Healthcare Data Analysis and Visualization Report

Headings in This Report

1. Objective
2. Data Cleaning Process
3. Findings and Answers (with insights and advice)
4. Key Visualizations (explaining insights)
5. Conclusion
6. Future Recommendations

Objective

This analysis aims to explore healthcare data and answer the following key questions:

1. **Which procedures have the highest risk-adjusted mortality rates?**
2. **How do deaths and cases vary by procedure?**
3. **Which hospitals have the highest risk-adjusted mortality rates?**
4. **How do deaths and cases change over the years?**
5. **What trends can be observed in risk-adjusted mortality rates over time?**

The goal is to provide actionable insights to improve patient outcomes, allocate resources effectively, and optimize hospital

Data Cleaning Process

Steps Taken

1. **Data Importation**:
   a. Imported the dataset into Python using `pandas` for data manipulation.
2. **Renaming Columns**:
   a. Corrected typos in column names, such as renaming `'Risk Adjuested Mortality Rate'` to `'Risk Adjusted Mortality Rate'`.
3. **Data Type Adjustments**:
   a. Converted columns to appropriate data types, e.g., `'YEAR'` to integer and `'Deaths'` to float.
4. **Handling Missing Values**:
   a. **Example**: Missing values in `'Risk Adjusted Mortality Rate'` were filled with the median for each `'Procedure/Condition'`.
   b. Other categorical columns like `'Hospital Ratings'` were filled with `'Unknown'`, while geographic columns (`'Longitude'`, `'Latitude'`) were imputed with the median by County.
5. **Validation**:
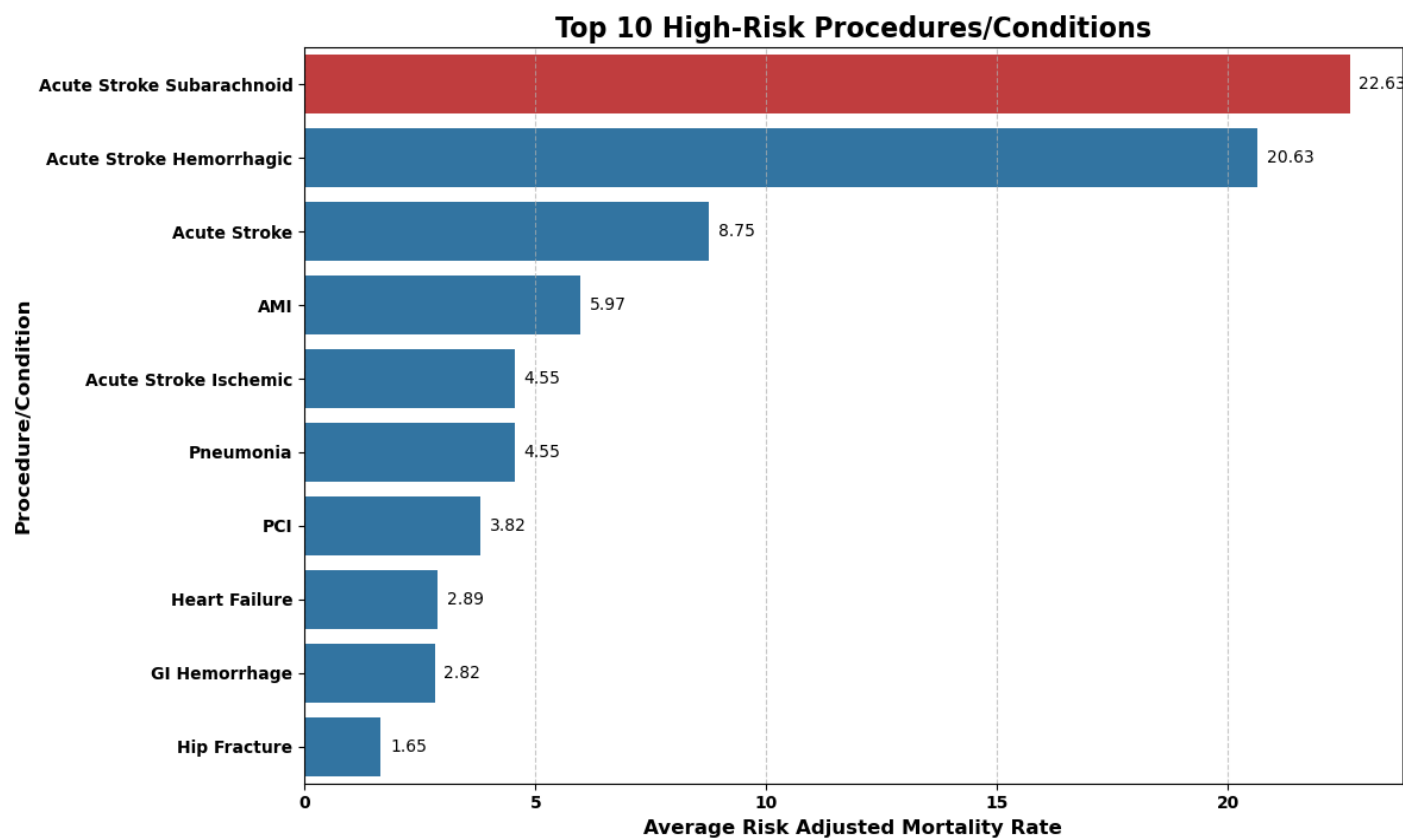   a. Verified no missing values remained in critical columns, ensuring the dataset was clean and ready for analysis.

This concise cleaning process ensured consistency and reliability for visualization and insights generation.

Findings and Answers

1. Which procedures have the highest risk-adjusted mortality rates?

**A. Finding**: Acute Stroke Subarachnoid has the highest risk-adjusted mortality rate, highlighted in red.

B. Advice: Allocate additional resources, staff training, and specialized care to manage high-risk procedures effectively.

Top 10 High-Risk Procedures/Conditions

Procedure/Condition	Average Risk Adjusted Mortality Rate
Acute Stroke Subarachnoid	22.63
Acute Stroke Hemorrhagic	20.63
Acute Stroke	8.75
AMI	5.97
Acute Stroke Ischemic	4.55
Pneumonia	4.55
PCI	3.82
Heart Failure	2.89
GI Hemorrhage	2.82
Hip Fracture	1.65

2. How do deaths and cases vary by procedure?

A. Finding: Heart Failure and Pneumonia showed high deaths relative to cases, visualized in a stacked bar chart.
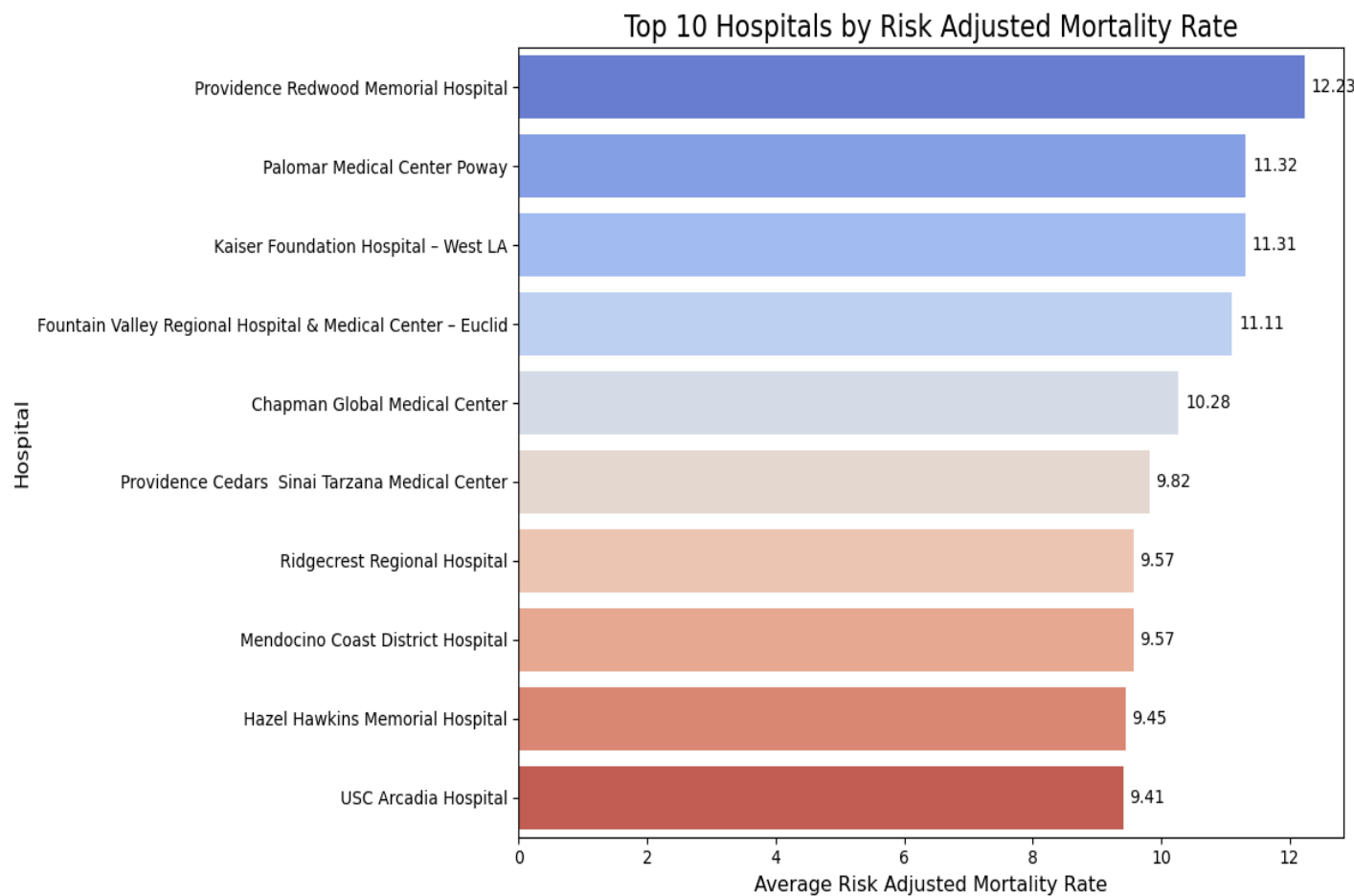
B. Advice: Focus on improving early detection and treatment protocols to reduce the death-to-case ratios for these procedures.

Stacked Bar Chart of Deaths and Cases by Procedure
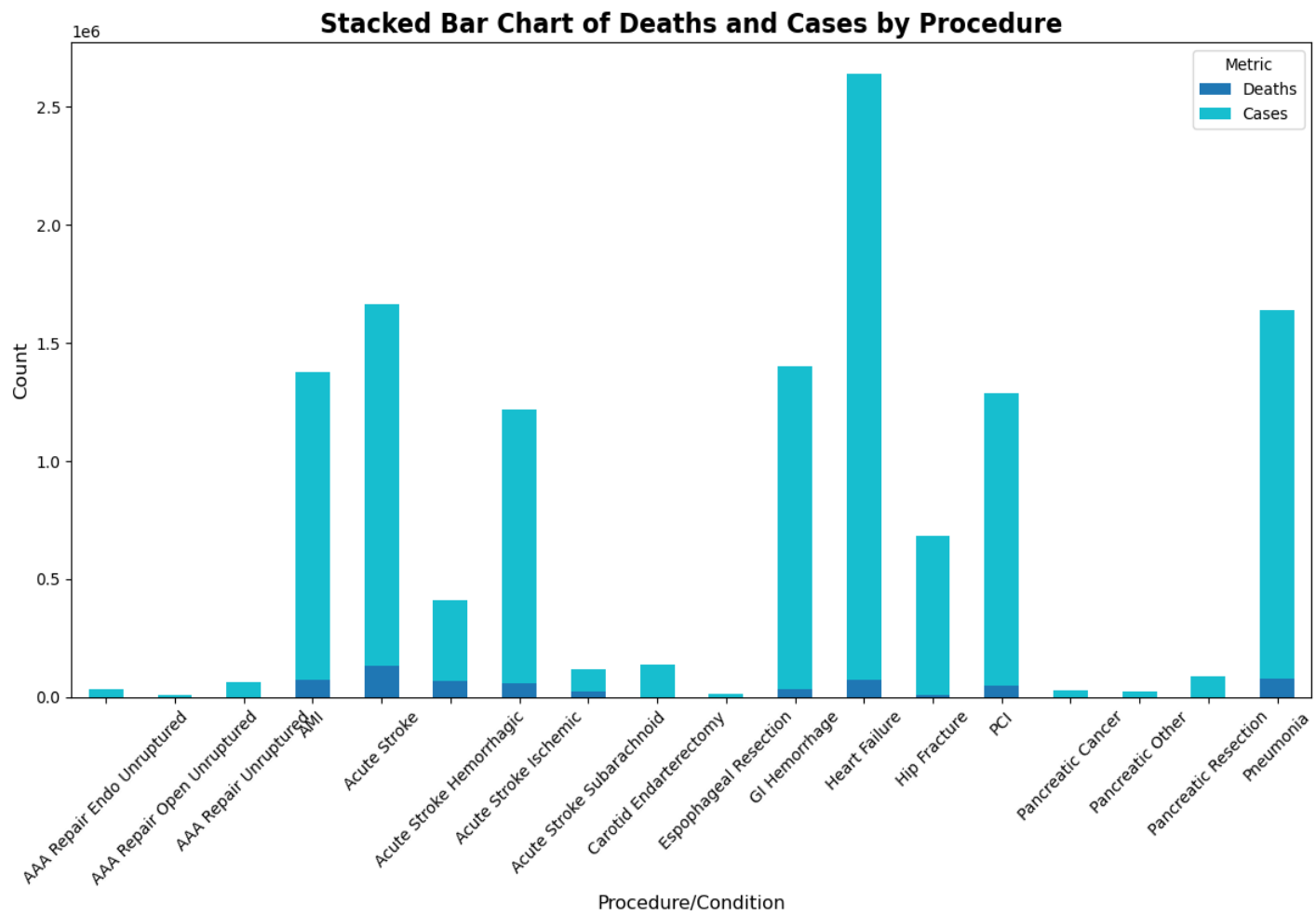
3. Which hospitals have the highest risk-adjusted mortality rates?

A. Finding: Providence Redwood Memorial Hospital exhibits the highest mortality rate, highlighted in red.

B. Advice: Conduct performance reviews for high-risk hospitals and implement best practices to improve outcomes.

Top 10 Hospitals by Risk Adjusted Mortality Rate

Hospital	Average Risk Adjusted Mortality Rate
Providence Redwood Memorial Hospital	12.23
Palomar Medical Center Poway	11.32
Kaiser Foundation Hospital – West LA	11.31
Fountain Valley Regional Hospital & Medical Center – Euclid	11.11
Chapman Global Medical Center	10.28
Providence Cedars Sinai Tarzana Medical Center	9.82
Ridgecrest Regional Hospital	9.57
Mendocino Coast District Hospital	9.57
Hazel Hawkins Memorial Hospital	9.45
USC Arcadia Hospital	9.41

4. How do deaths and cases change over the years?

A. Finding: 2019 had the highest number of cases, highlighted in orange on the grouped bar chart.

B. Advice: Investigate contributing factors to the 2019 spike and allocate resources to manage similar trends.

Yearly Comparison of Deaths and Cases
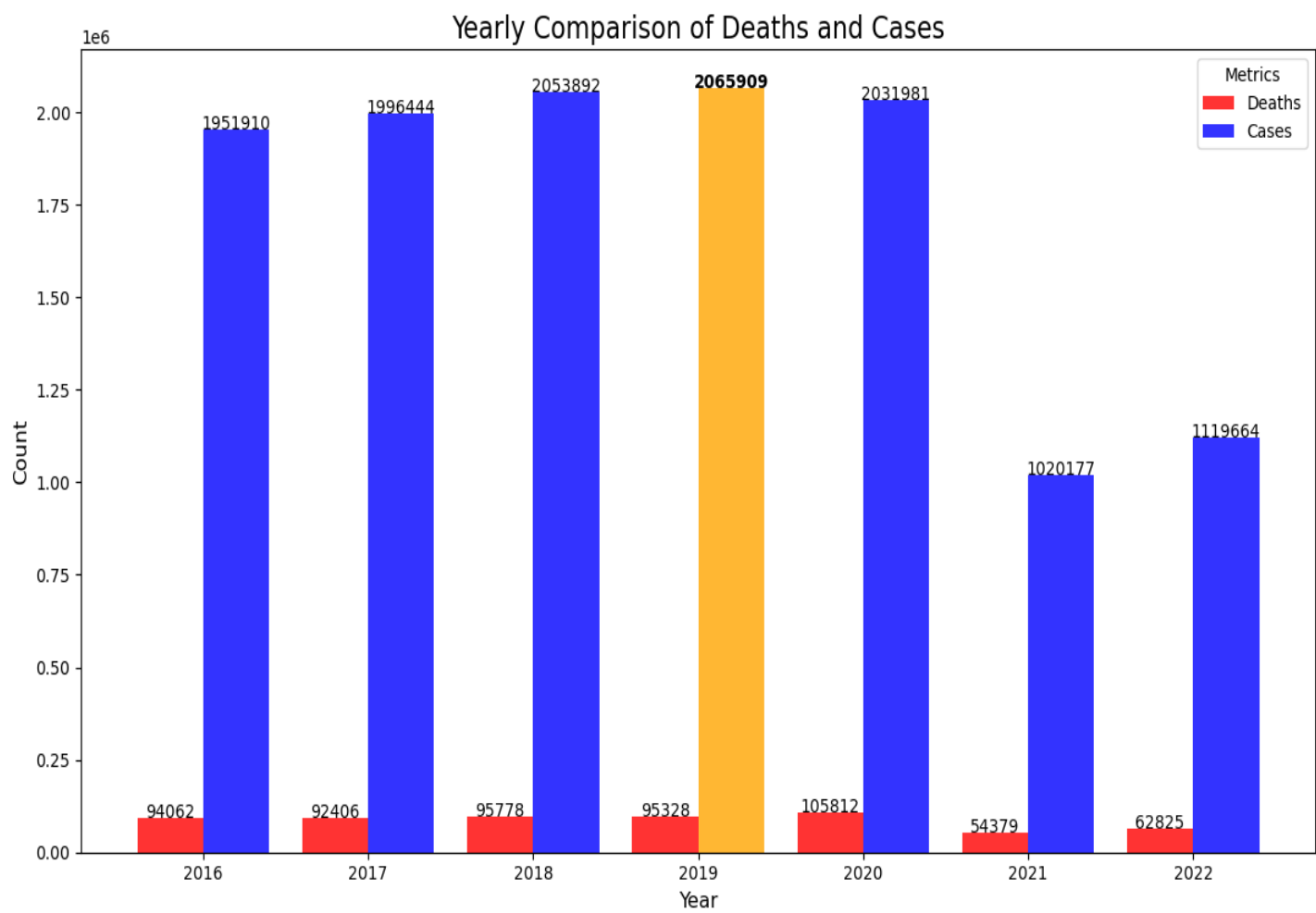
5. What trends can be observed in risk-adjusted mortality rates over time?

A. Finding: The mortality rate saw a significant dip in 2018, highlighted with a green dot.

B. Advice: Study successful practices from 2018 and replicate them to sustain lower mortality rates.
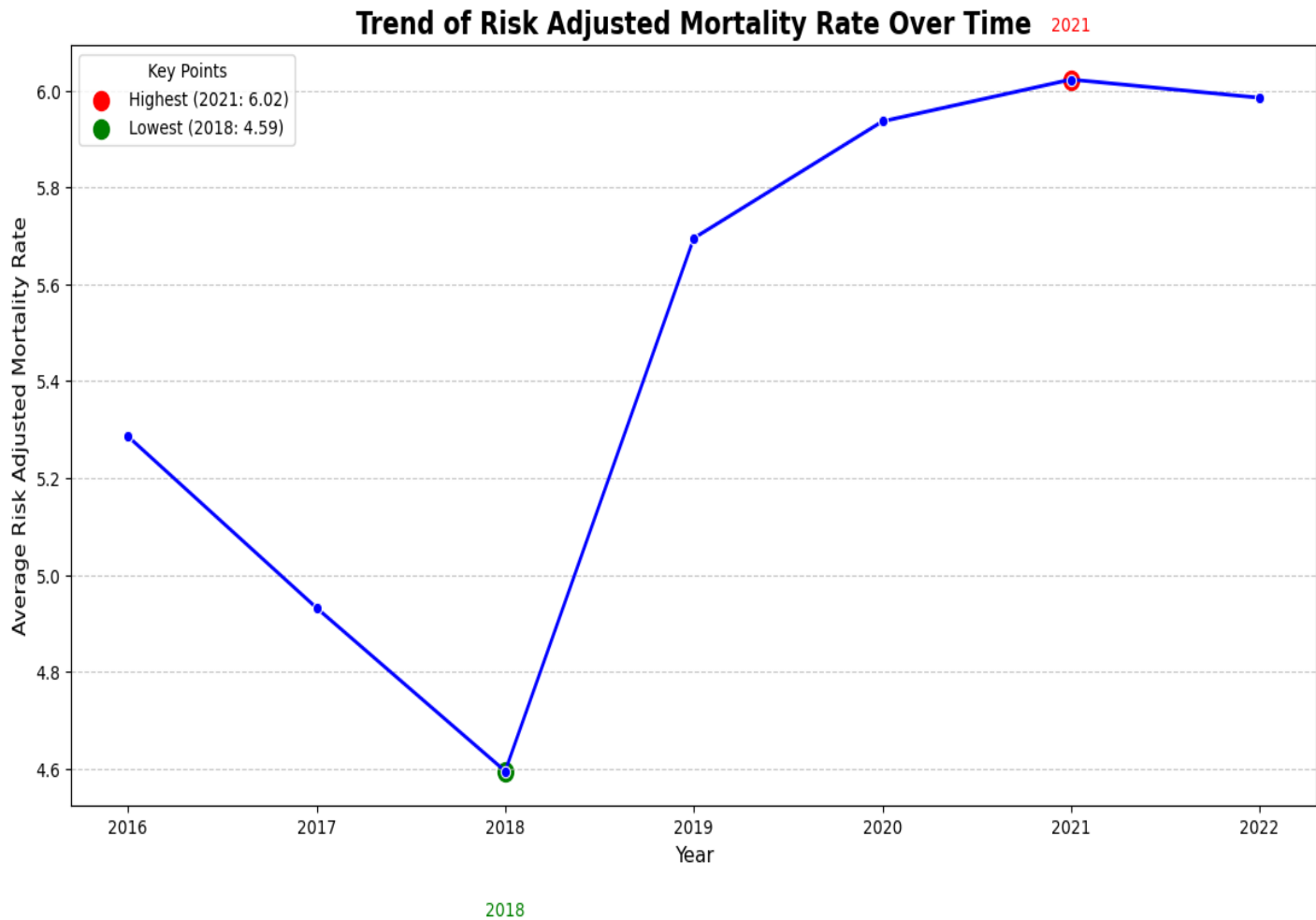
Key Visualizations

1. **Top 10 High-Risk Procedures**:
    a. Horizontal bar chart with the highest procedure highlighted in red.
2. **Deaths and Cases by Procedure**:
    a. Stacked bar chart with deaths in red and cases in blue.
3. **Top 10 High-Risk Hospitals**:
    a. Horizontal bar chart highlighting the highest-risk hospital in red.
4. **Yearly Comparison of Deaths and Cases**:
    a. Grouped bar chart with 2019 (highest cases) highlighted in orange.
5. **Trend of Risk-Adjusted Mortality Rates Over Time**:
    a. Line chart with the lowest point (2018) marked in green.

Conclusion

The analysis successfully identified high-risk procedures, hospitals with high mortality rates, and trends in healthcare data. These findings enable targeted actions to improve patient care, optimize resource allocation, and enhance hospital performance.

Recommendations Future

1. **Focus on High-Risk Procedures**:

     a. Implement specialized training and allocate more resources for procedures with high mortality rates.

2. **Support High-Risk Hospitals**:
   a. Conduct audits and provide performance reviews for hospitals with high mortality rates.

3. **Resource Planning**:
   a. Leverage yearly trends to anticipate and prepare for patient loads effectively.

4. **Continuous Monitoring**:
   a. Regularly update the dataset to track new trends and ensure proactive interventions.

5. **Explore Demographic Factors**:
   a. Include demographic data in future analyses to identify disparities and improve healthcare equity.