diversas opciones disponibles. Se describen los procedimientos para navegar por las diferentes secciones, filtrar y visualizar los datos, interpretar los gráficos y obtener la información significativa

Se incluyen ciertas recomendaciones de uso y buenas prácticas para garantizar que la experiencia de los usuarios sea óptima y eficiente, aprovechando todo el potencial de la herramienta.

Con un enfoque sencillo y directo, este manual está pensado como un recurso complementario tanto para los usuarios de nuevo ingreso o sin experiencia en el uso de dashboard o herramientas analíticas, como para los que usuarios con mayor experiencia o más avanzadas que deseen profundizar en las funcionalidades disponibles, brindándoles la confianza y la guía necesaria para utilizar la herramienta con éxito.

14. Instalación

Para el correcto uso por parte del usuario de la aplicación desarrollada en este proyecto, es fundamental seguir una serie de pasos de instalación que garantizan su

funcionamiento adecuado en tu equipo local. La herramienta está basada en el lenguaje de Python, y hace uso de diversas bibliotecas que facilitan el procesamiento de datos y su visualización interactiva.

En este proyecto no ha hecho falta ningún coste, ya que no se basa en ninguna inversión económica, El uso de esta herramienta es totalmente gratuito y, a parte de Python, la plataforma se ha construido mediante la biblioteca Dash.

El objetivo de esta sección del proyecto es el proporcionar al usuario una guía exhaustiva y ordenada para que el ingreso en la plataforma sea lo más exitoso posible.

14.1 Requisitos previos

Antes de ejecutar la herramienta, el usuario tiene que asegurarse de cumplir los siguientes requisitos:

13. Tener instalado Python 3.8 o superior en su equipo. En caso de no tenerlo, puedes descargarlo desde *python.org*.
14. Instalar las librerías requeridas: este paso puede llevarse a cabo simplemente ejecutando el siguiente comando en la terminal para instalar las librerías necesarias:

```
pip install -r requirements.txt
```

15. Google Drive: el archivo CSV principal, del que se realiza el análisis, al ser de un gran tamaño para ser subido a GitHub directamente. Como paso primordial en el proceso hay que descargarlo desde el enlace proporcionado en la carpeta README.md:

Descargar CSV bruto

14.2 Descarga del código

Para la descarga de y preparación del código es trascendental seguir estos pasos:

25. Descarga el proyecto como archivo comprimido ZIP.

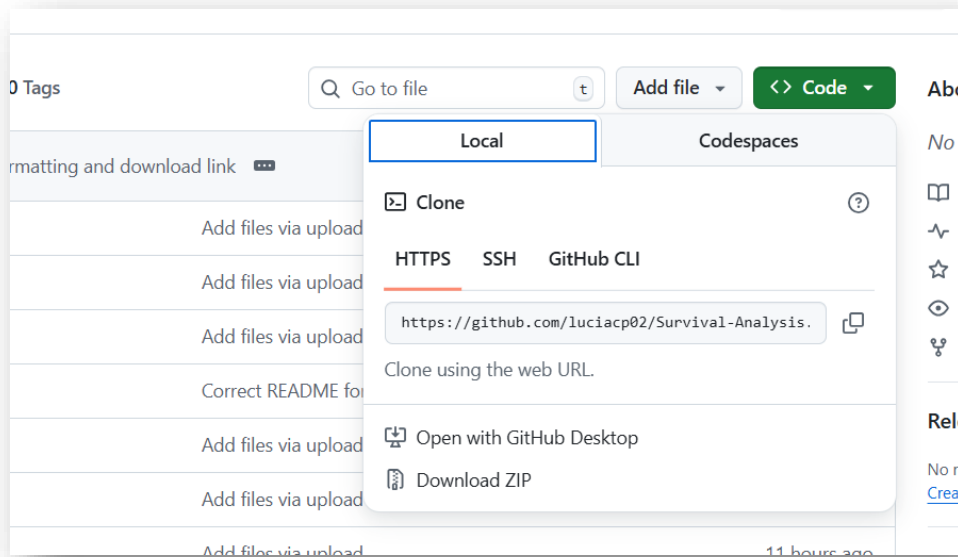


Ilustración 1: Descarga manual

26. En el caso de querer clonar el repositorio, el usuario tendrá que ejecutar en la terminal lo siguiente:

```
git clone https://github.com/luciap02/Survival-Analysis.git
```

27. **Descarga el archivo CSV:** Como se ha mencionado anteriormente. Este archivo está disponible en Google Drive y se puede acceder mediante README.md. El usuario puede hacer clic en el enlace de descarga que se proporciona en el manual o en el repositorio de GitHub.

28. Una vez instalado el archivo CSV, **sítualo en la carpeta principal**, donde se encuentran los archivos .py del proyecto para asegurar que el código funcione de manera correcta. El archivo CSV debe estar en el mismo directorio que el script *cargaDataset.py* para que la ejecución se pueda llevar a cabo con éxito.

29. **Ejecuta el código:** Una vez que hayas descargado el archivo CSV y colocado en la carpeta principal, ejecuta el script principal *cargaDataset.py* para iniciar el análisis:

```
python cargaDataset.py
```

30. Tras esperar unos segundos de la ejecución, o en el caso de error, volver a ejecutar,

el navegador abrirá la interfaz web del dashboard para comenzar con el análisis de los datos.

14.3 Instalación de las librerías de Python requeridas

13. En primer lugar, el usuario tendrá que asegurarse de tener instalado *pip*. Para ello, se tendrá que ejecutar en la terminal:

```
pip --version
```

En caso de estar instalado, debería aparecer la versión correspondiente, en caso de no tenerla instalado, puedes acceder a través de tu navegador a su instalación.

14. Seguidamente, hay que descargar las dependencias a través del archivo *requirements.txt*. Este archivo contiene todas las librerías necesarias y facilita su instalación de una sola vez ejecutando el comando en la carpeta del proyecto:

```
pip install -r requirements.txt
```

15. Una vez realizado este proceso, el usuario debe asegurarse de su instalación con la siguiente ejecución que mostrará las librerías descargadas:

```
pip list
```

14.4 Desinstalación del proyecto

Si el usuario desea de dejar de usar la herramienta, es posible proceder con su desinstalación de manera sencilla.

13. Eliminar repositorio.

Este proceso se puede realizar a mano, navegando a la carpeta del proyecto, seleccionándola y suprimiéndola.

14. Eliminar las librerías instaladas

Al haber instalado las librerías de forma global, sin entorno virtual, se puede realizar el proceso de desinstalación mediante *pip* con los siguientes comandos:

Esta ejecución listará las librerías:

```
pip freeze
```

Mediante estos comandos se desinstalarán todas ellas:

```
pip freeze | xargs pip uninstall -y
```

15. Eliminar Python

Como último paso opcional, el usuario puede desinstalar Python, si no desea mantenerlo en el sistema, de la siguiente manera:

- 17) Abrir el Panel de control
- 18) Hacer clic en “Programas”—“Desinstalar un programa”
- 19) Busca **Python** en la lista de programas instalados
- 20) Seleccionar en “Desinstalar”

15. Interfaz

Como ya se ha mencionado y detallado antes, la interfaz ha sido diseñada con el objetivo de ser clara, intuitiva y fácil de utilizar, teniendo en cuenta que muchos usuarios pueden acceder sin conocimientos técnicos avanzados.

15.1 Home

El acceso a la plataforma se considera bastante sencillo, debido a que al ejecutar el código del archivo principal de la aplicación (cargaDataset.py) y acceder a la dirección IP que aparece, la máquina donde se esté realizando el proceso muestra por pantalla la página inicial del Dashboard, abriendo una nueva pestaña en el navegador de forma

instantánea.

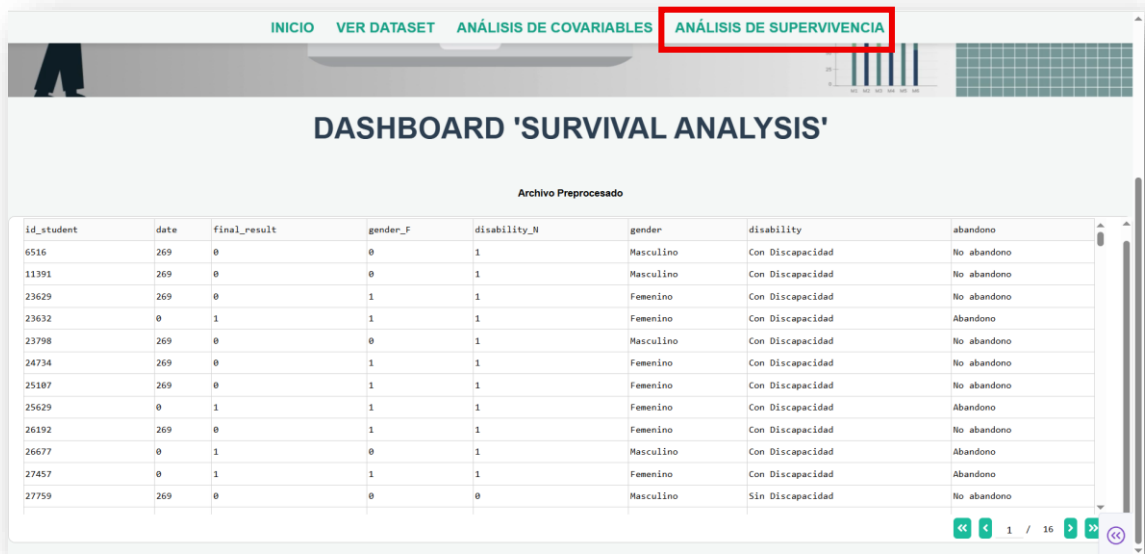


Ilustración 12: Sección Análisis de supervivencia

Al ingresar el usuario en esta sección, se muestran los tres modelos llevados a cabo en el análisis, los cuales proporcionan estudios enfocados en diferentes aspectos para cada una de las covariables y propuestos los resultados visualmente con gráficas dinámicas/ tablas de valores.

Todo lo que se presenta en este apartado, es generado directamente del dataset preprocesado, lo que garantiza coherencia, transparencia y trazabilidad en el proceso analítico.



Ilustración 13: Modelos del análisis de supervivencia

15.1.1 Kaplan-Meier

Error! Reference source not found.

En este primer modelo analítico disponible en la sección del Análisis de supervivencia, se estima la probabilidad de supervivencia lo largo del tiempo, sin asumir ninguna distribución específica para los datos.

El usuario principalmente verá la curva general de Kaplan-Meier sobre la variable evento, al comienzo de la página a la que se le deriva.



Ilustración 14: Sección de Kaplan-Meier

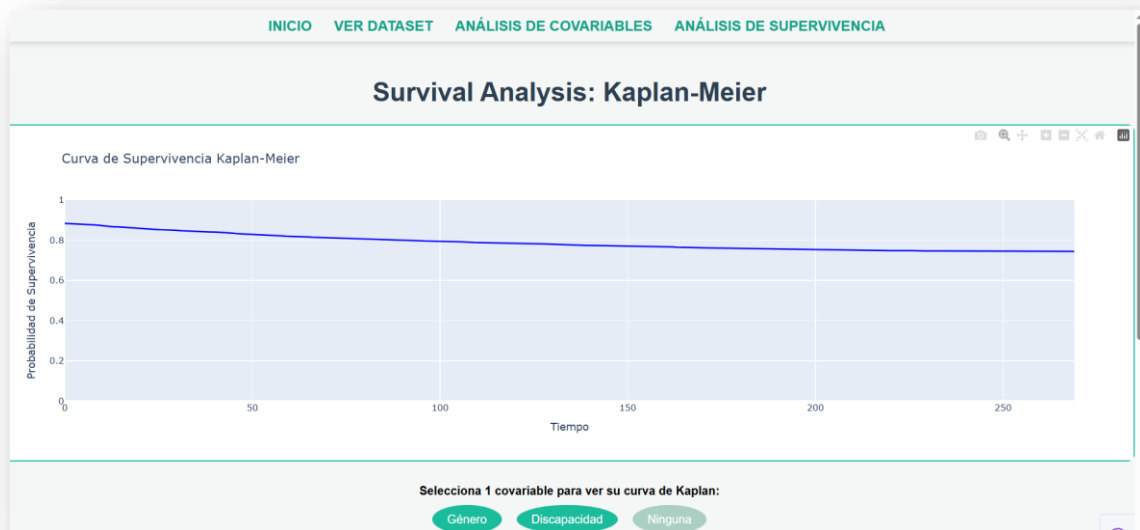


Ilustración 15: Curva Supervivencia del Abandono

Para profundizar en el análisis, seguida de esta gráfica, se aprecian los botones con los que el usuario filtrará para ilustrar cada curva específica para cada covariable del conjunto de datos. En el momento que el usuario haga clic en alguno de ellos, el dashboard actualizará dinámicamente para mostrar las curvas de Kaplan-Meier separadas por grupo.

Estas gráficas han sido implementadas de forma dinámica, lo que permite un análisis más detallado en cada instante de tiempo para cada uno de los datos estudiados.

- **Curva para la variable Género**

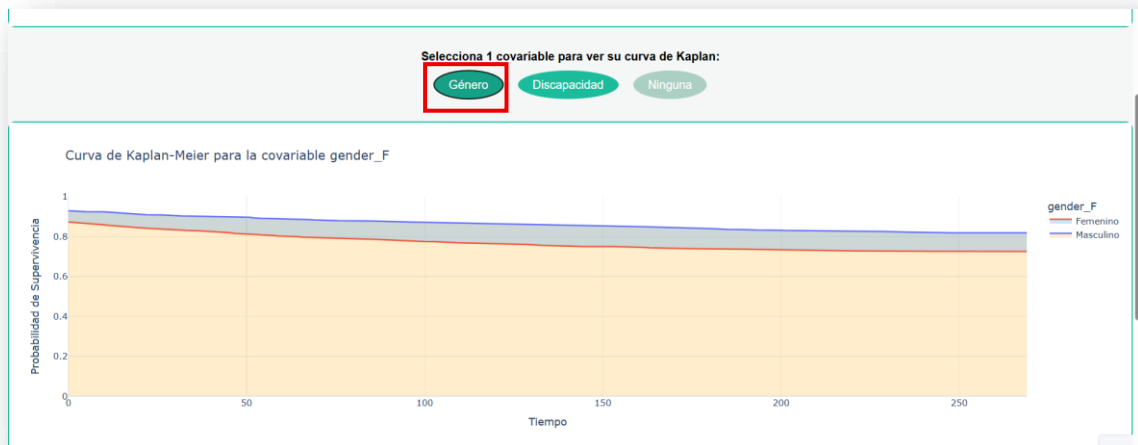


Ilustración 16: Curva de Kaplan para el Género

- **Curva para la variable Discapacidad:**



Ilustración 17: Curva de Kaplan para la Discapacidad

Entre las opciones proporcionadas en el proceso de filtro, el usuario cuenta con el botón “Ninguna”, el cual deja la pantalla como se muestra por defecto, en blanco y visualizando únicamente la gráfica general de la variable evento.

Selecciona 1 covariable para ver su curva de Kaplan:

Género

Discapacidad

Ninguna

Ilustración 18: Botón "Ninguna"

15.1.2 Regresión de Cox

Error! Reference source not found. DE COX

En este segundo apartado, se proporciona otro modelo analítico en la pantalla de análisis de supervivencia.

Mediante la regresión de Cox, el sistema permite analizar el tiempo hasta que ocurre un evento de interés, considerando uno o más factores que podrían influir en ese tiempo. Este modelo permite estimar cómo afectan diferentes variables al riesgo de que ocurra el evento, manteniendo otras variables constantes.

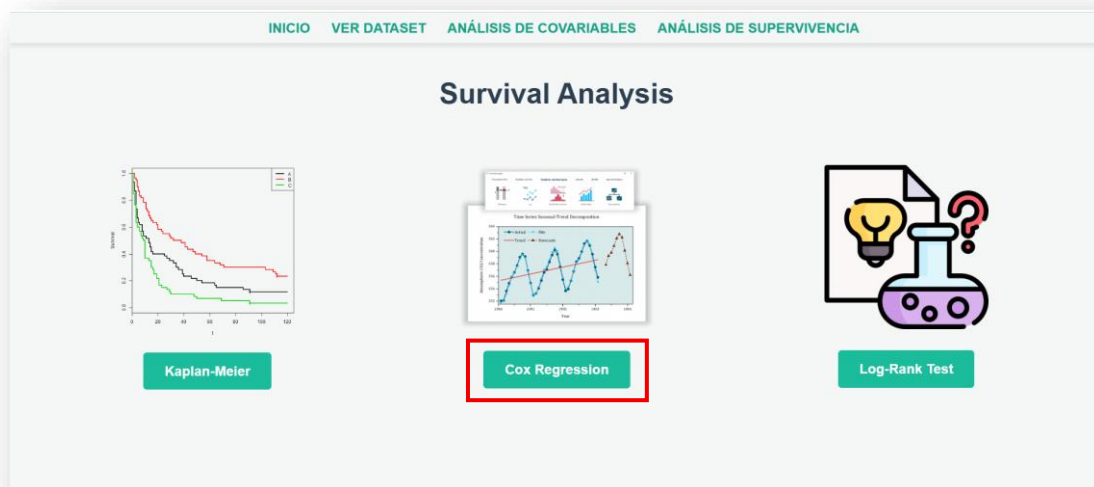


Ilustración 19: Sección de Regresión de Cox





Ilustración 20: Interfaz de la Regresión de Cox

Los resultados de este proceso van a estar representados en forma de tabla, dependiendo a las selecciones que lleve a cabo el usuario.



Ilustración 21: Barra de selección Cox

En este modelo, las variables estudiadas son esenciales para analizar la relación entre ciertas características y el tiempo hasta la ocurrencia del abandono.

Las posibles tablas de datos resultantes pueden ser:

1. Covariable Género

Se representa mediante una variable binaria que distingue entre Masculino y Femenino.

INICIO VER DATASET ANÁLISIS DE COVARIABLES ANÁLISIS DE SUPERVIVENCIA											
Survival Analysis: Regresión de Cox											
Género											
Covariable	Coef.	exp(Coef.)	SE(Coef.)	Coef. lower 95%	Coef. upper 95%	exp(Coef.) lower 95%	exp(Coef.) upper 95%	cmp to	z	p	-log2(p)
gender_F	0.4841010407131901	1.6227155977105028	0.09134348962461478	0.305071090826737	0.6631309905996432	1.3567214500274685	1.940859644482056	0	5.299787020428624	1.1593784358860951e-7	23.040145106960264

Ilustración 22:Opción covariable Género

2. Covariable Discapacidad

Se representa mediante una variable binaria que distingue entre Con Discapacidad y Sin Discapacidad.

INICIO VER DATASET ANÁLISIS DE COVARIABLES ANÁLISIS DE SUPERVIVENCIA											
Survival Analysis: Regresión de Cox											
Discapacidad											
Covariable	Coef.	exp(Coef.)	SE(Coef.)	Coef. lower 95%	Coef. upper 95%	exp(Coef.) lower 95%	exp(Coef.) upper 95%	cmp to	z	p	-log2(p)
disability_N	-0.22838028202951907	0.7958215656444027	0.10123909644600408	-0.4268052648910641	-0.029955299167974098	0.6525906221609074	0.9704889142408629	0	-2.255850655001902	0.024079985731566868	5.37602165253

Ilustración 23: Opción covariable Discapacidad

3. Covariables Género y Discapacidad

INICIO VER DATASET ANÁLISIS DE COVARIABLES ANÁLISIS DE SUPERVIVENCIA											
Género Discapacidad											
Covariable	Coef.	exp(Coef.)	SE(Coef.)	Coef. lower 95%	Coef. upper 95%	exp(Coef.) lower 95%	exp(Coef.) upper 95%	cmp to	z	p	-log2(p)
gender_F	0.47779783959184857	1.6125194628377506	0.09140571050601859	0.2986459390187577	0.6569497401649395	1.3480322533464637	1.9288997066468958	0	5.227220891854323	1.7207687606507652e-7	22.47044342
disability_N	-0.20511199000879732	0.8145560891596486	0.10130759357054939	-0.4036712247674956	-0.0065527552500989905	0.6678636622079632	0.9934686672329484	0	-2.0246457622740768	0.04290375164586685	4.542752382

Explicar

Ilustración 24: Opción covariables Género y Discapacidad

Según las tablas propuestas obtenemos las variables estudiadas que nos ayudan a

interpretar y analizar el impacto de las covariables. Considerando que cada una de ellas le aportan al usuario una información determinada, tenemos:

Covariable	Variables independientes que se están analizando en el modelo de regresión de Cox
Coef. (Coeficiente)	Parámetro que representa la magnitud y dirección del efecto de la covariable sobre el evento. Un coeficiente positivo indica un aumento en el riesgo de que ocurra el evento, mientras que un coeficiente negativo indica una reducción del riesgo
Exp. (Coef)	Transformación exponencial del coeficiente, lo cual es importante para la interpretación. Su valor muestra la razón de riesgos (hazard ratio) y se interpreta como la multiplicación del riesgo del evento por cada unidad de aumento en la covariable.
SE (Coef)	Error estándar del coeficiente, siendo una medida de precisión del coeficiente estimado. Mide la variabilidad del coeficiente en diferentes muestras de la población
Coef. lower 95%	Indica el límite inferior del intervalo de confianza del 95% para el Coef. Si el intervalo de confianza no incluye el valor 0, esto sugiere que la covariable tiene un efecto significativo.
Coef. upper 95%	Similar al límite inferior, pero muestra el límite superior del intervalo de confianza del 95% para el coeficiente estimado.
exp(Coef.) lower 95%	Muestra el límite inferior del intervalo de confianza para la razón de riesgos, indicando el rango en el cual la razón de riesgos es estadísticamente confiable.
exp(Coef.) upper 95%	Similar al límite inferior, muestra el límite superior del intervalo de confianza, ayudando a interpretar la precisión de la razón de riesgos.
Cmp to	Valor con el que se compara la covariable
z	Valor estadístico que se utiliza para probar la hipótesis nula de que el coeficiente es igual a cero.
p	Valor crucial para determinar la significancia estadística del coeficiente.

$-\log_2(p)$	Transformación del valor p usando el logaritmo base 2, utilizado para mostrar la significancia de los coeficientes de manera más resaltada.
--------------------------------	---

Tabla 1: Variables estudiadas Regresión de Cox

3.5.3 Pruebas de Log Rank

En este apartado del sistema, se ofrece una prueba estadística que permite comparar las curvas de supervivencia de dos grupos definidos por una covariable. Este análisis ofrece al usuario una visión, en la que determina si existen diferencias significativas en el tiempo hasta que ocurre un evento entre los dos grupos.

Error! Reference source not found.

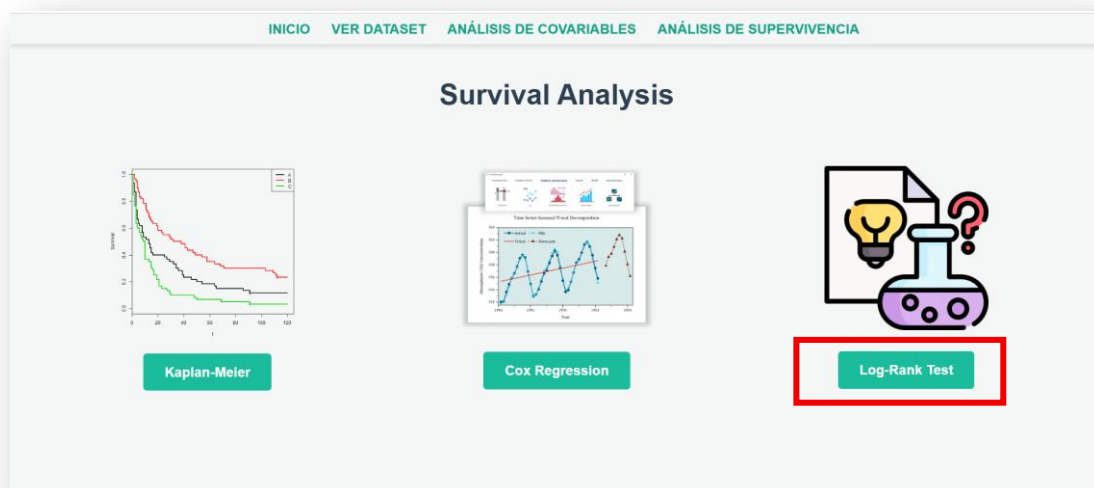


Ilustración 25: Sección Pruebas de Log Rank



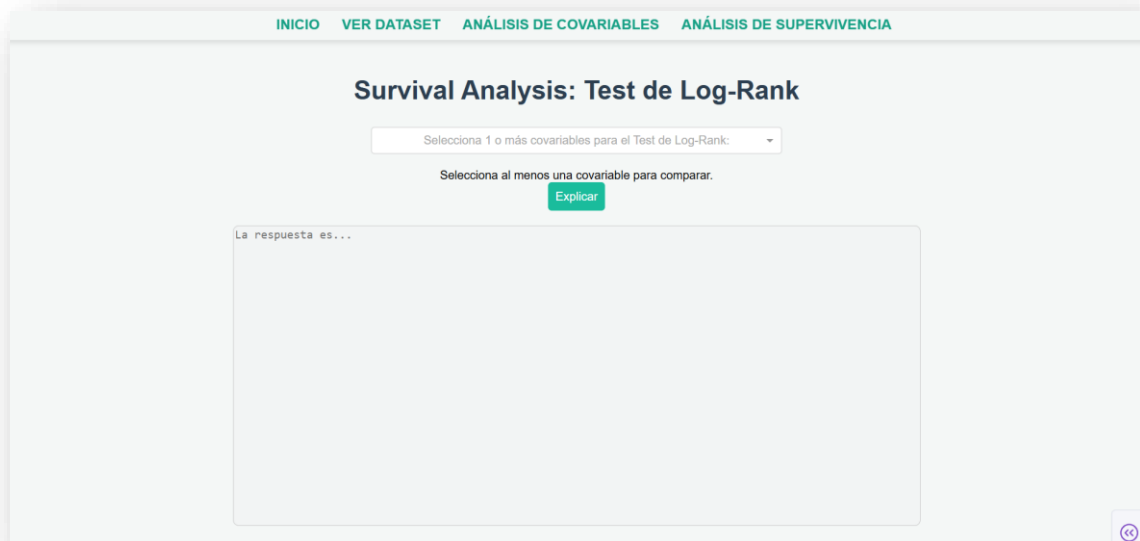


Ilustración 26: Interfaz Pruebas de Log Rank

Los resultados de este proceso van a estar representados en forma de tabla de igual manera que en la regresión de Cox, pero estudiando en base a otras variables estadísticas y obteniendo otros resultados y conclusiones, todo ello dependiendo a las selecciones que lleve a cabo el usuario.

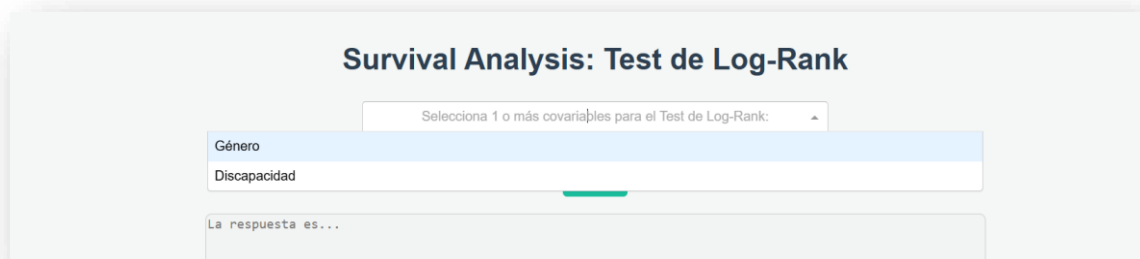


Ilustración 27: Opciones de selección de Log Rank

De forma similar al proceso de la regresión de Cox, las posibles tablas de datos generadas pueden ser:

1. Covariable Género

Análisis de una variable binaria que distingue entre Masculino y Femenino.



Ilustración 28: Opción covariable Género

2. Covariable Discapacidad

Se analiza la variable binaria que distingue entre el estudiante Con Discapacidad y Sin Discapacidad.



Ilustración 29: Opción covariable Discapacidad

3. Covariables Género y Discapacidad

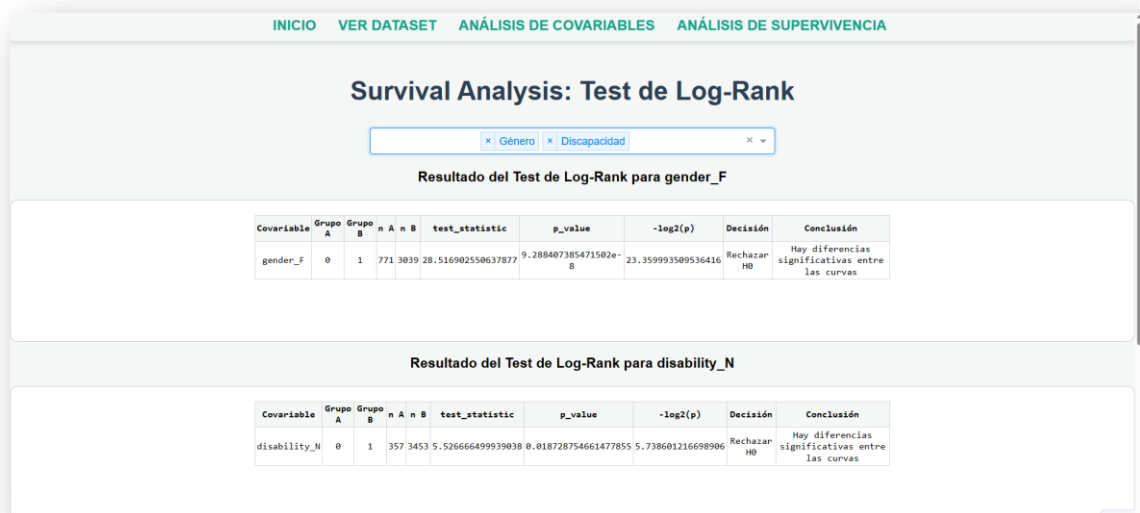


Ilustración 30: Opción covariables Género y Discapacidad

Basándonos en las tablas propuestas, las variables que se estudian dentro del Test son las siguientes:

Covariable	Variables independientes que se están analizando en el modelo de regresión de Cox
Grupo A y Grupo B	Representan los dos grupos comparados en el test de Log-Rank. Cada uno está determinado por la covariable seleccionada.
n A y n B	Número de observaciones en cada uno de los dos grupos, es decir, el recuento del grupo A y el total del grupo B.
Test_stadistic	Valor del estadístico de prueba calculado por el Test de Log-Rank. Valor utilizado para determinar si las curvas de supervivencia de los dos grupos son significativamente diferentes. Cuanto más alto sea el valor, mayor será la diferencia entre ambos grupos.
p_value	Valor que indica la probabilidad de obtener los resultados observados si la hipótesis nula fuera verdadera. Si el valor p es menor que el nivel de significancia, se rechaza la hipótesis nula.
-log2(p)	Valor transformado del valor p usando el logaritmo en base 2. Esto produce que los valores p más pequeños

	sean más fácilmente interpretables. Por tanto, un valor alto, da lugar a un valor p muy pequeño.
Decisión	Basado en el valor de P y el nivel de significancia, se toma una decisión nula y se concluye que hay diferencias significativas entre las curvas de supervivencia de los dos grupos.
Conclusión	Interpretación del resultado del test. Si se rechaza la H0, se concluye que existe una diferencia significativa entre las curvas. Si no se rechaza, las diferencias no son notables.

Tabla 2: Variables estudiadas Pruebas de Log Rank

3.5.4 Prompt IA

Inmerso en la sección del Análisis de Supervivencia, este sistema ofrece una opción de explicación automática de los resultados generados por los tres modelos de análisis: Kaplan-Meier, Regresión de Cox y Test de Log-Rank.

Esta funcionalidad permite a los usuarios obtener una interpretación clara y detallada de los datos expuestos en las tablas y gráficas tras realizar cualquiera de los análisis.

El proceso del usuario es sencillo, ya que únicamente puede hacer clic en el botón “Explicar” que se observa en la interfaz en la parte superior del cuadro de explicación.

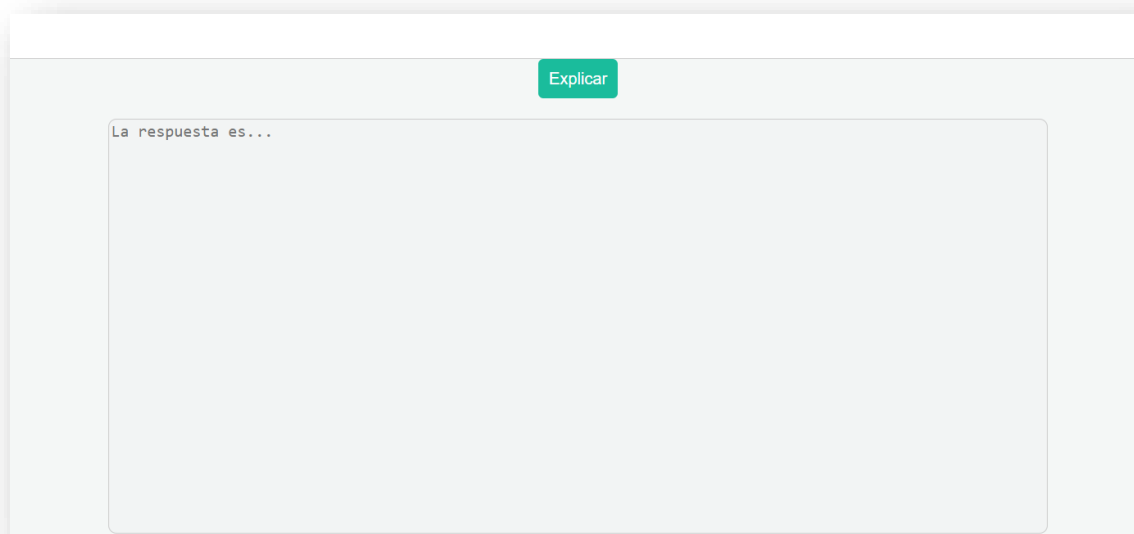


Ilustración 31: Prompt IA

El proceso en el que se genera la explicación consta de los siguientes pasos:

1. El sistema toma los resultados visualizados en la pantalla. En el caso de Kaplan, recoge la gráfica; y en el de Cox y Log Rank, recoge los resultados de la tabla.
2. Dependiendo del tipo de análisis, el sistema envía los datos relevantes al modelo Llama3 a través de la IA integrada de Ollama.
3. El modelo de IA genera un resumen interpretativo en español, proporcionando una explicación que describe de manera concisa lo que muestran los resultados
4. La respuesta generada por la IA se muestra automáticamente en la pantalla del usuario en el cuadro de texto que se aprecia en la imagen, donde se visualiza de forma instantánea.