DEPARTMENT OF COMPUTER SCIENCE, IT AND ANIMATION
DEOGIRI COLLEGE CHATRAPTI SAMBHAJINAGAR 431005.

-

Project Report

ON

"INDIA MORTALITY ANALYSIS AND LIFE EXPECTANCY"

Submitted by

Abhishek Kailas Gore (Roll No:43)

Vaibhav Dadasaheb Jadhav (Roll No:46)

Pradip Suresh Kakde (Roll No: 54)

Suraj Santosh Kadam (Roll No:53)

BCA (Sci.) III Year IV Semester

Academic Year 2024-2025

Guided by

Ms. Rajashri G. Kanke

Submitted To

Department of Computer Science, IT and Animation

Deogiri College,Chatrapati Sambhajinagar 431005

DEPARTMENT OF COMPUTER SCIENCE, IT AND ANIMATION
DEOGIRI COLLEGE CHATRAPATI SAMBHAJINAGAR 431005.

CERTIFICATE

It is the certified work of candidates of ***BCA(Sci.) III Year V/Sem*** *who have satisfactorily completed the project entitle* ***"India Mortality Analysis And Life Expectancy"*** *for the fulfillment of the Bachelor Degree from Dr. Babasaheb Ambedkar Marathwada University, Chatrapti Samabhajinagar for the academic year 2024-25.*

Submitted By

Pradip Suresh Kakde.

Suraj Santosh Kadam

Vaibhav Dadasaheb Jadhav

Abhishek Kailas Gore

In-Charge	Project Guide	HOD
Dr. Rajashri A. Joshi	Ms. Rajashri G. Kanke	Dr. Ritesh A. Magare

External Examiner

ACKNOWLEDGEMENT

Many people have contributed in some or the other forms to this project, although our name appears on the cover of this project. We could not have done this Case Study without the assistance or support of each of the following. We thank all of you.

We wish to place on record our deep sense of gratitude towards our project guide **Ms. Rajashri G. Kanke** and Incharge **Dr. Rajashri A. Joshi** for the constant motivation and valuable help through the project. We also want to express our gratitude towards our Head of the department **Dr. Ritesh A. Magare,** for his invaluable guidance, insightful advice and continuous encouragement. We also extend our thanks to other faculty members for their co-operation.

We also thank those who have directly as well as indirectly supported us for the completion of this Project.

Pradip Suresh Kakde.

Suraj Santosh Kadam.

Abhishek Kailas Gore.

Vaibhav D. Jadhav.

INDEX

INTRODUCTION TO POWER BI DEKSTOP

Power BI Desktop is a powerful business intelligence (BI) tool developed by Microsoft. It is used for data visualization, data analysis, and creating interactive reports. Power BI Desktop enables users to connect to various data sources, transform raw data, and build insightful dashboards. It provides a user-friendly interface for analyzing data, making it a popular tool among data analysts and business professionals.

Power BI Desktop integrates with cloud services and other Microsoft applications, allowing users to share reports and dashboards across organizations. With its drag-and-drop functionality and DAX (Data Analysis Expressions) language, Power BI Desktop simplifies complex data analysis tasks.

Power BI Desktop offers robust features such as data modeling, custom calculations using DAX (Data Analysis Expressions), and Power Query for data transformation. It seamlessly integrates with Microsoft services like Excel, Azure, and SQL Server, as well as other cloud and on-premise databases. Users can create real-time, dynamic reports that enhance business insights and drive data-driven decisions.

Additionally, Power BI Desktop allows users to publish reports to the Power BI Service, enabling collaboration and sharing within organizations. It is widely used across industries for financial analysis, sales reporting, market trends, and operational insights.

Key Features of Power BI Desktop

1. Connect to Multiple Data Sources

- Power BI Desktop can pull data from a variety of sources, including:
- Excel spreadsheets
- Databases like SQL Server, MySQL, and Oracle
- Cloud services like Azure, Google BigQuery, and AWS
- Online applications like Salesforce and Google Analyticsage
- CSV, JSON, and XML files

This flexibility means you can combine data from different sources into a single report, making analysis more comprehensive.

2. Clean and Transform Data with Power Query

Raw data is often messy. Power Query, built into Power BI Desktop, helps clean and prepare the data by:

- Removing duplicates and errors
- Filtering and sorting data
- Changing data formats
- Creating new columns based on calculations

This step ensures that your analysis is accurate and reliable.

3. Build Data Models and Relationships

Once the data is ready, Power BI allows you to create relationships between tables—just like in a database. This helps in:

- Connecting different datasets (e.g., sales data and customer data)
- Performing complex calculations using DAX (Data Analysis Expressions)
- Organizing data into hierarchies for better insights

4. Create Interactive Dashboards and Reports

Power BI Desktop comes with a wide range of visualization tools, such as:

- Bar charts, pie charts, and line graphs
- Maps for geographic analysis
- KPI indicators to track key business metrics
- Custom visuals from the Power BI Marketplace

You can filter, drill down, and interact with reports, making it easier to explore trends and patterns.

5. Use AI and Advanced Analytics

Power BI Desktop integrates Artificial Intelligence (AI) and Machine Learning (ML) for deeper insights, including:

- Predicting future trends
- Sentiment analysis from text data
- Smart insights using Microsoft's cognitive services
- Integration with Python and R for advanced analytics

Why Use Power BI Desktop?

1. Easy to Use, No Coding Required

Power BI's drag-and-drop interface makes it accessible even if you're not a programmer. However, for advanced users, DAX allows powerful calculations.

2. Handles Large Datasets

Power BI can process and analyze millions of rows of data without slowing down, making it great for big data projects.

3. Cost-Effective

Power BI Desktop is free!

Power BI Pro and Premium offer advanced features for teams and enterprises.

4. Seamless Integration with Microsoft & Other Tools

Since it's a Microsoft product, it integrates perfectly with Excel, SQL Server, Azure, and SharePoint. It also connects to third-party apps like Google Analytics and Salesforce.

5. Real-Time Data Analysis

Power BI supports live dashboards and real-time data streaming, which is crucial for businesses that

need up-to-the-minute updates.

Who Uses Power BI Desktop?

Power BI Desktop is used across industries for various purposes:

- Healthcare – Tracking patient health and hospital performance
- Finance – Monitoring investments, expenses, and fraud detection
- Retail – Analyzing sales trends and customer behavior
- Manufacturing – Managing supply chains and production efficiency
- Marketing – Measuring campaign performance and audience engagement

Advantages of Power BI Desktop:

1. **Cost-Effective:** Power BI Desktop is free to use, making it accessible for individuals and small teams without the need for additional financial investment.
2. **User-Friendly Interface:** The application shares similarities with other Microsoft Office products, such as Excel, which can shorten the learning curve for new users.
3. **Advanced Data Modeling:** Users can perform complex data transformations, create calculated columns, and utilize Data Analysis Expressions (DAX) for sophisticated data modeling.
4. **Interactive Visualizations:** Power BI Desktop provides a variety of customizable charts, graphs, maps, and other visual tools to create engaging and interactive reports.
5. **Offline Capabilities:** The application allows users to work offline, enabling report creation and data analysis without an internet connection.
6. **Extensive Data Connectivity:** Power BI Desktop supports connections to multiple data sources, including Excel files, SQL databases, cloud services, and more, facilitating comprehensive data integration.
7. **Custom Scripting Support:** The application supports R and Python scripting, allowing for advanced data manipulation and the creation of custom visuals.

INTRODUCTION TO POWER BI SERVICE

Power BI Service is Microsoft's cloud-based platform designed for hosting, sharing, and collaborating on business intelligence (BI) reports and dashboards. It extends the capabilities of Power BI Desktop by providing a centralized, online environment where users can access, interact with, and share their BI content securely.

Key Features of Power BI Service

1. Cloud-Based Hosting

Power BI Service serves as the cloud-based hosting environment for Power BI reports, enabling users to publish and access their reports online. This cloud infrastructure is part of Microsoft's Azure and Office 365 offerings, ensuring scalability and reliability.

2. Workspaces for Collaboration

Workspaces in Power BI Service act as collaborative environments where teams can share, review, and edit Power BI content. Each workspace can be tailored to specific departments or projects, with permissions set at the workspace level to control access to reports and datasets.

3. Report Publishing and Editing

Users can publish reports created in Power BI Desktop directly to the Power BI Service. Once published, these reports can be viewed and edited online, facilitating collaboration and iterative development.

4. Dataflows for ETL Processes

Dataflows in Power BI Service provide cloud-based Extract, Transform, Load (ETL) capabilities using the Power Query engine. They allow for centralized data transformation and storage, enabling multiple Power BI reports to utilize shared tables and ensuring consistency across reports.

5. Shared Datasets

With Power BI Service, datasets can be published once and reused across multiple reports. This

promotes a single source of truth and reduces redundancy, as calculations and models defined in a shared dataset are consistently applied wherever the dataset is used.

6. Deployment Pipelines

Deployment pipelines in Power BI Service facilitate the management of content across development, testing, and production environments. They allow for controlled and streamlined deployment processes, ensuring that updates are thoroughly tested before reaching end-users.

7. Apps for Organized Content Delivery

Power BI Apps enable the packaging and distribution of related reports and dashboards to users. They provide a structured and customizable way to deliver BI content, allowing users to access a curated set of reports relevant to their roles or departments.

8. Data Lineage and Governance

Power BI Service offers data lineage features that trace the flow of data from source to report, enhancing transparency and data governance. This capability helps users understand data dependencies and the impact of changes, supporting better data management practices.

Advantages of Power BI Service:

1. Scalability: Power BI Service can handle varying data volumes, making it suitable for both small businesses and large enterprises. Its cloud-based infrastructure allows organizations to scale their analytics capabilities without performance issues.

2. Regular Updates: Microsoft consistently releases updates for Power BI Service, introducing new features, performance enhancements, and security improvements. This ensures that users have access to the latest tools and functionalities without significant disruptions.

3. Personalization and Flexibility: Users can create customized dashboards and reports tailored to their specific business needs. Power BI Service supports various visualization options and allows integration with programming languages like Python and R for advanced analytics.

4. Cost Efficiency: Power BI Service offers flexible pricing plans, including free and paid versions, allowing organizations to choose options that fit their budgets. Its cloud-based nature eliminates the need for costly on-premises infrastructure, reducing overall expenses.

5. Data Security: Power BI Service prioritizes data protection with features like role-based access control, end-to-end encryption, and compliance with global standards such as GDPR and ISO 27001. Integration with Microsoft Azure enhances security through advanced threat assessment and multi-factor authentication.

6. Real-Time Data Analysis: Power BI Service allows users to connect to live data sources, enabling real-time monitoring of key metrics and performance indicators. This capability facilitates prompt responses to changing business conditions.

7. Mobile Access and Collaboration: With native mobile applications, Power BI Service ensures that users can access dashboards and reports on-the-go. Its collaboration features support real-time sharing and collaborative analysis among team members.

8. Integration with Microsoft Ecosystem: Power BI Service seamlessly integrates with other Microsoft products, such as Office 365 and Azure services, enhancing productivity and providing a unified experience across platforms.

INTRODUCTION TO DATA SOURCES

Power BI is a powerful business analytics tool that allows users to connect to a wide variety of data sources, enabling comprehensive data analysis and visualization. Understanding the types of data sources available is crucial for effectively leveraging Power BI in your projects.

Categories of Data Sources in Power BI:

1. File-Based Sources:

- **Excel Workbooks:** Import data directly from Excel files, utilizing existing spreadsheets and tables.
- **Text/CSV Files:** Ingest data from plain text or comma-separated values files for straightforward integration.
- **XML and JSON Files:** Handle hierarchical data structures commonly used in web services and configurations.
- **PDF Documents:** Extract tabular data from PDF files for analysis.
- **Parquet Files:** Work with columnar storage files optimized for performance.
- **Folders:** Import data from multiple files within a folder, useful for batch processing similar datasets.

2. Database Connections:

- **SQL Server:** Connect to Microsoft SQL Server databases for robust data querying.
- **Access Database:** Integrate data from Microsoft Access databases, commonly used in smaller-scale applications.
- **Oracle, MySQL, PostgreSQL:** Access data from various relational database management systems.
- **IBM Db2, Informix, Netezza:** Connect to IBM's suite of database solutions.
- **Sybase, Teradata:** Integrate data from these enterprise-level database systems.
- **SAP HANA, SAP Business Warehouse:** Access SAP's in-memory and data warehousing solutions.
- **Amazon Redshift:** Connect to Amazon's cloud-based data warehousing service.

3. Microsoft Fabric:

- **Power BI Semantic Models:** Utilize shared data models for consistent reporting.
- **Dataflows:** Leverage ETL processes defined in Power BI service for consistent data preparation.
- **Datamarts (Preview):** Utilize self-service analytics solutions combining data storage and ETL.
- **Warehouses:** Manage large-scale data storage for analytical processing.
- **Lakehouses:** Combine data lake and data warehouse capabilities for unified storage.
- **KQL Databases:** Query large datasets using Kusto Query Language.
- **Metric Sets:** Define and monitor key performance indicators.

4. Online Services:

- **SharePoint Online List:** Import data from SharePoint lists and libraries.
- **Microsoft Exchange Online:** Access data from Exchange mailboxes.
- **Dynamics 365:** Connect to Microsoft's suite of business applications.
- **Google Analytics:** Integrate web analytics data for comprehensive reporting.
- **Salesforce:** Import CRM data for sales and marketing analysis.

5. Other Sources:

- **Web Data:** Scrape and import data from web pages.
- **Data Feeds:** Connect to Open Data Protocol sources for standardized data access.
- **Active Directory:** Import organizational data from Active Directory services.
- **R and Python Scripts:** Execute scripts to import data, allowing for custom data processing.

INTRODUCTION TO VISUALS

Power BI provides a comprehensive suite of visualization tools designed to transform raw data into meaningful insights. Understanding the variety of available visuals and their appropriate applications is crucial for effective data analysis and presentation. Below is an in-depth overview of key Power BI visuals:

1. Card Visuals:

- **Card:** Displays a single data point, ideal for highlighting key metrics such as total sales or profit margins.
- **Multi-Row Card:** Presents multiple data points in a consolidated format, useful for showcasing several KPIs simultaneously.

2. Charts:

- **Bar and Column Charts:** Compare categorical data across different categories. Bar charts are horizontal, while column charts are vertical.
- **Stacked Bar and Column Charts:** Show the contribution of individual categories to the whole, emphasizing the part-to-whole relationships.
- **Line Chart:** Illustrates trends over time, making it suitable for time-series data analysis.
- **Area Chart:** Similar to line charts but with the area beneath the line filled, highlighting the magnitude of trends.
- **Ribbon Chart:** Depicts rank changes over time, useful for visualizing shifts in rankings among categories.
- **Waterfall Chart:** Displays the cumulative effect of sequential positive and negative values, often used in financial analysis.

3. Pie and Donut Charts:

- **Pie Chart:** Represents data as slices of a circle, illustrating part-to-whole relationships.
- **Donut Chart:** A variation of the pie chart with a blank center, providing a more modern aesthetic while conveying similar information.

4. Tables and Matrices:

- **Table:** Displays data in rows and columns, suitable for detailed data analysis.
- **Matrix:** Similar to a table but allows for data to be grouped by multiple dimensions, akin to a pivot table in Excel.

5. Maps:

- **Map Visuals:** Plot geographical data points, useful for location-based analyses.
- **Filled Map:** Highlights regions on a map based on data values, such as sales volume by state.
- **Shape Map:** Displays custom shapes and regions, allowing for specialized geographical representations.

6. Advanced Visuals:

- **Scatter Chart:** Shows the relationship between two numerical variables, aiding in correlation analysis.
- **Treemap:** Visualizes hierarchical data using nested rectangles, effectively displaying part-to-whole relationships within categories.
- **Gauge:** Represents progress toward a goal, such as actual sales versus a target.
- **Funnel Chart:** Illustrates stages in a process, commonly used in sales pipelines to show progression and drop-off rates.
- **Decomposition Tree:** Breaks down a metric into its contributing factors, facilitating root cause analysis.
- **Key Influencers:** Identifies factors that significantly impact a particular metric, aiding in understanding data drivers.

7. Slicers:

- **Slicer:** Provides interactive filtering options, allowing users to focus on specific data subsets within reports.

INTRODUCTION TO MICROSOFT SQL SERVER

Microsoft SQL Server is a relational database management system (RDBMS) developed by Microsoft, designed to store and retrieve data as requested by other software applications, whether on the same computer or across a network. It supports various workloads, from small applications to large-scale enterprise systems.

Key Features:

- **Relational Database Management:** SQL Server organizes data into tables with rows and columns, allowing for efficient storage and retrieval.

- **Transact-SQL (T-SQL):** An extension of SQL, T-SQL includes procedural programming and local variable support, enabling complex operations and business logic implementation within the database.

- **Data Warehousing:** SQL Server can function as a data warehouse, providing employees with insights into business operations and data. Its business intelligence platform offers a scalable infrastructure, promoting the use of business intelligence across the enterprise and delivering it where users need it.

- **Integration with .NET Framework:** Since SQL Server 2005, it supports programming languages running on the .NET Common Language Runtime (CLR) for creating stored procedures, allowing developers to use languages like C# or VB.NET to write database logic.

- **High Availability and Disaster Recovery:** Features like clustering, replication, log shipping, mirroring, and AlwaysOn Availability Groups ensure data redundancy and minimize downtime.

Editions:

SQL Server is available in various editions tailored to different audiences and workloads, ranging from small applications to large-scale enterprise systems.

Management Tools:

- **SQL Server Management Studio (SSMS):** A free, integrated environment for managing SQL infrastructure, providing tools to configure, monitor, and administer instances of SQL Server.

- **SQL Server Data Tools (SSDT):** A development environment for building SQL Server relational databases, Azure SQL databases, Integration Services packages, Analysis Services data models, and

  Reporting Services reports.

Programming Support:

SQL Server supports various programming approaches:

- **T-SQL:** For writing stored procedures and functions.

- **ODBC and OLE DB:** For data access.

- **SQL CLR:** Since SQL Server 2005, it supports creating stored procedures using .NET languages like C# or VB.NET.

- **sqlcmd Utility:** Provides the ability to execute SQL commands from batch files.

- **Integration with Development Platforms:** Such as ADO.NET for .NET applications, JDBC for Java applications, and PDO for PHP applications.

INTRODUCTION TO SQL SERVER MANAGEMENT STUDIO

SQL Server Management Studio (SSMS) is an integrated environment developed by Microsoft for configuring, managing, and administering all components within Microsoft SQL Server. First introduced with SQL Server 2005, SSMS has become the primary tool for database administrators and developers to interact with SQL Server instances.

Key Features:

- **Object Explorer:** Allows users to browse, select, and manage all objects within the server, such as databases, tables, views, and stored procedures.

- **Query Editor:** Provides a powerful interface for writing, executing, and optimizing SQL queries. It supports syntax highlighting, code completion, and debugging tools to enhance productivity.

- **Template Explorer:** Offers pre-defined templates for creating various SQL Server objects, aiding in standardizing and speeding up development tasks.

- **Solution Explorer:** Helps organize and manage SQL Server-related projects and files, facilitating better project management and version control.

- **Activity Monitor:** Enables real-time monitoring of SQL Server activities, including processes, resource waits, data file I/O, and recent expensive queries, assisting in performance tuning and troubleshooting.

Installation:

SSMS is available as a free download from Microsoft's official website. It can be installed independently of SQL Server, allowing users to manage SQL Server instances remotely. The installation process is straightforward, with default settings suitable for most environments.

Usage:

1. Connecting to a Server: Upon launching SSMS, the "Connect to Server" dialog prompts users to enter the server name and authentication details to establish a connection.

2. Navigating Object Explorer: After connecting, Object Explorer displays the hierarchical structure of the server, allowing users to interact with various components.

3. Executing Queries: The Query Editor facilitates the creation and execution of SQL queries, with features like syntax highlighting and execution plans to optimize performance.

4. Managing Security: SSMS provides interfaces to configure security settings, including logins, roles, and permissions, ensuring proper access control.

5. Backup and Restore: Users can perform backup and restore operations through SSMS, ensuring data protection and recovery capabilities.

Advantages:

- **User-Friendly Interface:** SSMS offers an intuitive graphical interface, making it accessible to both beginners and experienced professionals.

- **Comprehensive Toolset:** It combines a wide range of tools, from query writing to performance monitoring, within a single application.

- **Regular Updates:** Microsoft frequently updates SSMS, incorporating new features and improvements to align with the latest SQL Server capabilities.

INTRODUCTION TO PROJECT

The Mortality Analysis and Life Expectancy Project is a comprehensive study designed to evaluate mortality trends and life expectancy across different demographics, time periods, and influencing factors. By leveraging Power BI, this project utilizes data visualization techniques to provide in-depth insights into mortality rates, causes of death, demographic variations, and the impact of healthcare advancements over the years from 2000 to 2015**.**

The study focuses on understanding age-specific mortality rates, gender-based life expectancy differences, major causes of death, and the correlation between socio-economic factors and longevity. By analyzing patterns over time, the project aims to highlight key trends in life expectancy improvements and identify critical health challenges that persist.

Key Objectives of the Project:

- Analyze mortality trends across different age groups, genders, and regions.
- Evaluate life expectancy variations over the past century and their influencing factors.
- Identify leading causes of death and how they have changed over time due to medical advancements.
- Assess the impact of healthcare improvements, economic conditions, and social factors on longevity.
- Compare life expectancy differences among insured and uninsured populations.
- Understand the relationship between chronic diseases, lifestyle factors, and mortality rates**.**

Importance of the Project:

Understanding mortality patterns and life expectancy trends is crucial for healthcare policymakers, researchers, and public health officials to make data-driven decisions aimed at improving public health strategies. By identifying high-risk populations, major causes of death, and the impact of medical interventions, this study helps in developing better healthcare policies, preventive measures, and resource allocation strategies to enhance life expectancy and reduce premature deaths.

With the help of interactive dashboards, this project presents data in an easily interpretable format, allowing stakeholders to explore historical trends, detect emerging health threats, and formulate policies that contribute to better healthcare outcomes and increased life expectancy.

1. National Mortality Overview

The National Mortality Overview serves as a comprehensive summary of mortality trends across different regions and demographics. This page provides an in-depth analysis of death rates, major causes of mortality, life expectancy variations, and demographic influences on mortality trends over the years.

By leveraging Power BI, the dashboard visualizes key insights such as total deaths, age-specific mortality rates, gender-based mortality distribution, and regional comparisons. It also highlights historical patterns (2000-2015), improvements in healthcare, and socio-economic factors affecting mortality rates.

Key Insights Covered:

- Total Mortality Count over different time periods.
- Age and Gender-based Mortality Trends, identifying high-risk groups.
- Leading Causes of Death, including chronic diseases, infectious diseases, and external factors.
- Comparison of Life Expectancy across different demographics and regions.
- Impact of Healthcare Advancements on reducing mortality rates over the century.

Purpose and Importance:

This overview helps healthcare professionals, policymakers, and researchers understand mortality trends at a national level**,** enabling them to develop strategies for improving healthcare services, reducing premature deaths, and enhancing life expectancy**.** By identifying high-risk groups and the most common causes of death, this analysis supports data-driven decision-making for better public health outcomes**.**

2.State-wise Mortality Trends:

The State-wise Mortality Trends page provides an interactive and visual representation of mortality patterns across different states**,** allowing for a comparative analysis of death rates, causes of mortality, and life expectancy variations at a regional level. This page is designed to help identify geographical

disparities in health outcomes, assess state-level healthcare effectiveness**, and uncover** trends in mortality rates over time (2000-2015)**.**

Key Visual Elements:

I. **State-wise Mortality Heatmap**
   - A color-coded heatmap displaying mortality rates across different states.
   - Darker shades indicate higher mortality rates, while lighter shades indicate lower mortality rates.

II. **State-wise Mortality Comparison Chart**
   - A bar or line chart comparing mortality rates across multiple states.
   - Helps identify states with higher-than-average or lower-than-average death rates.

III. **Top Causes of Death by State**
   - A stacked bar chart or tree map categorizing the most common causes of death in each state.
   - Enables analysis of regional health risks and disease prevalence.

IV. **Life Expectancy by State Over Time**
   - A trend line chart showing life expectancy fluctuations in each state from 2000 to 2015**.**
   - Helps in understanding how healthcare improvements, lifestyle changes, and policies have impacted longevity.

V. **Demographic Breakdown by State**
   - Mortality rates segmented by age group, gender, and socio-economic status.
   - Identifies states with higher mortality risks for specific populations.

Purpose and Insights:

The State-wise Mortality Trends page provides a detailed, data-driven view of how mortality rates and life expectancy differ across states. This enables health officials, policymakers, and researchers to:

- Identify high-risk regions that require targeted healthcare interventions.
- Compare mortality trends across different time periods to assess improvements.
- Understand the impact of regional healthcare infrastructure on mortality outcomes.
- Develop policies for better resource allocation and improved public health.

3.Gender-Based Mortality Analysis :

The Gender-Based Mortality Analysis page examines mortality trends across different gender groups, highlighting disparities in death rates, life expectancy, and leading causes of death. This analysis helps identify high-risk groups, gender-specific health concerns, and the impact of medical advancements on mortality trends over time.

Key Visuals:

- Mortality Rate by Gender (Male vs. Female vs. Other)
- Life Expectancy Comparison Across Genders
- Leading Causes of Death by Gender
- Trends Over Time (1924-2024)

4.Cause-Specific Mortality Trends – Introduction

The Cause-Specific Mortality Trends page analyzes the leading causes of death over time, helping to identify patterns, emerging health threats, and the effectiveness of medical advancements in reducing mortality. This analysis provides insights into how different diseases, external factors, and socio-economic conditions contribute to overall death rates.

Key Visuals:

- Mortality Rates by Leading Causes (e.g., heart disease, cancer, respiratory diseases, accidents)
- Trends in Cause-Specific Deaths Over Time (1924-2024)
- Age and Gender Distribution for Specific Causes of Death
- Regional Variations in Cause-Specific Mortality

5.Population vs. Mortality Insights – Introduction

The Population vs. Mortality Insights page examines the relationship between population growth and mortality rates over time. This analysis helps in understanding how demographic changes, healthcare advancements, and socio-economic factors have influenced death rates and life expectancy from 1924 to 2024.

Key Visuals:

- Population Growth vs. Mortality Rate Trends
- Age Group Distribution and Corresponding Mortality Rates
- Impact of Urbanization and Rural Population Shifts on Mortality
- Correlation Between Population Density and Life Expectancy

.

SYSTEM ANALYSIS STUDY REPORT

The Mortality Analysis and Life Expectancy Project is a data-driven system utilizing Power BI to examine mortality trends, life expectancy, and related demographic factors. This system provides valuable insights into population health, regional mortality disparities, and socio-economic influences on life expectancy.

1. System Overview

1.1 Purpose of the System

- To analyze mortality trends and life expectancy variations over time.
- To study the impact of demographic factors (age, gender, region) on mortality.
- To identify leading causes of death and their trends over decades.
- To assist policymakers and healthcare professionals in decision-making.
- To provide interactive dashboards for in-depth data visualization.

1.2 Scope of the System

- Data Processing: Extracting, transforming, and visualizing mortality and life expectancy data.
- Data Storage: Storing historical and demographic mortality records.
- Visualization: Using Power BI dashboards for trend analysis and insights.
- User Access: Accessible to healthcare professionals, researchers, and policymakers.

1.3 Stakeholders

- Government & Public Health Agencies: Policy formulation based on mortality trends.
- Healthcare Institutions: Assessing healthcare effectiveness and planning interventions.
- Researchers & Academics: Analyzing socio-economic impacts on life expectancy.
- Insurance & Actuarial Analysts: Evaluating life expectancy for financial planning.

2. System Analysis

2.1 Existing System & Challenges Traditional mortality analysis relies on:

- Fragmented datasets with limited accessibility.
- Manual report generation leading to inefficiencies.
- Lack of real-time visualization for immediate decision-making.

- Inconsistent data sources affecting trend accuracy.

2.2 Proposed System Features The Power BI-based system enhances analysis by:

- Automating data integration for real-time insights.
- Providing interactive dashboards for mortality trends.
- Offering regional and demographic mortality comparisons.
- Improving decision-making with predictive analytics.

2.3 Functional Requirements The system should:

- Ingest and process mortality data from multiple sources (CSV, databases).
- Analyze demographic trends and causes of death.
- Generate Power BI dashboards with interactive reports.
- Provide secure access to authorized users.
- Ensure data privacy and confidentiality.

2.4 Non-Functional Requirements

- Scalability: Handle large historical datasets efficiently.
- Security: Protect sensitive mortality data.
- Performance: Deliver quick queries and visualization updates.
- Usability: Ensure an intuitive interface for users.

3. System Design

3.1 System Architecture

- Data Sources: Mortality records (CSV, SQL databases, public health data).
- Data Processing Layer: Power BI transformations for structuring data.
- Visualization Layer: Power BI dashboards for trend insights.
- User Interface: Accessible via web-based or desktop platforms.

3.2 Data Flow Diagram (DFD) Level 0 (Context Diagram):

- Users interact with Power BI to analyze mortality trends. Level 1 (High-Level Data Flow):
- Data is imported from historical mortality records.
- Data is cleaned, transformed, and structured.

- Power BI generates reports and dashboards.

4. Implementation Strategy

4.1 Data Collection & Processing

Step 1: Import data (e.g., population, mortality rates, cause of death, demographics).

Step 2: Clean and structure the data using Power Query in Power BI.

Step 3: Create calculated columns and measures for analytical insights.

4.2 Dashboard Development

- National Mortality Overview: Summary of death rates and life expectancy.
- State-wise Mortality Trends: Regional mortality rate comparisons.
- Gender-Based Mortality Analysis: Life expectancy and mortality variations by gender.
- Cause-Specific Mortality Trends: Leading causes of death over decades.
- Population vs. Mortality Insights: Relationship between population growth and mortality rates.

4.3 System Testing & Validation

- Data Accuracy Testing: Cross-checking results with official records.
- Performance Testing: Ensuring fast response times in Power BI.
- Security Testing: Verifying data protection and privacy compliance.

5. Benefits & Expected Outcomes

- Improved Policy Decisions: Data-driven insights for public health strategies.
- Enhanced Healthcare Planning: Identification of high-risk groups for interventions.
- Efficient Data Analysis: Real-time mortality trends using Power BI.
- Better Resource Allocation: Governments and healthcare providers can optimize resources effectively.

SCOPE OF SYSTEM

The Mortality Analysis and Life Expectancy System leverages **Power BI** to analyze mortality trends, demographic influences, and life expectancy variations. The system's scope spans key areas to enhance public health research, policy formulation, and healthcare planning.

1. Data Centralization and Integration

- **Unified Data Repository:** Aggregates mortality data from diverse sources—such as government records, census data, health surveys, and epidemiological studies—into a centralized platform for comprehensive analysis.

2. Predictive Analytics and Trend Analysis

- **Forecasting:** Uses predictive models to estimate future mortality rates, identify high-risk age groups, and analyze life expectancy trends based on historical data.

3. Real-Time Data Visualization

- **Interactive Dashboards:** Provides real-time insights into mortality patterns, regional disparities, and demographic-specific life expectancy trends to support informed decision-making.

4. Public Health and Policy Planning

- **Health Strategy Development:** Aids policymakers in identifying healthcare priorities, resource allocation, and targeted interventions to improve life expectancy.

5. Socio-Economic and Environmental Impact Analysis

- **Demographic and Economic Factors:** Examines how factors such as income levels, healthcare access, and environmental conditions influence mortality rates and life expectancy.

6. Cause-Specific Mortality Analysis

- **Disease and External Factors:** Analyzes leading causes of death, their long-term trends, and regional variations to guide medical research and preventive healthcare strategies.

7. Compliance and Data Security

- **Regulatory Adherence:** Ensures compliance with healthcare data protection laws and ethical standards while maintaining data confidentiality.

8. Risk Management and Actuarial Studies

- **Mortality Risk Assessment:** Supports insurance and actuarial studies by providing mortality probabilities and life expectancy insights for financial planning.

9. Historical and Future Trends Evaluation

- **Comparative Analysis:** Assesses changes in mortality patterns from 1924 to 2024, helping researchers and healthcare institutions understand historical shifts and future projections.

HARDWARE & SOFTWARE REQUIREMENTS

1. Hardware Requirements

The hardware components play a crucial role in supporting Power BI, SQL Server, and data processing tasks. The following specifications are recommended for optimal performance:

Hardware Specification	
RAM	**16GB**
Hard Disk	**512GB**
Processor	**Ryzen 7**
Monitor	**Monitor with Full HD resolution**

2.Software Requirements

A combination of software tools is required for data storage, transformation, modeling, visualization, and reporting

Software Specification	
Operating System	**Windows 11**
Web Browser	**Chrome**
Language	**SQL, DAX**
Software	**PBI Desktop, PBI Service, Microsoft SQL Server**

3.Additional Tools & Plugins:

DAX Studio: For advanced DAX formula debugging

Power Query: For ETL (Extract, Transform, Load) operations

Azure SQL Database (Optional): For cloud-based data storage

DFD (DATA FLOW DIAGRAM

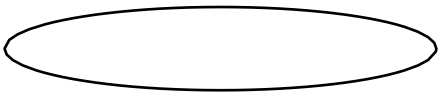
it is easy to understand the flow of data through system with the right data flow diagram software. A data flow diagram (DFD) maps out the flow of information for any process or system. It uses defined symbols like rectangles, circles, and arrows, plus short text labels, to show data inputs, outputs, storage point and the routes between each destination. Data flowchart can range from simple, even hand-drawn process overviews, to in-depth, multi-level DFDs that dig progressively deeper into how the data is handled. They can be used to analyses an existing system or model a new one.

Symbols and Notations Used in DFDs: The symbols depict the four components of data flow diagram: -

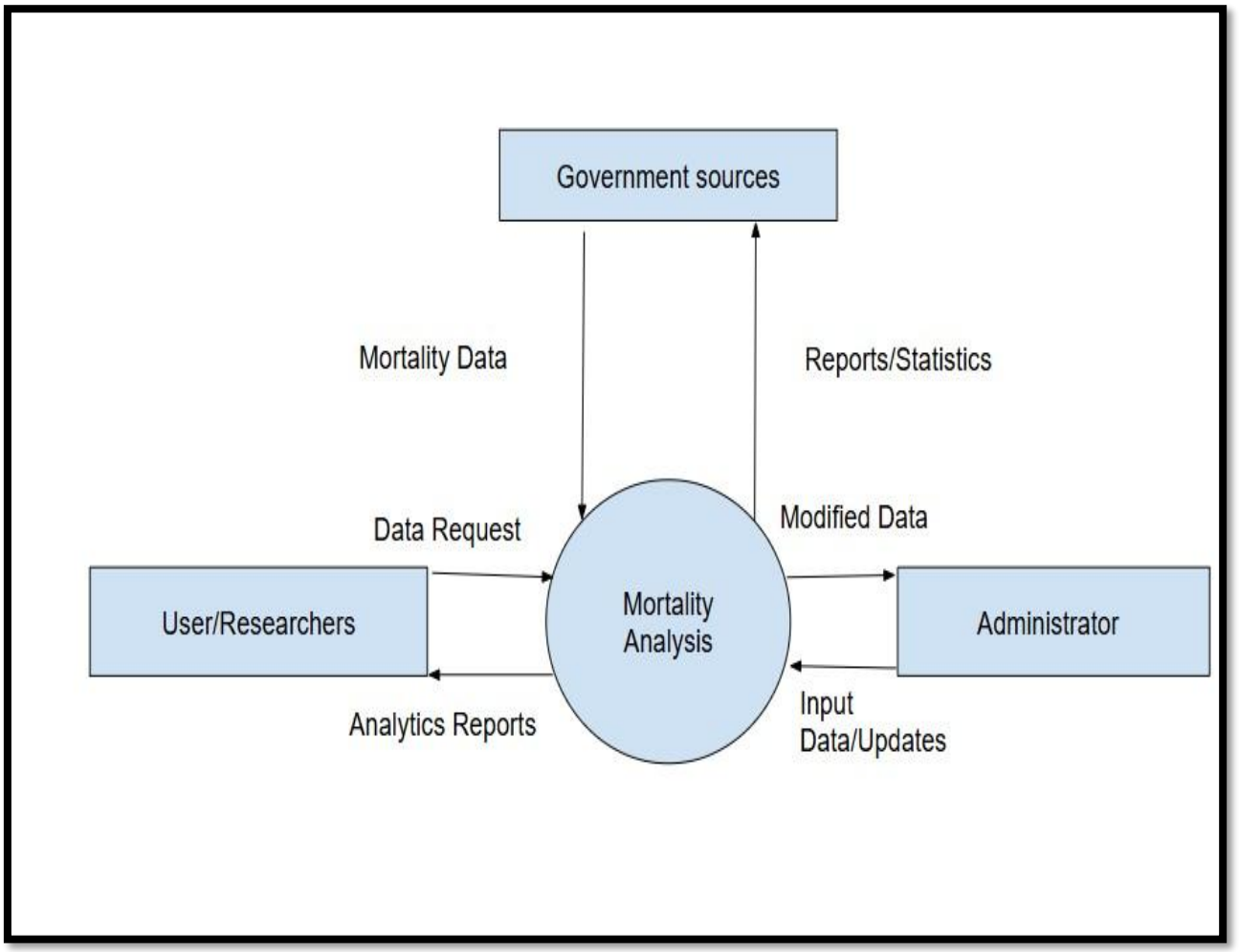
External entity :- An outside system that sends or receives data, communication with the system being diagrammed. They are the sources and destinations of information entering or leaving the system. They might be an outside organization or person, a computer system or a business system. They might also known as terminators, sources and sinks or actors. They are typically drawn on edges of diagram.

Data Store: Files or repositories that hold information for later use, such as a database table or only the register person can get appointment form.

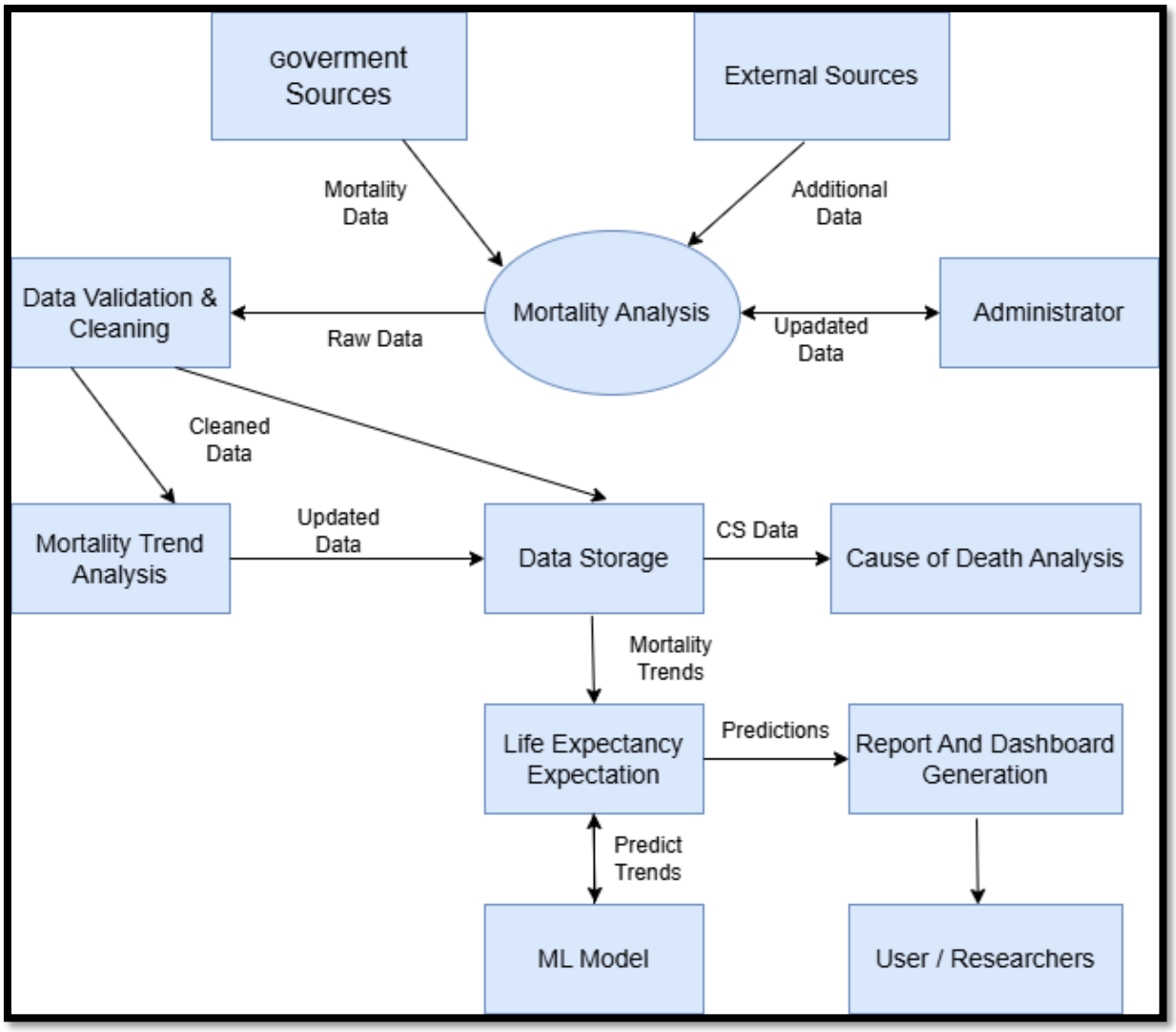
Data flow: the route that data takes between the external entities, processes and data stores. It portrays the interface between the other components and is shown with arrows, typically Labelled with a short data name, like “Appointment details.”

Notation	Symbol
External Entity	
Process	
Data Flow	

1.DATA FLOW DIAGRAM (0-Level):



2. DATA FLOW DIAGRAM (1-Level) :



ENTITY RELATIONSHIP DIAGRAM (E-R)

The ER or (Entity Relational Model) is a high-level conceptual data model diagram. Entity-Relation model is based on the notion of real-world entities and the relationship between them. ER modelling helps you to analyze data requirements systematically to produce a well-designed database. So, it is considered a best practice to complete ER modelling before implementing your database.

Components of the ER Diagram

This model is based on three basic concepts:

- Entities
- Attributes
- Relationships

Entities:

A real-world thing either living or non-living that is easily recognizable and not recognizable. It is anything in the enterprise that is to be represented in our database. It may be a physical thing or simply a fact about the enterprise or an event that happens in the real world.

Attributes:

An attribute describes the property of an entity. An attribute is represented as an oval in an ER diagram. There are four types of attributes:

1. Key attribute.
2. Composite attribute.
3. Multivalued attribute.
4. Derived attribute.

Relationships:

Defines the numerical attributes of the relationship between two entities or entity sets. Different types of cardinal relationships are:

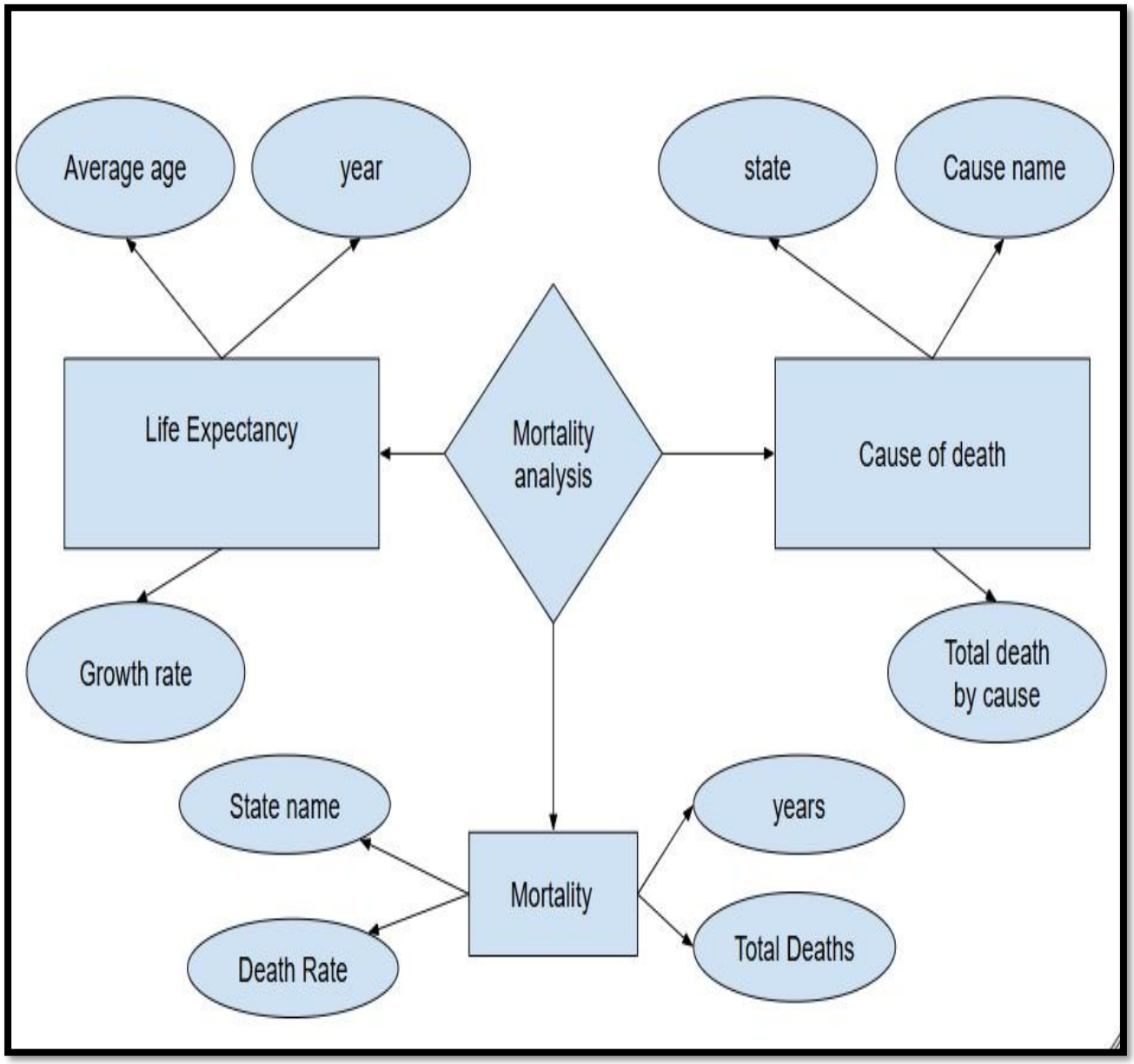
- One-to-One Relationships
- One-to-Many Relationships
- Many-to-One Relationships
- Many-to-Many Relationships

ER Diagram Notations

ER Diagram is a visual representation of data that describes how data is related to each other.

- Rectangles: This symbol represents entity types
- Ellipses: Symbol represents attributes
- Diamonds: This symbol represents relationship types
- Lines: It links attributes to entity types and entity types with other relationship types
- Primary key: Attributes are underlined
- Double Ellipses: Represent Multi-valued attribute
- Dashed Ellipses: Derived Attributes
- Double Rectangles: Weak Entity Sets
- Double Lines: Total participation of an entity in a relationship set

ER-DIAGRAM:


Average age
year
state
Cause name
Life Expectancy
Mortality analysis
Cause of death
Growth rate
Total death by cause
State name
years
Mortality
Death Rate
Total Deaths

FEASIBILITY STUDY

A feasibility study evaluates the practicality and viability of implementing the India Mortality Analytics and Life Expectancy Dashboard. It helps determine whether the project is technically, financially, and operationally feasible. This analysis ensures that the system can be successfully developed, deployed, and maintained.

The feasibility study is divided into the following key aspects:

1Technical Feasibility

2 Economic Feasibility

3 Operational Feasibility

4 Legal Feasibility

5 Schedule Feasibility

1. Technical Feasibility:

Objective: Assess whether the project can be implemented using existing technology and infrastructure.

Hardware & Software Compatibility:

Power BI, SQL Server, and data visualization tools are widely used and have strong industry support.

The project requires moderate-to-high computing power for data analysis, which is feasible with modern computers and cloud services.

The system can run on Windows-based platforms with minimal additional software requirements.

Data Storage & Processing Feasibility:

The dataset used (~40,000 rows) is manageable with Microsoft SQL Server and Power BI.

Advanced DAX and SQL queries ensure efficient data retrieval and processing.

Cloud-based Power BI Service ensures remote access and scalability.

Conclusion: The project is technically feasible, as the required tools and infrastructure are readily available.

2. Economic Feasibility

Objective: Determine whether the project is cost-effective and provides a good return on investment (ROI).

Development Cost:

Power BI Desktop is free, reducing software costs.

SQL Server Express (free edition) can be used, lowering database expenses.

No need for expensive third-party software.

Operational Cost:

Power BI Service requires a Pro license (₹780/user/month) for sharing dashboards, which is affordable.

Minimal hardware upgrades required for system implementation.

Cost-Benefit Analysis:

The dashboard provides valuable insights for healthcare policy-making.

The cost of implementation is low compared to the benefits of improved mortality analysis.

Conclusion: The project is economically feasible due to its low-cost implementation and high potential benefits.

3. Operational Feasibility:

Objective: Determine if the system will function effectively in real-world scenarios.

User Adoption & Training:

Power BI is user-friendly, making it easy for analysts and policymakers to use.

Training requirements are minimal, as basic Power BI skills are sufficient.

Scalability & Maintenance:

The system can easily handle additional data from future years.

Dashboards are easily updatable without major system changes.

Conclusion: The project is operationally feasible, as it is easy to use, scalable, and requires minimal training.

4. Legal Feasibility:

Objective: Ensure compliance with legal and ethical standards.

Data Privacy & Security:

The dataset does not contain personal or sensitive information, ensuring compliance with data protection laws.

Adhering to GDPR and Indian IT Act guidelines on data handling.

Intellectual Property & Licensing:

All tools used (Power BI, SQL Server) are properly licensed and compliant with software regulations.

The data used is either publicly available or ethically sourced.

5. Schedule Feasibility:

Objective: Assess whether the project can be completed within the proposed timeline.

Project Timeline & Milestones:

Data collection & cleaning: 1 week

Database setup & SQL integration: 2 weeks

Dashboard design & Power BI implementation: 4 weeks

Testing & optimization: 2 weeks

Final deployment & documentation: 1 week

Conclusion: The project is schedule feasible, as it can be completed within the planned timeframe.

DATA DICTIONARY

Column Name	Data Type	Description	Constraints
Region	VARCHAR(100)	Name of the Indian state/UT	Not Null
Year	INT	Year of data recording (2000-2015)	Range: 2000 – 2015
Gender	VARCHAR(10)	Gender category (Male/Female)	Allowed values: 'Male', 'Female'
Death_Rate	FLOAT(5,2)	Deaths per 1,000 people	Range: 5.0 – 15.0
Cause of Death	VARCHAR(100)	Leading cause of death	Not Null
Population	BIGINT	Total population of the state	Greater than 0
Total_Deaths	BIGINT	Total number of deaths in the given year/state	Calculated from Population Death_Rate / 1000
Life_Expectancy	FLOAT(5,2)	Average life expectancy in years	Range: 60 – 80
Age_Range	VARCHAR(10)	Age category of deceased persons	Values: '0-18', '19-35', '36-50', '51-65', '66+'

PROBLEM STATEMENTS WITH SQL QUERIES AND OUTPUT

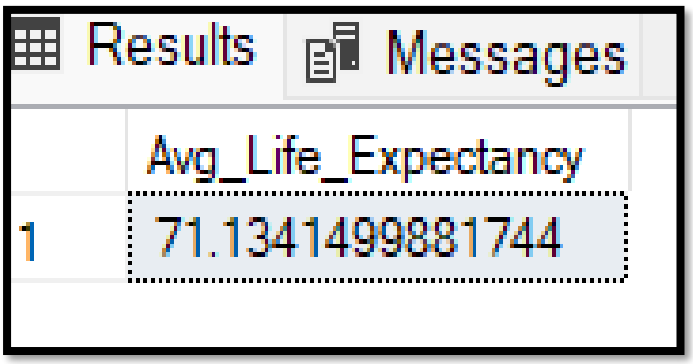
DASHBOARD 1 :National Mortality Overview

KPI Card 1: Avg of Life Expectancy

Problem Statement 1: What is the overall average life expectancy of the country?

Query:

```
SELECT
AVG(Life_Expectancy) AS Avg_Life_Expectancy
FROM finaldataset;
```

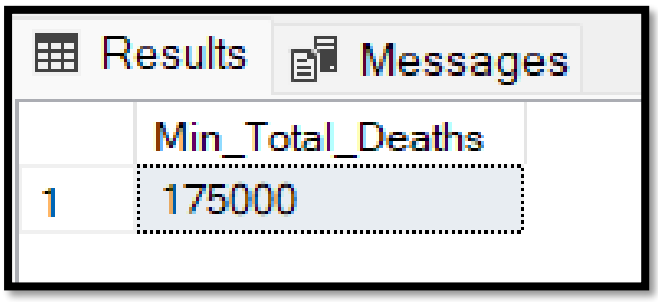


KPI Card 2: Min of Total Deaths

Problem Statement 2: What is the minimun numbers of total deaths recorded in the States?

Query:

```
SELECT
MIN(Total_Deaths) AS Min_Total_Deaths
FROM finaldataset;
```

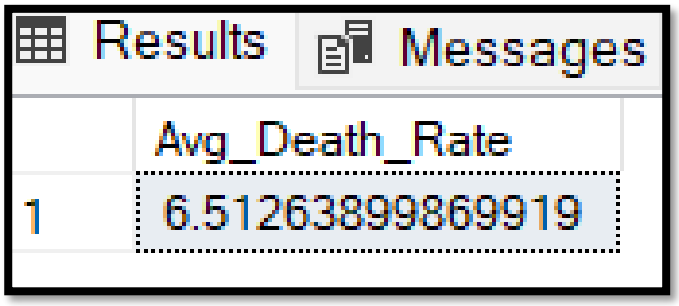


KPI Card 3: Average of total death rate

Problem Statement 3: What is the overall average death rate in all states?

Query:

```
SELECT
AVG(Death_Rate) AS Avg_Death_Rate
FROM finaldataset;
```

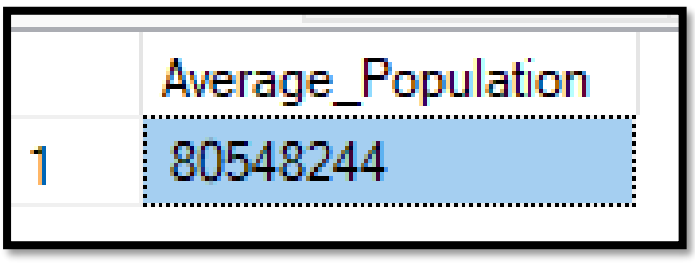


KPI Card 4: Average of whole population

Problem Statement 4: What is the overall average of population in all states?

Query:

```
SELECT
AVG(CAST(Population AS BIGINT)) AS Average_Population
FROM finaldataset;
```



KPI Card 5: Total Regions

Problem Statement 5: How many unique regions are present in the dataset, and what is their distribution in the analysis of mortality and life expectancy?

Query:

SELECT

COUNT (DISTINCT Region) AS Count_of_Region

FROM finaldataset;

	Count_of_Region
1	10

Chart 1: Line Chart

Problem Statement: Analyze how the average life expectancy has changed across different states over the years (2000-2015) ?

Query:

```
SELECT
    Year,
    Region,
    AVG([Life_Expectancy]) AS Avg_Life_Expectancy
FROM finaldataset
GROUP BY Year, Region
ORDER BY Year, Region;
```

	Year	Region	Avg_Life_Expectancy
1	2000	Andhra Pradesh	70.206779738604
2	2000	Bihar	61.9782608142798
3	2000	Gujarat	68.1520835558573
4	2000	Karnataka	70.9197364606355
5	2000	Kerala	76.992856888544
6	2000	Maharashtra	70.0791668362088
7	2000	Rajasthan	65.9838712138514
8	2000	Tamil Nadu	74.0104651340218
9	2000	Uttar Pradesh	63.8948278098271
10	2000	West Bengal	66.7359377145767
11	2001	Andhra Pradesh	70.3161291307019
12	2001	Bihar	62.292727244984
13	2001	Gujarat	68.3609371185303
14	2001	Karnataka	71.3822785389574

Chart 2 : Clustered Bar Chart

Problem Statement: Identify states with the highest and lowest total deaths to analyze mortality concentration across India."

Query:

```
SELECT
   Region,
   SUM(CAST([Total_Deaths] AS BIGINT)) AS Total_Deaths
FROM finaldataset
GROUP BY Region
ORDER BY Total_Deaths DESC;
```

	Cause_of_Death	Total_Deaths
1	Cardio	1497498221
2	Infect	1086603989
3	Accident	1017022732
4	Resp	724376116
5	Cancer	498191716
6	Stroke	384119964
7	Diabetes	286828336

Chart 3 : Pie Chart

Problem Statement: Understand the major causes of death in India by examining the percentage distribution of total deaths?

Query:

```
SELECT
   [Cause_of_Death],
   SUM(CAST([Total_Deaths] AS BIGINT)) AS Total_Deaths
FROM finaldataset
GROUP BY [Cause_of_Death]
```

ORDER BY Total_Deaths DESC;

	Region	Total_Deaths
1	Uttar Pradesh	1387869482
2	Bihar	876739802
3	West Bengal	632162598
4	Gujarat	527316135
5	Rajasthan	510478264
6	Karnataka	381771231
7	Tamil Nadu	367602648
8	Maharashtra	327954591
9	Andhra Pradesh	288211383
10	Kerala	194534940

Chart 4: Filled Map

Problem Statement: Visualize the geographic distribution of mortality rates across different states of India?

Query:

```
SELECT
   Region,
   AVG([Death_Rate]) AS Avg_Death_Rate,
   SUM(CAST(Population AS BIGINT)) AS Total_Population,
   SUM(CAST([Total_Deaths] AS BIGINT)) AS Total_Deaths,
   AVG([Life_Expectancy]) AS Avg_Life_Expectancy
FROM finaldataset
GROUP BY Region;
```

	Region	Avg_Death_Rate	Total_Population	Total_Deaths	Avg_Life_Expectancy
1	Maharashtra	5.90862384800031	56047184213	327954591	72.2256880205098
2	Rajasthan	7.01143148779138	73643665498	510478264	68.2436606128279
3	Kerala	5.0808124360493	38331103089	194534940	78.9482447280807
4	West Bengal	6.37370071131491	100154316642	632162598	69.1291202587594
5	Gujarat	6.49409949696479	82194944705	527316135	70.2597014925373
6	Tamil Nadu	5.46380198025467	67689451867	367602648	76.1999009689482
7	Uttar Pradesh	8.01780310467735	175345708174	1387869482	66.2302589772279
8	Karnataka	5.76520075339452	66817650592	381771231	73.1766573766144
9	Andhra Pradesh	5.80130129939204	50081850661	288211383	72.3282282221186
10	Bihar	9.30494527294861	95176571895	876739802	64.2313432721949

DASHBOARD 2 : State-Wise Mortality Trends

KPI Card 1: Avg of death rate

Problem Statement 1: Identify the states with the highest mortality rates to assess healthcare efficiency across regions?

Query:

```
 SELECT TOP 1
    Region,
    AVG(Death_Rate) AS Avg_Death_Rate
FROM finaldataset
GROUP BY Region
ORDER BY Avg_Death_Rate DESC;
```

	Region	Avg_Death_Rate
1	Bihar	9.3049452729486 1

KPI Card 2: Min of death rate

Problem Statement : Identify the states with the highest mortality rates to assess healthcare efficiency across regions?

Query:

```
SELECT TOP 1
    Region,
    AVG(Death_Rate) AS Avg_Death_Rate
FROM finaldataset
GROUP BY Region
ORDER BY Avg_Death_Rate ASC;
```

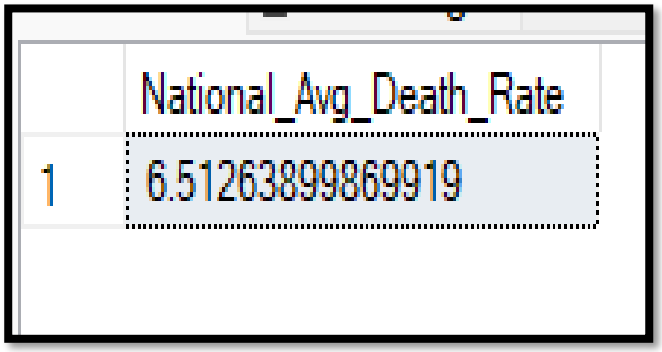
	Region	Avg_Death_Rate
1	Kerala	5.0808124360493

KPI Card 3: Min of death rate

Problem Statement : Identify the states with the national mortality rates to assess healthcare efficiency across regions?

Query:

```
SELECT
   AVG(Death_Rate) AS National_Avg_Death_Rate
FROM finaldataset;
```



	National_Avg_Death_Rate
1	6.51263899869919

Chart 1: Stacked Column Chart

Problem Statement : Analyze how total deaths have changed over time for each state to identify mortality trends?

Query:

```
SELECT
   Year,
   Region,
   SUM([Total_Deaths]) AS Total_Deaths
FROM finaldataset
GROUP BY Year, Region
ORDER BY Year, Region;
```

	Year	Region	Total_Deaths
1	2000	Andhra Pradesh	18838318
2	2000	Bihar	59952665
3	2000	Gujarat	26810439
4	2000	Karnataka	29559543
5	2000	Kerala	8239251
6	2000	Maharashtra	26086467
7	2000	Rajasthan	34401217
8	2000	Tamil Nadu	33696907
9	2000	Uttar Pradesh	82770929
10	2000	West Bengal	44972659
11	2001	Andhra Pradesh	19677749
12	2001	Bihar	47488377
13	2001	Gujarat	35034614
14	2001	Karnataka	29799489

Chart 2: Clustered Bar Chart

Problem Statement : Compare the average life expectancy across all states to identify variations in health outcomes?

Query:

```
SELECT
   Region,
   AVG([Life_Expectancy]) AS Avg_Life_Expectancy
FROM finaldataset
GROUP BY Region
ORDER BY Avg_Life_Expectancy DESC;
```

	Region	Avg_Life_Expectancy
1	Kerala	78.9482447280807
2	Tamil Nadu	76.1999009689482
3	Karnataka	73.1766573766144
4	Andhra Pradesh	72.3282282221186
5	Maharashtra	72.2256880205098
6	Gujarat	70.2597014925373
7	West Bengal	69.1291202587594
8	Rajasthan	68.2436606128279
9	Uttar Pradesh	66.2302589772279
10	Bihar	64.2313432721949

Chart 3: Clustered Column Chart

Problem Statement 3 : Analyze the correlation between literacy rates and death rates in different states?

Query:

```
SELECT
   Region,
   AVG(Death_Rate) AS Avg_Death_Rate,
   AVG(Life_Expectancy) AS Avg_LifeExpectancy_Rate
FROM finaldataset
GROUP BY Region
```

ORDER BY Avg_Death_Rate DESC;

	Region	Avg_Death_Rate	Avg_LifeExpectancy_Rate
1	Bihar	9.30494527294861	64.2313432721949
2	Uttar Pradesh	8.01780310467735	66.2302589772279
3	Rajasthan	7.01143148779138	68.2436606128279
4	Gujarat	6.49409949696479	70.2597014925373
5	West Bengal	6.37370071131491	69.1291202587594
6	Maharashtra	5.90862384800031	72.2256880205098
7	Andhra Pradesh	5.80130129939204	72.3282282221186
8	Karnataka	5.76520075339452	73.1766573766144
9	Tamil Nadu	5.46380198025467	76.1999009689482
10	Kerala	5.0808124360493	78.9482447280807

DASHBOARD 3: Gender-Based Mortality Analysis

Chart 1: Stacked Bar Chart

Problem Statement 1 : Compare the number of deaths for different causes across males and females to understand gender-based mortality differences?

Query:

```
SELECT
   Cause_of_Death,
   Gender,
   SUM(Total_Deaths) AS Total_Deaths
FROM finaldataset
GROUP BY Cause_of_Death, Gender
ORDER BY Cause_of_Death, Gender;
```

	Cause_of_Death	Gender	Total_Deaths
1	Accident	Female	367863976
2	Accident	Male	649158756
3	Cancer	Female	246551909
4	Cancer	Male	251639807
5	Cardio	Female	797126061
6	Cardio	Male	700372160
7	Diabetes	Female	143571321
8	Diabetes	Male	143257015
9	Infect	Female	596989650
10	Infect	Male	489614339
11	Resp	Female	461345289
12	Resp	Male	263030827
13	Stroke	Female	142672387
14	Stroke	Male	241447577

Chart 2: Scatter Chart

Problem Statement 2 : Examine how population size affects mortality rates across different genders?

Query:

```
SELECT
   Region,
   Gender,
   SUM(CAST(Population AS BIGINT)) AS TotalPopulation,
   AVG([Death_Rate]) AS AvgDeathRate,
```

```
    SUM(CAST([Total_Deaths] AS BIGINT)) AS TotalDeaths
FROM finaldataset
GROUP BY Region, Gender
ORDER BY Region, Gender;
```

	Region	Gender	TotalPopulation	AvgDeathRate	TotalDeaths
1	Andhra Pradesh	Female	25126272867	5.79145417840357	144360540
2	Andhra Pradesh	Male	24955577794	5.81124748598402	143850843
3	Bihar	Female	47784066269	9.32855731036823	441231039
4	Bihar	Male	47392505626	9.28100200454314	435508763
5	Gujarat	Female	39609030789	6.4856107555808	253700883
6	Gujarat	Male	42585913916	6.50195402203849	273615252
7	Karnataka	Female	35227679328	5.77139823542232	201465677
8	Karnataka	Male	31589971264	5.75828064006308	180305554
9	Kerala	Female	19091487239	5.07375757525666	96764793
10	Kerala	Male	19239615850	5.08776892228905	97770147
11	Maharashtra	Female	27497894418	5.94174273577963	161757706
12	Maharashtra	Male	28549289795	5.87663325900305	166196885
13	Rajasthan	Female	35715257691	7.01654008571609	247621526
14	Rajasthan	Male	37928407807	7.00662697307647	262856738

Chart 3: Line Chart

Problem Statement 3 : Analyze the trend of average death rates for males and females over time to identify long-term patterns?

Query:

```
SELECT
    Year,
    Gender,
    AVG(Death_Rate) AS Avg_Death_Rate
FROM finaldataset
GROUP BY Year, Gender
ORDER BY Year, Gender;
```

	Year	Gender	Avg_Death_Rate
1	2000	Female	7.85656346182336
2	2000	Male	7.82853035576427
3	2001	Female	7.56434250169573
4	2001	Male	7.42918493306749
5	2002	Female	7.39972694904731
6	2002	Male	7.36519480061221
7	2003	Female	7.23507740623073
8	2003	Male	7.16016284805944
9	2004	Female	7.04596089384843
10	2004	Male	6.99950769864596
11	2005	Female	6.82641509793839
12	2005	Male	6.93494117119733
13	2006	Female	6.63749999637845
14	2006	Male	6.77739001998454

Chart 4: Line Chart

Problem Statement 4 : Analyze which causes of death are more common among males and females?

Query:

```
SELECT
    Cause_of_Death,
    Gender,
    SUM(Total_Deaths) AS Total_Deaths
FROM finaldataset
WHERE Age_Range = '66+'
GROUP BY Cause_of_Death, Gender
ORDER BY Total_Deaths DESC;
```

	Cause_of_Death	Gender	Total_Deaths
1	Cardio	Male	76656773
2	Cardio	Female	71627638
3	Infect	Female	64307839
4	Accident	Male	60868737
5	Infect	Male	48841699
6	Resp	Female	42979690
7	Stroke	Male	34440542
8	Resp	Male	30718947
9	Accident	Female	28395446
10	Cancer	Female	22864498
11	Cancer	Male	21525961
12	Diabetes	Male	19066934
13	Stroke	Female	15436339
14	Diabetes	Female	10253172

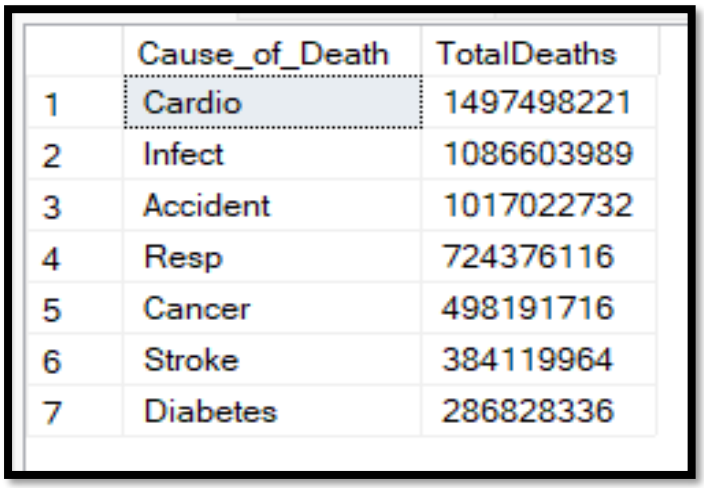
DASHBOARD 4: Cause-Specific Mortality Trends

Chart 1: Clustered Column Chart

Problem Statement 1 : Rank the leading causes of death in India to identify the most critical public health challenges?

Query:

```
SELECT
   [Cause_of_Death],
   SUM(CAST([Total_Deaths] AS BIGINT)) AS TotalDeaths
FROM finaldataset
GROUP BY [Cause_of_Death]
ORDER BY TotalDeaths DESC
```



	Cause_of_Death	TotalDeaths
1	Cardio	1497498221
2	Infect	1086603989
3	Accident	1017022732
4	Resp	724376116
5	Cancer	498191716
6	Stroke	384119964
7	Diabetes	286828336

Chart 2: 100% Stacked Column Chart

Problem Statement 2: Compare the percentage distribution of different causes of death across states to assess regional health issues?

Query:

```
SELECT
   Region,
   [Cause_of_Death],
   SUM(CAST([Total_Deaths] AS BIGINT)) * 100.0 /
     (SELECT SUM(CAST([Total_Deaths] AS BIGINT))
      FROM [finaldataset] AS X
      WHERE X.Region = M.Region) AS DeathPercentage
```

```
FROM [finaldataset] AS M
GROUP BY Region, [Cause_of_Death]
ORDER BY Region, Death Percentage DESC;
```

	Region	Cause_of_Death	DeathPercentage
1	Andhra Pradesh	Cardio	26.544837404982023
2	Andhra Pradesh	Accident	21.016789957945554
3	Andhra Pradesh	Infect	19.865632094066180
4	Andhra Pradesh	Resp	11.168859697675438
5	Andhra Pradesh	Cancer	8.918407639714910
6	Andhra Pradesh	Stroke	6.489409545631998
7	Andhra Pradesh	Diabetes	5.996063659983894
8	Bihar	Cardio	26.836397008926942
9	Bihar	Infect	20.127516350626454
10	Bihar	Accident	18.393368777387843
11	Bihar	Resp	15.272111143415386
12	Bihar	Cancer	8.435272794881051
13	Bihar	Stroke	6.869466044841431
14	Bihar	Diabetes	4.065867879920889

Chart 3: 100% Donut Chart

Problem Statement 3: Analyze how different causes of death impact males and females differently.?

Query:

```
SELECT
   Gender,
   [Cause_of_Death],
   SUM(CAST([Total_Deaths] AS BIGINT)) AS TotalDeaths
FROM finaldataset
GROUP BY Gender, [Cause_of_Death]
ORDER BY Gender, TotalDeaths DESC;
```

	Gender	Cause_of_Death	TotalDeaths
1	Female	Cardio	797126061
2	Female	Infect	596989650
3	Female	Resp	461345289
4	Female	Accident	367863976
5	Female	Cancer	246551909
6	Female	Diabetes	143571321
7	Female	Stroke	142672387
8	Male	Cardio	700372160
9	Male	Accident	649158756
10	Male	Infect	489614339
11	Male	Resp	263030827
12	Male	Cancer	251639807
13	Male	Stroke	241447577
14	Male	Diabetes	143257015

Chart 4: 100% Stacked Area Chart

Problem Statement 3: Identify how different causes of death affect various age groups over time?

Query:

```
SELECT
    [Age_Range],
    [Cause_of_Death],
    AVG([Death_Rate]) AS AvgDeathRate
FROM [finaldataset]
GROUP BY [Age_Range], [Cause_of_Death]
ORDER BY [Age_Range], AvgDeathRate DESC;
```

	Age_Range	Cause_of_Death	AvgDeathRate
1	0-18	Cardio	6.57125940627621
2	0-18	Infect	6.53497571852601
3	0-18	Accident	6.49645429965202
4	0-18	Cancer	6.47462427547212
5	0-18	Resp	6.47462184288923
6	0-18	Diabetes	6.41715789594148
7	0-18	Stroke	6.26088234957527
8	19-35	Cancer	6.57533334493637
9	19-35	Resp	6.55222223978194
10	19-35	Cardio	6.54827692398658
11	19-35	Infect	6.5386081242204
12	19-35	Diabetes	6.50568000793457
13	19-35	Stroke	6.4918918867369
14	19-35	Accident	6.46862745181148

DASHBOARD 5: Population & Mortality Insights

Chart 1: Scatter Chart

Problem Statement 1 : Explore the relationship between state population sizes and total mortality rates?

Query:

```
SELECT
    Region,
    SUM(CAST(Population AS BIGINT)) AS TotalPopulation,
    SUM(CAST([Total_Deaths] AS BIGINT)) AS TotalDeaths,
    AVG([Death_Rate]) AS AvgDeathRate
FROM finaldataset
GROUP BY Region
ORDER BY Region;
```

	Region	TotalPopulation	TotalDeaths	AvgDeathRate
1	Andhra Pradesh	50081850661	288211383	5.80130129939204
2	Bihar	95176571895	876739802	9.30494527294861
3	Gujarat	82194944705	527316135	6.49409949696479
4	Karnataka	66817650592	381771231	5.76520075339452
5	Kerala	38331103089	194534940	5.0808124360493
6	Maharashtra	56047184213	327954591	5.90862384800031
7	Rajasthan	73643665498	510478264	7.01143148779138
8	Tamil Nadu	67689451867	367602648	5.46380198025467
9	Uttar Pradesh	175345708174	1387869482	8.01780310467735
10	West Bengal	100154316642	632162598	6.37370071131491

Chart 2: Line Chart

Problem Statement 2 : Analyze how the population of each region has grown over time?

Query:

```
SELECT
    Year,
```

```
    Region,
    SUM(CAST(Population AS BIGINT)) AS TotalPopulation
FROM finaldataset
GROUP BY Year, Region
ORDER BY Year, Region;
```

	Year	Region	TotalPopulation
1	2000	Andhra Pradesh	2655000000
2	2000	Bihar	5520000000
3	2000	Gujarat	3360000000
4	2000	Karnataka	4180000000
5	2000	Kerala	1470000000
6	2000	Maharashtra	3600000000
7	2000	Rajasthan	4030000000
8	2000	Tamil Nadu	5160000000
9	2000	Uttar Pradesh	8700000000
10	2000	West Bengal	5760000000
11	2001	Andhra Pradesh	2829060000
12	2001	Bihar	4496800000
13	2001	Gujarat	4569600000
14	2001	Karnataka	4418864921

Chart 3: Stacked Area Chart

Problem Statement 3: Show the overall trend of total deaths across all states over time?

Query:

```
SELECT
    Year,
    Region,
    SUM(CAST([Total_Deaths] AS BIGINT)) AS TotalDeaths
FROM finaldataset
GROUP BY Year, Region
ORDER BY Year, Region;
```

	Year	Region	TotalDeaths
1	2000	Andhra Pradesh	18838318
2	2000	Bihar	59952665
3	2000	Gujarat	26810439
4	2000	Karnataka	29559543
5	2000	Kerala	8239251
6	2000	Maharashtra	26086467
7	2000	Rajasthan	34401217
8	2000	Tamil Nadu	33696907
9	2000	Uttar Pradesh	82770929
10	2000	West Bengal	44972659
11	2001	Andhra Pradesh	19677749
12	2001	Bihar	47488377
13	2001	Gujarat	35034614
14	2001	Karnataka	29799489

Chart 4: Field Map

Problem Statement 4: Map the relationship between population density and mortality rates across different states?

Query:

```
SELECT
    Region,
    SUM(CAST(Population AS BIGINT)) AS TotalPopulation,
    AVG([Death_Rate]) AS AvgDeathRate,
    AVG([Life_Expectancy]) AS AvgLifeExpectancy
FROM finaldataset
GROUP BY Region
ORDER BY Region;
```

	Region	TotalPopulation	AvgDeathRate	AvgLifeExpectancy
1	Andhra Pradesh	50081850661	5.80130129939204	72.3282282221186
2	Bihar	95176571895	9.30494527294861	64.2313432721949
3	Gujarat	82194944705	6.49409949696479	70.2597014925373
4	Karnataka	66817650592	5.76520075339452	73.1766573766144
5	Kerala	38331103089	5.0808124360493	78.9482447280807
6	Maharashtra	56047184213	5.90862384800031	72.2256880205098
7	Rajasthan	73643665498	7.01143148779138	68.2436606128279
8	Tamil Nadu	67689451867	5.46380198025467	76.1999009689482
9	Uttar Pradesh	175345708174	8.01780310467735	66.2302589772279
10	West Bengal	100154316642	6.37370071131491	69.1291202587594

DAX FUNCTIONS

1. Introduction:

DAX (Data Analysis Expressions) is a formula language used in Microsoft Power BI, Excel, and SQL Server Analysis Services (SSAS) to create custom calculations and aggregations. It enables users to analyze data efficiently, perform advanced calculations, and derive insights from datasets.

In the India Mortality Analytics and Life Expectancy Dashboard, DAX functions are used to calculate total deaths, average life expectancy, mortality trends, and other critical metrics that enhance the dashboard's functionality.

2. Key DAX Function Categories:

DAX functions are classified into different categories based on their usage:

Function Category	Purpose	Example Functions
Aggregation	Summarizes data values	SUM, AVERAGE, COUNT, MAX, MIN
Filter	Applies conditions to analyze specific data	FILTER, ALL, REMOVEFILTERS
Time Intelligence	Works with date/time data for trend analysis	TOTALYTD, PREVIOUSYEAR, DATEADD
Logical	Creates conditional statements	IF, SWITCH, AND, OR
Mathematical	Performs mathematical calculations	ROUND, ABS, EXP, DIVIDE
Text	Manipulates text values	CONCATENATE, LEFT, RIGHT, SEARCH

3. Used DAX Functions in the Project

3.1. Aggregation Functions

These functions help summarize death rates, total deaths, life expectancy, and population.

1.Total Deaths Calculation

DAX function:

```
Total_Deaths = SUM(finaldataset[Total_ Deaths])
```

Purpose: Calculates the total number of deaths in all regions.

2.Average Death Rate Calculation

DAX function:

```
Avg_Death_Rate = AVERAGE(finaldataset[Death_Rate])
```

Purpose: Computes the national average death rate.

3.Highest Death Rate State

DAX function:

```
Highest_Death_Rate_State = MAX(finaldataset[Death_Rate])
```

Purpose: Identifies the state with the highest mortality rate.

3.2. Filter Functions

Filter functions help refine the dataset for state-wise, gender-based, or time-based mortality trends.

1.Total Deaths in a Specific State

DAX function:

```
Deaths_in_State = CALCULATE(
    SUM(finaldataset[Total_Deaths]),
    finaldataset[Region] = "Maharashtra"
)
```

Purpose: Calculates total deaths in Maharashtra.

2.Deaths for a Specific Cause (e.g., Cardiovascular)

DAX function:

```
Deaths_Cardiovascular = CALCULATE(
    SUM(finaldataset[Total_Deaths]),
    finaldataset[Cause_of_Death] = "Cardiovascular"
)
```

Purpose: Computes total deaths due to cardiovascular diseases.

3..Deaths by Gender (Male/Female)

DAX function:

```
Deaths_Male = CALCULATE(
    SUM(finaldataset[Total_Deaths]),
    finaldataset[Gender] = "Male"
)
```

Purpose: Finds total male mortality.

3.3. Time Intelligence Functions

These functions analyze mortality trends over time and compare yearly data.

1.Total Deaths Year-to-Date (YTD)

DAX function:

```
Deaths_YTD = TOTALYTD(
    SUM(finaldataset[Total_Deaths]),
    finaldataset[Year]
```

```
)
```

Purpose: Calculates total deaths from the start of the year to the current date.

2.Previous Year's Death Rate

DAX function:

```
Prev_Year_Death_Rate = CALCULATE(
    AVERAGE(finaldataset[Death_Rate]),

 PREVIOUSYEAR(finaldataset[Year])
)
```

Purpose: Compares current death rate with the previous year.

3.Rolling 3-Year Average Life Expectancy

DAX function:

```
Rolling_Life_Expectancy = AVERAGEX(
    DATESINPERIOD(finaldataset[Year], LASTDATE(finaldataset[Year]), -3, YEAR),
    finaldataset[Life_Expectancy]
)
```

Purpose: Smooths fluctuations in life expectancy trends.

3.4. Logical Functions

These functions allow the creation of custom conditions and classifications.

1.Categorizing Death Rate (High/Medium/Low)

DAX function:

```
Death_Rate_Category =
SWITCH(
```

```
    TRUE(),
    finaldataset[Death_Rate] > 10, "High",
    finaldataset[Death_Rate] BETWEEN 7 AND 10, "Medium",
    "Low"
)
```

Purpose: Assigns a risk category based on mortality rates.

2. If a Region Has High Life Expectancy (>70 Years)

DAX function:

```
High_Life_Expectancy = IF(
    AVERAGE(finaldataset[Life_Expectancy]) > 70, "Yes", "No"
)
```

Purpose: Classifies states with high life expectancy.

3.5. Mathematical & Statistical Functions

These functions help in advanced calculations such as ratios and percentages.

1.Death Rate as Percentage of Population

DAX function:

```
Death_Rate_Percentage =
DIVIDE(
    SUM(finaldataset[Total_Deaths]),
    SUM(finaldataset[Population])
) * 100
```

Purpose: Computes death rate as a percentage of total population.

2.Mortality Index (Standardized Death Rate)

DAX function:

```
Mortality_Index =
```

```
ROUND(
    SUM(finaldataset[Total_Deaths]) / SUM(finaldataset[Population]) * 1000, 2
)
```

Purpose: Standardizes death rates for better comparison across regions.

3.Top 5 States with Highest Deaths

DAX function:

```
Top_5_States =
TOPN(5,
    SUMMARIZE(finaldataset,finaldataset[Region],"TotalDeaths",
SUM(finaldataset[Total_Deaths])),
    [Total Deaths], DESC
)
```

Purpose: Returns the top 5 states with the highest mortality.

4. Benefits of Using DAX in Power BI

Enhanced Data Analysis: Enables custom metrics beyond built-in Power BI functions.

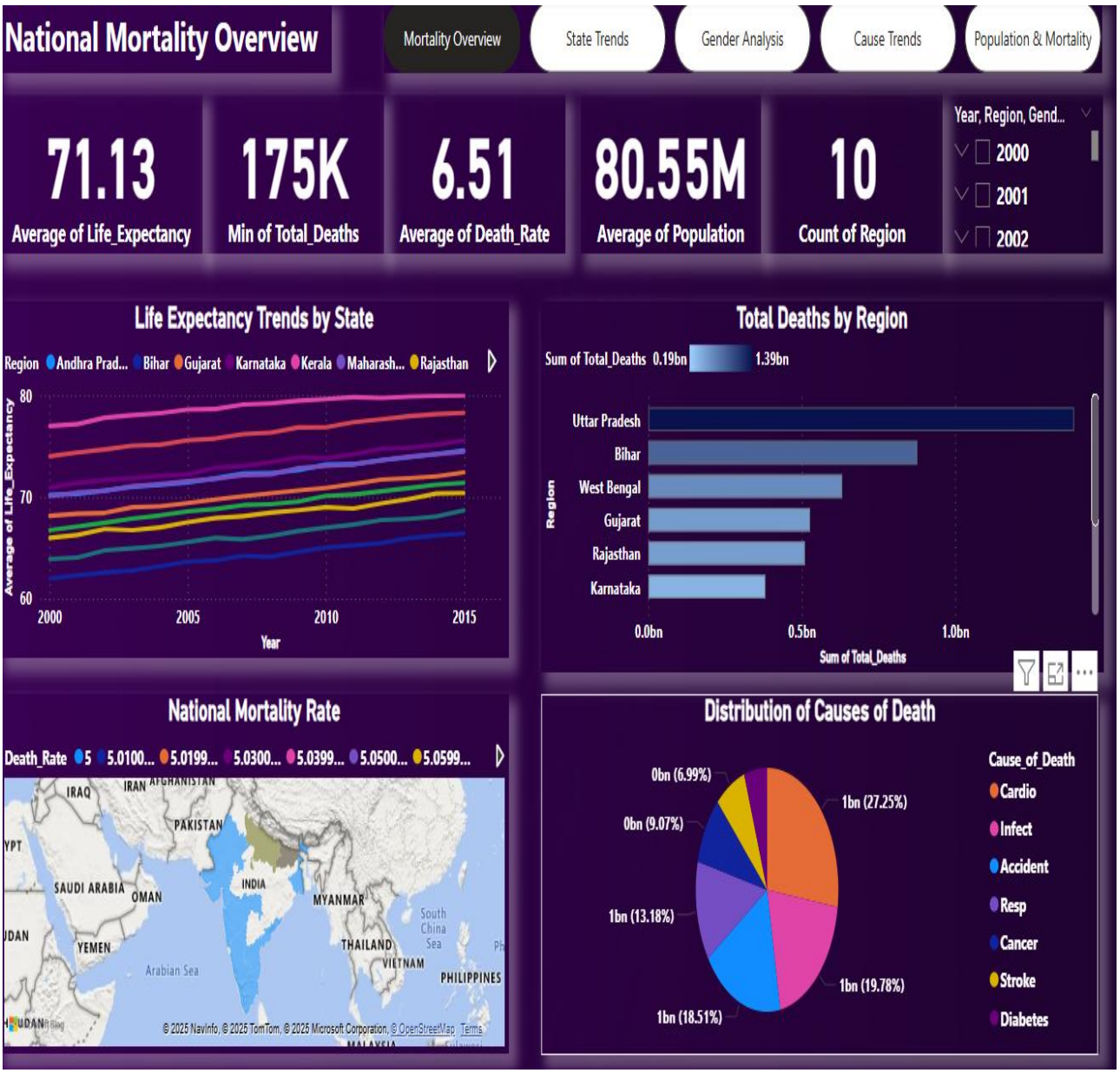
Improved Performance: Optimized data aggregation and calculations.

Flexibility in Reporting: Allows customized insights for different regions, years, and causes of death.

Seamless Integration: Works efficiently with SQL and Power BI Data Model

OUTPUT

DASHBOARD 1: National Mortality Overview



Overview of the "National Mortality Overview" Dashboard:

This Power BI dashboard provides a detailed analysis of mortality trends across different regions in India. It presents key insights into life expectancy, death rates, population trends, and causes of death through various visualizations.

Key Insights & Metrics:

1. **Top Summary Metrics (KPIs):**

71.13 → Average Life Expectancy

175K → Minimum Total Deaths

6.51 → Average Death Rate

80.55M → Average Population

10 → Count of Regions

2. **Life Expectancy Trends by State:**

   A line chart showing the average life expectancy trends from 2000 to 2015 for different states.

   Helps identify states with higher or lower improvements in life expectancy.

3. **Total Deaths by Region:**

   A bar chart displaying total deaths across different regions.

   Uttar Pradesh shows the highest mortality, followed by Bihar and West Bengal.

4. **National Mortality Rate (Geographical Map):**

   A heatmap displaying the death rate distribution across India.

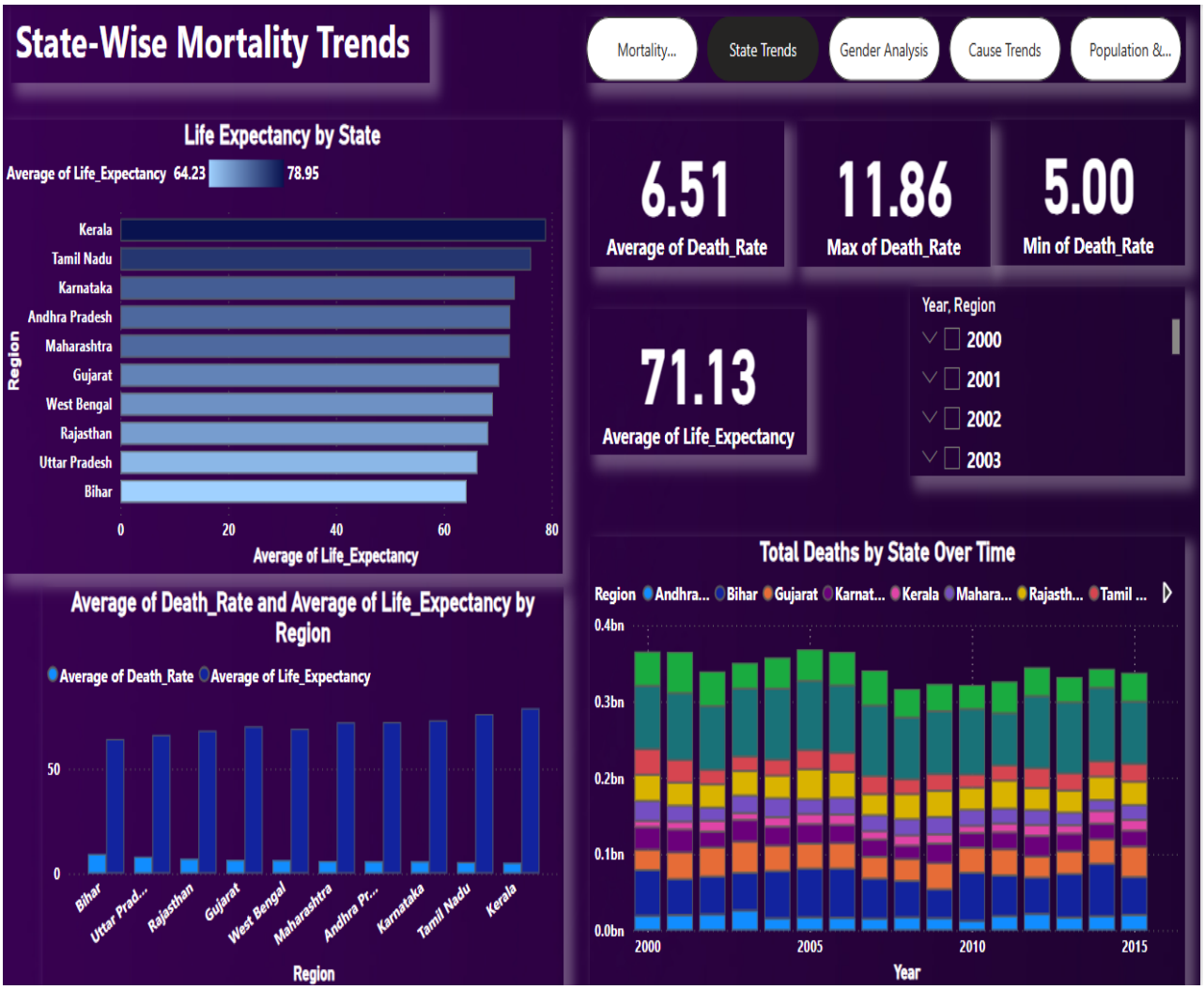
   Darker shades indicate higher mortality rates in specific regions.

5. Distribution of Causes of Death:

   A pie chart showing the proportion of deaths by cause.

   Cardiovascular diseases (Cardio), Infections, and Accidents appear as major causes.

DASHBOARD 2: State-Wise Mortality Trends



Overview of the "State-Wise Mortality Trends" Dashboard:

This Power BI dashboard provides insights into mortality trends at the state level, focusing on life expectancy, death rates, and total deaths over time across various regions in India.

Key Insights & Metrics:

1. **Life Expectancy by State (Bar Chart):**

   Kerala has the highest life expectancy (~79 years), while Bihar has the lowest (~64 years).

   Shows variations in healthcare, living conditions, and economic development across states.

2. **Top Summary Metrics (KPIs):**

   6.51 → Average Death Rate

11.86 → Maximum Death Rate

5.00 → Minimum Death Rate

71.13 → Average Life Expectancy

3. **Average of Death Rate vs Life Expectancy (Bar Chart):**

   Compares death rates and life expectancy across states.

   Higher life expectancy states (e.g., Kerala, Tamil Nadu) generally have lower death rates**.**

4. **Total Deaths by State Over Time (Stacked Bar Chart):**

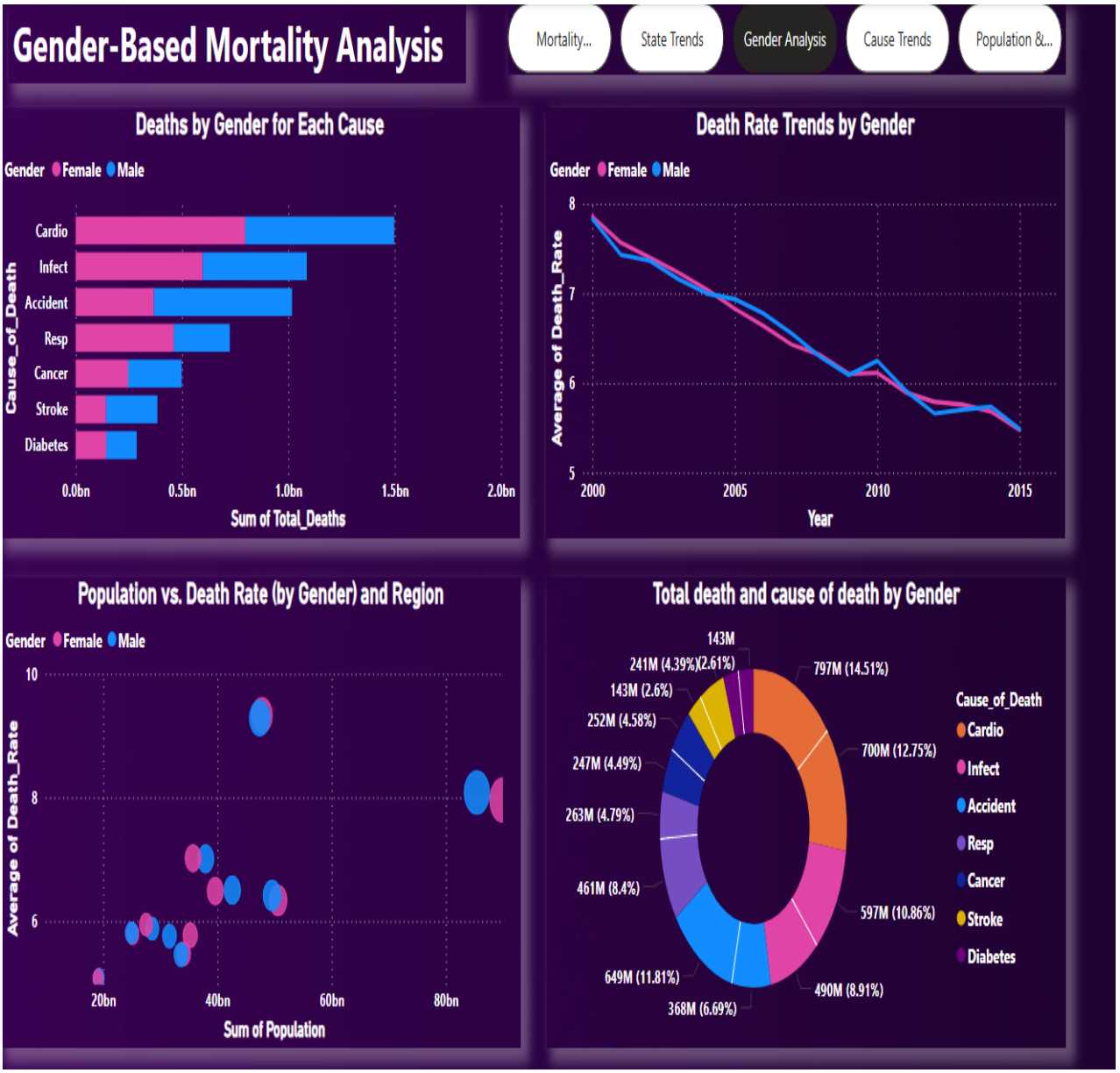
   Shows the distribution of total deaths from 2000 to 2015 across multiple states.

   Uttar Pradesh, Bihar, and Maharashtra seem to contribute significantly to total deaths.

5. **Filter Panel (Top Right):**

   Allows selection of specific years (2000–2003) to analyze trends over time.

DASHBOARD 3: Gender-Based Mortality Analysis



Overview of the "Gender-Based Mortality Analysis" Dashboard:

This dashboard provides a gender-based analysis of mortality trends, focusing on the causes of death, death rate trends over time, and the relationship between population and mortality.

Key Insights & Metrics:

1. **Deaths by Gender for Each Cause (Bar Chart - Top Left):**

Cardiovascular diseases (Cardio) have the highest mortality rate, followed by infectious diseases, accidents, and respiratory issues.

Males generally have higher death counts across most causes, except for diabetes and stroke, where female deaths are more significant.

2. Death Rate Trends by Gender (Line Chart - Top Right):

Death rates for both genders have shown a declining trend from 2000 to 2015.

Initially, female death rates were higher, but over time, both genders exhibit similar trends.

3.Population vs. Death Rate by Gender & Region (Bubble Chart - Bottom Left):

Shows the relationship between population size and death rate for males and females in different regions.

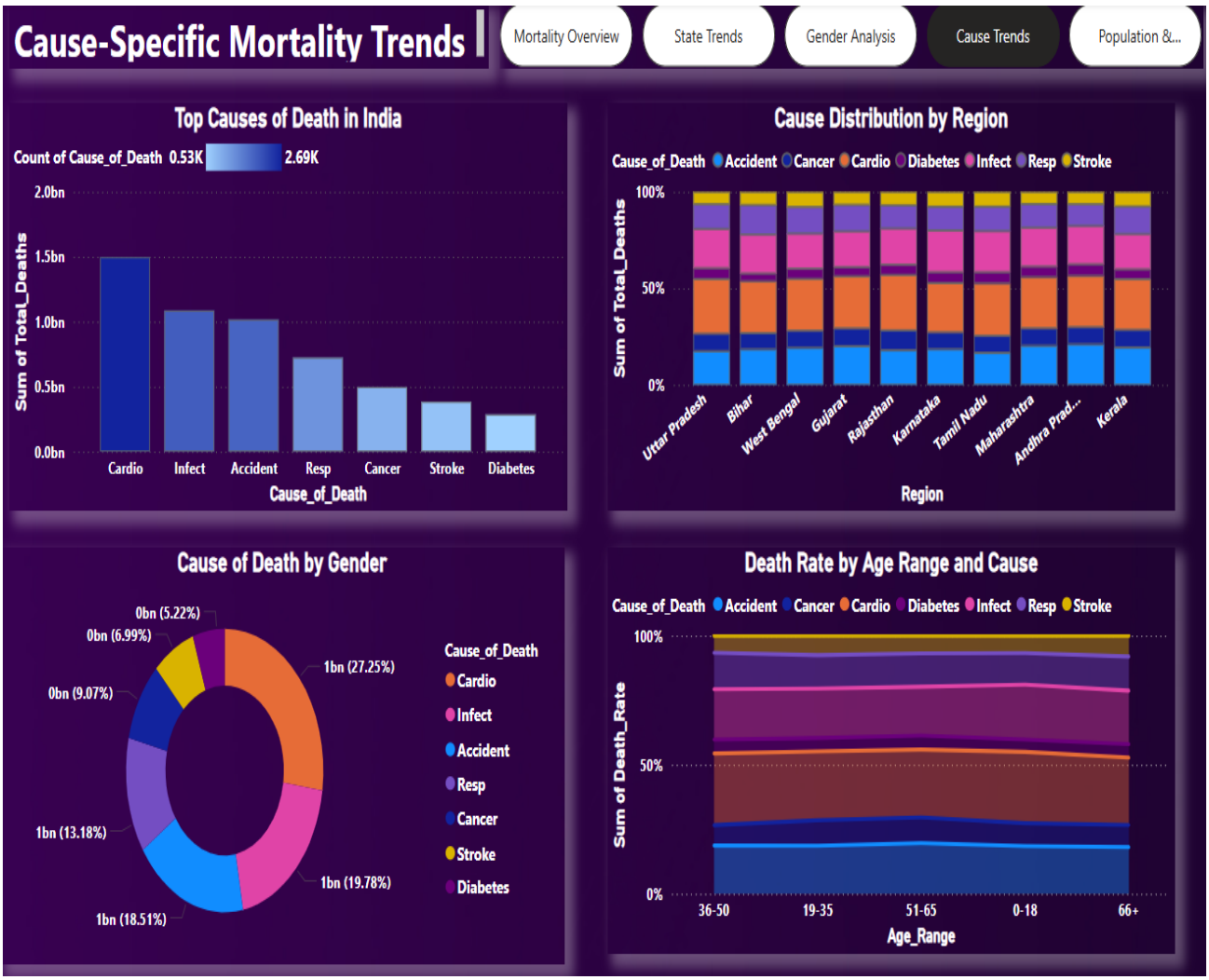
Larger populations tend to have lower death rates, while smaller regions may experience higher mortality rates.

4.Total Death & Cause of Death by Gender (Donut Chart - Bottom Right):

Cardiovascular diseases (14.51%) are the leading cause of death, followed by infections, cancer, and accidents.

Other significant causes include stroke and diabetes, which affect both genders but show slight variations.

DASHBOARD 4: Cause-Specific Mortality Trends



Overview of the "Cause-Specific Mortality Trends" Dashboard:

This dashboard provides a detailed analysis of mortality causes in India, showing how different causes contribute to deaths across regions, genders, and age groups.

Key Insights & Metrics:

1. **Top Causes of Death in India (Bar Chart - Top Left)**

   Cardiovascular diseases (Cardio) are the leading cause of death, followed by infectious diseases, accidents, and respiratory issues.

   Cancer, stroke, and diabetes contribute to a smaller but still significant portion of total deaths.

2. **Cause Distribution by Region (Stacked Bar Chart - Top Right)**

   Different states have varying proportions of deaths by cause.

Cardiovascular diseases (blue) and infections (orange) are dominant across most regions.

Accidents, cancer, and respiratory diseases have varying impacts depending on the state.

3. Cause of Death by Gender (Donut Chart - Bottom Left)

Cardiovascular diseases (27.25%) are the primary cause of death for both genders.

Other leading causes include infections (19.78%), accidents (18.51%), and respiratory diseases (13.18%).

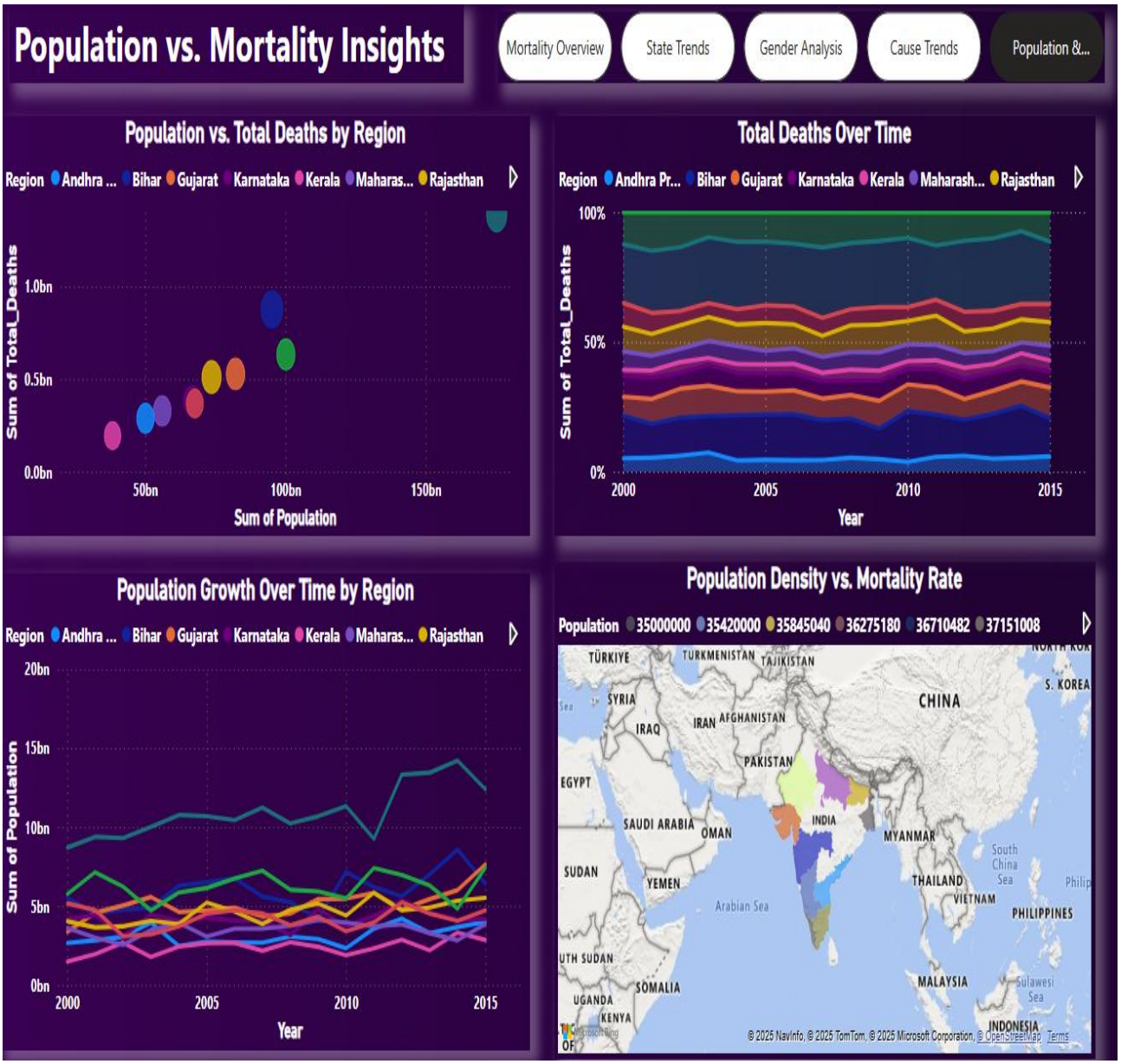
Cancer, stroke, and diabetes have lower but still notable shares of total deaths.

4. Death Rate by Age Range & Cause (Stacked Bar Chart - Bottom Right)

Older age groups (66+) have higher mortality rates from cardiovascular diseases and respiratory illnesses.

Younger age groups (0-18, 19-35) see higher deaths due to accidents and infections.

DASHBOARD 5: Population vs. Mortality Insights



Overview of the "Population vs. Mortality Insights" Dashboard:

This dashboard explores the relationship between population and mortality trends across different regions in India over time.

Key Insights & Metrics:

Population vs. Total Deaths by Region (Bubble Chart - Top Left)

Shows the correlation between total population and total deaths across various states.

Larger states with higher populations tend to have more deaths, but some states have

relatively lower death rates despite high populations.

1. **Total Deaths Over Time (Stacked Area Chart - Top Right)**

   Displays how total deaths have changed from 2000 to 2015 in different regions.

   Some states have seen a rise in mortality trends, while others remain stable.

2. **Population Growth Over Time by Region (Line Chart - Bottom Left)**

   Tracks the growth of population across different states over time.

   Some states have sharp population increases, while others show slow or stable growth.

3. **Population Density vs. Mortality Rate (Map Visualization - Bottom Right)**

   A geographical representation of how population density correlates with mortality rates across India.

   Higher population density areas may not always have the highest mortality rates.

CONCLUSION

The India Mortality Analytics and Life Expectancy Dashboard is a comprehensive data-driven project designed to analyze mortality trends, death rates, life expectancy, and causes of death across India from 2000 to 2015. The project leverages Microsoft Power BI, SQL Server, and DAX functions to provide interactive visualizations and data insights that help stakeholders, including government policymakers, healthcare professionals, and researchers, make informed decisions regarding public health and mortality trends.

Key Achievements of the Project:

1. In-Depth Mortality Analysis

One of the most significant achievements of this project is its ability to visualize mortality trends across different dimensions such as states, gender, and cause-specific mortality. The dashboard allows users to:

Identify high-risk regions with high death rates.

Track mortality trends over time to observe improvements or declines in public health.

Analyze life expectancy trends across states and genders, providing insights into the health conditions of different demographics.

The project successfully presents a data-driven view of mortality patterns, helping to identify areas where healthcare interventions and policies can be most effective.

2. Data-Driven Decision-Making for Policy & Healthcare

By providing clear and actionable insights, this project serves as a valuable resource for policymakers, public health officials, and researchers. The analysis helps in:

Understanding how different factors impact mortality, such as gender, region, and specific causes of death.

Developing healthcare strategies to address critical causes of death, such as cardiovascular diseases, respiratory illnesses, and accidents.

Allocating resources efficiently to improve life expectancy and reduce mortality rates.

This project highlights the power of data analytics in healthcare, allowing for evidence-based policymaking rather than assumptions or incomplete data.

3. Advanced Data Processing & Visualization

The project efficiently integrates large-scale mortality data using Microsoft SQL Server, ensuring data integrity and accessibility.

Power BI's data modeling and visualization capabilities were used to create interactive and dynamic reports.

DAX (Data Analysis Expressions) functions were applied to perform advanced calculations such as rolling averages, total deaths over time, and life expectancy comparisons.

Power Query transformations helped in data cleaning, validation, and structuring, ensuring accurate insights.

Through these tools and techniques, the project successfully creates a user-friendly dashboard that presents complex mortality data in an easy-to-understand manner.

Challenges & Solutions:

1. Data Quality Issues

Challenge: Missing values and inconsistencies in the dataset.

Solution: Used data validation techniques and applied imputation methods to estimate missing values accurately.

2.Large Dataset Performance Issues

Challenge: Processing large datasets in Power BI caused performance lag.

Solution: Applied optimized DAX formulas, data indexing, and aggregation techniques to enhance efficiency.

3.Ensuring Meaningful Visualizations

Challenge: Some visualizations displayed similar trends, making them redundant.

Solution: Improved visual diversity by incorporating filters, drill-through features, and comparative analyses.

Through these strategic solutions, the project was able to maintain accuracy, efficiency, and relevance in data analysis.

Future Scope & Scalability

This project provides a solid foundation for mortality and life expectancy analysis in India. However, it also presents opportunities for future expansion, including:

1Extending the Time Period (2016–2025):

By incorporating recent mortality data, we can analyze post-2015 trends, including the impact of pandemics or new health policies.

2.ncorporating Socioeconomic Factors:

Future versions can integrate data on literacy rates, healthcare accessibility, pollution levels, and economic conditions to find broader correlations.

3.Machine Learning for Predictive Analysis:

Using AI and predictive modeling, the system could forecast future mortality rates and life expectancy trends based on historical patterns.

4.Cloud Integration for Real-Time Insights:

Hosting the dashboard on Power BI Service with Azure SQL Database can enable real-time mortality tracking and dynamic updates.

These enhancements will make the project even more powerful and insightful, contributing to future healthcare advancements and data-driven policy-making.

Final Verdict:

The project is a major success in providing data-driven insights into mortality trends in India. By utilizing SQL for data processing, Power BI for visualization, and DAX for advanced calculations, the project ensures:

I. Accurate mortality analysis at the national and state levels.

II. Interactive and insightful dashboards that allow for deep data exploration.

III. Actionable insights for policymakers, researchers, and healthcare professionals.

The dashboard has significant potential for future scalability, making it a valuable tool for monitoring, analyzing, and improving public health outcomes in India.

Final Thought:

The India Mortality Analytics and Life Expectancy Dashboard is not just a data visualization tool; it is a step towards improving healthcare decision-making.

BIBLIOGRAPHY

1. Data Sources:

1.Government & Public Health Reports:

Registrar General of India, Census Reports (2001, 2011) – Mortality and demographic data

National Health Profile (NHP), Ministry of Health and Family Welfare (MoHFW)

Sample Registration System (SRS) Reports, Government of India

World Health Organization (WHO) Mortality Database.

2.Statistical Databases & Research Papers:

World Bank Open Data – Life Expectancy & Mortality Trends

United Nations (UN) Population Division Reports

Indian Council of Medical Research (ICMR) – Cause-Specific Mortality Trends

Lancet Global Health Studies on Mortality Trends in India.

2. Technical References:

1.Microsoft Power BI Documentation & Tutorials:

Microsoft Power BI Official Documentation: docs.microsoft.com/power-bi

Power Query M Language Reference: docs.microsoft.com/powerquery-m

DAX (Data Analysis Expressions) Guide: dax.guide

2.SQL Server & Database Management:

SQL Server 2019 Official Documentation: docs.microsoft.com/sql

SQL Server Management Studio (SSMS) User Guide

Best Practices for Indexing & Query Optimization in SQL Server

3.Data Science & Visualization Techniques:

Python Data Cleaning Techniques – Pandas & NumPy Documentation

Kaggle Datasets on Indian Mortality Trends

3. Books & Research Papers

Books on Data Analytics & Visualization:

"Data Science for Business" – Foster Provost & Tom Fawcett

"Practical Power BI" – Daniil Maslyuk

"SQL for Data Analytics" – Upom Malik, Matt Goldwasser, Benjamin Johnston

Healthcare & Public Health Research Papers:

"Epidemiological Transition in India" – Research paper on changing mortality patterns

"State-Level Disease Burden Initiative" – Journal published by the Indian Public Health Foundation.

4. Online Learning Resources & Tutorials:

YouTube & MOOCs (Massive Open Online Courses):

Power BI Training by Enterprise DNA

Microsoft Learn Power BI Courses

SQL & Power BI Online Courses – Coursera, Udemy, and EdX

Forums & Blogs:

Power BI Community: community.powerbi.com

SQL Server Central Forum: sqlservercentral.com

Stack Overflow for DAX & SQL Queries.

5. Other References & Acknowledgments:

Special thanks to academic mentors, faculty members, and research supervisors who provided valuable insights during project development.

Data insights were cross-verified using multiple sources to ensure accuracy and reliability.

WEB REFERENCES

1. Official Government & Public Health Portals:

Ministry of Health and Family Welfare (MoHFW), India, https://www.mohfw.gov.in

Provides official mortality and life expectancy reports.

Registrar General & Census Commissioner of India, https://censusindia.gov.in

Source for demographic and mortality statistics.

Sample Registration System (SRS) Reports, https://censusindia.gov.in/Vital_Statistics/SRS

Official death rate and life expectancy statistics.

World Health Organization (WHO) Mortality Database

https://www.who.int/data/gho/data/themes/mortality

Global insights on cause-specific mortality.

National Health Profile (NHP) – India

https://nhp.gov.in

National reports on public health and mortality trends.

World Bank Open Data – Life Expectancy & Mortality

https://data.worldbank.org/indicator/SP.DYN.LE00.IN

Reliable source for historical life expectancy trends.

2. Technical Documentation & Learning Resources

Microsoft Power BI Official Documentation

https://learn.microsoft.com/en-us/power-bi/

In-depth Power BI guides, tutorials, and best practices.

Power Query M Language Reference

https://learn.microsoft.com/en-us/powerquery-m/

Used for data transformations and cleaning.

DAX (Data Analysis Expressions) Guide

https://dax.guide

Official DAX function reference for Power BI.

Microsoft SQL Server Documentation

https://learn.microsoft.com/en-us/sql/

Comprehensive SQL query writing and optimization techniques.

Power BI Community & Forums

https://community.powerbi.com

Power BI user discussions, troubleshooting, and tutorials.

SQL Server Central

https://www.sqlservercentral.com

Expert advice and best practices for SQL database management.

3. Academic Research & Data Science Resources

Google Scholar – Research Papers on Indian Mortality Trends

https://scholar.google.com

Referenced academic research on life expectancy and mortality analytics.

Kaggle – Open Datasets on Indian Mortality Trends

https://www.kaggle.com/datasets

Used for testing and validating dataset accuracy.

Stack Overflow – SQL & Power BI Discussions

https://stackoverflow.com

Resolved technical issues related to Power BI and SQL queries.

DataCamp – Power BI & SQL Learning Resources

https://www.datacamp.com

Helped in learning advanced Power BI visualization techniques.

Coursera – Data Visualization & Analytics Courses

https://www.coursera.org

Used for self-paced learning on Power BI, SQL, and data analytics.

4. Other Online Tools & References

Python Pandas & NumPy Documentation

https://pandas.pydata.org/docs/

Used for data preprocessing before importing into Power BI.

Seaborn & Matplotlib Documentation

https://seaborn.pydata.org

Used for data validation & exploratory analysis.

India Stat – Socioeconomic Data

https://www.indiastat.com

Source for state-wise demographic insights.

Our World in Data – Global Mortality Statistics

https://ourworldindata.org

Provided global mortality trends for comparative analysis.

5. Acknowledgments & Contributions:

Faculty & Mentors: Provided valuable guidance and feedback during project development.

Power BI & SQL Experts: Helped refine queries and optimize dashboard performance.

Academic Journals & Research Papers: Contributed insights into mortality trends, healthcare, and life expectancy analysis.