TIME SERIES MODELING & ANALYSIS

Instructor Name: Reza Jafari

HW#: 8

Submitted by Dinesh Kumar Padmanabhan

Date: 11-Nov-2020

ANSWERS TO ASKED QUESTIONS

```
 1 C:\ProgramData\Anaconda3\python.exe "C:\Program Files\
   JetBrains\PyCharm 2019.3.1\plugins\python\helpers\pydev\
   pydevconsole.py" --mode=client --port=56694
 2
 3 import sys; print('Python %s on %s' % (sys.version, sys.
   platform))
 4 sys.path.extend(['C:\\Users\\nsree_000\\Desktop\\Python-
   Quiz', 'C:/Users/nsree_000/Desktop/Python-Quiz'])
 5
 6 Python 3.7.4 (default, Aug  9 2019, 18:34:13) [MSC v.1915
   64 bit (AMD64)]
 7 Type 'copyright', 'credits' or 'license' for more
   information
 8 IPython 7.8.0 -- An enhanced Interactive Python. Type '?'
   for help.
 9 PyDev console: using IPython 7.8.0
10
11 Python 3.7.4 (default, Aug  9 2019, 18:34:13) [MSC v.1915
   64 bit (AMD64)] on win32
12 In[2]: runfile('C:/Users/nsree_000/Desktop/Python-Quiz/TIME
    SERIES/HW8_.py', wdir='C:/Users/nsree_000/Desktop/Python-
   Quiz/TIME SERIES')
13
14
15 The correlation coefficient between the Sales value and
   AdBudget is:0.91
16
17
18 The correlation coefficient between the AdBudget and GDP is
   :-0.77
19
20
21 The correlation coefficient between the Sales value and GDP
    is:-0.64
22
23
24 Degree of Freedom = 98
25 t0-value =  -11.903523644706398
26 Absolute value of test statistic did not exceed the
   critical t-value from the table.
27 The correlation coefficient (partial correlation)is
   statistically insignificant.
28
29
30 t0-value =  -8.15797948861032
31 Absolute value of test statistic did not exceed the
```

```
31 critical t-value from the table.
32 The correlation coefficient (partial correlation)is
   statistically insignificant.
33
34
35 partial coefficient between sales and adbudget for given
   gdp: 0.3470017952959567
36 Degree of Freedom = 97
37 t0-value =  3.6439927810863364
38 Absolute value of test statistic exceeded the critical t-
   value from the table.
39 The correlation coefficient (partial correlation) is
   statistically significant.
40
41
42 partial coefficient between sales and gdp for given
   Adbudget: 0.051357197844445375
43 t0-value =  0.5064781133341545
44 Absolute value of test statistic did not exceed the
   critical t-value from the table.
45 The correlation coefficient (partial correlation)is
   statistically insignificant.
46
47
48 partial coefficient between Adbudget and gdp for given
   sales: -0.35241022252661686
49 t0-value =  -3.7087721248175556
50 Absolute value of test statistic did not exceed the
   critical t-value from the table.
51 The correlation coefficient (partial correlation)is
   statistically insignificant.
52
53
54
```

```
#%%==============================================================================
# 1: Load the "tute1.csv" dataset.
# Write a python program that calculate the correlation coefficient between Sales
and AdBuget
# and display the following message on the console.
# "Correlation Coefficient between Sales and AdBugdet is _____"
# %%-----------------------------------------------------------------------------

The correlation coefficient between the Sales value and AdBudget is:0.91


#%%==============================================================================
# 2: Write a python program that calculate the correlation coefficient between
AdBuget and GDP
# and display the following message on the console:
# "Correlation Coefficient between AdBugdet and GDP is _____"
# %%-----------------------------------------------------------------------------

The correlation coefficient between the AdBudget and GDP is:-0.77



#%%==============================================================================
# 3: Write a python program that calculate coefficient between Sales and GDP and
display the
# following message on the console:
# "Correlation Coefficient between Sales and GDP is _____"
# %%-----------------------------------------------------------------------------

The correlation coefficient between the Sales value and GDP is:-0.64

#%%==============================================================================
# 4: Using the hypothesis test (t-test) show whether the correlation coefficients
in step 2, 3, and 4
# are statistically significant? Assume the level of confident to be 95% with two
tails ( α = 0.05 ).
# %%-----------------------------------------------------------------------------

Degree of Freedom = 98
t0-value =  -11.903523644706398
Absolute value of test statistic did not exceed the critical t-value from the
table. The correlation coefficient (partial correlation)is statistically
insignificant.

t0-value =  -8.15797948861032
Absolute value of test statistic did not exceed the critical t-value from the
table. The correlation coefficient (partial correlation)is statistically
insignificant.
```

```
#%%============================================================================
# 5: Write a python program that calculate the partial correlation coefficient
between Sales and AdBudegt.
# Using the hypothesis test, step 5, shows whether the derived coefficient is
statistically significant.
# Write down your observation. Hint: Partial correlation coefficient
# between variable A and B with confounding variable C can be calculated as :
# %%--------------------------------------------------------------------------

partial coefficient between sales and adbudget for given gdp: 0.3470017952959567

Degree of Freedom = 97
t0-value =  3.6439927810863364

Absolute value of test statistic exceeded the critical t-value from the table.
The correlation coefficient (partial correlation) is statistically significant.


#%%============================================================================
# 6: Write a python program that calculate the partial correlation coefficient
between Sales and
# GDP. Using the hypothesis test, shows whether this coefficient is statistically
significant. Write
# down your observation. The t-value can be calculated as follow. The critical t-
value can be found
# from the t-table.
# Where r is the partial correlation coefficient and n is the number of
observations.
# %%--------------------------------------------------------------------------

partial coefficient between sales and gdp for given Adbudget: 0.051357197844445375

t0-value =  0.5064781133341545

Absolute value of test statistic did not exceed the critical t-value from the
table. The correlation coefficient (partial correlation)is statistically
insignificant.


#%%============================================================================
# 7:Write a python program that calculate the partial correlation coefficient
between AdBudegt and GDP.
# Using the hypothesis test, shows whether this coefficient is statistically
significant.
# Write down your observation.
# %%--------------------------------------------------------------------------

partial coefficient between Adbudget and gdp for given sales: -0.35241022252661686

t0-value =  -3.7087721248175556

Absolute value of test statistic did not exceed the critical t-value from the
table. The correlation coefficient (partial correlation)is statistically
insignificant.
```

```
#%%=============================================================================
# 8: Create a table and place all the results from step 2 through 8 inside the
table.
# Compare the correlation coefficients and partial correlation coefficients for
(Sales, AdBudget), (Sales, GDP)
# and (AdBudegt, GDP). Write down your observation.
# %%---------------------------------------------------------------------------
```

attributes	correlation_coefficients	partial_coefficients	t-value
Adbudget & gdp	-0.77		-11.1
sales & gdp	-0.64		-8.15
sales & AdBudget		0.35	3.64
sales & gdp		0.05	0.51
Adbudget & gdp		-0.35	-3.71

```
Negative values indicating a negative relationship between the sales value and GDP
and Adbudget & gdp respectively. The same goes with respect to t-value indicating
Absolute value of test statistic did not exceed the critical t-value from the
table. The correlation coefficient (partial correlation)is statistically
insignificant.

With respect to partial coefficients the absolute value of test statistic exceeded
the critical t-value from the table. The correlation coefficient (partial
correlation)is statistically significant for (Sales, AdBudget) however which is
not the case for  (AdBudegt, GDP) (Sales, GDP).
```

```
#%%=============================================================================
# 9: If you must drop one of the predictors (AdBudegt or GDP) which predictor do
you pick for
# elimination? You need to justify your answer using the results above.
# %%---------------------------------------------------------------------------
```

```
The test assumes that both samples Sales and Adbudget were drawn from a Gaussian
distribution however not gdp. In my opinion I would drop GDP since the correlation
coefficient (partial correlation) is statistically significant for sales and
adbudget.
```

APPENDIX

```python
#Homework 8 : Partial Correlation Coefficient

import pandas as pd
import numpy as np
from CorrelationCoefficient import *
from scipy.stats import ttest_ind
print('\n')


df = pd.read_csv("tute1.csv")
s = df["Sales"]
ad = df["AdBudget"]
r_s_ad = (correlation_Coefficient_cal(s, ad))
print('The correlation coefficient between the Sales value and AdBudget
is:{0:.2f}'.format(r_s_ad))
print('\n')

g = df["GDP"]
r_ad_g = (correlation_Coefficient_cal(ad, g))
print('The correlation coefficient between the AdBudget and GDP
is:{0:.2f}'.format(r_ad_g))
print('\n')

r_s_g = (correlation_Coefficient_cal(s, g))
print('The correlation coefficient between the Sales value and GDP
is:{0:.2f}'.format(r_s_g))
print('\n')

n = df.shape[0]
degreef = n-2

t0 = r_ad_g*np.sqrt((n-2)/(1-r_ad_g**2))
print("Degree of Freedom =",degreef)
print("t0-value = ",t0)
print("Absolute value of test statistic exceeded the critical t-value from the
table. \n"
"The correlation coefficient (partial correlation)is statistically significant.")
print('\n')


t0 = r_s_g*np.sqrt((n-2)/(1-r_s_g**2))
print("t0-value = ",t0)
print("Absolute value of test statistic exceeded the critical t-value from the
table.\n"
"The correlation coefficient (partial correlation)is statistically significant.")
print('\n')



n = df.shape[0]
k = 1
degreef = n-2-k

r_sad_g = (r_s_ad - r_s_g*r_ad_g)/np.sqrt(1-r_s_g**2)*np.sqrt(1-r_ad_g**2)
```

```python
print("partial coefficient between sales and adbudget for given gdp:",r_sad_g)
print("Degree of Freedom =",degreef)
t0 = r_sad_g*np.sqrt((n-2-k)/(1-r_sad_g**2))
print("t0-value = ",t0)
print("Absolute value of test statistic did not exceed the critical t-value from
the table.\n"
"The correlation coefficient (partial correlation) is statistically
insignificant.")
print('\n')




r_sg_ad = (r_s_g - r_s_ad*r_ad_g)/np.sqrt(1-r_s_g**2)*np.sqrt(1-r_ad_g**2)
print("partial coefficient between sales and gdp for given Adbudget:",r_sg_ad)

t0 = r_sg_ad*np.sqrt((n-2-k)/(1-r_sg_ad**2))
print("t0-value = ",t0)
print("Absolute value of test statistic exceeded the critical t-value from the
table.\n"
"The correlation coefficient (partial correlation)is statistically significant.")
print('\n')




r_adg_s = (r_ad_g - r_s_ad*r_s_g)/np.sqrt(1-r_s_ad**2)*np.sqrt(1-r_s_g**2)
print("partial coefficient between Adbudget and gdp for given sales:",r_adg_s)

t0 = r_adg_s*np.sqrt((n-2-k)/(1-r_adg_s**2))
print("t0-value = ",t0)
print("Absolute value of test statistic exceeded the critical t-value from the
table.\n"
"The correlation coefficient (partial correlation)is statistically significant.")
print('\n')
```