Crossfit athletes

Jordi Díaz Carné

Project Overview

Original dataset contains:

- 423.006 rows × 27 columns
- It is a database in which athletes from all over the world were asked things such as: age, weight, lifting marks, times of certain workouts and their training habits.

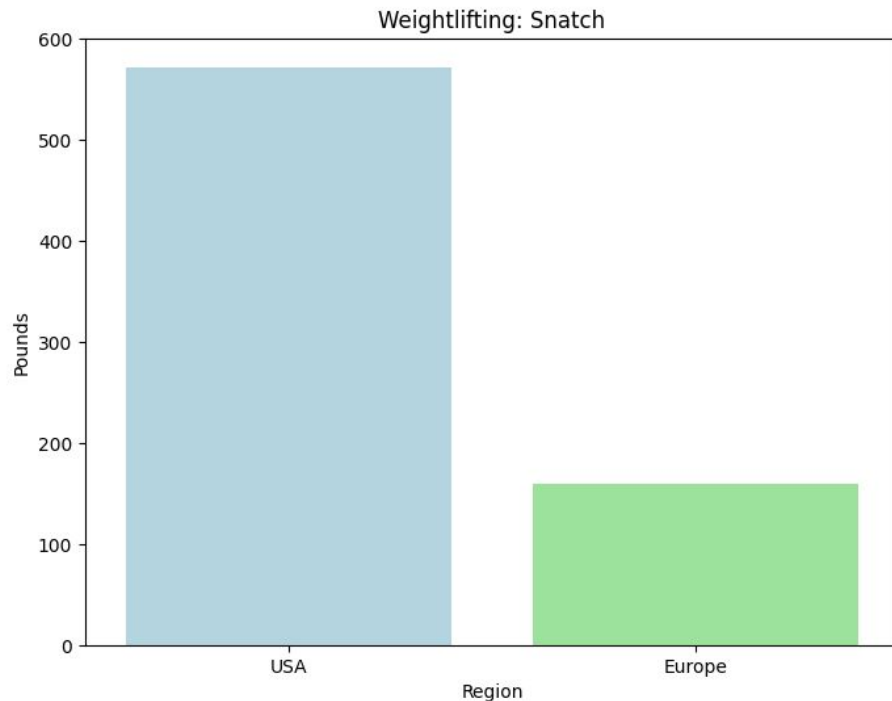# Data Wrangling and Cleaning
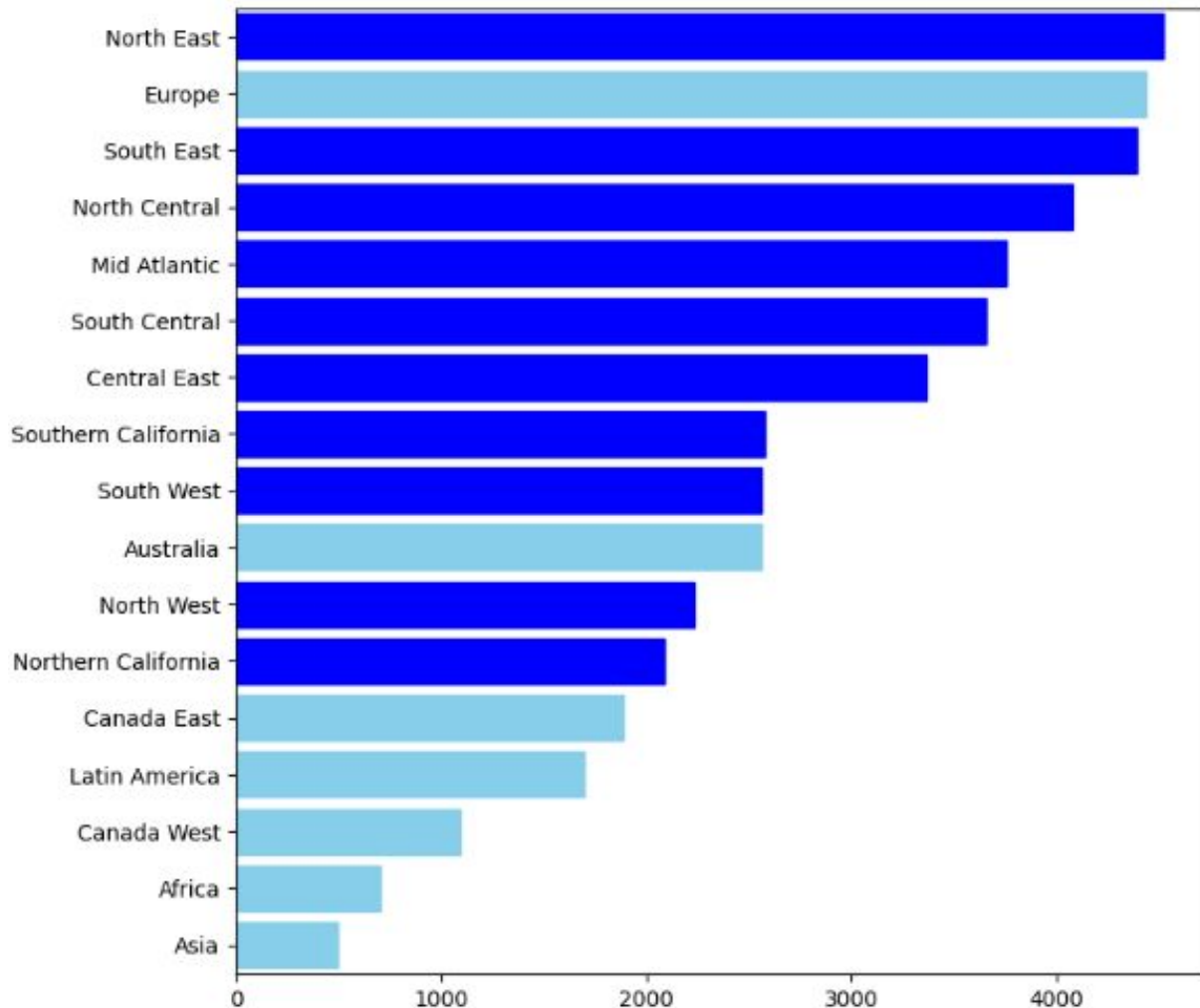
The methods I have used for data cleaning:

- df.**replace**(0, np.nan)
- Df.**drop**()
- Df.**dropna**(how='all') and (thresh=10)

The APIs I have used are:

- pandas
- numpy
- counter (from collections)
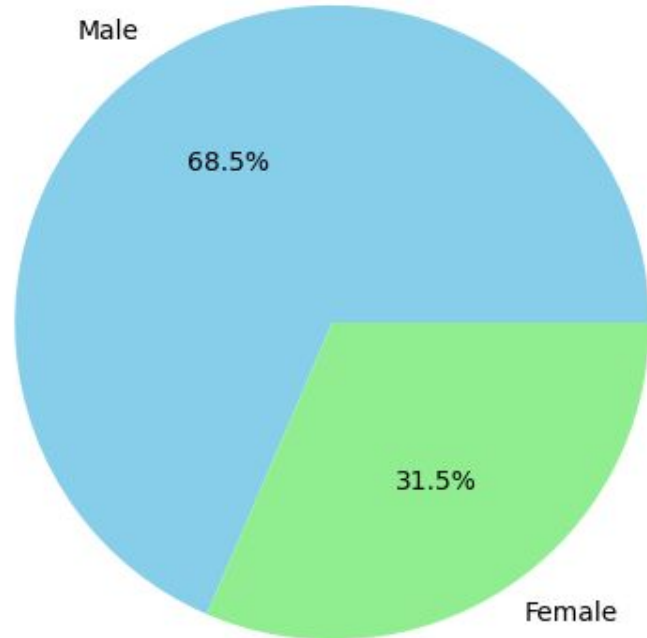- matplotlib.pyplot
- seaborn

Number of people doing Crossfit by region

Gender
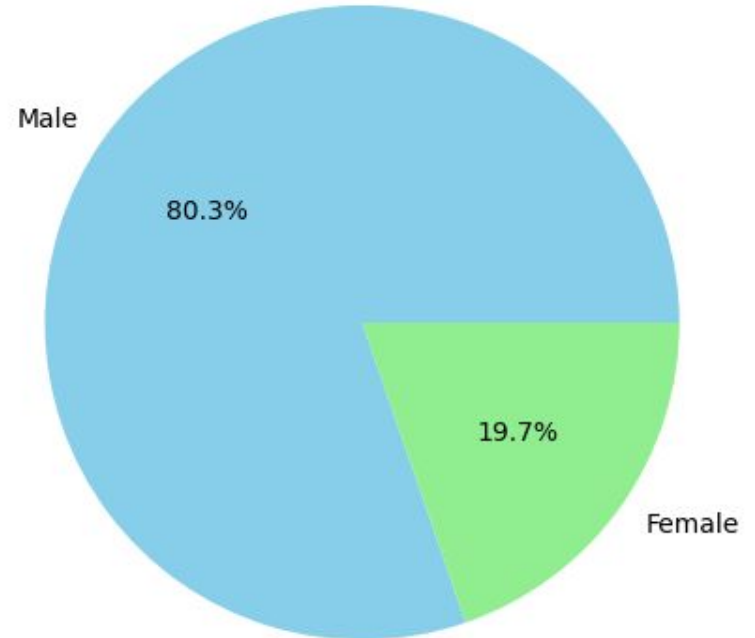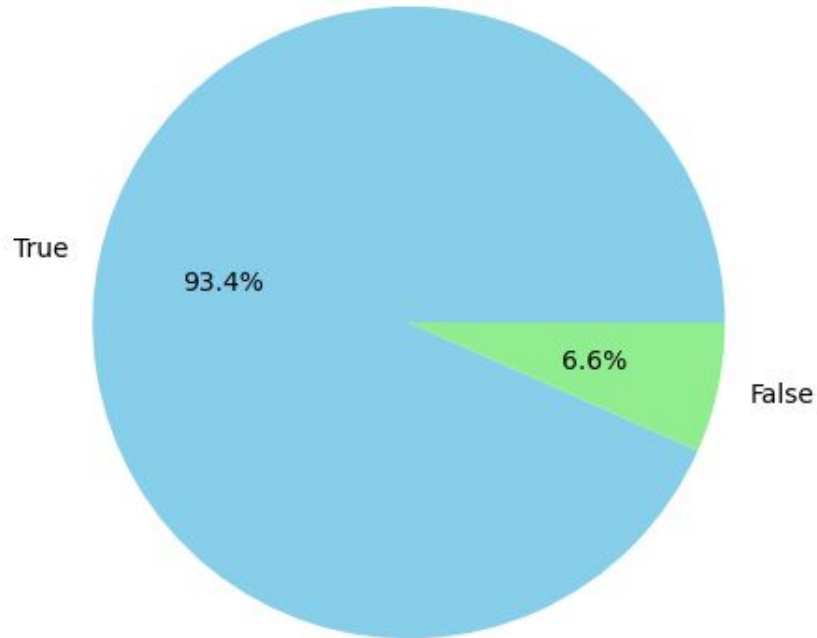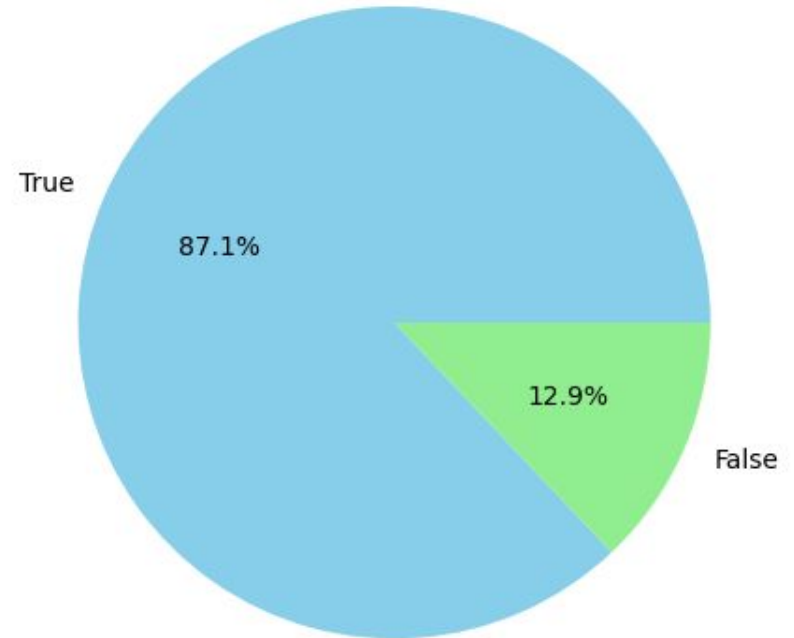

Gender Distribution in USA
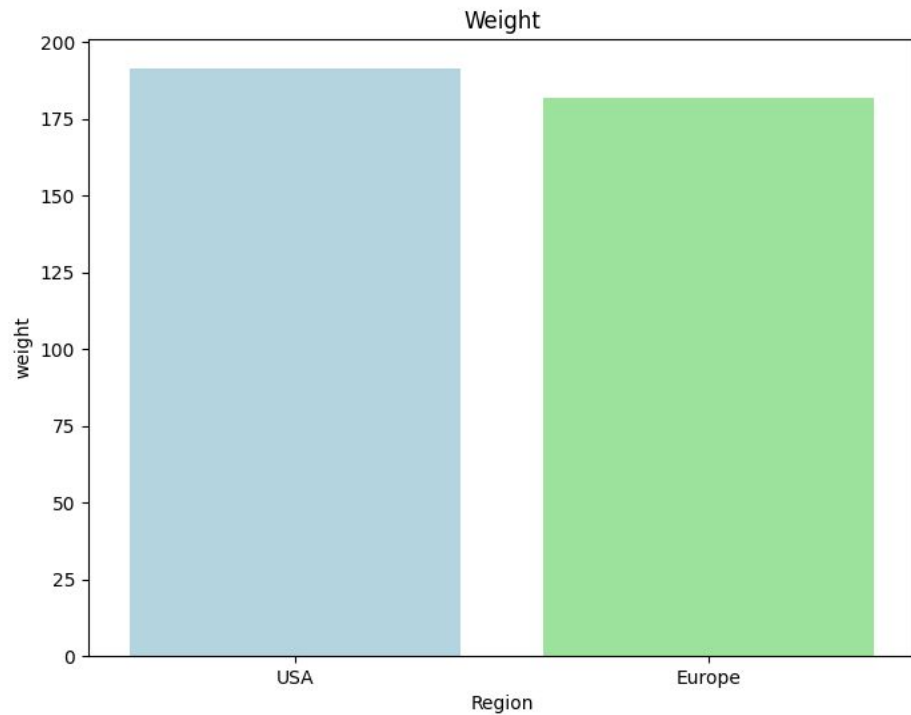
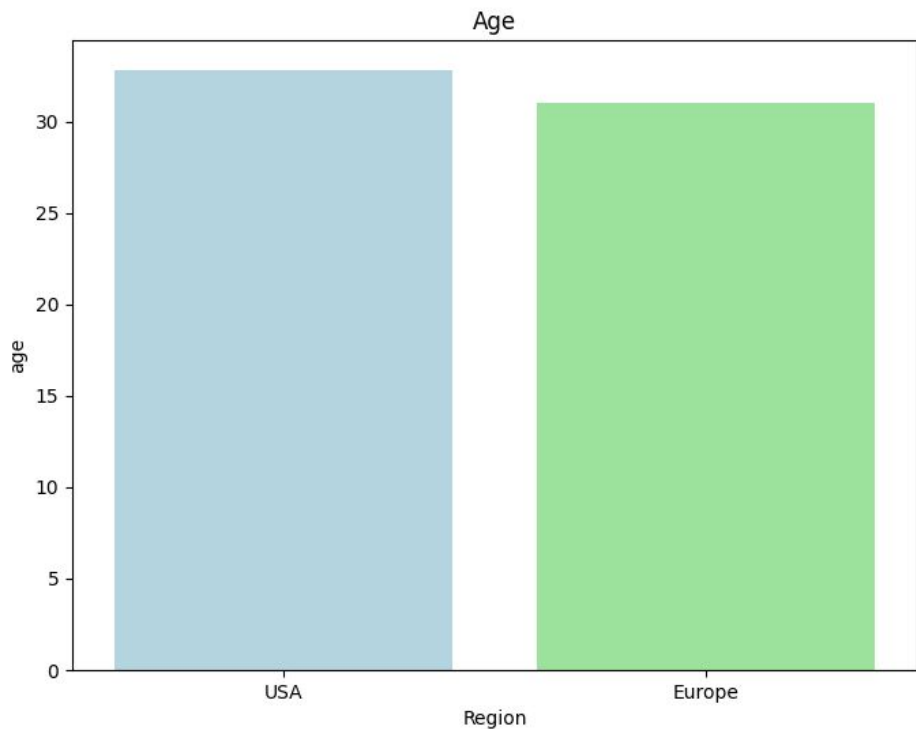Gender Distribution in Europe

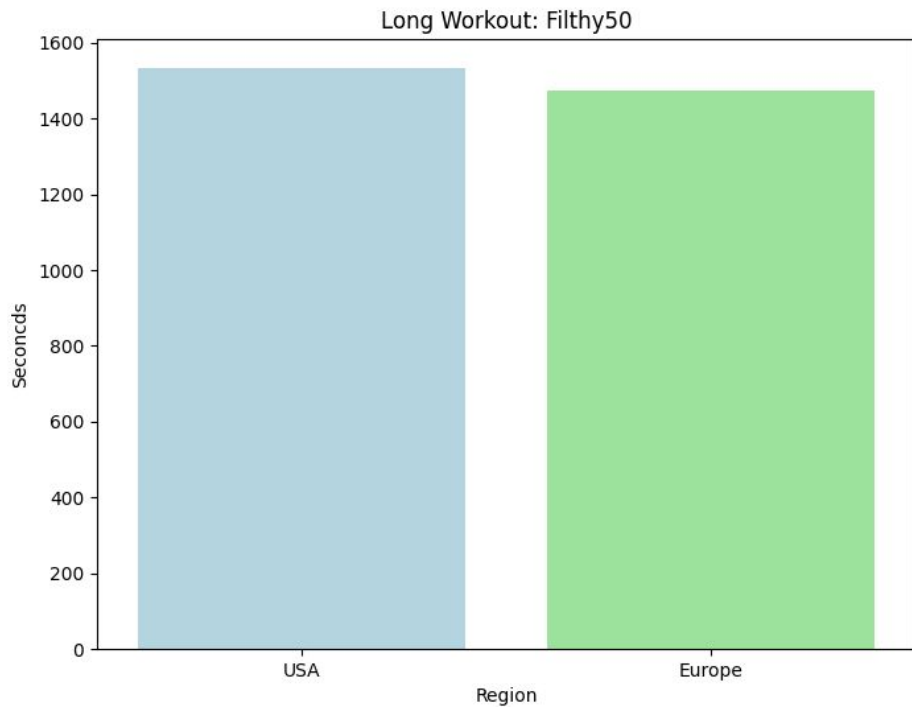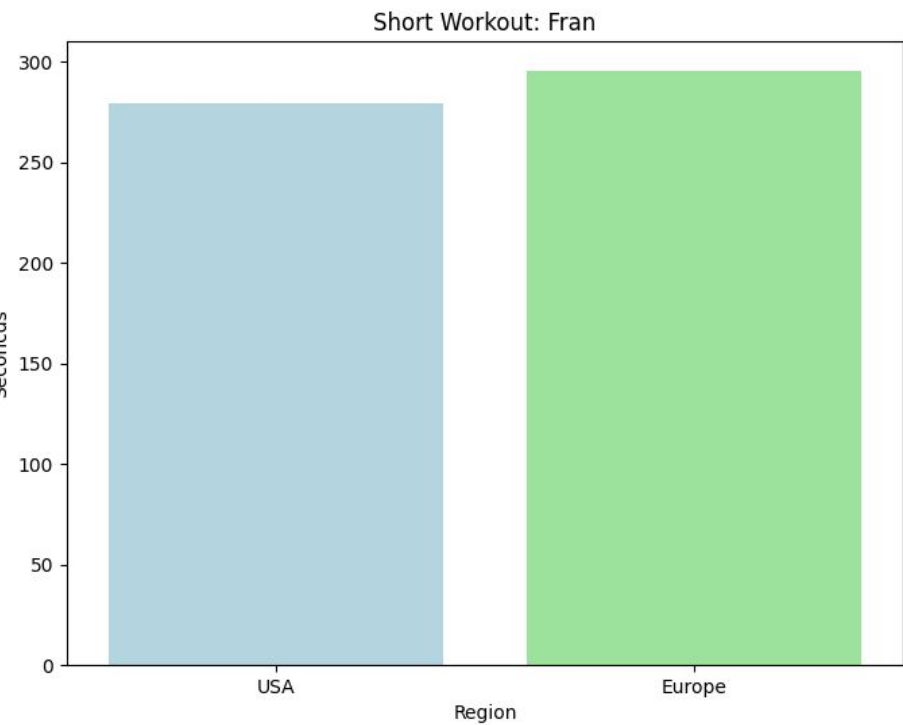# Crossfit Affiliate

Train in Affiliate Crossfit in USA

True 93.4%
6.6% False

Train in Affiliate Crossfit in Europe

True 87.1%
12.9% False

Crossfit athletes

Fitness

Fitness
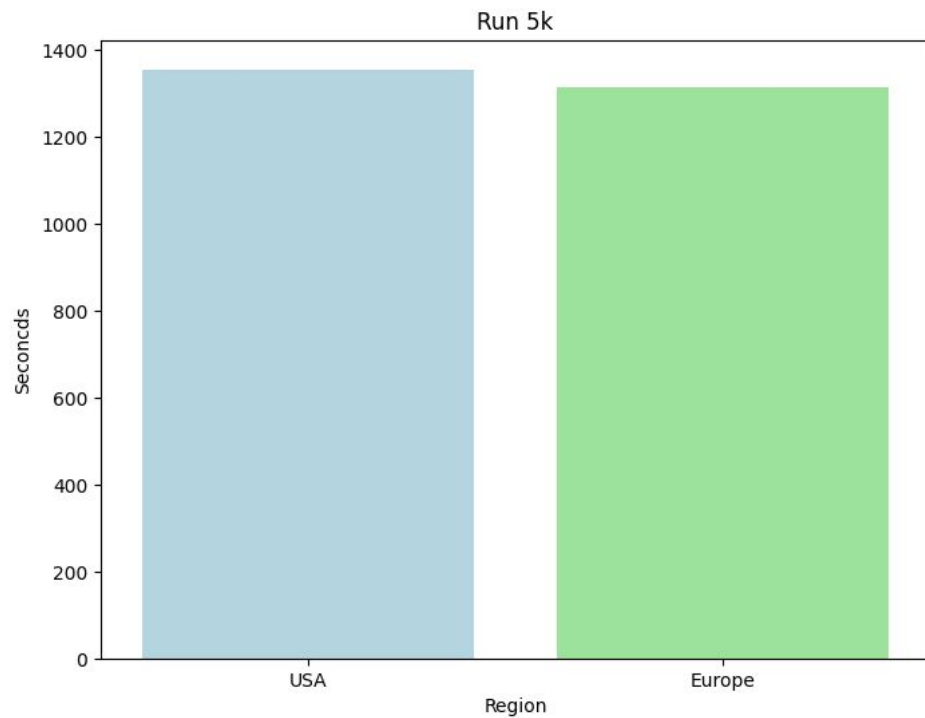
Strength
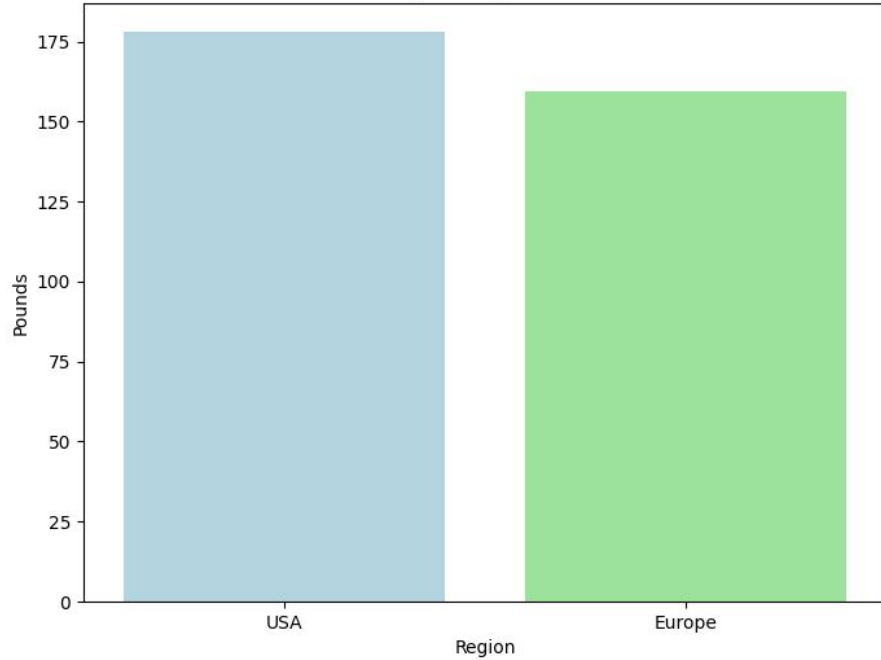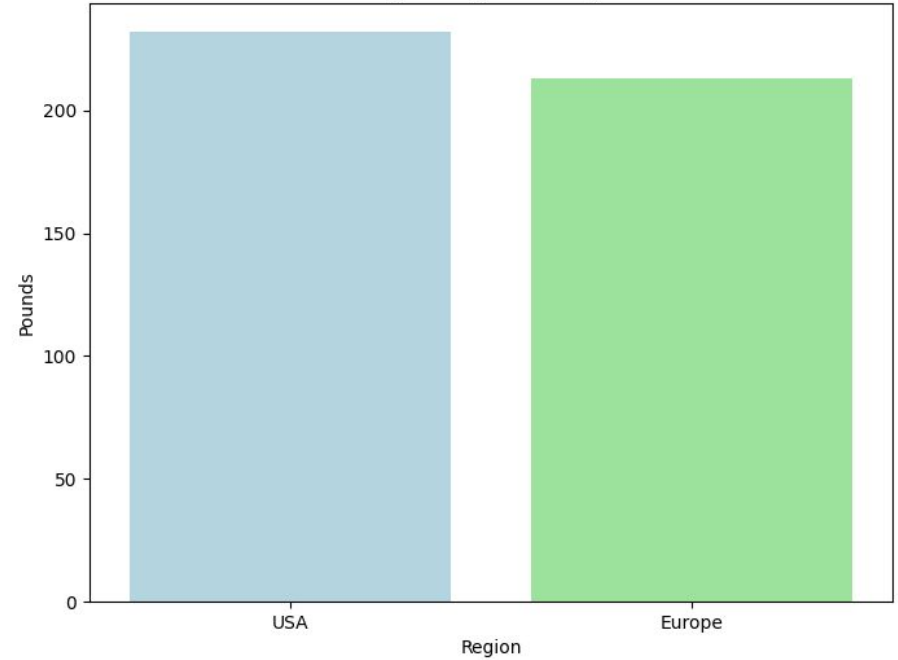
Major Obstacle

Discuss the biggest obstacle or mistake you encountered during this project.

Los principales obstáculos o errores que encontré durante este proyecto son:

- Not being able to download the table from a web page doing web scripting
- Reduce the number of results in a column by grouping them into groups
- Learn how to make graphics
- Delete outliers
- Start deleting column without using .copy()

Conclusions

- The majority of people who do Crossfit are from the USA
- In the USA there are more women who practice and more people practice it in Affiliate
- IN the USA the average athlete is slightly older and heavier. That probably correlates with them being stronger
- On the other hand, in Europe, athletes have shown greater fitness in the tests carried out.

Thank you for your time!

Jordi Díaz Carné