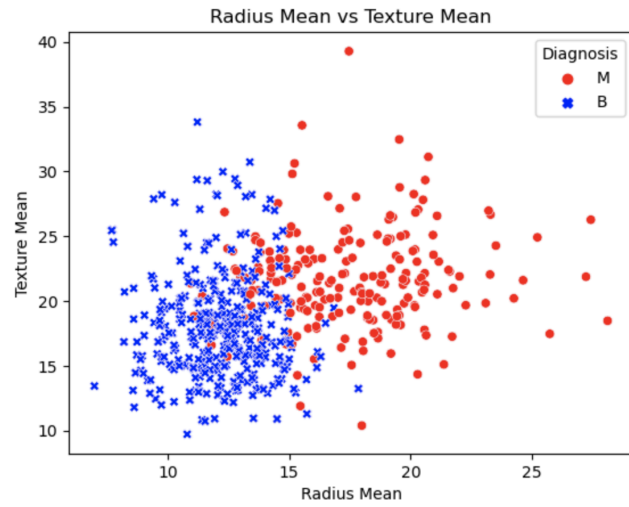


Xiaoming Huang  
1604905  
CSCI 184  
HW2, Part 2: Programming

1, 2, 4-6.

See code and output in [CSCI184\\_HW2\\_XiaomingHuang.ipynb](#)

3.



It seems like **it's linear separable**, as you can see the Benign samples are more on the left side of the graph and Malignant samples are more on the right side of the graph.

7.

I choose to use **Gaussian Naive Bayes**.

That's because all features are continuous numerical features, and the large sample size indicates that we roughly have a normal distribution on every feature.

8.

Result:

Accuracy: 0.69

Classification Report:

	precision	recall	f1-score	support
0	0.69	1.00	0.82	117
1	1.00	0.02	0.04	54
accuracy			0.69	171
macro avg	0.84	0.51	0.43	171
weighted avg	0.79	0.69	0.57	171

Confusion Matrix:

```
[[117  0]
 [ 53  1]]
```

```
array([[117,  0],
       [ 53,  1]])
```