PY0101EN_Coursera_FinalAssignment (1)

January 17, 2021

Analyzing US Economic Data and Building a Dashboard

Description

Extracting essential data from a dataset and displaying it is a necessary part of data science; therefore individuals can make correct decisions based on the data. In this assignment, you will extract some essential economic indicators from some data, you will then display these economic indicators in a Dashboard. You can then share the dashboard via an URL.

Gross domestic product (GDP) is a measure of the market value of all the final goods and services produced in a period. GDP is an indicator of how well the economy is doing. A drop in GDP indicates the economy is producing less; similarly an increase in GDP suggests the economy is performing better. In this lab, you will examine how changes in GDP impact the unemployment rate. You will take screen shots of every step, you will share the notebook and the URL pointing to the dashboard.

Table of Contents

```
<ul>
    <li><a href="#Section_1"> Define a Function that Makes a Dashboard </a></li>
<li><a href="#Section_2">Question 1: Create a dataframe that contains the GDP data and display
<li><a href="#Section_3">Question 2: Create a dataframe that contains the unemployment data an
<li><a href="#Section_4">Question 3: Display a dataframe where unemployment was greater than 8
<li><a href="#Section_5">Question 4: Use the function make_dashboard to make a dashboard</a></
    <li><a href="#Section_6"><b>(Optional not marked)</b> Save the dashboard on IBM cloud and
</ul>
```

Estimated Time Needed: 180 min

Define Function that Makes a Dashboard

We will import the following libraries.

[1]:
```
import pandas as pd
from bokeh.plotting import figure, output_file, show,output_notebook
output_notebook()
```

In this section, we define the function make_dashboard. You don't have to know how the function works, you should only care about the inputs. The function will produce a dashboard as well as an html file. You can then use this html file to share your dashboard. If you do not know what an html file is don't worry everything you need to know will be provided in the lab.

[2]:
```
def make_dashboard(x, gdp_change, unemployment, title, file_name):
    output_file(file_name)
    p = figure(title=title, x_axis_label='year', y_axis_label='%')
    p.line(x.squeeze(), gdp_change.squeeze(), color="firebrick", line_width=4,␣
 ↪legend="% GDP change")
    p.line(x.squeeze(), unemployment.squeeze(), line_width=4, legend="%␣
 ↪unemployed")
    show(p)
```

The dictionary links contain the CSV files with all the data. The value for the key GDP is the file that contains the GDP data. The value for the key unemployment contains the unemployment data.

[3]:
```
links={'GDP':'https://cf-courses-data.s3.us.cloud-object-storage.appdomain.
 ↪cloud/IBMDeveloperSkillsNetwork-PY0101EN-SkillsNetwork/labs/
 ↪FinalModule_Coursera/data/clean_gdp.csv',\
       'unemployment':'https://cf-courses-data.s3.us.cloud-object-storage.
 ↪appdomain.cloud/IBMDeveloperSkillsNetwork-PY0101EN-SkillsNetwork/labs/
 ↪FinalModule_Coursera/data/clean_unemployment.csv'}
```

Question 1: Create a dataframe that contains the GDP data and display the first five rows of the dataframe.

Use the dictionary links and the function pd.read_csv to create a Pandas dataframes that contains the GDP data.

Hint: links[“GDP”] contains the path or name of the file.

[4]:
```
# Type your code here
df = pd.read_csv(links["GDP"])
```

Use the method head() to display the first five rows of the GDP data, then take a screen-shot.

[5]:
```
# Type your code here
df.head()
```

[5]:

	date	level-current	level-chained	change-current	change-chained
0	1948	274.8	2020.0	-0.7	-0.6
1	1949	272.8	2008.9	10.0	8.7
2	1950	300.2	2184.0	15.7	8.0
3	1951	347.3	2360.0	5.9	4.1
4	1952	367.7	2456.1	6.0	4.7

Question 2: Create a dataframe that contains the unemployment data. Display the first five rows of the dataframe.

Use the dictionary links and the function pd.read_csv to create a Pandas dataframes that contains the unemployment data.

[6]:
```
# Type your code here
# 'https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/
 ↪IBMDeveloperSkillsNetwork-PY0101EN-SkillsNetwork/labs/FinalModule_Coursera/
 ↪data/clean_unemployment.csv'

df_employement = pd.read_csv(links["unemployment"])
```

Use the method head() to display the first five rows of the unemployment data, then take a screen-shot.

[7]:
```
# Type your code here
df_employement.head()
```

[7]:

	date	unemployment
0	1948	3.750000
1	1949	6.050000
2	1950	5.208333
3	1951	3.283333
4	1952	3.025000

Question 3: Display a dataframe where unemployment was greater than 8.5%. Take a screen-shot.

[8]:
```
# Type your code here
df_employement[df_employement['unemployment']>8.5]
```

[8]:

	date	unemployment
34	1982	9.708333
35	1983	9.600000
61	2009	9.283333
62	2010	9.608333
63	2011	8.933333

Question 4: Use the function make_dashboard to make a dashboard

In this section, you will call the function make_dashboard , to produce a dashboard. We will use the convention of giving each variable the same name as the function parameter.

Create a new dataframe with the column 'date' called x from the dataframe that contains the GDP data.

[9]:
```
x = df['date'] # Create your dataframe with column date
```

Create a new dataframe with the column 'change-current' called gdp_change from the dataframe that contains the GDP data.

[10]:
```
gdp_change = df['change-current'] # Create your dataframe with column␣
 ↪change-current
```

Create a new dataframe with the column 'unemployment' called unemployment from the dataframe that contains the unemployment data.

[11]:
```
unemployment = df_employement['unemployment'] # Create your dataframe with␣
↪column unemployment
```

Give your dashboard a string title, and assign it to the variable title

[12]:
```
title = 'Analyzing US Economic Data and Building a Dashboard'# Give your␣
↪dashboard a string title
```

Finally, the function make_dashboard will output an .html in your direictory, just like a csv file. The name of the file is "index.html" and it will be stored in the varable file_name.

[13]:
```
file_name = "index.html"
```

Call the function make_dashboard , to produce a dashboard. Assign the parameter values accordingly take a the , take a screen shot of the dashboard and submit it.

[15]:
```
# Fill up the parameters in the following function:
# make_dashboard(x=, gdp_change=, unemployment=, title=, file_name=)
make_dashboard(x=df['date'], gdp_change=df['change-current'],
               unemployment=df_employement['unemployment'],
               title='Analyzing US Economic Data and Building a Dashboard',
               file_name="index.html")
```

```
BokehDeprecationWarning: 'legend' keyword is deprecated, use explicit
'legend_label', 'legend_field', or 'legend_group' keywords instead
BokehDeprecationWarning: 'legend' keyword is deprecated, use explicit
'legend_label', 'legend_field', or 'legend_group' keywords instead
```

How to submit

Once you complete your notebook you will have to share it to be marked. Select the icon on the top right a marked in red in the image below, a dialogue box should open, select the option all content excluding sensitive code cells.

You can then share the notebook via a URL by scrolling down as shown in the following image:

About the Authors:

Joseph Santarcangelo has a PhD in Electrical Engineering, his research focused on using machine learning, signal processing, and computer vision to determine how videos impact human cognition. Joseph has been working for IBM since he completed his PhD.

Other contributors: Yi leng Yao, Mavis Zhou

References :

1) Economic Research at the St. Louis Fed : Civilian Unemployment Rate

2) Data Packaged Core Datasets

0.1 Change Log

Date (YYYY-MM-DD)	Version	Changed By	Change Description
2020-11-10	1.1	Malika Singla	Deleted the Optional part
2020-08-27	1.0	Malika Singla	Added lab to GitLab

##