Scatter plot – best to compare 2 numerical variables.

Y~X

~ is tilda

X is called explanatory variable or independent variable

Y is called predicted variable or dependent variable

Correlation requires linear relationship

Y = mx + c

M = slope = delta y / delta x

C = constant/intercept = y(0)

The line which best describes is the one which passes through mean of Y and X.

Sample date, we use b0 and b1

For population we use beta 1 and beta 2

Lm stands for linear modelling

HW

Pick 2 variables

Simple linear model

Write formula

Interpret the model

One Prediction

Check if the model is statistically significant