```
import pandas as pd
import scipy.stats as stats
import matplotlib.pyplot as plt
import seaborn as sns

data = pd.read_csv('iris.txt')
print(data)
```

```
     sepal_length  sepal_width  petal_length  petal_width
species
0             5.1          3.5           1.4          0.2     Iris-
setosa
1             4.9          3.0           1.4          0.2     Iris-
setosa
2             4.7          3.2           1.3          0.2     Iris-
setosa
3             4.6          3.1           1.5          0.2     Iris-
setosa
4             5.0          3.6           1.4          0.2     Iris-
setosa
..            ...          ...           ...          ...
...
145           6.7          3.0           5.2          2.3  Iris-
virginica
146           6.3          2.5           5.0          1.9  Iris-
virginica
147           6.5          3.0           5.2          2.0  Iris-
virginica
148           6.2          3.4           5.4          2.3  Iris-
virginica
149           5.9          3.0           5.1          1.8  Iris-
virginica

[150 rows x 5 columns]
```

```
setosa = data[data['species'] == 'Iris-setosa']['sepal_width']
virginica = data[data['species'] == 'Iris-virginica']['sepal_width']
print(f'Value of stosa is {setosa}')
print(f'Value of virginica is {virginica}')
```

```
Value of stosa is 0     3.5
1     3.0
2     3.2
3     3.1
4     3.6
5     3.9
6     3.4
7     3.4
8     2.9
9     3.1
```

```
10    3.7
11    3.4
12    3.0
13    3.0
14    4.0
15    4.4
16    3.9
17    3.5
18    3.8
19    3.8
20    3.4
21    3.7
22    3.6
23    3.3
24    3.4
25    3.0
26    3.4
27    3.5
28    3.4
29    3.2
30    3.1
31    3.4
32    4.1
33    4.2
34    3.1
35    3.2
36    3.5
37    3.1
38    3.0
39    3.4
40    3.5
41    2.3
42    3.2
43    3.5
44    3.8
45    3.0
46    3.8
47    3.2
48    3.7
49    3.3
Name: sepal_width, dtype: float64
Value of virginica is 100    3.3
101    2.7
102    3.0
103    2.9
104    3.0
105    3.0
106    2.5
107    2.9
```

```
108    2.5
109    3.6
110    3.2
111    2.7
112    3.0
113    2.5
114    2.8
115    3.2
116    3.0
117    3.8
118    2.6
119    2.2
120    3.2
121    2.8
122    2.8
123    2.7
124    3.3
125    3.2
126    2.8
127    3.0
128    2.8
129    3.0
130    2.8
131    3.8
132    2.8
133    2.8
134    2.6
135    3.0
136    3.4
137    3.1
138    3.0
139    3.1
140    3.1
141    3.1
142    2.7
143    3.2
144    3.3
145    3.0
146    2.5
147    3.0
148    3.4
149    3.0
Name: sepal_width, dtype: float64
```

Checkign normality using shapiro wilk test

```
shapiro_setosa, p_setosa = stats.shapiro(setosa)
print(f'Value of p_setosa is {p_setosa}')
```

```
shapiro_virginica, p_virginica = stats.shapiro(virginica)
print(f'Value of shapiro_virginica is {p_virginica}')
```

```
Value of p_setosa is 0.20465666067562632
Value of shapiro_virginica is 0.18089604109069918
```

Checking variance

```
levene_test, levene_p = stats.levene(setosa, virginica)
print(f'Value of levene_p is {levene_p}')
```

```
Value of levene_p is 0.30632327568956713
```

```
t_stat, p_value = stats.ttest_ind(setosa, virginica, equal_var=True)

print("Shapiro-Wilk Test for Normality (Setosa):", shapiro_setosa)
print("Shapiro-Wilk Test for Normality (Virginica):",
shapiro_virginica)
print("Levene's Test for Equal Variances:", levene_test)
print("Two-sample t-test p-value:", p_value)
```

```
Shapiro-Wilk Test for Normality (Setosa):
ShapiroResult(statistic=0.9686918964570316,
pvalue=0.20465666067562632)
Shapiro-Wilk Test for Normality (Virginica):
ShapiroResult(statistic=0.9673905311257006,
pvalue=0.18089604109069918)
Levene's Test for Equal Variances:
LeveneResult(statistic=1.0574747096290729, pvalue=0.30632327568956713)
Two-sample t-test p-value: 8.916634067006443e-09
```

F-Test for Comparing Variances

```
f_stat = setosa.var() / virginica.var()
dfn = len(setosa) - 1
dfd = len(virginica) - 1
f_p_value = 1 - stats.f.cdf(f_stat, dfn, dfd)

print("F-statistic for Variance Comparison:", f_stat)
print("F-test p-value:", f_p_value)
```

```
F-statistic for Variance Comparison: 1.3959028295592788
F-test p-value: 0.12327029558510993
```