Mid-semester assessment: Intro Statistics

1. What are 3 concepts or techniques you are struggling with so far this semester?

2. What are 3 concepts or techniques you feel like you've improved on so far this semester?

3. What parts of the course do you think are not working well?

4. What parts of the course do you think are working well?

5. What could we do differently to improve your learning?

Programming assessment

All questions below use the following data. Write R commands that will provide the desired results. The data.frame is named iris.

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
7	3.2	4.7	1.4	versicolor
6.4	3.2	4.5	1.5	versicolor
6.9	3.1	4.9	1.5	versicolor
6.3	3.3	6	2.5	virginica
5.8	2.7	5.1	1.9	virginica
7.1	3	5.9	2.1	virginica

1. Calculate the ratio of the average sepal length to the average petal length for all of the observations.

2. Calculate the average petal width for each species separately

3. Use ggplot to provide a scatterplot of sepal length against sepal width. Make sensible choices to help to visualize the relationship

4. Estimate a linear regression model for the relationship between sepal length and sepal width. Sepal length should be your outcome variable.

5. What is the probability that a randomly chosen observation will be of species Versicolor? You can do this one without R.