

Data Engineering and Visualization, Winter Semester 2023
Project Milestone 2
Submission Date: Monday 25/12/23

In this milestone you will perform

- Data Integration:
 - You will be integrating any data sets you prefer from the available data sets provided in the following link <https://www.kaggle.com/datasets/davidcariboo/player-scores/data?select=appearances.csv>. The integration should result in a new dataset and the newly data should be result in at least 2 research question per individual in the group. (note you do not have to deal with any outliers/missing values. If you have any you can drop it or replace with mean/mode.)
- Analysis
 - You should include the answers to all your research questions in the notebook. Answers should be provided as visualizations with a markdown cell explaining the insights we gain from the graph.

The grades will be given for:

- Implementation:
 - Descriptive markdown cells for your code cells. Your notebook is the place you describe and document the space of possibilities you explored at each step of your project. What visualizations did you use to look at your data in different ways? What are the different statistical methods you considered? Justify the decisions you made, and show any major changes to your ideas. How did you reach these conclusions?
 - Clean code. ex: descriptive variable names
 - Comment your code when relevant.
 - Implementation:
 - * Is the data integrated properly?
 - * Did you provide all the required research questions?
 - * Did you provide visualizatinns answering your research questions clearly?
- Example of things that can lead to **deduction of marks**:
 - Unuseful features
 - Wrong integration
 - Research questions with no answers
 - Inappropriate choice of plots/labels/title
 - Messy code
 - Incorrect conclusions
 - Trivial Insights

Deliverables

You should submit

- Your submitted jupyter/colab notebook

Submission Instructions

Submission will be handled through GitHub Classroom. A further notice will be sent later about google classroom.