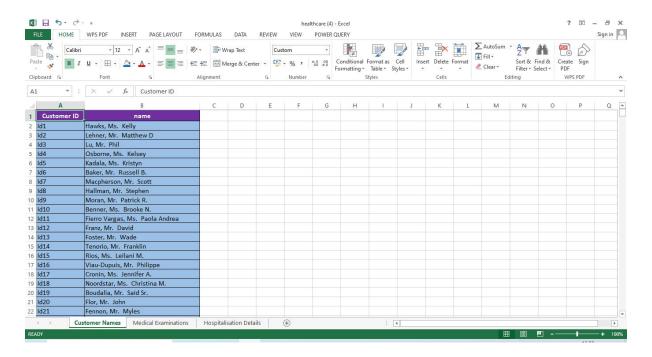
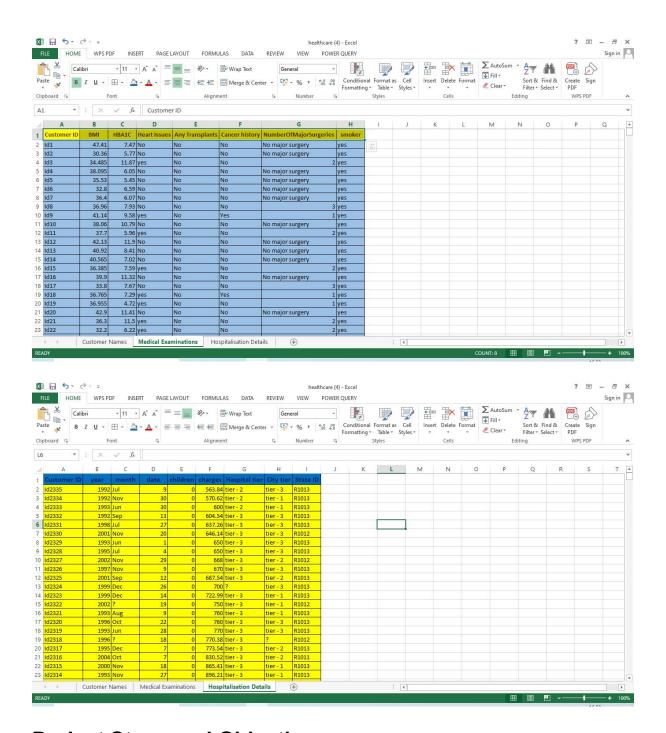
HEALTHCARE DATA ANALYSIS AND INSIGHTS

Problem Statement:

The healthcare industry generates vast amounts of data daily, providing valuable insights for healthcare providers and policymakers to improve patient care, allocate resources effectively, and manage healthcare costs. This project aims to analyze a comprehensive healthcare dataset comprising medical examinations, hospitalization details, and customer profiles to extract insights into patient health profiles, medical histories, and healthcare costs. By exploring relationships between various health metrics, identifying trends, and visualizing key patterns, we aim to deliver actionable insights to healthcare stakeholders for informed decision-making through rigorous data cleaning, transformation, exploration, and analysis.

Old Dataset





Project Steps and Objectives:

Data Cleaning:

- 1) Check for the number of missing values marked with '?' in each column of the "Medical Examinations" Table and "Hospitalization Details" Table.
- 2) Fill in the missing values of 'month' with Sep and 'year' with its average rounded to the nearest integer.
- 3) Determine the most frequently occurring values in the 'smoker', 'Hospital tier' and 'City tier' columns, and fill in the missing values accordingly.

4) If any 'State ID' values are missing, consider filling them with 'Unknown' or using another appropriate strategy.

Data Transformation:

- 1) Split the 'names' column in the "Customer Names" Table into 3 meaningful columns: 'Title', 'First Name', and 'Last Name'.
- 2) Convert the "NumberOfMajorSurgeries" column in the "Medical Examinations" Table to numerical data by replacing non-numeric characters with meaningful numerical values.
- 3) Check for inconsistencies in the 'Heart Issues' and 'smoker' columns and propose corrective actions if necessary.
- 4) Create a new column named "Weight Status" that categorizes BMI into different categories as below:

BMI Weight Status

Below - 18.5 - Underweight

18.5 – 24.9 - Normal Weight

25.0 - 29.9 - Overweight

30.0 and Above Obesity

5) Create a new column named "Diabetes Status" and fill it as per the information given below:

HbA1C Diabetes Status

Below - 5.7 Normal

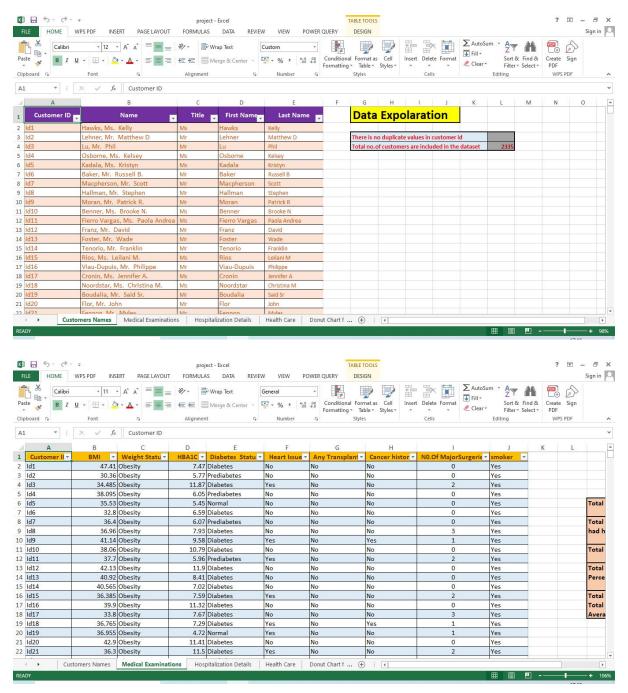
5.7 - 6.4 - Prediabetes

6.5 and Above Diabetes

6) Merge 'year', 'month' and 'date' columns in the "Hospitalization Details" Table into one column named 'Date of Birth' and format it in 'DD-MMM-YYYY' custom format.

7) Calculate the 'Age' of each customer based on their 'Date of Birth' and the date of collection of the dataset, which is 8th June 2023. (Hint: Use the DATEDIF function)



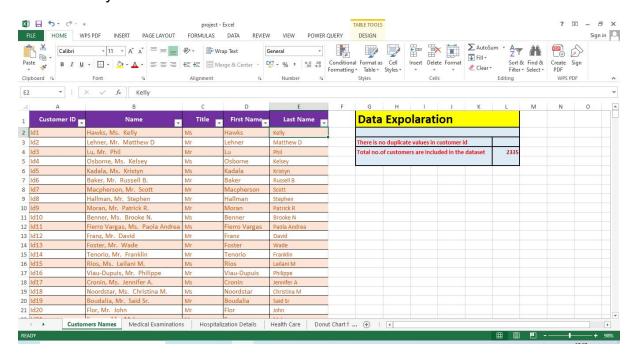


Data Exploration:

→ Customer Names Table:

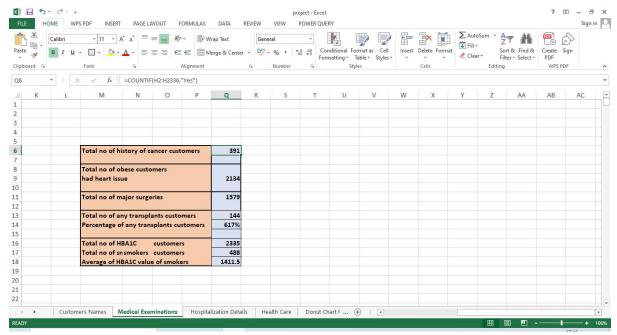
➤ Are there any duplicate Customer IDs in the dataset? If yes, how many?

> How many customers are included in the dataset?



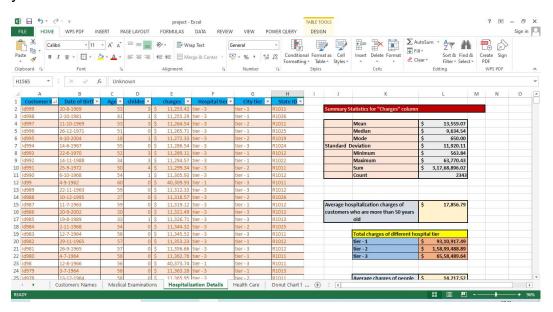
→ Medical Examination Table:

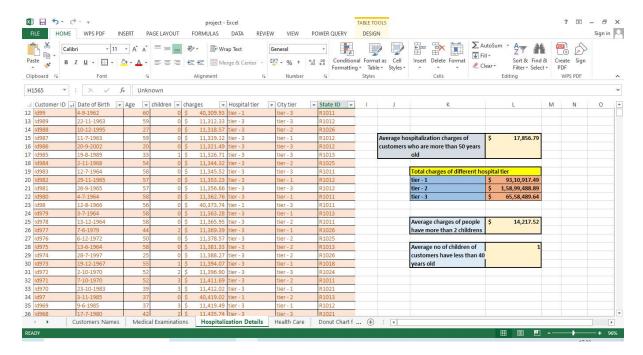
- ➤ How many customers have a history of cancer?
- ➤ How many obese customers have heart issues?
- > What is the total number of major surgeries performed on customers?
- > Calculate the percentage of customers who have undergone any transplants.
- ➤ Find the average HBA1C value of customers who are smokers.



→ Hospitalization details Table:

- ➤ Calculate all the Summary statistics for the 'charges' column.
- ➤ Find the average hospitalization charges for customers who are more than 50 years old.
- > Compare the total charges across different hospital tiers.
- ➤ Calculate the average charges for people who have more than 2 children.
- ➤ Find the integer average number of children of customers who are less than 40 years old.

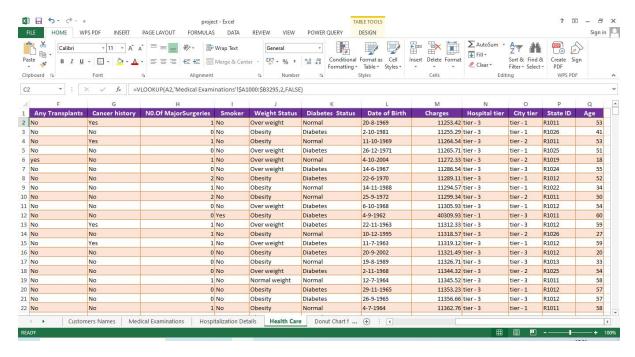




Data Analysis:

- ➤ Create a new sheet named "Healthcare", combine all three tables into one, using Customer ID as the common column, utilizing VLOOKUP.
- ➤ Retain the following necessary columns: Customer ID, First Name, BMI, HBA1C, Heart Issues, Any Transplants, Cancer history, NumberOfMajorSurgeries, smoker, Weight Status, Diabetes Status, Date of Birth, charges, Hospital tier, City tier, State ID,

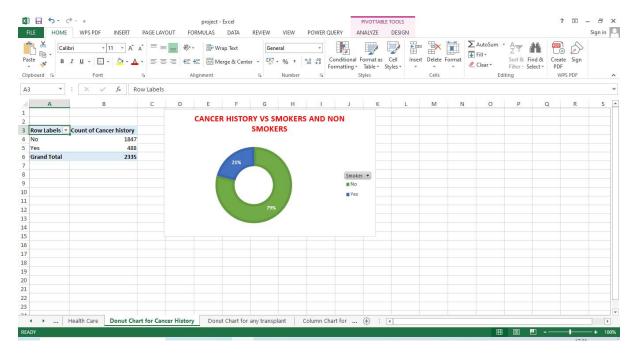
Age. ? 🗷 – 🗗 X Sign in FILE HOME WPS PDF INSERT PAGE LAYOUT POWER QUERY DESIGN FORMULAS DATA REVIEW VIEW Calibri - 11 - A A A = = = 89 - ₩ Wrap Text Br Br ∑ AutoSum → AT H General **▼** Fill + B I U → 🖽 → 🚵 → 🙇 → 🚍 등 등 등 🖅 🗒 Merge & Center Conditional Format as Cell Formatting Table Styles Create Sign PDF Insert Delete Format Sort & Find & Filter + Select + Font Fg. Alignment 5 fx =VLOOKUP(A2,'Medical Examinations'!\$A1000:\$B3295,2,FALSE) eries Smoker Weight Status Diabetes Status Date of Bir Campbell 28.6 1 No Over weight 36.6 5.2 Yes No Yes 1 No Obesity Normal 11-10-1969 Diabete Peters 28.31 4.55 No yes 1 No Over weight Normal 4-10-2004 Diabete Tufaro 29.83 14-6-1967 Chen 32.775 8.02 Yes No 2 No Obesity Diabetes 22-6-1970 Mitchell 38.93 5 64 Ves Ohesity 14-11-1988 33.7 4.01 No Obesity 25-9-1972 Rall Normal 2 No 11 ld990 12 ld99 Donnelly 27.645 10.56 No 0 No 6-10-1968 41.51 6.92 No No 0 Yes Diabetes 4-9-1962 Gage Obesity Over weight 24.91 7.33 Yes No 1 No Diabetes 22-11-1963 48.12 4.51 Yes 10-12-1995 1 No Obesity Chiappone Normal 9.18 Yes Anderson 24.93 1 No Diabetes 11-7-1963 54.47 7.48 No 0 No 20-9-2002 Webster Diabetes McCue 42.46 4.11 No No No 0 No Obesity Normal 19-8-1989 Winters 27.75 9.88 No 0 No Over weight Diabetes 2-11-1968 Blain 23.3 5.36 Yes 1 No Normal weight 12-7-1964 6.72 No 29-11-1965 Burger 34.01 No 0 No Obesity Diabete 26-9-1965 Robertson 4-7-1964 35.7 Obesity 1 No Normal Customers Names | Medical Examinations | Hospitalization Details | Health Care | Donut Chart f ... 🕂 : 🔻



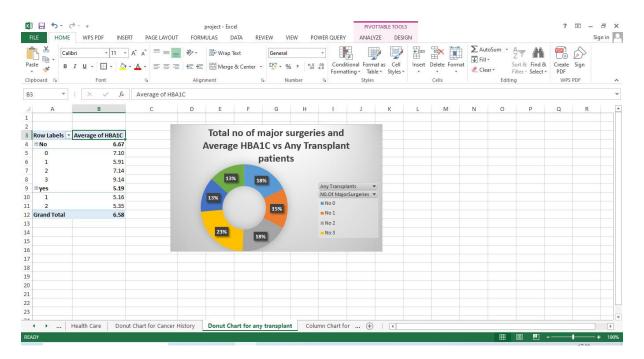
Create pivot tables if required to do the following analysis, then visualize through charts:

Analysis using Pie/Donut Chart:

> What is the distribution of cancer history among smokers and non-smokers?

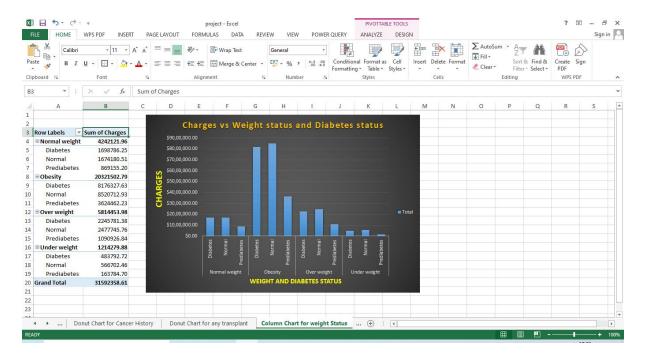


➤ How does the total number of major surgeries and average HbA1C differ between patients with and without a history of transplants?

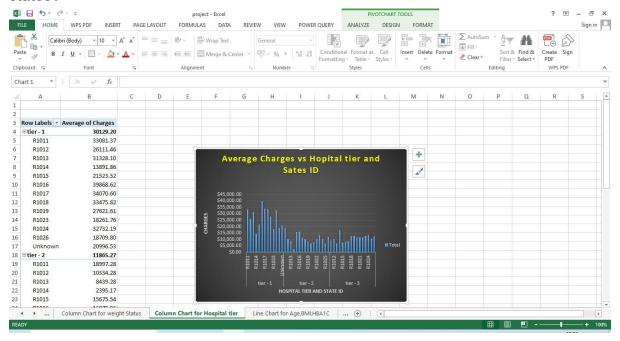


Analysis using Column/Bar Chart:

➤ How do healthcare charges vary based on different weight statuses and diabetes statuses?

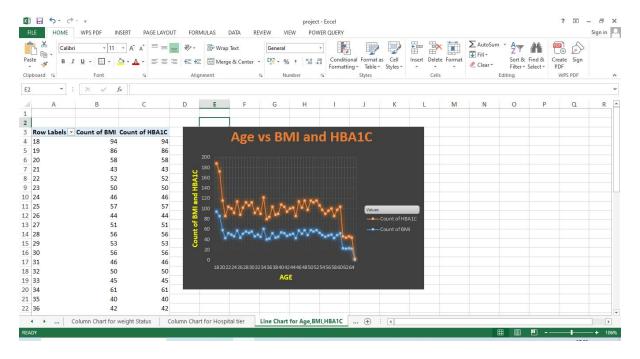


➤ Can you compare the average charges for each hospital tier within different states?

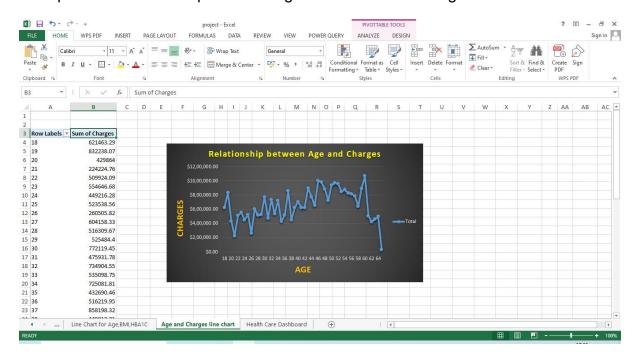


Analysis using Line/Scatter Plot:

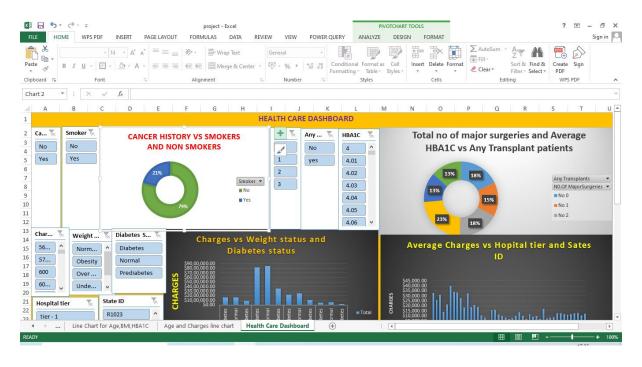
➤ Is there any correlation between age and both BMI and HbA1C in the dataset?

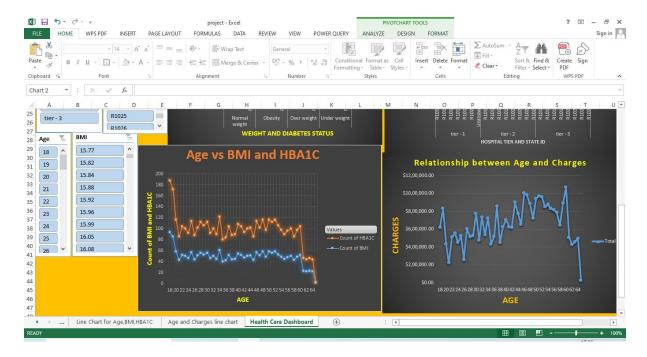


> Explore the relationship between age and healthcare charges.



Health Care Dashboard:





Conclusion:

Project management in healthcare is critical for the efficient delivery of medical services and the implementation of innovations that enhance patient care. By understanding and addressing the unique challenges of project management in healthcare, health professionals can lead projects that significantly improve healthcare outcomes and operational efficiencies.

As the healthcare landscape evolves, the role of project management will only grow in importance, driving forward the industry's ability to adapt and thrive in an increasingly complex environment.