+ Code + Text

```
import pandas as pd

import numpy as np

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

	sepal_length	sepal_width	petal_length	petal_width	species
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 × 5 columns

```
df.dtypes
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 150 entries, 0 to 149
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   sepal_length  150 non-null    float64
 1   sepal_width   150 non-null    float64
 2   petal_length  150 non-null    float64
 3   petal_width   150 non-null    float64
 4   species       150 non-null    object
dtypes: float64(4), object(1)
memory usage: 6.0+ KB
```

```
df.groupby("species").agg({
    'sepal_length': ['mean', 'median', 'min', 'max', 'std'],
    'sepal_width': ['mean', 'median', 'min', 'max', 'std'],
    'petal_length': ['mean', 'median', 'min', 'max', 'std'],
    'petal_width': ['mean', 'median', 'min', 'max', 'std']
})
```

	sepal_length					sepal_width					petal_length					petal_width				
	mean	median	min	max	std	mean	median	min	max	std	mean	median	min	max	std	mean	median	min	max	s
species																				
setosa	5.006	5.0	4.3	5.8	0.352490	3.418	3.4	2.3	4.4	0.381024	1.464	1.50	1.0	1.9	0.173511	0.244	0.2	0.1	0.6	(
versicolor	5.936	5.9	4.9	7.0	0.516171	2.770	2.8	2.0	3.4	0.313798	4.260	4.35	3.0	5.1	0.469911	1.326	1.3	1.0	1.8	(
virginica	6.588	6.5	4.9	7.9	0.635880	2.974	3.0	2.2	3.8	0.322497	5.552	5.55	4.5	6.9	0.551895	2.026	2.0	1.4	2.5	(

```
df1 = pd.read_csv("Customers.csv")
df1
```

	CustomerID	Age	Annual Income ($)
0	1	19	15000
1	2	21	35000
2	3	20	86000
3	4	23	59000
4	5	31	38000
...	...	...	...
1995	1996	71	184387
1996	1997	91	73158
1997	1998	87	90961
1998	1999	77	182109
1999	2000	90	110610

2000 rows × 3 columns

```
df1.dtypes
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 3 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   CustomerID         2000 non-null   int64
 1   Age                2000 non-null   int64
 2   Annual Income ($)  2000 non-null   int64
dtypes: int64(3)
memory usage: 47.0 KB
```

```
df1.groupby("Age").agg({
    'Annual Income ($)': ['mean', 'median', 'min', 'max', 'std']
})
```

	Annual Income ($)				
	mean	median	min	max	std
Age					
0	115200.291667	121412.5	22000	186002	43526.590200
1	113246.961538	109883.0	12000	187297	43162.282253
2	115497.600000	111618.0	52727	175208	36474.457652
3	122450.500000	130387.0	55634	188557	43182.814714
4	120743.266667	106816.0	53097	183282	44492.800825
...	...	...	...	...	...
95	106048.000000	91352.0	36000	178228	44969.675858
96	115824.272727	112052.0	1000	186882	43169.135795
97	131496.454545	131529.0	66312	183783	35559.148677
98	102939.444444	104249.0	51866	167197	34981.644266
99	126896.411765	144176.0	55972	184426	46830.757276

100 rows × 5 columns