

Analyzing US Economic Data and Building a Dashboard

- This study was a part of Coursera IBM Data Science Course. I had firstly made this study on January 22, 2020. Now I am making same study again to publish it on GitHub. All information including data were retrieved from the sources of that course.
- In order to realize this study, I used IBM Cloud. IBM provides free cloud accounts to make this kind of studies. By using my IBM Cloud account I created a new project and made this project in a notebook in IBM Cloud.
- This small project was prepared to show how to build a dashboard. For this project I used US Gross domestic product (GDP) and Unemployment data those were provided by IBM data science course. The data include information between the years 1948 and 2016.
- Here is some information from the course source:

" 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. "

In [1]:

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

(<https://bokeh.org>)2 successfully loaded.

In [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)
```

In [3]:

```
links={'GDP': 'https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/PY0101EN/projects/coursera_project/clean_gdp.csv', \
      'unemployment': 'https://s3-api.us-geo.objectstorage.softlayer.net/cf-courses-data/CognitiveClass/PY0101EN/projects/coursera_project/clean_unemployment.csv'}
```

In [4]:

```
csv_path='links[GDP]'
df=pd.read_csv(links["GDP"])
```

In [5]:

```
df.head()
```

Out[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

In [6]:

```
df.tail()
```

Out[6]:

	date	level-current	level-chained	change-current	change-chained
64	2012	16155.3	15354.6	3.6	1.8
65	2013	16691.5	15612.2	4.4	2.5
66	2014	17427.6	16013.3	4.0	2.9
67	2015	18120.7	16471.5	2.7	1.6
68	2016	18624.5	16716.2	4.2	2.2

In [7]:

```
csv_path='links[unemployment]'  
df=pd.read_csv(links["unemployment"])
```

In [8]:

```
df.head()
```

Out[8]:

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

In [9]:

```
df.tail()
```

Out[9]:

	date	unemployment
64	2012	8.075000
65	2013	7.358333
66	2014	6.158333
67	2015	5.275000
68	2016	4.875000

In [10]:

```
df = pd.read_csv(links['GDP'])  
x = df["date"]  
x.head()
```

Out[10]:

```
0    1948  
1    1949  
2    1950  
3    1951  
4    1952  
Name: date, dtype: int64
```

In [11]:

```
gdp_change = df['change-current']  
gdp_change.head()
```

Out[11]:

```
0    -0.7  
1    10.0  
2    15.7  
3     5.9  
4     6.0  
Name: change-current, dtype: float64
```

In [12]:

```
df = pd.read_csv(links['unemployment'])  
unemployment = df['unemployment']  
unemployment.head()
```

Out[12]:

```
0    3.750000  
1    6.050000  
2    5.208333  
3    3.283333  
4    3.025000  
Name: unemployment, dtype: float64
```

In [13]:

```
title = 'US Economic Data Analysis'
```

In [14]:

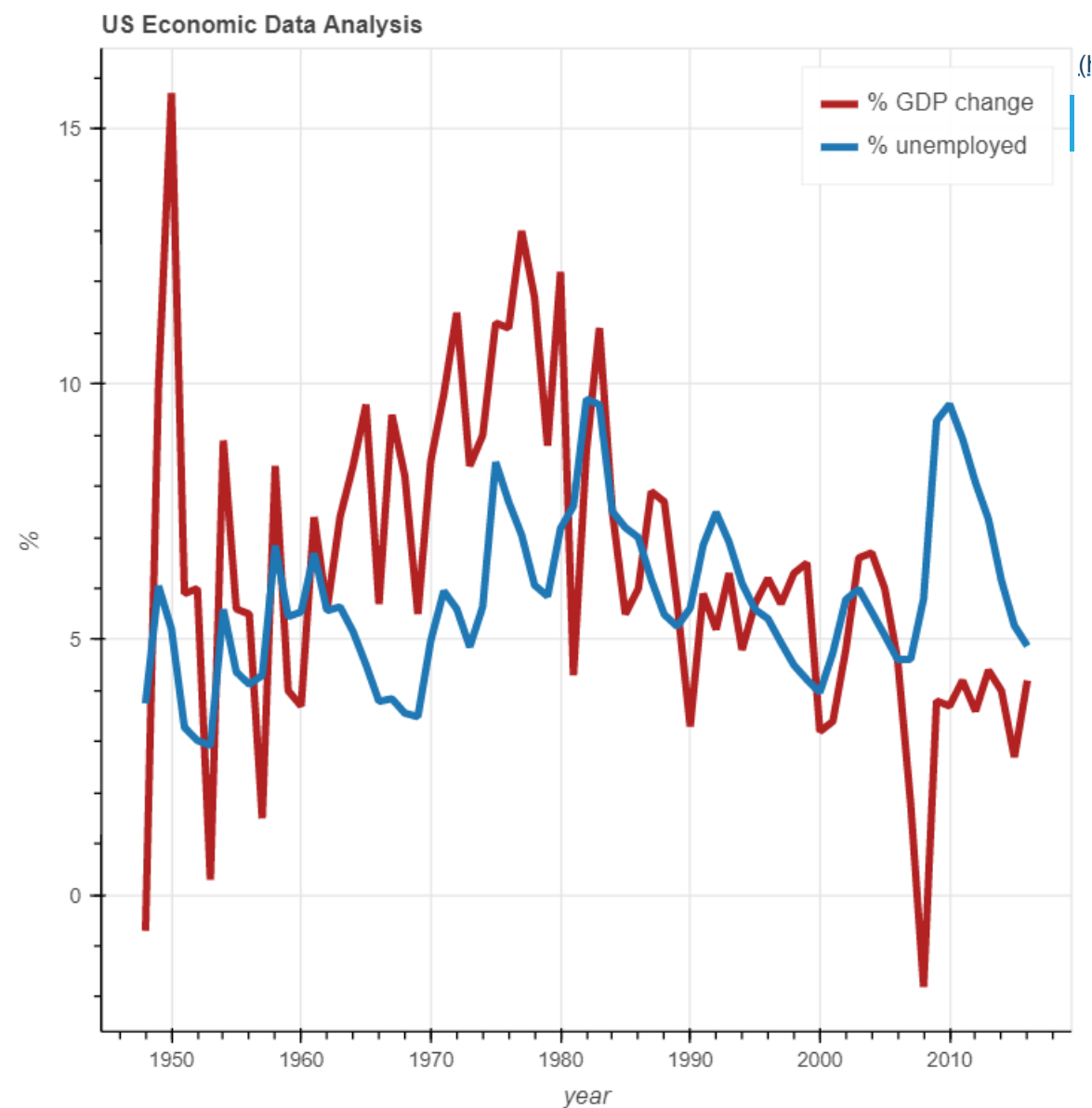
```
file_name = "index.html"
```

In [15]:

```
make_dashboard(x=x, gdp_change=gdp_change, unemployment=unemployment, title=title, file_name=file_name)
```

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



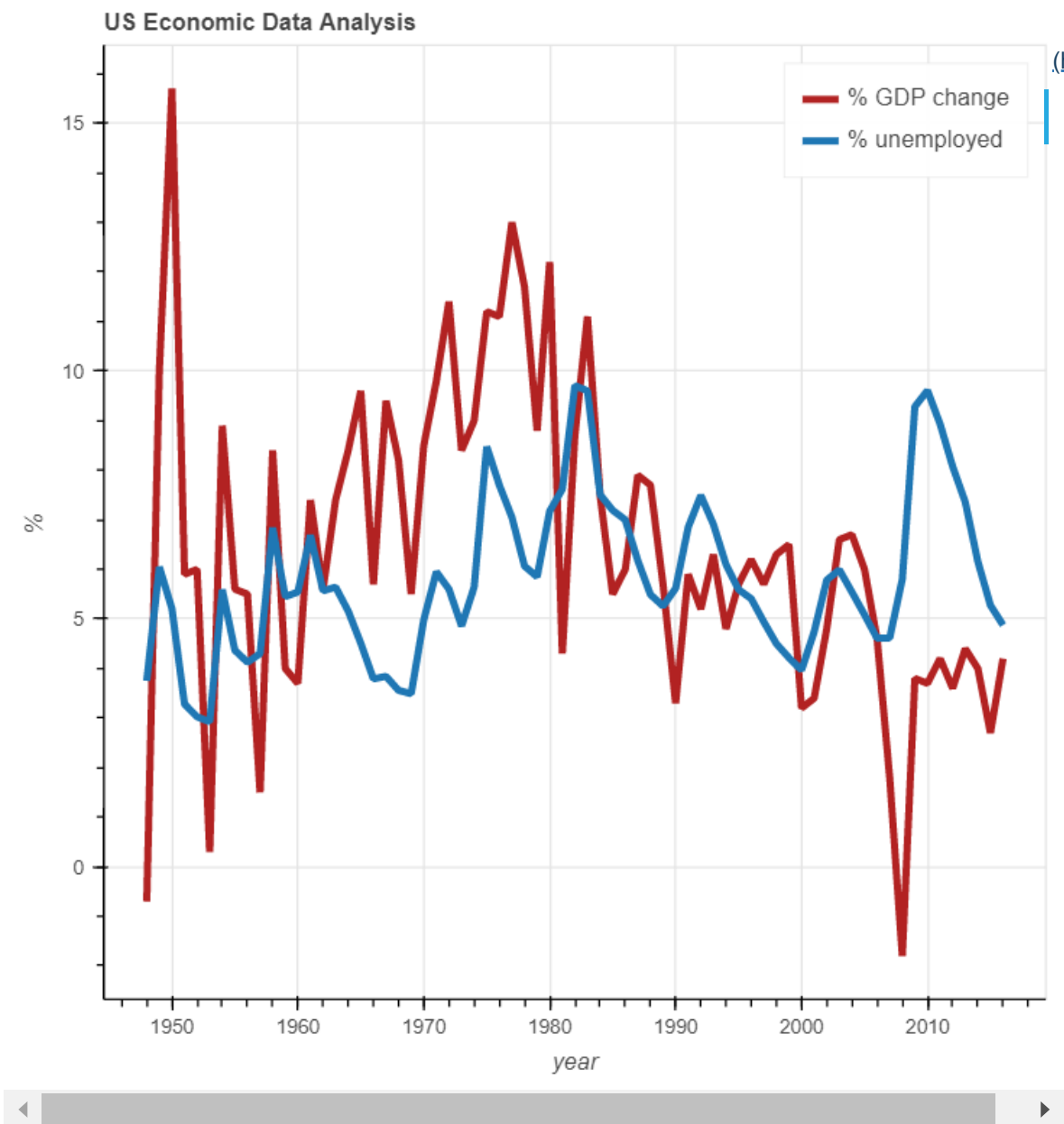
In []:

In [22]:

```
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_label="% GDP change")  
    p.line(x.squeeze(), unemployment.squeeze(), line_width=4, legend_label="% unemployed")  
    show(p)
```

In [23]:

```
make_dashboard(x=x, gdp_change=gdp_change, unemployment=unemployment, title=title, file_name=file_name)
```



In []: