

SPR^{CKET} CENTRAL

Data journey

I have summarised the data journey since the beginning and made it simple for stake holders to understand what is happening with the company.



Data exploration

In this step we did a screening of the dataset provided by the client. We performed the following activities:

1. Removal of illogical values.
2. Correcting and formatting the values.
3. Dropping empty rows and columns from the dataset.

The cleaning of dataset is the first part of the analysis as it will help us have clean and consistent data which will let us have better insights and hence better decisions.



Data transformation

In this step we are going to talk about the modifications we made to the dataset in form of feature engineering. It means we created new variables to aid in our analysis.

The variable Age, profit, profit% etc were created using the existing variables and using some formulae so that we can figure out what is happening in the business.

Insights gathered



The following insights have been gathered from the dataset:

1. The standard product line generates 75% of the profit and mountain product line generates the least profit.
2. The top ten customers are 3040, 1266, 3024, 3159, 363, 569, 2557, 1612, 287 and 2745 as they give the highest profit and revenue to the company.
3. WeareA2B is the most profitable brand and the least profitable brand is Norco.
4. The amount of online and offline purchases are almost same.
5. New South Wales is the state with the highest business volume.
6. Female and Male customers have almost the same age and tenure.
7. Medium product class is the most costliest.
8. Medium order size is having the highest average profit percentage.