AEROFIT
CASE STUDY
REVIEW
Class starts
@ 9:10pm

Agenda

1) Problem Statement

2) Exploring the data (EDA)

3) Conclusion

4) Q/A

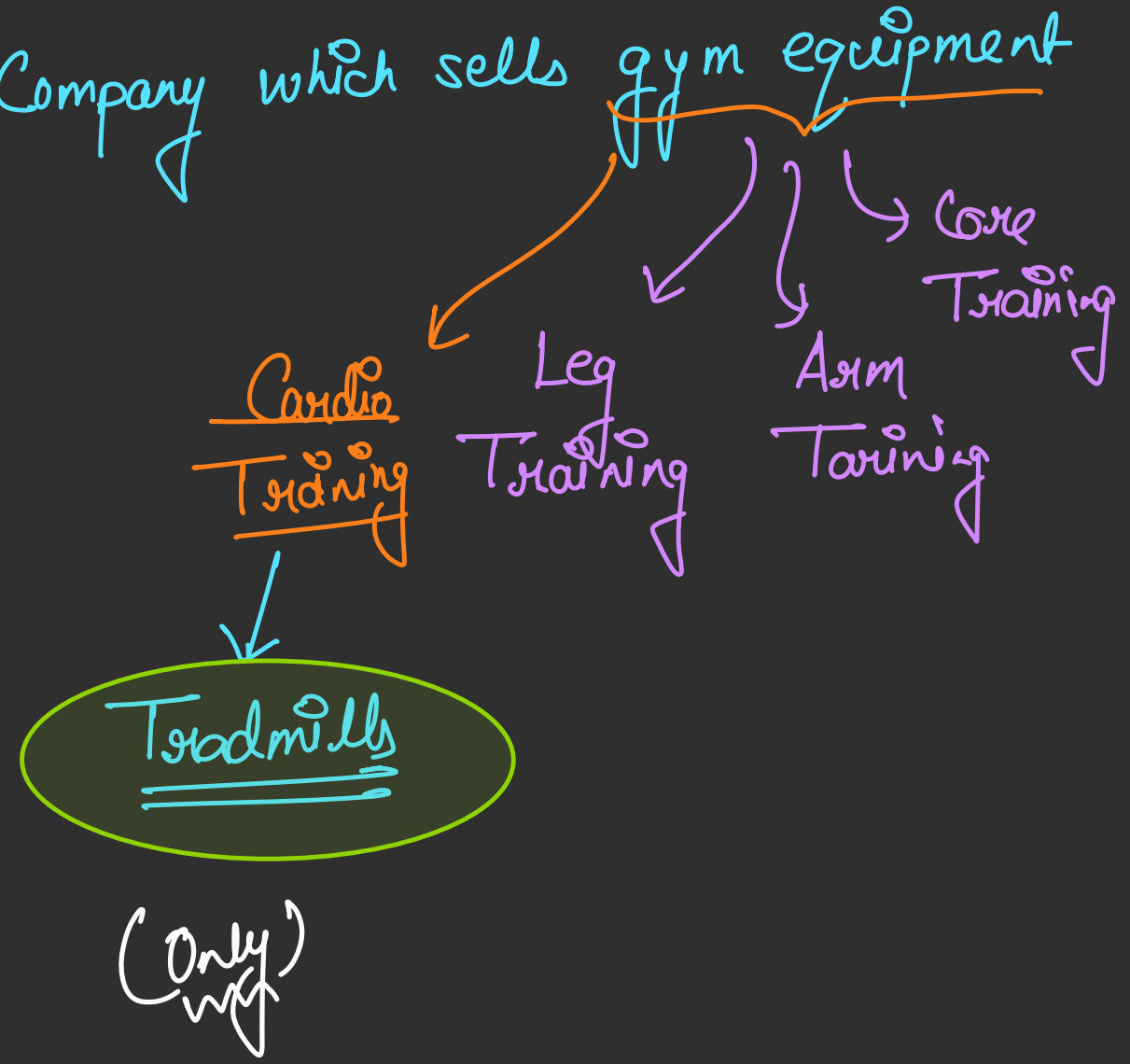
1.5 h

10:30pm

problem Statement

Business Understanding → Company which sells gym equipment

(Aerofit)



Products

- Kp781
- Kp481
- Kp281

problem statement ?

→ Understand the problem

1. Scoping → What is the problem Statement?
   → Do we have all the required data?
   → Are those data useful?

2. Data Analysis → EDA (column profiling, Null value identification/treatment, Outlier treatment/ Colcum creation, Viz)

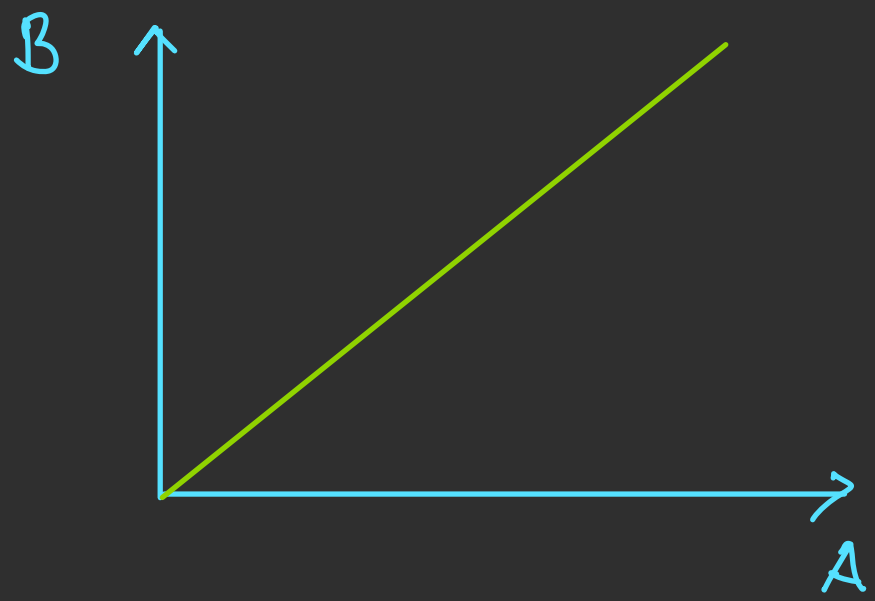
3. Observation/Insights & Recommendation

Your analysis result

Your suggestion Action Items for the leadership

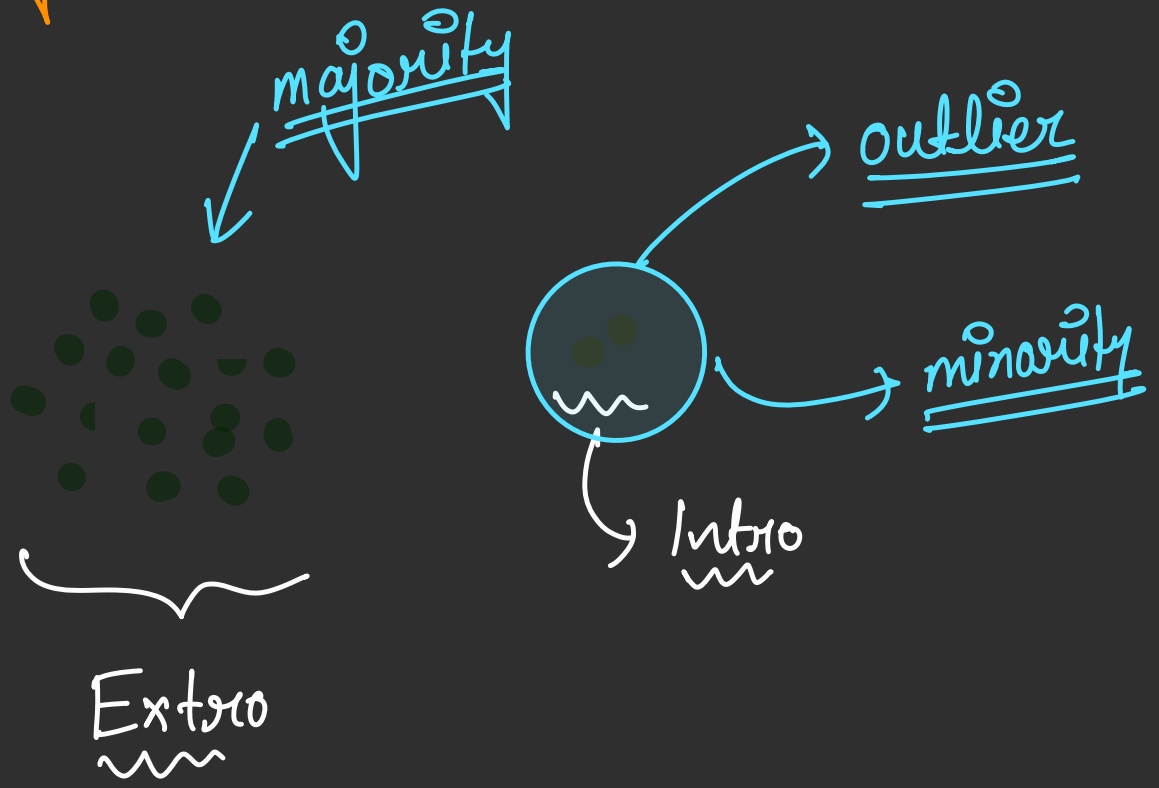
$A \propto B$

$A \uparrow \quad B \uparrow$



Outlier

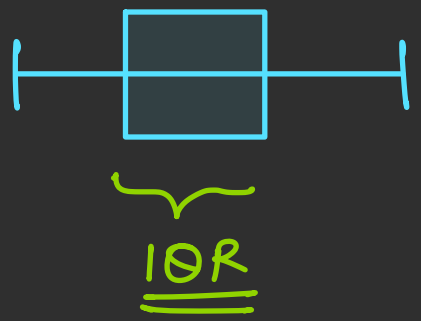
↳ Are lonely data points



1) IQR Method

$$\begin{cases} Q_1 \rightarrow 25^{th} \text{ per} \\ Q_2 \rightarrow 50^{th} \text{ per} \\ Q_3 \rightarrow 75^{th} \text{ per} \end{cases}$$

$$IQR = Q_3 - Q_1$$

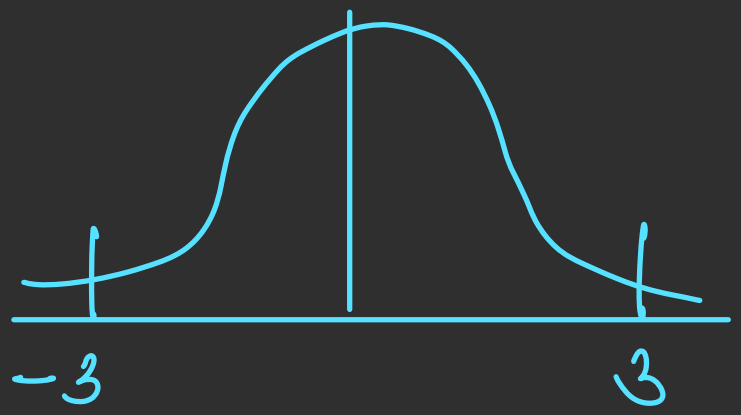


$$\begin{cases} x_1 = Q_1 - 1.5 \times IQR \\ x_2 = Q_3 + 1.5 \times IQR \end{cases}$$

2) Z score

$$z = \frac{x - \mu}{\sigma}$$

col → z score →



df

↳ Income

0 — 89000

Mean

OR

df_copy

↳ Inocne

0 — 6000

Mean'

1) If you have a large dataset, where you can sacrifice some data pts.

2) Outlier are anomlyies, rare cases. If the data pts are actually more then remove it