

Primer Proyecto – Explorando datos con R

Los estudiantes desarrollarán un estudio descriptivo de un set de datos a su elección. Para buscar el set de datos se sugiere utilizar el sitio <https://www.kaggle.com/> que tiene sets de datos de diferentes temas. Los requisitos para el dataset es que tenga al menos 1,500 observaciones y al menos tres variables numéricas. El proyecto a desarrollar debe ser aprobado en clase antes del 9-3-2023.

Objetivos

- Hacer uso de R y R Studio como herramientas para explorar conjuntos de datos
- Utilizar estadística descriptiva para comprender las características de un set de datos
- Utilizar con propiedad los paquetes dplyr y ggplot2

Requisitos de entrega

- Presentar un informe de investigación descriptivo
 - Introducción
 - Método
 - Análisis
 - Conclusiones
- Utilizar al menos las siguientes herramientas
 - R Studio
 - Estadística descriptiva
 - Gráficas (histogramas, gráficas de frecuencia, boxplot, etc.)
 - Uso de las siguientes librerías
 - Dplyr
 - Ggplot2
- Entregables (ideal que todo el proyecto esté en un ZIP)
 - Informe de su investigación
 - Scripts de R. Estos deben estar documentados y deben poder reproducirse en cualquier computadora que tenga R Studio según instrucciones que ustedes proveerán
 - Dataset(s). Deben subir los archivos separados por coma que contienen sus datos.

Sugerencia

- Explorar información relacionada con el tema seleccionado
- Explorar eventos relacionados

Fecha de entrega: 16 de Marzo

Modalidad: grupos de 3 personas máximo