

Coursera Data Scientist Track - Statistical Inference

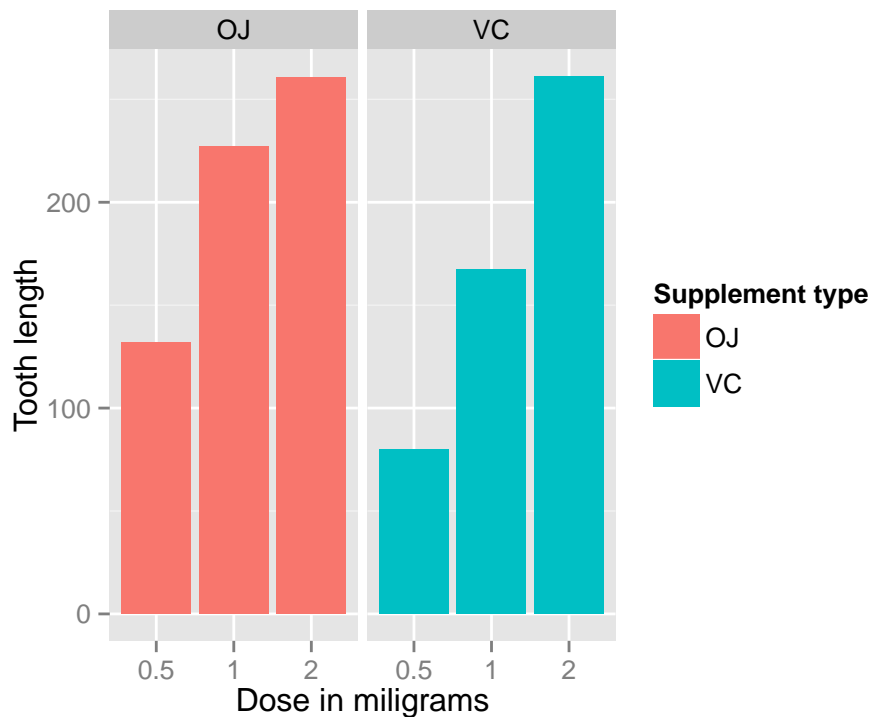
Course Project, Part Two: Basic inferential data analysis

Pieter Jonckheere

Now in the second portion of the class, we're going to analyze the ToothGrowth data in the R datasets package.

Perform data load and have a look at the data.

```
library(datasets)
library(ggplot2)
ggplot(data=ToothGrowth, aes(x=as.factor(dose), y=len, fill=supp)) +
  geom_bar(stat="identity",) +
  facet_grid(. ~ supp) +
  xlab("Dose in miligrams") +
  ylab("Tooth length") +
  guides(fill=guide_legend(title="Supplement type"))
```



As can be seen above, there seems to be a clear positive correlation between the tooth length and the dose levels of Vitamin C, for both supplement types.

[Not completed further hypothesis testing so far.]