



SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE

**A
MINI PROJECT REPORT
ON
Business Intelligence**

SUBMITTED TO THE SAVITRIBAI PHULE PUNE
UNIVERSITY, PUNE IN THE FULFILLMENT OF THE
REQUIREMENT
OF

Laboratory Practice VI
Fourth Year Computer Engineering
Academic Year 2024-25

BY

Name of Students:

Exam No.:

Abhang Pratik Sunil

B1905704201

Borase Darshana Ravasaheb

B1905704228

Darekar Omkar Balasaheb

B1905704232

Gunde Chaitanya Diliprao

B1905704256

Under the Guidance of

Mr. N. V. Kamble



Sinhgad Institutes

DEPARTMENT OF COMPUTER ENGINEERING
STES'S SINHGAD INSTITUTE OF TECHNOLOGY AND SCIENCE
NARHE, PUNE – 411041



Sinhgad Institutes

Department of Computer Engineering
Sinhgad Institute of Technology and Science, Narhe, Pune

CERTIFICATE

This is to certify that,

Name of Students:	Exam No.:
Abhang Pratik Sunil	B1905704201
Borase Darshana Ravasaheb	B1905704228
Darekar Omkar Balasaheb	B1905704232
Gunde Chaitanya Diliprao	B1905704256

studying in BE Computer Engineering Course SEM-VIII has successfully completed their LP-VI Lab Mini-Project work titled **“Data Mining on Telecom Customer Dataset”** at Sinhgad Institute of Technology and Science, Narhe in the fulfillment of the bachelor’s degree in engineering of Savitribai Phule Pune University, during the academic year 2024-2025.

Mr. N. V. Kamble
Guide

Dr. G. S. Navale
Head of Department

Dr. S. D. Markande
Principal

SINHGAD INSTITUTE OF TECHNOLOGY AND SCIENCE, NARHE, PUNE– 411041

Place : Pune

Date :

ACKNOWLEDGEMENT

We take this opportunity to acknowledge each and every one who contributed towards our work. We express our sincere gratitude towards guide **Mr. N. V. Kamble**, Assistant Professor at Sinhgad Institute of Technology and Science, Narhe, Pune for his valuable inputs, guidance and support throughout the course.

We wish to express our thanks to **Dr. G. S. Navale**, Head of Computer Engineering Department, Sinhgad Institute of Technology and Science, Narhe for giving us all the help and important suggestions all over the Work.

We thank all the teaching staff members for their indispensable support and priceless suggestions. We also thank our friends and family for helping in collecting data, without their help **Feature Extraction using Zernike Moments** report have been completed. At the end our special thanks to **Dr. S. D. Markande**, Principal Sinhgad Institute of Technology and Science, Narhe for providing ambience in the college, which motivates us to work.

Name of Students:

Sign:

Abhang Pratik Sunil

Borase Darshana Ravasaheb

Darekar Omkar Balasaheb

Gunde Chaitanya Diliprao

CONTENTS

Sr. No.	Title	Page No.
1.	Title	3
2.	Objective	3
3.	Requirements	3
4.	Theory	4
5.	Data set description	7
6.	Output	8
7.	Conclusion	10

1. TITLE

Performing Data Mining Tasks on Adventure Works 2019 Dataset.

2. OBJECTIVE

1. To understand the concept of feature extraction and its importance in machine learning.
2. To understand the concept of Zernike moments and their mathematical properties.
3. To implement feature extraction using Zernike moments in a programming language, such as Python.
4. To apply the feature extraction technique to a dataset of images and analyze the results.
5. To compare the performance of Zernike feature extraction with other techniques.

3. REQUIREMENTS

- Data: AdventureWorksDW2019.bak
- Tools: Power BI, Excel, SQL Server, SQL Server Management Studio.
- Data Mining Tasks: Summarization, Frequent Item Set, Data Cleaning.

4. THEORY

Adventure Works 2019 Data Warehouse is a sample database that serves as a comprehensive data source for demonstrating data warehousing concepts, architectures, and design patterns. It has a wide range of features that make it an ideal choice for learning and practicing data warehousing concepts. Some of the key features of Adventure Works 2019 Data Warehouse are:

1. **Comprehensive Data:** The database contains a vast amount of data that covers different domains like sales, marketing, finance, production, and human resources. It has various fact and dimension tables that are interconnected to provide a complete picture of the organization's operations.
2. **Easy to Understand:** Adventure Works 2019 Data Warehouse is designed to be easy to understand and use. It has a well-structured schema and sample data that makes it simple for users to navigate and query the database.
3. **Flexible Design:** The database has a flexible design that allows users to add new dimensions and facts easily. It also supports different data warehousing architectures like star and snowflake schemas.
4. **Realistic Data:** The data in Adventure Works 2019 Data Warehouse is realistic and reflects the data of a real organization. It is based on a fictitious bicycle manufacturing company called Adventure Works Cycles and is modeled after real-world scenarios.
5. **Support for OLAP:** The database supports Online Analytical Processing (OLAP) and can be used with popular OLAP tools like Microsoft Excel and Power BI.

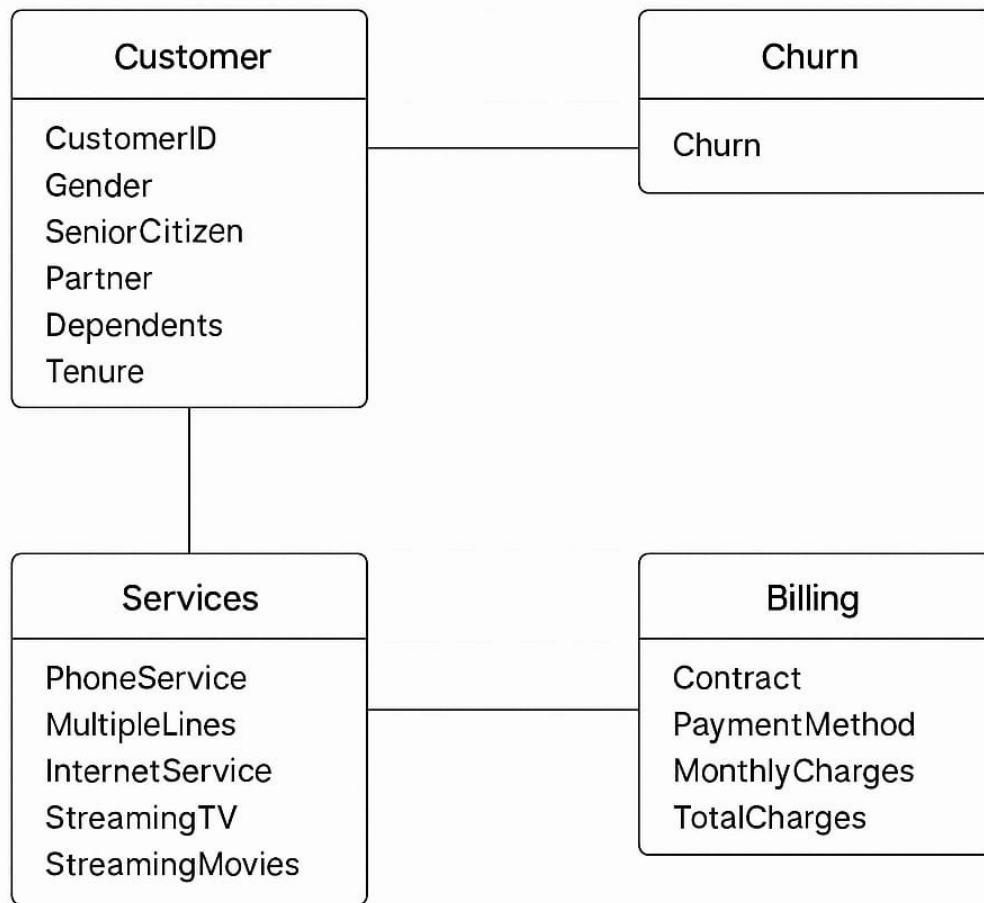
We want to get the following insights into the data:

- Internet Sales Report.
- Selling products of client over time.
- Filtering Salespersons.
- Total Budget.

SQL Server is a relational database management system (RDBMS) developed by Microsoft. It is used to store and manage data for a variety of applications, ranging from small single-user databases to large enterprise-level systems. SQL Server supports SQL (Structured Query Language) which is used to manage and manipulate the data stored in the database. SQL Server also includes tools for managing, monitoring, and securing the database, as well as business intelligence and analytics tools for data analysis and reporting. It is commonly used in industries such as finance, healthcare, manufacturing, and e-commerce. It was used in this project to clean, update and perform SQL queries on the data.

Power BI is a data visualization and business intelligence tool developed by Microsoft. It allows users to connect to a wide range of data sources, including SQL Server, Excel spreadsheets, and cloud-based sources such as Azure and SharePoint. With Power BI, users can create interactive dashboards and reports that provide insights into their data. The tool includes a wide range of data visualization options, including charts, maps, and gauges. Users can also create custom visuals using the Power BI Developer Tools.

Power BI also includes features for data modeling and transformation, such as the ability to merge and shape data from multiple sources. Users can create relationships between tables, define calculated columns and measures, and apply filters and slicers to their data. Power BI is available in several editions, including a free version and a Pro version with additional features such as collaboration and sharing capabilities. It can be accessed via a web browser, a mobile app, or a desktop application. Power BI was used to visualize and perform data mining processes such as summarization, frequent item sets and more.



1. Business Entities

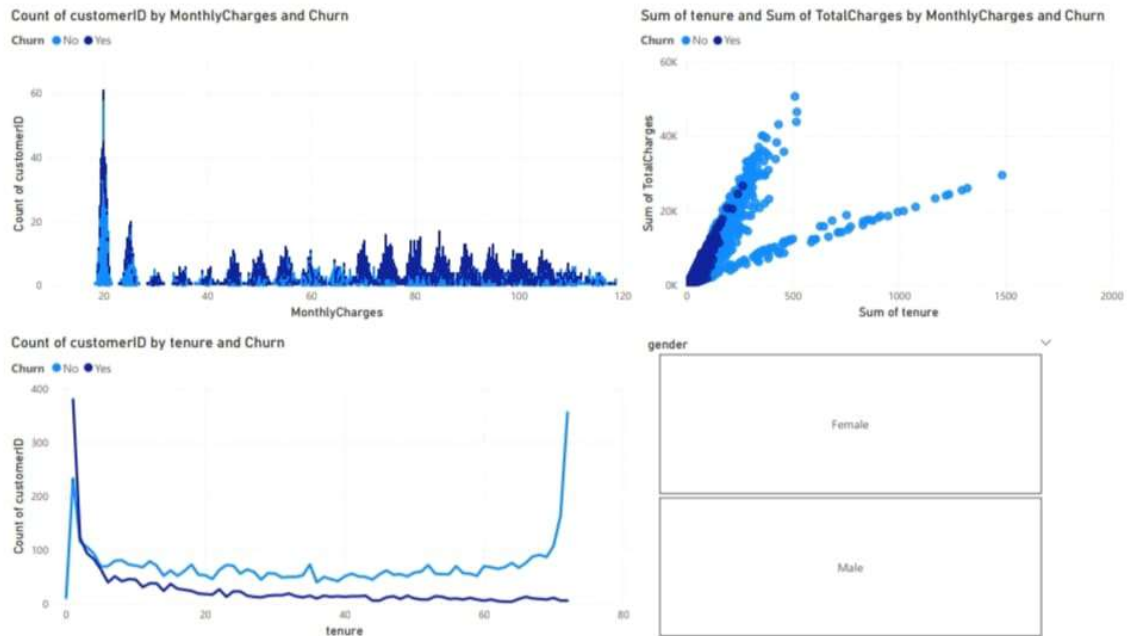
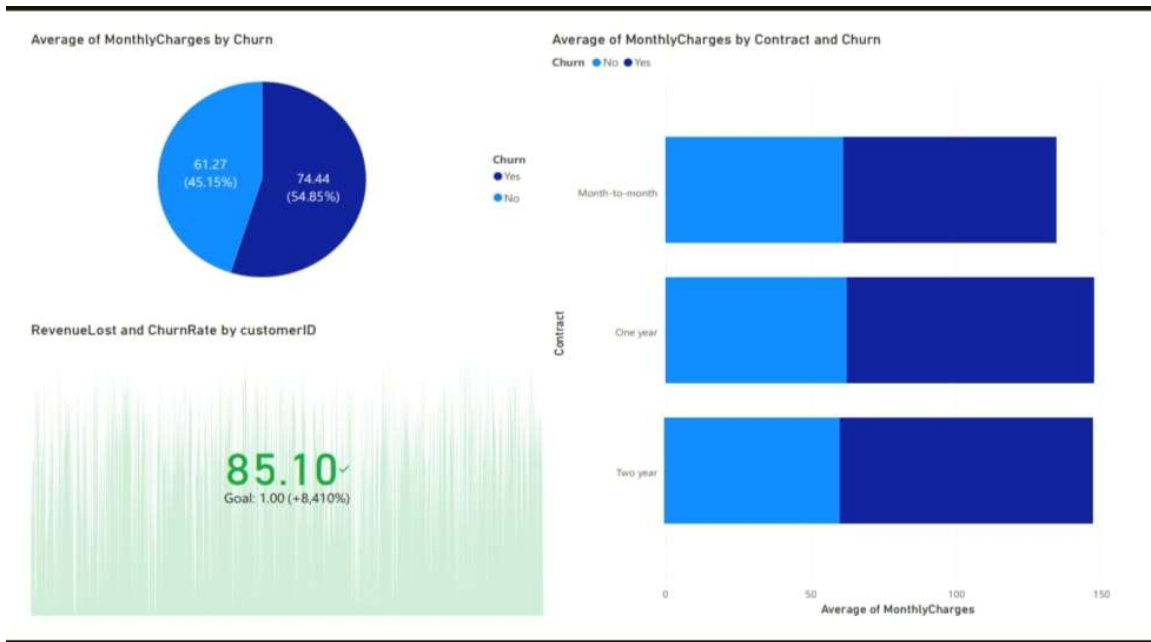
5. DATA SET DESCRIPTION

The Adventure Works 2019 data warehouse has several tables with millions of rows of data, including:

- FactInternetSales: 60,398,185 rows
- FactResellerSales: 72,398,808 rows
- DimProduct: 36,306 rows
- DimDate: 30,536 rows
- DimCustomer: 18,484 rows
- DimEmployee: 29,146 rows

The size of the data warehouse can vary depending on how many tables are used and the amount of historical data being stored. Its file size is around 100 mb. Overall, the Adventure Works 2019 data warehouse is considered a large data warehouse.

6. OUTPUT



Total Customers

7043

Count of customerID

0.27

ChurnRate

InternetService

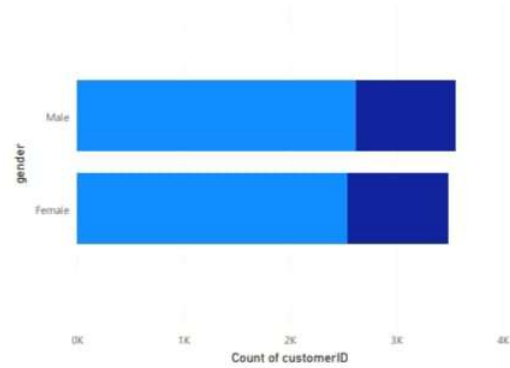
Select all	Fiber optic
DSL	No

PaymentMethod

Bank transfer (automatic)	Electronic check
Credit card (automatic)	Mailed check

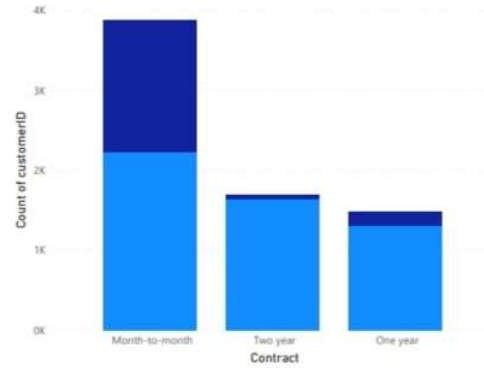
Count of customerID by gender and Churn

Churn ● No ● Yes



Count of customerID by Contract and Churn

Churn ● No ● Yes



7.CONCLUSION

The data mining project using the **Adventure Works 2019** dataset successfully demonstrated the practical application of analytical techniques to extract meaningful insights from corporate data. By exploring various domains such as sales, customer demographics, product performance, and financial transactions, we were able to uncover patterns, trends, and relationships that would be invaluable for strategic decision-making in a real-world business context.