STATISTICS FOR DATA SCIENCE

Power Test &
Simple Linear Regression

Dr. Karthiyayini
Department of Science and Humanities

STATISTICS FOR DATA SCIENCE

Unit 5 : Power Test & Simple Linear Regression

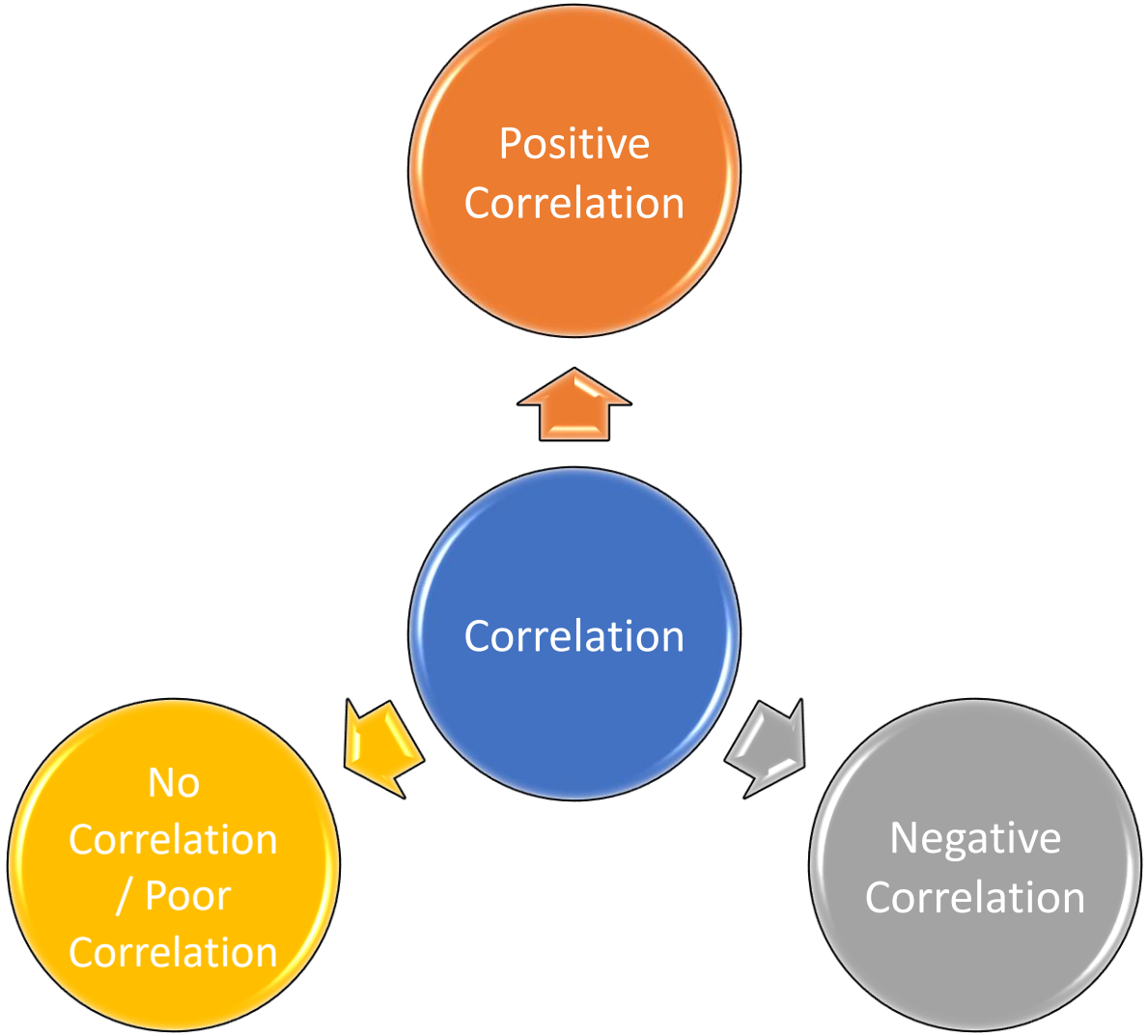
Session : Continued Session

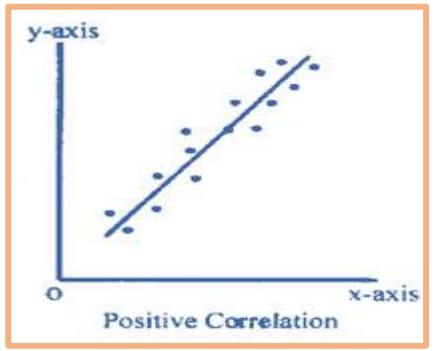
Sub Topic : Correlation

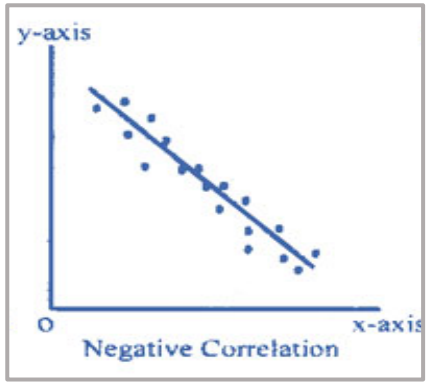
Dr. Karthiyayini
Department of Science & Humanities

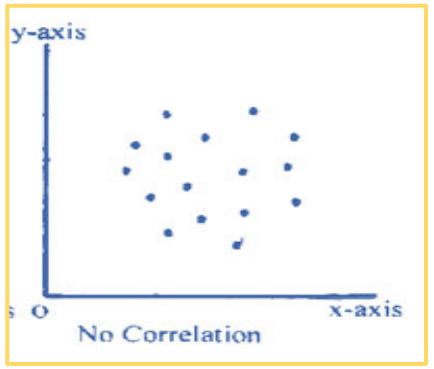
STATISTICS FOR DATA SCIENCE

Classification of Correlation









Source : emathzone.com

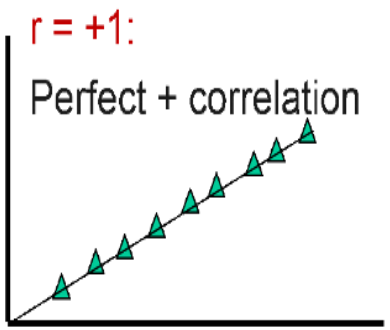
STATISTICS FOR DATA SCIENCE

Some Examples :

Oil Prices	Fight Rates	*Positive Correlation*
No. of hours spent on tread mill	No. of Calories Burnt	*Positive Correlation*
Self Esteem	Depression	*Negative Correlation*
Rise in Temperature	Ice Cream Sales	*Positive Correlation*
Height	IQ	*Zero Correlation*

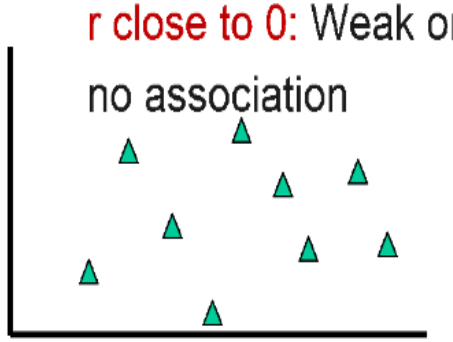
STATISTICS FOR DATA SCIENCE

Interpretation of Correlation Coefficient



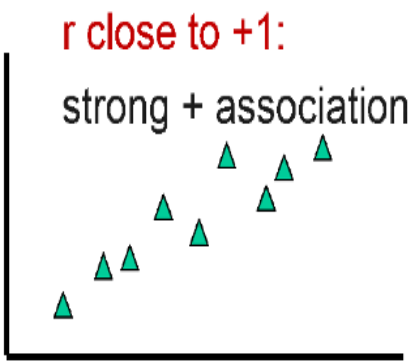
$r = \pm 1 \Longrightarrow$
Perfect Positive / Perfect Negative Correlation

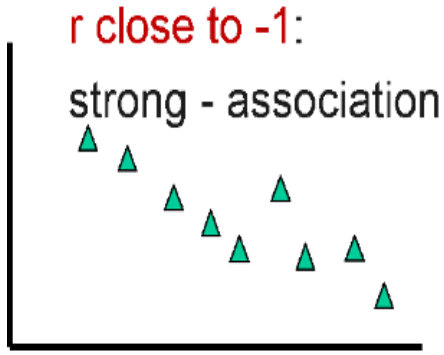
$r = 0 \Longrightarrow$
No Correlation



$0 < r < 1 \Longrightarrow$
Positive Correlation

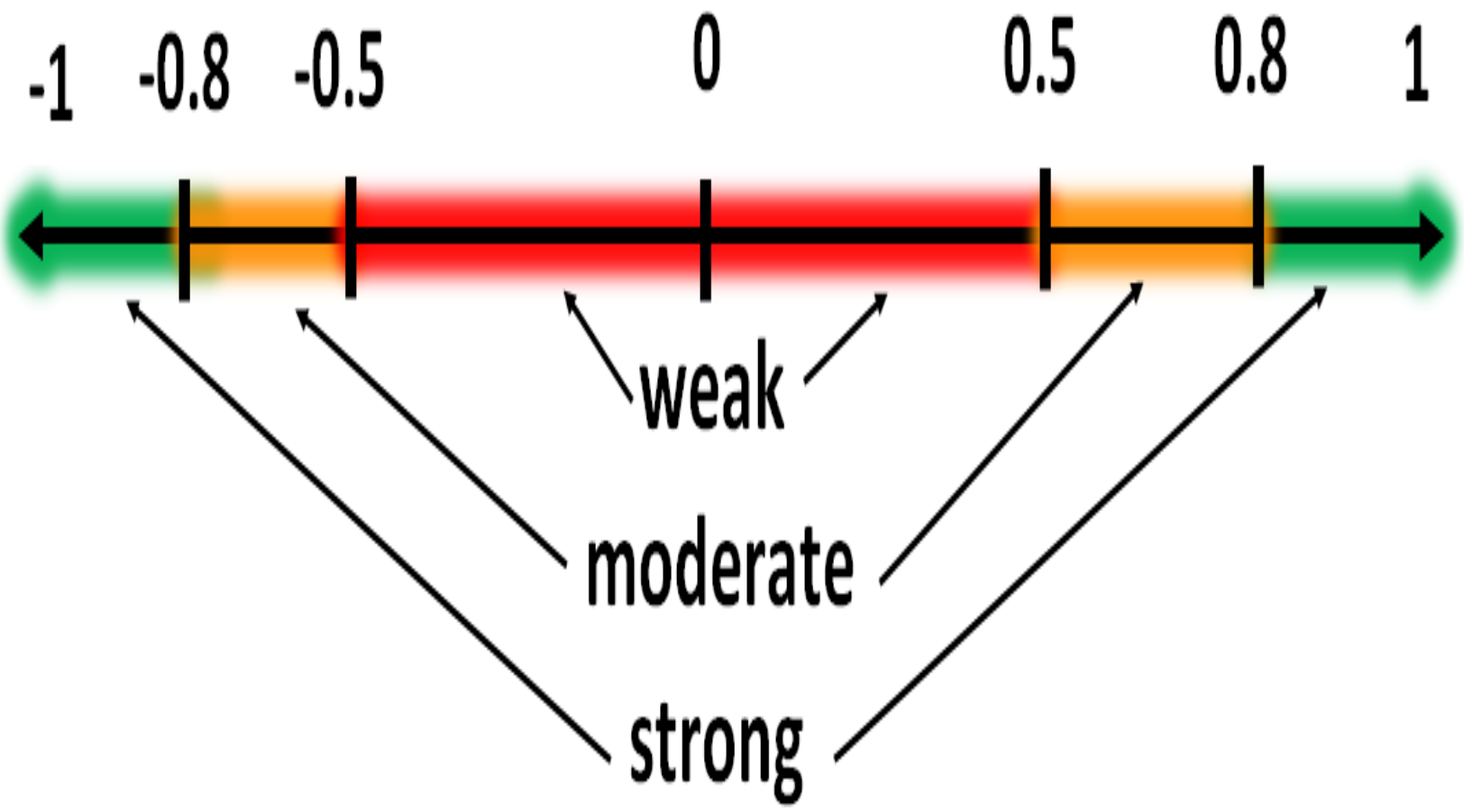
$-1 < r < 0 \Longrightarrow$
Negative Correlation





Source : sphweb.bumc.bu.edu

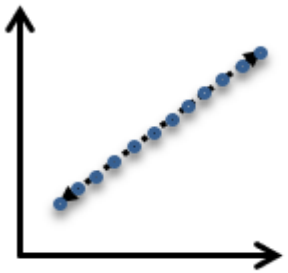
Interpretation of Correlation Coefficient



Source : datavedas.com

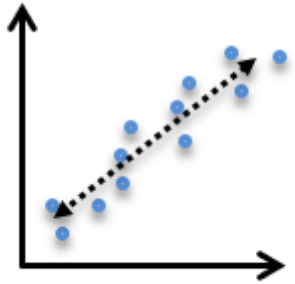
Examples of various levels of correlation

Perfect Positive Correlation



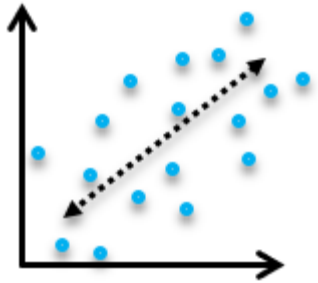
r =1

Highly Positive Correlation



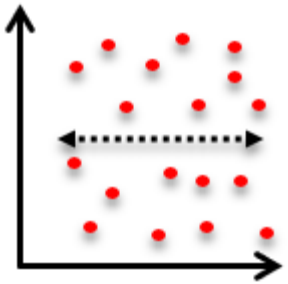
r =0.8

Low Positive Correlation



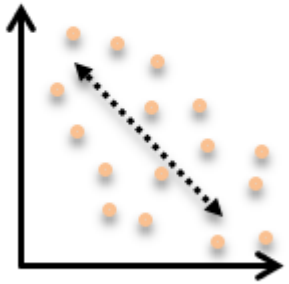
r =0.3

No Correlation



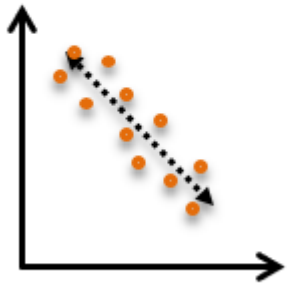
r =0

Low Negative Correlation



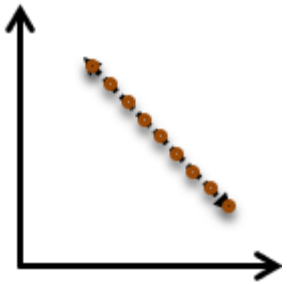
r =-0.3

Highly Negative Correlation



r =-0.8

Perfect Negative Correlation

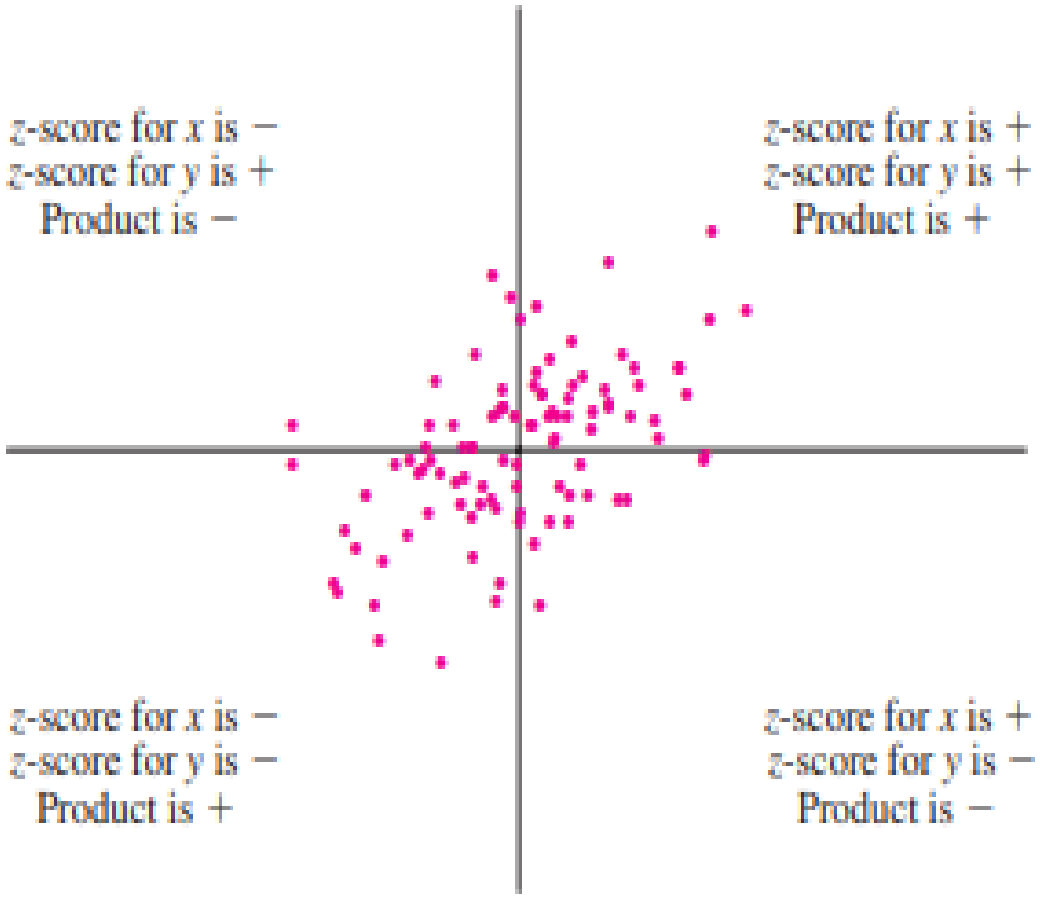


r =-1

Source : https://lytongblog.wordpress.com/

STATISTICS FOR DATA SCIENCE

How the Correlation coefficient works!!



Source : Statistics for Engineers & Scientists, William Navidi

STATISTICS FOR DATA SCIENCE

More about Correlation Coefficient

Correlation Coefficient

Sample Correlation ‘r’

Population Correlation $'\rho'$

Note that Sample Correlation is not only used to measure the strength of a relationship but is also used *to construct Confidence intervals and perform Hypothesis testing* on the population correlation.

Source : shutterstock.com, clker.com

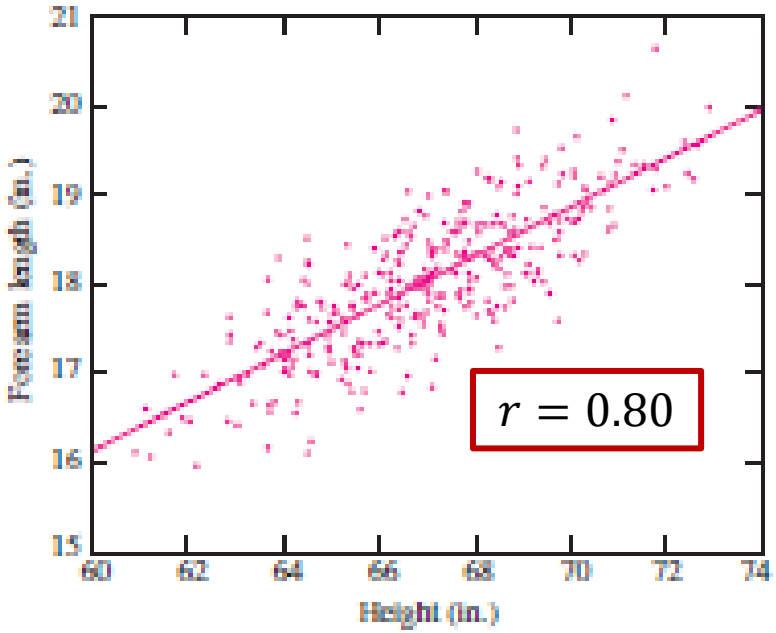
The Correlation Coefficient is unitless!!

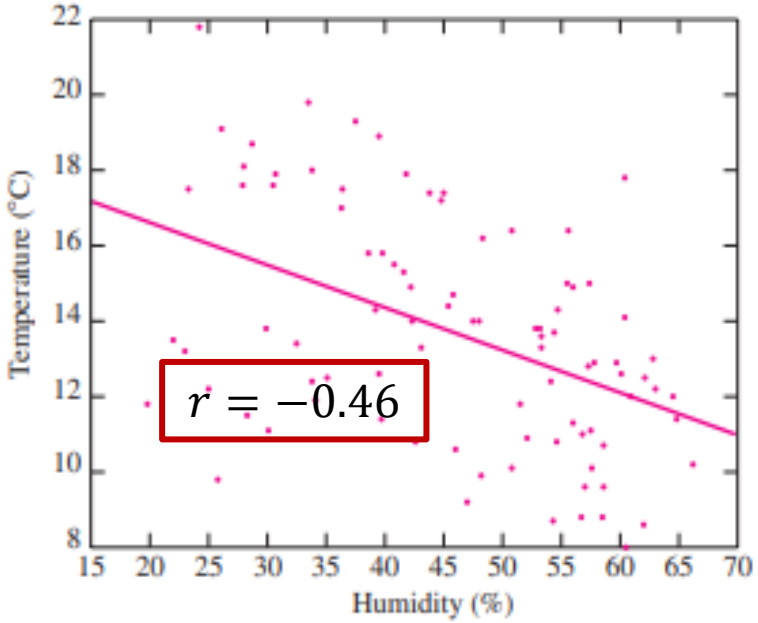
- Consider the Correlation Co-efficient given by,

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S_x}\right)\left(\frac{y_i-\bar{y}}{S_y}\right)$$

have the same units

have the same units.





Source : Statistics for Engineers & Scientists, William Navidi

Some more properties of the Correlation Coefficient

The Correlation coefficient remains unaltered in the following cases :

- Interchanging the values of x and y.
- Adding a constant to each value of a variable
- Multiplying each value of a variable by a positive constant.

PES UNIVERSITY ONLINE

THANK YOU

Dr. Karthiyayini
Department of Science & Humanities

Karthiyayini.roy@pes.edu

+91 80 6618 6651