

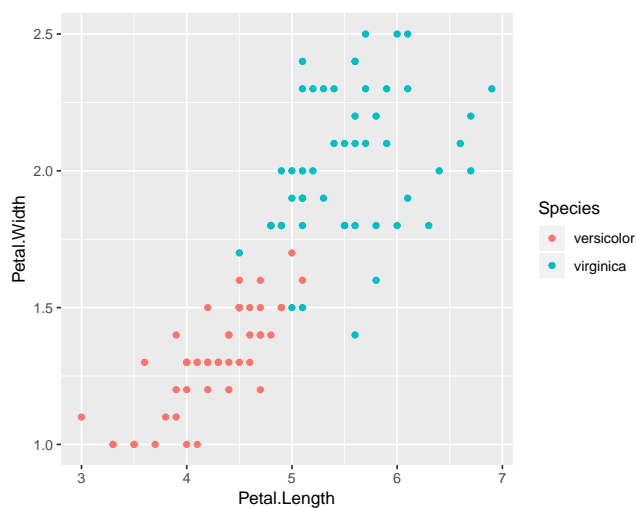
# Homework 2

1. Recall `iris` data set. Find the maximum value of the sum `Sepal.Length + Sepal.Width`. To which species does the sample with the maximum value you found above belong? *Hint*: you might want to start with defining a new variable `Sepal.Sum` using `mutate()`, for example.

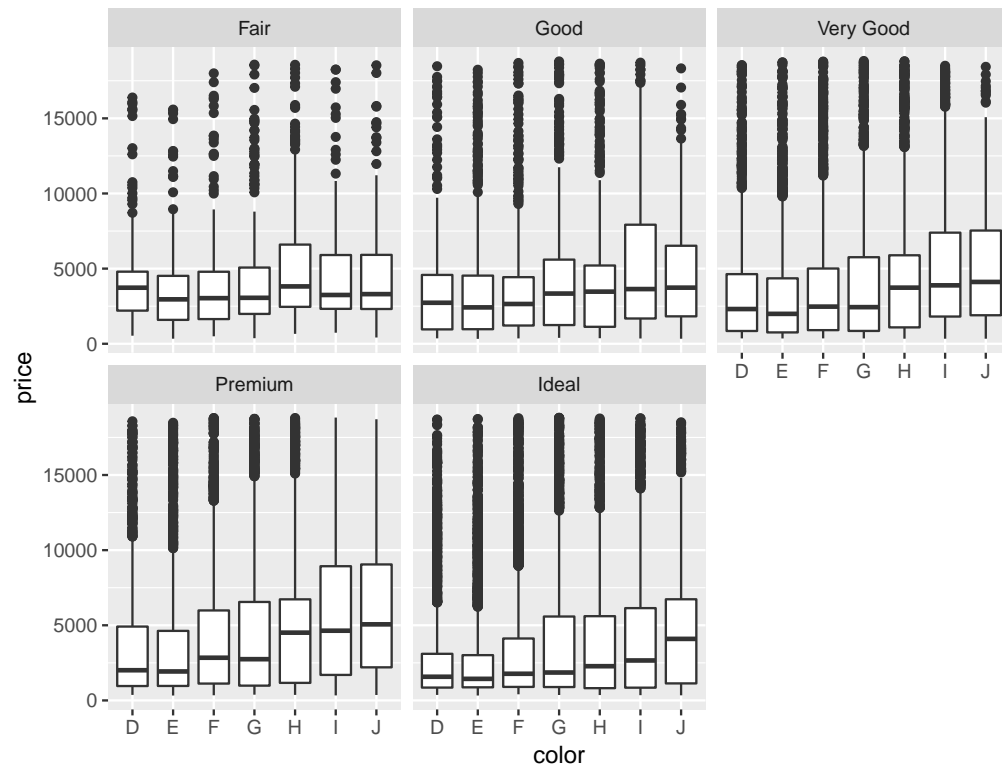
```
## [1] virginica
```

```
## Levels: setosa versicolor virginica
```

2. Create the following plot using `iris` data set. Note that `setosa` species do not show up. You do NOT need to adjust size or width/height ratio of the plot.



3. Create the following plot using `diamonds` data set.



4. Determine  $c$  so that the probability  $0 \leq Z \leq c$  is 0.45, where  $Z$  follows the standard normal distribution  $N(0, 1^2)$ .
5. Recall the `heights.txt` data we used in class with 1375 mom/daughter height pairs:

```
ht <- read.table(file = "heights.txt", header = T, sep = " ")
```

Assuming that UK women's heights follow a normal distribution  $N(\mu, \sigma^2)$ , compute the 95% confidence interval of  $\mu$  based on the mom's height data.