

```
import pandas as pd
import numpy as np
```

```
df=pd.read_csv('iris.csv')
```

```
df
```

	sepal.length	sepal.width	petal.length	petal.width	variety
0	5.1	3.5	1.4	0.2	Setosa
1	4.9	3.0	1.4	0.2	Setosa
2	4.7	3.2	1.3	0.2	Setosa
3	4.6	3.1	1.5	0.2	Setosa
4	5.0	3.6	1.4	0.2	Setosa
...
145	6.7	3.0	5.2	2.3	Virginica
146	6.3	2.5	5.0	1.9	Virginica
147	6.5	3.0	5.2	2.0	Virginica
148	6.2	3.4	5.4	2.3	Virginica
149	5.9	3.0	5.1	1.8	Virginica

```
[150 rows x 5 columns]
```

```
df.isnull().sum()
```

```
sepal.length    0
sepal.width     0
petal.length    0
petal.width     0
variety         0
dtype: int64
```

```
df.describe()
```

	sepal.length	sepal.width	petal.length	petal.width
count	150.000000	150.000000	150.000000	150.000000
mean	5.843333	3.057333	3.758000	1.199333
std	0.828066	0.435866	1.765298	0.762238
min	4.300000	2.000000	1.000000	0.100000
25%	5.100000	2.800000	1.600000	0.300000
50%	5.800000	3.000000	4.350000	1.300000
75%	6.400000	3.300000	5.100000	1.800000
max	7.900000	4.400000	6.900000	2.500000

```
df['sepal.width'].describe()
```

```
count    150.000000
mean      3.057333
std       0.435866
min       2.000000
25%      2.800000
50%      3.000000
```

```

75%      3.300000
max      4.400000
Name: sepal.width, dtype: float64

```

```
df.groupby('variety').describe()
```

	sepal.length								
	count	mean	std	min	25%	50%	75%	max	\
variety									
Setosa	50.0	5.006	0.352490	4.3	4.800	5.0	5.2	5.8	
Versicolor	50.0	5.936	0.516171	4.9	5.600	5.9	6.3	7.0	
Virginica	50.0	6.588	0.635880	4.9	6.225	6.5	6.9	7.9	

	sepal.width		...	petal.length		petal.width	
\	count	mean	...	75%	max	count	
mean							
variety			...				
Setosa	50.0	3.428	...	1.575	1.9	50.0	
0.246							
Versicolor	50.0	2.770	...	4.600	5.1	50.0	
1.326							
Virginica	50.0	2.974	...	5.875	6.9	50.0	
2.026							

	std	min	25%	50%	75%	max
variety						
Setosa	0.105386	0.1	0.2	0.2	0.3	0.6
Versicolor	0.197753	1.0	1.2	1.3	1.5	1.8
Virginica	0.274650	1.4	1.8	2.0	2.3	2.5

```
[3 rows x 32 columns]
```

```
variety_list = ['iris-Setosa', 'iris-Versicolor', 'iris-Virginica']
```

```

for variety in variety_list:
    print(f"Summary Statistics for {variety}:\n")
    print(df[df['variety'] == variety].describe()) # Summary stats
    for each species
    print("\n" + "-"*50 + "\n") # Separator for readability

```

```
Summary Statistics for iris-Setosa:
```

	sepal.length	sepal.width	petal.length	petal.width
count	0.0	0.0	0.0	0.0
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN

50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
max	NaN	NaN	NaN	NaN

Summary Statistics for iris-Versicolor:

	sepal.length	sepal.width	petal.length	petal.width
count	0.0	0.0	0.0	0.0
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
max	NaN	NaN	NaN	NaN

Summary Statistics for iris-Virginica:

	sepal.length	sepal.width	petal.length	petal.width
count	0.0	0.0	0.0	0.0
mean	NaN	NaN	NaN	NaN
std	NaN	NaN	NaN	NaN
min	NaN	NaN	NaN	NaN
25%	NaN	NaN	NaN	NaN
50%	NaN	NaN	NaN	NaN
75%	NaN	NaN	NaN	NaN
max	NaN	NaN	NaN	NaN
