INT353

EDA PROJECT

Name   :   Aditya Kumar Pandey

Reg No.   :   12018602

Course code   :   INT-353

Data set   :   Sample super store

Project introduction: -

I have chosen a data set from Kaggle. It is about a super store in the United States. It contains 13 columns, namely ship mode, segment, country, city, state, postal code, region, categories of office supplies, furniture, and technology, subcategories like Binder's paper, phones, art, furnishings, envelops, bookcases, chairs, phones, storage, labels, accessories, fasteners, tables, supplies, machines, appliances, sales column, quantity, discount, and finally profit.

Why I chosen this data set: -

I mainly chose this data set to analyse how a store works under different situations and how they tackle profit or loss. While they are in profit, how they get profit and how to increase it further, and if they are in loss, how they tend to increase the sales and background work like how they analyse to make profit and improve sales. Not only for one store, a huge number of stores near 10,000.

Domain: -

A superstore is a large retail store operated on a self-service basis, selling groceries, fresh produce, bakery, and dairy products, and sometimes an assortment of non-food goods.

About my data set: -

It is all about sample superstores in the United States. It contains 13 columns Those are as follows:-

- Ship mode it contains the class of shipping modes they are standard class, second class, and first class.
- Segment it contains segment categories they are consumer, corporate, and home office.
- Country it contains the country of the super store in this data set all the super stores are taken from united states only.
- City it contains city of the super store.
- State it contains state of the super store.
- Postal code.
- Region, it contains the region that is from West or east or central or south or north.
- Category it contains the type of products, In it three categories are there office supplies, furniture, and technology.
- Sub-category it contains different products depend on that category of the super store.
- Sales it contains number of times sold depend on sub-category.

- Quantity it contains the number products sold at a time depend on sub-category.
- Discount it contains how much discount got on an average of sales, quantity.
- Profit.

Analysis I will do: -

- I will analysis the sales of a certain products and its profit and discount that the super storeproviding.
- I will analysis categories and sub-categories depend on sales and profit so that I will get the relation on sales and profit, and in which categories the sales is more and profit also more, orsales less and profit more, or sales less and profit also less.
- I will analysis the profit or lose by giving discounts.
- I will analysis sub-categories that which products are more in sales and in which products getting more profits, and how the super stores giving the discount depend on the product andcategory.
- I will analysis profit and sales depend on shipping mode.
- I will analysis segment that in which segment category how the sales and profits going on.
- I will analysis which region is attractive by seeing sales and in which region the profits aremore.