Lab1_AML

January 7, 2024

```python
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import seaborn as sns
```

```python
x = np.array([1,2,4,8,16,32,64,128])
y = np.array([128,64,32,16,8,4,2,1])
```
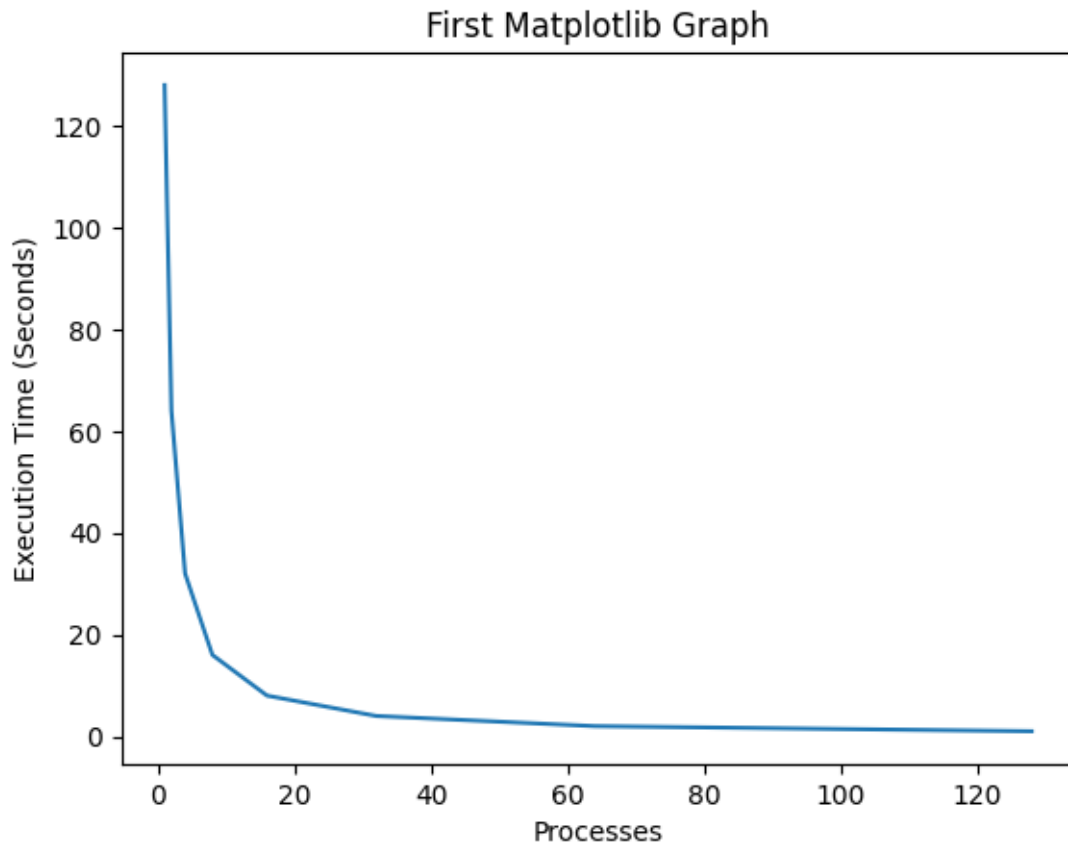
```python
print(x.dtype)
print(y.dtype)
```

```
int64
int64
```

```python
plt.plot(x,y)
plt.xlabel("Processes")
plt.ylabel("Execution Time (Seconds)")
plt.title("First Matplotlib Graph")
```

```
Text(0.5, 1.0, 'First Matplotlib Graph')
```

First Matplotlib Graph


```python
df = pd.read_csv("G11_prices.csv")
df
```

```
      area  bedrooms  age  price
0     1125         4    2   2.80
1     1000         3    2   2.30
2     1575         5    5   3.15
3     1125         5    1   2.92
4      900         4    8   1.85
5     1125         5    3   2.85
6     1125         4    7   1.60
7     1125         4    5   2.10
8     1800         5    4   3.00
9     4500         8    3   7.00
10    1800         5    4   2.80
11    3204         6    3   5.60
12    3204         6    3   5.96
13    1575         6    3   3.40
14    1800         5    5   3.25
15    1800         5    5   3.30
```

```
16  1250        4   2   2.10
17  4950        9   4   7.10
18   900        4   8   1.68
```

[ ]: df.head(10)

[ ]:
```
    area  bedrooms  age  price
0   1125         4    2   2.80
1   1000         3    2   2.30
2   1575         5    5   3.15
3   1125         5    1   2.92
4    900         4    8   1.85
5   1125         5    3   2.85
6   1125         4    7   1.60
7   1125         4    5   2.10
8   1800         5    4   3.00
9   4500         8    3   7.00
```

[ ]: df.tail(4)

[ ]:
```
     area  bedrooms  age  price
15   1800         5    5   3.30
16   1250         4    2   2.10
17   4950         9    4   7.10
18    900         4    8   1.68
```

[ ]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 19 entries, 0 to 18
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   area      19 non-null     int64
 1   bedrooms  19 non-null     int64
 2   age       19 non-null     int64
 3   price     19 non-null     float64
dtypes: float64(1), int64(3)
memory usage: 736.0 bytes
```

[ ]: corr = df.corr()
     print(corr)

```
              area  bedrooms       age     price
area      1.000000  0.928692 -0.203855  0.968575
bedrooms  0.928692  1.000000 -0.195161  0.914198
age      -0.203855 -0.195161  1.000000 -0.337718
price     0.968575  0.914198 -0.337718  1.000000
```

```
[ ]: sns.heatmap(corr, annot= True)
```

```
[ ]: <Axes: >
```



```
[ ]: from sklearn.model_selection import train_test_split
     train,test = train_test_split(df,test_size=0.2, random_state=41)
     xtrain = train[['area', 'bedrooms','age']]
     ytrain = train['price']
```

```
[ ]: print(xtrain.shape)
     print(ytrain.shape)
```

```
(15, 3)
(15,)
```

```
[ ]: xtest = test[['area','bedrooms','age']]
     ytest = test['price']
```

```
[ ]: print(xtest.shape)
     print(ytest.shape)
```

```
(4, 3)
```

```
(4,)
```

```python
from sklearn.linear_model import LinearRegression
```

```python
model = LinearRegression()
model.fit(xtrain,ytrain)
```

```python
LinearRegression()
```

```python
y_pred = model.predict(xtest)
```

```python
ytest
```

```
6      1.60
13     3.40
11     5.60
4      1.85
Name: price, dtype: float64
```

```python
y_pred
```

```
array([2.01906075, 3.22264273, 5.18347443, 1.63437317])
```