INTERNSHIP REPORT

HEALTH CARE LANDSCAPE ANALYSIS FOR NORTH INDIA

D.Yeshwanth
Roll No. 23955A6720

An INTERNSHIP REPORT

On

HEALTH CARE LANDSCAPE ANALYSIS FOR NORTH INDIA

Bachelor of Technology
In

Computer Science and Engineering (DATA SCIENCE)

By

Dandu Yeshwanth

23955A6720

Department of CSE (Data Science)

INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal, Hyderabad – 500043, Telangana

June 2024

DECLARATION

I certify that

a. The work contained in this report is original and has been done by me under the guidance of my supervisor (s).

b. The work has not been submitted to any other Institute for any degree or diploma.

c. I have followed the guidelines provided by the Institute for preparing the report.

d. I have conformed to the norms and guidelines given in the Code of Conduct of the Institute.

e. Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the report and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

D. Yashwanth

Signature of the Student

Roll No. 23955A6720

Place : Hyderabad

Date: 29/06/2024

CERTIFICATE

MAKRI
SOLUTIONS (OPC) PVT. LTD.

CIN: U72900TG20200PC144816

Date: 01-06-2024

To Whom It May Concern

Sub: Letter of Internship Acknowledgement

This is to certify that **Dandu Yeshwanth** bearing Roll Number **23955A6720** has completed fieldwork in the TechnologyTrack of **Python & Power BI** with the project Titled **Healthcare Indicators and Demographic Data Analysis** from 13-May-2024 to 01-June-2024.

During this period, the student demonstrated dedication to their project work. They participated actively in the tasks, collaborated well with team members, and understood the subject well. Their ability to apply theoretical knowledge to practical scenarios wasnotable.

We wish them all the best in their future endeavors and look forward to their continued success.

With Best Regards,

D. Murali Madhavi.

Mrs. Murali Madhavi Devineni
Director

MAKRI SOLUTIONS (OPC) PVT. LTD. Hyderabad

APPROVAL SHEET

This INTERNSHIP REPORT entitled on **Healthcare Landscape Analysis of North India** by **DANDU YESHWANTH** is approved for the award of the Degree Bachelor of Technology in **Computer Science Engineering in Data Science.**

Class Incharge

Head of the Department

Head of the Department
Data Science
INSTITUTE OF AERONAUTICAL ENGINEERING
Dundigal, Hyderabad - 500 043

Date: 29/06/2024

Place: Dundigal

ABSTRACT

This internship report presents an in-depth analysis of data collected through the India Annual Health Survey (AHS) 2012-13, focusing on the Empowered Action Group (EAG) states of Uttarakhand, Rajasthan, Uttar Pradesh, Bihar, Jharkhand, Odisha, Chhattisgarh, Madhya Pradesh, and Assam. These states, comprising nearly half of India's population, are pivotal due to their high fertility and mortality rates, contributing significantly to national vital statistics. The survey encompassed a representative sample of approximately 21 million individuals and 4.32 million households, distributed across urban and rural areas within these states. Its primary objective was to compile a comprehensive, reliable dataset featuring critical indicators such as Infant Mortality Rate (IMR), Maternal Mortality Ratio (MMR), and Total Fertility Rate (TFR), alongside relevant covariates at the district level. This dataset facilitates annual monitoring and mapping of health indicators, enabling a nuanced understanding of factors impacting reproductive and child health. Through rigorous data analysis techniques, this report explores trends, correlations, and disparities revealed by the AHS data.

ACKNOWLEDGEMENT

I extend my heartfelt gratitude to Makri Solutions (OPC) Private Limited, Hyderabad - 500038, for providing me with the opportunity to undertake this internship project. I am deeply thankful to Dr. D V Ramana, Project Manager, for his invaluable guidance, mentorship, and unwavering support throughout the internship. His expertise and insights have been instrumental in shaping my understanding and skills in Healthcare Landscape Analysis of North India. I am also grateful to the entire team at Makri Solutions for their encouragement and assistance, which significantly contributed to the success of this project. This internship experience has been invaluable in my professional development, and I am privileged to have been part of Makri Solutions (OPC) Private Limited during this transformative journey.

CONTENTS

‘

CHAPTER I - INTRODUCTION

CHAPTER 1: Introduction

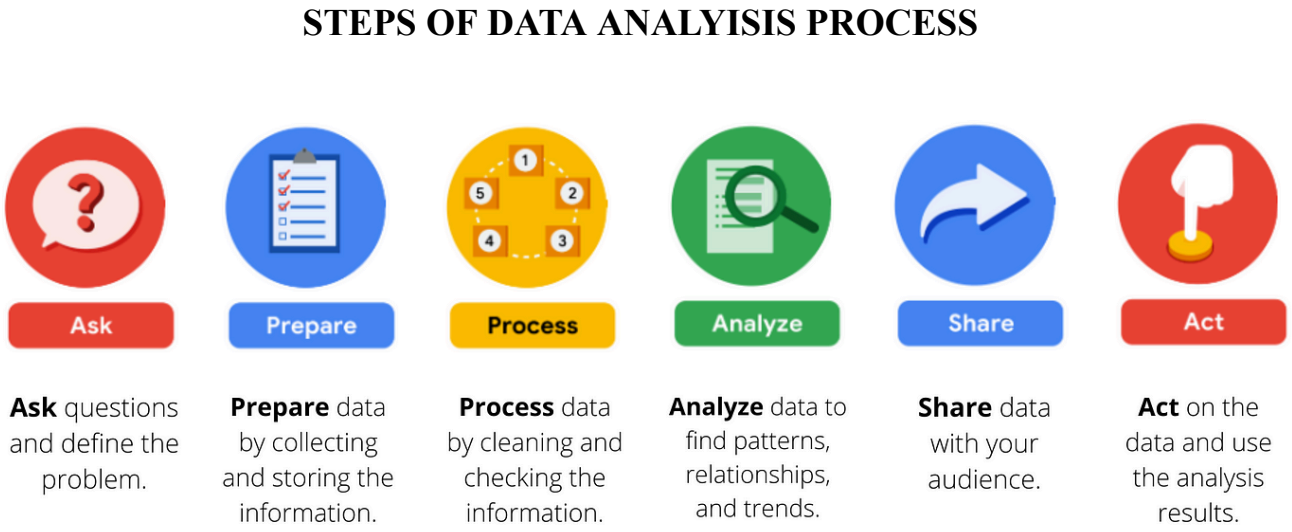
1.1 Overview of Visualization and Data Analytics

In the contemporary digital age, the sheer volume and complexity of data available to businesses, researchers, and policymakers have exponentially increased. This surge in data, commonly referred to as "big data," necessitates sophisticated tools and techniques for analysis and visualization. Data analytics and visualization are critical components that empower organizations to extract meaningful insights from vast datasets, facilitating informed decision-making and strategic planning

Data Analytics: Unlocking Hidden Insights

Data analytics involves the systematic examination of large datasets to identify hidden patterns, correlations, and trends. This process is fundamental in uncovering actionable insights that drive decision-making across various domains. By analyzing data, organizations can:

- **Identify Trends:** Detect emerging patterns that may signal market shifts or operational changes.
- **Enhance Decision-Making:** Provide evidence-based insights that support strategic decisions, minimizing reliance on intuition.
- **Optimize Operations:** Highlight inefficiencies and areas for improvement, fostering operational excellence.
- **Predict Future Outcomes:** Utilize historical data to forecast future trends, aiding in proactive planning.
- **Communicate Insights:** Present findings in a clear and compelling manner, enhancing communication with stakeholders.



Data Visualization: Bridging Understanding with Clarity

Definition

Data visualization refers to the graphical representation of information and data through visual elements such as charts, graphs, maps, and infographics. It transforms complex datasets into accessible visual formats, enabling users to quickly grasp and interpret the underlying data.

Advantages

- **Simplifies Complex Data:** Converts large and complex datasets into more digestible visual forms.
- **Reveals Hidden Patterns:** Helps identify trends, outliers, and patterns not immediately apparent in raw data.
- **Enhances Memory Retention:** Visual representations are often more memorable than textual data, aiding in better recall.
- **Engages Audience:** Interactive visuals make data exploration more engaging and insightful.

Disadvantages

- **Misleading Representations:** Poorly designed visuals can distort data, leading to incorrect interpretations.
- **Over-Simplification:** Nuances of complex data might be lost in overly simplified visuals.
- **Resource Intensive:** Creating effective visualizations can be time-consuming and requires specialized skills and tools.

1.2 Features and Capabilities of Power BI and Python

Power BI

Features and Capabilities

- **User-Friendly Interface:** Power BI offers an intuitive drag-and-drop interface, accessible to users of varying technical expertise.
- **Integration with Multiple Data Sources:** It connects to a diverse range of data sources, including Excel, SQL Server, and cloud services.
- **Real-Time Data Monitoring:** Supports real-time data analytics and monitoring, providing up-to-date insights.
- **Interactive Dashboards:** Facilitates the creation of interactive reports and dashboards that can be shared and collaborated on across an organization.

Python

Features and Capabilities

- **Versatility and Flexibility:** Python is a highly versatile language used for data manipulation, statistical analysis, and machine learning.
- **Rich Ecosystem of Libraries:** Libraries such as Pandas, NumPy, Matplotlib, Seaborn, and Plotly provide extensive tools for data analysis and visualization.
- **Scalability:** Capable of handling large datasets and complex calculations efficiently.
- **Customizability:** Offers extensive customization options for bespoke visualizations and analytical models.
- **Integration with Other Technologies:** Easily integrates with various data sources, databases, and web applications.

1.3 Comparison of Power BI and Python

Power BI Visualization

- **Ease of Use:** Features a drag-and-drop interface and pre-built templates, making it user-friendly, especially for non-technical users.
- **Interactivity:** Excels in creating interactive and shareable dashboards with real-time data updates.
- **Business Intelligence Focus:** Designed for business users, seamlessly integrating with Microsoft Office Suite and other enterprise tools.
- **Customization:** Limited compared to Python; users rely on built-in functionalities.
- **Performance:** Handles moderate-sized datasets efficiently but may struggle with extremely large datasets.

Python Visualization

- **Flexibility and Control:** Provides granular control over every aspect of visualization, allowing for highly customized and intricate visualizations.
- **Advanced Analytics:** Integrates seamlessly with advanced analytical and machine learning frameworks for deeper data analysis.
- **Library Variety:** Multiple libraries (e.g., Matplotlib, Seaborn, Plotly, Bokeh) cater to different visualization needs and preferences.
- **Scalability:** Better suited for large-scale data analysis and visualization projects.
- **Technical Skill Requirement:** Requires programming knowledge, making it less accessible to non-technical users.

CHAPTER II – INSTALLATION PROCESS

CHAPTER 2: Installation Process

2.1 Configuring Power BI

Getting Started

Setting Up Power BI: Follow these step-by-step instructions to download, install, and set up Power BI Desktop, including any necessary configurations.

For installation guide follow this link: Power BI Installation Guide

2.2 Installing the Required Libraries and Python

Installing Python and Necessary Libraries

Installing Python and Essential Libraries: Install Python and the essential libraries for data analysis and visualization, such as Pandas, NumPy, Matplotlib, Seaborn, and Plotly.

For detailed information, visit: Python Downloads

Library Installation: Libraries can be installed using the pip command in CMD (e.g., `pip install pandas`). For virtual labs like Google Colab, these libraries are pre-installed in the backend system.

System Configurations

- **Operating System:** Windows 10 or above
- **Python:** Ensure Python is installed
- **Jupyter Notebook:** Recommended for data analysis and visualization
- **Required Libraries:** Install modules like Seaborn, Matplotlib, Plotly, NumPy, etc., as needed
- **Power BI:** Ensure Power BI Desktop is installed and configured

CHAPTER III - PREPARING DATA

CHAPTER 3: Preparing Data

3.1 Data Acquisition

Overview

Data acquisition involves collecting relevant data from various sources for analysis. This step is critical as the quality and reliability of the data directly impact the outcomes of the analysis.

Sources of Data

- **Internal Databases:** Company records, sales data, employee data
- **External Databases:** Public databases, research datasets, APIs
- **Web Scraping:** Extracting data from websites
- **Manual Data Entry:** Collecting data through surveys or forms

Steps in Data Acquisition

1. **Identify Data Requirements:** Determine the type and scope of data needed for analysis.
2. **Source Data:** Access data from identified sources, ensuring reliability and relevance.
3. **Extract Data:** Download or extract data in a usable format (e.g., CSV, Excel, JSON).

3.2 Data Cleaning and Preprocessing

Overview

Data cleaning and preprocessing involve preparing raw data for analysis by handling inconsistencies, missing values, and irrelevant information.

Steps in Data Cleaning

4. **Remove Duplicates:** Identify and eliminate duplicate entries to ensure data integrity.
5. **Handle Missing Values:** Replace, remove, or impute missing data points using appropriate methods.
6. **Correct Errors:** Fix data entry errors, such as typos or incorrect values.
7. **Standardize Data:** Ensure consistent formatting (e.g., date formats, unit measurements).

Steps in Data Preprocessing

8. **Normalize Data:** Scale numerical data to a standard range for uniformity.
9. **Encode Categorical Data:** Convert categorical variables into numerical formats using techniques like one-hot encoding.
10. **Feature Engineering:** Create new features or modify existing ones to improve model performance.

3.3 Data Transformations

Overview

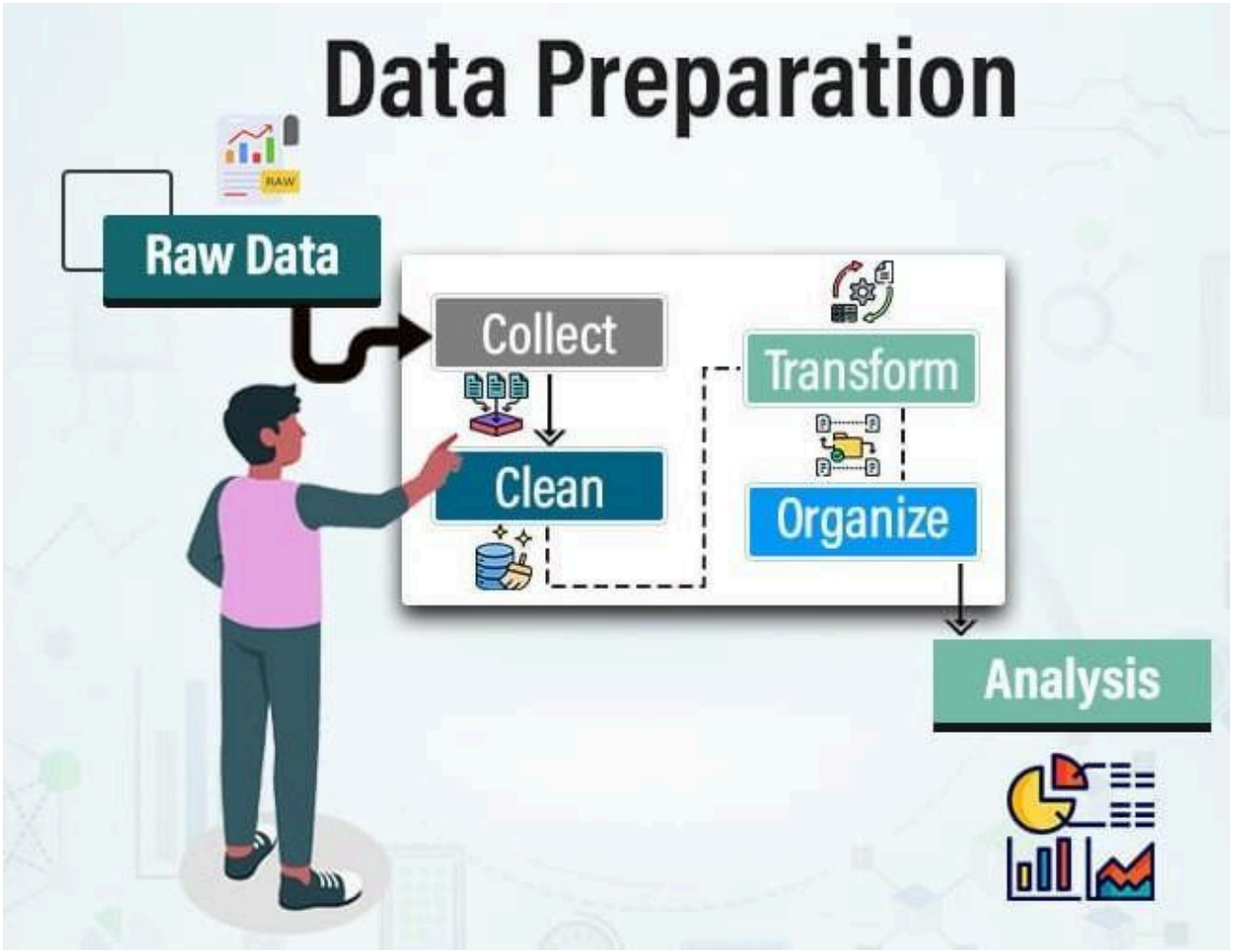
Data transformation involves modifying the dataset to enhance its suitability for analysis and to uncover hidden patterns.

Types of Data Transformations

11. **Aggregation:** Summarize data by grouping and calculating aggregate metrics (e.g., mean, sum).
12. **Filtering:** Select a subset of data based on specific criteria to focus the analysis.
13. **Merging:** Combine multiple datasets into a single cohesive dataset.
14. **Pivoting:** Reshape data to make it easier to analyze, often converting rows into columns or vice versa.
15. **Scaling:** Adjust the range of data to bring all features onto a comparable scale.

Importance of Data Transformation

- **Enhances Data Quality:** Improves the reliability and accuracy of the data.
- **Facilitates Analysis:** Makes data easier to work with and more insightful.
- **Uncovers Insights:** Reveals hidden patterns and relationships within the data.



CHAPTER IV – ANALYSIS WITH PYTHON

4.1 Data Loading and Exploration

Libraries

- **Pandas:** For data manipulation and analysis.
- **NumPy:** For numerical operations and array handling.
- **Matplotlib:** For creating static visualizations.
- **Seaborn:** For statistical data visualization.
- **Plotly:** For interactive visualizations.

Data Loading

Steps to Load Data

- **Using Pandas:** Load data from CSV files into DataFrames with `pandas.read_csv()`.

Exploratory Data Analysis (EDA)

EDA Process

- **Summary Statistics:** Calculate mean, median, and standard deviation to understand data distribution.
- **Data Visualization:** Plot histograms, scatter plots, and other visualizations to identify patterns and relationships.
- **Identify Missing Values:** Detect and handle missing values appropriately, either by removing or imputing them.

4.2 Data Visualization with Python

Data Visualization (Matplotlib/Seaborn)

Key Plots

- **Histograms:** Visualize the distribution of data.
- **Scatter Plots:** Show relationships between two variables.
- **Bar Charts:** Compare different categories.
- **Box Plots:** Display data spread and identify outliers.
- **Error Plots:** Indicate variability and uncertainty in data.

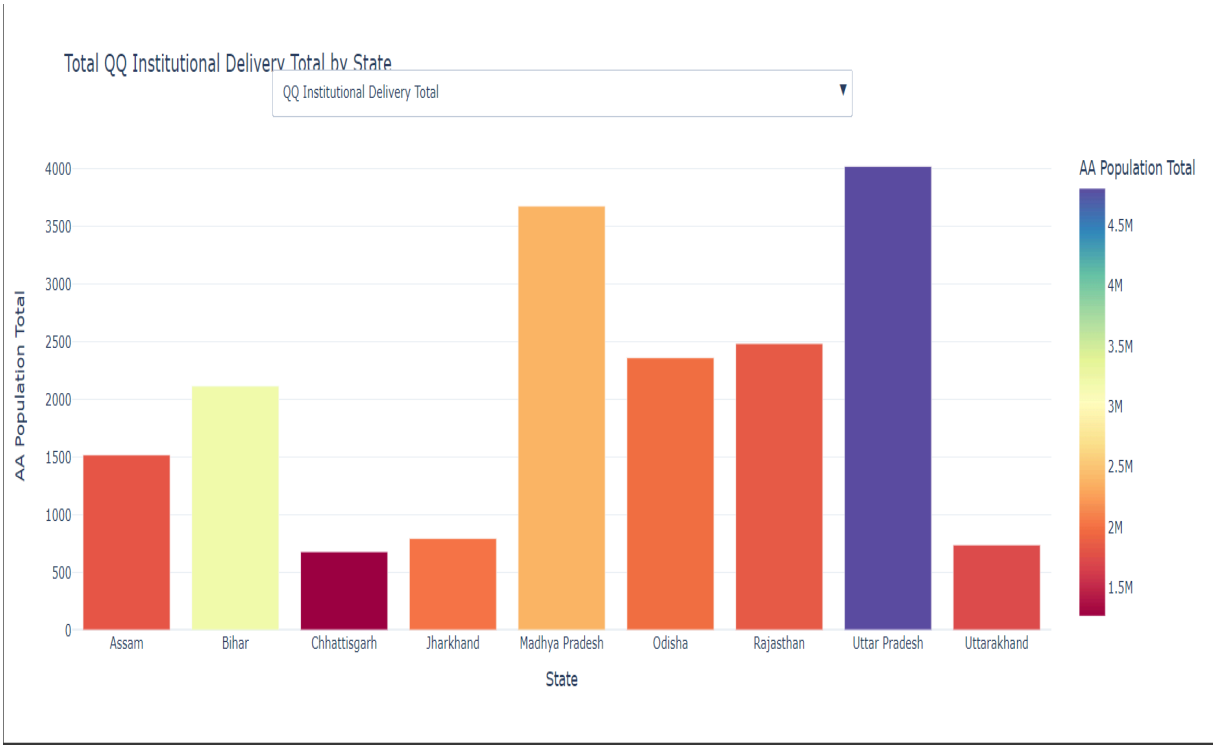
Data Visualization with Plotly

Creating Interactive Visualizations

- **Line Charts:** Show trends over time with interactive features for detailed exploration.
- **Scatter Plots:** Dynamically explore relationships between variables.
- **Other Visualizations:** Enhance data exploration with interactive bar charts, box plots, and more to gain deeper insights.

PROBLEM STATEMENT – 1

Total Institutional Deliveries by State: Using Bar Chart

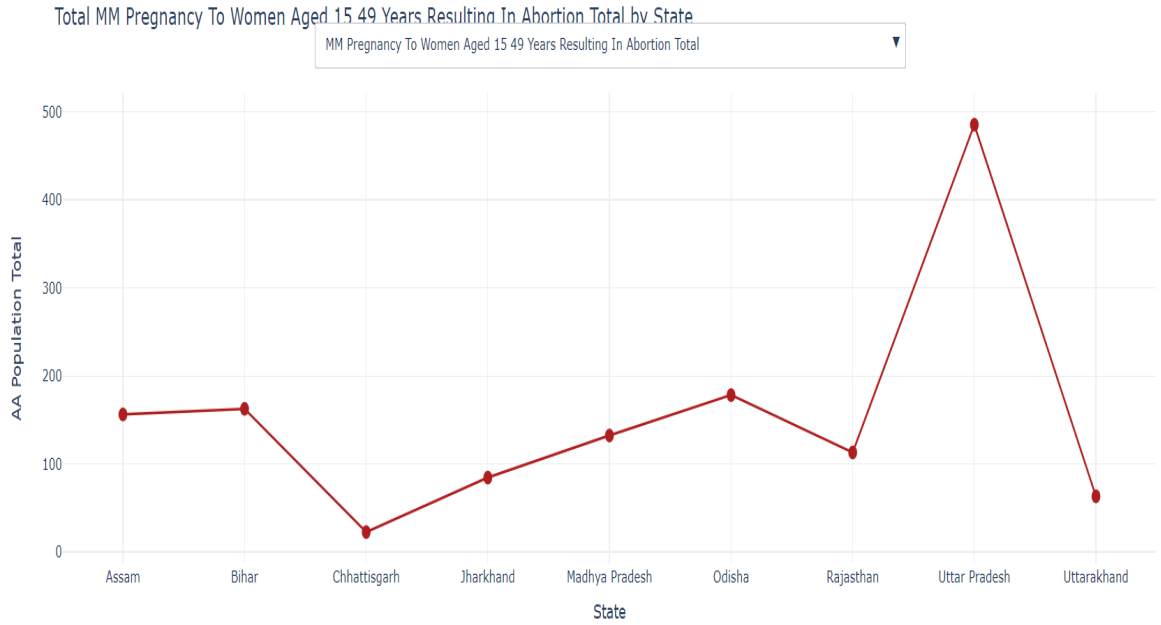


Explanation:

A bar chart visually represents categorical data by using the length of bars. The total institutional deliveries by state bar chart displays the distribution of institutional deliveries across various states, with" Uttar Pradesh "having the highest, followed by states like "Madhya Pradesh" and "Rajasthan". Each bar represents the count of occurrences for a specific job title. The colors show the count of deliveries, ranging from 1.5 million to 4.5 million. The x-axis lists States, while the y-axis shows the Population

PROBLEM STATEMENT – 2

Total Number of Women Aged 15 to 49 years Resulting in Abortion in Each State : Using Line Chart

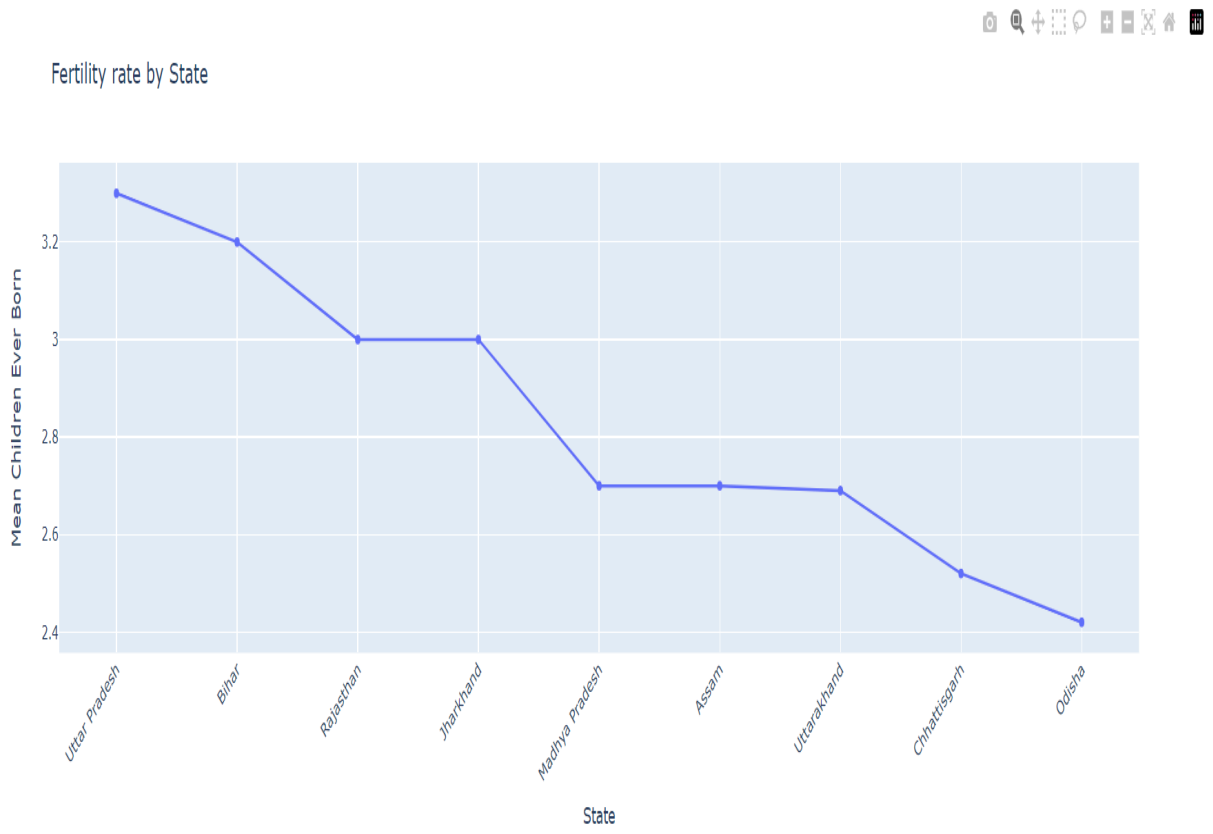


Explanation:

Interactive line charts allow users to dynamically interact with data. The interactive line chart displays the total number of women aged 15 to 49years resulting in abortion in each state. Each dot represents a different state, with the height indicating the number of women aged 15 to 49years resulting in abortion. The plot includes labels for the x-axis (State) and y-axis (Total Population).The highest abortions are observed in Uttar Pradesh and least goes to Chhattisgarh according to the graphical data analysis.

PROBLEM STATEMENT – 3

Fertility Rate by State: Using Line Chart

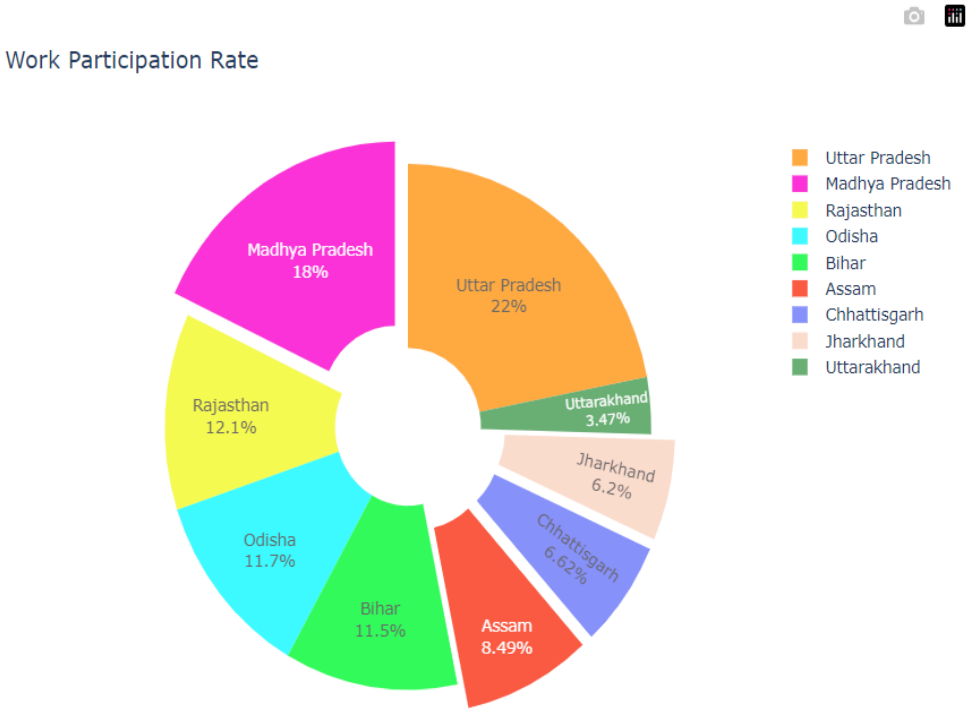


Explanation :

The above fertility rate by state chart displays the mean children ever born in each state. Each dot represents a different state, with the height indicating the number mean children ever born. The plot includes labels for the x-axis (State) and y-axis (Mean Children Ever Born).So according to the graphical analysis of the data above highest fertility rate is observed in Utttar Pradesh and least fertility rate is observed in Odisha.

PROBLEM STATEMENT – 4

Work Participation Rate: Using Donut Chart



Explanation:

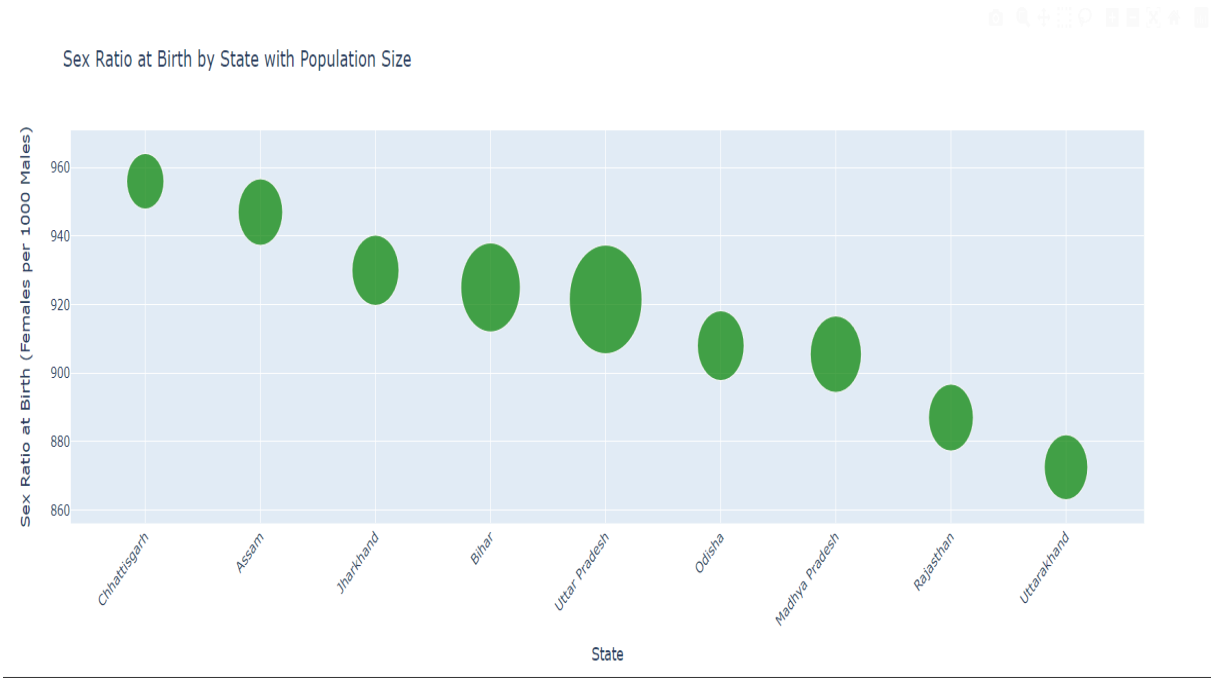
The donut chart shows the work participation rate across various Indian states, with Uttar Pradesh at 22%, Madhya Pradesh at 18%, Rajasthan at 12.1%, Odisha at 11.7%, Bihar at 11.5%, Assam at 8.49%, Chhattisgarh at 6.62%, Jharkhand at 6.2%, and Uttarakhand at 3.47%.

Key Findings:

Highest work participation rate is observed in Uttar Pradesh with 22%
Least work participation is observed in Jharkhand and Chhattisgarh with slight difference of equal proportion with 6.2% and 6.62%

PROBLEM STATEMENT – 5

Sex Ratio at Birth by State with Population Size: Using Bubble Chart

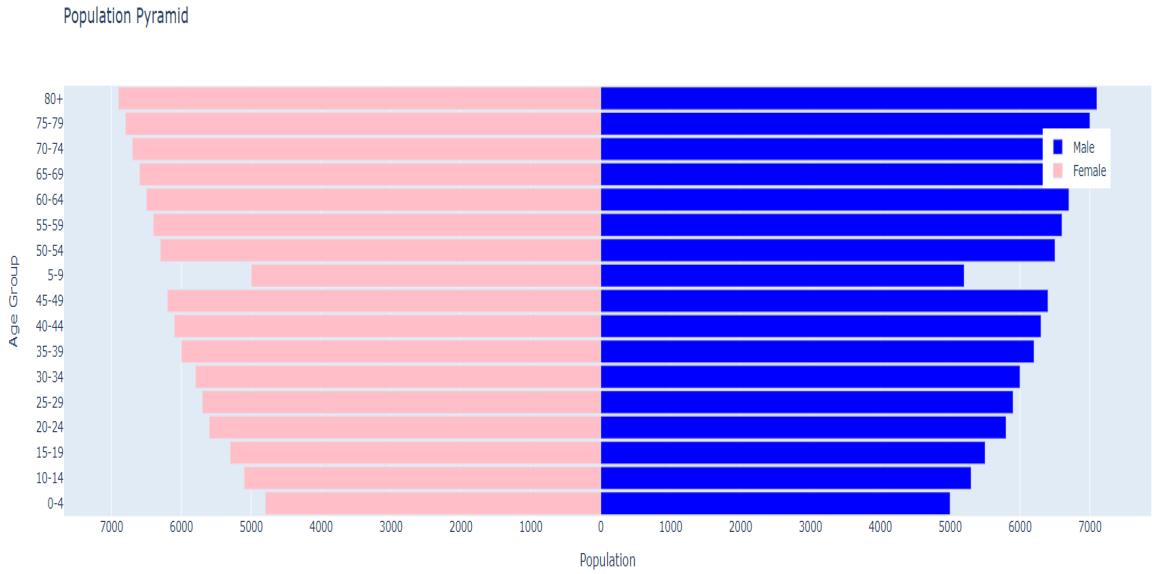


Explanation:

The bubble chart illustrates the sex ratio at birth by state with population size, showing states like Chhattisgarh and Assam with higher sex ratios, while Rajasthan and Uttarakhand have lower ratios. The x-axis lists states, the y-axis represents the sex ratio (females per 1,000 males), and bubble size indicates population size, with Uttar Pradesh having the largest population.

PROBLEM STATEMENT – 6

Population by Age Group and Gender: Using Population Pyramid

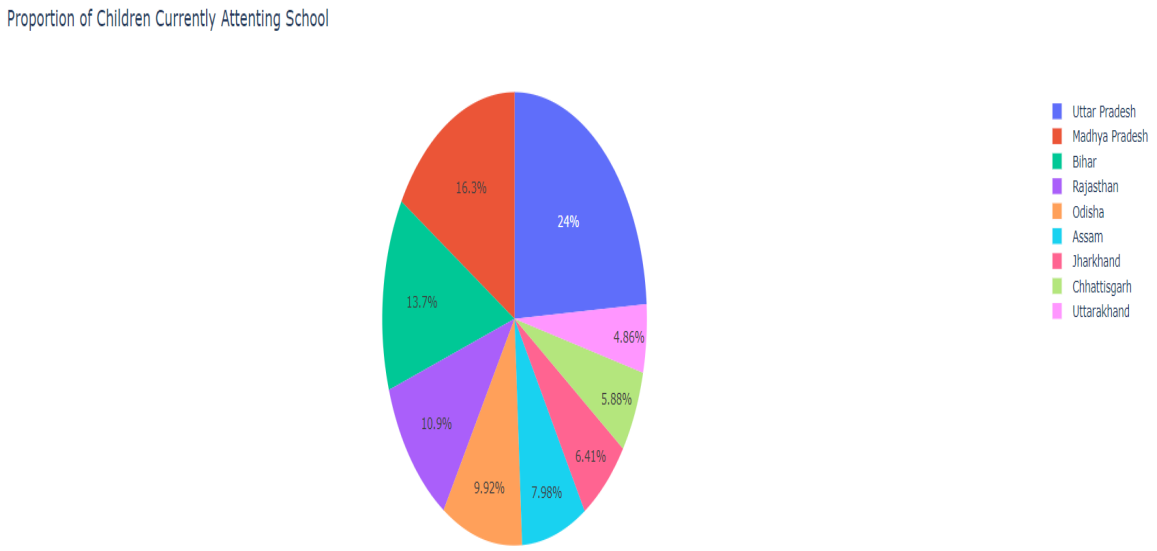


Explanation:

The population pyramid shows the distribution of population by age group and gender, with blue bars for males and pink bars for females. The x-axis represents the population count, and the y-axis represents age groups from 0-4 to 80+. Younger age groups (0-34) show a balanced gender distribution, middle age groups (35-59) have a slightly larger male population, and older age groups (60+) have more females, reflecting higher female life expectancy. The pyramid shape, with a broad base, indicates a growing population, highlighting the need for targeted age-specific healthcare and support services

PROBLEM STATEMENT – 7

Children Currently Attending School: Using Pie Chart



Explanation:

The pie chart shows the proportion of children currently attending school in various states, with Uttar Pradesh having the highest at 24%, followed by Madhya Pradesh (16.3%) and Bihar (13.7%). Rajasthan, Odisha, and Assam have proportions of 10.9%, 9.92%, and 7.98%, respectively, while Jharkhand, Chhattisgarh, and Uttarakhand have lower proportions of 6.41%, 5.88%, and 4.86%. This indicates notable disparities in school attendance rates, with states having lower proportions potentially needing targeted interventions to improve education access.

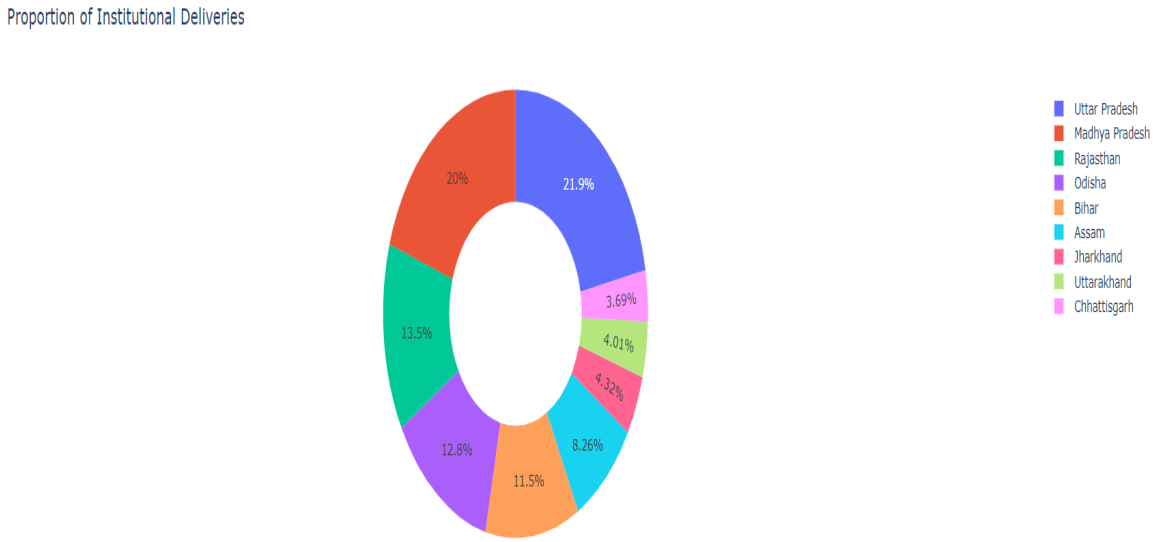
Key Findings:

Highest(most of the Children Attending School): In Uttar Pradesh most of the children are attending the school with 24% of the total proportion which is a good indicator for development in education

Least(less number of Children Attending School): Less number of children are attending the school from Uttarakhand with only 4.86% which indicates lack of education and literacy rate

PROBLEM STATEMENT – 8

Proportion of Institutional Deliveries: Donut Chart



Explanation:

The donut chart displays the proportion of institutional deliveries across various Indian states, with Uttar Pradesh at 21.9%, Madhya Pradesh at 20%, Rajasthan at 13.5%, Odisha at 12.8%, Bihar at 11.5%, Assam at 8.26%, Jharkhand at 4.32%, Uttarakhand at 4.01%, and Chhattisgarh at 3.69%.

Key Findings:

Highest Institutional Deliveries: The highest institutional deliveries were done at Uttar Pradesh with 21.9% of the total proportion because of its more number of Private Institutions(Hospitals) with vast area and population.

Lowest Institutional Deliveries: The lowest institutional deliveries were done at Chhattisgarh with 3.69% of the total proportion because of its least population and less infrastructure.

CHAPTER IV – ANALYSIS USING POWER BI

5.1 Data Import and Preparation

Steps to Import Data

- **Connect to Data Sources:** Utilize Power BI's capabilities to connect to various data sources such as Excel and SQL Server.

Example

- **Excel and SQL Database:** Import data from an Excel file and a SQL database for integration into Power BI.

Data Transformation

- **Data Cleaning:** Use Power Query Editor to remove duplicates, filter data, and add calculated columns.

Example

- **Cleaning Tasks:** Remove rows with missing values, filter irrelevant data, and add calculated columns for derived metrics.

5.2 Data Modeling

Creating Data Models

- **Define Relationships:** Establish relationships between different data tables.
- **Create Calculated Columns and Measures:** Use DAX (Data Analysis Expressions) to create custom calculations.

Example

- **Relationships:** Link crop production data with climatic data.
- **Calculated Measures:** Create measures for total production, average yield, etc.

5.3 Data Visualization with Power BI

Develop Interactive Dashboards

- **Dashboards:** Display key metrics like total crop production, production by region, and yield trends.

Example

- **Dashboard Example:** Show total production over time with filters for different crops and regions.

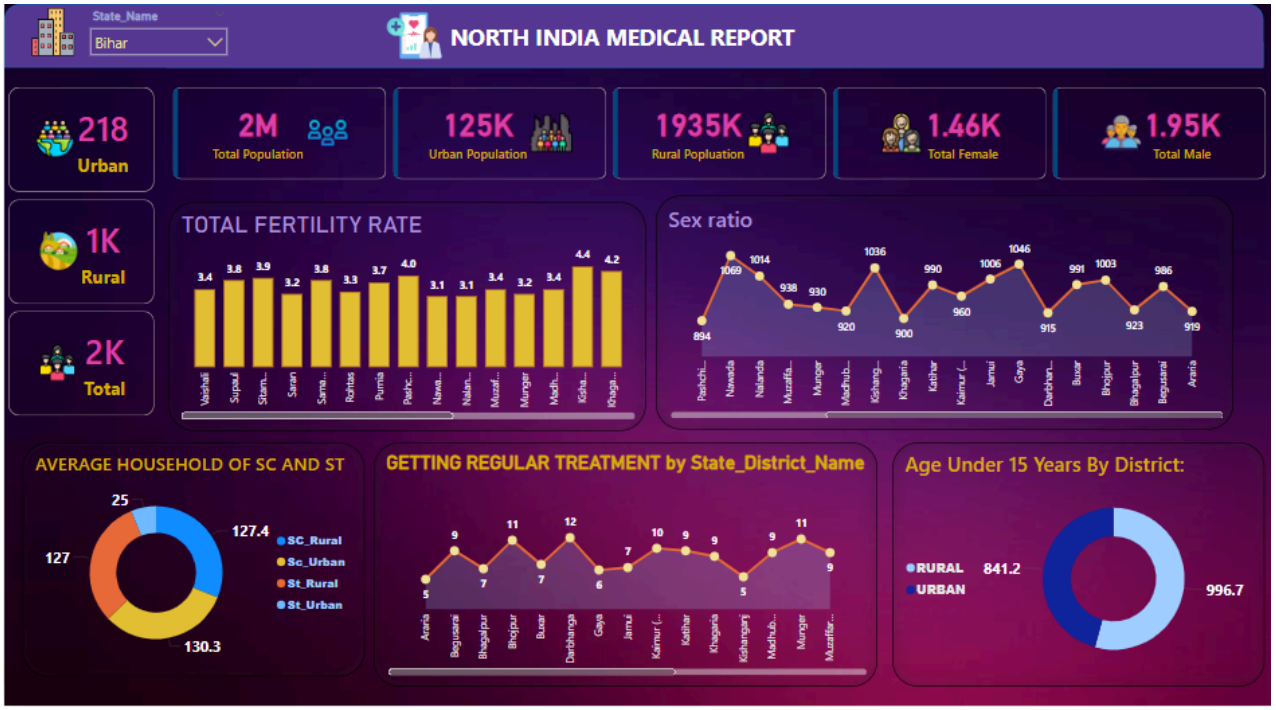
Generate Detailed Reports

- **Reports:** Create detailed reports for comprehensive analysis, such as monthly production and comparative yield reports.

Example

- **Report Example:** Compare crop yields across different countries and over various years.

POWER BI DASHBOARD



Explanation:

- **Title:** North India Medical Report

- **Objective:** To analyse north India medical report with total population by urban and rural divisions as well. The report configures the total fertility by state, sex ratio ,average household of SC and ST, getting regular treatment by state, district-name and age Under 15 years By District

Key Insights from the dashboard:

1. Total Population:

Total: 2M
Urban: 125K
Rural: 1935K

2. Gender Distribution:

Total Female: 1.46K
Total Male: 1.95K

3. Urban and Rural Population:

Urban: 218
Rural: 1K
Total: 2K

4. Total Fertility Rate: (shown as a bar graph for different regions)

Values range approximately between 3.1 to 4.2 for various districts.
Sex Ratio: (shown as a line graph for different regions)
Values range approximately between 894 to 1069 for different districts.

5. Average Household of SC and ST:

SC Rural: 127
SC Urban: 130.3
ST Rural: 127.4
ST Urban: 25 (possibly indicating a different measure or data point)

6. Getting Regular Treatment:

Example: Buxar has a value of 7.20

7. Age Under 15 Years by District: (shown as pie charts for Urban and Rural)

Urban: 841.2
Rural: 996.7

These metrics provide a comprehensive overview of the population distribution, gender ratio, fertility rates, healthcare access, and age demographics in the region of Bihar.

Various Plots/ graphs used in this dashboard:

1. Bar Chart

Bar chart is a type of graph that uses rectangular bars to represent data. The length or height of each bar corresponds to the value it represents. Bar graphs are a great way to compare data between different categories. They are simple to understand and effective at visually highlighting differences and relationships between the data points.

2. Donut Chart

A donut chart is a variation of a pie chart with a blank center. It shows categorical data with segments representing parts of a whole. The size of each segment corresponds to the proportion of each category.

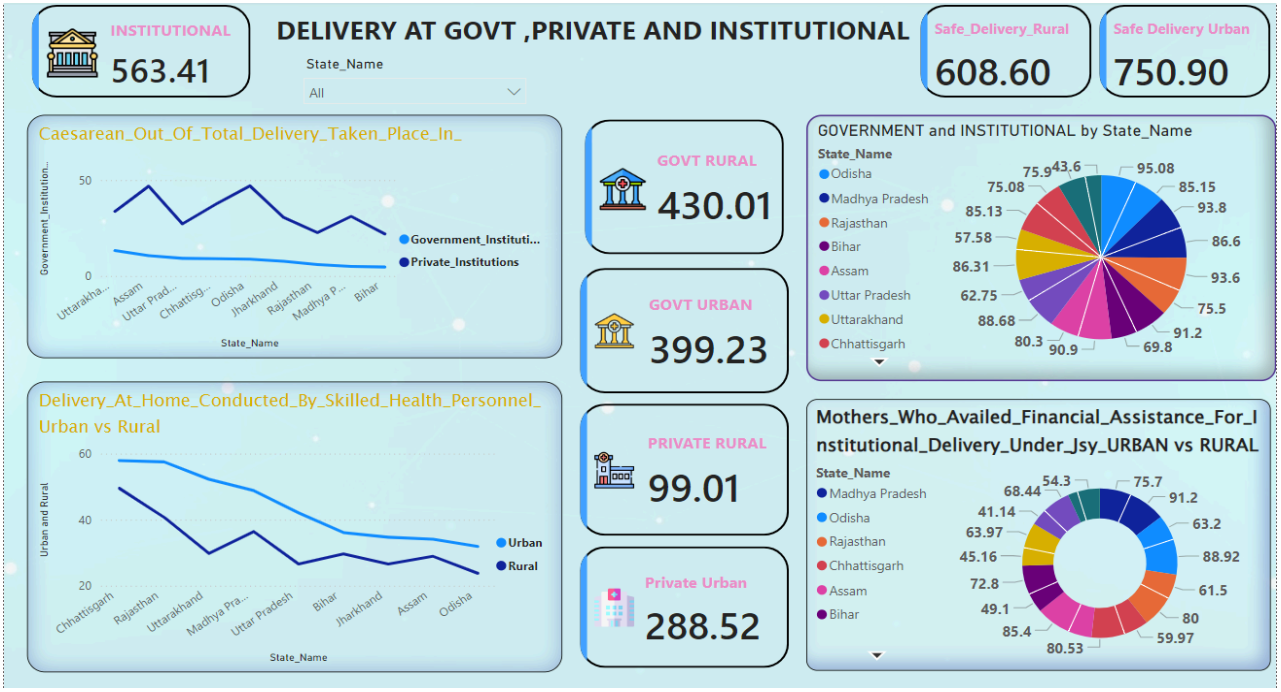
3. Line Chart:

A funnel chart is a graph that depicts stages in a process, with the number of participants typically shrinking as they progress through the funnel. Just like a real funnel, the chart is wide at the top and narrows towards the bottom.

4.Area Chart

An area chart is a graphical representation that displays quantitative data over time. It is similar to a line chart but with the area below the line filled in with color. This makes it useful for showing trends over time and comparing multiple data series.

DELIVERY COUNT AT VARIOUS ORGANIZATIONS:



Objective:

Delivery count at different organizations provides a comprehensive overview of safe deliveries at Govt, Private and Institutional organizations.

Key Insights from the dashboard:

1. Overall Institutional Deliveries:

563.41: This represents the total number of deliveries at government, private, and institutional facilities.

2.Safe Deliveries:

Rural: 608.60
Urban: 750.90
These numbers indicate the number of safe deliveries in rural and urban areas.

3. Delivery Breakdown by Type and Area:

Govt Rural: 430.01
Govt Urban: 399.23
Private Rural: 99.01
Private Urban: 288.52
These figures show the number of deliveries conducted in government and private institutions, divided between rural and urban areas.

4. Caesarean Section Deliveries:
A line chart compares Caesarean deliveries in government institutions versus private institutions across various states.

5. Home Deliveries by Skilled Personnel:
Another line chart compares the number of home deliveries conducted by skilled health personnel in urban versus rural settings.

6. Government and Institutional Deliveries by State:

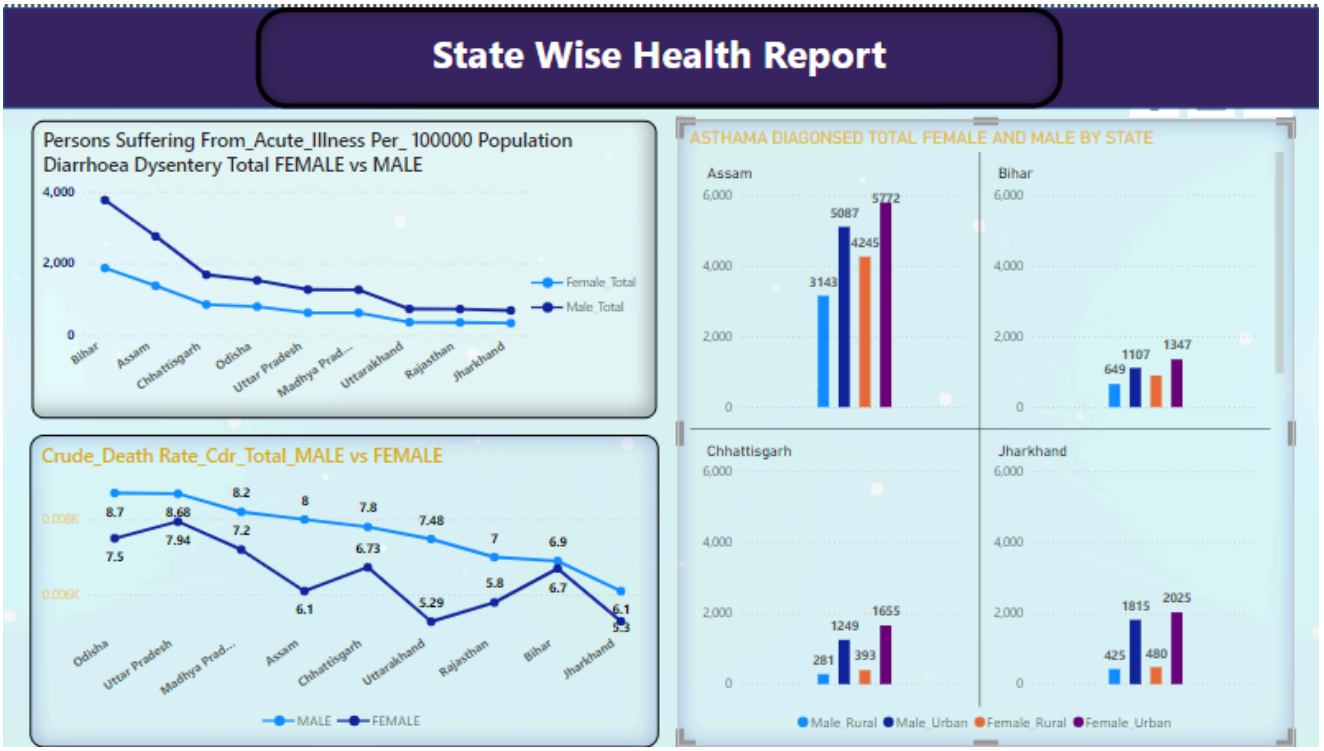
A pie chart shows the distribution of government and institutional deliveries across different states, with percentages.

7. Financial Assistance for Institutional Deliveries:

This section shows the percentage of mothers who availed financial assistance for institutional deliveries under the Janani Suraksha Yojana (JSY) scheme, divided into urban and rural categories.

These key metrics give a comprehensive view of maternal health services, highlighting the distribution and safety of deliveries across different regions and types of facilities.06:56 PM
a comprehensive view of maternal health services, highlighting the distribution and safety of deliveries across different regions and types of facilities.

STATE WISE HEALTH REPORT:



Key Insights from the Dashboard:

1. Persons Suffering from Acute Illness per 100,000 Population (Diarrhea and Dysentery) Female vs. Male:

Line chart shows the number of females and males suffering from acute illnesses per 100,000 population across various states.

Trend Insight: The chart helps in understanding the gender-specific burden of acute illnesses like diarrhea and dysentery in each state.

2.Crude Death Rate (CDR) per 1,000 Population:

Total, Male vs. Female:
Line chart compares the crude death rate for males and females across different states.
Trend Insight: It shows the mortality rate differences between genders and helps identify states with higher or lower death rates.

3. Asthma Diagnosed Total (Female and Male) by State:

Bar Charts:
Bar charts represent the number of females and males diagnosed with asthma in each state.
Assam:
Male: 5,772
Female: 5,087
Male Rural: 4,245
Female Rural: 3,143
Bihar:
Male: 1,347
Female: 1,107
Male Rural: 649
Female Rural: 249
Chhattisgarh:
Male: 1,655
Female: 1,249
Male Rural: 489
Female Rural: 281
Jharkhand:
Male: 2,025
Female: 1,815
Male Rural: 480
Female Rural: 425
Comparison Insight: These charts help compare the prevalence of asthma diagnoses across genders and states.

Summary of Key Insights:

1. Gender-Specific Health Issues:
Acute illnesses like diarrhea and dysentery show a noticeable gender disparity across states, with trends typically higher in females or males depending on the state.
The crude death rate also varies by gender, providing insights into mortality differences.

2. State-Specific Health Trends:
Some states have higher rates of acute illness or crude death rates, which can indicate public health issues needing attention

Asthma diagnosis rates vary significantly between states and genders, highlighting regions where asthma may be a more prevalent health concern.
These metrics collectively provide a comprehensive view of health issues at the state level, enabling targeted public health interventions and resource allocation.

CHAPTER VI – CONCLUSION

CHAPTER 6: Conclusion

1. Delivery Count at Various Organizations:

- **Delivery Count at Various Organizations** bar chart provides a breakdown of delivery counts at government, private, and institutional organizations in rural and urban areas. The analysis shows that institutional deliveries are higher in urban areas compared to rural areas. Additionally, government institutions handle more deliveries than private ones in rural settings. This information is crucial for healthcare administrators to allocate resources effectively. By understanding the distribution of deliveries, healthcare services can be optimized to ensure that both rural and urban areas have adequate facilities and trained personnel.

2. Total Population and Gender Distribution:

- **Total Population and Gender Distribution** The area chart illustrates the total population distribution and gender ratio across various districts in Bihar. It highlights significant differences in population density and gender distribution between rural and urban areas. This demographic information helps policymakers and businesses in planning and delivering services. For instance, healthcare providers can use this data to tailor health programs to specific populations, and businesses can target products and services more effectively based on demographic trends.

3. Age Under 15 Years by District:

- **Age Under 15 Years by District** pie charts provide a visual representation of the proportion of the population under 15 years old in both urban and rural settings. It shows that rural areas have a higher percentage of younger populations compared to urban areas.This insight is valuable for educational institutions and child health services. It indicates a need for more schools, pediatric healthcare services, and child welfare programs in rural areas.

4. Persons Suffering from Acute Illness:

- **Persons Suffering from Acute Illness** chart compares the number of males and females suffering from acute illnesses such as diarrhea and dysentery across various states. It highlights gender-specific health issues and the prevalence of these conditions. For healthcare providers and public health officials, this data is critical for developing targeted health interventions and allocating medical resources efficiently to address gender-specific health needs.

5.Safe Deliveries Dashboard:

- **Safe Deliveries Dashboard** dashboard provides an overview of safe deliveries at government, private, and institutional facilities. It includes a breakdown of deliveries by type (e.g., caesarean section) and location (rural vs. urban), and highlights the financial assistance received under the Janani Suraksha Yojana (JSY) scheme.

The dashboard is highly beneficial for healthcare policymakers and hospital administrators. It helps in monitoring and improving maternal health services, planning resource allocation, and assessing the impact of financial assistance programs on institutional deliveries.

6. Health and Demographic Overview of Bihar:

- **Health and Demographic Overview of Bihar** dashboard presents a detailed demographic and health profile of Bihar, including population distribution, fertility rates, and healthcare access. It identifies key health indicators and their variations across different districts. This comprehensive view aids government agencies and NGOs in developing targeted health interventions and policies. Businesses can also use this data for market analysis and to tailor health-related products and services to meet the specific needs of different regions.

7. State-wise Health Report:

- **State-wise Health Report** dashboard compares health indicators such as crude death rate (CDR), asthma prevalence, and acute illness rates across different states. It shows significant variations in health outcomes between states and genders. For public health officials, this dashboard is essential for identifying states with higher health risks and developing state-specific health programs. Businesses in the healthcare industry can use this information to expand their services in regions with higher health needs.

Business Industries

1. Healthcare Providers:

- **Benefit:** Detailed health data analysis helps in optimizing service delivery, improving patient outcomes, and efficiently allocating medical resources.
- **Application:** Hospitals and clinics can use these insights to expand their facilities, hire necessary medical staff, and improve healthcare access in underserved areas.

2. Pharmaceutical Companies:

- **Benefit:** Understanding regional health trends allows pharmaceutical companies to tailor their product offerings and marketing strategies.
- **Application:** Companies can focus on areas with high prevalence of specific illnesses, ensuring the availability of relevant medications and conducting targeted health awareness campaigns.

3. Insurance Companies:

- **Benefit:** Health data analytics assist in risk assessment, pricing insurance products, and identifying regions with higher healthcare needs.
- **Application:** Insurance providers can design customized health plans and offer better coverage options based on regional health data.

4. Public Health Agencies:

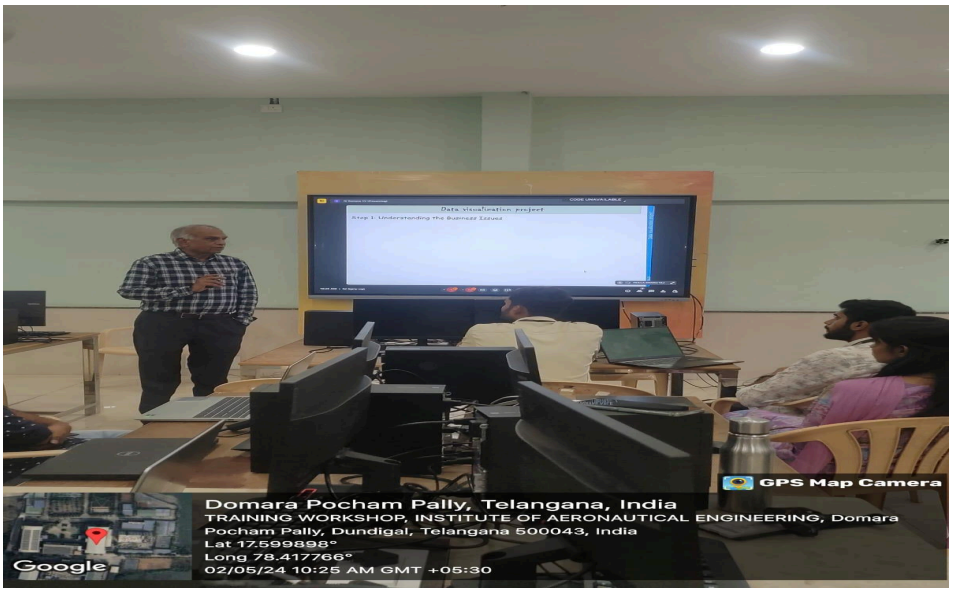
- **Benefit:** Comprehensive health data helps in formulating effective public health policies, monitoring health outcomes, and evaluating the impact of health interventions.
- **Application:** Agencies can allocate resources, plan public health campaigns, and implement programs targeting specific health issues highlighted by the data.

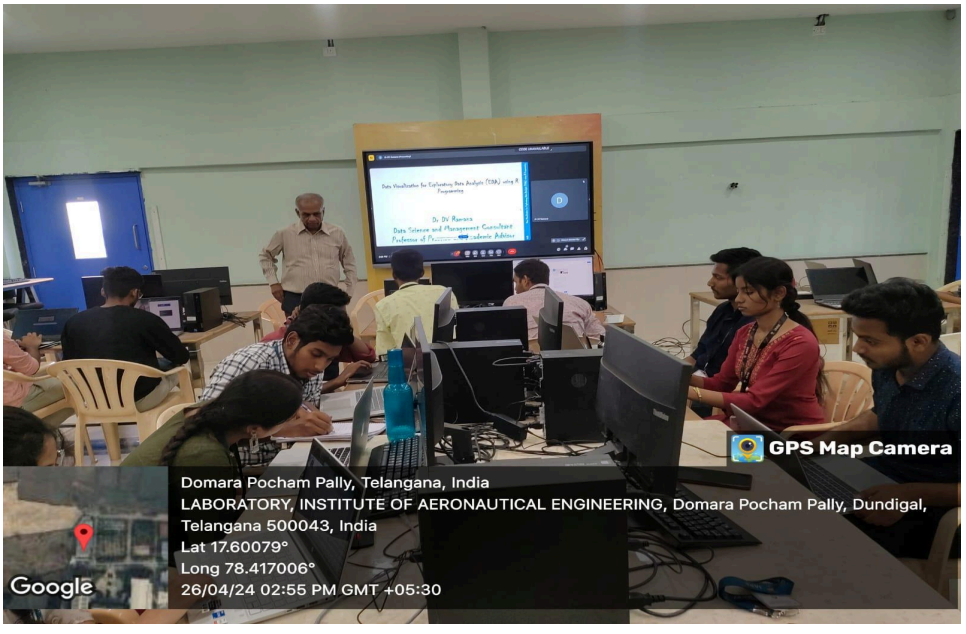
5. Market Research Firms:

- **Benefit:** Detailed demographic and health insights are valuable for conducting market analysis and identifying business opportunities.
- **Application:** Firms can advise clients on market entry strategies, product development, and marketing initiatives based on regional health and demographic data.

By integrating Python and Power BI for data analysis and visualization, this project not only provides actionable insights into health indicators but also demonstrates the potential for enhancing decision-making processes across various sectors in the business industry.

IMAGES FROM THE INTERNSHIP:





CHAPTER VII – References

Dataset:

Accessed from govt website; https://data.world/rajanand/key-indicators-of-annual-health-survey

Python:

Official Python Website: https://www.python.org/

Power BI:

Microsoft Power BI:

https://www.microsoft.com/en-us/power-platform/products/power-bi#tabs-pill-bar-ocb9d418_tab1

- **Libraries:**
- **Pandas ("Panel Data"):**
    - Data Structures and Tools
    - Version: pandas: 2.1.3
    - Official Documentation: Pandas
- **NumPy ("Numerical Python"):**
    - Arrays and Matrices
    - Version: numpy: 1.26.1
    - Official Documentation: NumPy
- **SciPy ("Scientific Python"):**
    - Integrals, Solving Differential Equations, Optimization
    - Version: SciPy: 1.11.4
    - Official Documentation: SciPy
- **Matplotlib ("Matlab-like Plotting Library"):**
    - Plots and Graphs, Most Popular
    - Version: matplotlib:3.8.1
    - Official Documentation: Matplotlib
- **Seaborn:**
    - Plots: Heat Maps, Time Series, Violin Plots
    - Version: seaborn: 0.13.2
    - Official Documentation: Seaborn
- **Plotly Open-Source Graphing Library for Python ("Plotly Graphing Library"):**
    - Interactive Graphs and Dashboards
    - Version: plotly:5.20.0
    - Official Documentation: Plotly