Scenario 1: Medical Readmission

In the medical industry, readmission of patients is such a problem that an external organization, the Centers for Medicare and Medicaid Services (CMS), penalizes hospitals for excessive readmissions. When it comes to readmission penalties, studies show that many hospitals are overconfident and underprepared. The percentage of hospitals penalized for readmissions has increased each year since CMS began imposing penalties, and according to the CMS reporting, as much as 78 percent of hospitals were fined in fiscal year 2015. However, three-quarters of hospitals feel confident in their ability to reduce readmissions, and only 55 percent of them anticipate receiving a penalty this year. Given the historical trend and the addition of COPD and Hip and Knee replacement to the list of medical conditions measured, the percentage of hospitals penalized will likely be much higher than 55 percent. Additionally, although hospitals are applying various reduction strategies, fewer than 1 in 5 utilize technology that is specific to reducing their readmissions, so they may not be doing all that they can.

You are a data analyst on a team of analysts for a large medical hospital chain with patients in almost every state in the United States. You have been provided with a data set and asked to investigate the extent to which readmission is a problem for this chain of hospitals. In addition to your analysis, you have asked to build a data dashboard to enable executive leaders to explore the data, identify trends, and compare key metrics.

Note that, although the original reason for hospitalization is not provided in the data, the data does present the opportunity to predict readmission based on other conditions and factors of the patient.

Audience Background

Executive Leaders

You will build a Tableau data dashboard to be used by the executive leaders of your organization to guide their decision-making. As described below, each leader has a specific focus and area of expertise in the organization, but they do not have a technical data analysis background like yours. As a result, the dashboard you build must be easy to navigate and should present broad and understandable insights from the data that are relevant to their perspective.

Senior Vice President of Hospital Operations (SVP)

- This SVP oversees operations across all hospital locations.
- They are responsible for developing new initiatives to improve patient outcomes based on observed trends.
- The SVP is interested in broad categorization of patient treatments and outcomes as a function of demographics, and how these trends play out across regions.

Vice President of Research (VP)

- A key focus of this VP is to oversee research initiatives to identify patterns in patient care and drive improvements in patient outcomes through strategic initiatives.

Panel of Regional Vice Presidents (Regional VPs)

- This panel consists of every Regional VP across the organization.
- Each Regional VP is responsible for executing policies and managing operations in conjunction with the SVP.

Data Analytics Peers

You work on a team of data analysts with a technical background similar to yours. Members of this team are research-minded and have a specific interest in how the design, methodology, and results of a data analysis can be translated to specific business insights. Your peers are eager to hear you tell an engaging story about the data and offer actionable recommendations that are backed by evidence.

Data File being used:
medical_clean.csv

Data Dictionary:

The data set consists of the following categories of information:

- patients who are readmitted to the hospital within a month of release (the "ReAdmis" column)
- patient medical conditions (high blood pressure, stroke, obesity, arthritis, diabetes, etc.)
- patient information (service they received while hospitalized, days in hospital, type of initial admission, etc.)
- patient demographic information (gender, age, job, education level, etc.)

The data set consists of 10,000 customers and 50 columns/variables:

- **CaseOrder:** A placeholder variable to preserve the original order of the raw data file
- **Customer_id:** Unique patient ID
- **Interaction, UID:** Unique IDs related to patient transactions, procedures, and admissions

The following variables represent customer demographic data:

- **City:** Patient city of residence as listed on the billing statement
- **State:** Patient state of residence as listed on the billing statement
- **County:** Patient county of residence as listed on the billing statement
- **Zip:** Patient zip code of residence as listed on the billing statement
- **Lat, Lng:** GPS coordinates of patient residence as listed on the billing statement
- **Population:** Population within a mile radius of patient, based on census data
- **Area:** Area type (rural, urban, suburban), based on unofficial census data
- **TimeZone:** Time zone of patient residence based on patient's sign-up information
- **Job:** Job of the patient (or primary insurance holder) as reported in the admissions information

    - **Children:** Number of children in the patient's household as reported in the admissions information (might not be children of the patient)
    - **Age:** Age of the patient as reported in admissions information
    - **Income:** Annual income of the patient (or primary insurance holder) as reported at time of admission
    - **Marital:** Marital status of the patient (or primary insurance holder) as reported on admission information
    - **Gender:** Customer self-identification as male, female, or nonbinary

- **ReAdmis:** Whether the patient was readmitted within a month of release or not (yes, no)
- **VitD_levels:** The patient's vitamin D levels as measured in ng/mL
- **Doc_visits:** Number of times the primary physician visited the patient during the initial hospitalization
- **Full_meals_eaten:** Number of full meals the patient ate while hospitalized (partial meals count as 0, and some patients had more than three meals in a day if requested)
- **VitD_supp:** The number of times that vitamin D supplements were administered to the patient
- **Soft_drink:** Whether the patient habitually drinks three or more sodas in a day (yes, no)
- **Initial_admin:** The means by which the patient was admitted into the hospital initially (emergency admission, elective admission, observation)
- **HighBlood:** Whether the patient has high blood pressure (yes, no)
- **Stroke**: Whether the patient has had a stroke (yes, no)
- **Complication_risk:** Level of complication risk for the patient as assessed by a primary patient assessment (high, medium, low)
- **Overweight:** Whether the patient is considered overweight based on age, gender, and height (yes, no)
- **Arthritis:** Whether the patient has arthritis (yes, no)
- **Diabetes:** Whether the patient has diabetes (yes, no)
- **Hyperlipidemia:** Whether the patient has hyperlipidemia (yes, no)
- **BackPain:** Whether the patient has chronic back pain (yes, no)
- **Anxiety:** Whether the patient has an anxiety disorder (yes, no)
- **Allergic_rhinitis:** Whether the patient has allergic rhinitis (yes, no)
- **Reflux_esophagitis:** Whether the patient has reflux esophagitis (yes, no)
- **Asthma:** Whether the patient has asthma (yes, no)
- **Services:** Primary service the patient received while hospitalized (blood work, intravenous, CT scan, MRI) (Note: the patient may have received more services, but only the primary service is reported)
- **Initial_days:** The number of days the patient stayed in the hospital during the initial visit
- **TotalCharge:** The amount charged to the patient daily. This value reflects an average per patient based on the total charge divided by the number of days hospitalized. This amount reflects the typical charges billed to patients, not including specialized treatments.

- **Additional_charges:** The average amount charged to the patient for miscellaneous procedures, treatments, medicines, anesthesiology, etc.

The following variables represent responses to an eight-question survey asking customers to rate the importance of various factors/surfaces on a scale of 1 to 8 (1 = most important, 8 = least important)

  - **Item1:** Timely admission
  - **Item2:** Timely treatment
  - **Item3:** Timely visits
  - **Item4:** Reliability
  - **Item5:** Options
  - **Item6:** Hours of treatment
  - **Item7:** Courteous staff
  - **Item8:** Evidence of active listening from doctor