In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import scipy
import scipy.stats as stats
import pylab
```

In [2]:
```
LabTAT = pd.read_csv('/Users/acer/Sandesh Pal/Data Science Assgn/Hypothesis/LabTAT.csv')
```

In [3]:
```
LabTAT
```

Out[3]:

	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

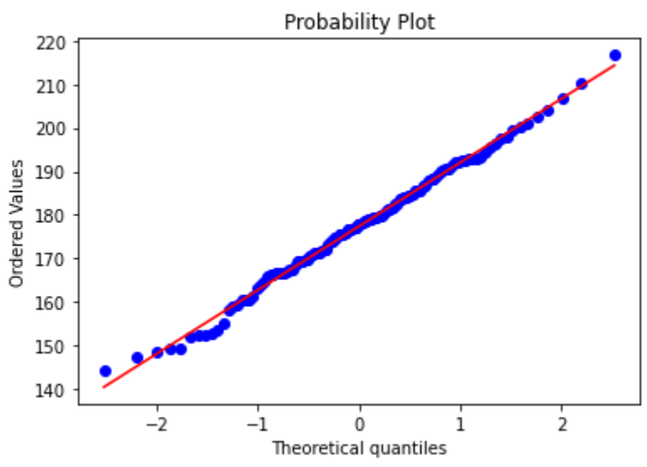
120 rows × 4 columns

In [4]:
```
LabTAT.describe()
```

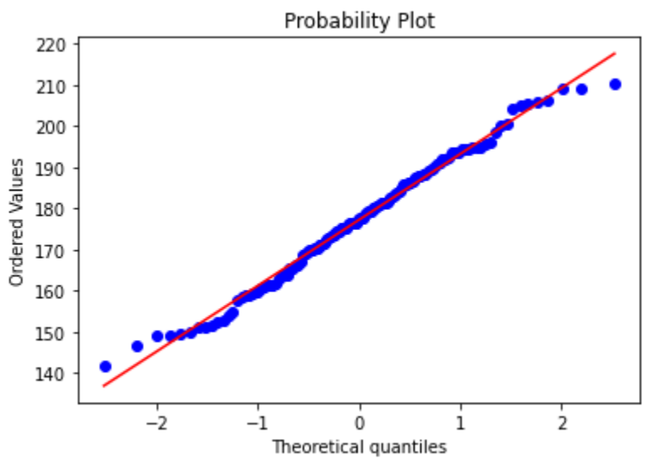
Out[4]:

	Laboratory 1	Laboratory 2	Laboratory 3	Laboratory 4
count	120.000000	120.000000	120.000000	120.00000
mean	178.361583	178.902917	199.913250	163.68275
std	13.173594	14.957114	16.539033	15.08508
min	138.300000	140.550000	159.690000	124.06000
25%	170.335000	168.025000	188.232500	154.05000
50%	178.530000	178.870000	199.805000	164.42500
75%	186.535000	189.112500	211.332500	172.88250
max	216.390000	217.860000	238.700000	205.18000

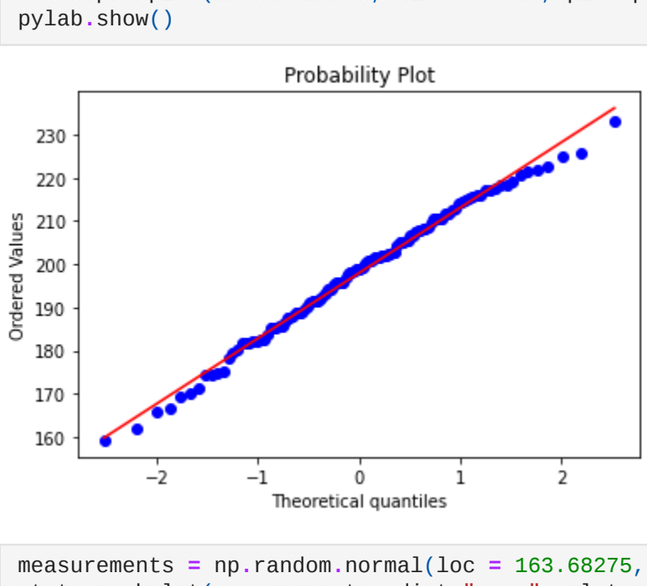
In [5]:
```
measurements = np.random.normal(loc = 178.361583, scale = 13.173594  , size=120)
stats.probplot(measurements, dist="norm", plot=pylab)
pylab.show()
```



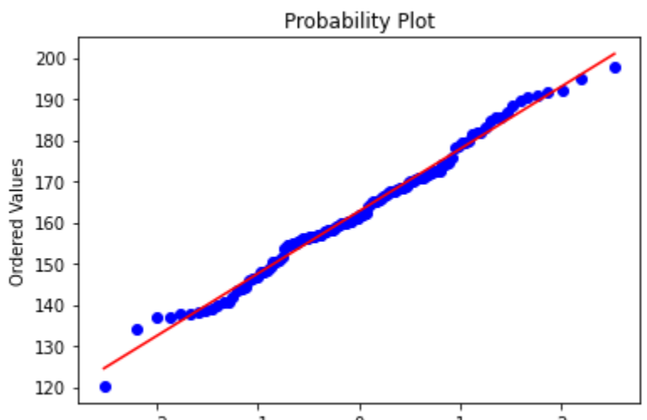
In [6]:
```
measurements = np.random.normal(loc = 178.902917, scale = 14.957114  , size=120)
stats.probplot(measurements, dist="norm", plot=pylab)
pylab.show()
```



In [7]:
```
measurements = np.random.normal(loc = 199.913250, scale = 16.539033  , size=120)
stats.probplot(measurements, dist="norm", plot=pylab)
pylab.show()
```



In [8]:
```
measurements = np.random.normal(loc = 163.68275, scale = 15.08508  , size=120)
stats.probplot(measurements, dist="norm", plot=pylab)
pylab.show()
```



In [10]:
```
#Hence, all the samples are normally distributed
```

In [9]:
```
#Since more than 2 samples are involved, we'll go for 1 way F test i.e. Anova Test
```

In [11]:
```
#Ho = All means are equal
#Ha = All means are not equal
```

In [12]:
```
stats.f_oneway(LabTAT['Laboratory 1'],LabTAT['Laboratory 2'],LabTAT['Laboratory 3'],LabTAT['Laboratory 4'])
```

Out[12]:
```
F_onewayResult(statistic=118.70421654401437, pvalue=2.1156708949992414e-57)
```

In [13]:
```
#Since p-value(2.115e-57)< alpha (0.05), hence reject the null hypothesis
#Conculsion: there is a difference in average TAT among the different laboratories at 5% significance level.
```

In [ ]:
```

```