Churn Analytics

High Level Document

Rashika Singhal

DOCUMENT VERSION CONTROL

Date Issued	Version	Description	Author
21st September 2024	1.0	First Version of Complete HLD	Rashika Singhal

ABSTRACT

Customer Churn refers to the loss of customers or subscribers for any reason at all. Businesses measure and track churn as a percentage of lost customers compared to total number of customers over a given time period.

Customer churn rate is an important metric to understand because lost customers mean lost revenue—and the higher the churn rate, the more your bottom line will suffer.

Churn happens in almost all business, and almost all of them want to prevent it. Many brands are fix on decreasing churn since it hinders business growth.

Many companies place more emphasis on acquiring new clients than keeping their current customers. And they are unaware of the cost of customer turnover until it has significantly eroded their earnings.

1. Introduction

. Customer churn refers to the rate at which customers stop using a company's services. This project focuses on analyzing and predicting customer churn using customer data. The cleaning and transformation process uses NumPy**,** while Power BI is employed to visualize insights and key metrics, enabling better business decisions.

1.1 Why this High-Level Design Document?

This document outlines the overall architecture and key components involved in the churn analysis project. It describes how the data is processed using NumPy and how it is visualized in Power BI, providing a roadmap for implementation.

1.2 Scope

The scope of this document includes:

- Data ingestion and preprocessing using NumPy.
- Data export and integration with Power BI.
- Visualization of churn-related insights for actionable recommendations.

2. General Description

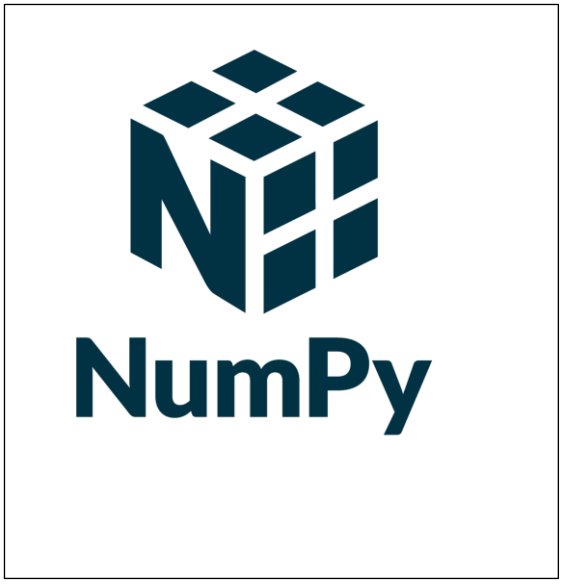
2.1 Problem Statement

. The business faces a high rate of customer churn, leading to a negative impact on revenue. This project aims to analyze the available customer data to understand the causes of churn and predict potential churners, allowing the business to take preemptive actions.

2.2 Tools Used

NumPy: For data cleaning and transformation (e.g., handling missing values, performing statistical analysis).

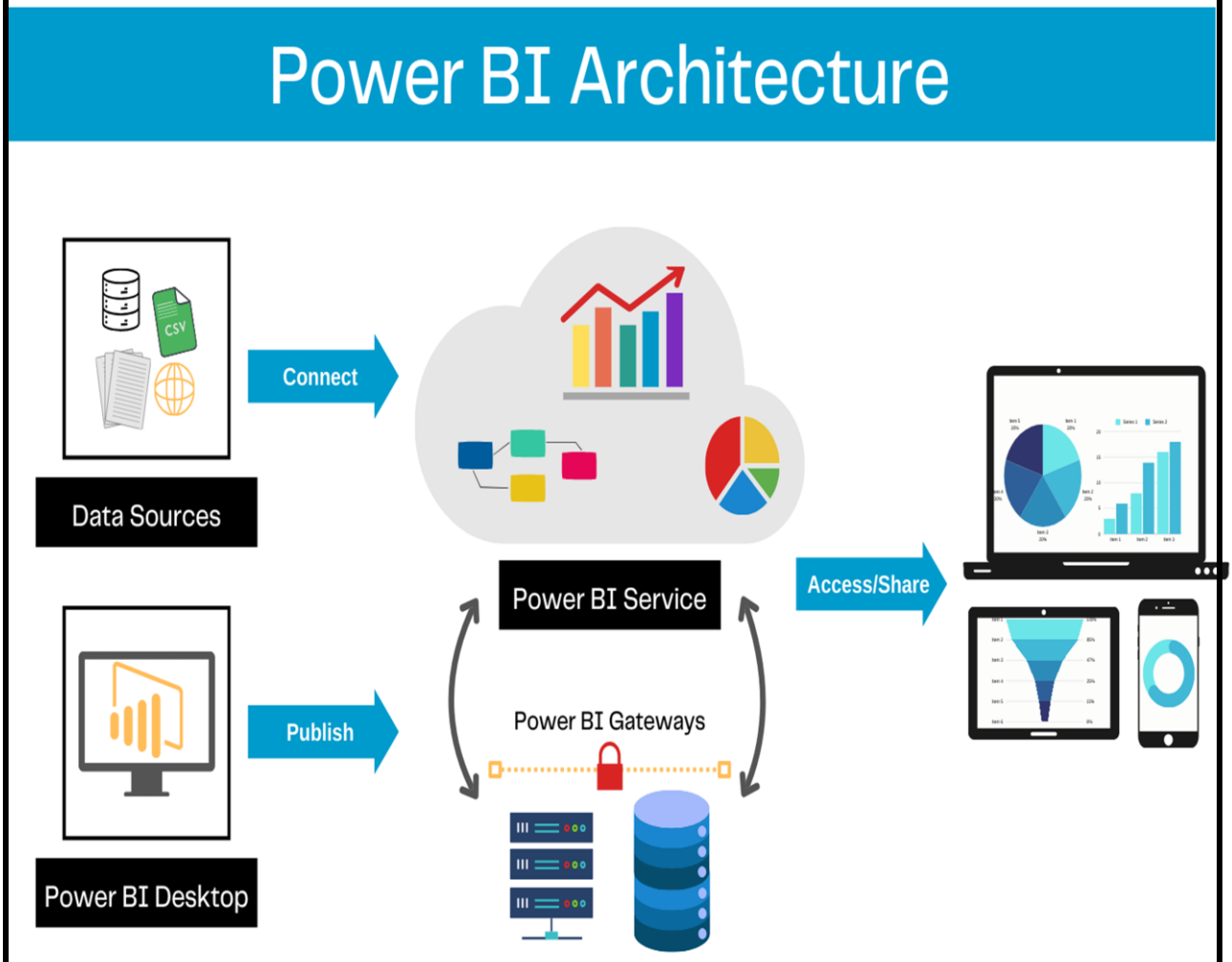
Power BI: For creating dynamic dashboards and reports to visualize churn patterns and trends.

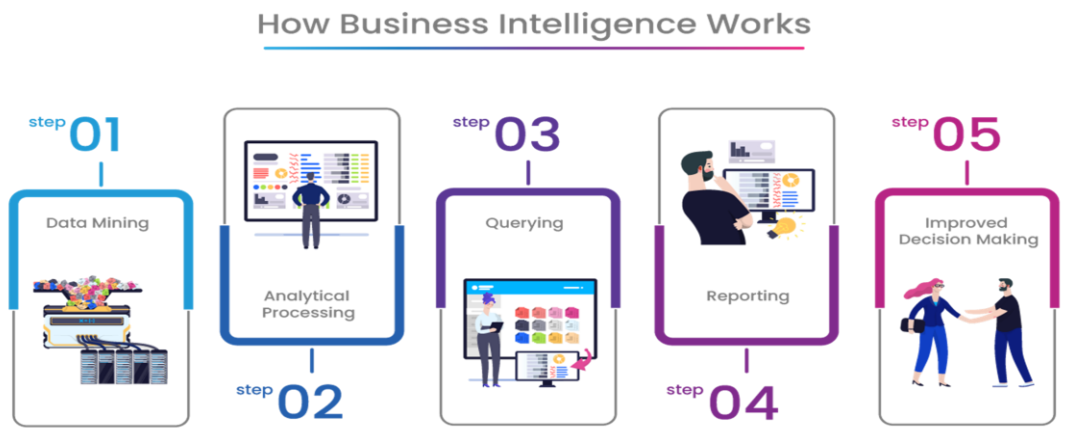


3. Design Details

3.1 Functional Architecture

1. **Data Ingestion**:
   - Load the customer data into Python using NumPy arrays.
   - The data may come from CSV files, databases, or other sources.





2. **Data Cleaning & Transformation** (NumPy):
   - Handle missing values using NumPy functions like np.nan_to_num().
   - Perform data normalization or scaling using statistical functions in NumPy.
   - Implement feature engineering to calculate new attributes, such as customer tenure and churn status.
3. **Data Export**:
   - After preprocessing, save the cleaned data to CSV format, making it ready for import into Power BI.
   - Create charts (bar charts, line graphs, pie charts) to track churn trends across different dimensions (e.g., region, service usage).
4. **Visualization in Power BI**:
   - Import the cleaned CSV into Power BI.
   - Build interactive dashboards to visualize key metrics like churn rate, retention rate, and customer segments.

3.2 Optimization

3.2.1 NumPy Optimization

- **Efficient Array Operations**: Use vectorized operations for fast, memory-efficient data cleaning and transformations.
- **Memory Management**: Optimize memory usage by leveraging NumPy's compact array structures.

3.2.2 Power BI Optimization

- **Data Model**: Design an optimized data model in Power BI for smooth interaction with large datasets.
- **Visual Efficiency**: Use optimized visuals and limit unnecessary calculations in Power BI reports to improve load times.

4. KPIs

- **Churn Rate**: Percentage of customers who have churned in a given period.
- **Retention Rate**: Percentage of customers retained.
- **Customer Segmentation**: Insights into customer behavior, segmented by factors such as region or product usage.
- **Churn Prediction Accuracy**: Metrics showing the accuracy of the churn prediction model.

5. Deployment

- The final solution will be a set of Power BI dashboards that provide real-time monitoring of churn metrics. These dashboards will allow stakeholders to identify trends and take proactive steps to reduce churn.