Multivariate Modeling

DATS 6450

Homework # 2

Using the Python program and “pandas”, “matplotlib.pyplot” ,“numpy” and “statsmodels” library perform the following tasks:

1. Load the time series data called tute1. The tute1 dataset is the same dataset used in LAB#1. Graph the scatter plot for Sales and GDP. (y-axis plot Sales and x-axis plot GDP). Add the appropriate x-label and y-label. Do not add any title in this step. This needs to be updated in step 7.
2. Graph the scatter plot for Sales and AdBudget. (y-axis plot Sales and x-axis plot AdBudget). Add the appropriate x-label and y-label. Do not add any title in this step. This needs to be updated in step 7.
3. Graph the scatter plot for GDP and AdBudget. (y-axis plot GDP and x-axis plot AdBudget). Add the appropriate x-label and y-label. Do not add any title in this step. This needs to be updated in step 7.
4. Call the function “correlation\_coefficent\_cal(x,y)” developed in the LAB#2 with y as the Sales data and the x as the GDP data. Save the correlation coefficient between these two variables as r\_xy. Display the following message on the console:

“The correlation coefficient between the Sales value and GDP is \_\_\_\_\_\_\_\_\_”.

Does the r\_xy value make sense with respect to the scatter plot graphed in step 7. Explain why?

1. Call the function “correlation\_coefficent\_cal(x,z)” developed in LAB#2 with x as the Sales data and the z as the AdBudget data. Save the correlation coefficient between these two variables as r\_yz. Display the following message on the console:

“The correlation coefficient between the Sales value and AdBudget is \_\_\_\_\_\_\_\_\_”.

Does the r\_yz value make sense with respect to the scatter plot graphed in step 8. Explain why?

1. Call the function “correlation\_coefficent\_cal(y,z)” developed in LAB#2 with y as the GDP data and the z as the AdBudget data. Save the correlation coefficient between these two variables as r\_yz. Display the following message on the console:

“The correlation coefficient between the GDP value and AdBudget is \_\_\_\_\_\_\_\_\_”.

Does the r\_yz value make sense with respect to the scatter plot graphed in step 8. Explain why?

1. Include the r\_xy, r\_yz and r\_xz in the title of the graphs developed in step 5 and 6. Write your code in a way that anytime r\_xy, r\_yz and r\_xz value changes it automatically updated on the figure title. Hint: you can use the following python command:

*plt.title("Scatter plot of GDP and Sales with r ={}".format(r\_xy))*

1. By looking at the correlation coefficients, write down your observation about the effect of AdBudget data and GDP data on the Sales revenue?

Be ready to upload the **solution report (as a single pdf**) plus **the .py file** through BB by the due date.