[1]:
```
import pandas as pd
```

[5]:
```
import numpy as np
```

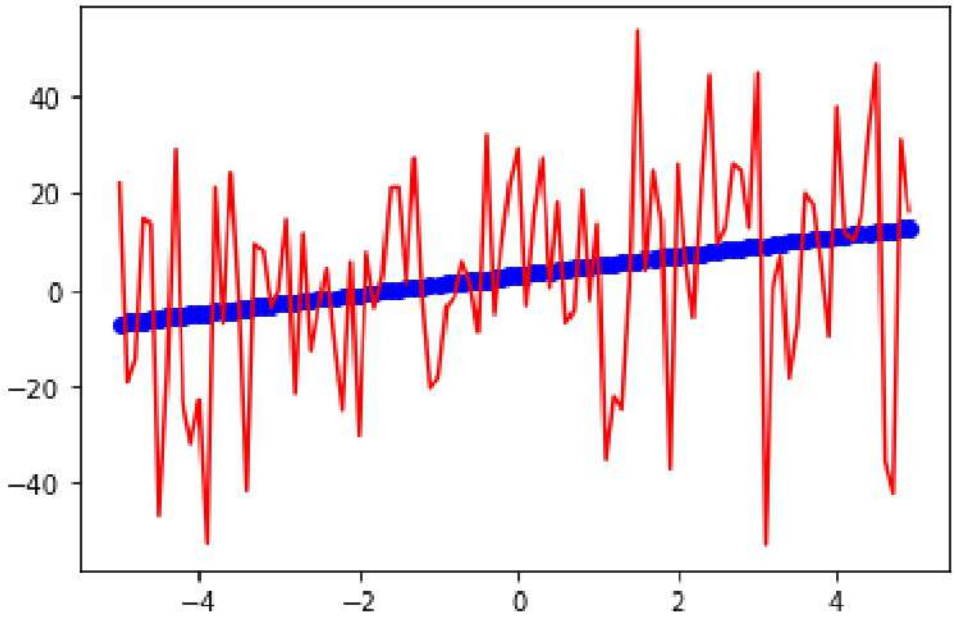
[3]:
```
import matplotlib.pyplot as plt
```

[4]:
```
%matplotlib inline
```

x = np.arange(-5.0, 5.0, 0.1)

##You can adjust the slope and intercept to verify the changes in the graph y = 2*(x) + 3 y_noise = 2* np.random.normal(size=x.size) ydata = y + y_noise #plt.figure(figsize=(8,6)) plt.plot(x, ydata, 'bo') plt.plot(x,y, 'r') plt.ylabel('Dependent Variable') plt.xlabel('Indepdendent Variable') plt.show()

[13]:
```
y= 2*(x) + 3
noise = 20*np.random.normal(size=y.size)
y_data= y + noise
plt.plot(x,y,'bo')
plt.plot(x,y_data,'r')
plt.show()
```



[14]:
```
!wget -nv -O china_gdp.csv https://s3-api.us-geo.objectstorage.softlayer.net/
↪cf-courses-data/CognitiveClass/ML0101ENv3/labs/china_gdp.csv
```

```
2019-09-22 06:26:13 URL:https://s3-api.us-geo.objectstorage.softlayer.net/cf-
courses-data/CognitiveClass/ML0101ENv3/labs/china_gdp.csv [1218/1218] ->
"china_gdp.csv" [1]
```

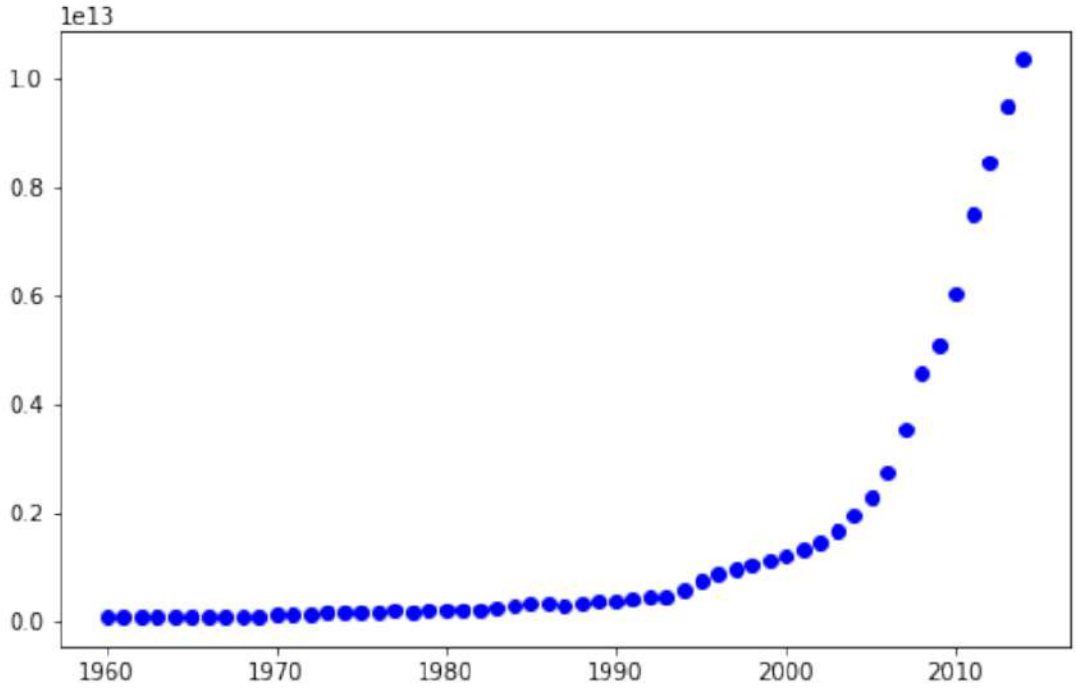
[15]:
```
df = pd.read_csv("china_gdp.csv")
df.head(9)
```

[15]:

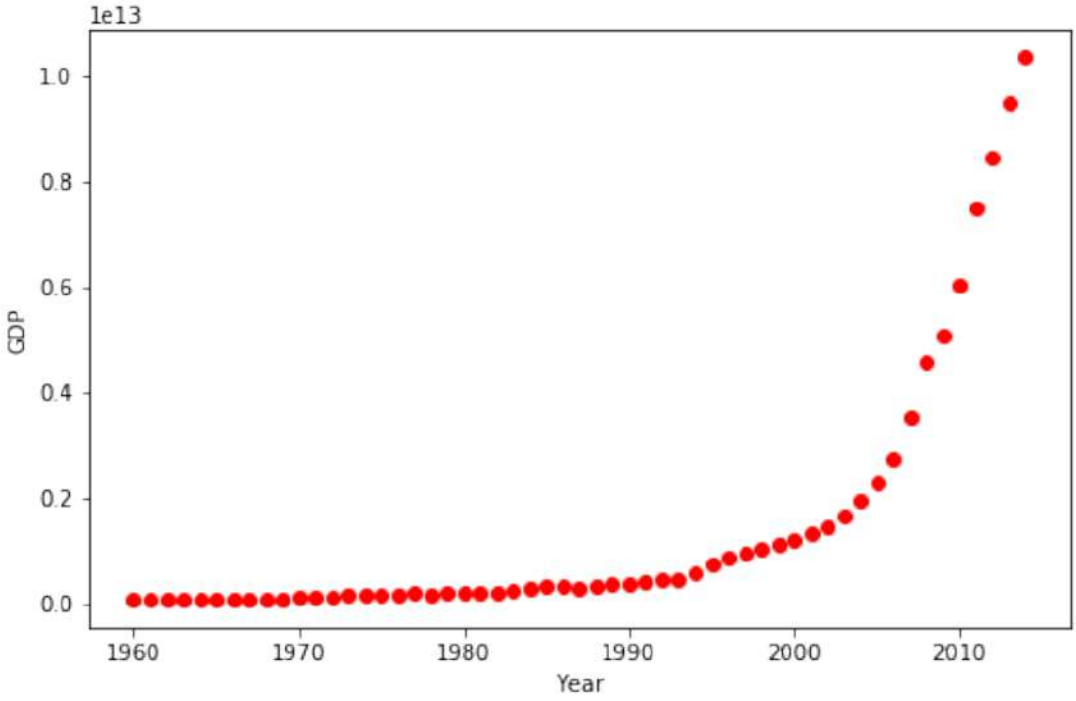
	Year	Value
0	1960	5.918412e+10
1	1961	4.955705e+10
2	1962	4.668518e+10
3	1963	5.009730e+10
4	1964	5.906225e+10
5	1965	6.970915e+10
6	1966	7.587943e+10
7	1967	7.205703e+10
8	1968	6.999350e+10

[19]:
```
plt.figure(figsize=(8,5))
plt.scatter(df.Year, df.Value, color='blue')
```

[19]:
```
<matplotlib.collections.PathCollection at 0x7f1318587e80>
```



[18]:
```
plt.figure(figsize=(8,5))
x_data, y_data = (df["Year"].values, df["Value"].values)
plt.plot(x_data, y_data, 'ro')
plt.ylabel('GDP')
plt.xlabel('Year')
plt.show()
```



[22]:
```
def sigmoid(data, beta_1, beta_2):
    y= 1/(1+np.exp(-beta_1*(data-beta_2)))
    return y
```

[23]:
```
x_data= df['Year'].values
```

[24]:
```
x_data.head()
```

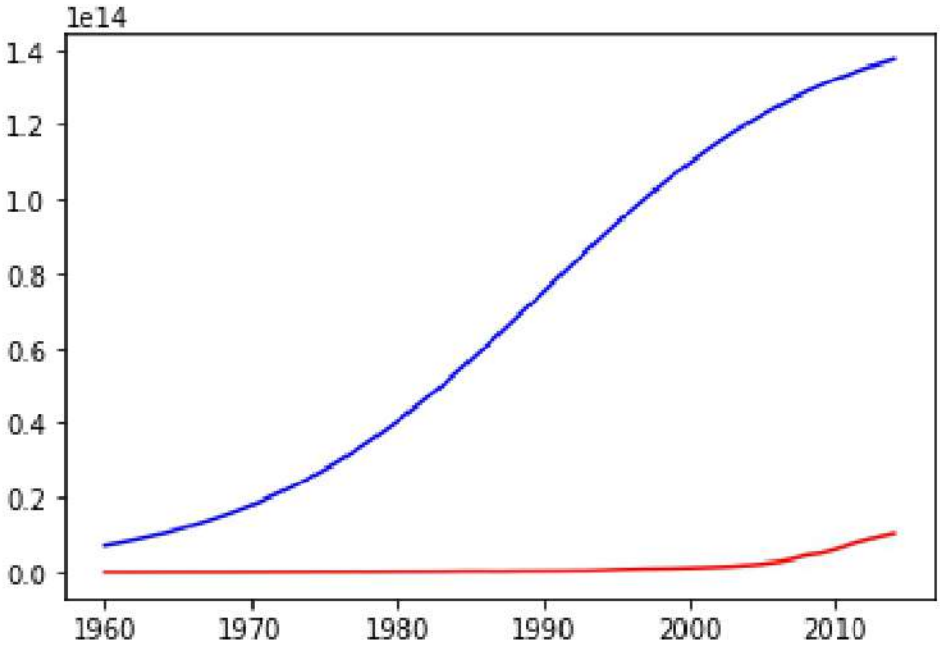
```
␣
↪---------------------------------------------------------------------------

       AttributeError                            Traceback (most recent call␣
↪last)

       <ipython-input-24-864c1a255ed1> in <module>
   ----> 1 x_data.head()


       AttributeError: 'numpy.ndarray' object has no attribute 'head'
```

[25]:
```
x_data
```

[25]:
```
array([1960, 1961, 1962, 1963, 1964, 1965, 1966, 1967, 1968, 1969, 1970,
       1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979, 1980, 1981,
       1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992,
       1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003,
       2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014])
```

[27]:
```
y_data = df['Value'].values
beta_1= 0.10
beta_2= 1990.0
y_pred= sigmoid(x_data, beta_1, beta_2)
plt.plot(x_data, y_pred*150000000000000, color='blue')
plt.plot(x_data, y_data, color='red')
```

[27]:
```
[<matplotlib.lines.Line2D at 0x7f13184b2208>]
```



[28]:
```
x_data= x_data/max(x_data)
y_data= y_data/max(y_data)
```

[29]:
```
x_data
```

[29]:
```
array([0.97318769, 0.97368421, 0.97418073, 0.97467726, 0.97517378,
       0.97567031, 0.97616683, 0.97666336, 0.97715988, 0.97765641,
       0.97815293, 0.97864945, 0.97914598, 0.9796425 , 0.98013903,
       0.98063555, 0.98113208, 0.9816286 , 0.98212512, 0.98262165,
       0.98311817, 0.9836147 , 0.98411122, 0.98460775, 0.98510427,
```

```
       0.98560079, 0.98609732, 0.98659384, 0.98709037, 0.98758689,
       0.98808342, 0.98857994, 0.98907646, 0.98957299, 0.99006951,
       0.99056604, 0.99106256, 0.99155909, 0.99205561, 0.99255214,
       0.99304866, 0.99354518, 0.99404171, 0.99453823, 0.99503476,
       0.99553128, 0.99602781, 0.99652433, 0.99702085, 0.99751738,
       0.9980139 , 0.99851043, 0.99900695, 0.99950348, 1.        ])
```

[30]:
```
from scipy.optimize import curve_fit
```

[31]:
```
popt, pcov= curve_fit(sigmoid, x_data, y_data)
```

[32]:
```
popt
```

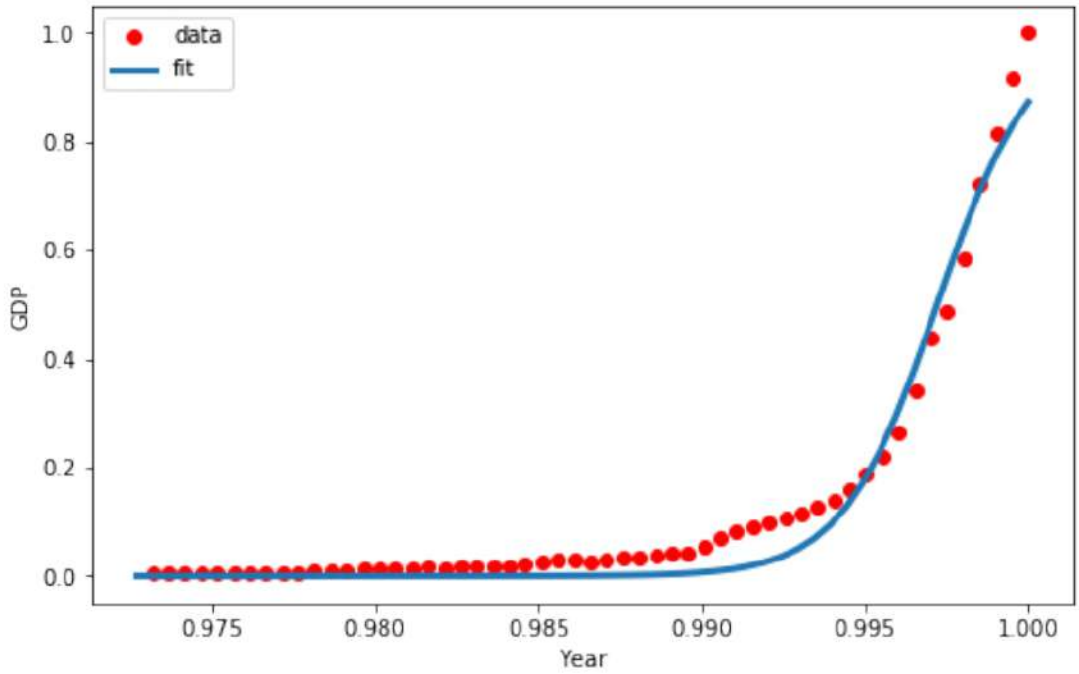
[32]:
```
array([690.4475266 ,   0.99720712])
```

[33]:
```
pcov
```

[33]:
```
array([[ 1.52263822e+03, -2.88128654e-04],
       [-2.88128654e-04,  7.25977514e-09]])
```

[41]:
```
x = np.linspace(1960, 2015, 55)
plt.figure(figsize=(8,5))
x=x/max(x)
y = sigmoid(x, popt[0], popt[1])
plt.plot(x_data, y_data, 'ro', label='data')
plt.plot(x,y, linewidth=3.0, label='fit')
plt.legend(loc='bottomleft')
plt.ylabel('GDP')
plt.xlabel('Year')
plt.show()
```

```
/home/jupyterlab/conda/envs/python/lib/python3.6/site-
packages/ipykernel_launcher.py:7: MatplotlibDeprecationWarning: Unrecognized
location 'bottomleft'. Falling back on 'best'; valid locations are
        best
        upper right
        upper left
        lower left
        lower right
        right
        center left
        center right
        lower center
        upper center
        center
This will raise an exception in 3.3.
  import sys
```


data
fit
1.0
0.8
0.6
0.4
0.2
0.0
GDP
0.975
0.980
0.985
0.990
0.995
1.000
Year