A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch. Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.

In [1]:
```
import pandas as pd
import numpy as np
from scipy import stats
from scipy.stats import norm
import seaborn as sns
from matplotlib import pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

1.Import the data

In [2]:
```
lab_data = pd.read_csv('LabTAT.csv')
lab_data
```

Out[2]:

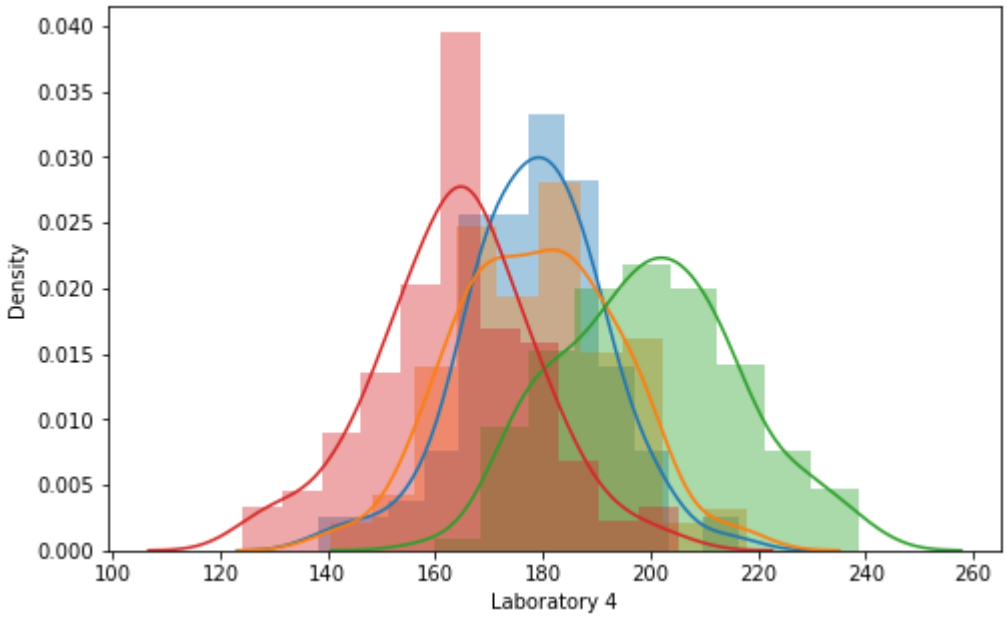
	Laboratory 1	Laboratory 2	Laboratory 3	Laboratory 4
0	185.35	165.53	176.70	166.13
1	170.49	185.91	198.45	160.79
2	192.77	194.92	201.23	185.18
3	177.33	183.00	199.61	176.42
4	193.41	169.57	204.63	152.60
...	...	...	...	...
115	178.49	170.66	193.80	172.68
116	176.08	183.98	215.25	177.64
117	202.48	174.54	203.99	170.27
118	182.40	197.18	194.52	150.87
119	182.09	215.17	221.49	162.21

120 rows × 4 columns

In [3]:
```
lab1 = lab_data['Laboratory 1'].mean()
lab2 = lab_data['Laboratory 2'].mean()
lab3 = lab_data['Laboratory 3'].mean()
lab4 = lab_data['Laboratory 4'].mean()
print('Laboratory 1 mean = ',lab1)
print('Laboratory 2 mean = ',lab2)
print('Laboratory 3 mean = ',lab3)
print('Laboratory 4 mean = ',lab4)
```

```
Laboratory 1 mean =  178.36158333333339
Laboratory 2 mean =  178.9029166666668
Laboratory 3 mean =  199.91325000000003
Laboratory 4 mean =  163.6827499999999
```

In [4]:
```
plt.figure(figsize = [8,5])
sns.distplot(lab_data['Laboratory 1'])
sns.distplot(lab_data['Laboratory 2'])
sns.distplot(lab_data['Laboratory 3'])
sns.distplot(lab_data['Laboratory 4'])
plt.show()
```



In [5]:
```
lab1 = pd.DataFrame(lab_data["Laboratory 1"])
lab2 = pd.DataFrame(lab_data["Laboratory 2"])
lab3 = pd.DataFrame(lab_data["Laboratory 3"])
lab4 = pd.DataFrame(lab_data["Laboratory 4"])
```

In [8]:
```
import scipy.stats as stats
stats.f_oneway(lab1,lab2,lab3,lab4)
```

Out[8]:
```
F_onewayResult(statistic=array([118.70421654]), pvalue=array([2.11567089e-57]))
```

compare p_value at 5% significance level i.e. 0.05