w3_assessment

January 7, 2022

In this assignment we'll ask you to plot multiple variables.

You will use what you find in this assignment to answer the questions in the quiz that follows. It may be useful to keep this notebook side-by-side with this week's quiz on your screen.

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import scipy.stats as stats
%matplotlib inline
import matplotlib.pyplot as plt
pd.set_option('display.max_columns', 100)

path = "Cartwheeldata.csv"
```

In [2]:
```
# First, you must import the cartwheel data from the path given above
df = pd.read_csv(path)
```

In [3]:
```
# Next, look at the 'head' of our DataFrame 'df'.
df.head()
```

Out[3]:

	ID	Age	Gender	GenderGroup	Glasses	GlassesGroup	Height	Wingspan	CWDistance	Complete	CompleteGroup	Score
0	1	56	F	1	Y	1	62.0	61.0	79	Y	1	7
1	2	26	F	1	Y	1	62.0	60.0	70	Y	1	8
2	3	33	F	1	Y	1	66.0	64.0	85	Y	1	7
3	4	39	F	1	N	0	64.0	63.0	87	Y	1	10
4	5	27	M	2	N	0	73.0	75.0	72	N	0	4

If you can't remember a function, open a previous notebook or video as a reference, or use your favorite search engine to look for a solution.

0.1 Scatter plots

First, let's looks at two variables that we expect to have a strong relationship, 'Height' and 'Wingspan'.

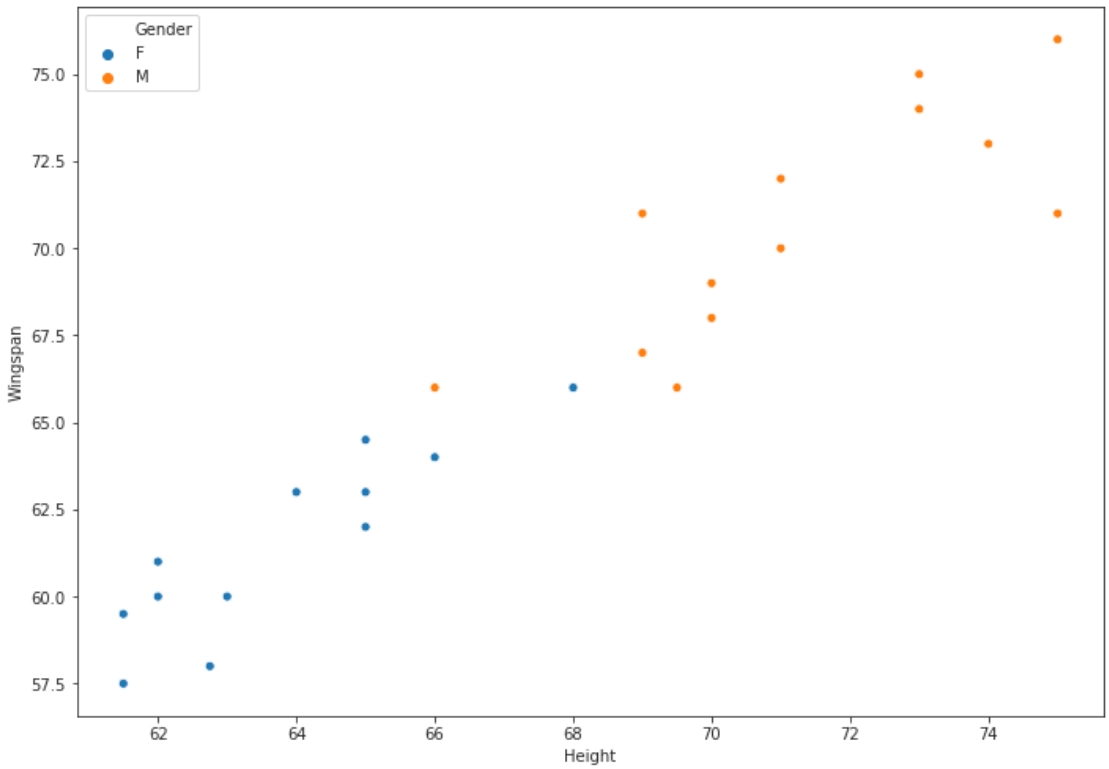
```
In [10]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25 entries, 0 to 24
Data columns (total 12 columns):
ID               25 non-null int64
Age              25 non-null int64
Gender           25 non-null object
GenderGroup      25 non-null int64
Glasses          25 non-null object
GlassesGroup     25 non-null int64
Height           25 non-null float64
Wingspan         25 non-null float64
CWDistance       25 non-null int64
Complete         25 non-null object
CompleteGroup    25 non-null int64
Score            25 non-null int64
dtypes: float64(2), int64(7), object(3)
memory usage: 2.4+ KB
```

```
In [29]: # Make a Seaborn scatter plot with x = height and y = wingspan using sns.scatterplot(
         # plot
         sns.set_style('ticks')
         fig, ax = plt.subplots()

         # the size of A4 paper
         fig.set_size_inches(11.7, 8.27)
         sns.scatterplot(data= df, x="Height", y="Wingspan", hue = 'Gender')
```

```
Out[29]: <matplotlib.axes._subplots.AxesSubplot at 0x7fccb4de2668>
```



How would you describe the relationship between 'Height' and 'Wingspan'?
Questions you can ask: * Is it linear? * Are there outliers? * Are their ranges similar or different?

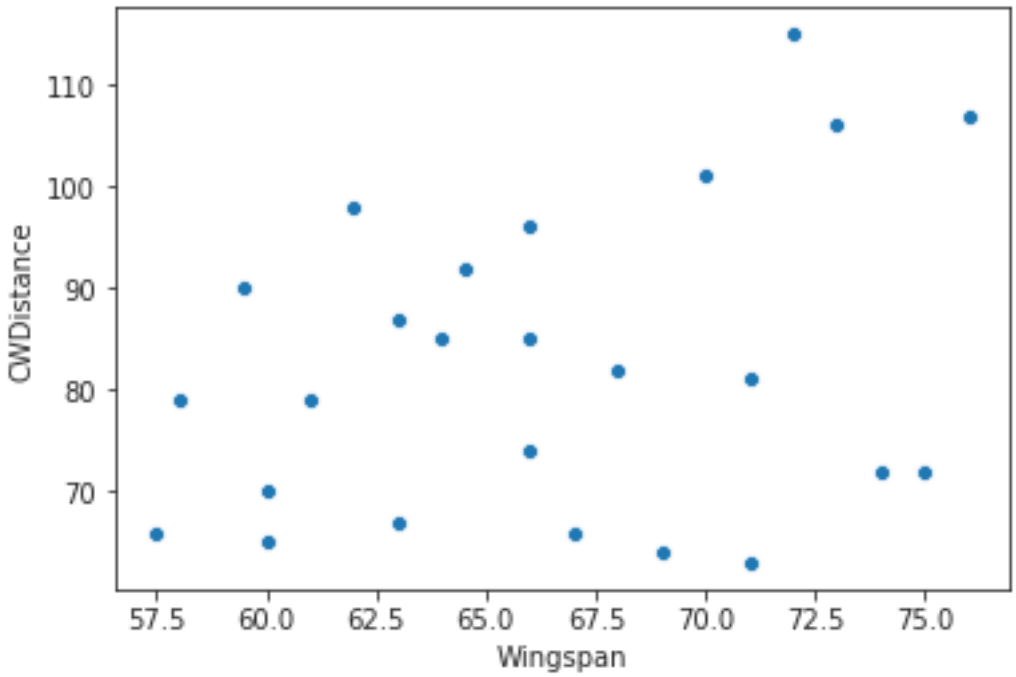
How else could you describe the relationship?

Now let's look at two variables that we don't yet assume have a strong relationship, 'Wingspan' and 'CWDistance'

```
In [20]: # Make a Seaborn scatter plot with x = wingspan and y = cartwheel distance

         sns.scatterplot(data = df, x = 'Wingspan', y = 'CWDistance')

Out[20]: <matplotlib.axes._subplots.AxesSubplot at 0x7fccb4fd7da0>
```



How would you describe the relationship between 'Wingspan' and 'CWDistance'? * Is it linear? * Are there outliers? * Are their ranges similar or different?

How else could you describe the relationship?

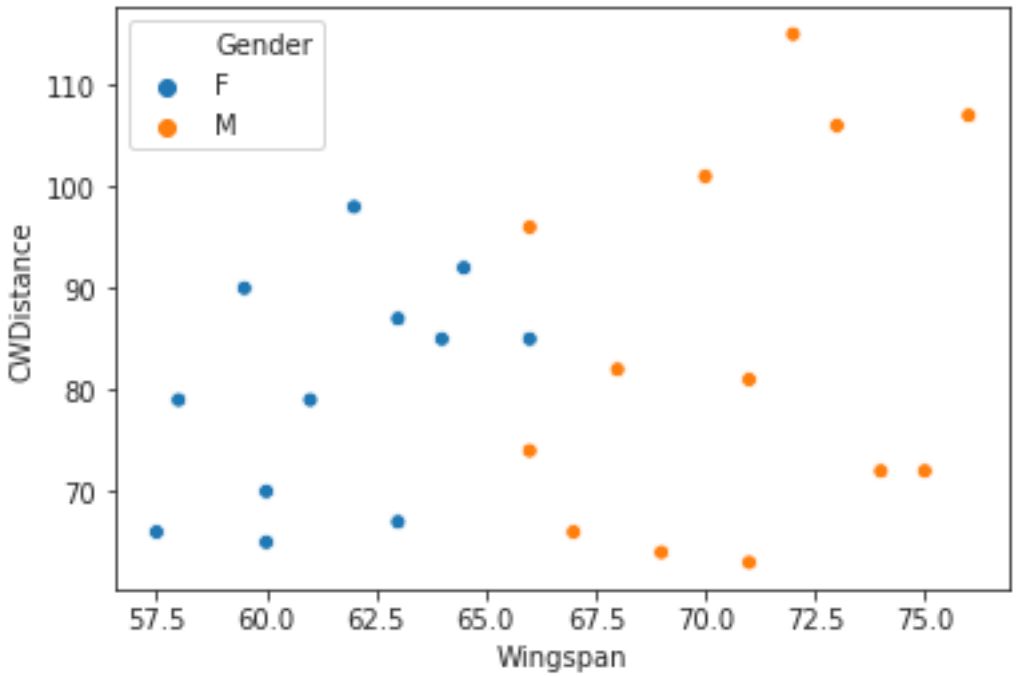
Let makes the same plot as above, but now include 'Gender' as the color scheme by including the argument

```
hue=df['Gender']
```

in the Seaborn function

```
In [22]: # Make a Seaborn scatter plot with x = wingspan and y = cartwheel distance, and hue =
         sns.scatterplot(data = df, x = 'Wingspan', y = 'CWDistance', hue = 'Gender')

Out[22]: <matplotlib.axes._subplots.AxesSubplot at 0x7fccb51a76d8>
```

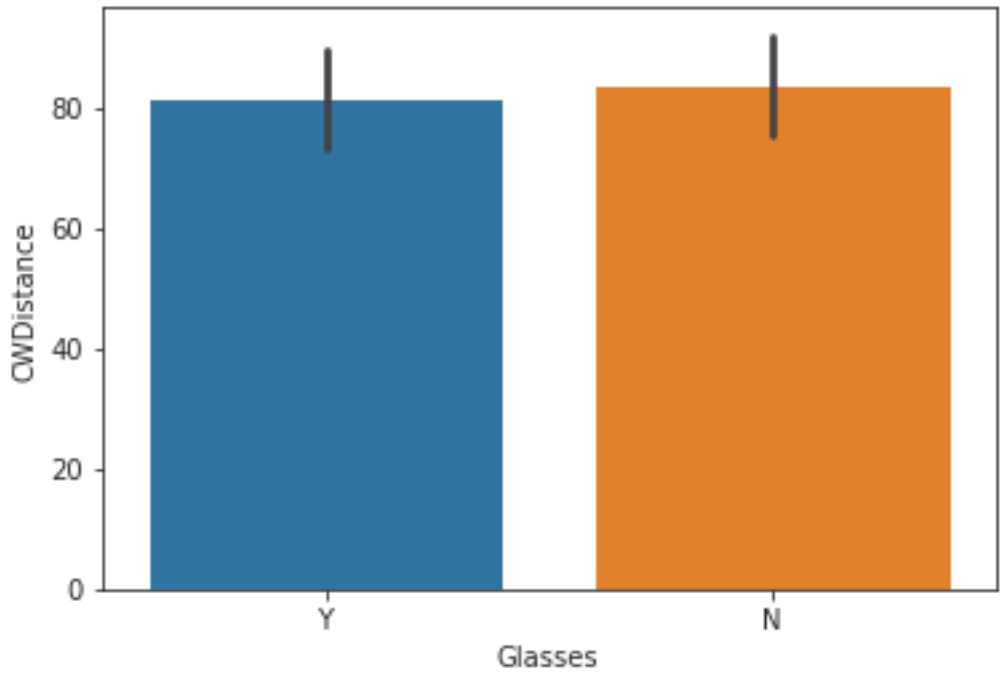


Does does this new information on the plot change your interpretation of the relationship between 'Wingspan' and 'CWDistance'?

0.2 Barcharts

Now lets plot barplots of 'Glasses'

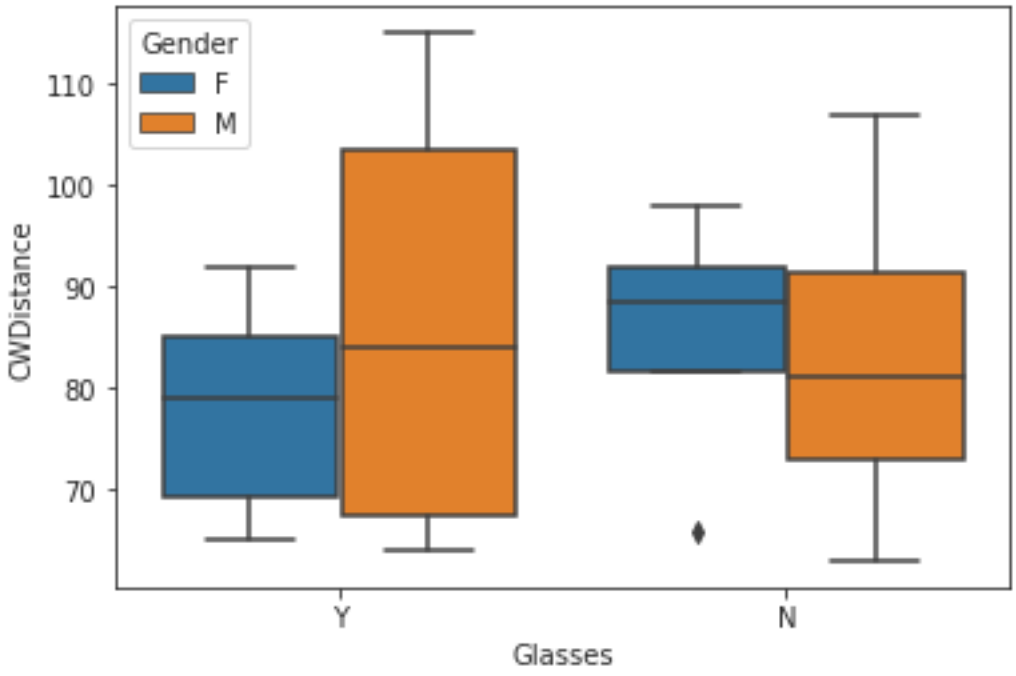
```
In [26]: # Make a Seaborn barplot with x = glasses and y = cartwheel distance
         ax = sns.barplot(data = df, x = 'Glasses', y = 'CWDistance')
```



What can you say about the relationship of 'Glasses' and 'CWDistance'?

```
In [28]: # Make the same Seaborn boxplot as above, but include gender for the hue argument
         sns.boxplot(data = df, x = 'Glasses', y ='CWDistance', hue = 'Gender')
```

```
Out[28]: <matplotlib.axes._subplots.AxesSubplot at 0x7fccb4e57240>
```



How does this new plot change your interpretation about the relationship of 'Glasses' and 'CWDistance'?