

Nombre del Programa: Magíster en Ciencia de Datos

Curso: Python para Ciencia de Datos

Profesor: Alejandro Cataldo

Tarea 1

Objetivo

El objetivo de esta tarea es que apliquen algunos de los conceptos aprendidos durante su revisión del MOOC, y con esto, puedan aplicar algunas de las herramientas de Python que ya han aprendido. Para esto, deberán utilizar el archivo “data_football_ratings.csv”, que contiene datos sobre algunos partidos de fútbol disputados entre los años 2016 a 2018. Así, en cada fila de este conjunto de datos se entregan estadísticas sobre el rendimiento de un jugador en un partido, pudiendo ser más de un jugador por partido, según la calificación de algunas de seis publicaciones deportivas.

Preguntas que deben responder

Tomando el archivo entregado, y cargándolo como Dataframe en Python, deben entregar un archivo en Jupyter Notebook que entregue el código y la respuesta/resultado a cada una de las siguientes 6 preguntas.

- (1) ¿Cuál es el promedio, mediana y desviación estándar de los datos registrados en la columna “original_rating”?
- (2) ¿Cuántos de los partidos de la base de datos tuvieron una duración (“game_duration”) de 120 minutos?
- (3) Considerando solamente los partidos de la Copa del Mundo 2018 (“World Cup 2018”), ¿cuántos partidos disputó Croacia?
- (4) Construya un gráfico de torta con los goles marcados por Robert Lewandowski, agrupando los goles que marcó por su selección y por su club.
- (5) ¿Cuántas tarjetas rojas fueron mostradas en la Eurocopa 2016 (“Euro 2016”)?
- (6) Construya un Dataframe que contenga, para cada jugador del archivo, la cantidad total de partidos en los que fue evaluado. Muestre solamente los 10 primeros registros de este Dataframe. Considere que un jugador no es evaluado cuando original_rating = 1 y los minutos jugados por el jugador es menor que 20 (minutos).

Reglas

Para realizar esta tarea debe utilizar y cargar el archivo “data_football_ratings.csv” sin ninguna modificación. Debe entregar un archivo Python que se ejecute en Jupyter Notebook, y que contenga el número de la pregunta, el código y la respuesta/resultado de cada pregunta.