Lab 1

Grace Randall

2024-01-22

Setup

```r
x <- c(0,1,2,3,4,5,6,7,8,9,10)
y <- c(2.96, 4.20, 2.84, 3.84, 6.57, 6.95, 9.32, 10.57, 9.72, 11.57, 11.53)

tree_data <- data.frame(height=x,diameter=y)
```

Question 1

```r
#a
mean(tree_data$height)
```

```
## [1] 5
```

```r
mean(tree_data$diameter)
```

```
## [1] 7.279091
```

```r
#b
quantile(tree_data$height,probs=0.5)
```

```
## 50%
##   5
```

```r
quantile(tree_data$diameter,probs=0.5)
```

```
##  50%
## 6.95
```

    a) the mean tree height is 5m. The mean tree diameter is 7cm.

    b) the median tree height is 5m. The median tree diameter is 7cm.

Question 2

```r
# a)
range(tree_data$height)
```

```
## [1]  0 10
```

```r
range(tree_data$diameter)
```

```
## [1]  2.84 11.57
```

```r
max(tree_data$diameter)-min(tree_data$diameter)
```

```
## [1] 8.73
```

a) The range of the tree height is 10m from 0m to 10m. the range of tree diameter is 8.73cm from 2.84cm to 11.57cm.

```
# b)
```

```
var(tree_data$height)
```

```
## [1] 11
```

```
var(tree_data$diameter)
```

```
## [1] 11.77149
```

b) The variance in tree height is 11 square meters. The variance in tree diameter is 11.77 square centimeters

```
# c)
```

```
sd(tree_data$height)
```

```
## [1] 3.316625
```
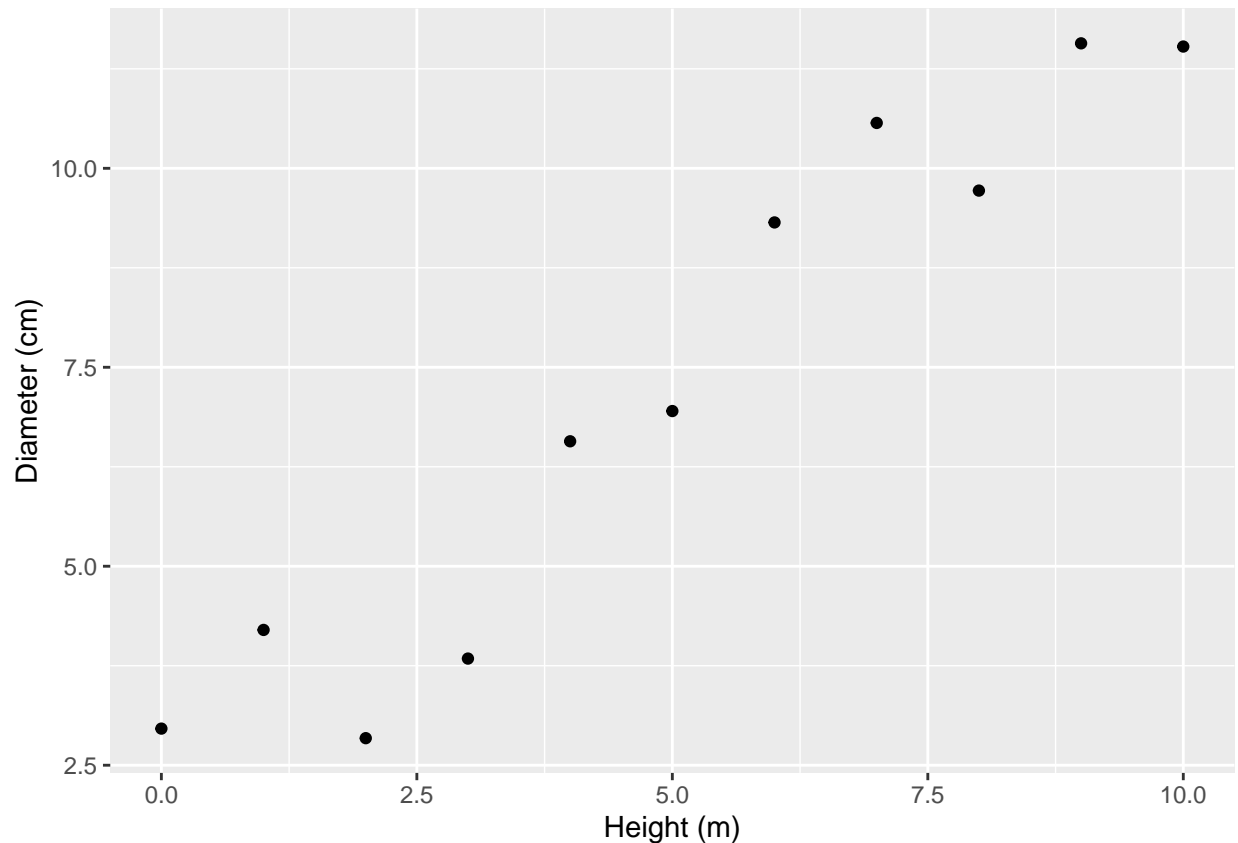
```
sd(tree_data$diameter)
```

```
## [1] 3.43096
```

c) The standard deviation of tree height is 3.31m. The standard deviation of tree diameter is 3.43cm.

Question 3

```
# a)
```

```
ggplot(tree_data,aes(x = height, y = diameter))+
  geom_point()+
  xlab("Height (m)")+
  ylab("Diameter (cm)")
```

b) I hypothesize that as tree height increases so does as well

c) A linear regression with m=1 and b=2 would mean that as tree height increases by 1m the tree diameter would increase by approximately 1cm and that trees of 0m of height would have a diameter of 2cm.

```
# d)

ggplot(tree_data,aes(x = height, y = diameter))+
  geom_abline(intercept = 2, slope = 1, color="lightblue",size=1)+
  geom_point()+
  xlab("Height (m)")+
  ylab("Diameter (cm)")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```