



Data Exploration

Stages of Data Exploration



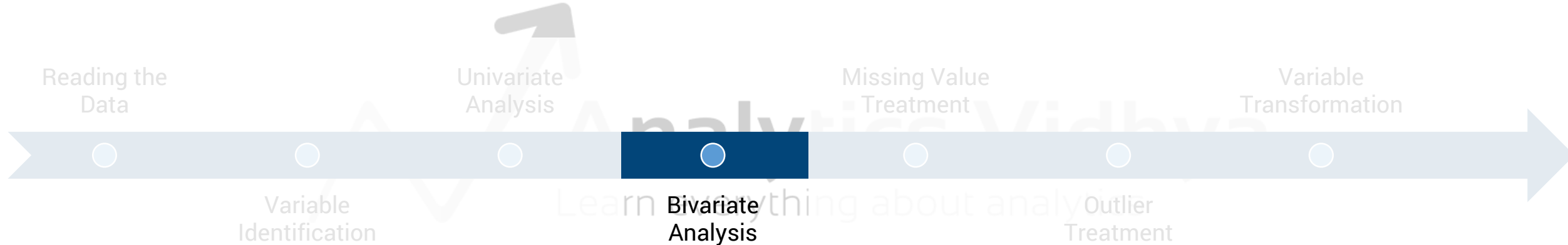
Stages of Data Exploration



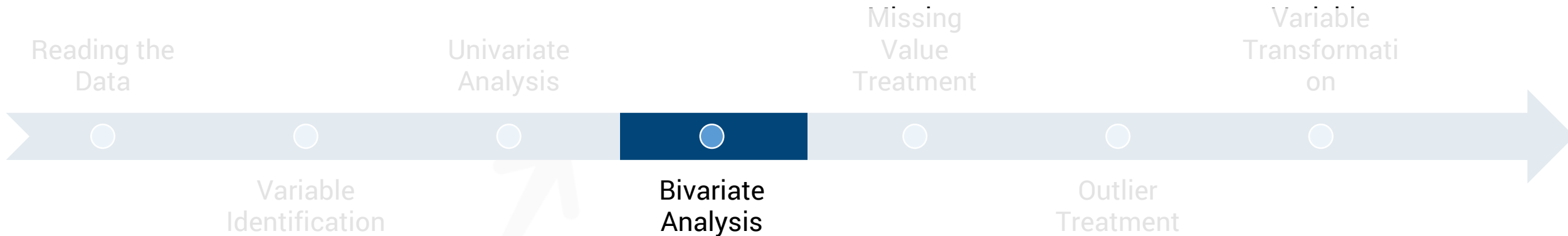
Stages of Data Exploration



Stages of Data Exploration



Stages of Data Exploration



Topics to be covered

1. What is Bivariate Analysis?
2. Why do we need Bivariate Analysis?
3. How do we perform Bivariate Analysis?

Analytics Vidhya
Learn everything about analytics

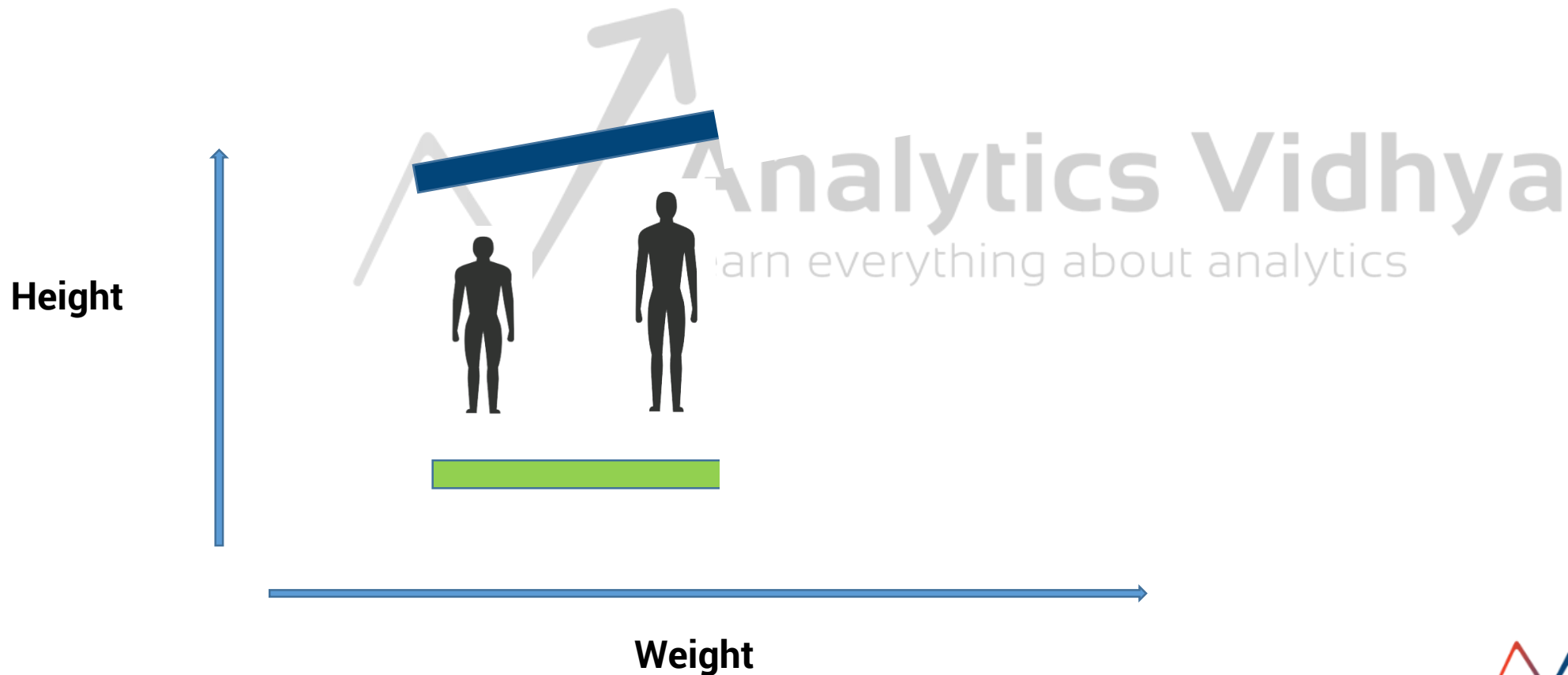
Bivariate Analysis

What is Bivariate Analysis?



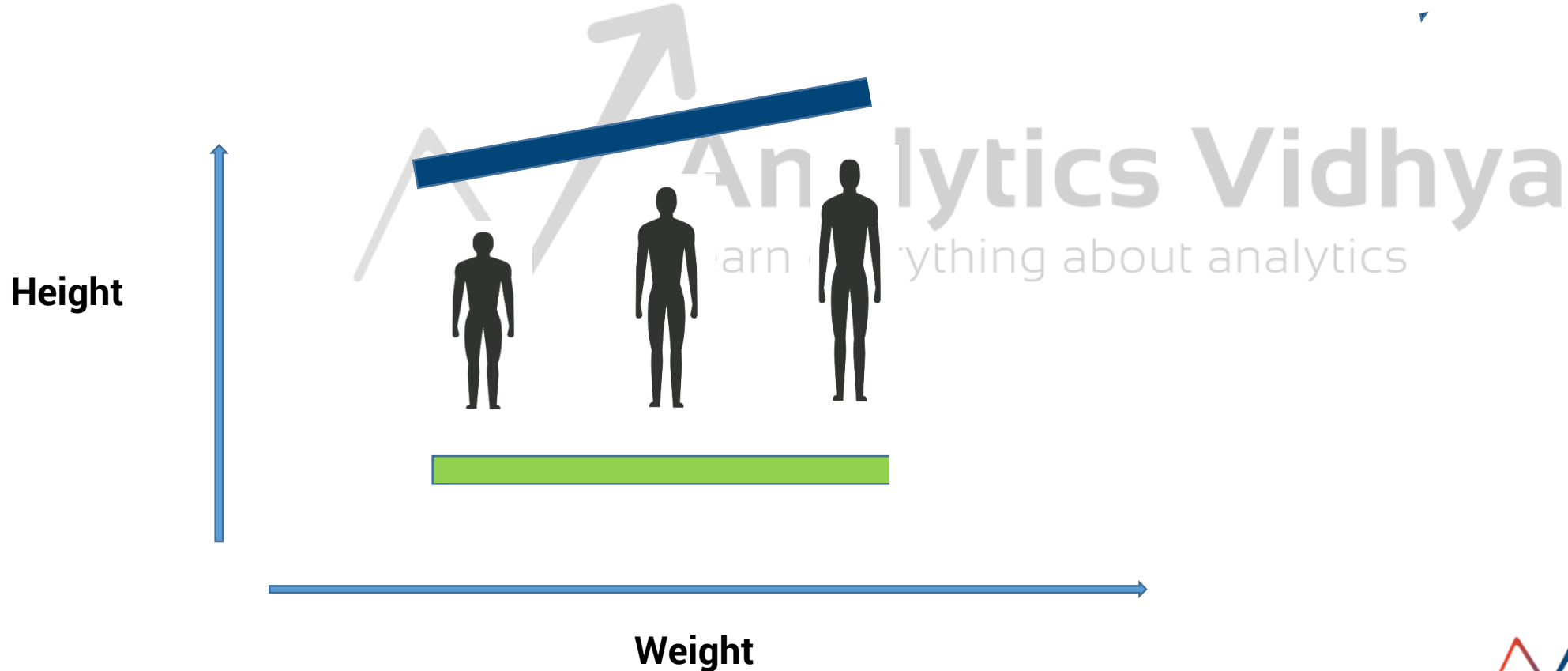
Bivariate Analysis

What is Bivariate Analysis?



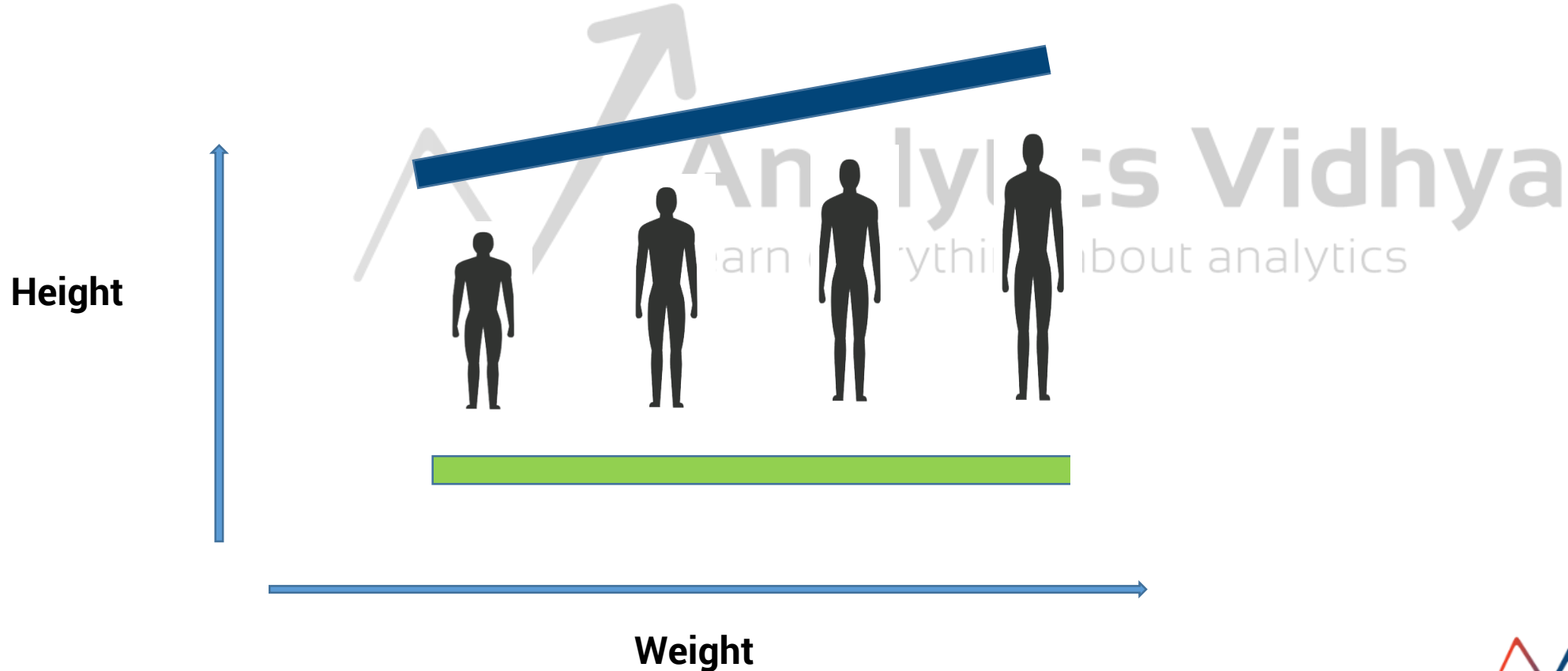
Bivariate Analysis

What is Bivariate Analysis?



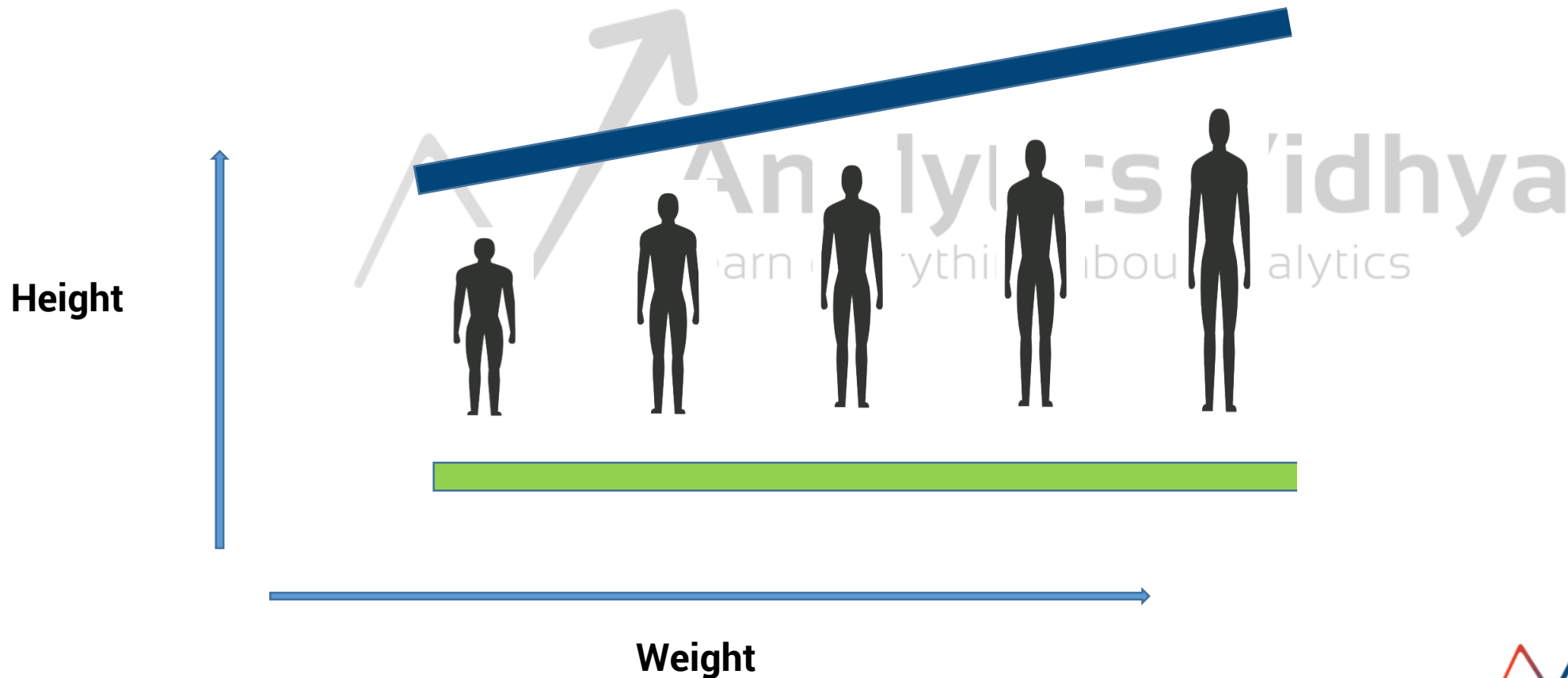
Bivariate Analysis

What is Bivariate Analysis?



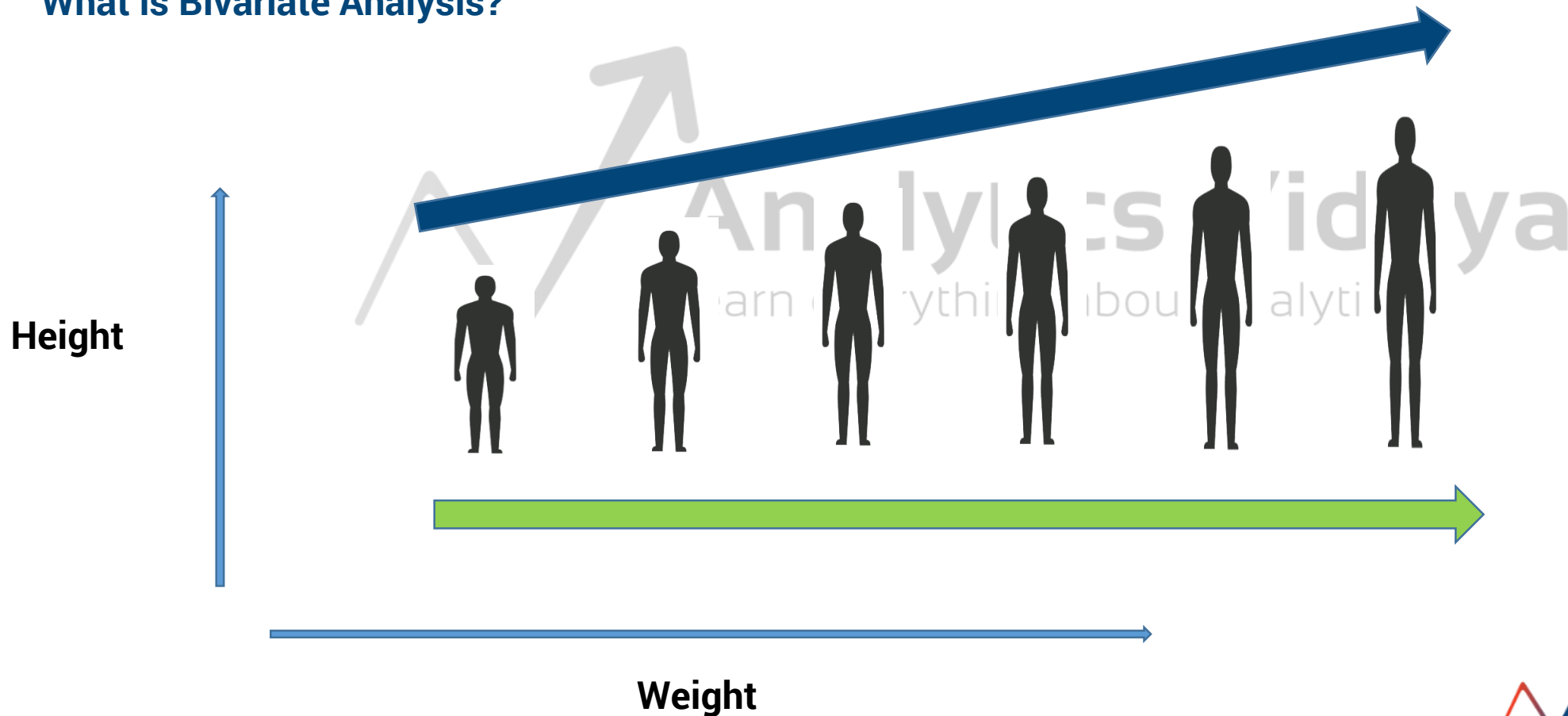
Bivariate Analysis

What is Bivariate Analysis?



Bivariate Analysis

What is Bivariate Analysis?

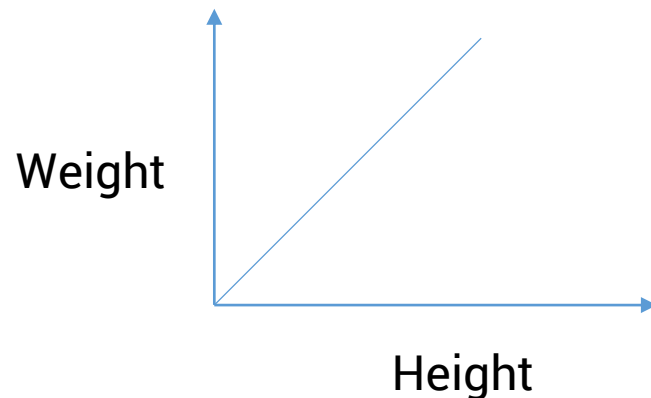


Bivariate Analysis

What is Bivariate Analysis?

- When two variables are studied together for their empirical relationship.
- When you want to see whether the two variables are associated with each other.

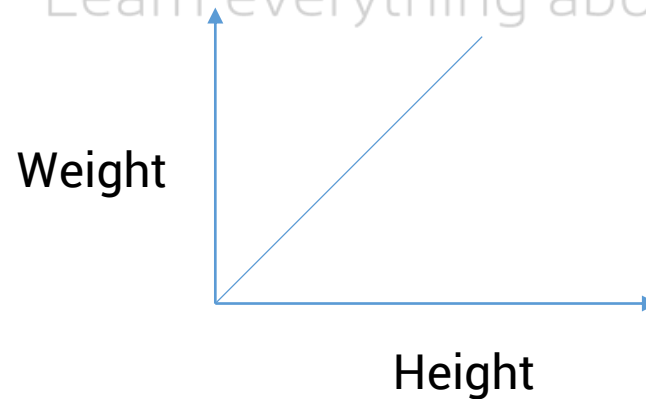
For example – Increase in height is associated with increase in weight and vice versa.



Bivariate Analysis

Why do we need Bivariate Analysis?

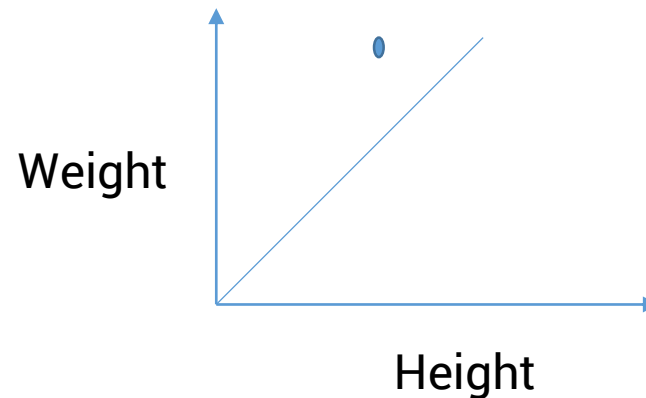
- It helps in prediction. [when two variables are associated one may be used to infer the other.]



Bivariate Analysis

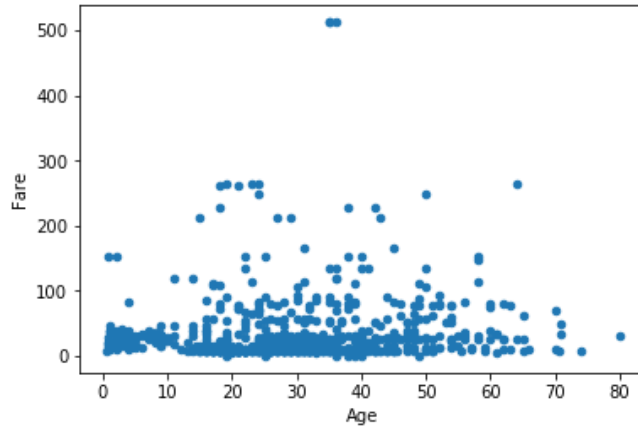
Why do we need Bivariate Analysis?

- It helps in detecting anomalies.

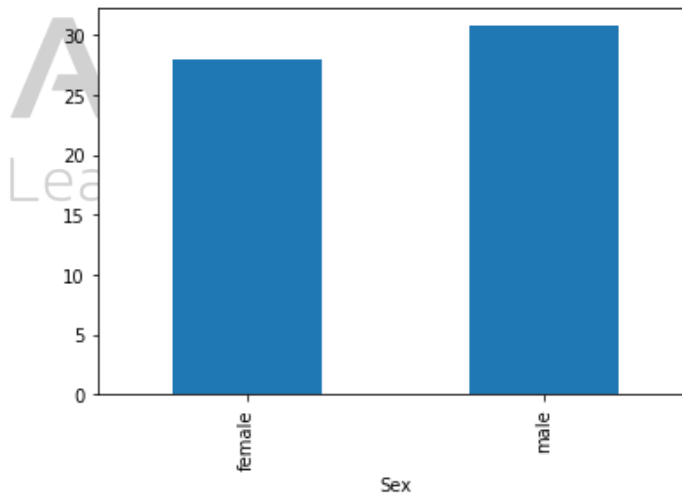


Bivariate Analysis

Types of Bivariate Analysis



Age-Fare price



Gender – Mean Age

Survived	0	1
Sex		
female	81	233
male	468	109

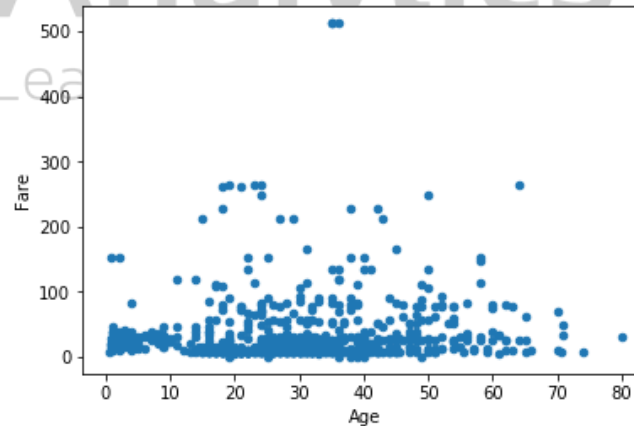
Gender – survived

How do we perform Bivariate Analysis?

Continuous – Continuous Variables

Does the Fare of ticket increases with increase in age?

Graphical Method



Scatter plot

How do we perform Bivariate Analysis?

Continuous – Continuous Analysis

Does the price of ticket increases with increase in age?

Analysis Test

$$r = \frac{\text{Cov}(x, y)}{\sqrt{s_x^2 * s_y^2}}$$

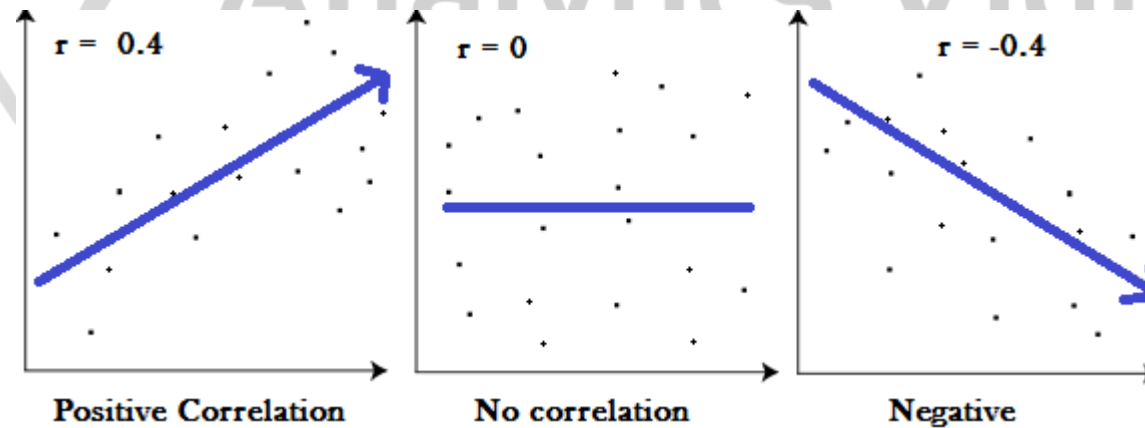
Correlation

How do we perform Bivariate Analysis?

Continuous – Continuous Analysis

Does the price of ticket increases with increase in age?

Analysis Test



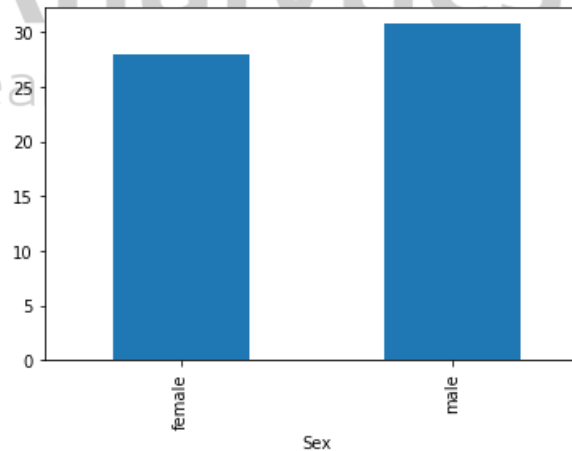
Correlation

How do we perform Bivariate Analysis?

Categorical – Continuous Analysis

Is the mean age of males different from the mean age of females?

Graphical Method



Gender – Age mean bar plot

How do we perform Bivariate Analysis?

Categorical – Continuous Analysis

Is the mean age of males different from the mean age of females?

Analysis Test

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

2- sample T-test

\bar{y}_1 - Mean age of males

\bar{y}_2 - Mean age of females

s_1^2 - Variance of males

s_2^2 - Variance of females

How do we perform Bivariate Analysis?

Categorical – Categorical Analysis

Does gender have any effect on the survival rates?

Graphical Method

Sex	Survived	
	0	1
female	81	233
male	468	109

Two-way table

How do we perform Bivariate Analysis?

Categorical – Categorical Analysis

Does gender have any effect on the survival rates?

Analysis Test

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

χ^2 = the test statistic \sum = the sum of

O = Observed frequencies E = Expected frequencies

Chi-square test

