**ANANYA DOSHI – 60004200015, A1**

# DATA MINING AND WAREHOUSE

## PRACTICAL 2

Yashvi Satam (60004200014)

Ananya Doshi (60004200015)

Syed Mohammed Raza Rizvi (60004200019)

Batch - A/A1

Computer Engineering

## Aim

Build Data Warehouse/Data Mart for a given problem state

## Theory

Data Warehouse: A data warehouse is a large collection of business data used to help an organization make decisions. The concept of the data warehouse has existed since the 1980s, when it was developedto help transition data from merely powering operations to fueling decision support systems that revealbusiness intelligence.A Data Warehousing (DW) is a process for collecting and managing data from varied sources to providemeaningful business insights. A data warehouse is typically used to connect and analyze business data from heterogeneous sources. The data warehouse is the core of the BI system which is built for data analysis and reporting.It is a blend of technologies and components which aids the strategic use of data. It is electronic storageof a large amount of information by a business which is designed for query and analysis instead of transaction processing. It is a process of transforming data into information and making it available to users in a timely manner to make a difference.

Star Schema: Each dimension is represented with only one-dimensional table. This table contains a setof attributes. There is a fact table at the center, It contains keys to other dimensions.
Snowflake Schema: Some dimension tables normalized. The normalization splits up the data into additional tables. Due to normalization, the snowflake schema, the redundancy is reduced and therefore, it becomes easy to maintain and save storage space.

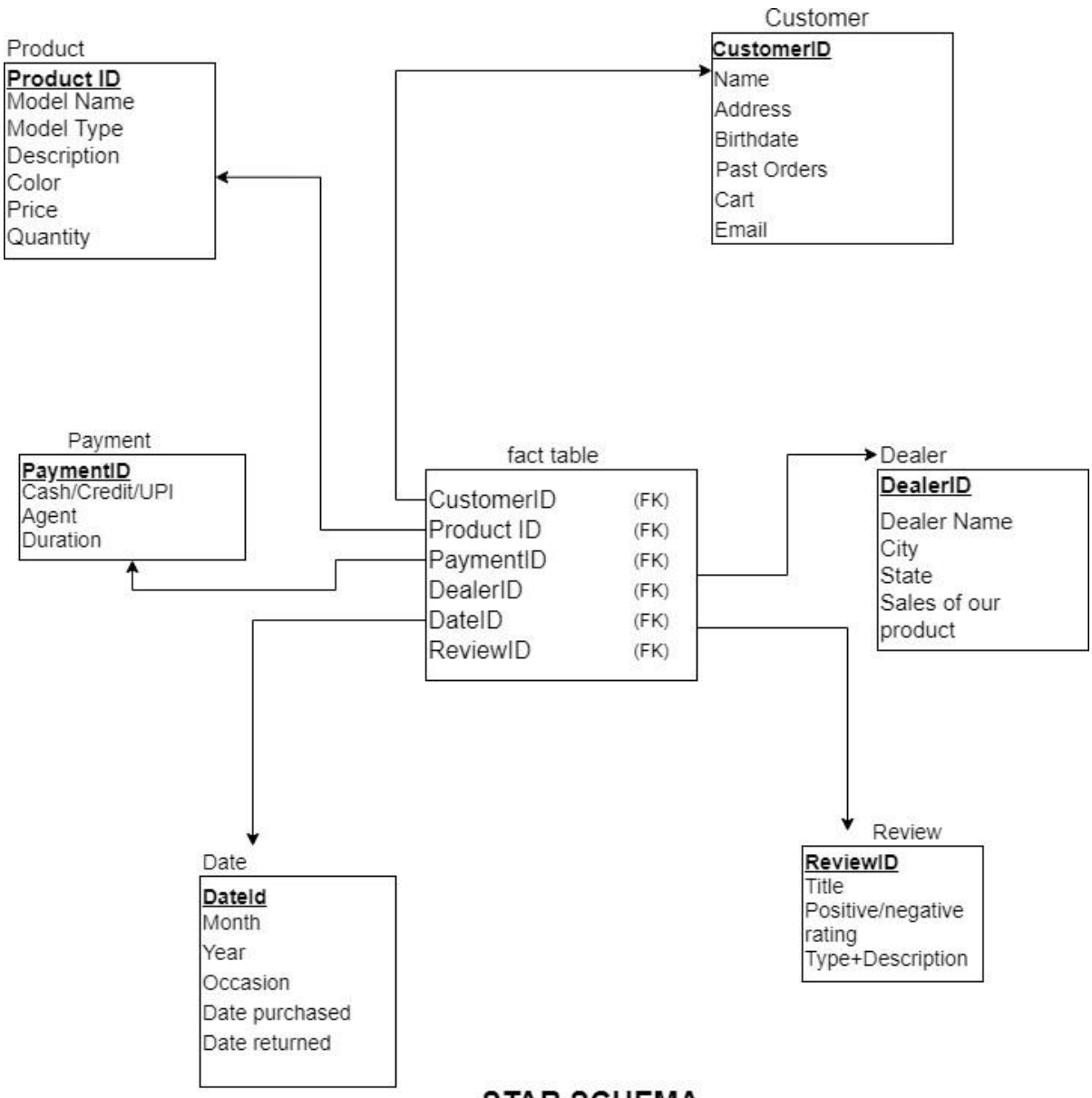
**Information Package:** An Information Package is a conceptual container of two types of informationcalled Content Information and Preservation Description Information (PDI). The Content Information and PDI are viewed as being encapsulated and identifiable by the Packaging Information. The resultingpackage is viewed as being discoverable by virtue of the Descriptive Information

# Dimensions

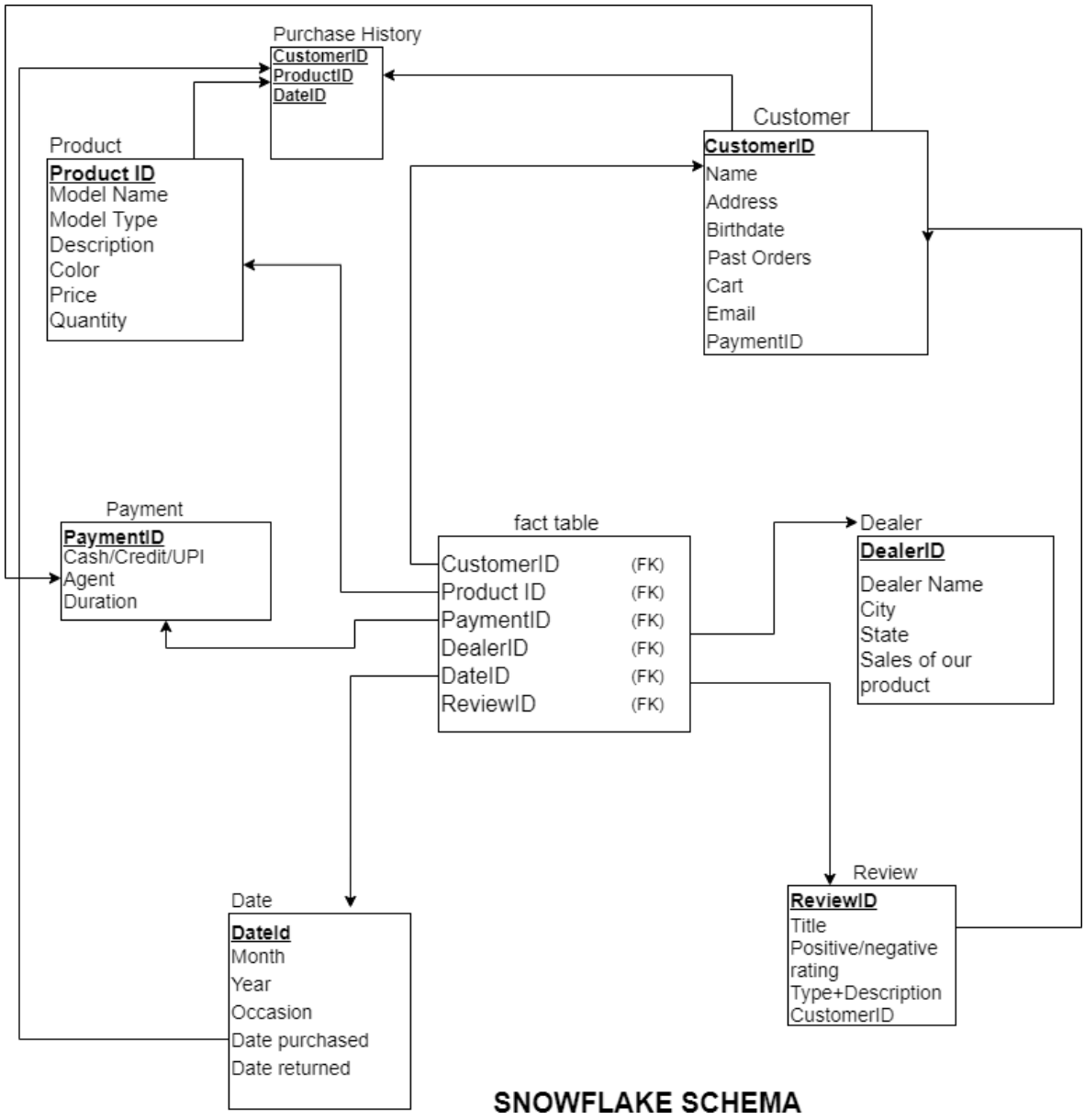
| | Time | Product | Customer Info | Dealer | Payment Method | Review |
|---|---|---|---|---|---|---|
| **Categories** | Product launch date | Model name | Name | Dealer name | Cash, Credit, EMI or UPI | Title |
| | Year | Model type | Address | City | Agent | Positive<br>negative rating |
| | Month | Description | Birthdate | State | Duration | Complaints type+ Description |
| | Quarter | Color | Past complications if any | Sales of our product | | |
| | Day of week | Price | Cart | | | |
| | Day of month<br>Occasion flag<br>Product return flag | Quantity | Email | | | |

## Questions:

**Q.1)** What is the percent increase in sales during IPO Announcement/Dividend Declaration
**Q.2)** Compare the Percent Increase in Customers during advertising Campaign and subsequently calculate the ROI (Return on Investment ) with respect to advertising Investment
**Q.3)** Which schemes influenced maximum sales?
**Q.4)** What is the quarter-wise sales distribution of every product?
**Q.5)** Which payment method is common among the customers and for which product?
**Q.6)** Which product and sold by which agent has the most negative reviews/complaints?
**Q.7)** What product has the most customers opting for EMI?
**Q.8)** How many existing India Bulls customers have consistently renewed their membership year over year
**Q.9)** In which locations do we have customers greater than a specific number where wedon't have dealers?
**Q.10)** What age groups prefer which type of membership
**Q.11)** How many Dhani App Members have enrolled for finance lending scheme?


Product
Product ID
Model Name
Model Type
Description
Color
Price
Quantity
Customer
CustomerID
Name
Address
Birthdate
Past Orders
Cart
Email
Payment
PaymentID
Cash/Credit/UPI
Agent
Duration
fact table
CustomerID (FK)
Product ID (FK)
PaymentID (FK)
DealerID (FK)
DateID (FK)
ReviewID (FK)
Dealer
DealerID
Dealer Name
City
State
Sales of our product
Date
DateId
Month
Year
Occasion
Date purchased
Date returned
Review
ReviewID
Title
Positive/negative rating
Type+Description

**STAR SCHEMA**




## Conclusion:

In this experiment we explored and built a data mart,schema and information package for India bulls policy packages.