Lab Group Assignment – 2

Module Development on Statistical Functions

Instructions:
- In this assignment, you have to develop a module that contains the definition of the following statistical functions. You must name your module after your group id such as Group1, Group2 and so on.
- You have to also create a main.py file that imports your module and show the comparison of the output of your module with the output of Pandas statistical functions.
- You have to use the dataset you used for your Lab Group Assignment – 1.
- You must submit a report that shows the screenshot of the output of different function calls of your module and Pandas statistical functions. The report template is also provided.
- **You must implement the following Pandas Statistical Functions in your module.** Further, you have to define a few extra statistical functions, that you learnt from your course, which are not implemented in Pandas. The implementation of extra functions solely depends on you. You will choose what to define.

Sample Script:

Hello.py

```python
# Statistical Functions
# Mean
def mean(a):
    sum = 0;
    for i in a:
        sum+=i;
    n = len(a)
    return sum/n;
```

main.py

```python
#%%
import Hello
import pandas as pd
df = pd.read_csv('adult.data')
df.head()
df.columns = ['age','workclass','fnlwgt','education','education-
num','marital-status','occupation','relationship','race','sex','capital-
gain', 'capital-loss','hours-per-week','native-country', 'salary']

#%%
series1 = df['age']
print("Module Output:", "%.6f"%hello.mean(series1))
print("Pandas Output:", df[['age']].mean())
```

List of Statistical Functions must be implemented:

count	Number of non-NA values
describe	Compute set of summary statistics for Series or each DataFrame column
min, max	Compute minimum and maximum values
argmin, argmax	Compute index locations (integers) at which minimum or maximum value obtained, respectively
idxmin, idxmax	Compute index labels at which minimum or maximum value obtained, respectively
quantile	Compute sample quantile ranging from 0 to 1
sum	Sum of values
mean	Mean of values
median	Arithmetic median (50% quantile) of values
mad	Mean absolute deviation from mean value
prod	Product of all values
var	Sample variance of values
std	Sample standard deviation of values
skew	Sample skewness (third moment) of values
kurt	Sample kurtosis (fourth moment) of values
cumsum	Cumulative sum of values
cummin, cummax	Cumulative minimum or maximum of values, respectively
cumprod	Cumulative product of values
diff	Compute first arithmetic difference (useful for time series)
pct_change	Compute percent changes

Submission Link:

Google Classroom Assignment Submission.

Submission Deadline:

Sunday, 25 April 2021, by 11:50 AM (Before Sunday's Class).

Late Submission will be automatically graded as ZERO.