Business & Data Understaning

MANAI MOHAMED MORTADHA (3GII-SSE)

```
import pandas as pd
df=pd.read_csv("/content/AirPassengers.csv")
df.head()
```

	Month	#Passengers
0	1949-01	112
1	1949-02	118
2	1949-03	132
3	1949-04	129
4	1949-05	121

```
df.tail()
```

	Month	#Passengers
139	1960-08	606
140	1960-09	508
141	1960-10	461
142	1960-11	390
143	1960-12	432

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 144 entries, 0 to 143
Data columns (total 2 columns):
 #   Column       Non-Null Count  Dtype 
---  ------       --------------  ----- 
 0   Month        144 non-null    object
 1   #Passengers  144 non-null    int64 
dtypes: int64(1), object(1)
memory usage: 2.4+ KB
```

This dataset include history of 12 years from **1949-01** to **1960-12**

```
import matplotlib.pyplot as plt
plt.xlabel('Months')
plt.ylabel('Number of Passengers')
plt.title('Distribution of the number of passengers over 12 years')
plt.plot(df['#Passengers'])
plt.show()
```

Distribution of the number of passengers over 12 years

Data Preparation

```
def load_data(data, seq_len):
  X = []
  y = []
  for i in range(seq_len, len(data)):
    X.append(data.iloc[i-seq_len : i, 1])
    y.append(data.iloc[i,1])
  return X,y
```

```
X,y=load_data(df,20)
```

```
len(y)
```

```
124
```

```
X_train=X[:100]
y_train=y[:100]
X_test=X[100:]
y_test=y[100:]
```

```
import numpy as np
X_train=np.array(X_train)
y_train=np.array(y_train)
X_test=np.array(X_test)
y_test=np.array(y_test)
```

```
X_train
```

```
array([[112, 118, 132, ..., 149, 170, 170],
       [118, 132, 129, ..., 170, 170, 158],
       [132, 129, 121, ..., 170, 158, 133],
       ...,
       [301, 356, 348, ..., 491, 505, 404],
       [356, 348, 355, ..., 505, 404, 359],
       [348, 355, 422, ..., 404, 359, 310]])
```

```
print('x_train.shape = ',X_train.shape)
print('y_train.shape = ', y_train.shape)
print('x_test.shape = ', X_test.shape)
print('y_test.shape = ', y_test.shape)
```

```
x_train.shape =  (100, 20)
y_train.shape =  (100,)
x_test.shape =  (24, 20)
y_test.shape =  (24,)
```

```
X_train = np.reshape(X_train, (100, 20, 1))
X_test = np.reshape(X_test, (24, 20, 1))
```

```
print('x_train.shape = ',X_train.shape)
print('y_train.shape = ', y_train.shape)
print('x_test.shape = ', X_test.shape)
print('y_test.shape = ', y_test.shape)
```

```
x_train.shape =  (100, 20, 1)
y_train.shape =  (100,)
x_test.shape =  (24, 20, 1)
y_test.shape =  (24,)
```

Neural Network from Keras

```
from keras.models import Sequential
from keras.layers import Dense
```

```
model = Sequential()
model.add(Dense(40, input_shape=(20,),activation='relu'))
model.add(Dense(40,activation='relu'))
model.add(Dense(1))
```

```
model.compile(loss='mean_squared_error', optimizer='adam')
model.fit(X_train, y_train, epochs=500, batch_size=2, verbose=2)
```

```
50/50 - 0s - loss: 53.7200 - 104ms/epoch - 2ms/step
Epoch 473/500
50/50 - 0s - loss: 60.3407 - 108ms/epoch - 2ms/step
Epoch 474/500
50/50 - 0s - loss: 56.6721 - 123ms/epoch - 2ms/step
Epoch 475/500
50/50 - 0s - loss: 89.6014 - 117ms/epoch - 2ms/step
Epoch 476/500
50/50 - 0s - loss: 59.4444 - 126ms/epoch - 3ms/step
Epoch 477/500
50/50 - 0s - loss: 62.7189 - 119ms/epoch - 2ms/step
Epoch 478/500
50/50 - 0s - loss: 61.8451 - 103ms/epoch - 2ms/step
Epoch 479/500
50/50 - 0s - loss: 66.5758 - 105ms/epoch - 2ms/step
Epoch 480/500
50/50 - 0s - loss: 57.3906 - 109ms/epoch - 2ms/step
Epoch 481/500
50/50 - 0s - loss: 83.2053 - 106ms/epoch - 2ms/step
Epoch 482/500
50/50 - 0s - loss: 114.6352 - 95ms/epoch - 2ms/step
Epoch 483/500
50/50 - 0s - loss: 96.4560 - 124ms/epoch - 2ms/step
Epoch 484/500
50/50 - 0s - loss: 117.0442 - 113ms/epoch - 2ms/step
Epoch 485/500
50/50 - 0s - loss: 82.1807 - 114ms/epoch - 2ms/step
Epoch 486/500
50/50 - 0s - loss: 86.4101 - 117ms/epoch - 2ms/step
Epoch 487/500
50/50 - 0s - loss: 74.5774 - 118ms/epoch - 2ms/step
Epoch 488/500
50/50 - 0s - loss: 63.9437 - 105ms/epoch - 2ms/step
Epoch 489/500
50/50 - 0s - loss: 86.3503 - 113ms/epoch - 2ms/step
Epoch 490/500
50/50 - 0s - loss: 139.2220 - 113ms/epoch - 2ms/step
Epoch 491/500
50/50 - 0s - loss: 296.7442 - 115ms/epoch - 2ms/step
Epoch 492/500
50/50 - 0s - loss: 140.0320 - 107ms/epoch - 2ms/step
Epoch 493/500
50/50 - 0s - loss: 156.4481 - 108ms/epoch - 2ms/step
Epoch 494/500
50/50 - 0s - loss: 98.1925 - 110ms/epoch - 2ms/step
Epoch 495/500
50/50 - 0s - loss: 88.9676 - 112ms/epoch - 2ms/step
Epoch 496/500
50/50 - 0s - loss: 80.3688 - 111ms/epoch - 2ms/step
Epoch 497/500
50/50 - 0s - loss: 131.3052 - 106ms/epoch - 2ms/step
Epoch 498/500
50/50 - 0s - loss: 89.6818 - 113ms/epoch - 2ms/step
Epoch 499/500
50/50 - 0s - loss: 95.3477 - 113ms/epoch - 2ms/step
Epoch 500/500
50/50 - 0s - loss: 70.0680 - 116ms/epoch - 2ms/step
<keras.src.callbacks.History at 0x7b88c1a7d660>
```

```
y_pred=model.predict(X_test)
```

```
1/1 [==============================] - 0s 89ms/step
```

```
y_pred
```

```
array([[351.3825 ],
       [334.73746],
       [374.88666],
       [375.75555],
       [396.2213 ],
       [489.9771 ],
       [540.4173 ],
       [549.36145],
       [450.62213],
       [389.89178],
       [336.34064],
       [383.15656],
       [413.22577],
       [401.8603 ],
       [457.02637],
       [422.505  ],
       [457.09988],
       [538.73566],
       [611.03424],
       [619.6796 ],
```

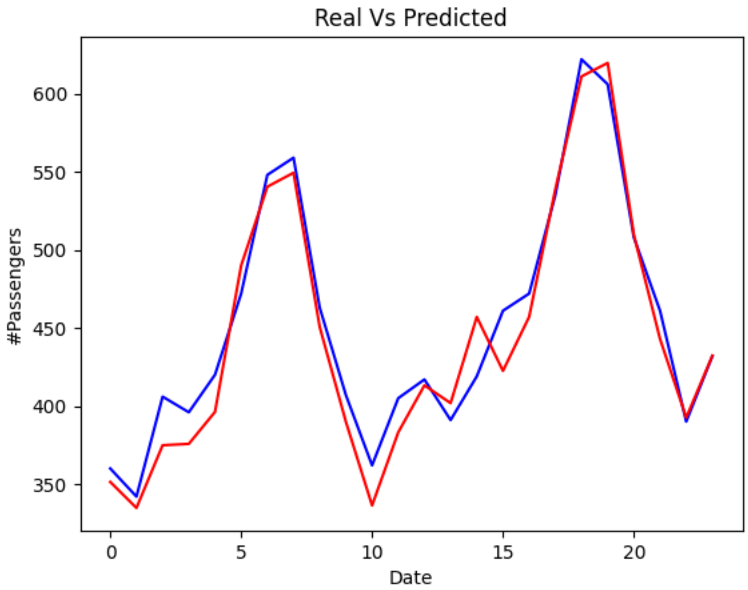
```
           [509.58765],
           [442.91708],
           [392.5771 ],
           [432.24    ]], dtype=float32)
```

```
y_test
```

```
    array([360, 342, 406, 396, 420, 472, 548, 559, 463, 407, 362, 405, 417,
           391, 419, 461, 472, 535, 622, 606, 508, 461, 390, 432])
```

```
plt.plot(y_test, color='blue',label='Real')
plt.plot(y_pred, color='red',label='Predicted')
plt.title('Real Vs Predicted')
plt.xlabel('Date')
plt.ylabel('#Passengers')
```

```
    Text(0, 0.5, '#Passengers')
```




```
from sklearn.metrics import mean_squared_error
rmse = np.sqrt(mean_squared_error(y_test, y_pred)).round(2)
mae = np.round(np.mean(np.abs(y_test-y_pred))*100,2)
mape = np.round(np.mean(np.abs(y_test-y_pred)/y_test)*100,2)
```

```
print("RMSE=",rmse)
print("MAE=",mae)
print("MAPE=",mape)
```

```
    RMSE= 18.33
    MAE= 8624.88
    MAPE= 18.89
```