

```

1  """
2  Pandas - Mini Project 3
3
4  # Part A:  Alcohol Consumption
5  A1 Upload the data file, view the data file in Colab
6  A2 Read and Assign it to a variable called users and use the 'user_id' as index
7  A3 Which continent drinks more beer on average?
8  A4 For each continent print the statistics for wine consumption.
9  A5 Print the mean alcohol consumption per continent for every column
10 A6 Print the median alcohol consumption per continent for every column
11 A7 Print the mean, min and max values for spirit consumption.
12
13 # Part B :  Student Alcohol Consumption
14
15 B1 Upload new data file 'student-mat'
16 B2 For the purpose of this exercise slice the dataframe from 'school' until the 'guardian' column
17 B3 Create a lambda function that will capitalize strings.
18 B4 Capitalize both Mjob and Fjob
19 B5 Print the last elements of the data set.
20 B6 Create a function called majority that returns a boolean value to a new column called majority
21 B7 Multiply every number of the dataset by 10
22 """

```

↳ '\nPandas - Mini Project 3\n\n# Part A: Alcohol Consumption\nUpload the data file, view the data file in Colab\nRead and Assign it to a variable called users and use the 'user_id' as index\nWhich continent drinks more beer on average?\nFor each continent print the statistics for wine consumption.\nPrint the mean alcohol consumption per continent for every column\nPrint the median alcohol consumption per continent for every column\nPrint the mean, min and max values for spirit consumption.\n\n# Part B : Student Alcohol Consumption\n\nUpload new data file 'student-mat'\nFor the purpose of this exercise slice the dataframe from 'school' until the 'guardian' column\n

A1 Upload the data file, view the data file in Colab

Done

A2 Read and Assign it to a variable called users and use the 'user_id' as index

user_id do not exists

```

1 import numpy as np
2 import pandas as pd
3 users= pd.read_csv("drinks.csv")
4 users

```

	country	beer_servings	spirit_servings	wine_servings	total_litres_of_pure.
0	Afghanistan	0	0	0	
1	Albania	89	132	54	
2	Algeria	25	0	14	
3	Andorra	245	138	312	
4	Angola	217	57	45	
...
188	Venezuela	333	100	3	
189	Vietnam	111	2	1	
190	Yemen	6	0	0	

Basics

```
1 users.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 193 entries, 0 to 192
Data columns (total 6 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   country                               193 non-null    object
1   beer_servings                         193 non-null    int64
2   spirit_servings                       193 non-null    int64
3   wine_servings                        193 non-null    int64
4   total_litres_of_pure_alcohol         193 non-null    float64
5   continent                             170 non-null    object
dtypes: float64(1), int64(3), object(2)
memory usage: 9.2+ KB
```

```
1 users.describe(include='all')
```

	country	beer_servings	spirit_servings	wine_servings	total_litres_of_pu
count	193	193.000000	193.000000	193.000000	

A3 Which continent drinks more beer on average?

```

continent
top    Afghanistan      India      India      India

```

```

1 cont=users.groupby('continent')['beer_servings'].sum()
2 cont

```

```

continent
AF      3258
AS      1630
EU      8720
OC      1435
SA      2101
Name: beer_servings, dtype: int64
75%      NaN      188.000000      128.000000      59.000000

```

```

1 cont[cont==cont.max()]

```

```

continent
EU      8720
Name: beer_servings, dtype: int64

```

A4 For each continent print the statistics for wine consumption.

```

1 contwine=users.groupby('continent')['wine_servings'].sum()
2 contwine

```

```

continent
AF      862
AS      399
EU     6400
OC      570
SA      749
Name: wine_servings, dtype: int64

```

A5 Print the mean alcohol consumption per continent for every column

```

1 contwine=users.groupby('continent')['total_litres_of_pure_alcohol'].mean()
2 contwine

```

```

continent
AF      3.007547
AS      2.170455
EU      8.617778
OC      3.381250
SA      6.308333
Name: total_litres_of_pure_alcohol, dtype: float64

```

A6 Print the median alcohol consumption per continent for every column

```
1 contwinealc = users.groupby('continent')['total_litres_of_pure_alcohol'].median()  
2 contwinealc
```

```
continent  
AF      2.30  
AS      1.20  
EU     10.00  
OC      1.75  
SA      6.85  
Name: total_litres_of_pure_alcohol, dtype: float64
```

A7 Print the mean, min and max values for spirit consumption.

```
1 round(users.spirit_servings.mean())
```

```
81
```

```
1 round(users.spirit_servings.min())
```

```
0
```

```
1 round(users.spirit_servings.max())
```

```
438
```