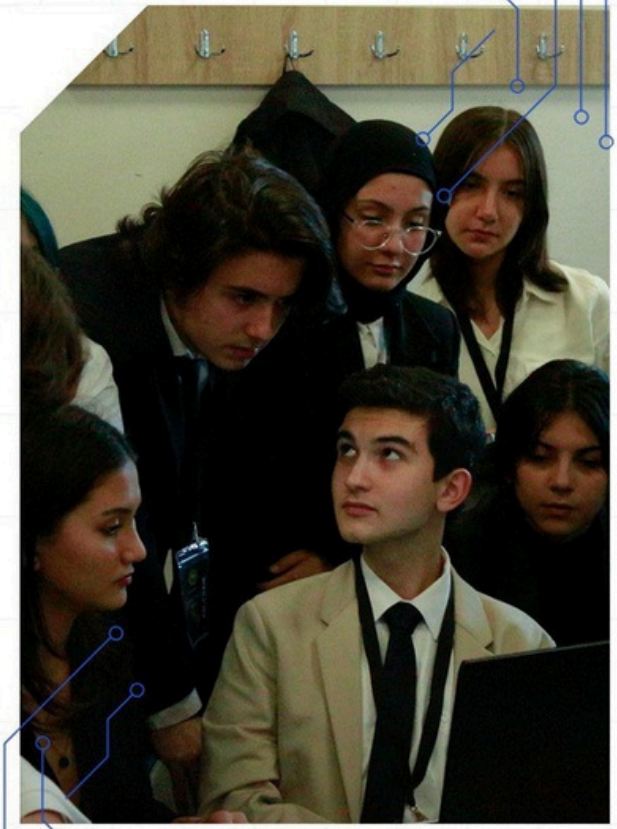
VISUALIZING SHOPPING TRENDS WITH POWER BI

This presentation showcases our project aimed at creating an insightful Power BI dashboard to analyze shopping trends.

RAGINI KATIYAR



OUR TEAM MEMBERS

Meet the talented individuals driving our success

RAGINI KATIYAR
MOHAMED YASIR
ABHINAYA
SELSHIYA
VIDHI JAIN

INTRODUCTION TO POWER BI

Power BI is a Microsoft business analytics tool that helps visualize data, create interactive dashboards, and generate reports for data-driven decision-making. It connects to various data sources, allows for data transformation, and integrates seamlessly with Microsoft tools. Users can share insights, collaborate in the cloud, or deploy on-premises. Power BI is widely used for tasks like financial reporting, sales analysis, and operational monitoring.

OVERVIEW

INTRODUCTION

This section provides an overview and sets the stage for the upcoming topics and discussions.

Customer segmentation is the process of dividing a customer base into distinct groups based on shared characteristics, behaviors, or preferences. It helps businesses tailor their marketing, products, and services to meet the specific needs of each group. Common segmentation criteria include demographics, geographic location, psychographics, and purchase behavior. By understanding these groups, companies can enhance customer satisfaction and improve profitability.

DATASET ANALYSIS

EXPLORING SHOPPING TRENDS DATASET

In-depth examination of shopping trends utilizing data collected from Kaggle for insightful analysis.

Dataset Summary for Shopping Trends

The Shopping Trends dataset includes detailed consumer purchasing information. Key columns are:

- **Demographics:** Customer ID, age, gender (male/female).
- **Product Details:** Item purchased, category (accessories, clothing, footwear, outerwear), size, color, and season.
- **Transaction Details:** Purchase amount (USD), location, review ratings, subscription status (yes/no), payment methods (bank transfer, cash, credit card, debit card, PayPal, Venmo), shipping type, discount applied, and promo code usage.

- **Customer Behavior:** Previous purchases (integer), preferred payment methods, and frequency of purchases categorized (bi-weekly, monthly, annually, etc.) and numerically processed.
- **Processed Features:** An added "age2" column categorizes customers into young, middle-aged, and old based on age, and a numerically processed "frequency of purchases" for analysis.

DATA PREPROCESSING

Essential Techniques for Preparing Data for Analysis

REMOVING NULL VALUES

One of the essential steps in data preprocessing involves identifying and removing null values from the dataset. Null values can skew analysis and lead to incorrect conclusions, making it crucial to address them before proceeding with data analysis.

CHANGING DATA TYPES

Adjusting data types is vital for ensuring that each column in the dataset is represented in the appropriate format. For example, converting string representations of dates into actual date objects allows for time-series analysis and other time-dependent operations.

HANDLING OUTLIERS

Outliers are data points that deviate significantly from other observations. Identifying and managing outliers is crucial as they can distort statistical analyses and lead to misleading interpretations. Techniques such as trimming or transforming data can be applied.

NORMALIZATION AND STANDARDIZATION

Normalization and standardization are techniques used to scale data. Normalization adjusts the values in the dataset to a common scale without distorting differences in the ranges of values, while standardization transforms data to have a mean of zero and a standard deviation of one.

ENCODING CATEGORICAL VARIABLES

Categorical variables need to be converted into a numerical format for machine learning models. Techniques like one-hot encoding or label encoding allow algorithms to process categorical data effectively.

SPLITTING DATA INTO TRAINING AND TESTING SETS

To assess the performance of machine learning models, it is crucial to split the data into training and testing sets. This ensures that the model is trained on one subset of the data and validated on another, reducing the risk of overfitting.

FEATURE SCALING

Feature scaling ensures that each feature contributes equally to the distance computations in algorithms. Techniques such as Min-Max scaling or Z-score standardization can be employed to adjust the scale of features.

DATA TRANSFORMATION

Data transformation involves applying mathematical functions to the dataset, such as logarithmic or square root transformations, to stabilize variance and make the data more normally distributed for analysis.

DATA IMPUTATION

Imputation fills in missing values with substituted values, allowing for more complete datasets. Methods include using the mean, median, or mode of the variable, or more sophisticated techniques like K-Nearest Neighbors imputation.

DATA TRANSFORMATION TECHNIQUES

Exploring essential techniques for effective data transformation and analysis.

ADDING NEW COLUMNS

Adding new columns to a dataset allows analysts to introduce additional attributes or metrics that can provide insights and improve data analysis. For instance, you might add a 'Sales Growth' column that calculates the percentage increase in sales over time, enabling more detailed performance tracking.

CREATING NEW MEASURES

Creating new measures involves defining calculations that summarize data in meaningful ways. This can include aggregations like sums, averages, or more complex calculations like profit margins, which help stakeholders understand performance metrics at a glance.

UTILIZING DAX FUNCTIONS

DAX (Data Analysis Expressions) functions are powerful tools used in data transformation to perform calculations and data manipulation in tools like Power BI. For instance, using the CALCULATE function allows users to modify filter contexts when performing calculations, enhancing the analytical capabilities of the data model.

PRESENTATION OVERVIEW

PROJECT PRESENTATION OVERVIEW

An in-depth look at a Power BI Dashboard analyzing shopping trends and consumer behavior.

CREATING VISUALIZATIONS

Explore different visualization techniques to gain better customer insights

LINE CHART

A line chart is used to display information that changes over time, making it ideal for showing trends and patterns in data. For example, tracking sales growth month over month.

DONUT CHART

Similar to pie charts, donut charts visually represent data as a percentage of a whole, but with a central hole. This makes it easier to read and compare between different segments.

BAR CHART

Bar charts are effective for comparing quantities across different categories. They can be oriented vertically or horizontally, making them versatile for various data presentations.

PIE CHART

Pie charts provide a visual representation of data in a circular format, where each slice represents a proportion of the whole. They are best used when you have a limited number of categories.

CARDS

Cards are a compact way to present data insights, often used in dashboards. Each card can display key metrics, trends, or visualizations for quick assessments.

MAP TREE

Map trees visualize hierarchical data, showing relationships between categories and subcategories. This is particularly useful for understanding complex datasets.

COMPOSITION TREE

A composition tree breaks down data into components and subcomponents, allowing users to see how the overall data is structured and where different segments fit.

MAP

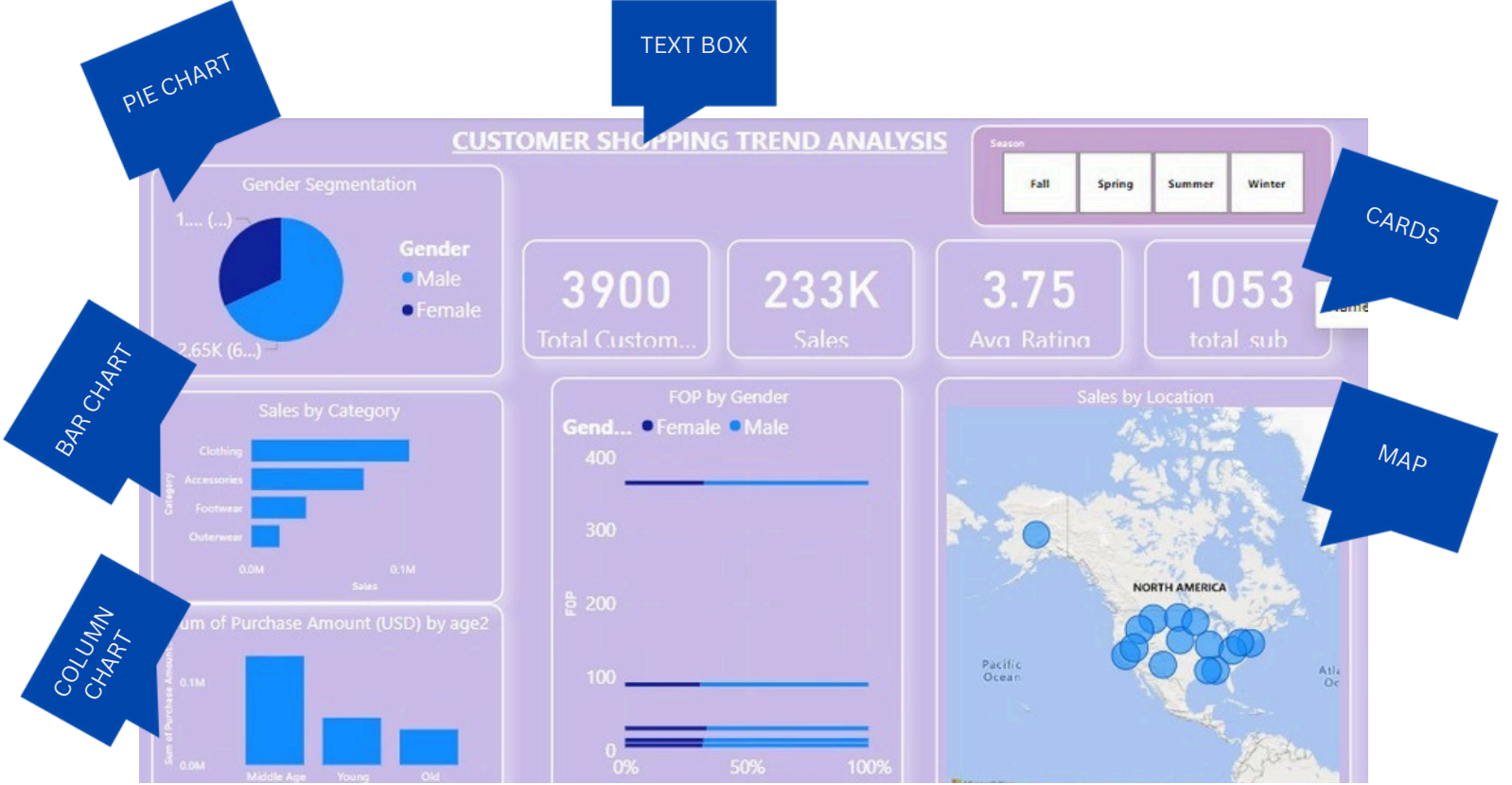
Maps are essential for geographical data representation, helping to visualize data points based on location, such as sales regions or customer demographics.

MATRIX

Matrices allow for the representation of data in a grid format, making it easy to see relationships and correlations between different variables.

RIBBON CHART

Ribbon charts display relationships between categories over time, showing how values change and flow from one category to another, useful for trend analysis.


PIE CHART
TEXT BOX
CARDS
BAR CHART
MAP
COLUMN CHART
CUSTOMER SHOPPING TREND ANALYSIS
Gender Segmentation
Gender
Male
Female
Season
Fall
Spring
Summer
Winter
3900
Total Custom...
233K
Sales
3.75
Avg Rating
1053
total sub
Sales by Category
Clothing
Accessories
Footwear
Outerwear
FOP by Gender
Female
Male
Sales by Location
NORTH AMERICA
Pacific Ocean
Middle Age
Young
Old

MATRIX

Category	2-Day Shipping	Express	Free Shipping	Next Day Air	Standard	Store Pickup	Total
Accessories	12770	12419	11676	12335	12085	12915	**74200**
Clothing	16890	17702	18224	17185	17839	16424	**104264**
Footwear	5593	5746	7283	5669	5758	6044	**36093**
Outerwear	2827	3200	3594	2804	2551	3548	**18524**
Total	**38080**	**39067**	**40777**	**37993**	**38233**	**38931**	**233081**

Q&A

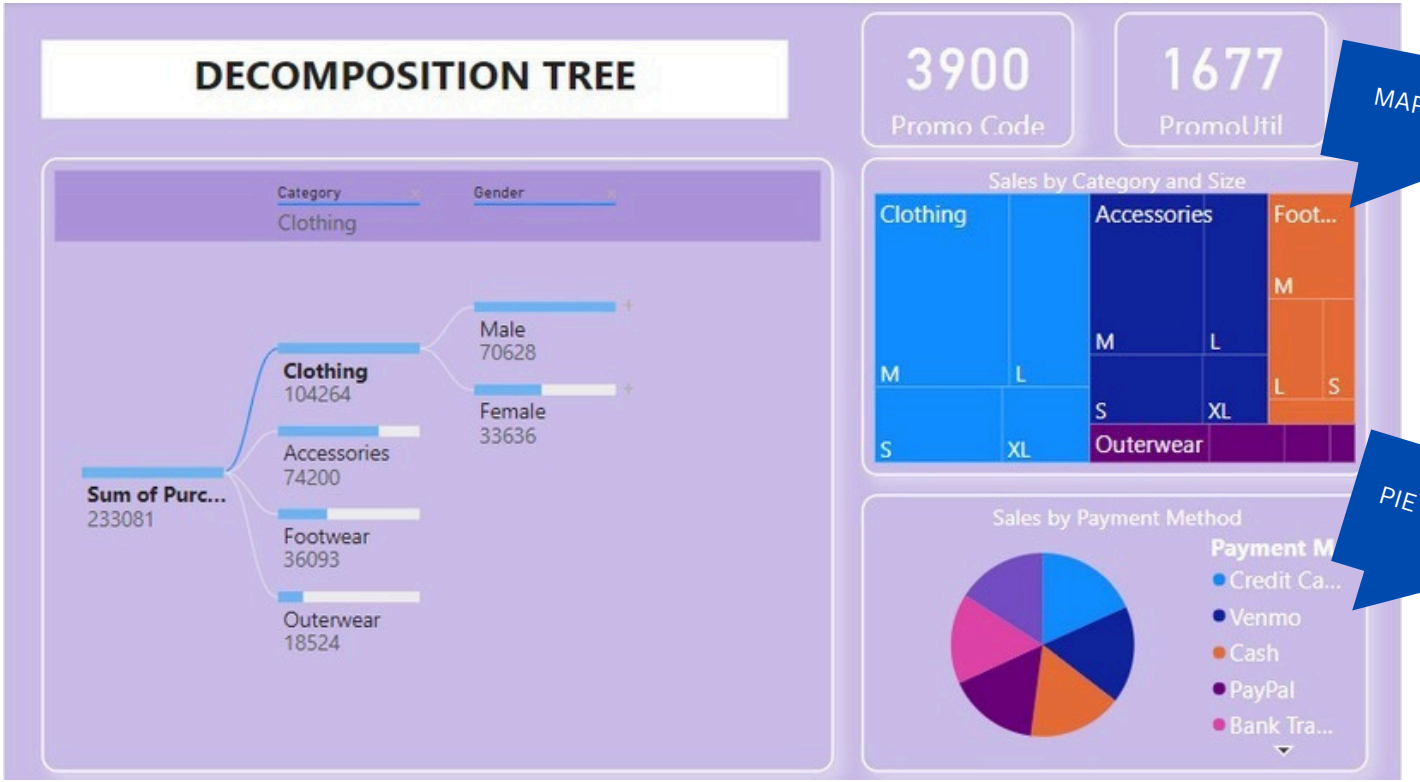
Ask a question about your data

Try one of these to get started

top locations by total sub

RIBBON CHART

LINE CHART


DECOMPOSITION TREE
3900
Promo Code
1677
PromoUtil
MAP TREE
Category
Clothing
Gender
Sum of Purc...
233081
Clothing
104264
Accessories
74200
Footwear
36093
Outerwear
18524
Male
70628
Female
33636
Sales by Category and Size
Clothing
Accessories
Foot...
M
L
S
XL
M
L
S
XL
M
L
S
Outerwear
PIE CHART
Sales by Payment Method
Payment M
Credit Ca...
Venmo
Cash
PayPal
Bank Tra...