# OVERVIEW

This repo contains files and related project data for a Code Louisville bootcamp on Python programming and data analysis. It uses data from Kaggle located at https://www.kaggle.com/adamjyu/major-league-pitching-pandas-tutorial-and-eda/notebook

# SCOPE

The website has an excellent tutorial for information on major league pitching. The data from the Kaggle site contains data for hitting, pitching, fielding, etc., and has additional data for managers, postseason, and Hall of Fame personnel. The time period is from 1871 to 2015, but for this project, we will only be looking at data from 1965 to 1972.

After 1950, Major League Baseball set the standard height of the pitching mound to 15 inches. After “The Year of the Pitcher” (1968), MLB lowered the height from 15 to 10 inches.

The purpose of this analysis is to see if there is an actual change in the pitcher’s Earned Run Average (ERA), Strikeouts per Nine Innings, Walks per Nine Innings, and Strikeout to Walk Ratio. The data is focused on the four years before and four years after the implementation of the lowered mound height.

# INSTRUCTIONS

CODE LOUISVILLE project requirements are called out in the code, just search for "Code Louisville" in the py files.

This program was built using Python version 3.10.1 and pip version 21.3.1, with the following installed packages:

To install prerequisites:

- >pip install pandas

- >pip install numpy

- >pip install logging

To run the program:

- >git clone https://github.com/Joe-Mart/BaseballHistory.git

- >cd BaseballHistory

- >python Pitching\_Analysis.py

# Code Louisville Requirements

Category 1: Python Programming Basics:

- [x] Create a dictionary or list, populate it with several values, retrieve at least one value, and use it in your program.

Category 2: Utilize External Data:

- [x] Read data from an external file, such as text, JSON, CSV, etc, and use that data in your application.

Category 3: Data Display

- [x] Visualize data in a graph, chart, or other visual representation of data.

Category 4: Best Practices

- [x] Use logging to track the progress of the execution of the program.

"Stretch" list:

- [x] Use pandas, matplotlib, and/or numpy to perform a data analysis project. Ingest 2 or more pieces of data, analyze that data in some manner, and display a new result to a graph, chart, or other display.