

# Project: Exploratory Data Analysis

## Overview

For this project, you will choose a dataset that is interesting to you and conduct exploratory data analysis (EDA) to gain insights into the data. The goal of this project is to practice EDA techniques and apply them to a real-world dataset.

## Dataset locations

Kaggle, UCI, sklearn-datasets are good places to look.

## Requirements

1. Choose a dataset that has at least 1000 observations and roughly 10 variables or more.
2. Perform data cleaning and preprocessing as necessary.
3. Conduct exploratory data analysis using some EDA techniques, including visualizations, correlation analysis, (outlier detection, feature engineering, and dimensionality reduction).
4. Write a short report summarizing your findings and insights.

## Deliverables

1. A Jupyter notebook or Python script with your EDA code and visualizations.
2. A report summarizing your findings and insights. The report should include the following sections:
  - Introduction: Describe the dataset and its variables.
  - Data cleaning and preprocessing: Describe the steps you took to clean and preprocess the data.
  - Exploratory data analysis: Summarize your findings and insights from each of the EDA techniques you applied.
  - Conclusion: Summarize your overall findings and insights.