

w3_assessment

December 14, 2021

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 [2]: 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 [3]: # First, you must import the cartwheel data from the path given above
df = pd.read_csv(path)
```

```
In [4]: # Next, look at the 'head' of our DataFrame 'df'.
df.head()
```

```
Out[4]:
```

	ID	Age	Gender	GenderGroup	Glasses	GlassesGroup	Height	Wingspan	\
0	1	56	F	1	Y	1	62.0	61.0	
1	2	26	F	1	Y	1	62.0	60.0	
2	3	33	F	1	Y	1	66.0	64.0	
3	4	39	F	1	N	0	64.0	63.0	
4	5	27	M	2	N	0	73.0	75.0	

	CWDistance	Complete	CompleteGroup	Score
0	79	Y	1	7
1	70	Y	1	8
2	85	Y	1	7
3	87	Y	1	10
4	72	N	0	4

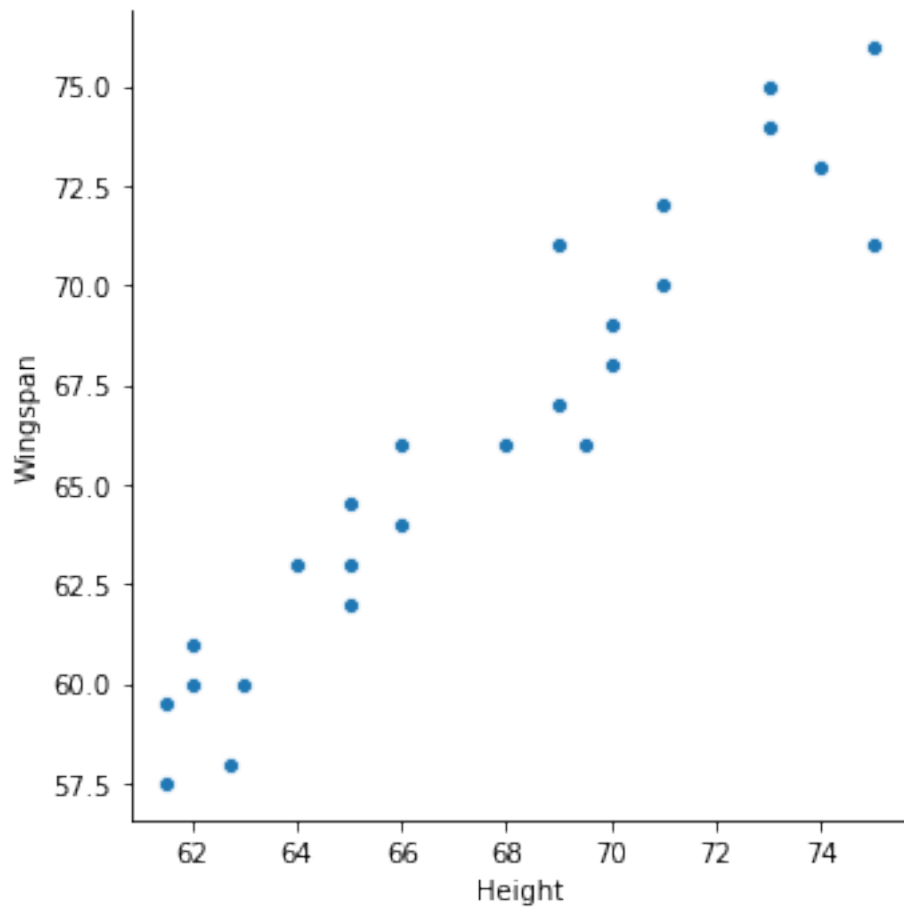
If you can't remember a function, open a previous notebook or video as a reference, or use your favorite search engine to look for a solution.

0.1 Scatter plots

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

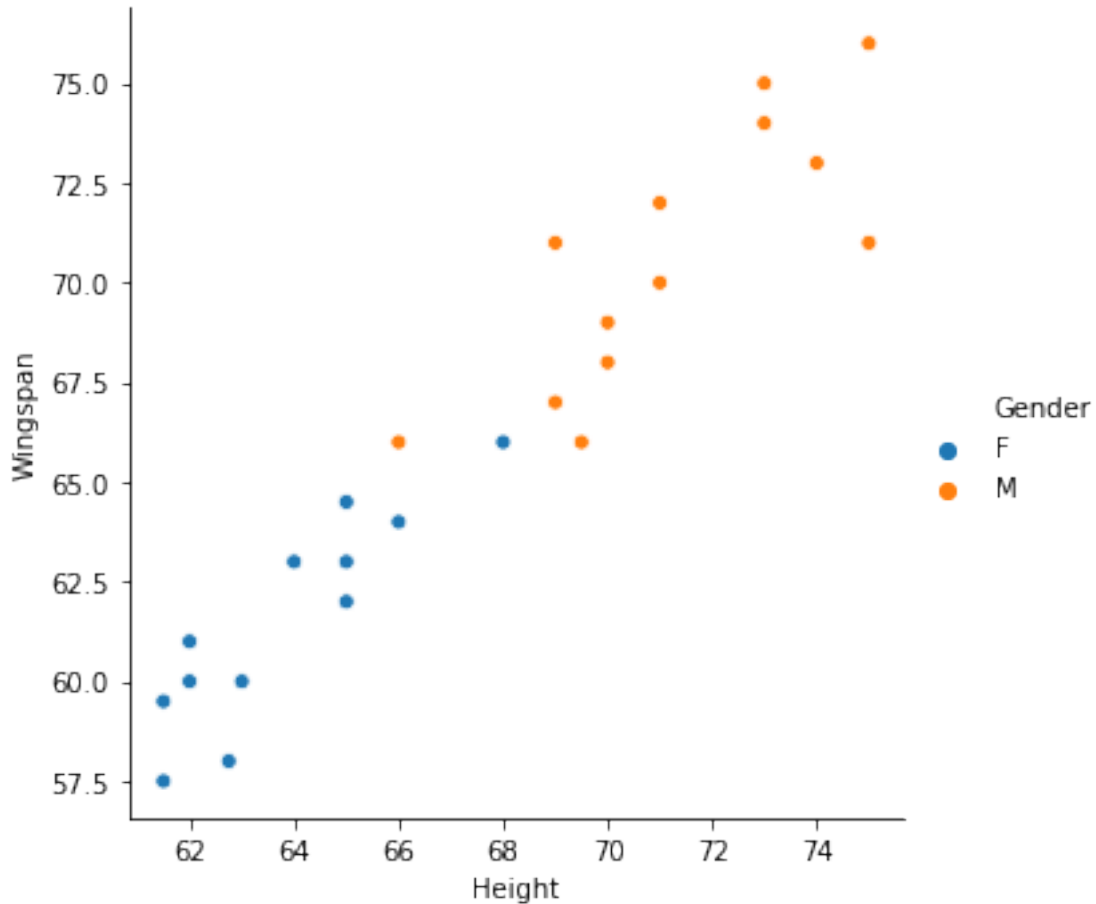
```
In [5]: # Make a Seaborn scatter plot with x = height and y = wingspan using sns.scatterplot(x=  
sns.relplot(x="Height", y="Wingspan", data=df)
```

```
Out[5]: <seaborn.axisgrid.FacetGrid at 0x7ff89c32fd68>
```



```
In [6]: sns.relplot(x="Height", y="Wingspan", hue=df['Gender'], data=df)
```

```
Out[6]: <seaborn.axisgrid.FacetGrid at 0x7ff89a2774a8>
```



How would you describe the relationship between 'Height' and 'Wingspan'?

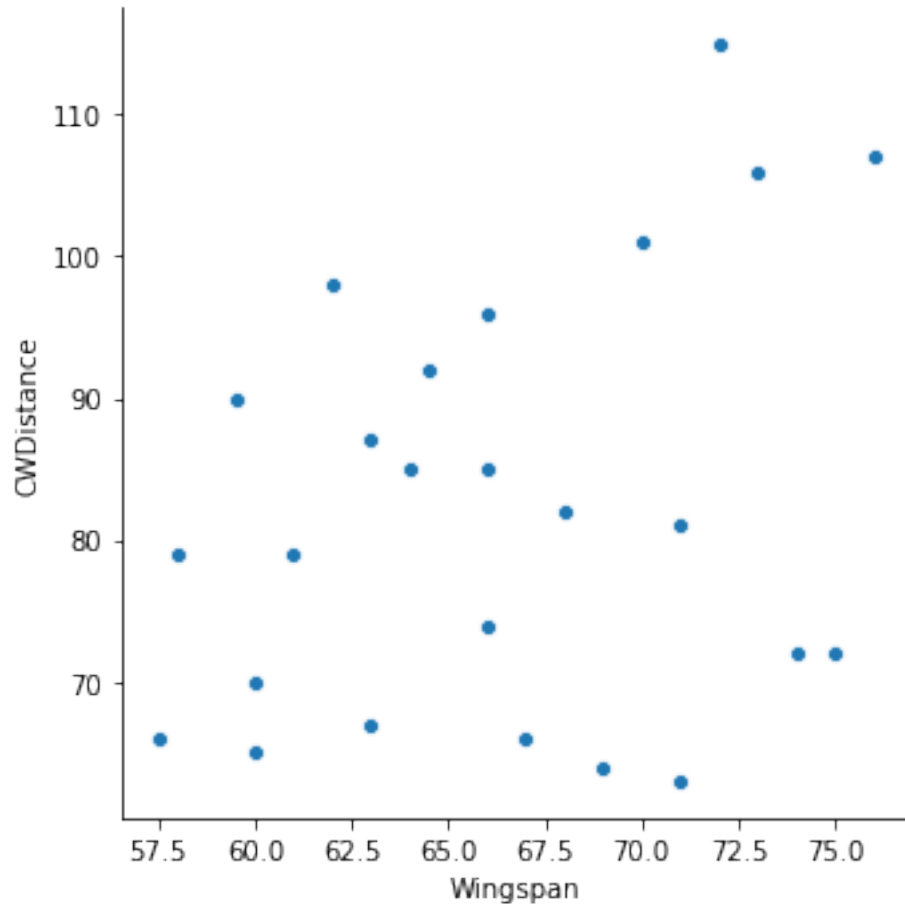
Questions you can ask: * Is it linear? yes * Are there outliers? yes * Are their ranges similar or different? Similar

How else could you describe the relationship? positive correlation

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

```
In [7]: # Make a Seaborn scatter plot with x = wingspan and y = cartwheel distance
sns.relplot(x="Wingspan", y="CWDistance", data=df)
```

```
Out[7]: <seaborn.axisgrid.FacetGrid at 0x7ff89a2715f8>
```



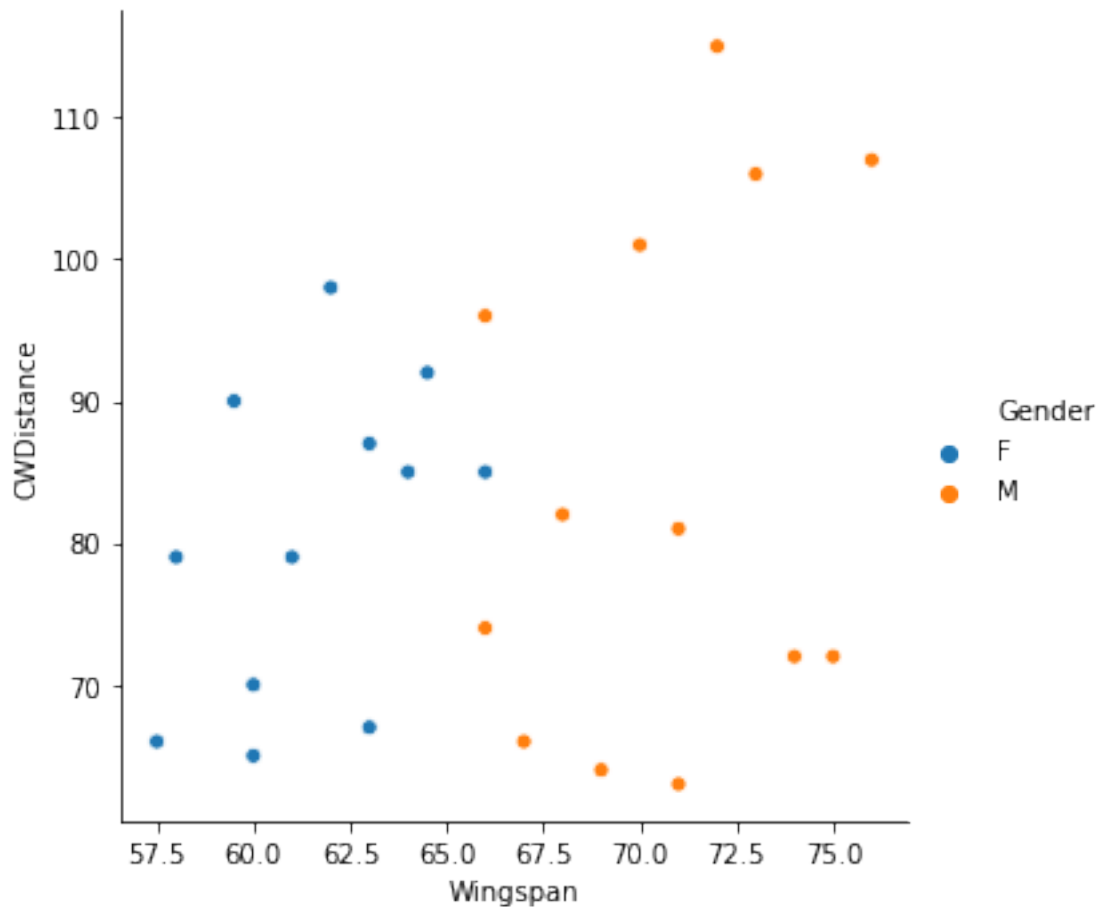
How would you describe the relationship between 'Wingspan' and 'CWDistance'?
 * Is it linear? No * Are there outliers? Yes * Are their ranges similar or different? Different
 How else could you describe the relationship? No correlation
 Let makes the same plot as above, but now include 'Gender' as the color scheme by including the argument

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

in the Seaborn function

```
In [8]: # Make a Seaborn scatter plot with x = wingspan and y = cartwheel distance, and hue = gender
sns.relplot(x="Wingspan", y="CWDistance", hue=df['Gender'], data=df)
```

```
Out[8]: <seaborn.axisgrid.FacetGrid at 0x7ff968cbcc18>
```



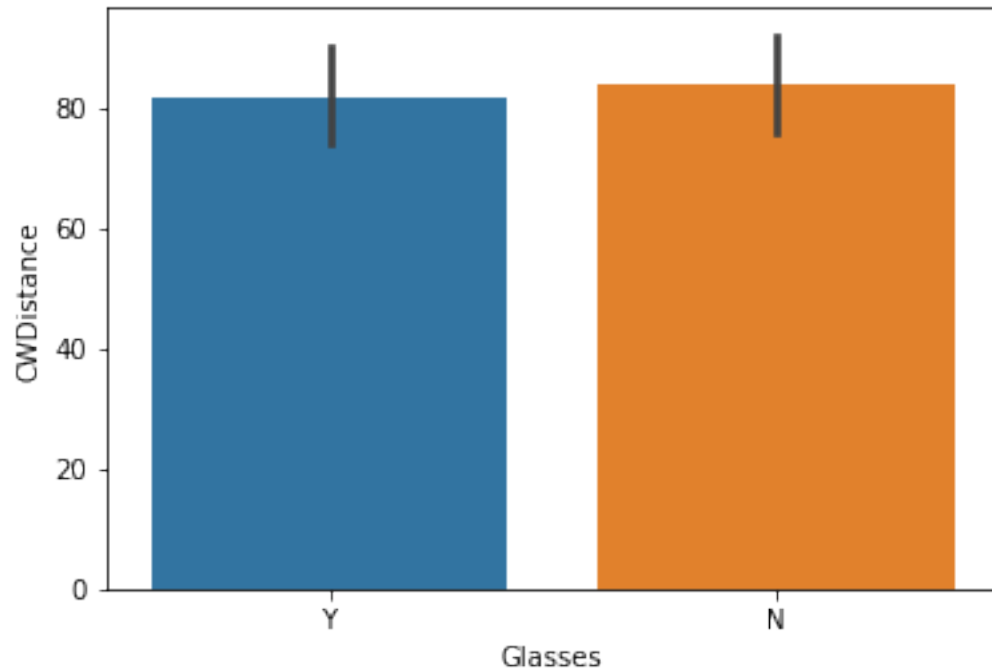
Does does this new information on the plot change your interpretation of the relationship between 'Wingspan' and 'CWDistance'? Yes. Seems to have some correlation for Female, but no for Males

0.2 Barcharts

Now lets plot barplots of 'Glasses'

```
In [9]: # Make a Seaborn barplot with x = glasses and y = cartwheel distance
sns.barplot(x="Glasses", y="CWDistance", data=df)
```

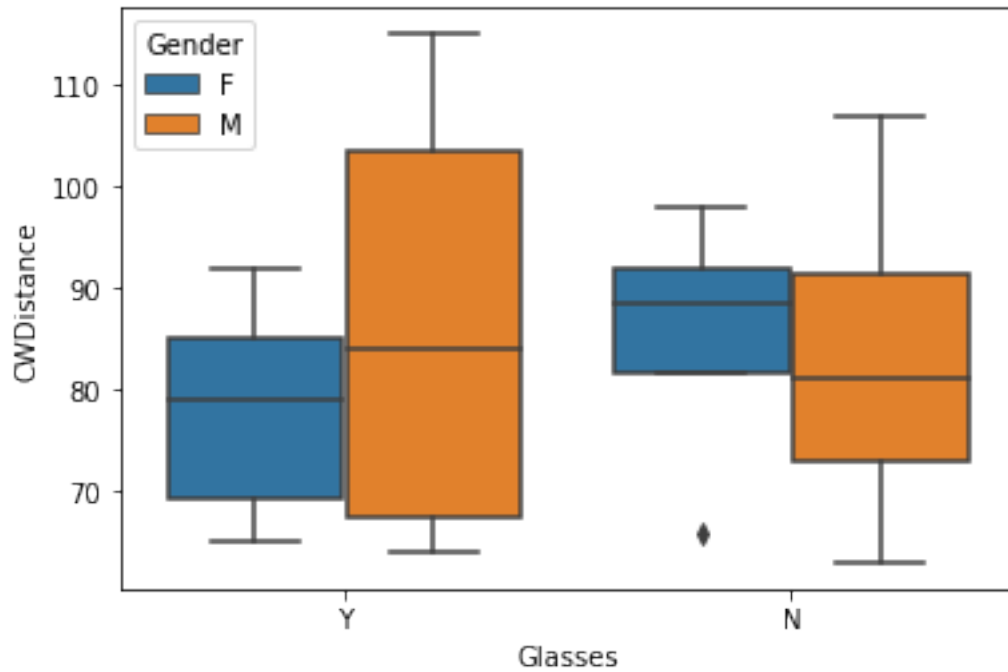
```
Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0x7ff89a17fa20>
```



What can you say about the relationship of 'Glasses' and 'CWDistance'? Almost no difference. Glasses N slightly higher.

```
In [10]: # Make the same Seaborn boxplot as above, but include gender for the hue argument
sns.boxplot(x="Glasses", y="CWDistance", hue="Gender", data=df)
```

```
Out[10]: <matplotlib.axes._subplots.AxesSubplot at 0x7ff89a0c5d30>
```



How does this new plot change your interpretation about the relationship of 'Glasses' and 'CWDistance'? Glasses Y higher for female than male. For the Glasses N the results are different. So results change by gender.