Name : Bryan Rendra Pramana

Class : CIS 3120

Instructor : Prof. Jairam

Project 1

The objectives of the project 1 is to get a deep insight of volleyball and swimming player team in different various colleges. The method that I am using is leveraging the capability of Python programming languages to scrap data from 5 colleges. The library that being used in this project is pandas and beautiful soup. There are 2 data that being scrapped, the first one contains string which is name of the player team and the other one is their height. Comparing to the previous assignment that given in the class, it has similar concept, however the dataset in this project is much larger and wider. The purpose of having larger data set in a data analytic project is to get more accurate, new insight, and to make any predictive analytical strategy. The way I scrapped the data is unlike as instructed in the project guideline. Student was asked to scrap every single website and create a data frame for every single one of them. Later once all rosters have been scraped, the data frames is expected to be merged as single big data frame. However, I am using a list to scrap all the website at once and append the data directly into a big data frame. To find the average, first I needed to convert the height of the player from feet to cm and by using .mean() I could get the average of the players easily. Below is the list of the founding :

* Average height Men Swimmers Team 179.60622222222233 cm
* Average height Women Swimmers Team: 164.7190000000001 cm
* Average height Men volleyball Team 182.88000000000005 cm
* Average height Women volleyball Team 168.44210526315797 cm

After analyzing the data that have scrapped, It can be concluded that the average height of men’s swimmers is shorter compare to men’s volleyball height. The women’s swimmers team average high is also shorter than women’s volleyball team. From these data set, we know that volleyball player is requiring a taller player where the height is matter when you are playing volleyball. There are some difficulties to scrap the data in various website, some of the website has different structures than the other. The difference forced me to redesign the algorithm on how to read the html. Below are the list of the top 5 highest player and 5 shortest player for each game by gender.

5 highest Men Swimmers height

Name Height

12 Gregory Becker 195.58

16 Kasper Gacek 195.58

9 Mark MacEachen 193.04

27 Evan Nikolic 190.50

34 Morgan Welling 190.50

5 smallest Men Swimmers height

Name Height

30 Ben Sionov 165.10

22 Fernando Marroquin 170.18

33 Joseph Vasile 170.18

38 Anthony Jordan 170.18

40 hristopher Mendez 170.18

5 highest Women Swimmers height

Name Height

4 Shenaika Eliassaint 180.34

13 Onika George 177.80

8 Alyssa Taylor 175.26

12 Andrea Gasic 175.26

15 Asimina Hamakiotes 175.26

5 smallest Women Swimmers height

Name Height

17 Amanda Lee 152.40

5 Julie Huang 154.94

10 Stephanie Antonova 154.94

16 Mirtho-Myra Lamonier 154.94

24 Alicia Soares 154.94

5 highest Men volleyball height

Name Height

18 Artem Zinkin 195.58

21 Leon Petrovitsky 195.58

30 Akil Vaughn 195.58

10 Andrew Tsororos 193.04

19 Michael Higgins 193.04

5 smallest Men volleyball height

Name Height

32 Jayden Deokinanan 162.56

0 Snigdho Hasan 165.10

26 John Vitor 167.64

27 Juan Rodriguez 167.64

28 Shawn Nemoto 170.18

5 highest Women volleyball height

Name Height

34 Autumn Reeves 187.96

16 Claudia Daporta 182.88

31 Mirialis Almanzar 182.88

35 Sophia Wheelan 182.88

5 Madison Cronk 180.34

5 smallest Women volleyball height

Name Height

18 Young Gi Go 152.40

9 Savanna Koutsakis-Keener 154.94

24 Cady Xia 154.94

6 Laurie Lee 157.48

14 Jasmin Vidal Ventura 157.48