MAVEN HEALTHCARE CHALLENGE

MAVEN ANALYTICS

HOSPITAL CONSUMER ASSESSMENT OF HEALTHCARE PROVIDERS AND SYSTEMS (HCAHPS) SURVEY

DATA INSIGHTS AND PRESENTATION

SYED ASAD ALI KAZMI

DATA ANALYST, AMERICAN HOSPITAL ASSOCIATION (AHA)

DATA QUALITY ASSESSMENT FOR HCAHPS SURVEY DATA

- Data Profiling: Examined key statistics, data distribution, and identified gaps in the survey responses.
- Data Cleansing: Removed duplicate entries, addressed missing responses, and standardized variable formats.
- Data Accuracy: Validated responses against source documents, corrected inaccuracies, and resolved outliers.
- Consistency and Integrity: Checked for logical consistency between responses, ensuring data coherence.
- Timelines: Verified survey period alignment, excluded outdated data points.
- Documentation: Ensured complete metadata, established clear traceability of survey questions.
- Quality Metrics: Defined and measured quality metrics, including response validity and internal consistency.
- Results: Strengthened data reliability, contributing to more accurate healthcare insights and improved patient experiences.

DATA OVERVIEW

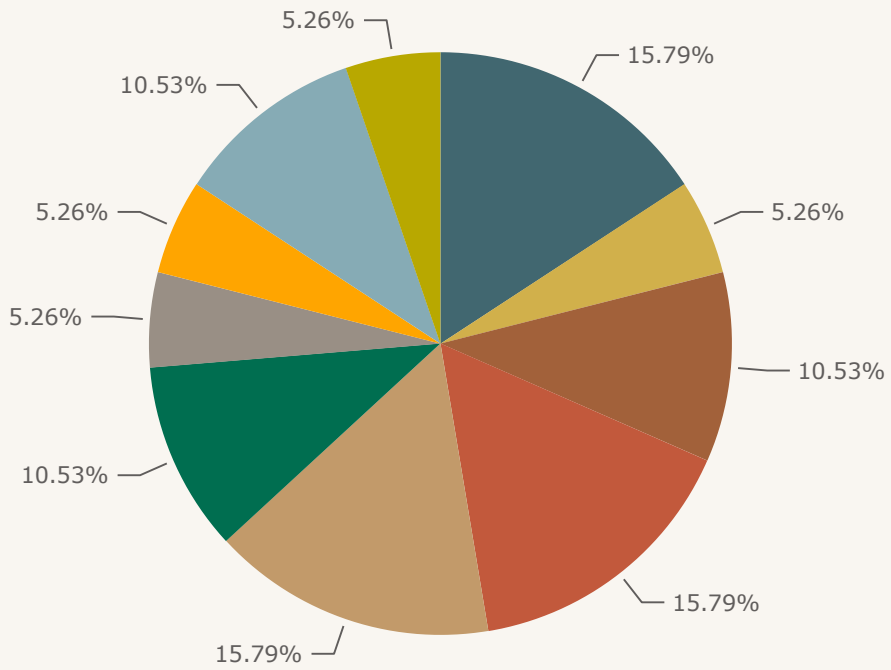
- Dataset Source: National & state-level scores from 2013 to 2022 for the HCAHPS survey, a standardized assessment of inpatient hospital experiences.
- Data Structure: The dataset is a consolidation of seven CSV files: `reports.csv`, `states.csv`, `measures.csv`, `questions.csv`, `national_results.csv`, `state_results.csv`, and `responses.csv`.
- Objective: Analyze and derive meaningful insights to enhance healthcare services and patient satisfaction.
- Data Preparation: Merged and integrated the seven CSV files to create a unified dataset for streamlined analysis.
- Key Features:
  - Patient feedback scores and responses.
  - Hospital performance across states and regions.
  - Bottom, middle, and top-box percentage metrics.
  - Survey questions and measures.
- Importance: Supports evidence-based decision-making for healthcare improvements and patient-centric services.
- Next Steps: Exploratory data analysis, data visualization, and advanced analytics to uncover patterns, trends, and actionable insights.

TRANSFORMING HEALTHCARE DATA INTO INSIGHTS

- Data Collection: Imported diverse CSV files using Python's `pd.read_csv`.
- Consistency Ensured: Standardized column names for uniformity across DataFrames to lowercase.
- Data Fusion: Integrated DataFrames into a comprehensive dataset named 'merged.'
- Valuable Output: Created 'hcahps.csv,' a harmonized dataset for deeper analysis.
- Insight Preview: Examined 'merged' DataFrame using `.info()` to assess structure, memory usage, and potential data quality.

HCAHPS SURVEY DATA INSIGHTS

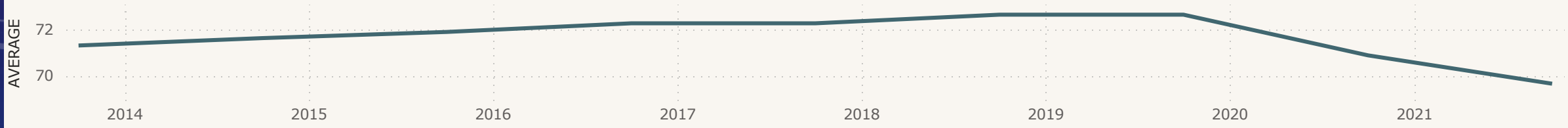
9	71.69	5016	19	9	51
Years	Average Patient Satisfaction S…	Total Facilities	No. of Questions Asked	Regions	States



Measure
- Care Transition
- Cleanliness of Hospital Environment
- Communication about Medicines
- Communication with Doctors
- Communication with Nurses
- Discharge Information
- Overall Hospital Rating
- Quietness of Hospital Environment
- Responsiveness of Hospital Staff
- Willingness to Recommend the Hospital

release pe…
- 07_2015
- 07_2016
- 07_2017
- 07_2018
- 07_2019
- 07_2020
- 07_2021
- 07_2022
- 07_2023

Patient Satisfaction Over Years

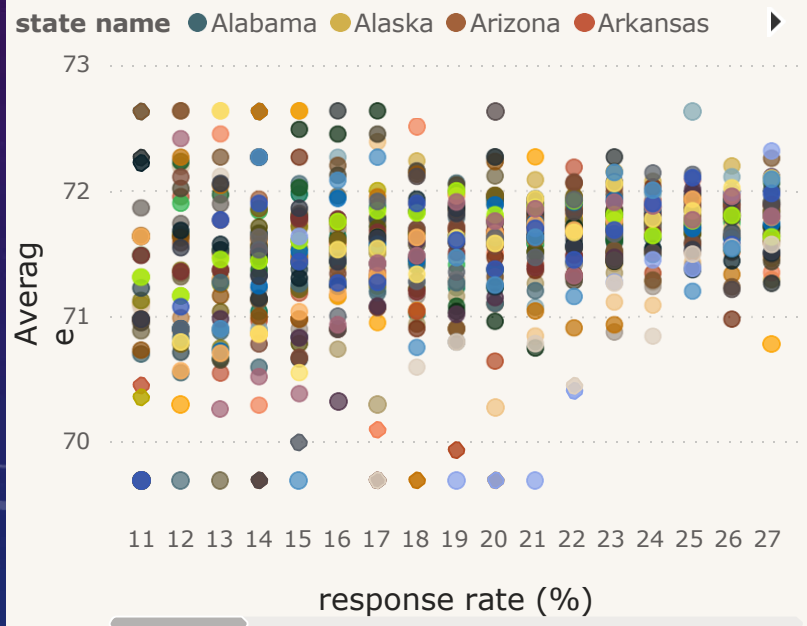


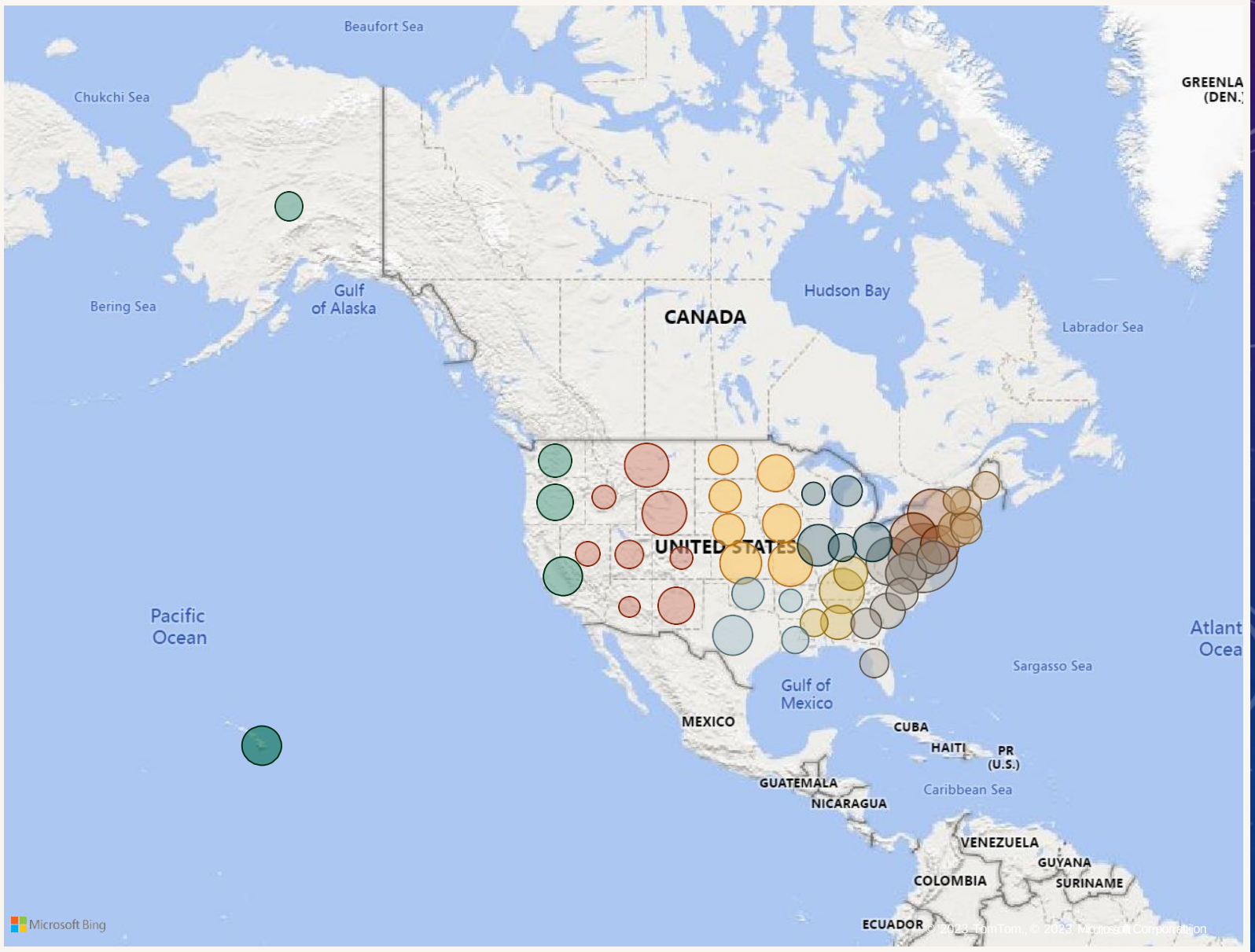
AVERAGE TOP BOX PERCENTAGE

state name

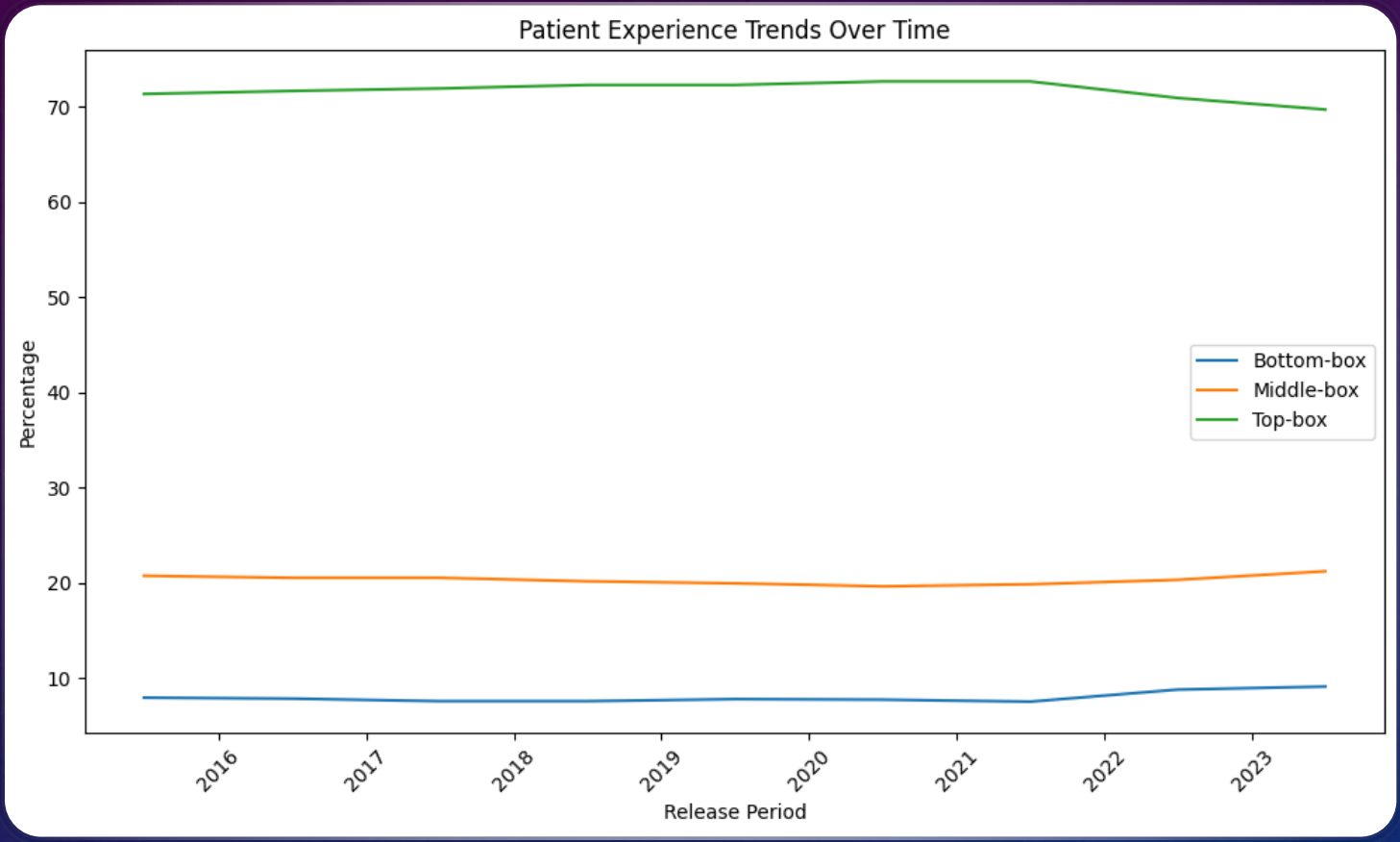
- Alabama
- Alaska
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Delaware
- District of Colum...

state name	Average
Maryland	71.72
New York	71.70
West Virginia	71.70
Pennsylvania	71.70
Tennessee	71.70
Missouri	71.70
Wyoming	71.70
Montana	71.70
Kansas	71.70
District of Columbia	71.70
Illinois	71.70
Virginia	71.70
Total	**71.69**





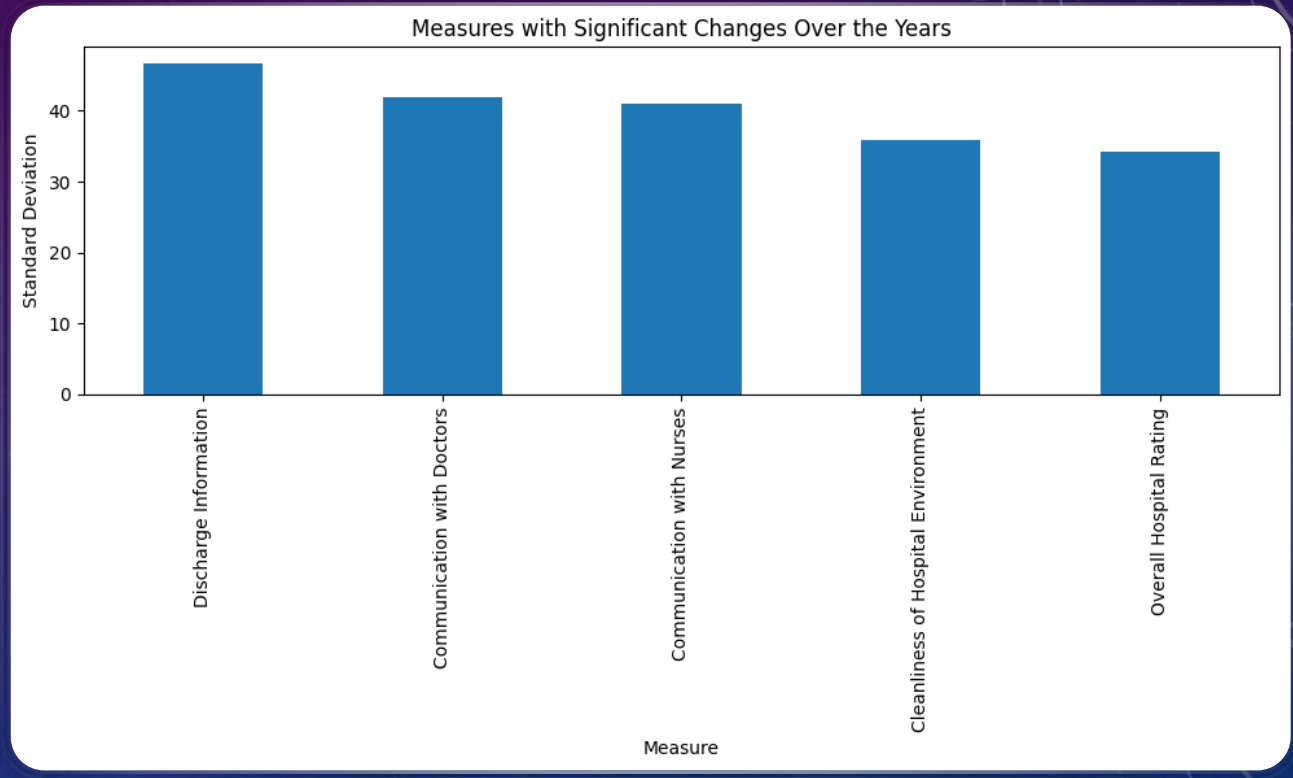
TRENDS IN PATIENT EXPERIENCE SCORES OVER TIME

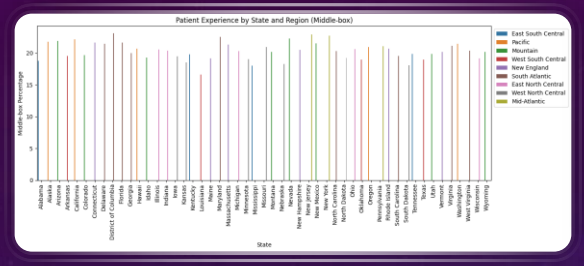


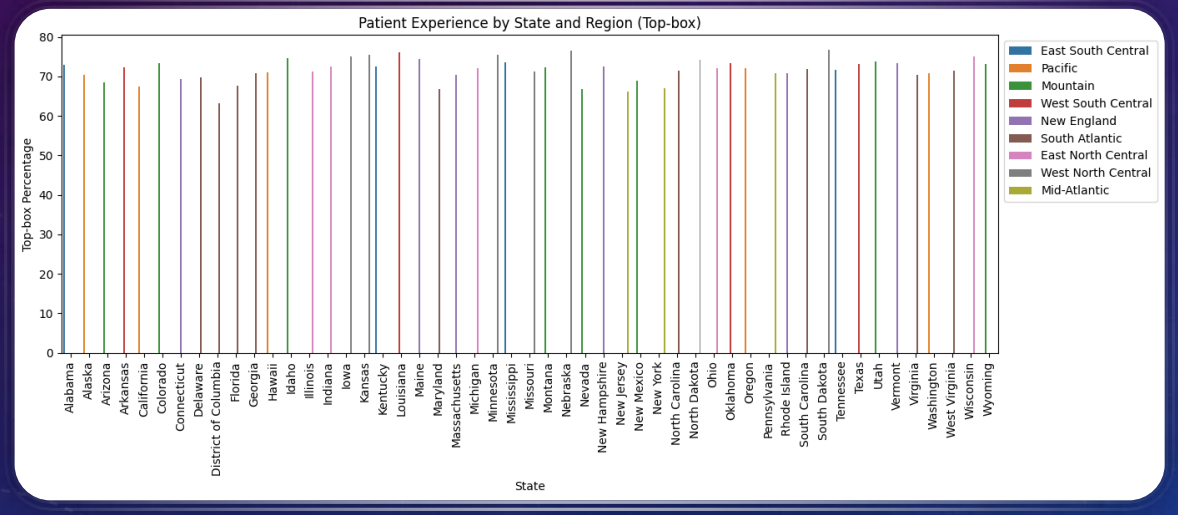
- The line plots showcase mean percentages for 'bottom-box', 'middle-box', and 'top-box' scores across release periods (years).
- 'Top-box' percentages exhibit a decline in recent years (2022 and 2023) after steady levels.
- The analysis indicates a potential decline in patient satisfaction, especially in the 'top-box' category, in recent years—highlighting a potential area for improvement or decision-making.

IDENTIFYING CHANGING MEASURES IN PATIENT EXPERIENCE

- Highlights the top 5 measures with the highest standard deviations, indicating significant score changes over years.
- Measures listed in descending order of standard deviations.
- Larger values suggest increased variability and potential score changes over time.
- Indicate focus areas for analysis and improvement to address patient satisfaction changes.







PATIENT EXPERIENCE ANALYSIS BY STATE AND REGION

Top-box Percentage Analysis

- Generally high top-box percentages indicate positive patient experiences in most states and regions.
- 'West South Central' and 'East South Central' regions show slightly lower top-box percentages.
- States like 'Montana,' 'Nebraska,' and 'North Dakota' display higher top-box percentages, implying better patient experiences.

Overall Insights

- Comparative view of patient experiences across states and regions.
- Identifies areas for enhanced patient experiences, particularly where bottom-box percentages are relatively high.
- Consistently high top-box percentages reflect positive patient experiences overall, with some regional variations.

OVERALL PATIENT EXPERIENCE

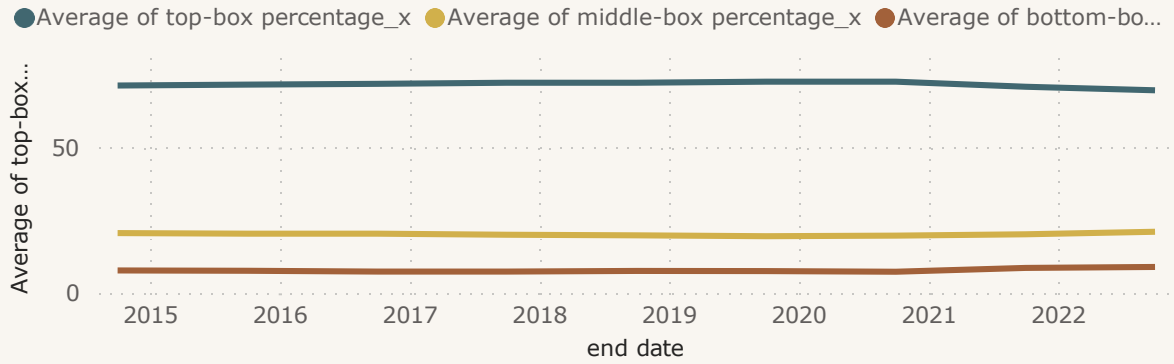
state name

- Alabama
- Alaska
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Delaware
- District of Colum…

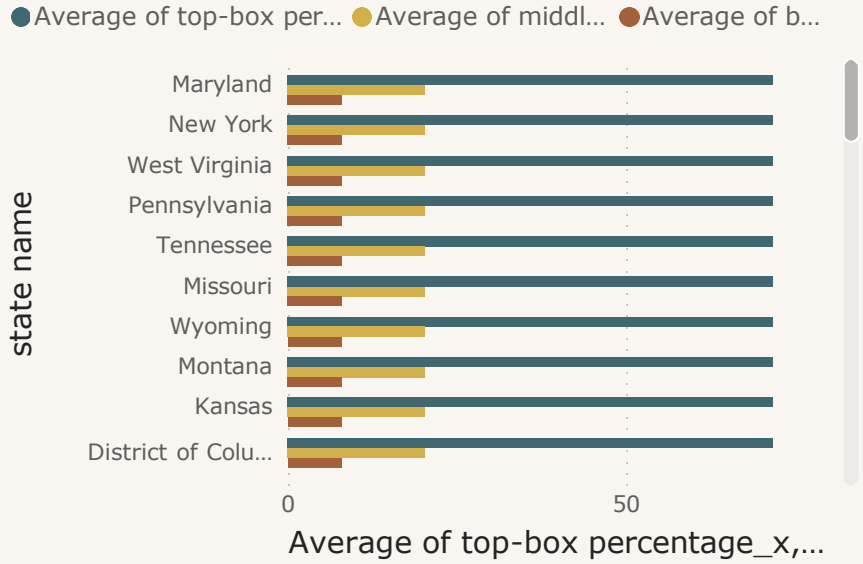
release peri…

- 07_2015
- 07_2016
- 07_2017
- 07_2018
- 07_2019
- 07_2020
- 07_2021
- 07_2022
- 07_2023

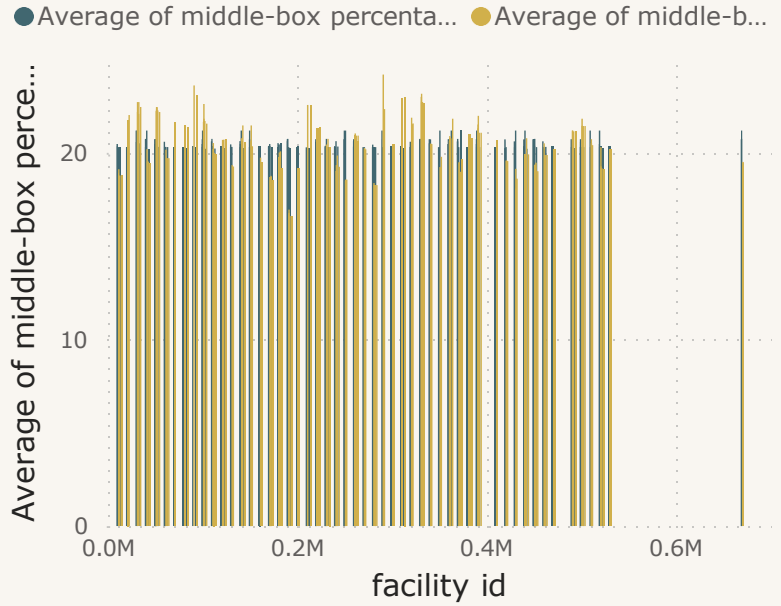
Overall Patient Experience Trend



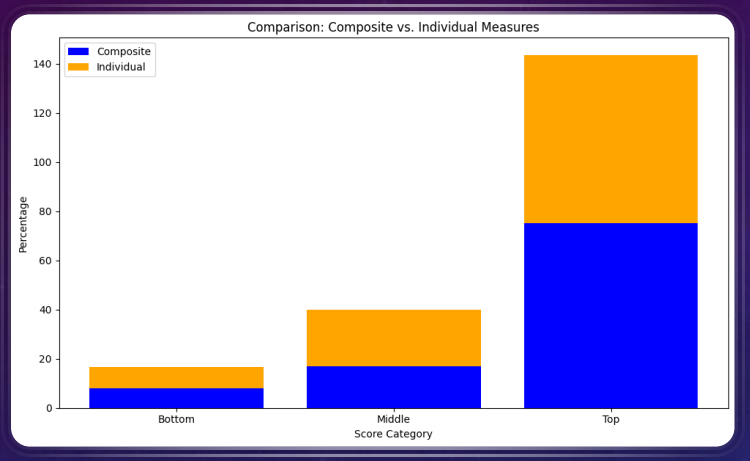
Highest and Lowest Patient Experience Scores By State



Communication and Responsiveness



COMPARISON: COMPOSITE VS. INDIVIDUAL MEASURES IN PATIENT EXPERIENCE



Data categorized into composite measures and individual measures.

Mean Composite Data:

- Bottom-box: Avg. 7.91% (lower satisfaction).
- Middle-box: Avg. 16.99% (moderate satisfaction).
- Top-box: Avg. 75.10% (high satisfaction).

Mean Individual Data:

- Bottom-box: Avg. 8.67%.
- Middle-box: Avg. 23.05%.
- Top-box: Avg. 68.28%.

Interpretation and Insights:

- Composites show higher top-box (better satisfaction).
- Individual measures have more middle-range responses.
- Higher bottom-box/middle-box in individuals suggests improvement areas.
- Helps understand patient responses distribution.

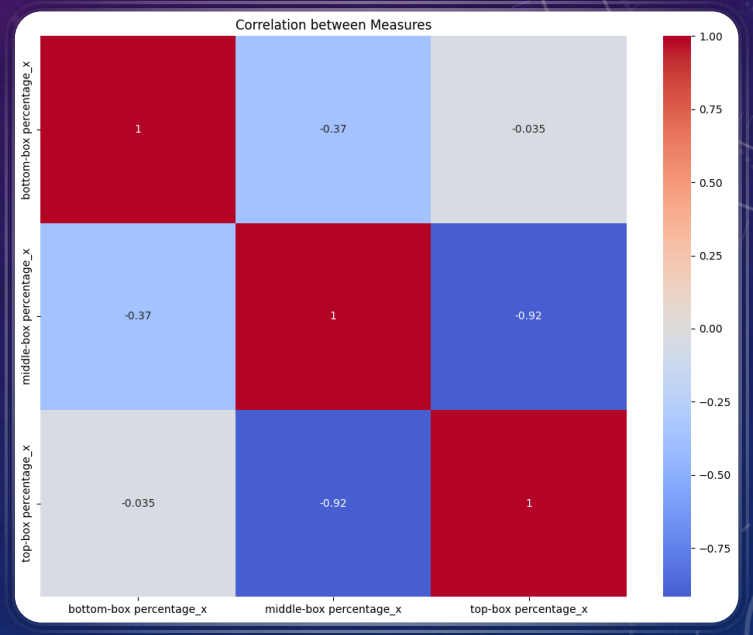
CORRELATION ANALYSIS: PATIENT EXPERIENCE MEASURES

Explores relationships between satisfaction measures.

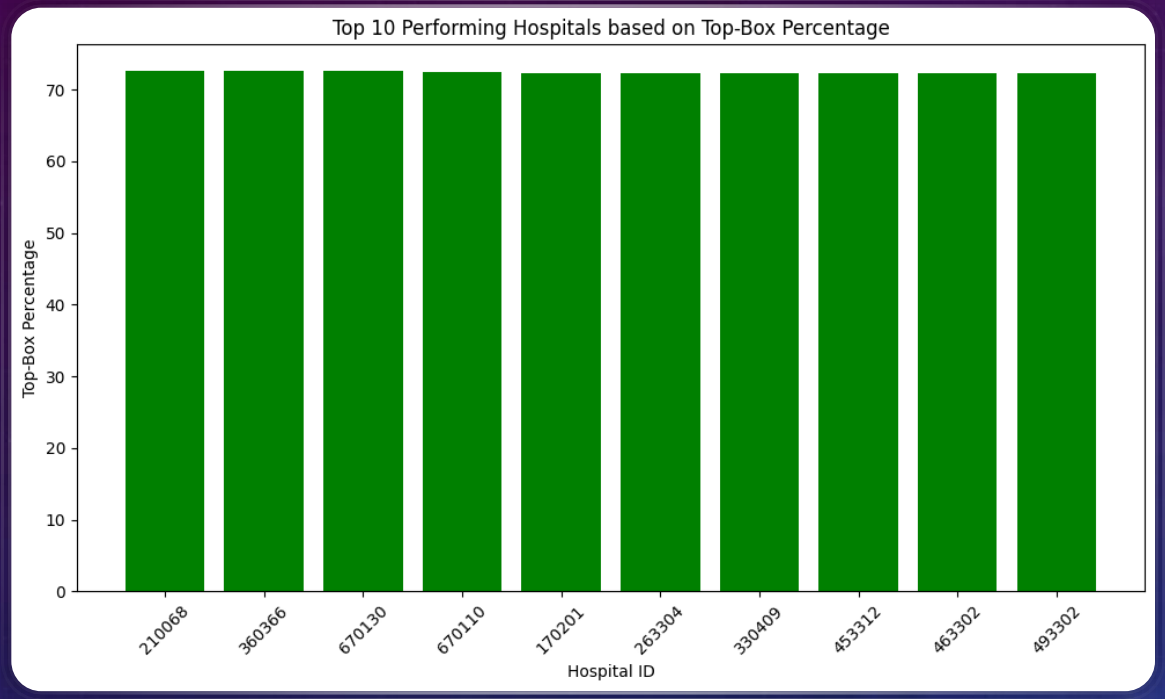
- 'bottom-box %' positively correlates with 'middle-box %' (-0.37).
- 'middle-box %' strongly negatively correlates with 'top-box %' (-0.92).
- Darker colors indicate stronger correlations.
- Lighter colors imply weaker or no correlations.

Interpretation:

- Negative correlation between 'middle-box %' and 'top-box %'.
- As moderate satisfaction drops, high satisfaction rises.

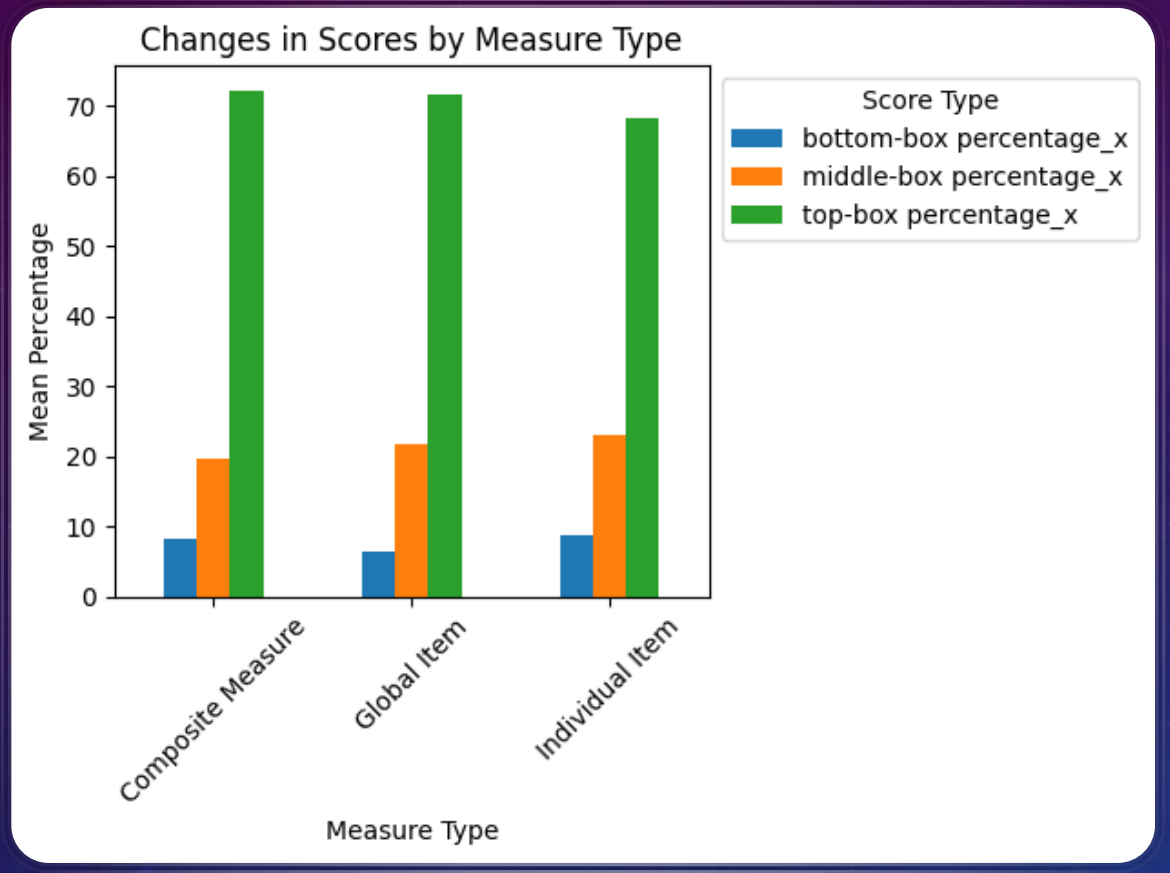


TOP-PERFORMING HOSPITALS: PATIENT EXPERIENCE ANALYSIS



- Identifying top-performing hospitals based on patient experience scores.
- Highlights hospitals in the top 10% for highest 'top-box %'.
- Offers insights for learning from high-performing hospitals.
- Enhances patient experiences industry-wide.

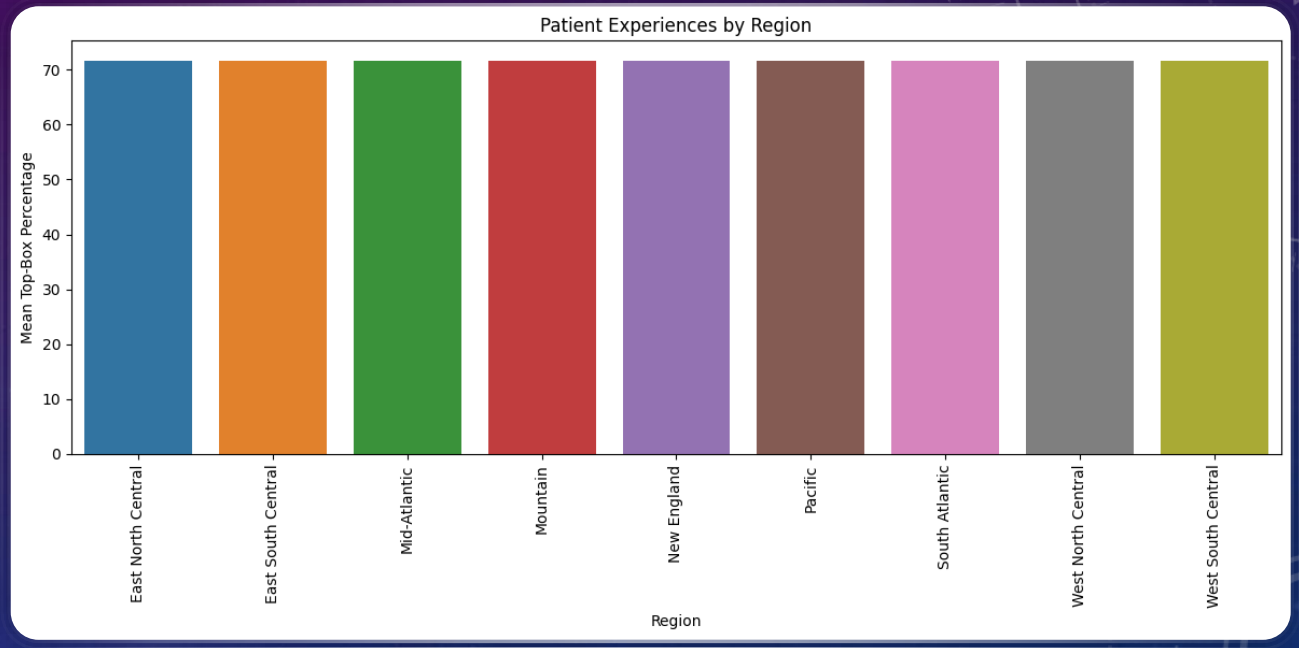
PATIENT EXPERIENCE ANALYSIS BY MEASURE TYPES



- Evaluating scores across measure types.
- Visualizing changes using a bar plot.
- Compares satisfaction across composite, global, and individual measures.
- Aids understanding of contribution to overall satisfaction.

PATIENT SATISFACTION ANALYSIS BY REGIONS AND STATES

- Evaluating scores across regions and states.
- Visualizing using a bar plot.
- Compares satisfaction among regions and states.
- Identifies variations for targeted improvement strategies.



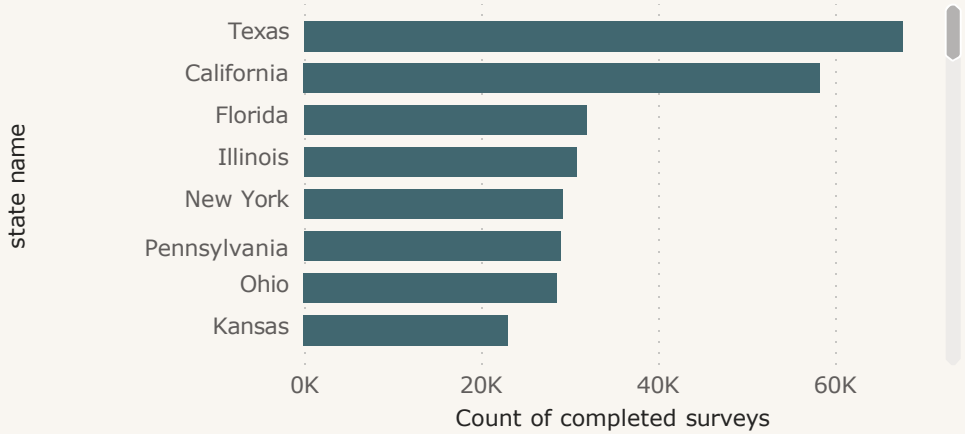
PATIENT DEMOGRAPHIC ANALYSIS

region	Count of completed surveys	Average of top-box perce
East North Central	124526	
East South Central	65778	
Mid-Atlantic	69521	
Mountain	71668	
New England	30780	
Pacific	92112	
South Atlantic	117249	
West North Central	116755	
West South Central	122018	
Total	**810407**	

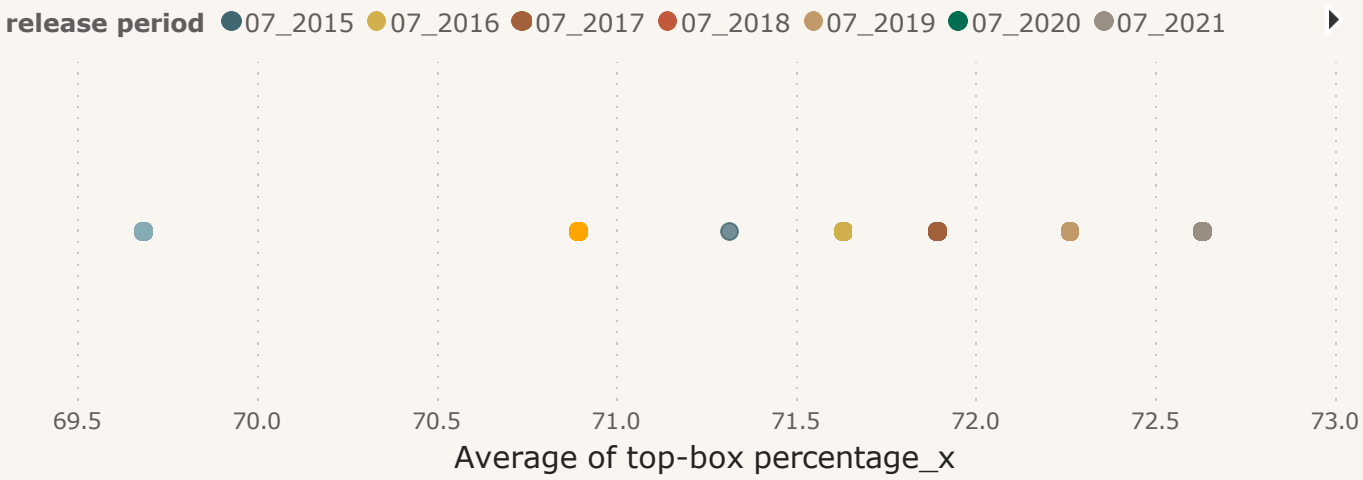
region

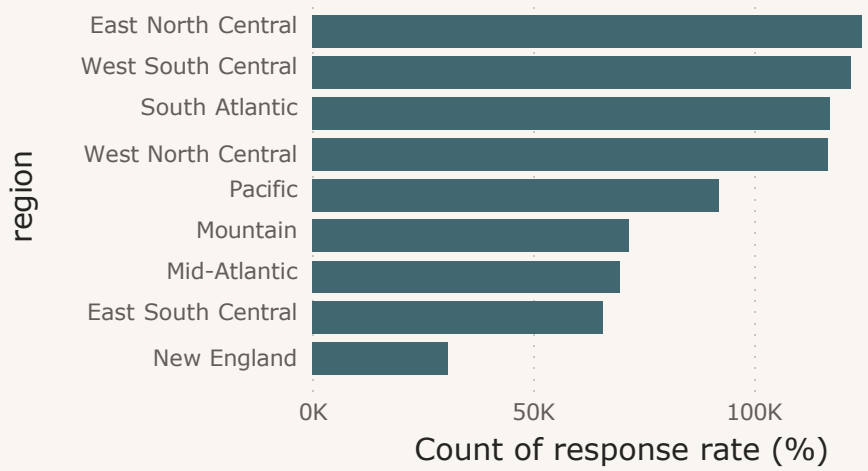
- East North Central
- East South Central
- Mid-Atlantic
- Mountain
- New England
- Pacific
- South Atlantic
- West North Central
- West South Central

RESPONSE RATES AND COMPLETION



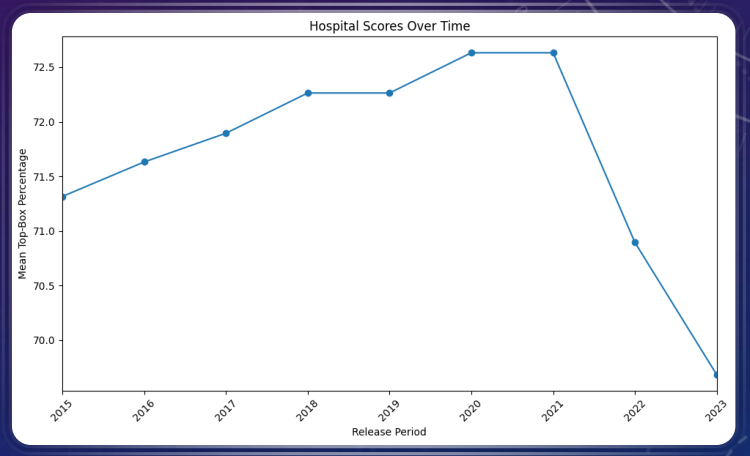
IMPACT OF SURVEY COMPLETION ON PATIENT EXPERIENCE



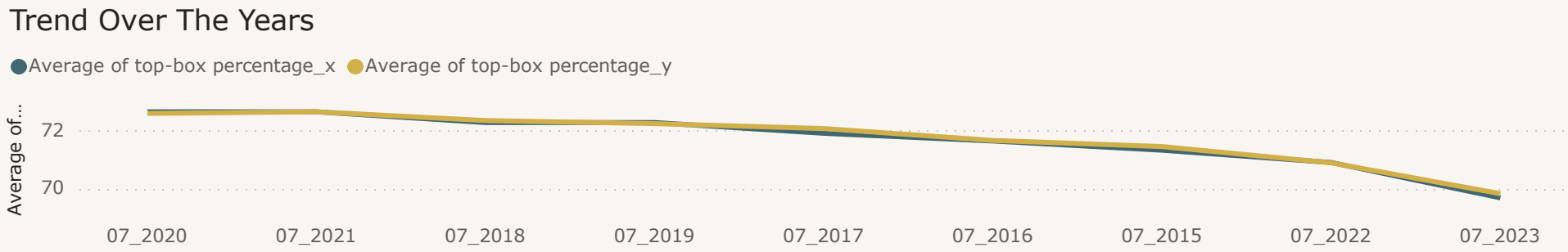


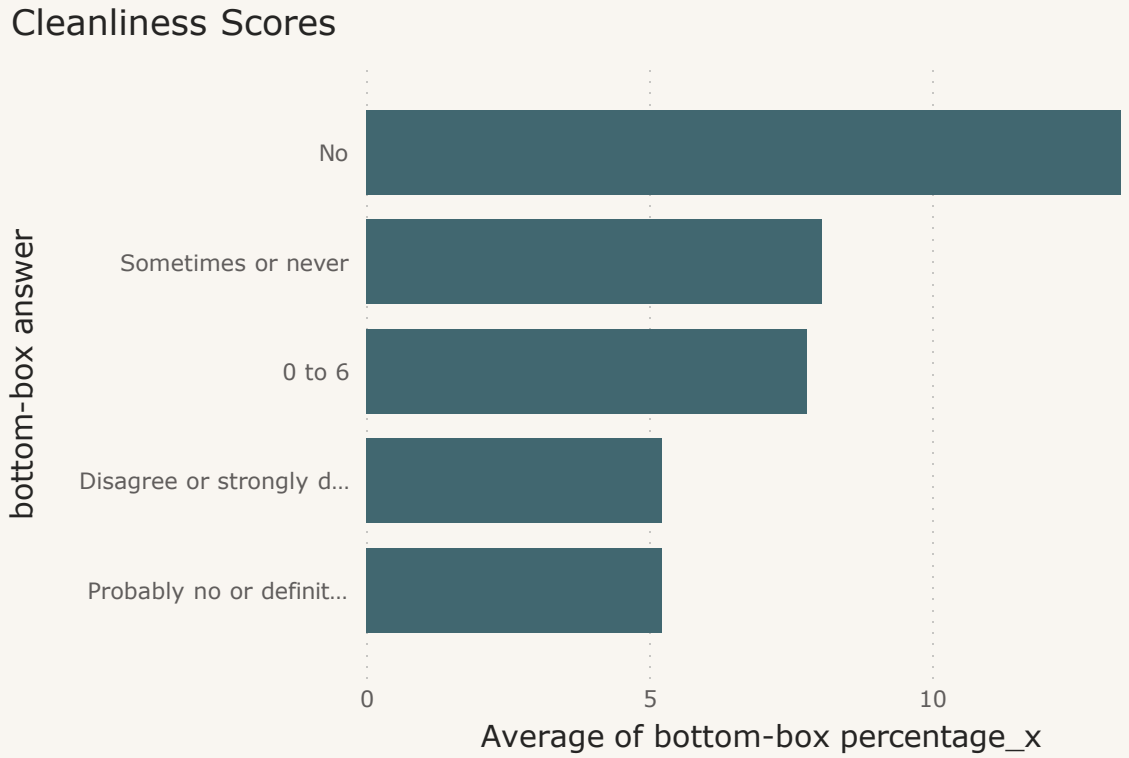
TIME SERIES ANALYSIS

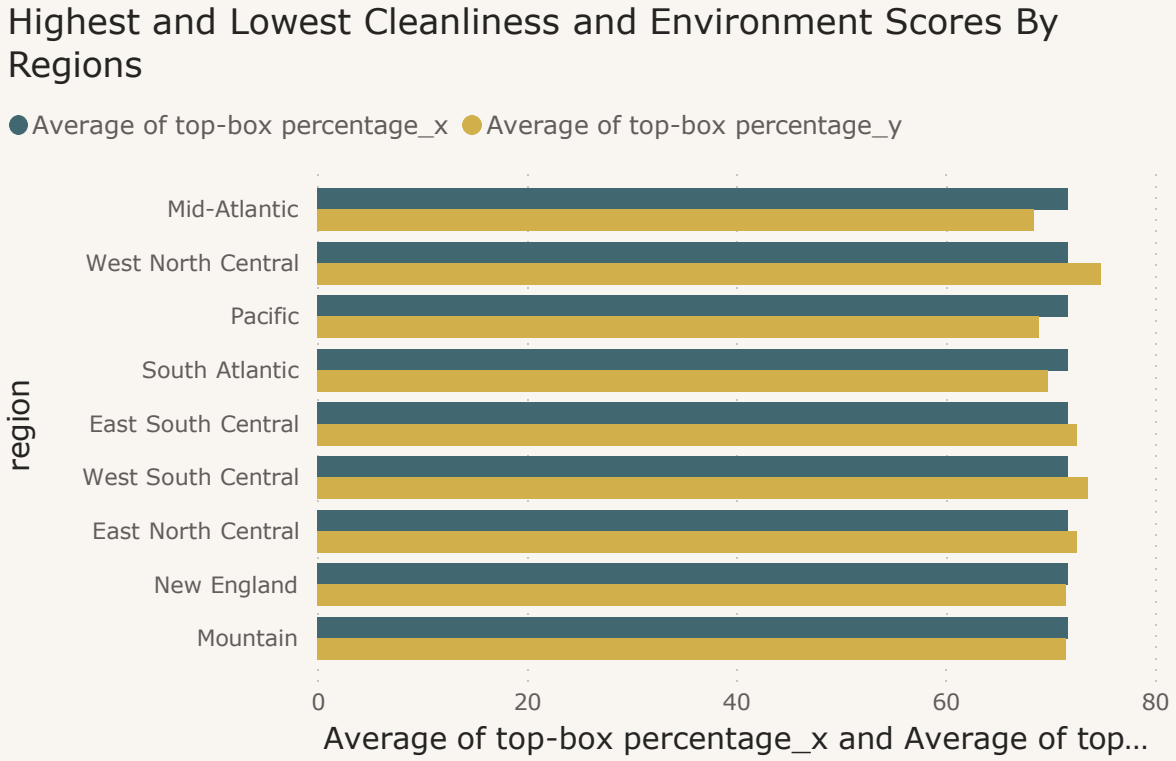
- Tracking scores across release periods.
- Visualizing using a line plot.
- Observing trends and fluctuations in satisfaction scores.
- Identifying patterns, seasonality, or significant changes.

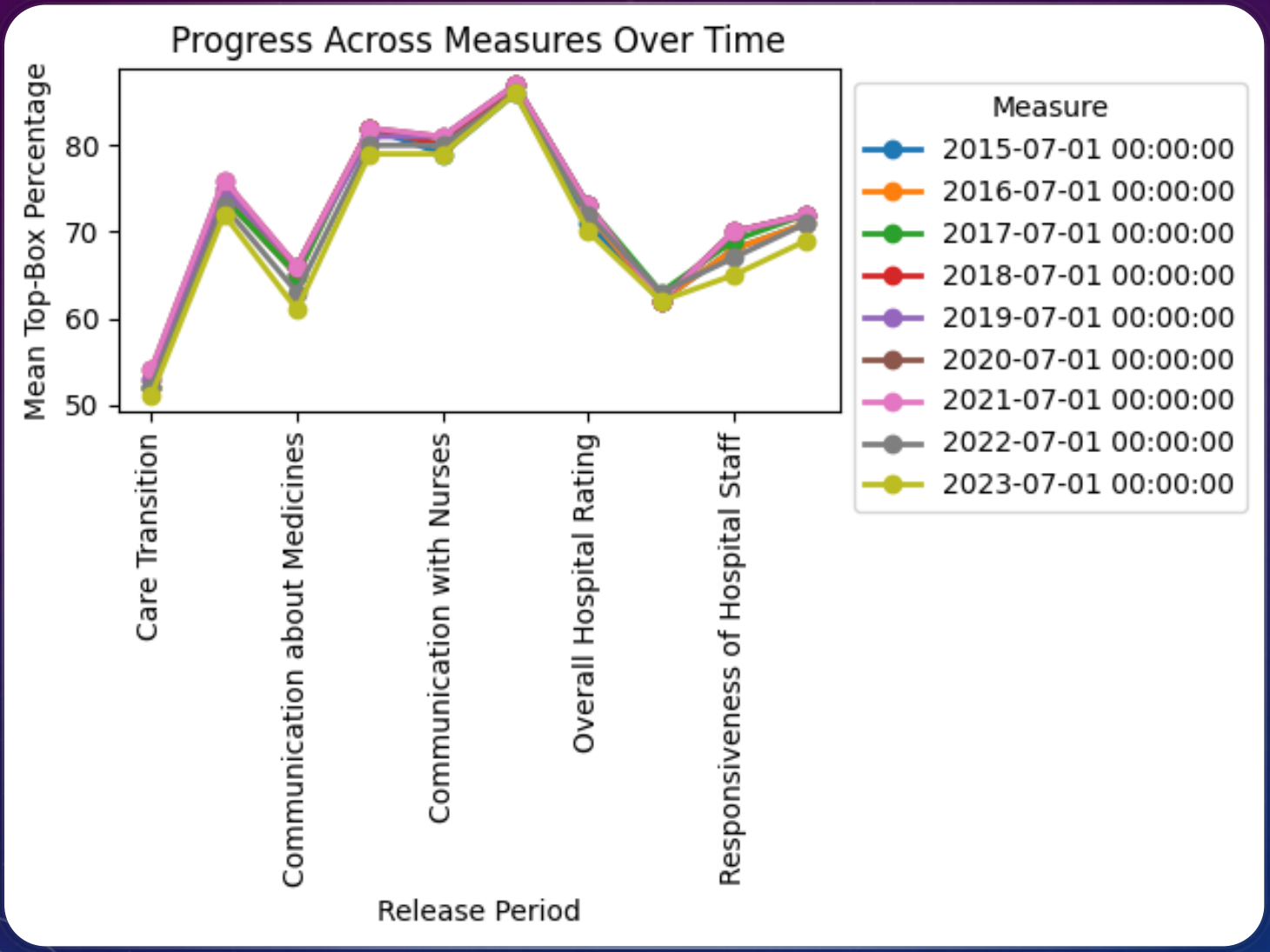


CLEANLINESS AND ENVITONMENT









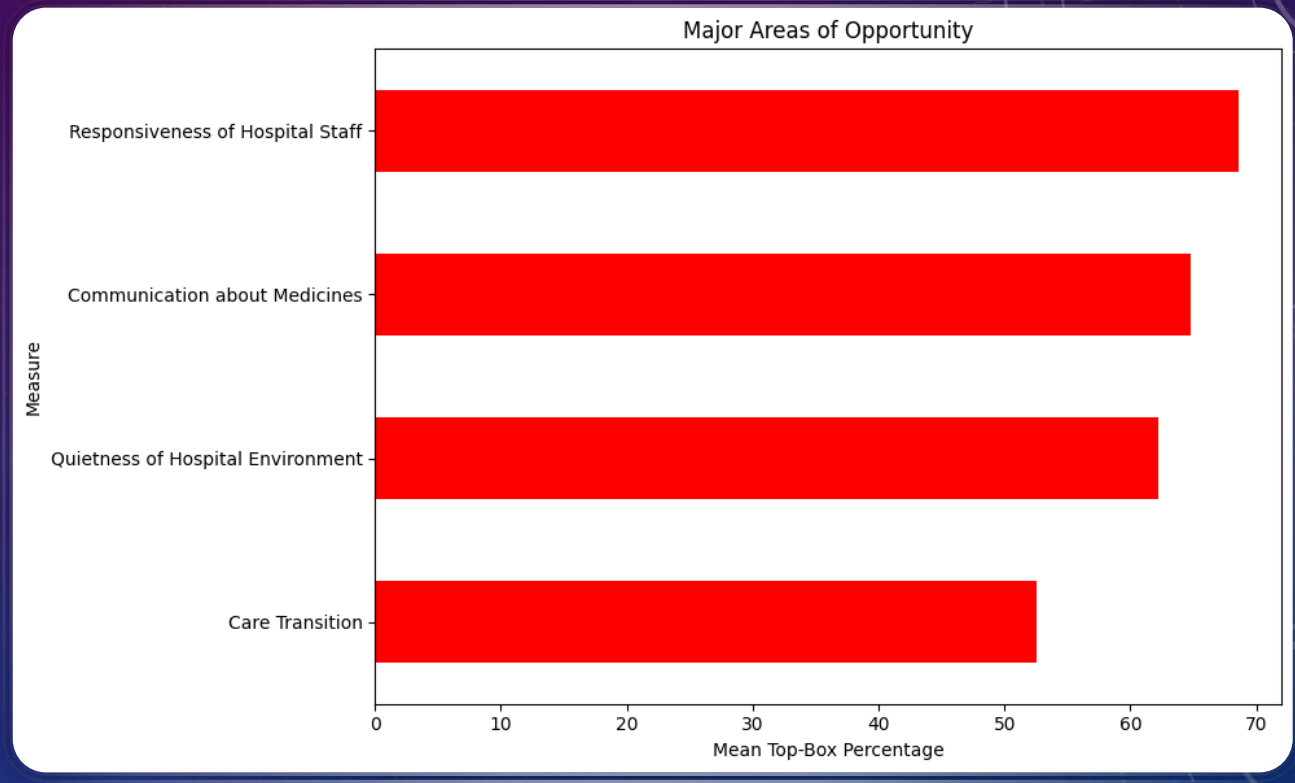
DETAILED TIME SERIES ANALYSIS: MEASURE PROGRESS

- Mean top-box percentage scores per measure over release periods.
- Line plot showcases progress across measures.
- X-axis: Release periods, Y-axis: Mean top-box % scores.
- Legend outside plot for clarity.

OPPORTUNITY ANALYSIS: IDENTIFYING IMPROVEMENT AREAS

Identifies measures with consistently low scores or no significant improvement (mean top-box % below 70).

- Filters and identifies areas for improvement.
- Creates `opportunity_measures` Series.
- Focus on enhancing patient experiences and satisfaction.



RECOMMENDATIONS

1

Focus on measures like "Communication about Medicines" and "Care Transition," which display consistent opportunities for improvement. Implement targeted initiatives to elevate patient satisfaction.

2

Allocate resources based on regional analysis. Regions like "South Atlantic" and "East South Central" have potential for improved patient experiences. Tailor strategies to address specific regional needs.

3

Leverage time series analysis to monitor and respond to changing patient satisfaction trends over the years. Continuous improvement efforts should align with evolving patient expectations.

THANK YOU