PROJECT REPORT

ON

SALES DATA ANALYSIS

BY

SANTHOSH S

1. **Introduction**

**Background**

In today's fast-paced and competitive business environment, having a deep understanding of sales performance is crucial. Businesses generate vast amounts of sales data daily from various channels such as physical stores, online platforms, and third-party vendors. This data, if analyzed effectively, can provide valuable insights into customer preferences, sales trends, and market dynamics. Leveraging these insights can help businesses optimize their strategies, enhance customer satisfaction, and drive revenue growth.

Sales data analysis involves examining historical sales data to uncover patterns, trends, and anomalies. By applying analytical techniques, businesses can identify their best-selling products, peak sales periods, and the impact of different marketing campaigns. This analysis is instrumental in making informed decisions about inventory management, pricing strategies, and promotional activities.

**Aim**

The primary aim of this project is to conduct a comprehensive analysis of sales data to identify significant patterns and trends. Specifically, the project seeks to:

* Calculate key revenue metrics such as total sales.
* Analyze sales trends over different time periods.
* Identify best-selling products.
* Create visualizations to effectively communicate the findings.

These objectives are geared towards providing actionable insights that can inform business decisions and strategies aimed at increasing sales and profitability.

**Purpose of the System**

The purpose of this sales data analysis system is to enable businesses to transform raw sales data into meaningful insights. The system is designed to:

* **Collect and Store Sales Data**: Integrate and organize sales data from various sources into a central database.
* **Analyze Sales Performance**: Use statistical and computational techniques to evaluate sales data.
* **Visualize Data**: Create intuitive and interactive visualizations that make it easier to understand complex data sets.
* **Support Decision-Making**: Provide a solid foundation of data-driven insights that help businesses make strategic decisions regarding inventory, marketing, and sales tactics.

**2. Analysis**

**Hardware and Software Requirements**

* **Hardware System Configuration**
  + processor - Pentium – IV
  + RAM - 4 GB (min)
  + Hard Disk - 20 GB
* **Software System Configuration**
  + Operating System : Windows 7 or 8
  + Software : MySQL Workbench

**Feasibility Study**

**Technical Feasibility**

The project requires knowledge of MySQL for database management and PowerBI for visualization. Both technologies are widely used and have ample documentation and community support.

**Operational Feasibility**

Businesses can easily integrate this system into their existing workflows. The use of common software tools like MySQL and PowerBI ensures compatibility with most business environments.

**Economic Feasibility**

The cost involved includes the subscription for PowerBI (if using the Pro version) and potential costs for MySQL hosting if cloud services are used. These costs are minimal compared to the potential benefits gained from data-driven decision-making.

**3. Design**

**Block Diagram**

**The block diagram below illustrates the flow of the data cleansing process:**

+------------+ +-------------+ +-------------+

| Sales Data | -----> | MySQL | -----> | PowerBI |

+------------+ +-------------+ +-------------+

1. **Sales Data**: Raw sales data collected from various sources.
2. **MySQL**: Data storage and querying.
3. **PowerBI**: Data visualization and reporting.

**4. Methodology**

**Methodology**

The following steps outline the methodology used for the sales data analysis process:

**4.1 Data Collection and Import**

1. **Data Import**: Load the raw sales dataset into a MySQL database.
   1. **Data Storage**
2. **Store Cleaned Data**: Store the prepared data in a structured format within the MySQL database. **Calculate Key Metrics**: Use SQL queries to calculate key revenue metrics:
3. **Total Sales**: Calculate the sum of the sales amount to determine the overall revenue generated.
4. **Sales Trends**: Group data by time periods (e.g., daily, monthly, quarterly) to observe trends and patterns over time.

**5. Best-Selling Products**: Identify top products by sales volume to understand which items contribute most to revenue.

* 1. **Data Visualization**

1. **Import Data into PowerBI**: Connect PowerBI to the MySQL database and import the cleaned and prepared sales data.
2. **Create Visualizations**: Design various charts and graphs to represent total sales, sales trends, and best-selling products. Use visual tools such as bar charts, line graphs, and pie charts to make the data easily interpretable.
3. **CODING AND IMPLEMENTATION**

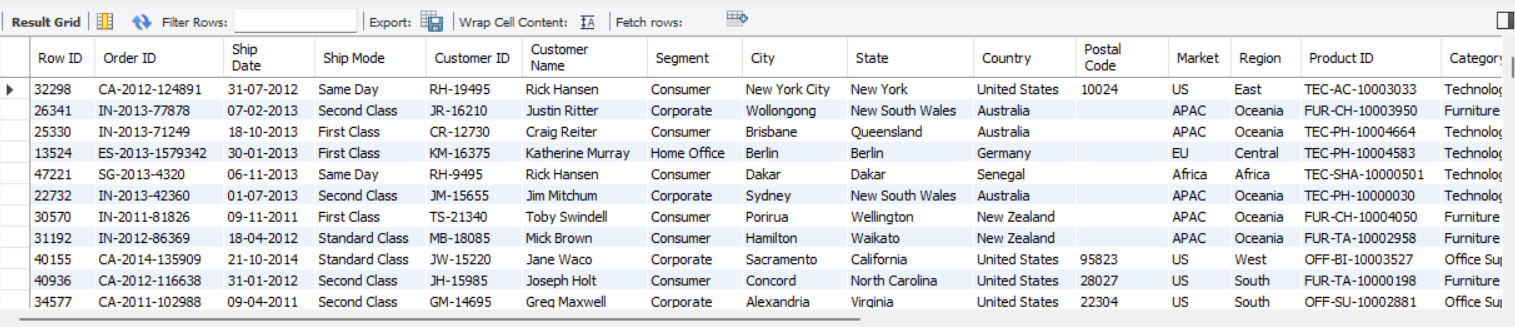
**1. Create Database**

create database sales;

use sales;

**2. Import .CSV file**

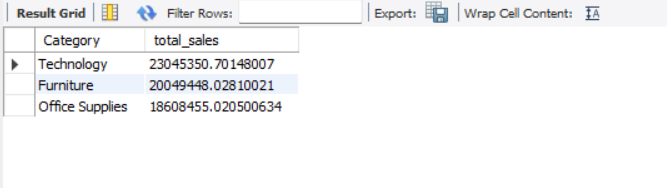
select \* from orders;



**3. SQL queries to calculate Total Sales:**

select Category,

sum(Quantity\*Sales) as total\_sales from orders group by Category;



**4. Group data by time periods**

select

year(order\_date) as year,

sum(case when month(order\_date)=1 then Sales else 0 end) as jan\_sales,

sum(case when month(order\_date)=2 then Sales else 0 end) as feb\_sales,

sum(case when month(order\_date)=3 then Sales else 0 end) as mar\_sales,

sum(case when month(order\_date)=4 then sales else 0 end) as apr\_sales,

sum(case when month(order\_date)=5 then sales else 0 end) as may\_sales,

sum(case when month(order\_date)=6 then sales else 0 end) as jun\_sales,

sum(case when month(order\_date)=7 then sales else 0 end) as jul\_sales,

sum(case when month(order\_date)=8 then sales else 0 end) as aug\_sales,

sum(case when month(order\_date)=8 then sales else 0 end) as sep\_sales,

sum(case when month(order\_date)=10 then sales else 0 end) as oct\_sales,

sum(case when month(order\_date)=11 then sales else 0 end) as nov\_sales,

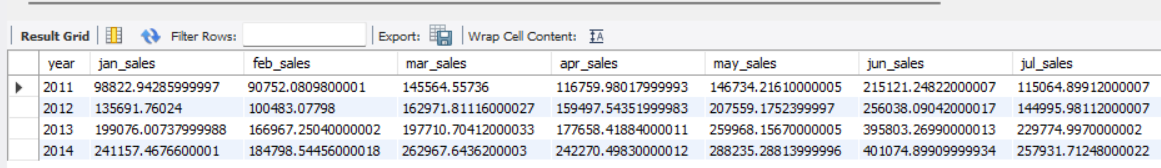
sum(case when month(order\_date)=12 then Sales else 0 end) as dec\_sales,

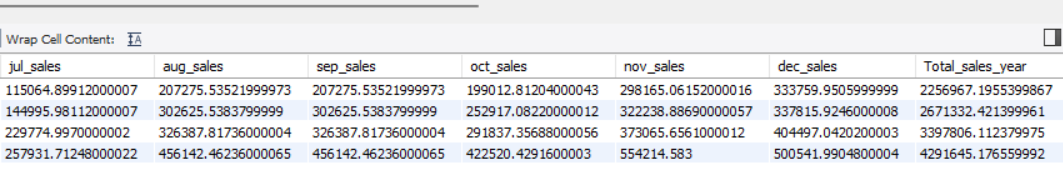
sum(sales) as Total\_sales\_year from orders

where year(order\_date) in (2011,2012,2013,2014)

group by year(order\_date)

order by year(order\_date);

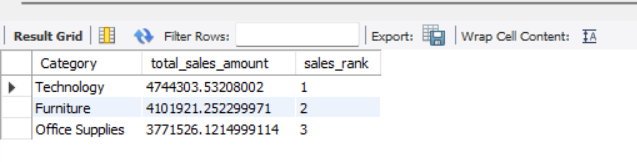




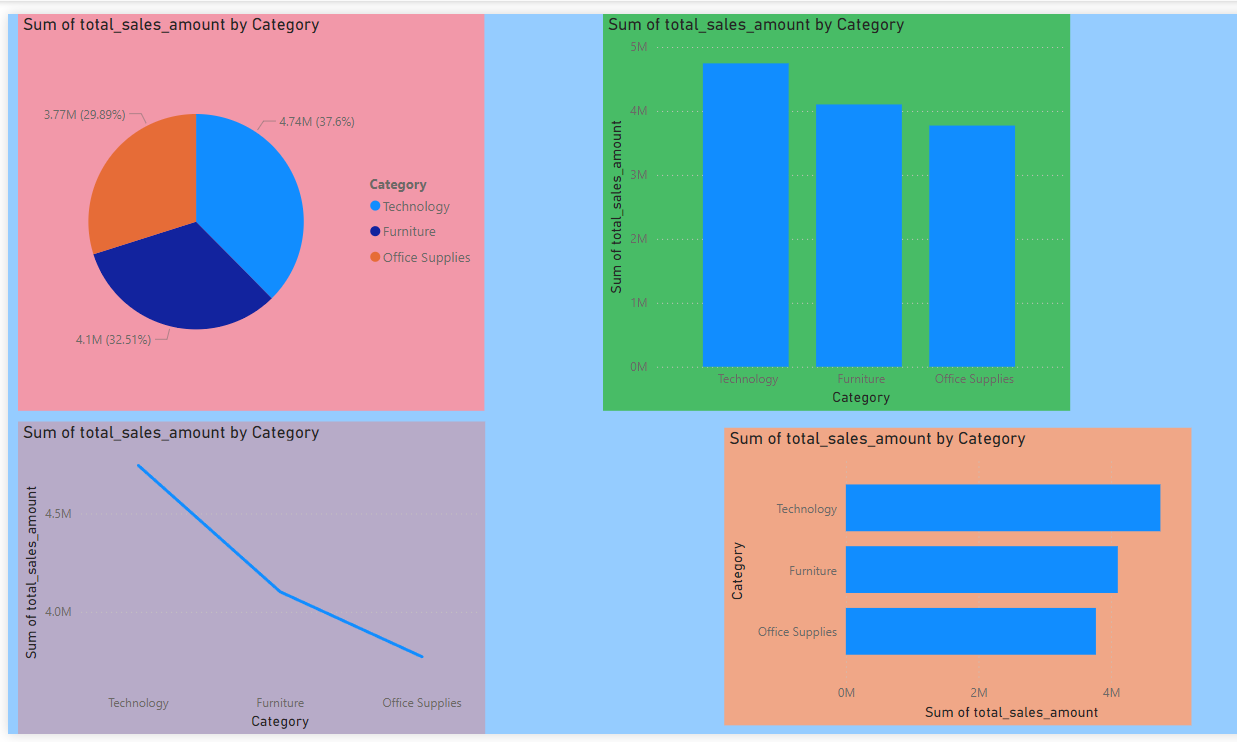
5. Identify **Best-Selling Products**

select Category, total\_sales\_amount, RANK() over(order by total\_sales\_amount DESC) AS sales\_rank

from ( select category, sum(sales) as total\_sales\_amount from orders group by Category) as sales\_summary order by total\_sales\_amount DESC ;

****

**6. Visualisation using Power BI**

****

**5. Discussion**

**Results and Observations**

* **Total Sales**: The system provides a clear calculation of total sales.
* **Sales Trends**: Trends over time help in identifying peak sales periods and seasonal effects.
* **Best-Selling Products**: Highlighting top products aids in inventory and marketing strategy planning.

**Challenges and Solutions**

* **Data Quality**: Ensuring data accuracy was challenging. Implemented thorough data cleaning processes.
* **Integration**: Combining MySQL and PowerBI required ensuring compatible data formats and structures. Used consistent data schemas and data transformation tools.

**Future Enhancements**

* **Predictive Analytics**: Incorporate machine learning models to forecast future sales.
* **Real-Time Data Analysis**: Implement real-time data processing for up-to-the-minute insights.
* **Enhanced Visualizations**: Utilize advanced PowerBI features for more interactive and detailed visualizations.