

Nurse Staffing Optimization for Healthcare Providers

Objective

The objective of this analysis was to explore the relationship between nurse staffing hours and various factors such as facility ownership type, staffing levels during weekdays vs. weekends, and facility ratings. The goal was to generate actionable insights for a healthcare staffing provider's sales team to better understand staffing patterns and enhance their strategy for healthcare staffing optimization.

Overview

The analysis was conducted through the following steps:

- 1. Data Acquisition and Storage**
 - Downloaded the [Payroll Based Journal Daily Nurse Staffing](#) (`pbj_transformed`) and [Provider Information](#) datasets (`provider_transformed`).
 - Stored the CSV files in an S3 bucket on AWS for easy access and processing.
- 2. Data Cleaning and Preprocessing**
 - Used AWS Lambda to clean and standardize data (e.g., column formatting, duplicate removal, missing value handling).
- 3. Data Transformation and Cataloging**
 - Uploaded the cleaned data into a new S3 bucket and used AWS Glue to crawl the tables.
 - Configured AWS Athena to query and analyze the data. Tables were created in Athena for easier querying.
- 4. Analysis**
 - Wrote SQL queries in Athena to analyze staffing data across different parameters such as:
 - Comparison of staffing levels by ownership type.
 - Weekend vs. weekday staffing levels by state.
 - RN staffing hours by facility rating.
- 5. Machine Learning Model Implementation**
 - Connected Athena to Jupyter Notebook in SageMaker to query the data for deeper analysis.
 - Used Scikit-learn to predict weekend staffing hours from weekday hours, applying multiple regression models (e.g., Linear Regression, Random Forest).
- 6. Visualization**
 - Created visualizations using Matplotlib to communicate insights and trends found in the data.

Analysis and Results

1. Comparison of Ownership Type Staffing Levels

Purpose: To understand how different ownership types (government, non-profit, for-profit) affect nurse staffing levels in healthcare facilities.

```
SELECT prov.ownership_type,
       CASE
         WHEN prov.ownership_type LIKE 'Government%' THEN 'Government'
         WHEN prov.ownership_type LIKE 'Non profit%' THEN 'Non-Profit'
         WHEN prov.ownership_type LIKE 'For profit%' THEN 'For-Profit'
         ELSE 'Other'
       END AS ownership_category,
       AVG(pbj.hrs_rn + pbj.hrs_lpn + pbj.hrs_cna) / AVG(NULLIF(pbj.mdscensus, 0))
```

```
AS avg_nurse_hours_per_resident
FROM pbj_transformed pbj
JOIN provider_transformed prov
ON pbj.provnum = prov.cms_certification_number_ccn_
GROUP BY prov.ownership_type
ORDER BY avg_nurse_hours_per_resident DESC;
```

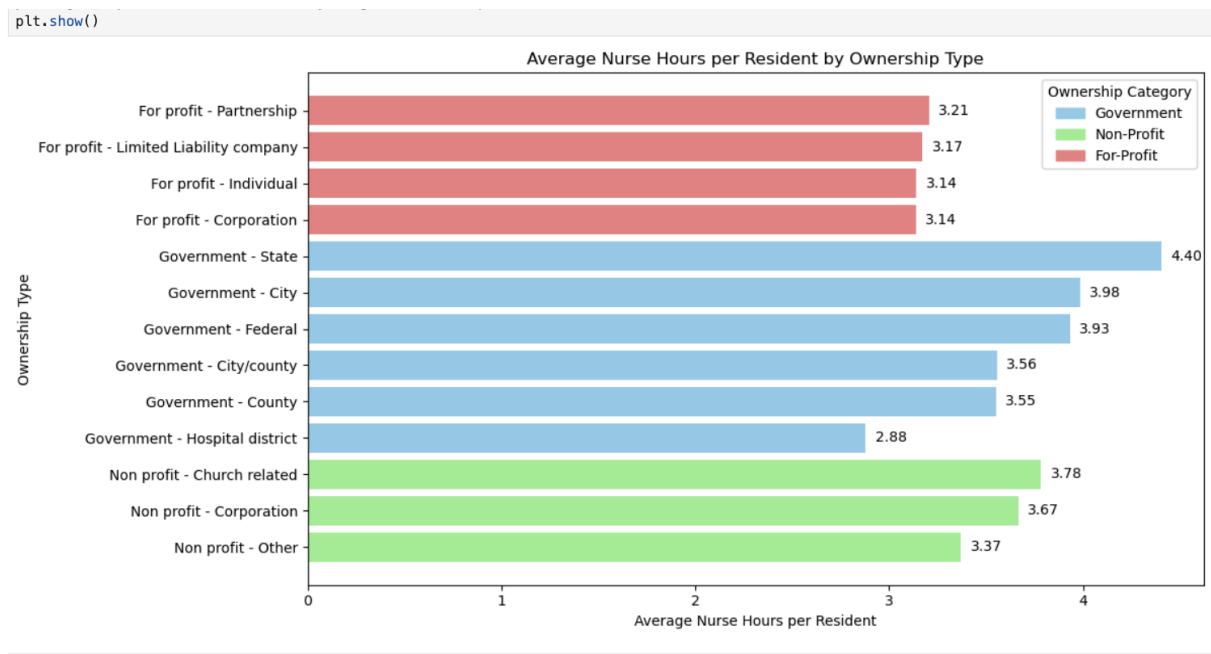
Findings:

- Government facilities have the highest nurse staffing levels (4.40 hours per resident), suggesting strict regulations and better funding.
- Non-profit facilities have moderate staffing levels (3.37 - 3.78), likely due to budget constraints.
- For-profit facilities have the lowest staffing levels (3.14 - 3.21), indicating cost-cutting measures and a higher likelihood of needing contract staffing.

Recommendations:

- **Target for-profit facilities for contract staffing solutions:** For-profit facilities have the lowest staffing levels, meaning they are more likely to need flexible staffing options to meet their care demands and control costs. Offering contract staffing gives them the flexibility to adjust staffing as needed without the commitment of permanent hires. By providing tailored staffing solutions, a the healthcare staffing provider can be a valuable partner in helping these facilities manage their workforce more effectively.

Data Visualization:



2. Weekend vs. Weekday Staffing Levels by State

Purpose: To show significant gaps in some regions and identify facilities with low weekend staffing, which can impact patient safety and regulatory compliance.

```
SELECT pbj.state,
       AVG((pbj.hrs_rn + pbj.hrs_lpn + pbj.hrs_cna) / NULLIF(pbj.mdscensus, 0)) AS weekday_hours,
       AVG(prov.total_number_of_nurse_staff_hours_per_resident_per_day_on_the_weekend) AS weekend_hours,
```

```

(AVG((pbj.hrs_rn + pbj.hrs_lpn + pbj.hrs_cna) / NULLIF(pbj.mdscensus, 0)) -
AVG(prov.total_number_of_nurse_staff_hours_per_resident_per_day_on_the_weekend)) AS staffing_gap
FROM pbj_transformed pbj
JOIN provider_transformed prov
ON pbj.state = prov.state
GROUP BY pbj.state
ORDER BY staffing_gap DESC;

```

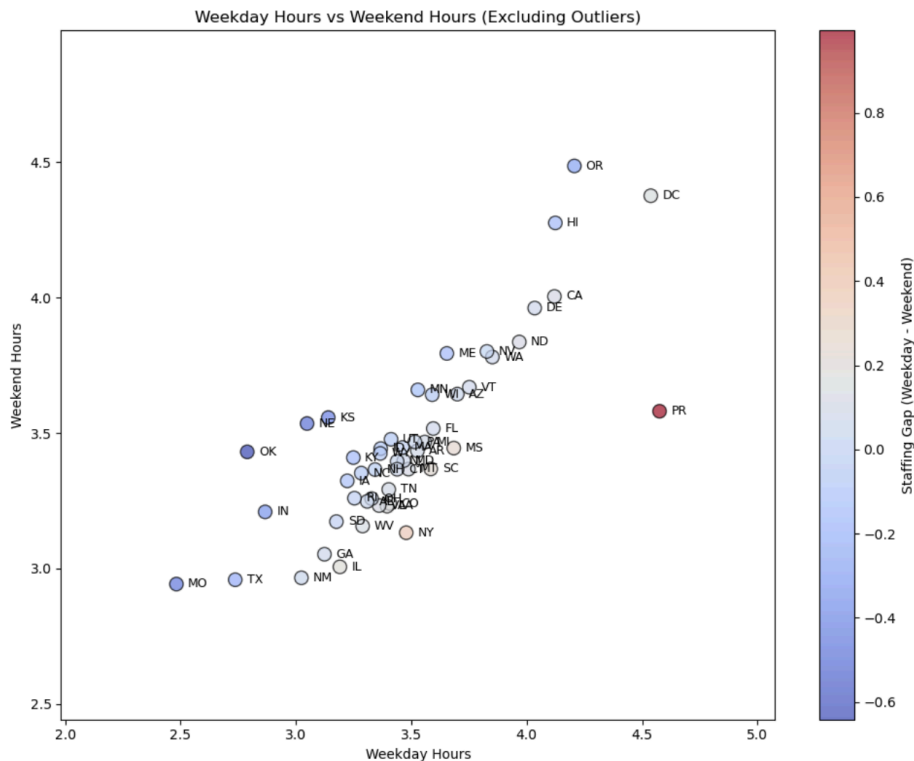
Findings:

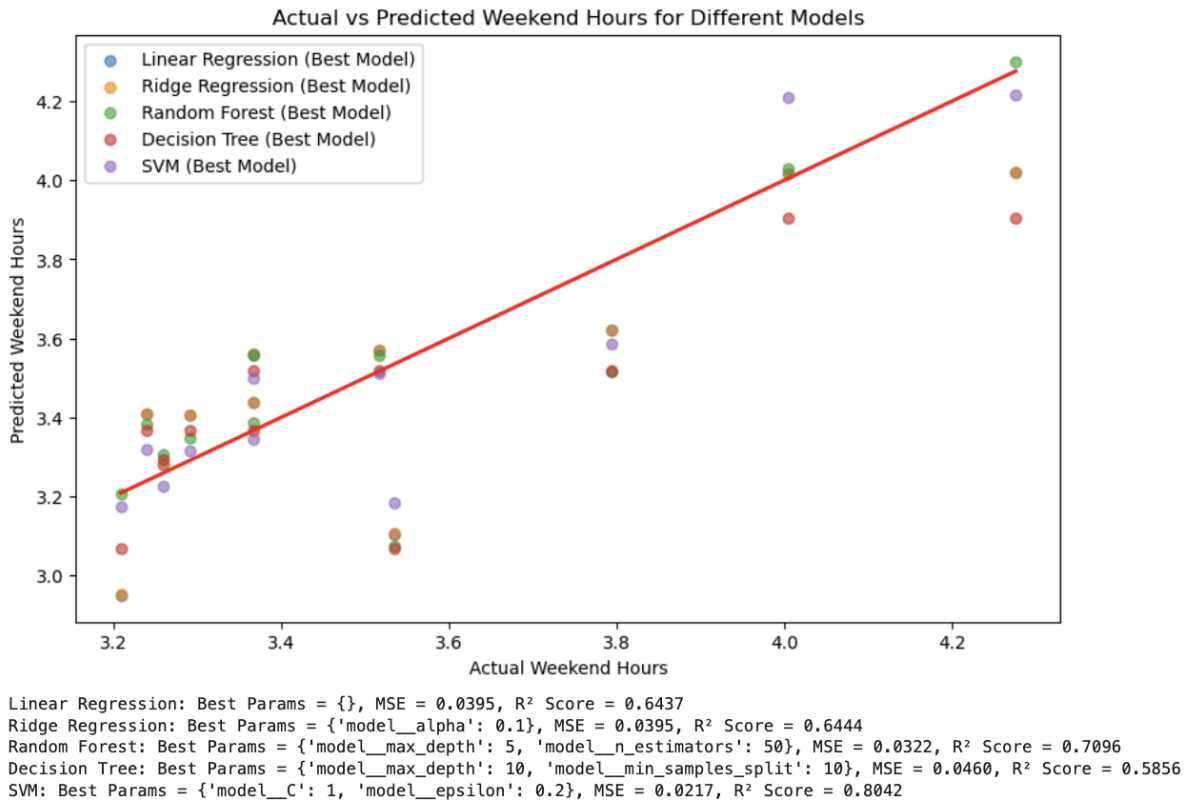
- Certain states, like Puerto Rico, New York, Mississippi, experience significant reductions in staffing levels on weekends, highlighting a clear need for additional weekend staff support. The machine learning models, especially Linear Regression and Ridge Regression (R^2 scores around 0.86), indicate that weekend staffing in these states is much lower compared to weekdays.
- These gaps highlight the need for additional weekend staffing support in certain states to maintain consistent care.

Recommendation:

- **Focus on increasing weekend staffing in states like Puerto Rico, New York, and Mississippi:** These states have the biggest drops in staffing levels over the weekend. By offering more weekend staff in these areas, the healthcare staffing provider can help ensure care stays consistent and reliable, even on the weekends when staffing is lower. This will help prevent any disruptions in service and improve the overall care quality for patients.
- Ensuring adequate staffing on weekends will not only enhance care but also lead to higher ratings for these facilities. This will highlight the healthcare staffing provider as a key partner in maintaining high standards of patient care, especially during critical weekend shifts.

Data Visualization:





3. RN Staffing Hours by Facility Rating

Purpose: To compare nurse staffing hours across facilities with different overall ratings, analyzing how staffing levels impact ratings and overall care quality.

```
SELECT prov.overall_rating,
       AVG(pbj.hrs_rn / NULLIF(pbj.mdscensus, 0)) AS avg_rn_hours_per_resident,
       AVG(pbj.hrs_lpn / NULLIF(pbj.mdscensus, 0)) AS avg_lpn_hours_per_resident,
       AVG(pbj.hrs_cna / NULLIF(pbj.mdscensus, 0)) AS avg_cna_hours_per_resident
FROM pbj_transformed pbj
JOIN provider_transformed prov
ON pbj.provnum = prov.cms_certification_number_ccn_
WHERE prov.overall_rating IS NOT NULL
GROUP BY prov.overall_rating
ORDER BY prov.overall_rating DESC;
```

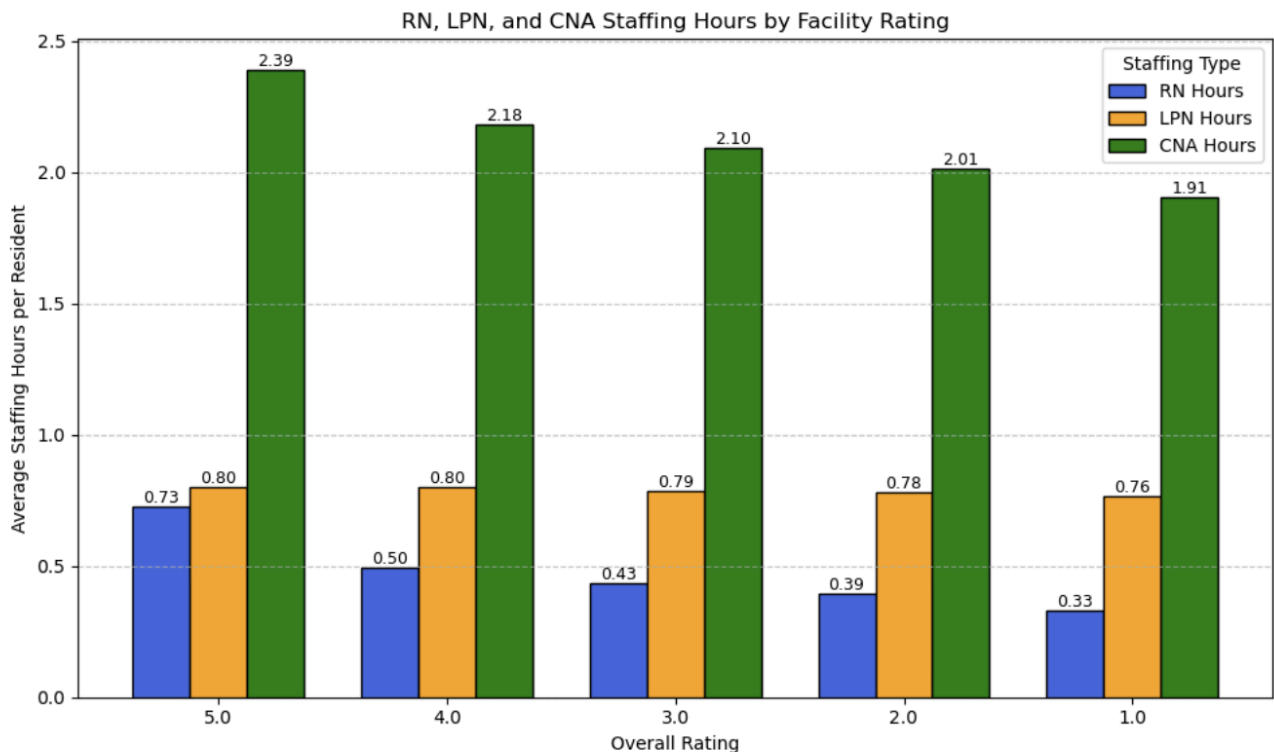
Findings:

- Higher-rated facilities (5 stars) have significantly higher average RN hours per resident (0.73 hours), compared to lower-rated facilities (1-2 stars), which have the lowest RN staffing levels (0.33-0.39 hours).
- LPN staffing remains relatively consistent across all ratings, ranging from 0.76 to 0.80 hours per resident, indicating that LPN staffing is not as directly tied to facility ratings as RN staffing.
- CNA staffing also follows a similar pattern, with higher-rated facilities having more CNA hours (2.39 hours) compared to lower-rated ones (1.91 hours). This suggests that a higher overall staffing level is associated with better ratings.

Recommendation:

- If the healthcare staffing provider partners with lower-rated (1-2 star) facilities, the focus should be on increasing RN and CNA staffing. These facilities often struggle with low staffing levels, especially for RNs and CNAs, which affects care quality. By focusing on increasing staffing levels in these areas, the healthcare staffing provider can position itself as a strategic partner in helping these facilities improve their ratings and patient care. Since higher RN and CNA hours are linked to better ratings, prioritizing these roles in lower-rated facilities could improve care consistency and lead to higher ratings over time.

Data Visualization:



Conclusion

This analysis highlights critical staffing challenges and opportunities across different facility types, states, and rating levels. By targeting **for-profit facilities, states with weekend staffing gaps, and low-rated facilities**, the sales team of the healthcare staffing provider can drive meaningful improvements in patient care while positioning our services as essential solutions. Tailoring sales messaging based on these insights will enhance engagement and lead to better business outcomes.