Diagnosis of Impacting Factors on Hospital Readmission

Problem Statement and Objective

- Hospital readmission occurs when a patient who has been discharged is readmitted within a certain timeframe.
- High readmission rates for specific conditions are now seen as a quality indicator for hospitals and contribute to increased healthcare costs. Consequently, the Centers for Medicare & Medicaid Services (CMS) introduced the Hospital Readmissions Reduction Program.
- This program aims to enhance patient care quality and reduce healthcare expenses by imposing financial penalties on hospitals that exceed expected readmission rates for specific conditions.
- In 2011, American hospitals incurred over $41 billion in expenses due to diabetic patients being readmitted within 30 days of discharge.
- Identifying the factors that lead to higher readmission rates among these patients and predicting which patients are likely to be readmitted can help hospitals save millions of dollars while improving patient care.
- With this context, we utilised a dataset of diabetic patient records to diagnose the strongest factor causing hospital readmission in diabetic patients.

Dataset Description:

Race: Categories include Caucasian, Asian, African American, Hispanic, and other.
Gender: Categories include male, female
Age: Age grouped in 8 intervals: (20-30), ..., (90-100).
Admission type ID: Integer identifier corresponding to 5 distinct values (e.g., emergency, urgent, elective, newborn, other).
Discharge disposition ID: Integer identifier corresponding to 9 distinct values (e.g., discharged to home, expired, etc).
Admission source ID: Integer identifier corresponding to 6 distinct values (e.g., physician referral, emergency room, transfer from a hospital).
Time in hospital: Integer number of days between admission and discharge.
Number of lab procedures: Number of lab tests performed during the encounter.
Number of procedures: Number of procedures (other than lab tests) performed during the encounter.
Number of medications: Number of distinct generic names administered during the encounter.
Number of diagnoses: Number of diagnoses entered into the system.
Max glucose serum: Indicates the range of the glucose serum test result.
A1C result: Indicates the range of the A1C test result.
Change of medications: Indicates if there was a change in diabetic medications

Diabetes med: Indicates if there was any diabetic medication prescribed. Values: “yes” and “no”.
Readmitted: “Readmitted” or “Not Readmitted”
Service utilisation: A custom feature representing a composite measure of healthcare services utilized.
Numchange: A custom feature representing the number of changes in the patient's medication regimen (specific details not provided in the dataset description).
Level1_diag1: Primary diagnosis at level 1 (coded as the first three digits of ICD9).

Data Cleaning Operations:

1. Replace the values in “Admission type ID” as follows:
   { 1: “emergency”, 2: “urgent”, 3: “elective”, 5: “newborn”, 6: “other”}

   - Create a new calculated field as follows:

   ```
   CASE [Admission Type Id]
   WHEN 1 THEN "emergency"
   WHEN 2 THEN "urgent"
   WHEN 3 THEN "elective"
   WHEN 5 THEN "newborn"
   WHEN 6 THEN "other"
   ELSE "unknown"
   END
   ```

2. Replace the values “Admission source ID” as follows:
   { 1: “emergency”, 2: “referral”, 4: “transfer”, 6: “Scheduled”, 7: “Maternity”, 17: “Outpatient”}

   - Create a new calculated field as follows:

   ```
   CASE [Admission source ID]
   WHEN 1 THEN "emergency"
   WHEN 2 THEN "referral"
   WHEN 4 THEN "transfer"
   WHEN 6 THEN "Scheduled"
   WHEN 7 THEN "Maternity"
   WHEN 17 THEN "Outpatient"
   ELSE "unknown"
   END
   ```

3. Replace the null values in “Number of procedures” by zero.

    - Create a new calculated field as follows:
    - IFNULL([Number of procedures], 0)

4. Replace the null values in “Number of medications” by median.

    - Create a new calculated field as follows:
    - IFNULL([Num Medications], { FIXED : MEDIAN([Num Medications]) })

5. Rename column name “DiebetesMed” to “Prescribed_Diabetes_Medication”

    - Feature can be renamed in the data pane.

6. Replace the values in “Service_Utilization” as follows:
   { 0: “Never”, 1: “Once”, 2: “Twice”}

    - Create a new calculated field as follows:

   ```
   CASE [Service_Utilization]
   WHEN 0 THEN "Never"
   WHEN 1 THEN "Once"
   WHEN 2 THEN "Twice"
   ELSE "Other"
   END
   ```

7. Rename the column “level1_diag1” to “Category of Disease” . Replace the values as follows:
   { 0: “Circulatory”, 1: “Respiratory”, 2: “Digestive”, 3: “Digestive”, 4: “Injury”, 5: “Musculoskeletal”, 6:”Genitourinary”, 7: “Neoplasms”, 8: “Other”}

   Repeat the same procedure as mentioned above.

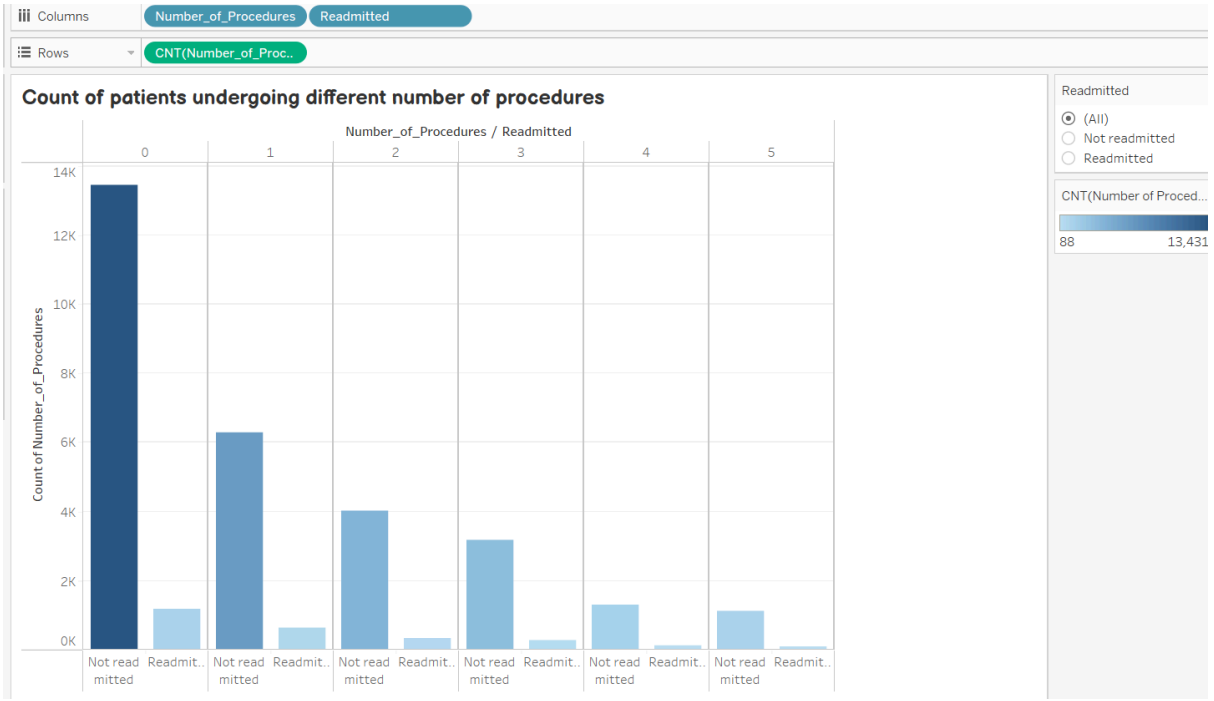
   ```
   CASE [Category of Disease]
   WHEN 0 THEN "Circulatory"
   WHEN 1 THEN "Respiratory"
   WHEN 2 THEN "Digestive"
   WHEN 3 THEN "Digestive"
   WHEN 4 THEN "Injury"
   WHEN 5 THEN "Musculoskeletal"
   WHEN 6 THEN "Genitourinary"
   ```

```
WHEN 7 THEN "Neoplasms"
WHEN 8 THEN "Other"
ELSE "Unknown"
END
```

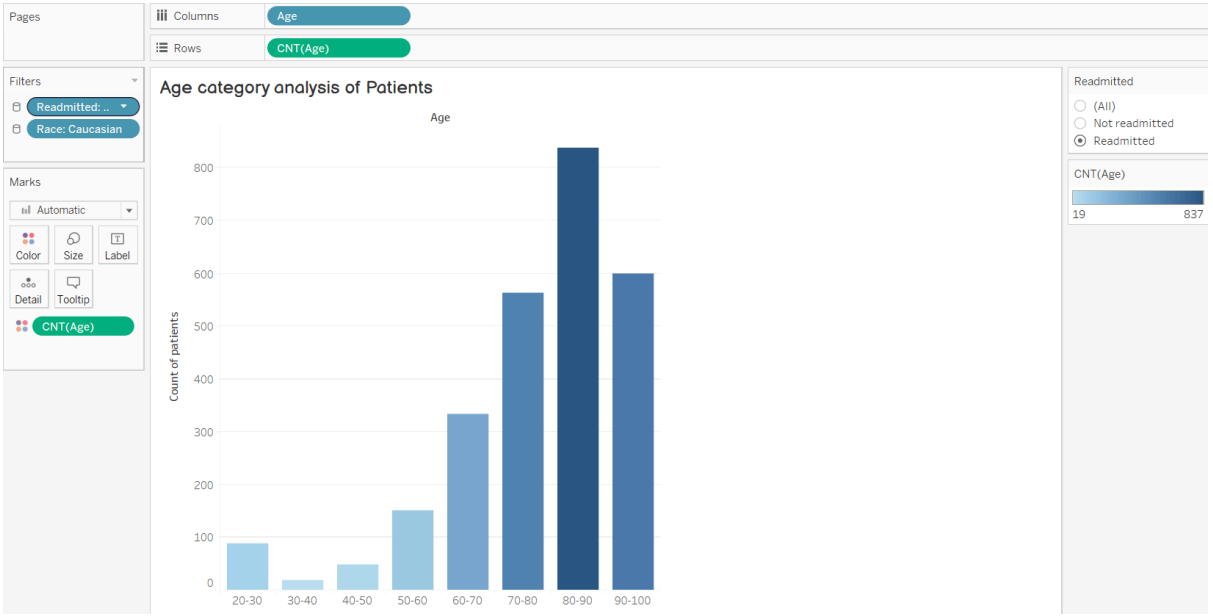
8. Check whether the features are correctly separated as attributes and dimensions. If not, relocate the features in their correct category.

Worksheets:

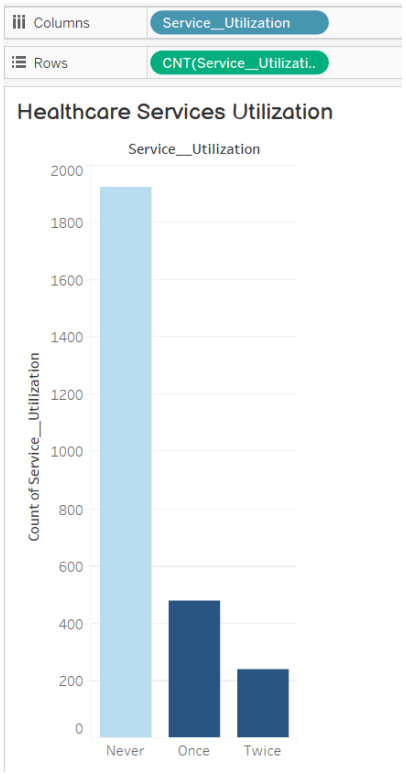
1. Visualize the count of patients undergoing different numbers of procedures in report 1.



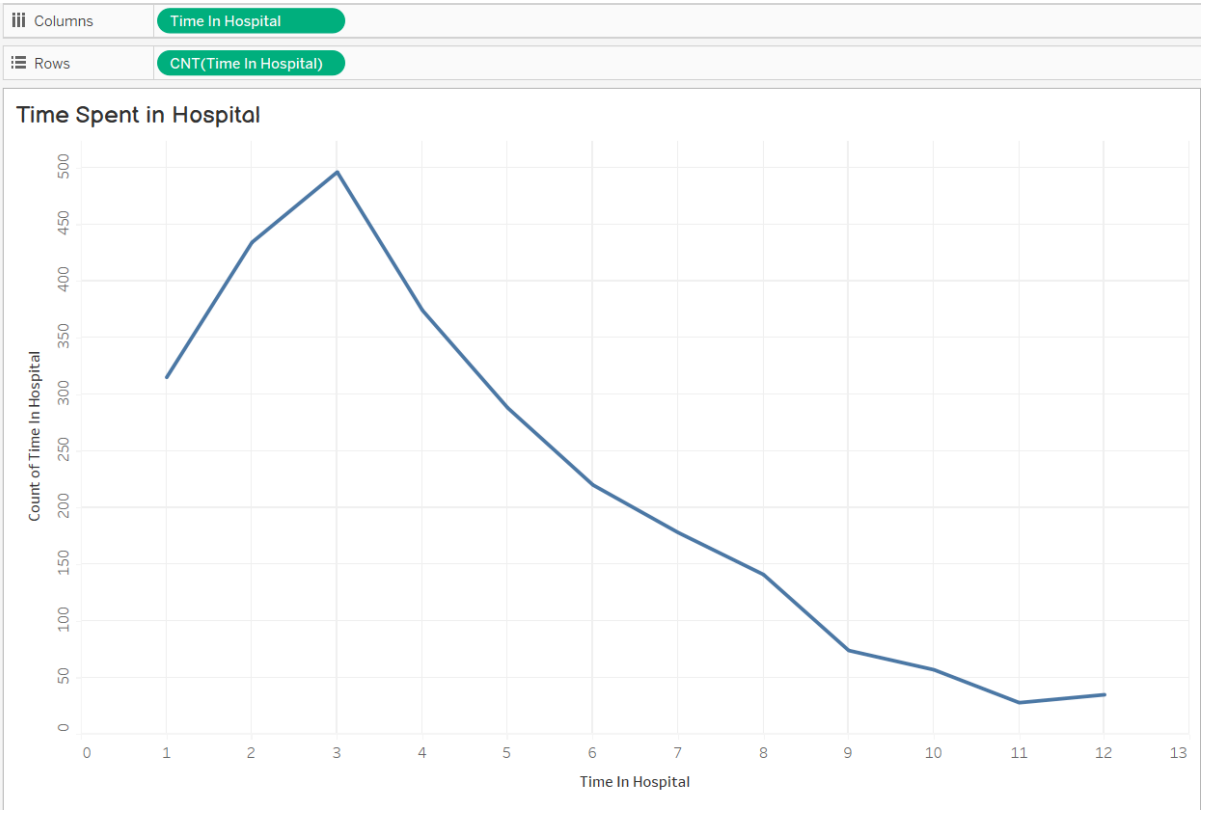
2. Perform age category analysis in worksheet 2 to find the count of patients from different age groups. Represent the age categories with different colours. Name the report with the appropriate title.



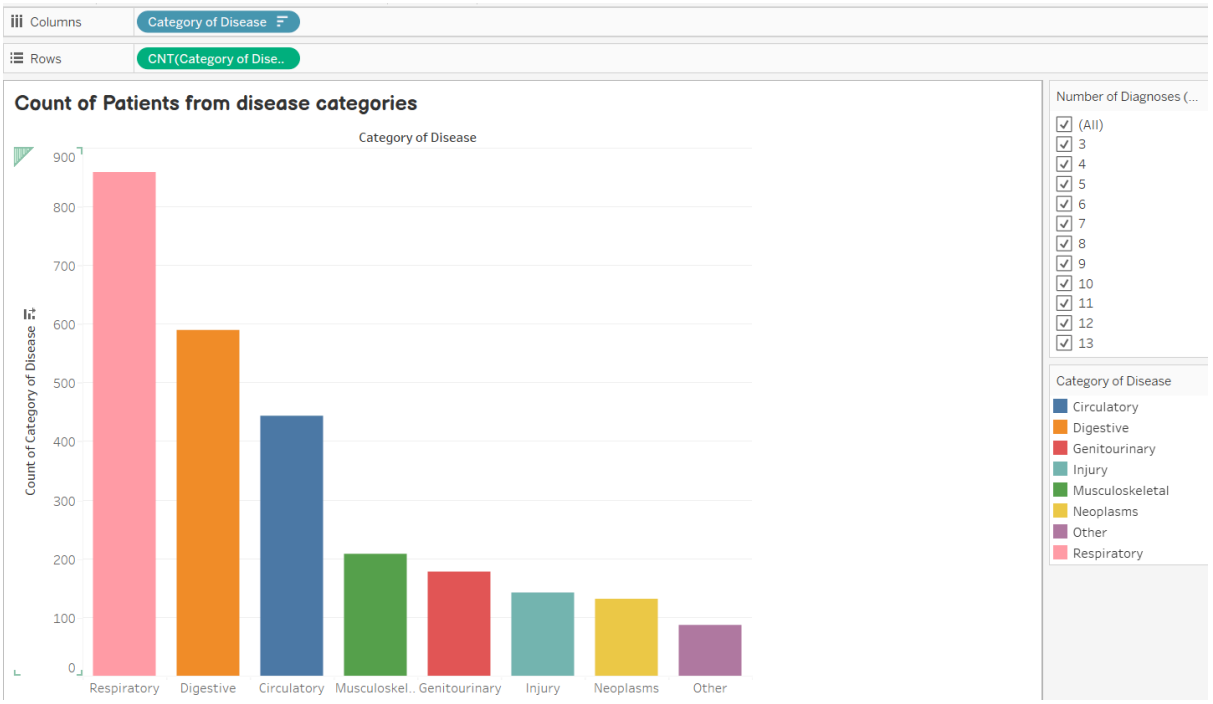
3. In worksheet 3, summarise the percentage of patients who have utilised healthcare services in the past using appropriate visualisation methods.



4. Visualize time spent in hospitals by patients in worksheet 4.

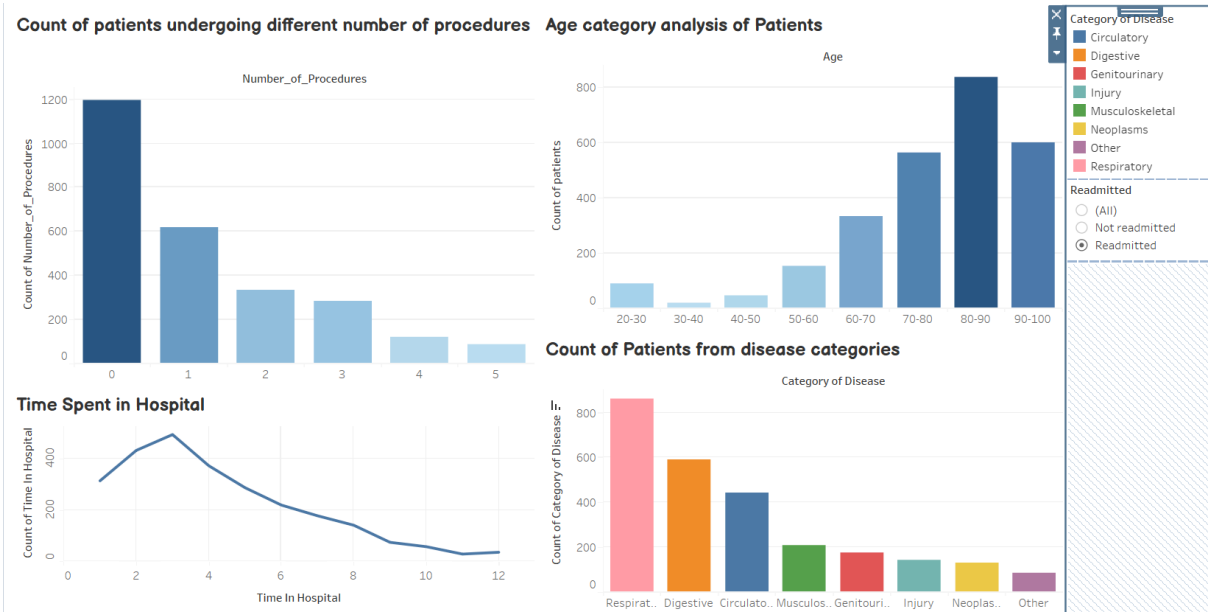


5. In worksheet 5, visualize the count of patients belonging to different categories of diseases.



Dashboard Creation:

1. Create a dashboard using all the 5 worksheets generated above.



2. Apply filter by “Race” category and “Readmitted” Features.

3. Summarize the impact of various factors on hospital readmission.