**Stat 109 Project – Spring 2023**

**Impact of Social Determinant of Health Drivers on Hospital Quality Measures:**

**Team Members:** Erin Rebholz, Nadia Zafar, Imran Naskani, Max Yanover

**Project Questions:**

1. What influence does census Social Determinant of Health Data have on hospital quality and outcomes data?
2. Which factors are most important in creating this prediction?

**Motivation:**

In this paper, we will explore how United States Census demographic data and Medicare Quality reporting data can be used to predict health quality and account for regional disparities in care across the United States. Diagram

Description automatically generated

Increasingly, the United States has acknowledged that social determinants of health factors, such as education access, economic stability, community context and health care access, play a role in disparate outcomes within the United States health care system. The Healthy People 2030 program defines Social determinants of health (SDOH) as , “the conditions in the environments where people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks.” **[Figure 1] [1].**

For our outcome of interest, we will use a hospital’s Overall Star Rating, a value calculated by the Centers for Medicare and Medicaid Services (CMS) based on measures reported by hospitals.

The Overall Star Rating is a five-point scale based on a composite of measures, including mortality, safety, readmission, patient experience, and timely & effective care. The Overall Star Rating uses a weighted average across measure groups. Given the breadth of measures, we chose to use the Overall Star Rating as a proxy for quality of care received by patients **[Figure 2]**.

Table

Description automatically generated

In our project, we will prepare a multivariable model to predict hospital quality using hospital features such as census region, urban/rural location, percentage of non-white population, average education attainment, and median income. We will explore various approaches and determine which model produces the most accurate and interpretable results.

**Data Sources and Descriptions:**

**CMS Hospital Quality Data Set:**  The Centers for Medicare & Medicaid Services (CMS) manages a program of quality metric reporting where hospitals submit standardized metrics pertaining to hospital quality. CMS uses these datasets as part of its quality management and payment programs and publicizes the data sets to create transparency around health care quality within the United States:

**Overview:** <https://data.cms.gov/provider-data/topics/hospitals>

**Data Dictionary and Data Download:** [Hospitals | Provider Data Catalog (cms.gov)](https://data.cms.gov/provider-data/topics/hospitals)

**Background on Overall Star Rating calculation:** <https://data.cms.gov/provider-data/topics/hospitals/overall-hospital-quality-star-rating/>

**US Census, Zip Level Demographic Data:** Every 10 years, the United States government collects demographic information about its population as part of the US Census. The demographic data collected is used in this project to understand SDoH factors such as educational attainment, income, age of population and proportion of population which is in minority demographics. The data is available at the zip code level and is mapped to the Hospital Quality Data set by the hospital’s zip code:

Data available to download from: <https://data.census.gov/cedsci/table?g=0100000US%248600000>

**Notes on Data Cleansing:** Hospitals where zip codes don’t match up to census data will be omitted from the analysis. In practice, hospitals sometimes have their own zip codes and do not have census data available. We will focus only on entities where the Overall Star Rating is available and narrow are results to Short Term Acute Care and Critical Access Health facilities. Other facility types such as DoD/VA, Psychiatric, Rehabilitation and Cancer Centers will be excluded from the analysis, as they are subject to different types of quality programs and not well covered by the Overall Star Rating metric. Critical Access facilities have up to 25 beds, may provide a reduced range of services and serve primarily rural areas. We will compare performance on Overall Star Ratings between Short Term Acute and Critical Access facilities.

**In R:** need to drop the NA values, filter

**Exploration of Data:**

<Insert findings of EDA>

Overall Star Ratings by Census Region/Division- Note regional disparities in healthcare, especially within the south.

Table of average value for SdOH factor by STAR Rating Measure (1-5).

Maybe a prop test result?

Average STAR Rating Measure for STC (Acute) vs. (CAH) – See if there is a difference.

**Analytic Approach:**

<Discuss Approach>

One-Hot Encode (Make Binary) any non-continuous variables. (e.g. census region)

Question regarding appropriate model to use for integer type variables (1-5).

Logistic vs. Linear models?

Decision Tree Model has the advantage of working for both continuous and categorical variables.

**Discussion of Results:**

<Insert regression results>

**Conclusions:**

<Highlight any key findings>

**List of References:**

**[1]** Healthy People 2030, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. Retrieved March 15, 2023, from <https://health.gov/healthypeople/objectives-and-data/social-determinants-health>

**[2]** <https://www.cms.gov/about-cms/agencyAc-information/omh/health-equity-programs/cms-framework-for-health-equity>

**[3]** https://data.cms.gov/provider-data/topics/hospitals/overall-hospital-quality-star-rating/

**[4]** Braveman P, Gottlieb L. The Social Determinants of Health: It’s Time to Consider the Causes of the Causes. Public Health Reports. 2014;129(1\_suppl2):19-31. doi: [10.1177/00333549141291S206](https://doi.org/10.1177/00333549141291S206)

**Data Dictionary:**

|  |  |  |
| --- | --- | --- |
| **Field Name** | **Field Description** | **Data Source** |
| Facility ID | Medicare Provider Number, can be used to link across CMS data sets. Unique to a hosptial | Medicare Quality Reporting - Hospital General File |
| Facility Name | Facility Name | Medicare Quality Reporting - Hospital General File |
| Address | Address | Medicare Quality Reporting - Hospital General File |
| City | City | Medicare Quality Reporting - Hospital General File |
| State | State | Medicare Quality Reporting - Hospital General File |
| ZIP Code | ZIP Code | Medicare Quality Reporting - Hospital General File |
| County Name | County Name | Medicare Quality Reporting - Hospital General File |
| Hospital Type | Acute Care Hospitals, Critical Access Hospitals, Children’s, Psychiatric, etc | Medicare Quality Reporting - Hospital General File |
| Hospital Ownership | Various types of government owned, privately held and non-profit entity types | Medicare Quality Reporting - Hospital General File |
| Hospital overall rating | A five point scale based on a composite of measures, including mortality, safety, readmission, patient experience, and timely & effective care. 5 is best, 1 is lowest rating. Value not available for all providers. | Medicare Quality Reporting - Hospital General File |
| census\_region | Mapped by state to census region | Aligns to Census Region |
| census\_division | Mapped by state to census division | Aligns to Census Division |
| median\_age | Median age in hospital zip | Census Demographic Info |
| per\_white\_non\_hisp | Percentage of white, non-hispanic population in hospital zip | Census Demographic Info |
| med\_inc\_15plus\_12mo | Median income in last 12 months for individuals 15 or older in hospital zip | Census Demographic Info |
| per\_below\_poverty | Percent of population below poverty in hospital zip | Census Demographic Info |
| per\_college\_grad\_deg\_25\_plus | Percent of population who are 25 or older and have a college or gradulate level degreee, within hospital zip | Census Demographic Info |

**Bibtex File Start:**

**List of References:**

@article{ <https://www.cms.gov/about-cms/agencyAc-information/omh/health-equity-programs/cms-framework-for-health-equity>,   
author = {Centers for Medicare and Medicaid Services},

title = {

}

@article{doi:10.1177/00333549141291S206,

author = {Paula Braveman and Laura Gottlieb},

title ={The Social Determinants of Health: It's Time to Consider the Causes of the Causes},

journal = {Public Health Reports},

volume = {129},

number = {1\\_suppl2},

pages = {19-31},

year = {2014},

doi = {10.1177/00333549141291S206},

note ={PMID: 24385661},

URL = {

https://doi.org/10.1177/00333549141291S206

},

eprint = {

https://doi.org/10.1177/00333549141291S206

}

,

abstract = { During the past two decades, the public health community's attention has been drawn increasingly to the social determinants of health (SDH)—the factors apart from medical care that can be influenced by social policies and shape health in powerful ways. We use “medical care” rather than “health care” to refer to clinical services, to avoid potential confusion between “health” and “health care.” The World Health Organization's Commission on the Social Determinants of Health has defined SDH as “the conditions in which people are born, grow, live, work and age” and “the fundamental drivers of these conditions.” The term “social determinants” often evokes factors such as health-related features of neighborhoods (e g., walkability, recreational areas, and accessibility of healthful foods), which can influence health-related behaviors. Evidence has accumulated, however, pointing to socioeconomic factors such as income, wealth, and education as the fundamental causes of a wide range of health outcomes This article broadly reviews some of the knowledge accumulated to date that highlights the importance of social—and particularly socioeconomic—factors in shaping health, and plausible pathways and biological mechanisms that may explain their effects. We also discuss challenges to advancing this knowledge and how they might be overcome. }

}