

ITP 449

HOMEWORK 3

20 POINTS

For each one of the following questions, write Python code in PyCharm. For each question, create a new Python file. Name it *HW3_Q1_lastname_firstname.py*. Create a header in each file using *comments* to display your name and HW information. After that write your Python code.

Tommy Trojan

ITP 449

HW3

Question 1

- Apart from the above comments, include single line comments describing the core logic of your algorithm / code.

As an example,

Printing the names of the team Quarterbacks

Question 1:

Consider the following DataFrame:

	attempts	name	qualify	score
0	1	Anastasia	yes	12.5
1	3	Dima	no	9.0
2	2	Katherine	yes	16.5
3	3	James	no	NaN
4	2	Emily	no	9.0
5	3	Michael	yes	20.0
6	1	Matthew	yes	14.5
7	1	Laura	no	NaN
8	2	Kevin	no	8.0
9	1	Jonas	yes	19.0

1. Create a dictionary in Python, which contains this information. Then define a DataFrame in pandas using this dictionary.
Hint: Use *np.nan* to define NaN entries
2. Write a single-line code in Python to print the *name* and the *attempts* of qualified contestants.
3. Write a single-line code in Python to print the *name* and the *score* of those contestants, who qualified with a single attempt.
Hint: Check out *np.logical_and()* in *numpy*
4. Write a single-line code in Python to replace all the NaN values with Zero's in the score column of a dataframe. Then print the dataframe to confirm the change.
Hint: Check out *np.isnan()* in *numpy*
5. Write a single-line code in Python to print the dataframe such that it is sorted the by *attempts* in ascending order (and *score* in descending order if 2 contestants have the same number of *attempts*.)
Note: Make sure to leave *inplace=False*

Question 2:

Trojans_roster.csv – Roster dataset for 2022 USC Trojan football team, containing 9 aspects (attributes) of each player: #, NAME, POS., HT., WT., YEAR, HOMETOWN, LAST SCHOOL, MAJOR

1. Read the csv file using Pandas. Store the output into a dataframe named *DF*. Then print *DF*
2. You notice that the index is 0 ... 109. There is a column #. Set the index of the dataframe to #. In other words, make the player's number column the index of *DF*. Print *DF* to make sure that the change happened.
Hint: Check out *set_index()* method for dataframe. Make sure to make the change *inplace=True*
3. Remove the 'LAST SCHOOL' and 'MAJOR' columns from the dataframe. Print *DF* to make sure that the change happened.

4. Write a single-line code in Python to print the names of all the Quarterbacks in the team.
5. Write a single-line code in Python to print the name, position, height, and weight of the tallest player in the team.
6. Write a single-line code in Python to print how many players are local (i.e., their hometown is 'LOS ANGELES, CA'). Note that the answer is a number.
7. Write a single-line code in Python to print the info of 3 heaviest players.
8. Define a new column for *DF* named BMI, which contains the BMI of the players. Print *DF* to make sure that the change happened.

$$\text{BMI} = 703 \times \frac{\text{Weight (lb)}}{\text{Height}^2 \text{ (in)}}$$

9. Write single-line codes in Python to print the mean and median of players' height, weight, and BMI.
10. Write single-line codes in Python to print the mean and median of players' height, weight, and BMI for each position.
Hint: Check out *groupby()* method for dataframe.
11. Write a single-line code in Python to print the number of players in each position.
Hint: Check out *count()* method for dataframe.
12. Write a single-line code in Python to print the names of the players whose BMI is below the team average (mean).
13. Write a single-line code in Python to print all the unique players' numbers. (unique numbers in *DF* index)
Hint: Check out *unique()* method for dataframe.

Zip all files together (name the zip file *lastname_firstname_hw3.zip*) and submit it on Blackboard.