Universidad Francisco Marroquín

Data Wrangling - 2019

# Laboratorio #10 - Exploratory Data Analysis

Las tablas “movies.csv” y “ratings.csv” contienen data de películas y sus respectivos ratings de usuarios de IMDB. Utilizando R o Python, realice un EDA. (Notar que debe de incluir tanto las features de la pelicula como el rating de la tabla “ratings.csv”)

El reporte deberá de incluir las gráficas necesarias para las variables numéricas y categóricas, con su respectivo análisis y conclusiones.

**Rúbrica de calificación:**

1. Detalle de tipos de datos (5%)
2. Manipulación de las tablas (10%)
3. Análisis de las variables numéricas (20%)
4. Análisis de las variables categóricas (20%)
5. Correlación entre variables (20%)
6. 3 gráficas de correlación entre dos variables a su elección (algunas variables pueden repetirse, por ejemplo revenue vs rating, revenue vs votes, votes vs rating) (5% cada una)
7. Conclusiones (10%)

El laboratorio deberá de ser entregado por medio de MiU a más tardar el domingo, 3 de noviembre a las 11:59pm. No estaremos aceptando entregas tarde ni por correo electrónico. La entrega sera el link al documento en GitHub, en formato markdown, estén trabajando en R o en Python.