

Data cube – data used for analysis; whether proper cleaning have been done/ consistent or missing data

Exploratory Analysis First to gain enough insights; then we do proceed to other phase (i.e association rule mining)

Which has highest sale value/volume by product categories or by sales channel or by time pattern.

Are they consistent or varies?

What is the sales revenue by region / categories.

DO DESCRIPTIVE STATISTICS FIRST. Fit X by Y analysis by then... Example:

Within SH have diff outlet, when sales of shoes, not all outlet give same revenue, at least 2 stores with sales revenue higher than the rest... use fit x by y to show (annova). Understand variation.

create any month column