

Homework 1 Solutions

Brian Lois

1/21/2017

1.

a)

```
is_prime <- function(n){  
  if (n == 1){  
    return(FALSE)  
  }  
  if (n == 2){  
    return(TRUE)  
  }  
  for (i in 2:(n-1)){  
    if (n %% i == 0){  
      return(FALSE)  
    }  
  }  
  return(TRUE)  
}
```

b)

```
nth_prime <- function(n){  
  primes <- c()  
  i <- 2  
  while (length(primes) < n){  
    if (is_prime(i)){  
      primes <- c(primes, i)  
    }  
    i <- i + 1  
  }  
  return(primes[length(primes)])  
}
```

c)

```
sapply(seq(20), nth_prime)
```

```
## [1] 2 3 5 7 11 13 17 19 23 29 31 37 41 43 47 53 59 61 67 71
```

2.

```
x <- c(91, NA, 90, 7, 67, NA, 87, 36, 2, 93, 27, 16)
```

a)

```
x <- x[!is.na(x)]
```

b)

```
x[c(1, 3, 8)]
```

```
## [1] 91 7 93
```

c)

```
x[x > 50]
```

```
## [1] 91 90 67 87 93
```

d)

```
x[x %% 2 == 1]
```

```
## [1] 91 7 67 87 93 27
```

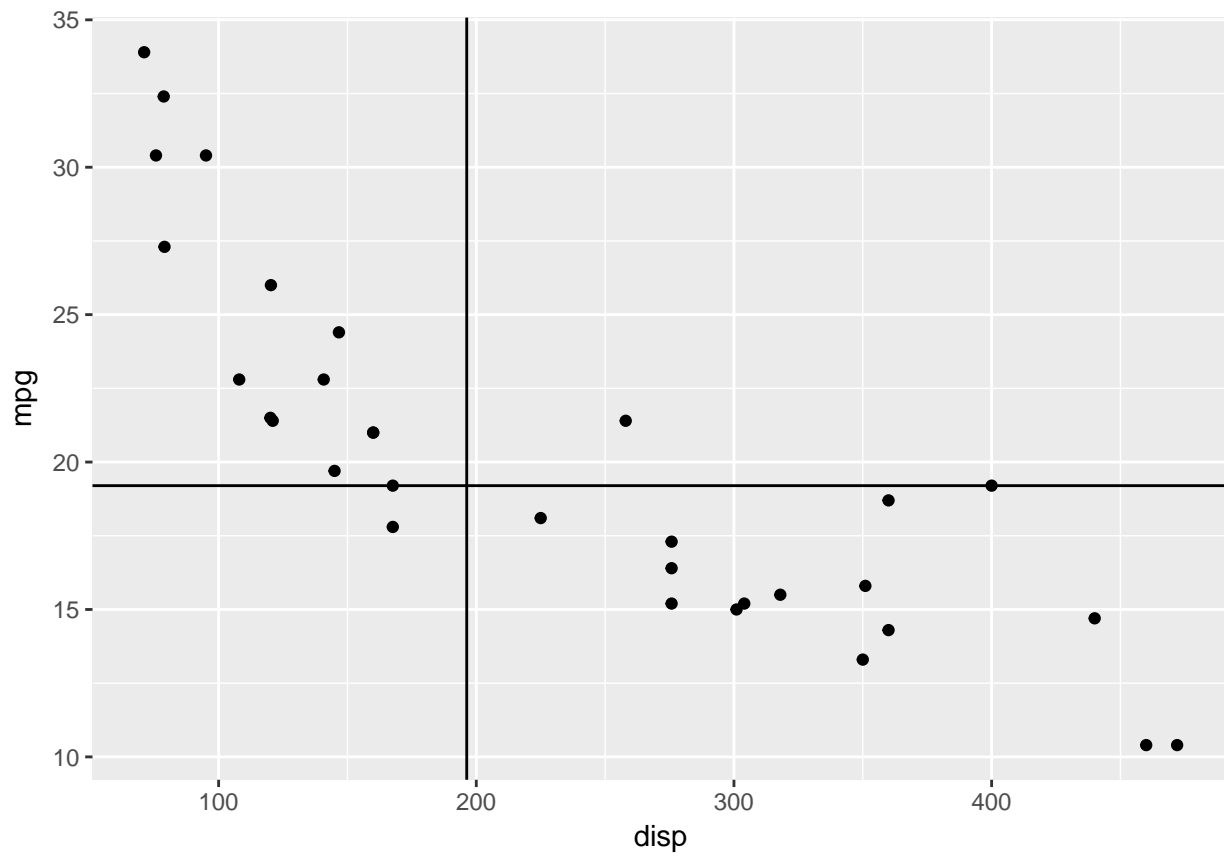
3.

a), b), c)

```
library(ggplot2)
```

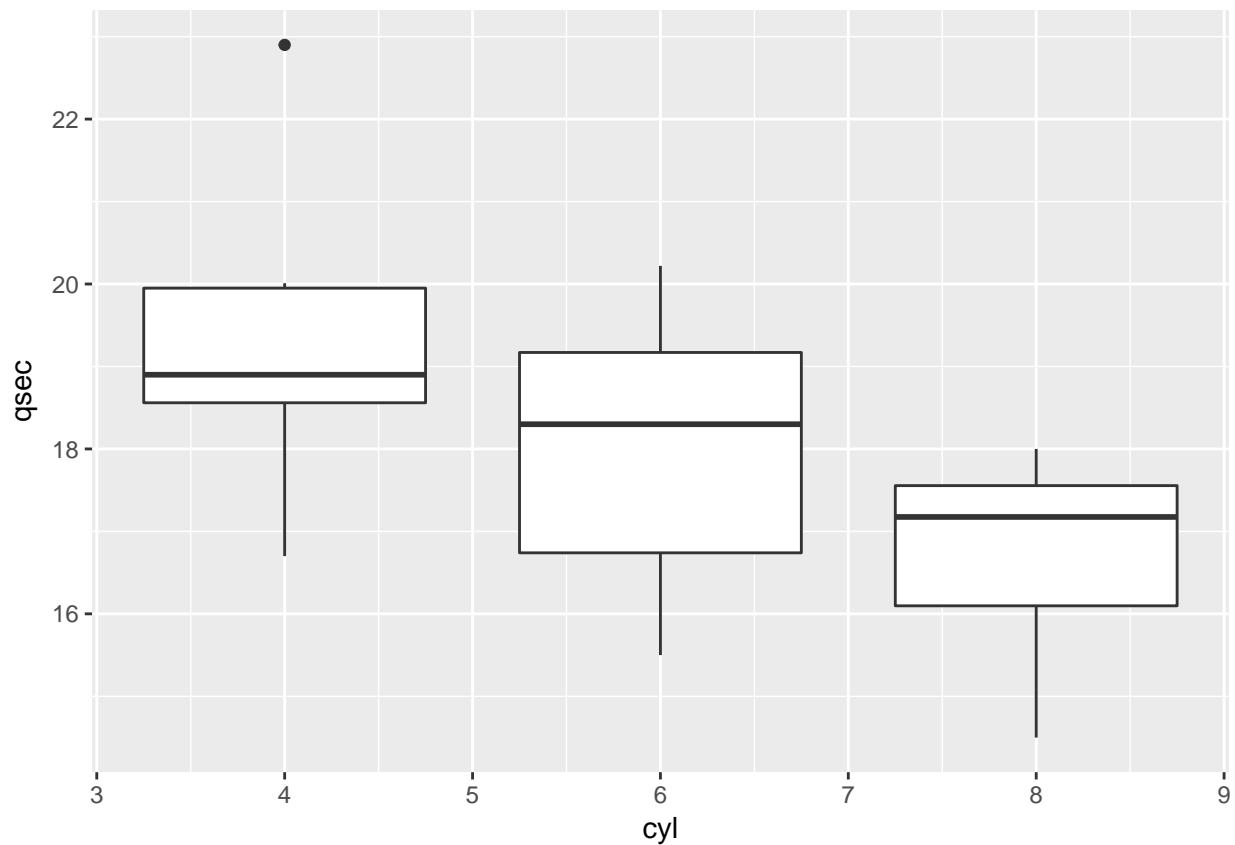
```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```
ggplot(data = mtcars) + geom_point(mapping = aes(x = disp, y = mpg)) +  
  geom_abline(slope = 0, intercept = median(mtcars$mpg)) +  
  geom_vline(xintercept = median(mtcars$disp))
```



d)

```
ggplot(data = mtcars) + geom_boxplot(mapping = aes(x = cyl, y = qsec, group = cyl))
```



```
?geom_boxplot
```

4.

a)

Suppose someone wants to gamble on the outcome of sporting events. He builds a model to predict the winner and score of each game. In this case all the modeler cares about is the predictions. If the predictions are good he will win money.

b)

Suppose a data scientist is working for a sports team and builds a model to predict the winner and score of his team's games. Here the model is only useful if there is some insight gained that can improve the team's chance of winning.