

Módulo 2– Procesamiento de datos en R

Trabajo Practico N° 2



Objetivos

El objetivo del presente trabajo es que puedan integrar todos los conceptos vistos durante el curso, pero principalmente los del módulo #2.



Consigna

La idea principal es que puedan realizar un análisis de un dataset a elección, mediante el tratamiento de los datos y algunas visualizaciones con el fin de poder responder las hipótesis planteadas.

- 1) Deberán seleccionar un dataset (puede ser público o de sus trabajos) que contenga como mínimo 50.000 filas y 10 columnas con variedad en sus tipos, por ej: numéricas, fecha, categóricas, etc.
- 2) Una vez obtenido el dataset, la idea es que planteen 4 o 5 preguntas de interés sobre el mismo para poder responderlas posteriormente.
- 3) Sera necesario que carguen el archivo y realicen un chequeo/limpieza de los datos, es decir, verificar la integridad de los mismos.
- 4) Realizar un análisis exploratorio de datos considerando cuestiones generales como, por ejemplo: tipos de datos, distribuciones, valores atípicos, correlación entre las variables, etc.
- 5) Responder las preguntas planteadas en el punto 2 mediante el uso de tablas de contingencia y gráficos relevantes. (Serán necesarios como mínimo 5 gráficos, 3 de ellos distintos).

Formato de presentación:

El archivo final puede ser un RScript en conjunto al dataset elegido o un archivo .Rdata para poder cargar todo el entorno de trabajo con el código corrido por ustedes.

Si lo desean pueden utilizar las R Notebooks, pero no es obligatorio. Si suben su proyecto a un repositorio de Github es un plus!

Fecha límite de entrega:

La fecha límite de entrega se establece como una semana después de haber finalizado la última clase del segundo módulo. Como es el último trabajo del curso, se dispone de una semana extra para los casos donde no se encuentre aprobado en la primera instancia.

Criterios de evaluación

Se contemplará fuertemente la originalidad del código como también el formato y la prolijidad, como también si el mismo se encuentra ordenado y debidamente comentado. Cualquier actividad extra que se quiera adicionar es bienvenida, siempre y cuando se cumplan las consignas previamente establecidas.

Información adicional

Les adjunto algunos links donde podrán encontrar datasets públicos bastante interesantes:

- **Buscador de Dataset de Google:**

<https://datasetsearch.research.google.com>

- **Plataforma de datos abiertos del Gobierno de la Ciudad:**

<https://data.buenosaires.gob.ar/dataset>

- **Buscador de dataset Kaggle:**

<https://www.kaggle.com/datasets>

- **Datos Abiertos de la OMS:**



<https://www.who.int/data>

- **Datos Abiertos del Gobierno de EEUU:**

<https://www.data.gov/>

- **Oficina del Censo de EEUU:**

<https://www.census.gov/data.html>

- **Conjunto de Datos Abiertos de FreeCodeCamp:**

<https://github.com/freeCodeCamp/open-data>

- **Buscador de Datos de UNICEF:**

<https://data.unicef.org/>

- **Data Science Doho:**

<https://datasciencedojo.com/blog/datasets-data-science-skills/>

- **UCI Machine Learning:**

<https://archive.ics.uci.edu/ml/datasets.php>

Y en el siguiente link tienen un tutorial bastante completo donde se especifican los pasos a seguir para cargar sus archivos en un repositorio de Github.

[https://es.khanacademy.org/computing/computer-programming/html-css/web-development-tools/a/hosting-your-website-on-github#:~:text=Abre%20tu%20proyecto%20en%20Github,files%22%20\(Subir%20archivos\).&text=Ver%C3%A1s%20la%20interfaz%20de%20Github,los%20archivos%20que%20quieras%20subir.](https://es.khanacademy.org/computing/computer-programming/html-css/web-development-tools/a/hosting-your-website-on-github#:~:text=Abre%20tu%20proyecto%20en%20Github,files%22%20(Subir%20archivos).&text=Ver%C3%A1s%20la%20interfaz%20de%20Github,los%20archivos%20que%20quieras%20subir.)

[github#:~:text=Abre%20tu%20proyecto%20en%20Github,files%22%20\(Subir%20archivos\).&text=Ver%C3%A1s%20la%20interfaz%20de%20Github,los%20archivos%20que%20quieras%20subir.](https://es.khanacademy.org/computing/computer-programming/html-css/web-development-tools/a/hosting-your-website-on-github#:~:text=Abre%20tu%20proyecto%20en%20Github,files%22%20(Subir%20archivos).&text=Ver%C3%A1s%20la%20interfaz%20de%20Github,los%20archivos%20que%20quieras%20subir.)