STATISTICS WITH R

HOME ASSESSMENT-I

DATA SCIECNE WITH R
CASE STUDY

P SRI SASHANK
208W1A12B3
IT-II

Ch.Nanda Krishna
Assistant Professor
Information technology

PROBLEM :Do planes with a delayed departure fly with a faster average speed to make up for the delay?

>First we have to calculate the average speed of delay planes which was not given in the data

>we have to compute it from the distance and the air time.

>Next, we have to plot the graph between the **Depdelay(given)** and average speed to know the differences in-between them

>Then we have to apply the suitable model to that to know the time difference between them

>Here, we are going to apply the linear model .

>By this we can come to a conclusion with how much speed the plane should when the plan was delayed so that the plan will reach its destination on time.

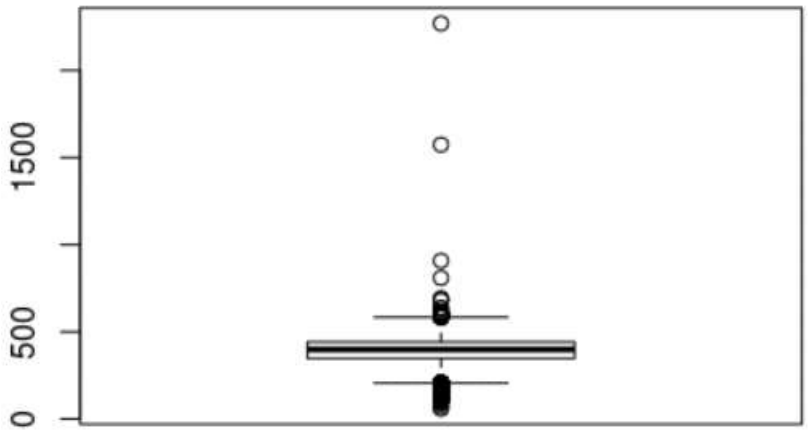
Calculating average speed

```
install.packages("dplyr")
install.packages("ggplot2")
install.packages("mgcv")
library(dplyr)    # Easy data cleaning, and the very convenient pipe ope
library(ggplot2) # Beautiful plots
library(mgcv)     # Package to fit generalized additive models

path="C:/Users/sasha/OneDrive/Desktop/etjical hacking/flight_2008.csv"
flights <- read.csv(path, header=TRUE)
head(flights,n=3)
flights$speed &amp;lt;- flights$Distance / (flights$AirTime/60)
boxplot(flights$speed)
```

Average speed = distance/ change in time

Output : graph

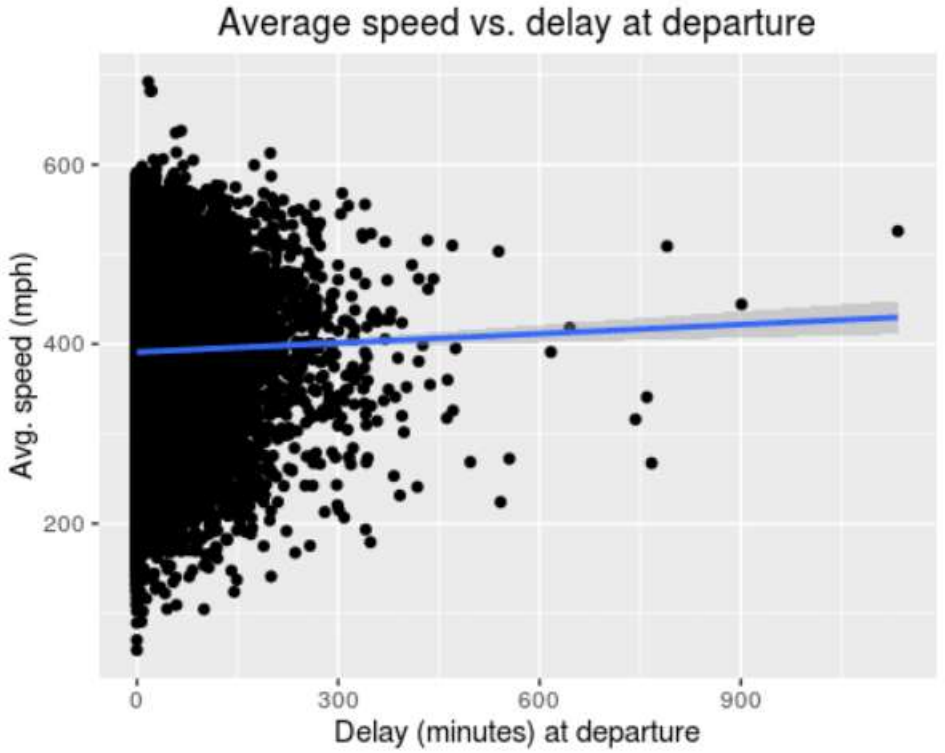


1)Display the graph.
2)calculate the speed.

```
ggplot(flights, aes(x=DepDelay, y=speed)) + geom_point() +
  geom_smooth(method="lm") +
labs(x="Delay (minutes) at departure", y="Avg. speed (mph)",
     title="Average speed vs. delay at departure")
```

Display the graph between the **average speed** and **Depdelay** by using ggplot->geom_points

Output :



It seems like there is a slight increase in average speed for planes that leave with a larger delay. Let's fit a **linear model** to quantify the effect

Linear model can be used when ever the goal can be prediction and when ever we need to know the relationship between two variables.(average speed and Depdelay)

Applying linear model and calculating the speed

```
mod &amp;lt;- lm(speed ~ DepDelay, data=flights) summary(mod)
```

Mod function : returns the remainder of the value
Lm : linear model
Summary :used to produce result summaries of various model fitting functions.

Conclusion :

From the above we can conclude that if the flight is delayed for 60 min it should travel at 2.04mph faster on average.