

Project Proposal

(Superstore Sales)

Description:

The Super Store dataset contains data on order details of customers and Statistics on the products purchased, the cities and the date on which they were purchased

The dataset getting Kaggle website:

<https://www.kaggle.com/rohitsahoo/sales-forecasting>

Data:

Superstore sales dataset contain 18 features and 9800 unique values.

A few feature highlights include the row id ,order id, order date , ship date, ship mode, customer id, customer name, segment , country , city , state, postal code , region, product id , category, sub category, product name , sales.

Goal:

The data shows the types of products that people buy and date of the ship and sales of each product and the goal of these is statistics most products buy and solve questions.

Problem:

This dataset it is not statistics on what's the most purchased product and which city shop the most.

I need I want to solve these two problems:

What is the most purchased product?

People from city shop the most?

well, going to solve this through the dataset.

Tools:

Numpy and Pandas for data manipulation , Scikit-learn for modeling , Matplotlib and Seaborn for plotting , Tableau for interactive visualizations.