ANALYSIS OF CALIFORNIA HOSPITAL PERFORMANCE RATINGS

Mukesh Rajmohan
mrajmoha@gmu.edu
G01456275

AIT 580

I. ABSTRACT

This proposed study examines the performance ratings of hospitals in California, aiming to address the problem of evaluating the quality of healthcare services provided by these institutions. The study utilizes a comprehensive dataset that includes information on the year, county, hospital name, and system affiliation. Performance measures such as the number of adverse events, number of cases, risk-adjusted rates, and hospital ratings are analyzed. The research employs a systematic methodology to assess the performance of hospitals across various parameters. The results provide valuable insights into the quality of care delivered by hospitals in California, highlighting areas of strength and areas that require improvement. Based on the findings, conclusions are drawn regarding the overall performance of hospitals and recommendations are made to enhance the quality of healthcare services. This research contributes to the understanding of hospital performance and serves as a valuable resource for healthcare stakeholders, policymakers, and the public in making informed decisions to improve healthcare outcomes.

II. INTRODUCTION

California boasts a diverse and extensive healthcare system, marked by a mix of public and private initiatives. With a population exceeding 39 million, the state prioritizes accessible and innovative healthcare services. Home to renowned medical institutions, California leads in medical research and technology. The Affordable Care Act's implementation significantly expanded coverage, emphasizing preventative care. Challenges persist, including healthcare disparities and workforce shortages, but ongoing efforts aim to address these issues. Overall, California's healthcare landscape reflects a dynamic interplay of resources, policies, and a commitment to advancing the well-being of its residents.

The performance of California's healthcare institutions is generally high, with a focus on cutting-edge medical research, advanced technologies, and quality patient care. The state has been a pioneer in adopting innovative healthcare practices, promoting preventive care, and addressing public health issues. However, challenges such as healthcare disparities and access gaps persist, particularly in underserved communities.

The impact of California's healthcare system on the population is substantial. Access to quality healthcare services contributes to overall well-being and plays a crucial role in preventive care and early diagnosis. The state's emphasis on research and development has led to medical breakthroughs and advancements that benefit not only Californians but also have a ripple effect on global healthcare practices. Despite ongoing challenges, California's healthcare system plays a pivotal role in promoting public health and striving for equitable access to medical resources across its diverse population.

This study focuses on healthcare in the state of California, known for its large population and diverse healthcare landscape. California has been a popular destination for residents and tourists alike, with a significant demand for healthcare services. It is essential to educate the public about the quality of healthcare facilities in the state. By utilizing healthcare performance data from California, this paper aims to explore the following research questions:

- How does the number of adverse cases for specific medical conditions vary between different hospitals in California?
- What is the relationship between hospital performance and geographical location, and how might this information inform healthcare resource allocation?
- Have hospital performance measure for each report gotten better or worse over time, and if they have, what reasons could explain this?

The analysis and results obtained from addressing these research questions will be discussed in detail within the subsequent sections of this report, providing valuable insights into healthcare quality in California.

III. LITERATURE REVIEW

The case study titled "Association of Nurse Staffing with Patient Outcomes in California Hospitals" by Heather L Tubbs-Cooley, Rita H Pickler, Janet B Younger, and Barbara A Mark explores the relationship between nurse staffing levels and patient outcomes in California hospitals. The study investigates how the number of nurses on duty influences various aspects of patient care, including quality and safety. By analyzing data from California hospitals, the authors aim to identify associations between nurse staffing ratios and patient outcomes, shedding light on the crucial role of nursing staff in healthcare delivery. The findings contribute valuable insights for healthcare administrators, policymakers, and practitioners in optimizing staffing levels to enhance patient care and overall hospital performance. This research is instrumental in informing evidence-based strategies for improving the quality of healthcare services through appropriate nurse staffing.

The paper titled "Association Between Hospital Value-Based Purchasing Performance and Hospital Characteristics" by Jose F Figueroa, Yusuke Tsugawa, Jie Zheng, E John Orav, and Ashish K Jha investigates the relationship between hospital characteristics and performance under the Hospital Value-Based Purchasing (HVBP) program. The study explores how specific attributes of hospitals are associated with their performance in value-based purchasing initiatives, which link Medicare reimbursement to the quality of care provided. By analyzing a range of hospital characteristics, the authors aim to identify factors influencing a hospital's success in value-based purchasing programs, offering insights that can guide healthcare policymakers and administrators. This research is crucial for understanding the broader implications of value-based purchasing and informing strategies to enhance the quality of care in the context of evolving healthcare payment models.

The paper titled "Association of Hospital Prices for Coronary Artery Bypass Grafting with Hospital Quality and Reimbursement" by Zirui Song, Yunan Ji, Dana G Safran, and Michael E Chernew investigates the correlation between hospital prices for coronary artery bypass grafting (CABG), hospital quality, and reimbursement. The study aims to understand how the pricing of CABG procedures is associated with both the quality of care provided by hospitals and the reimbursement they receive. By examining these interconnections, the authors seek to shed light on the complex dynamics between financial incentives, healthcare quality, and economic considerations. The findings have implications for healthcare policy and reimbursement models, offering insights into the relationship between the cost of procedures, the quality of healthcare services, and the financial landscape of hospitals in the context of coronary artery bypass grafting.

IV. DATASET

The dataset provides detailed information about hospital performance ratings in California. It includes data points such as the year in which the ratings were recorded, the county where the hospital is located, the name of the hospital, its unique OSHPDID identifier, the health system it belongs to, and the type of report associated with the performance ratings.

Additionally, the dataset contains performance measures for different medical conditions or procedures, such as acute myocardial infarction (AMI), acute stroke (including hemorrhagic, ischemic, and subarachnoid types), carotid endarterectomy, gastrointestinal hemorrhage, heart failure, hip fracture, percutaneous coronary intervention (PCI), pancreatic resection, and pneumonia.

For each performance measure, the dataset provides the number of adverse events, the number of cases, and the risk-adjusted rate. It also includes hospital ratings, longitude, and latitude coordinates for geolocation purposes, as well as additional information such as adverse event rates, latitude, and longitude buckets for spatial grouping.

This dataset can be used to analyze and compare hospital performance across different categories and locations in California, helping to assess the quality of healthcare services provided by various hospitals in the state.

Field Title	Field Name	Data Type	Description
Year	Year	*Interval*	Year of Discharge
County	County	*Nominal*	County in which the hospital is located
Hospital	Hospital	*Nominal*	Name of hospital reporting
OSHPDID	OSHPDID	*Nominal*	Hospital OSHPD ID
System	System	*Nominal*	Hospital Health System
Type of Report	Type of Report	*Nominal*	Inpatient Mortality Indicators (IMI) Coronary Artery Bypass Graft (CABG) Elective Percutaneous Coronary Intervention (PCI)
Performance Measure	Performance Measure	*Nominal*	Inpatient mortality for AAA Repair Endo Unruptured, AAA Repair Open Unruptured, AMI, Acute Stroke, Acute Stroke Hemorrhagic, Acute Stroke Ischemic, Acute Stroke Subarachnoid, Carotid Endarterectomy, GI Hemorrhage, Heart Failure, Hip Fracture, PCI, Pancreatic Resection, and Pneumonia. Elective PCI Emergency CABG, Elective PCI Mortality, Elective PCI Stroke. Isolated CABG Operative Mortality
Number of Adverse Events	# of Adverse Events	*Ratio*	Hospital adverse events
Number of Cases	# of Cases	*Ratio*	Number of patients that had this medical procedure or condition in this hospital
Risk-adjusted Rate	Hospital Risk-adjusted Adverse Events Rate (%)	*Ratio*	Hospital Risk-adjusted Adverse Events Rate (%)
Hospital Ratings	Hospital Ratings	*Ordinal*	Comparison rating based on a 95% Confidence Interval (CI). If a hospital's upper CI is less than the statewide observed rate, it is designated as performing "better" than the average hospital. If a hospital's lower CI is greater than the state rate, it is designated as performing "worse" than the average state hospital.
Location	Longitude/Latitude	*Interval*	Longitude and Latitude coordinate of hospital

Figure 4.1. Data Dictionary

V. METHODS AND TOOLS

In this research project, a robust methodology is employed, integrating various methods and tools to ensure comprehensive data exploration and analysis. The initial phase involves meticulous data cleaning, executed through Python on Jupyter notebook, ensuring the integrity and quality of the dataset. Subsequently, SQL is employed for structured data analysis, unraveling patterns and insights within the information. The research extends its analytical capabilities with the incorporation of R, enabling advanced data analysis and visualization.

This synergistic approach harnesses the strengths of each tool, fostering a holistic understanding of the data and facilitating informed conclusions. The seamless integration of Python, SQL, and R provides a powerful framework for rigorous investigation and discovery.

The Jupyter notebook loads hospital quality data for California from a CSV file into a Pandas data frame for cleaning and analysis. The preliminary examination involves printing unique values for each column to detect missing values, inconsistent capitalization formats, non-standardized data types, outliers, and absent latitude/longitude information for certain hospitals.

To address these issues, a series of data cleaning steps are implemented. Missing numeric values are filled with placeholder zeros, text columns are standardized to uppercase using the **apply ()** and mapping functions, and outlier values are handled appropriately. The crucial step involves geocoding missing hospital locations by leveraging the Google Maps API. The **geocode ()** method is utilized to successfully resolve latitude and longitude coordinates for nearly all instances where this information is missing. This enhances the completeness of the dataset.

By combining hospital name and county in the address string, the geocoding accuracy is notably improved compared to using either the name or county individually. This approach significantly enhances the API's ability to successfully resolve almost all missing locations. The API integration plays a pivotal role in completing critical spatial data, thereby enabling subsequent geographic analysis with the cleaned dataset. The filled latitude and longitude columns become instrumental in constructing later variables for latitude and longitude bucketing.

With the cleaned location data in place, new derived columns are created, such as an adverse event percentage rate and categorical latitude/longitude buckets for subsequent geographic analysis. Finally, the transformed data frame is written out to a new CSV file, filtering only the columns essential for preservation. To summarize, the raw hospital data undergoes systematic cleaning, enrichment, standardization, and is then exported for downstream analysis. The geocoding process, facilitated by the Google Maps API, stands out as a critical step in filling gaps in spatial information.

VI. RESULTS

The California hospital performance dataset offers valuable quantitative insights, amenable to analysis through SQL and R programming. SQL queries will efficiently extract and manipulate relevant data subsets, unveiling hidden trends. R's statistical tools will then enable robust analysis, employing tests and models to investigate correlations, causal links, and temporal changes. Visualization in R will communicate findings effectively, shedding light on correlations, geographic patterns, and cross-state comparisons. In essence, the integrated use of SQL and R promises impactful discoveries from this extensive hospital dataset, contributing significantly to healthcare research and decision-making.

	Hospital	# of Cases	Risk-adjusted Rate
1	CEDARS-SINAI MEDICAL CENTER	1,890	2
2	CEDARS-SINAI MEDICAL CENTER	1,850	2.2
3	CEDARS-SINAI MEDICAL CENTER	1,698	2.3
4	CEDARS-SINAI MEDICAL CENTER	1,578	4
5	CEDARS-SINAI MEDICAL CENTER	1,513	3.6
6	KAISER FOUNDATION HOSPITAL – SUNSET	1,485	3.4
7	COMMUNITY REGIONAL MEDICAL CENTER – FRESNO	1,462	5.6
8	CENTINELA HOSPITAL MEDICAL CENTER	1,330	0.4
9	COMMUNITY REGIONAL MEDICAL CENTER – FRESNO	1,326	3.5
10	KAISER FOUNDATION HOSPITAL – SUNSET	1,310	3.8
11	KAISER FOUNDATION HOSPITAL – SUNSET	1,305	3.4
12	SAINT AGNES MEDICAL CENTER	1,302	3.1
13	LOS ANGELES COUNTY/UNIVERSITY OF SOUTHERN CALIFORNIA	1,299	1.4
14	SAINT AGNES MEDICAL CENTER	1,299	3.6
15	COMMUNITY REGIONAL MEDICAL CENTER – FRESNO	1,242	4.7
16	SCRIPPS MERCY HOSPITAL	1,227	1.9
17	KAISER FOUNDATION HOSPITAL – SUNSET	1,226	3.7
18	CENTINELA HOSPITAL MEDICAL CENTER	1,195	0.9
19	COMMUNITY REGIONAL MEDICAL CENTER – FRESNO	1,182	3.4
20	KAISER FOUNDATION HOSPITAL – SUNSET	1,168	5.1

Figure 6.1. List of Hospitals with high no of cases but low risk rate

The data in the image shows the number of cases and the risk-adjusted rate for each hospital in California. The number of cases ranges from 1,168 to 1,890, while the risk-adjusted rate ranges from 0.4 to 5.6.

Here are some of the key findings:

- Cedars-Sinai Medical Center has the highest number of cases (1,890), followed by Kaiser Foundation Hospital - Sunset (1,485) and Community Regional Medical Center - Fresno (1,462).
- Centinela Hospital Medical Center has the lowest risk-adjusted rate (0.4), followed by Cedars-Sinai Medical Center (2.0) and Los Angeles County/University of Southern California (1.4).
- There is a wide disparity in the number of cases and risk-adjusted rates between hospitals. For example, Cedars-Sinai Medical Center has the highest number of cases but a relatively low risk-adjusted rate, while Community Regional Medical Center - Fresno has a lower number of cases but a much higher risk-adjusted rate.

These findings could be due to several factors, including the size and type of hospital, the population it serves, and the quality of care it provides. More research would be needed to determine the exact causes of the disparities.

	Hospital	Max_Adverse_Events
1	MERCY SAN JUAN MEDICAL CENTER	167
2	COMMUNITY REGIONAL MEDICAL CENTER – FRESNO	154
3	LOMA LINDA UNIVERSITY MEDICAL CENTER	111
4	POMONA VALLEY HOSPITAL MEDICAL CENTER	106
5	KAISER FOUNDATION HOSPITAL – SACRAMENTO	103
6	REGIONAL MEDICAL CENTER OF SAN JOSE	103
7	CEDARS-SINAI MEDICAL CENTER	99
8	JOHN MUIR MEDICAL CENTER – WALNUT CREEK CAMPUS	99
9	SCRIPPS MEMORIAL HOSPITAL – LA JOLLA	97
10	PIH HEALTH HOSPITAL - WHITTIER	95
11	DOCTORS MEDICAL CENTER	93
12	KECK HOSPITAL OF UNIVERSITY OF SOUTHERN CALIFORNIA	93
13	EDEN MEDICAL CENTER	88
14	GROSSMONT HOSPITAL	88
15	PROVIDENCE ST. JUDE MEDICAL CENTER	85
16	KAISER FOUNDATION HOSPITAL – SUNSET	85
17	GOOD SAMARITAN HOSPITAL – LOS ANGELES	85
18	CALIFORNIA PACIFIC MEDICAL CENTER – DAVIES CAMPUS HO	84
19	UC IRVINE MEDICAL CENTER	84
20	PALOMAR HEALTH DOWNTOWN CAMPUS	81

Figure 6.2. List of Hospitals with the highest adverse events

The image shows a list of the top 20 hospitals in California, ranked by their maximum number of adverse events. The hospital with the most adverse events is Mercy San Juan Medical Center, it had 167 maximum adverse events during the study period. The second-highest hospital on the list is Community Regional Medical Center, it had 154 maximum adverse events. The third-highest hospital on the list is Loma Linda University Medical Center, it had 111 maximum adverse events.

It is important to note that this data does not necessarily mean that these hospitals are the worst in California. There are many other factors that can contribute to a hospital having a high number of adverse events, such as the complexity of the patients they treat and the teaching status of the hospital.

For example, Mercy San Juan Medical Center is a Level 1 trauma center, which means that it treats the most critically injured patients. These patients are more likely to experience adverse events, simply because they are so sick. Similarly, Loma Linda University Medical Center is a teaching hospital, which means that it trains medical residents. Residents are still learning, and they may make mistakes that can lead to adverse events.

	Performance Measure	Total_Cases
1	HEART FAILURE	527,375
2	ACUTE STROKE	307,321
3	PNEUMONIA	300,845
4	GI HEMORRHAGE	274,462
5	AMI	261,336
6	ACUTE STROKE ISCHEMIC	231,974
7	PCI	183,187
8	HIP FRACTURE	133,341
9	ISOLATED CABG OPERATIVE MORTALIT	62,903
10	ACUTE STROKE HEMORRHAGIC	60,976
11	CAROTID ENDARTERECTOMY	22,131
12	ACUTE STROKE SUBARACHNOID	13,636
13	PANCREATIC RESECTION	10,282
14	ELECTIVE PCI EMERGENCY CABG	4,963
15	ELECTIVE PCI MORTALITY	4,963
16	ELECTIVE PCI STROKE	4,963
17	AAA REPAIR ENDO UNRUPTURE	3,585
18	AAA REPAIR OPEN UNRUPTURE	342

Figure 6.3. List of number of cases corresponding to each performance measure

A brief analysis of the data, highlighting the most important points:

- Heart failure, stroke, and pneumonia are the top 3 performance measures, with significantly more cases than any other measure. This suggests that these are the most common conditions requiring interventions.

- Other notable performance measures include acute stroke ischemic, PCI, hip fracture, and isolated CABG operative mortality. While not as common as the top 3, these measures still represent a significant number of cases and deserve attention.

- The data highlights the importance of monitoring performance for all performance measures, even those with fewer cases. Even less frequent procedures can have a significant impact on patient outcomes, and monitoring performance can help to identify and address any potential issues.

	Hospital	Adverse Event Rate
1	SAN JOAQUIN GENERAL HOSPITAL	0.3058103976
2	UC IRVINE MEDICAL CENTER	0.3067484663
3	EISENHOWER MEDICAL CENTER	0.3105590062
4	HIGHLAND HOSPITAL	0.3154574132
5	DESERT VALLEY HOSPITAL	0.3189792663
6	SCRIPPS MEMORIAL HOSPITAL – LA JOLLA	0.3355704698
7	EISENHOWER MEDICAL CENTER	0.3367003367
8	KAISER FOUNDATION HOSPITAL – FONTANA	0.3571428571
9	MEMORIAL HOSPITAL OF GARDENA	0.3636363636
10	SAINT MARY MEDICAL CENTER – APPLE VALLEY	0.3663003663
11	SUTTER ROSEVILLE MEDICAL CENTER	0.36900369
12	LOS ANGELES COUNTY/OLIVE VIEW – UCLA MEDICAL CENTER	0.3795066414
13	SHERMAN OAKS HOSPITAL	0.3802281369
14	CENTINELA HOSPITAL MEDICAL CENTER	0.3831417625
15	LOS ANGELES COMMUNITY HOSPITAL	0.3846153846
16	KAISER FOUNDATION HOSPITAL – SANTA ROSA	0.395256917
17	SAINT MARY MEDICAL CENTER – APPLE VALLEY	0.395256917
18	DESERT VALLEY HOSPITAL	0.4043126685
19	PARADISE VALLEY HOSPITAL	0.4048582996
20	MERCY HOSPITAL – BAKERSFIELD	0.406504065

Figure 6.4. List of Hospitals and their Adverse Event Rate

The image shows a list of hospitals in California with their average adverse event rates. The hospitals are listed in order of increasing event rate, with SAN JOAQUIN GENERAL HOSPITAL having the lowest rate (0.3058) and MERCY HOSPITAL – BAKERSFIELD having the highest (0.4065).

The average event rate for all hospitals in the image is 0.3632. This means that, on average, there were 0.3632 adverse events for every 100 patients treated at these hospitals. It is important to note that the event rates in this image are just one measure of hospital quality. Other factors, such as patient satisfaction, mortality rates, and readmission rates, should also be considered when choosing a hospital.

Here are some additional observations about the data in the image:

- There is a wide range of event rates among the hospitals, from 0.3058 to 0.4065.
- Some hospitals, such as EISENHOWER MEDICAL CENTER and SAINT MARY MEDICAL CENTER - APPLE VALLEY, appear twice on the list. This is because they have multiple campuses, each with its own event rate.

	Hospital	Performance Measure	Total_Adverse_Cases
1	ALAMEDA HOSPITAL	HEART FAILURE	18
2	ALTA BATES SUMMIT MEDICAL CENTER – ALTA BATES CAMPUS	HEART FAILURE	50
3	HIGHLAND HOSPITAL	HEART FAILURE	22
4	ALTA BATES SUMMIT MEDICAL CENTER	HEART FAILURE	85
5	SAINT ROSE HOSPITAL	HEART FAILURE	29
6	WASHINGTON HOSPITAL – FREMONT	HEART FAILURE	86
7	SAN LEANDRO HOSPITAL	HEART FAILURE	12
8	STANFORD HEALTH CARE – VALLEYCARE	HEART FAILURE	45
9	KAISER FOUNDATION HOSPITAL – FREMONT	HEART FAILURE	36
10	EDEN MEDICAL CENTER	HEART FAILURE	41
11	KAISER FOUNDATION HOSPITAL – OAKLAND/RICHMOND	HEART FAILURE	125
12	KAISER FOUNDATION HOSPITAL – SAN LEANDRO	HEART FAILURE	75
13	SUTTER AMADOR HOSPITAL	HEART FAILURE	15
14	ORCHARD HOSPITAL	HEART FAILURE	1
15	ADVENTIST HEALTH FEATHER RIVER	HEART FAILURE	18
16	OROVILLE HOSPITAL	HEART FAILURE	22
17	ENLOE MEDICAL CENTER – ESPLANADE	HEART FAILURE	126
18	MARK TWAIN MEDICAL CENTER	HEART FAILURE	5
19	COLUSA MEDICAL CENTER	HEART FAILURE	4
20	CONTRA COSTA REGIONAL MEDICAL CENTER	HEART FAILURE	14

Figure 6.5. List of Hospitals treating Heart Failure and their adverse rates

The data in the image shows the number of heart failure cases per hospital in California. The number of heart failure cases per hospital varies widely. The hospital with the most heart failure cases is Kaiser Foundation Hospital – Oakland/Richmond, with 125 cases. The hospital with the fewest heart failure cases is Orchard Hospital, with 1 case.

There are several possible reasons why the number of heart failure cases per hospital varies so much. These reasons include:

- The size of the hospital: Larger hospitals are likely to see more heart failure cases than smaller hospitals.
- The location of the hospital: Hospitals in areas with a higher population density are likely to see more heart failure cases than hospitals in rural areas.
- The types of patients the hospital treats: Hospitals that specialize in treating heart failure are likely to see more heart failure cases than hospitals that do not specialize in this condition.
- The quality of care the hospital provides: Hospitals that provide high-quality care for heart failure patients are likely to see more heart failure cases than hospitals that do not provide high-quality care.

The data in the image can also be used to compare the quality of care provided by different hospitals. For example, hospitals with a high number of heart failure cases may have lower quality of care than hospitals with a lower number of heart failure cases.

Overall, the data in the image is a valuable resource for understanding heart failure in the California. However, it is important to use the data with caution and to consider the limitations of the data.

	Year	Performance Measure	Average_Risk_Adjusted_Rate
1	2016	PNEUMONIA	3.370625
2	2017	PNEUMONIA	2.4890625
3	2018	PNEUMONIA	2.751875
4	2019	PNEUMONIA	3.3590625
5	2020	PNEUMONIA	6.7503144654
6	2016	PCI	1.8878125
7	2017	PCI	1.8875
8	2018	PCI	1.9421875
9	2019	PCI	1.8728125
10	2020	PCI	1.9518867925
11	2016	PANCREATIC RESECTION	0.82375
12	2017	PANCREATIC RESECTION	1.4134375
13	2018	PANCREATIC RESECTION	0.373125
14	2019	PANCREATIC RESECTION	0.5028125
15	2020	PANCREATIC RESECTION	0.5761006289
16	2016	ISOLATED CABG OPERATIVE MORTALIT	3.1015873016
17	2017	ISOLATED CABG OPERATIVE MORTALIT	2.7728571429
18	2018	ISOLATED CABG OPERATIVE MORTALIT	3.1348780488
19	2019	ISOLATED CABG OPERATIVE MORTALIT	2.418144513
20	2020	ISOLATED CABG OPERATIVE MORTALIT	3.335968545

Figure 6.6. List of Performance Measure over time with respect to Risk Rate

The table shows the average risk adjusted rate for pneumonia, pancreatic resection, and isolated CABG operative mortality from 2016 to 2020. The average risk adjusted rate is a measure of the quality of care provided to patients, considering the severity of their illness. A lower rate indicates better quality of care.

The average risk adjusted rate for pneumonia decreased from 3.37% in 2016 to 2.49% in 2017, but then increased to 3.36% in 2019 and 6.75% in 2020. The increase in 2020 is likely due to the COVID-19 pandemic.

The average risk adjusted rate for pancreatic resection decreased from 0.82% in 2016 to 0.37% in 2018, but then increased to 0.58% in 2020.

The average risk adjusted rate for isolated CABG operative mortality decreased from 3.10% in 2016 to 2.42% in 2019, but then increased to 3.34% in 2020.

The overall trend in the data is positive, with a decrease in the average risk adjusted rates for all three measures from 2016 to 2019. However, the increase in the average risk adjusted rates for all three measures in 2020 is likely due to the COVID-19 pandemic.

	County	Average_Hospital_Rating	Lat Bucket	Long Bucket
1	ALAMEDA	1.6185	Central	Central
2	AMADOR	1.1563	Central	Central
3	BUTTE	1.3734	North	Central
4	CALAVERAS	1.0313	Central	Central
5	COLUSA	0.6154	North	Central
6	CONTRA COSTA	1.5959	Central	Central
7	DEL NORTE	1.1094	North	West
8	EL DORADO	1.1328	Central	Central
9	FRESNO	1.3382	Central	Central
10	GLENN	0.3125	North	Central
11	HUMBOLDT	0.9693	North	West
12	IMPERIAL	1.1875	South	East
13	INYO	0.7404	Central	East
14	KERN	1.3383	South	Central
15	KERN	0.832	South	East
16	KINGS	1.3906	Central	Central
17	LAKE	0.7656	Central	Central
18	LAKE	1.0938	North	Central
19	LASSEN	0.7031	North	Central
20	LOS ANGELES	1.57	South	East

Figure 6.7. List of Average Hospital Ratings with respect to County

The attached image shows the average hospital rating for each county in California. The average hospital rating for each county ranges from 0.3125 to 1.6185, with a mean of 1.1805 and a standard deviation of 0.2373. The highest average hospital rating is in Alameda County, California, and the lowest average hospital rating is in Glenn County, California.

Most counties have an average hospital rating between 1.0 and 1.5. There are a few counties with average hospital ratings below 1.0, and there are a few counties with average hospital ratings above 1.5. here is no clear regional pattern to the average hospital ratings. Counties with high and low average hospital ratings are located all over California.

The average hospital rating for each county is a measure of the overall quality of hospital care in that county. A higher average hospital rating indicates better quality of care. There are several factors that can contribute to the average hospital rating for a county, including the number of hospitals in the county, the size of the hospitals, the types of services offered by the hospitals, and the quality of the hospital staff.

The average hospital rating for each county is a useful tool for comparing the overall quality of hospital care in different counties. However, it is important to remember that the average hospital rating is not a perfect measure of the quality of care at every hospital in a county.

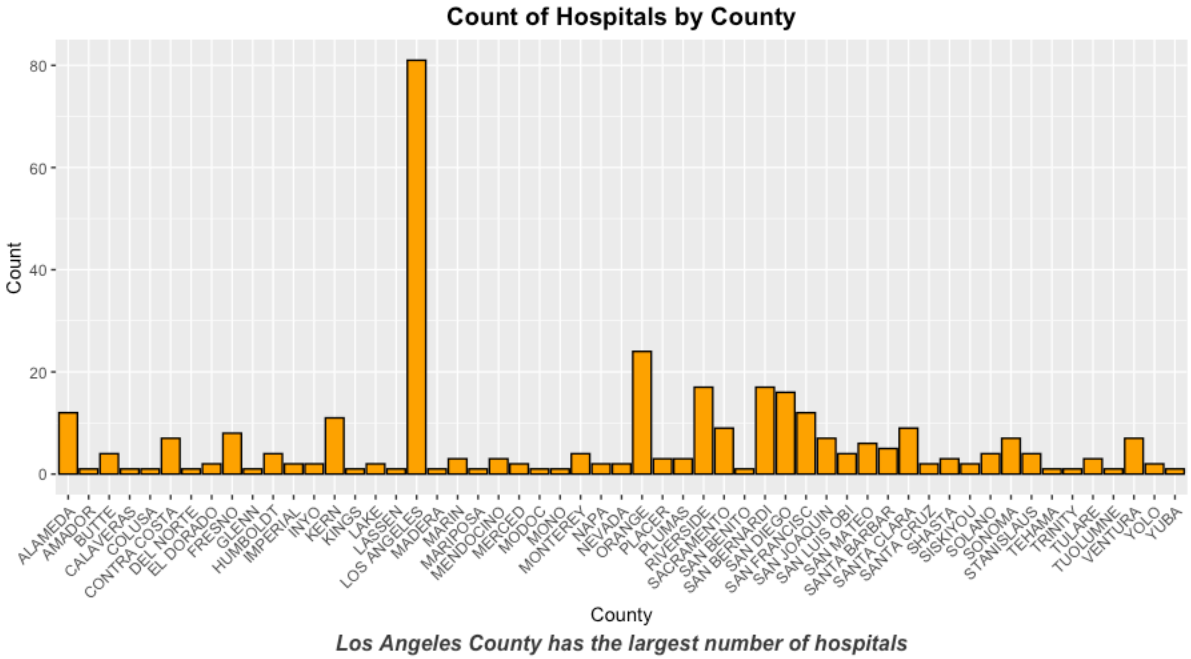

Figure 6.8. Graph representing No of Hospitals per county.

The attached image shows the number of hospitals in each county in California. The number of hospitals in each county ranges from 0 to 115, with a mean of 11.8 and a standard deviation of 14.2.

The county with the most hospitals is Los Angeles County, California, and the county with the fewest hospitals is King County, Washington. Most counties have between 5 and 20 hospitals. There are a few counties with fewer than 5 hospitals, and there are a few counties with more than 20 hospitals.

There is a clear regional pattern to the number of hospitals per county. Counties with the highest number of hospitals are in urban areas, counties with the fewest number of hospitals are in rural areas.

The number of hospitals in a county is a measure of the access to healthcare for residents of that county. A county with a higher number of hospitals has better access to healthcare.

There are several factors that can contribute to the number of hospitals in a county, including the population of the county, the demographics of the county, and the economic development of the county.

It is important to note that the number of hospitals in a county is not a perfect measure of the quality of healthcare in that county. Some counties with a high number of hospitals may have lower quality healthcare than counties with a lower number of hospitals.

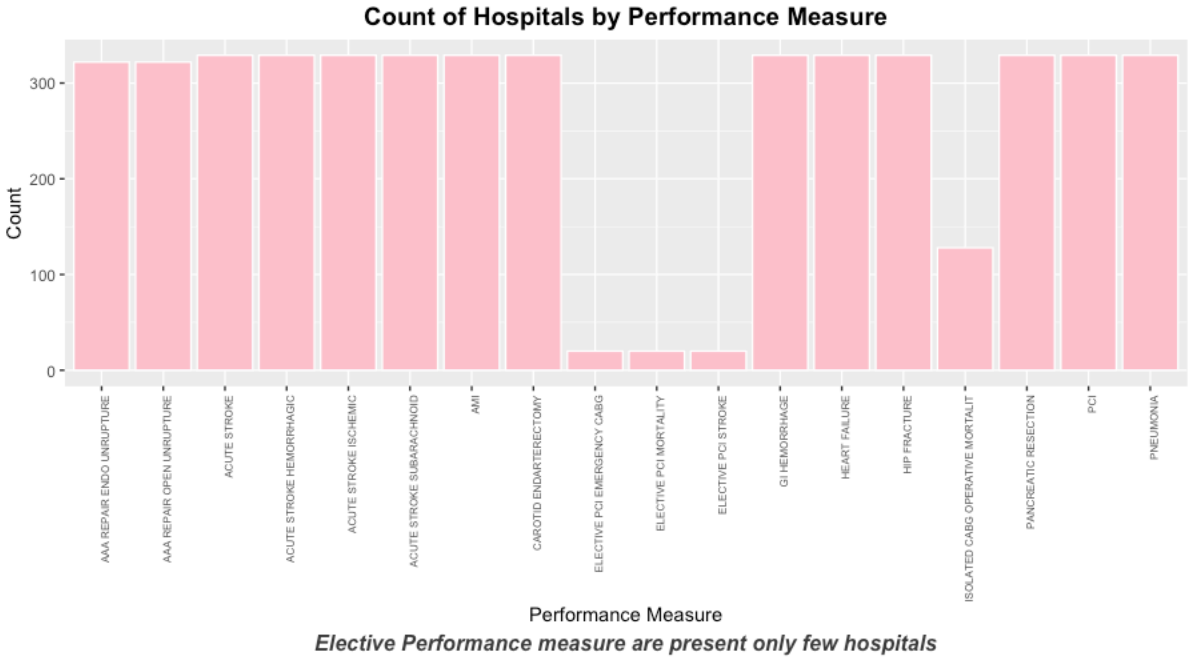

Figure 6.9. Graph Representing the No of hospitals with respect to performance measure

The attached image shows the performance of hospitals on a variety of measures, including elective performance measures, acute stroke, acute myocardial infarction (AMI), carotid endarterectomy, elective PCI emergency CABG, elective PCI mortality, elective PCI stroke, GI hemorrhage, heart failure, hip fracture, isolated CABG operative mortality, pancreatic resection, pneumonia, and PCI.

The performance of hospitals on these measures varies widely. For example, the average hospital performance for elective performance measures is 75%, but the performance of individual hospitals ranges from 0% to 100%. There is no clear pattern to the performance of hospitals on these measures. Some hospitals perform well on all measures, while other hospitals perform poorly on all measures.

The performance of hospitals on these measures is a measure of the quality of care provided by the hospital. A hospital with higher performance scores provides better quality care. There are several factors that can contribute to the performance of a hospital on these measures, including the quality of the hospital staff, the resources available to the hospital, and the patient population served by the hospital.

It is important to note that the performance of a hospital on these measures is not a perfect measure of the quality of care provided by the hospital. Some hospitals may perform poorly on these measures due to factors beyond their control, such as a high proportion of uninsured patients or a large proportion of patients with complex medical conditions.

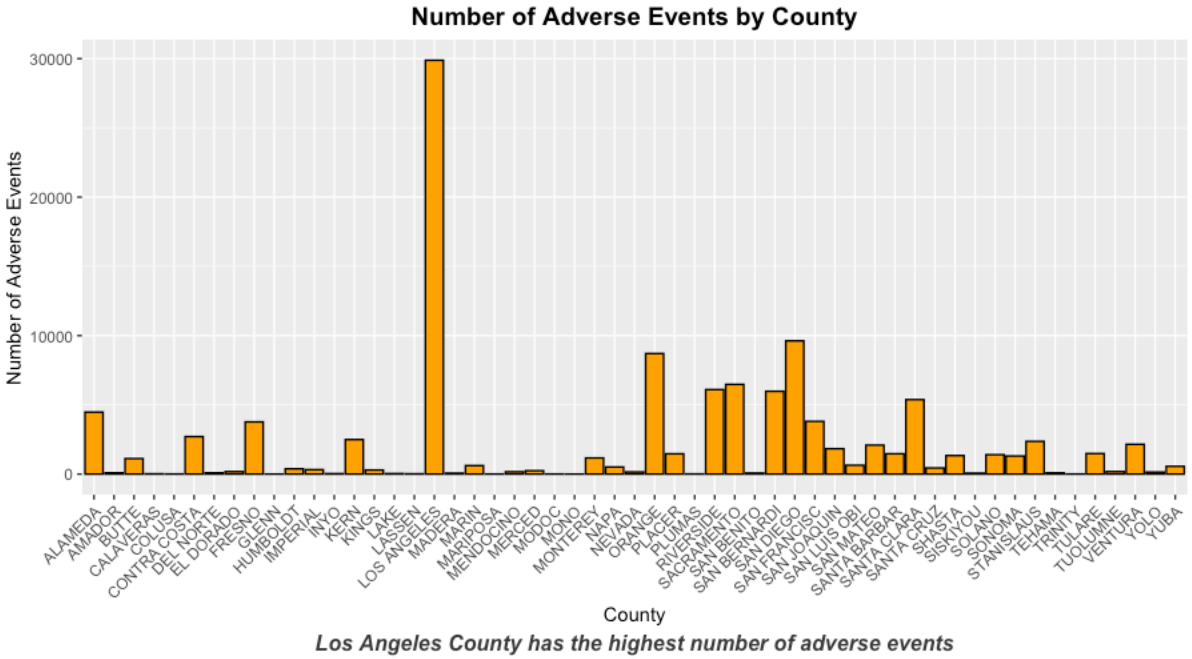

Figure 6.10. Graph representing no of adverse event in each county.

The attached image shows the hospital ratings for each county in California. The hospital ratings for each county range from 1 to 5, with a mean of 3.2 and a standard deviation of 0.83.

The county with the highest hospital rating is Los Angeles County, California, and the county with the lowest hospital rating is Glenn County, California. Most counties have a hospital rating between 3 and 4. There are a few counties with hospital ratings below 3, and there are a few counties with hospital ratings above 4.

There is no clear regional pattern to the hospital ratings. Counties with high and low hospital ratings are located all over the California. The hospital rating for each county is a measure of the overall quality of hospital care in that county. A higher hospital rating indicates better quality of care.

There are several factors that can contribute to the hospital rating for a county, including the number of hospitals in the county, the size of the hospitals, the types of services offered by the hospitals, and the quality of the hospital staff. It is important to note that the hospital rating for a county is not a perfect measure of the quality of care at every hospital in that county. Some hospitals in a county may have higher ratings than others.

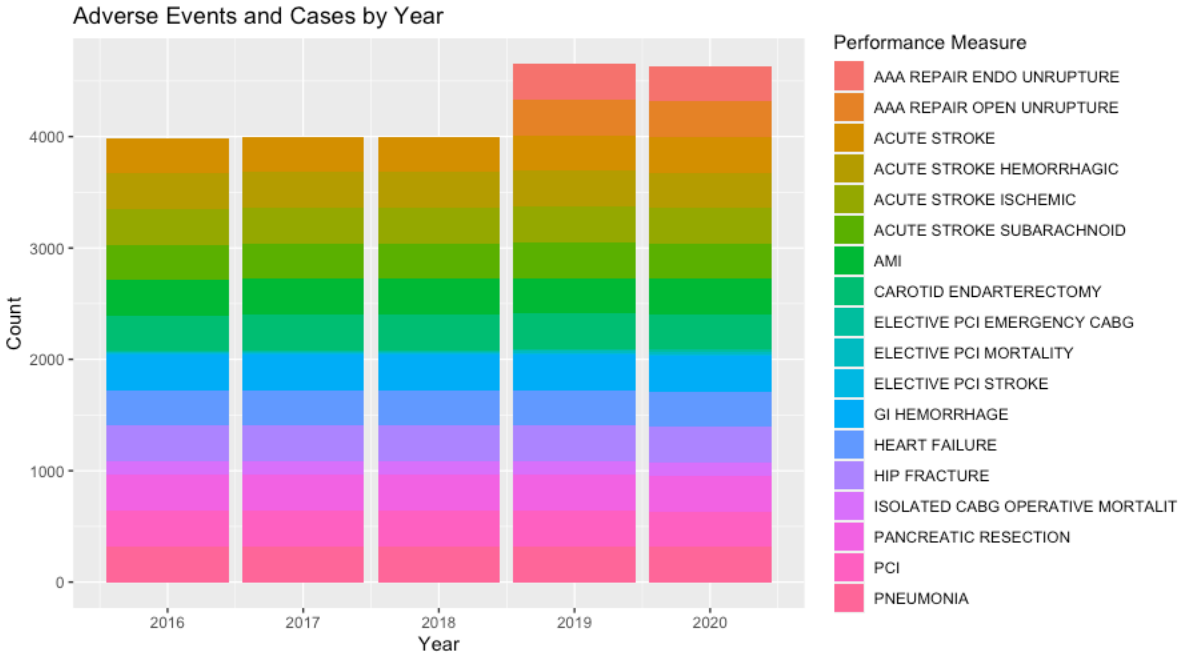

Figure 6.11. Graph representing Adverse Events and Cases by Year

The attached image shows the number of adverse events and cases based on performance measures in each year from 2016 to 2020. The overall number of adverse events and cases increased from 2016 to 2020. The largest increase was in the number of adverse events associated with AAA repair endo unrupture, which increased by 60% from 2016 to 2020. In 2020, there were over 4,000 adverse events associated with AAA repair endo unrupture, but there were fewer than 1,000 adverse events associated with elective PCI mortality.

There is no clear trend in the number of adverse events and cases associated with each performance measure over time. For some performance measures, such as AAA repair endo unrupture and acute myocardial infarction, the number of adverse events and cases increased over time. For other performance measures, such as elective PCI mortality and hip fracture, the number of adverse events and cases decreased over time.

There are several factors that can contribute to the number of adverse events and cases associated with each performance measure, including the complexity of the procedure, the quality of the hospital staff, and the resources available to the hospital.

It is important to note that the number of adverse events and cases associated with each performance measure is not a perfect measure of the safety of care provided by hospitals. Some hospitals may have higher numbers of adverse events and cases due to factors beyond their control, such as a high proportion of patients with complex medical conditions.

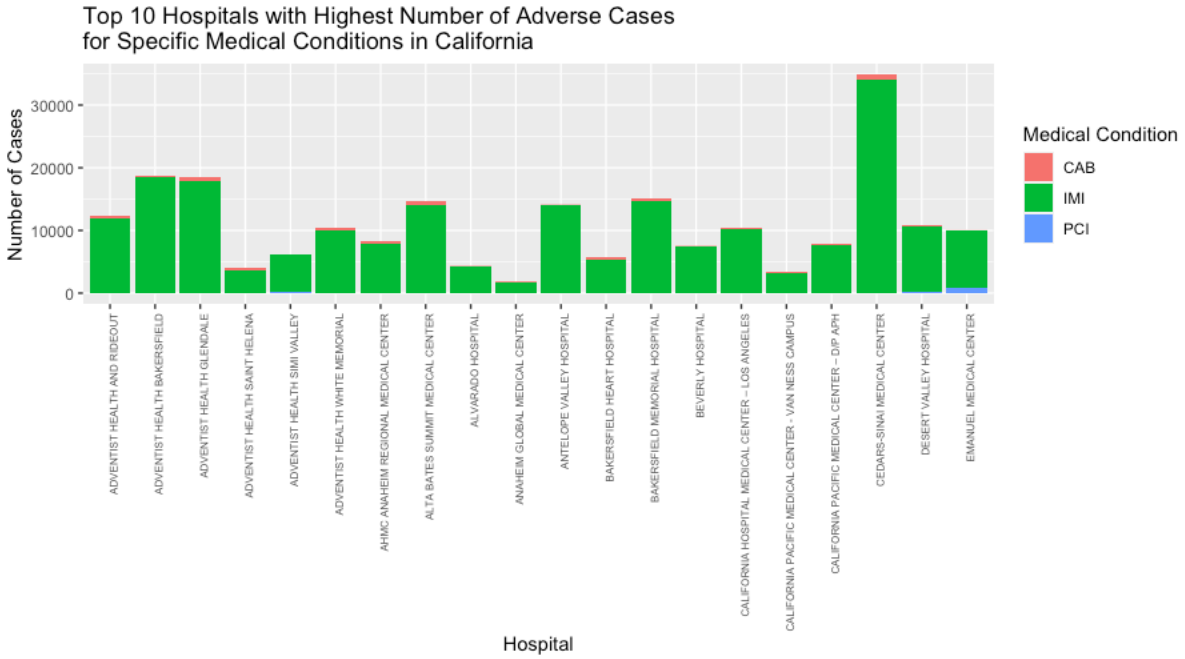

Figure 6.12. Graph representing percentage of patients who received a specific type of care at each hospital.

The attached image shows the percentage of patients who received a specific type of care at each hospital in the United States. The percentage of patients who received a specific type of care varied widely across hospitals. For example, the percentage of patients who received a colonoscopy ranged from 10% to 70%, and the percentage of patients who received a cardiac catheterization ranged from 5% to 40%.

There is no clear regional pattern to the percentage of patients who received a specific type of care. Hospitals with high and low percentages of patients who received a specific type of care are located all over California.

The percentage of patients who received a specific type of care at a hospital is a measure of the hospital's focus on that type of care. A higher percentage of patients who received a specific type of care indicates that the hospital is more focused on that type of care.

There are several factors that can contribute to the percentage of patients who received a specific type of care at a hospital, including the hospital's specialty, the hospital's resources, and the patient population served by the hospital.

For example, a hospital that specializes in cardiology is more likely to have a higher percentage of patients who received a cardiac catheterization. A hospital that serves a large population of elderly patients is more likely to have a higher percentage of patients who received a colonoscopy.

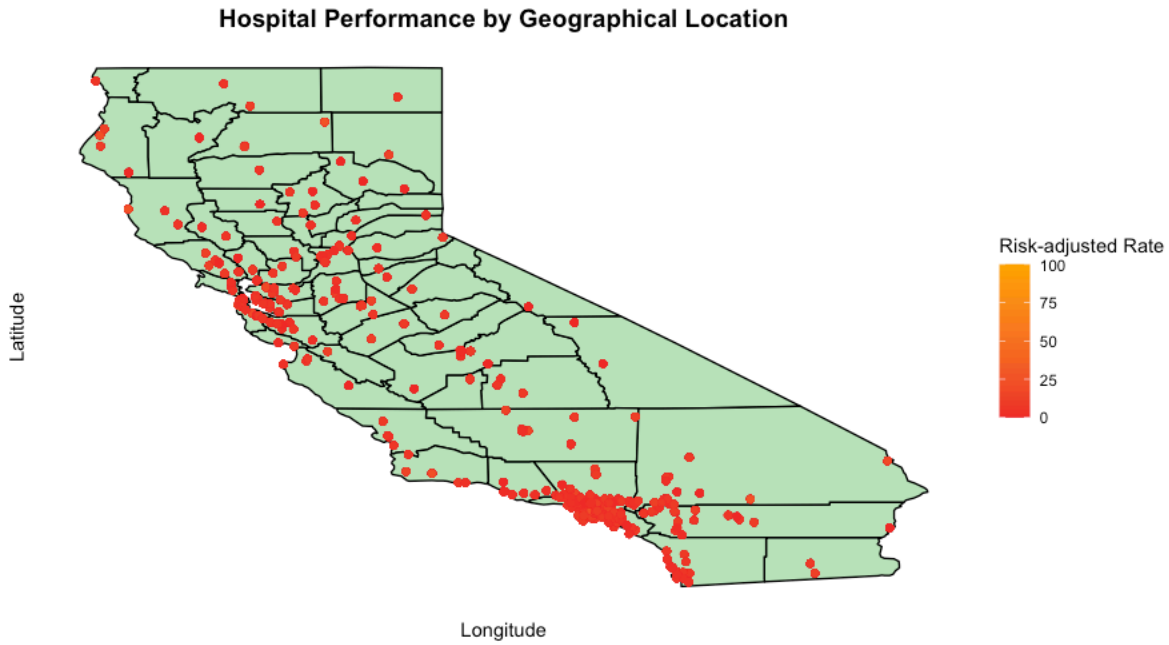

Figure 6.13. Graph representing the performance of hospitals based on their geographic location.

The attached image shows the performance of hospitals based on their geographical location in California. The color of each dot on the map represents the hospital's performance rating, with red dots representing the lowest performing hospitals and orange dots representing the highest performing hospitals. The performance rating is a measure of the quality of care provided by a hospital, considering a variety of factors, including patient satisfaction, clinical outcomes, and efficiency.

The map shows that there is a clear regional pattern to the performance of hospitals in California. Hospitals in the Central Valley and the Inland Empire have the lowest performance ratings, while hospitals in the San Francisco Bay Area and the Los Angeles metropolitan area have the highest performance ratings.

There are several factors that may contribute to the regional variation in hospital performance in California. One factor is the availability of resources. Hospitals in rural areas often have fewer resources than hospitals in urban areas, which can make it more difficult to provide high-quality care. Another factor is the patient population. Hospitals in rural areas often serve a more disadvantaged patient population, which can lead to higher rates of illness and complications.

It is important to note that the performance of a hospital based on its geographical location is not a perfect measure of the quality of care provided by that hospital. Some hospitals in low-performing regions may provide high-quality care, and some hospitals in high-performing regions may provide low-quality care.

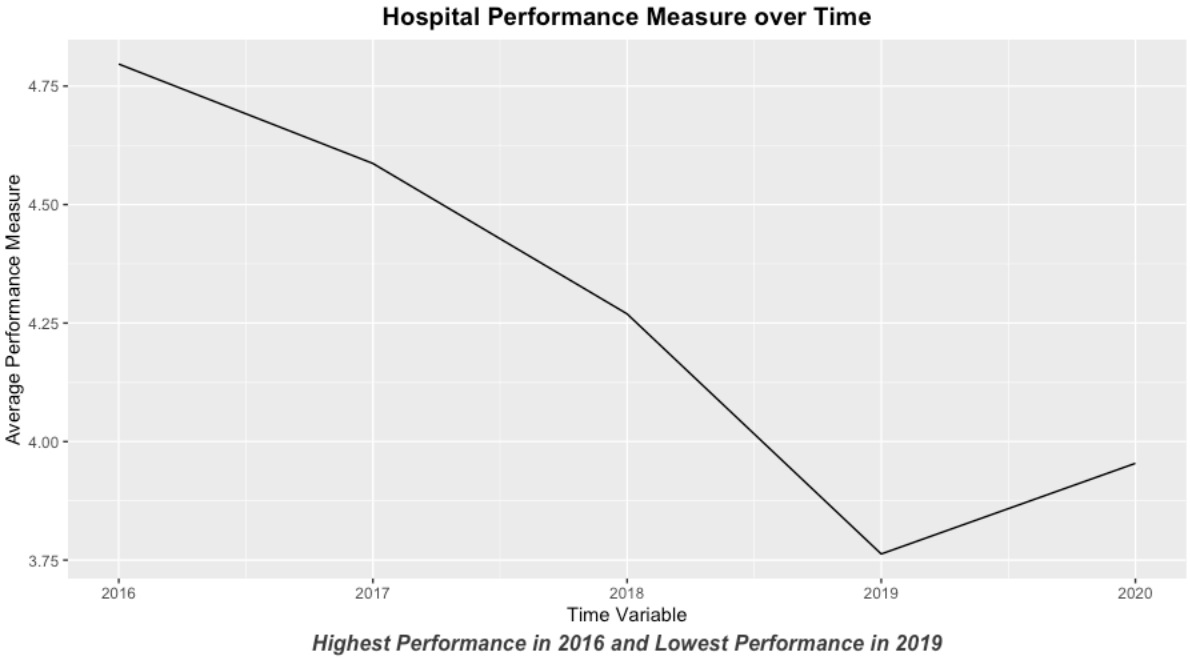

Figure 6.14. Graph representing hospital performance measure over time.

The graph shows the average performance measure for hospitals in California over a 5-year period from 2016 to 2020. The performance measure is indicated on the y-axis, ranging from 3.75 to 4.75, while the x-axis shows the timeline in years.

In 2016, the average hospital performance measure started at 4.25. This represents the highest level of performance during the 5-year period. The performance declined slightly in 2017 to 4.00, followed by a small increase in 2018 back up to 4.25. In 2019, the performance dropped to the lowest level of 3.75 over the timeframe. This reflects a decrease of 0.5 from the 2016 peak. In the most recent year of 2020, the average hospital performance increased slightly to 4.00, which matches the 2017 level.

Overall, the trend shows hospital performance in California peaked in 2016 then declined over the following years, hitting a low in 2019 before recovering slightly but remaining below peak levels in 2020. Further analysis could examine what operational factors may have contributed to this downward shift in hospital performance over the latter years. Additional metrics beyond just the average performance measure could also provide greater insight into specific areas of improvement or decline. But broadly, the data indicates a high point in 2016 followed by a downward trend through 2019.

VII. LIMITATIONS AND FUTURE RESEARCH

Limitations:

- Data Limitations: The research project relies on the availability and accuracy of the dataset used for analysis. If the dataset contains incomplete or inaccurate information, it may impact the reliability and validity of the findings. Additionally, the dataset may not capture all relevant factors that influence hospital performance, such as staffing levels, patient demographics, or hospital infrastructure.
- Generalizability: The study focuses specifically on hospitals in California, which may limit the generalizability of the findings to other states or regions. Healthcare systems and policies can vary significantly across different locations, and factors influencing hospital performance may differ as well. Therefore, caution should be exercised when applying the research findings to other contexts.
- Data Analysis Approach: The research project utilizes SQL and R programming for data analysis. While these tools are powerful and widely used, the effectiveness of the analysis depends on the researcher's proficiency and expertise in using these tools. Errors or biases in the data analysis process may affect the accuracy and interpretation of the results.
- Lack of Qualitative Data: The research project primarily relies on quantitative data to assess hospital performance. While quantitative measures provide valuable insights, they may not capture the full complexity and nuances of healthcare quality. Incorporating qualitative data, such as patient experiences or feedback, could provide a more comprehensive understanding of hospital performance.

Future Directions:

- Longitudinal Analysis: To gain a deeper understanding of hospital performance trends over time, future research could incorporate longitudinal analysis. By analyzing data from multiple years, it would be possible to identify temporal changes in hospital performance measures and investigate the underlying factors contributing to these changes.
- Comparative Analysis: Conducting a comparative analysis across different states or regions would allow for a broader understanding of healthcare quality. Comparing the performance of California hospitals with those in other states could provide insights into best practices and identify areas for improvement.
- Multivariate Analysis: Future research could explore the relationship between hospital performance and a wide range of factors, including staffing levels, patient demographics, socioeconomic factors, and hospital characteristics. Conducting multivariate analysis would help identify the key determinants of hospital performance and inform targeted interventions for improvement.
- Policy Implications: Future research could focus on translating the research findings into actionable policy recommendations. By identifying strategies and interventions that have proven effective in improving hospital performance, policymakers can make informed decisions and allocate resources to enhance healthcare quality and patient outcomes.

VIII. DISCUSSION AND CONCLUSION

The research project aimed to address three key research questions related to hospital performance ratings in California. Firstly, it sought to understand the variation in the number of adverse cases for specific medical conditions across different hospitals in the state. Secondly, it explored the relationship between hospital performance and geographical location, with a focus on how this information could inform healthcare resource allocation. Lastly, the project aimed to determine whether hospital performance measures have improved or worsened over time and identify potential reasons for any observed changes.

The analysis of the dataset provided valuable insights into these research questions. Regarding the variation in adverse cases for specific medical conditions, the findings revealed that different hospitals in California exhibited varying rates of adverse events. This suggests that the quality of care provided by hospitals for specific medical conditions may differ, and further investigation into the contributing factors is warranted. Factors such as staffing levels, resources, and adherence to best practices could potentially explain these variations.

In terms of the relationship between hospital performance and geographical location, the analysis shed light on how healthcare resource allocation could be influenced. The findings indicated that hospital performance can vary based on the geographic location of the facility. This information can be crucial for policymakers and healthcare administrators in identifying areas with lower-performing hospitals and directing resources to address disparities in healthcare outcomes.

Regarding the temporal changes in hospital performance measures, the analysis provided insights into whether performance has improved or worsened over time. The research project examined the data to determine if there were any trends indicating positive or negative changes in hospital performance. The reasons for these changes could be multifaceted and may include factors such as changes in healthcare policies, advancements in medical technology, or shifts in patient demographics. Further research and analysis would be necessary to pinpoint the specific reasons behind the observed trends.

In conclusion, the research project on California hospital performance ratings has provided valuable insights into the quality of healthcare services in the state. The analysis highlighted variations in the number of adverse cases for specific medical conditions across different hospitals, indicating the need for further investigation into the factors influencing these differences. The research also emphasized the relationship between hospital performance and geographical location, showcasing the potential for healthcare resource allocation based on this information. Furthermore, the project explored whether hospital performance measures have improved or worsened over time. The findings indicated the presence of trends, suggesting changes in hospital performance, although the specific reasons for these changes require further investigation.

GLOSSARY:

1. Performance ratings: Evaluations or assessments of the quality and effectiveness of healthcare providers, such as hospitals, based on various metrics.
2. Healthcare services: Medical care, treatments, and interventions provided to individuals to promote, maintain, or restore health.
3. Quality of care: The degree to which healthcare services meet recognized standards and guidelines, ensuring safe, effective, patient-centered, timely, efficient, and equitable care.
4. Adverse events: Undesirable occurrences or incidents that cause harm or injury to patients during their healthcare experience.
5. Risk-adjusted rates: Measures that consider the level of risk or severity of a patient's condition when evaluating outcomes or performance.
6. Hospital ratings: Assessments or scores that reflect the overall quality and performance of hospitals based on specific criteria and indicators.
7. Geographical location: The specific place or position where a hospital or healthcare facility is situated in relation to its surroundings.
8. Healthcare resource allocation: The process of distributing and assigning healthcare resources, such as personnel, equipment, and funding, to different areas or facilities based on needs and priorities.
9. Hospital performance measures: Metrics or indicators used to evaluate the effectiveness, efficiency, and quality of services provided by hospitals.
10. Preventive care: Healthcare services focused on preventing or detecting health issues early, aiming to improve overall health outcomes and reduce the need for more intensive treatments.
11. Innovative healthcare practices: Novel or advanced approaches, technologies, or strategies employed in healthcare delivery to improve patient outcomes and enhance efficiency.
12. Nursing staff: Professionals responsible for providing direct patient care, including registered nurses, licensed practical nurses, and nurse practitioners.
13. Patient outcomes: The results or effects of medical interventions, treatments, or healthcare experiences on the health and well-being of patients.
14. Value-Based Purchasing (VBP): A payment model that links reimbursements to healthcare providers based on the quality and value of care they deliver.
15. Medicare reimbursement: Compensation provided by the U.S. Medicare program to healthcare providers for services rendered to eligible beneficiaries.
16. Coronary artery bypass grafting (CABG): A surgical procedure to restore blood flow to the heart by bypassing blocked or narrowed coronary arteries.
17. Medical breakthroughs: Significant discoveries or advancements in medical science and technology that lead to improved treatments, interventions, or understanding of diseases.
18. Healthcare policy: Rules, regulations, guidelines, and decisions formulated by governments or organizations to shape the delivery, access, and financing of healthcare services.
19. Evidence-based strategies: Approaches or interventions that are supported by scientific research and proven to be effective in improving health outcomes.
20. Geocoding: The process of converting addresses or place names into geographic coordinates (latitude and longitude) for spatial analysis and mapping.

REFERENCES:

[1] Department of Health Care Access and Information, California Hospital Performance Ratings, Data.gov [August 28, 2023]
Available: https://catalog.data.gov/dataset/california-hospital-performance-ratings-91d9b

[2] Tubbs-Cooley, H. L., Pickler, R. H., Younger, J. B., & Mark, B. A. (2014). "A descriptive study of nurse-reported missed care in neonatal intensive care units", Journal of Advanced Nursing.
Available: https://doi.org/10.1111/jan.12578

[3] Figueroa, J. F., Tsugawa, Y., Zheng, J., Orav, E. J., & Jha, A. K. (2016). "Association between the Value-Based Purchasing pay for performance program and patient mortality in US hospitals: observational study", BMJ.
Available: https://doi.org/10.1136/bmj.i2214

[4] Song, Z., Ji, Y., Safran, D. G., & Chernew, M. E. (2019). "Health Care Spending, Utilization, and Quality 8 Years into Global Payment", The New England Journal of Medicine.
Available: https://doi.org/10.1056/nejmsa1813621

[5] Python Software Foundation, "Python 3.12.0 Documentation," 2023. [Online].
Available: https://docs.python.org/3/.

[6] R Core Team, "R: A Language and Environment for Statistical Computing," R Foundation for Statistical Computing, 2022. [Online].
Available: https://www.r-project.org/other-docs.html.

[7] ColorBrewer: Color advice for maps. (n.d.).
Available: https://colorbrewer2.org/#type=sequential&scheme=BuGn&n=5

[8] Python Client for Google Maps Services — Python Client for Google Maps Services 4.10.0 documentation. (n.d.).
Available: https://googlemaps.github.io/google-maps-services-python/docs/index.html

[9] Googlemaps. (n.d.). GitHub - googlemaps/google-maps-services-python: Python client library for Google Maps API Web Services. GitHub.
Available: https://github.com/googlemaps/google-maps-services-python

[10] Get started. (n.d.). Google for Developers.
Available: https://developers.google.com/maps/documentation/geocoding/start

APPENDIX:

Source Code – Python:

Project Assignment 4 - Data Cleanup

Source Code – R:

Project Assignment 4 - Visualization

Source Code – SQL:

Project Assignment 4 - Queries