



# Tele-Health Data Analytics

*By Madan Thevar*

# Overview and Purpose

The Telehealth Services Data Analysis project utilizes Python, PostgreSQL, Power BI and ERD Builders to manage and analyze telehealth data, build queries and dashboards for data visualization. By uncovering insights into service utilization and patient satisfaction, the analysis helps healthcare providers enhance service delivery and make data-driven decisions. This project highlights the critical role of data management in optimizing telehealth services and improving patient outcomes.

# Key Findings

- **Service Utilization Patterns:** Analysis reveals peak usage times and most common service types.
- **Patient Satisfaction:** Identifies factors influencing patient satisfaction and areas for improvement.
- **Cost Analysis:** Breaks down healthcare costs by service type and insurance coverage.
- **Follow-Up Trends:** Examines the impact of follow-up requirements on healthcare costs and satisfaction.
- Most Common Health Diagnosis making frequent visits for Mental Health Diseases



# About the Dataset

- **Source:** The dataset includes telehealth service records from providers such as Teladoc Health and Amwell, with additional data web scraped from sources like SourceTelehealth and RemoteHealth.
- **Content:** Contains fields such as ServiceType, PatientID, SatisfactionScore, HealthcareCost, and FollowUpRequired.
- **Modification:** Fictional Patient IDs and other IDs were created to protect privacy. All other fields and values reflect genuine telehealth service records.
- **Data Cleaning:** The dataset was cleaned using Python to remove duplicates and ensure data integrity, providing accurate insights into service usage and patient satisfaction.

# SQL Queries Using PostGresQL




**Question 1: How many telehealth visits were conducted by gender?**

	<b>gender</b> character varying (10) 	<b>visitcount</b> bigint 
1	Other	29
2	Male	35
3	Female	36




**Question 2: What is the average satisfaction score for each service type?**

	<b>servicetype</b> character varying (100) 	<b>averagesatisfaction</b> numeric 
1	General Consultation	2.94
2	Mental Health	2.86
3	Chronic Disease Management	3.18
4	Follow-up Visit	2.62

**Question - 3: What is the average duration of visits for each gender and socioeconomic status, only for visits where the satisfaction score is greater than 3?**

	gender character varying (10) 	socioeconomicstatus character varying (50) 	averageduration numeric 
1	Female	High	41.20
2	Female	Low	29.25
3	Female	Medium	25.00
4	Male	High	25.00
5	Male	Low	29.17
6	Male	Medium	31.20
7	Other	High	32.71
8	Other	Low	38.00
9	Other	Medium	23.00

## Question - 4: How many patients had technical issues during their telehealth visits, grouped by primary diagnosis and insurance type?

	primarydiagnosis character varying (100) 	insurancetype character varying (50) 	patientswithtechnicalissues bigint 
1	Anxiety	Medicaid	4
2	Anxiety	Medicare	2
3	Anxiety	Private	4
4	Anxiety	Uninsured	2
5	Asthma	Medicaid	2
6	Asthma	Medicare	2
7	Asthma	Uninsured	2
8	Depression	Medicaid	2
9	Depression	Medicare	2
10	Depression	Private	5
11	Depression	Uninsured	2
12	Diabetes	Medicaid	1
13	Diabetes	Medicare	1
14	Diabetes	Private	2
15	Diabetes	Uninsured	4
16	Hypertension	Medicare	1
17	Hypertension	Private	2
18	Hypertension	Uninsured	3



**Question - 5) Determine the top 3 primary diagnoses with the highest average healthcare costs for patients aged 60 and above, and calculate the average satisfaction score and the percentage of visits that required follow-up for each diagnosis?**

	primarydiagnosis character varying (100) 🔒	avghealthcarecost text 🔒	avgsatisfactionscore numeric 🔒	followupvisitpercentage numeric 🔒
1	Depression	\$165.71	3.71	28.57
2	Anxiety	\$154.29	1.71	42.86
3	Hypertension	\$150.00	4.75	25.00

# ERD Builder Using dbdiagram.io Using a SQL Script

```
// Define the Patients table
Table Patients {
  PatientID INT [pk]
  Age INT
  Gender VARCHAR(10)
  Ethnicity VARCHAR(50)
  SocioeconomicStatus VARCHAR(50)
}

// Define the Visits table
Table Visits {
  VisitID INT [pk]
  PatientID INT [ref: > Patients.PatientID]
  TelehealthVisitDate DATE
  ServiceType VARCHAR(100)
  DurationOfVisit INT
  SatisfactionScore INT
  TechnicalIssueID INT [ref: > TechnicalIssueDetails.IssueID]
  FollowUpRequired BOOLEAN
  HealthcareCost NUMERIC
  InsuranceID INT [ref: > Insurance.InsuranceID]
  PrimaryDiagnosisID INT [ref: > Diagnoses.DiagnosisID]
  TelehealthPlatformID INT [ref: > TelehealthPlatforms.PlatformID]
}

// Define the Diagnoses table
Table Diagnoses {
  DiagnosisID INT [pk]
  PrimaryDiagnosis VARCHAR(100)
}

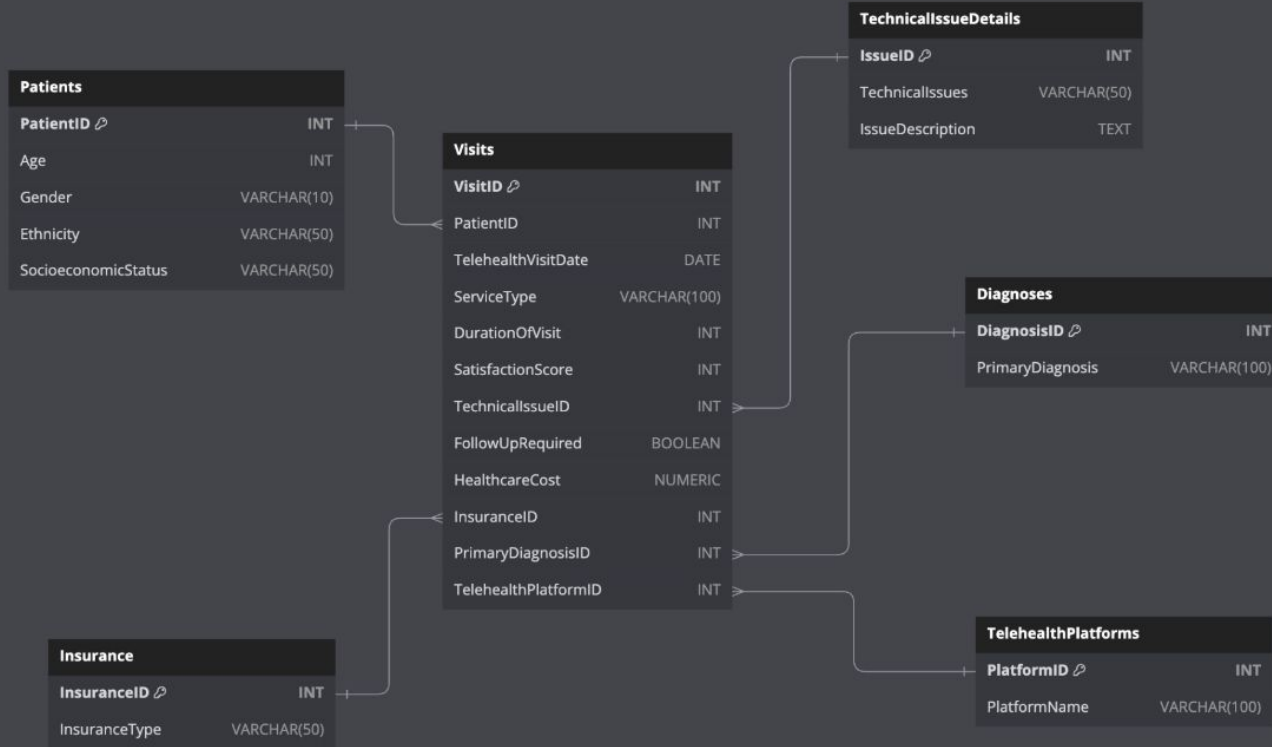
// Define the TelehealthPlatforms table
Table TelehealthPlatforms {
```

```
  PlatformID INT [pk]
  PlatformName VARCHAR(100)
}

// Define the Insurance table
Table Insurance {
  InsuranceID INT [pk]
  InsuranceType VARCHAR(50)
}

// Define the TechnicalIssueDetails table
Table TechnicalIssueDetails {
  IssueID INT [pk]
  TechnicalIssues VARCHAR(50)
  IssueDescription TEXT
}
```

# ERD Diagram



# Power-BI DAX Functions

- a) Used COUNTROWS Function (Total Visits =  
`COUNTROWS('telehealth_services_usage_extended_cleaned')`)
  
- b) To calculate Total Healthcare Costs (Total Healthcare Cost =  
`SUM('telehealth_services_usage_extended_cleaned'[Healthcare_Cost])`)
  
- c) To calculate Average Satisfaction Score (Average Satisfaction Score =  
`AVERAGE('telehealth_services_usage_extended_cleaned'[Satisfaction_Score])`)

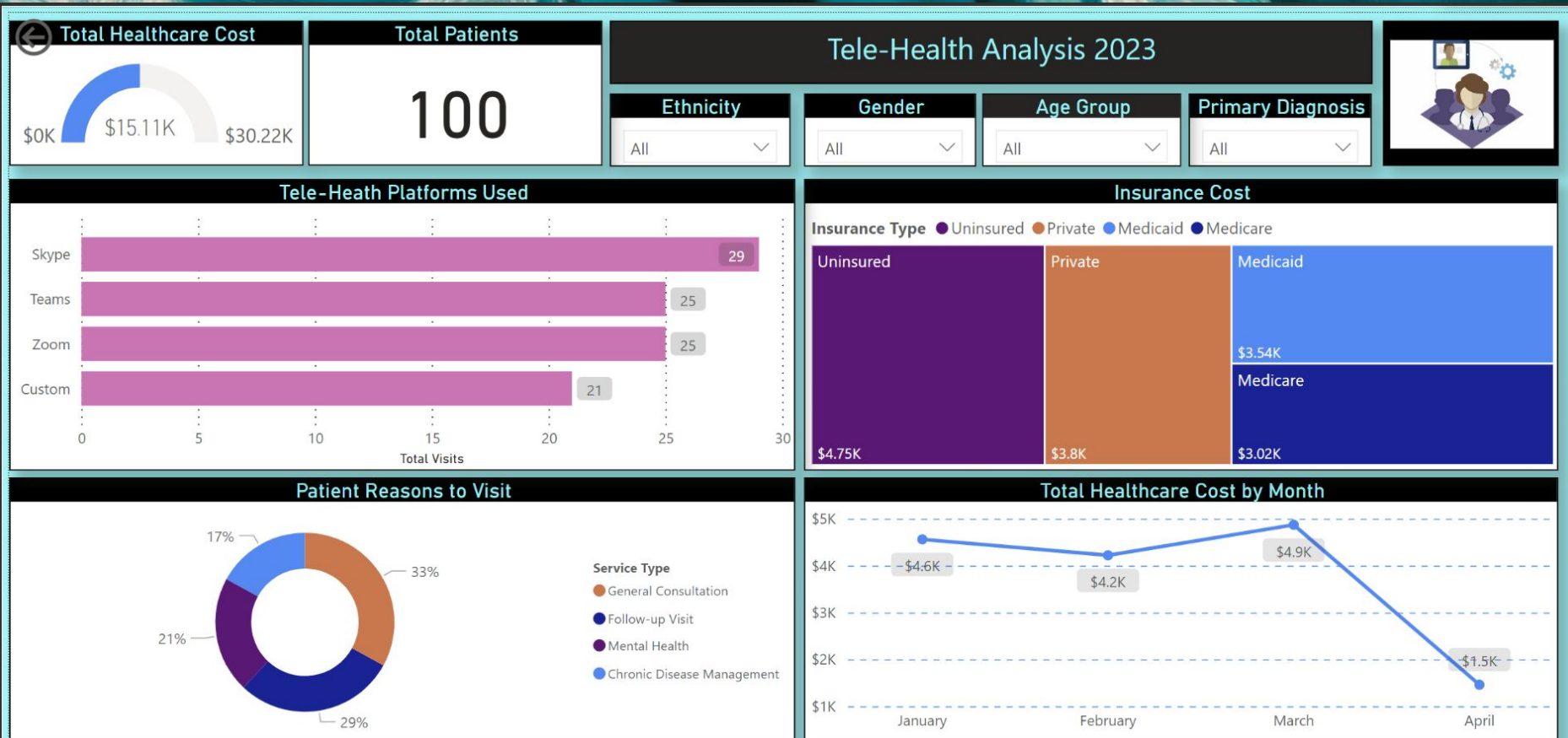
# More Power-BI DAX Functions

d) Used Distinct Count Function for Unique Patients (Unique Patients =  
`DISTINCTCOUNT('telehealth_services_usage_extended_cleaned'[Patient_ID])`)

e) Higher Satisfaction Flag using IF statement used (High Satisfaction Flag =  
`IF('telehealth_services_usage_extended_cleaned'[Satisfaction_Score] >= 4, "Yes", "No")`)



# Power-BI Interactive Dashboard



# Key Findings from the Analytics

## **Platform Utilization Insights:**

- Analyzing telehealth platform usage reveals which platforms are preferred, guiding resource allocation and partnership strategies to enhance service delivery based on user preferences.

## **Service Type Demand:**

- Understanding the distribution of visits by service type helps in effectively managing resources and staffing, ensuring high-demand services are well-supported and less-utilized services are appropriately promoted.

## **Satisfaction and Performance Metrics:**

- Monitoring satisfaction scores and performance trends enables targeted quality improvement initiatives, leading to higher patient satisfaction and better health outcomes by addressing specific areas of concern.

# Room for Improvements

## 1. **Enhanced Resource Allocation:**

- Insights from platform utilization trends help allocate resources more efficiently, ensuring popular platforms receive adequate support and training, improving the overall user experience.

## 2. **Targeted Marketing and Outreach:**

- Identifying high-demand and underutilized services enables targeted marketing efforts, optimizing promotional strategies and managing service demand more effectively.

## 3. **Continuous Quality Improvement:**

- Regularly analyzing satisfaction scores and performance metrics facilitates continuous quality improvement, ensuring high standards of care and better patient outcomes through data-driven decision-making.