Scatter Plots:

讀檔

```python
import pandas as pd
pd.plotting.register_matplotlib_converters()
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
print("Setup Complete")
```

```python
# Set up code checking
import os
if not os.path.exists("../input/candy.csv"):
    os.symlink("../input/data-for-datavis/candy.csv", "../input/candy.csv")
from learntools.core import binder
binder.bind(globals())
from learntools.data_viz_to_coder.ex4 import *
print("Setup Complete")
```

Setup Complete

```python
# Path of the file to read
candy_filepath = "../input/candy.csv"

# Fill in the line below to read the file into a variable candy_data
candy_data = pd.read_csv(candy_filepath,index_col="id")

# Run the line below with no changes to check that you've loaded the data correctly
step_1.check()
```

Print 前五筆資料

```python
# Print the first five rows of the data
candy_data.head() # Your code here
```

| id	competitorname	chocolate	fruity	caramel	peanutyalmondy	nougat	crispedricewafer	hard	bar	pluribus	sugarpercent	pricepercent	winpercent
0	100 Grand	Yes	No	Yes	No	No	Yes	No	Yes	No	0.732	0.860	66.971725
1	3 Musketeers	Yes	No	No	No	Yes	No	No	Yes	No	0.604	0.511	67.602936
2	Air Heads	No	Yes	No	No	No	No	No	No	No	0.906	0.511	52.341465
3	Almond Joy	Yes	No	No	Yes	No	No	No	Yes	No	0.465	0.767	50.347546
4	Baby Ruth	Yes	No	Yes	Yes	Yes	No	No	Yes	No	0.604	0.767	56.914547

Step3:建立一個散佈圖，顯示「糖百分比」（橫軸 x 軸）和「勝率百分比」（縱軸 y 軸）之間的關係。暫時不要加入迴歸線

```
# Scatter plot showing the relationship between 'sugarpercent' and 'winpercent'
sns.scatterplot(x=candy_data['sugarpercent'],y=candy_data['winpercent'])

# Check your answer
step_3.a.check()
```
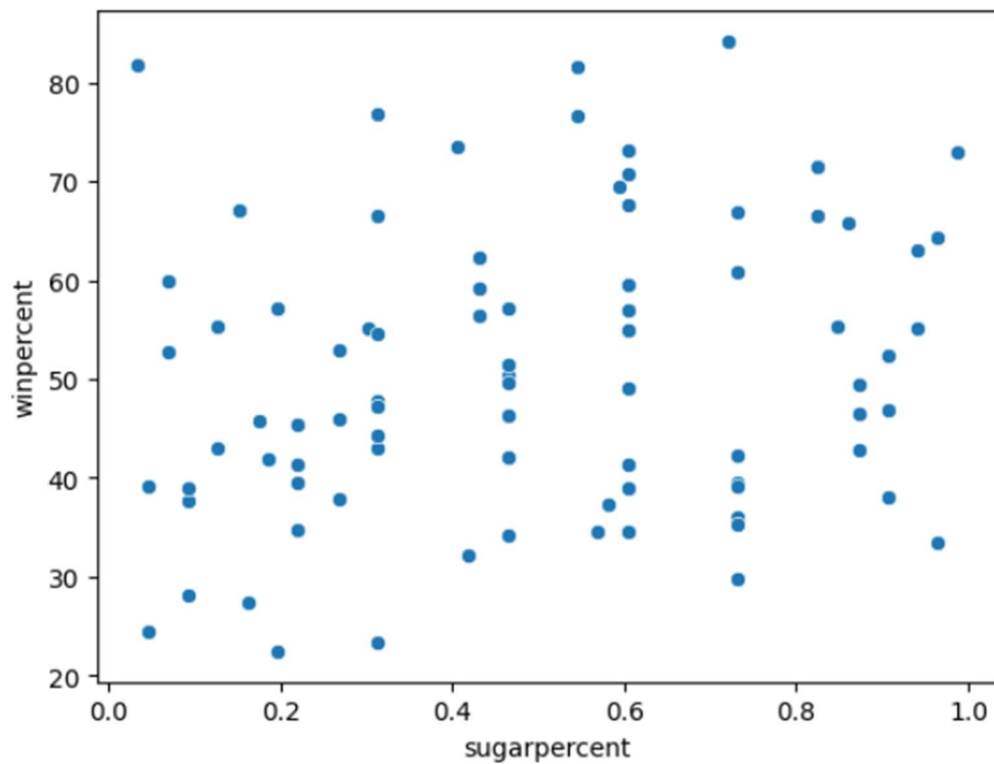

Step4 建立與步驟 3 中建立的相同的散點圖，但現在帶有迴歸線！

```
# Scatter plot w/ regression line showing the relationship between 'sugarpercent' and 'winpercent'
sns.regplot(x=candy_data['sugarpercent'], y=candy_data['winpercent'])

# Check your answer
step_4.a.check()
```
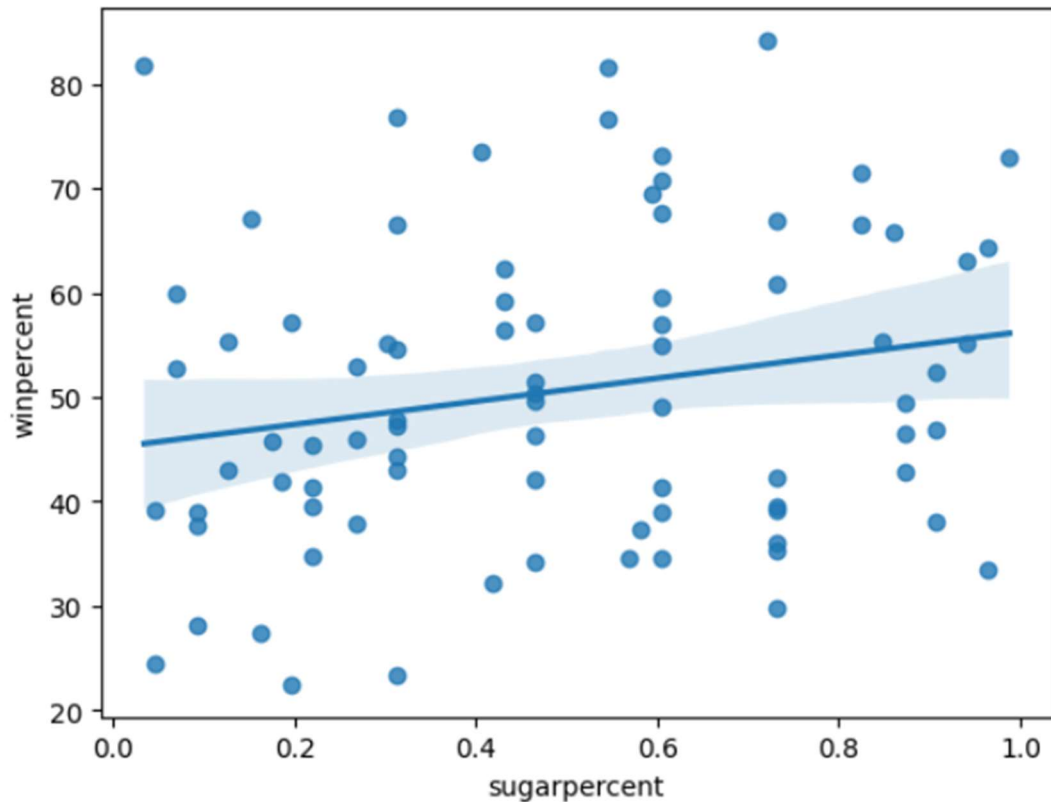
Step5:在下面的程式碼單元中，建立一個散佈圖來展示「pricepercent」（水平 x 軸）和「winpercent」（垂直 y 軸）之間的關係。使用 "chocolate" 列對點進行顏色編碼。暫時不要添加任何回歸線

(散佈圖顯示了「pricepercent」、「winpercent」和「chocolate」之間的關係)

```
# Scatter plot showing the relationship between 'pricepercent', 'winpercent', and 'chocolate'
sns.scatterplot(x=candy_data['pricepercent'],y=candy_data['winpercent'],hue=candy_data['chocolate'])

# Check your answer
step_5.check()
```
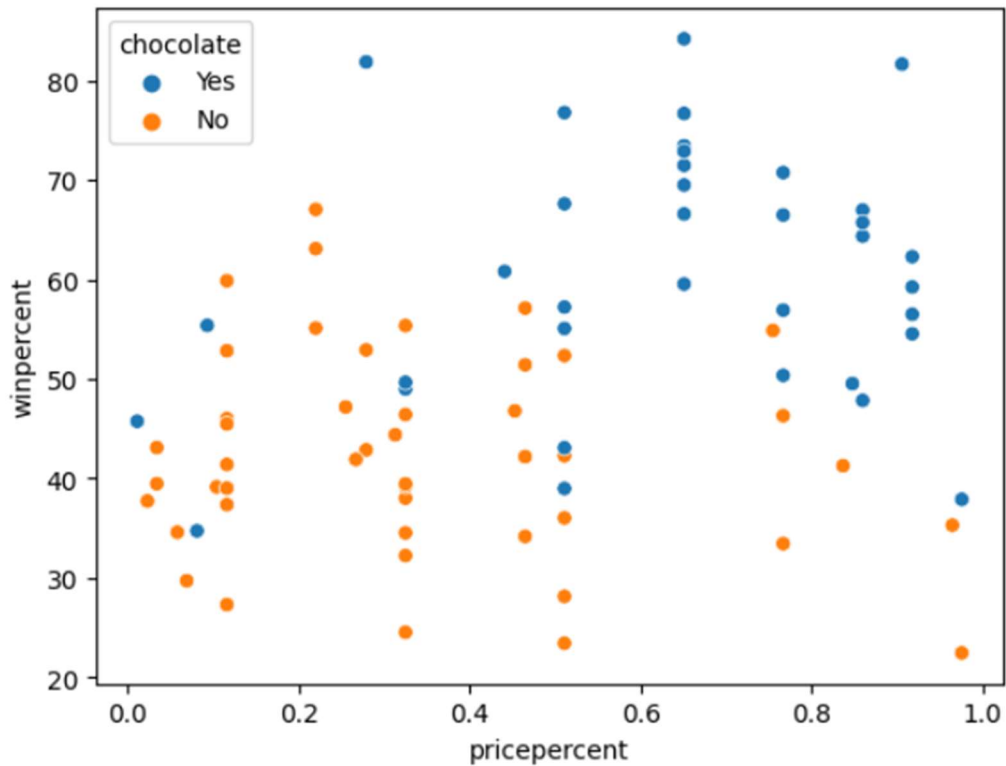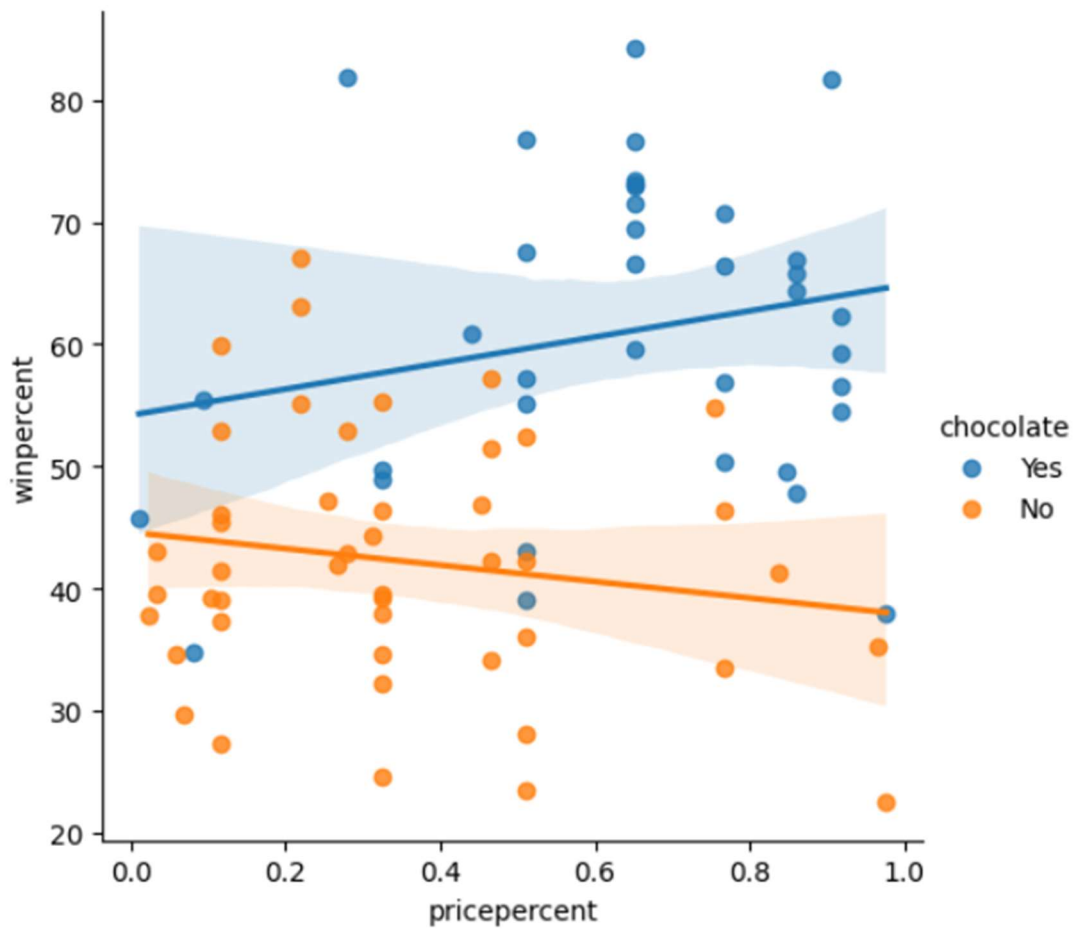
Step6: 建立與步驟 5 相同的散佈圖，但現在有兩條迴歸線，分別對應 (1) 巧克力糖果和 (2) 不含巧克力的糖果。

```python
# Color-coded scatter plot w/ regression lines
sns.lmplot(x="pricepercent",y="winpercent",hue="chocolate",data=candy_data)

# Check your answer
step_6.a.check()
```

Step7: 建立一個分類散點圖，突出顯示「巧克力」和「勝率」之間的關係。將「巧克力」放在（水平）x 軸上，「勝率」放在（垂直）y 軸上。

```python
# Scatter plot showing the relationship between 'chocolate' and 'winpercent'
sns.swarmplot(x=candy_data['chocolate'],y=candy_data['winpercent'])
```