Data-Assignment

Charlotte Xu, Ruiyang Dong, Xuyuan Zhang, Zihan Wang

15 December, 2024

Cost Analysis and Forecasting for Hospital Financial Performance

General Format of the report

Charlotte Xu; Ruiyang Dong; Xuyuan Zhang; Zihan Wang

- Charlotte Xu: Shiny Results and Statistical Summary.
- Ruiyang Dong: Introduction and co-lead on Data Cleaning.
- Xuyuan Zhang: Co-lead on Data Cleaning and lead on Linear Models and Spatial Analysis.
- Zihan Wang: Machine Learning implementation and analysis.

The Github Repository is listed in: https://github.com/sergiozxy/BIOSTAT625-Project

Here are the packages we used in this project

Introduction

The financial health of hospitals is a critical concern in the U.S. healthcare system, influencing both patient care quality and operational sustainability of institutions. This report aims to analyze historical operating costs and revenue trends of hospitals using the CMS Hospital Provider Cost Report dataset. The project employs statistical models and machine learning tools to guide hospital administrators in making informed decisions, optimizing budgets, and improving resource allocation efficiency.

The study focuses on two key dependent variables: Cost-to-Revenue Ratio, which measures operating efficiency by comparing operating costs to total revenue, and Revenue per Bed, which evaluates revenue generation relative to hospital capacity. These metrics provide a comprehensive view of hospital financial performance and allow for deeper insights into factors that influence profitability.

The dataset spans from 2011-2022, includes a wide range of variables, such as total discharges, hospital total days, total salaries, inpatient and outpatient charges, total income, liabilities, current and fixed assets, and inventory. This comprehensive set of variables enables robust analysis and modeling to capture the diverse factors impacting hospital financial health. By incorporating both operational and financial measures, the study provides actionable insights into optimizing resource allocation and improving decision-making processes.

This research comes at a critical time, as healthcare financial challenges continue to rise. In 2023, healthcare bankruptcies reached their highest level in five years, with 79 Chapter 11 bankruptcy filings recorded. This marked a significant increase from 51 cases in 2019, primarily driven by large liabilities and pandemic-related economic shifts (Payerchin, 2024). These challenges underline the importance of understanding and addressing hospital profitability determinants.

Revenue cycle management (RCM) also plays a pivotal role in improving hospital profitability by streamlining claim submission, billing, and reimbursement processes. Effective RCM improves cash flow, enhances patient experience, and reduces billing errors (Chandawarkar et al., 2024). Previous research has identified factors affecting hospital profitability, including organizational structure, reimbursement mechanisms, and patient mix (Nevola et al., 2016; Ly & Cutler, 2018). This study focuses on these established variables to assess their influence on financial performance.

By leveraging these dependent variables and exploring various financial and operational covariates, this report aims to provide actionable insights for healthcare leaders and policymakers striving to maintain profitability while delivering high-quality care.

Data Cleaning

A thorough data cleaning process was conducted to prepare the CMS Hospital Provider Cost Report dataset for analysis. This step ensured the data's reliability for understanding hospital financial performance and supporting decision-making.

Duplicates were identified based on the Provider CMS Certification Number (CCN) and year. For numeric variables, the mean was retained, while for non-numeric variables, the first occurrence was preserved. Numeric columns fully missing within duplicate groups were removed.

To address outliers, data points with a Cost-to-Revenue Ratio greater than 100 and Revenue per Bed exceeding $100 million (scaled by dividing by 1,000,000) were excluded. These entries were deemed extreme and could distort the analysis.

Missing numeric values were interpolated using the zoo::na.approx() function when sufficient data points (more than two) were available. Sparse variables, where interpolation was unreliable, were left untouched to avoid introducing artificial patterns.

Several financial and operational metrics, such as Total Salaries, Charges, Income, and Liabilities, were scaled to millions for consistency. Two key metrics—Cost-to-Revenue Ratio and Revenue per Bed—were calculated to measure operating efficiency and revenue generation.

This cleaning process ensured the dataset was optimized for robust analysis. By removing inconsistencies, addressing outliers, and filling gaps, it became suitable for deriving actionable insights into hospital profitability, addressing financial challenges, and improving decision-making as outlined in the introduction.

Statistical Summary

R Shiny Application

The CMS Hospital Provider Cost Report dataset, with its wealth of information, provides a unique opportunity to explore financial performance metrics at both granular (hospital-level) and aggregated (state-level) scales. Effectively communicating these insights requires an intuitive and customizable tool, which R Shiny provides through its interactive and dynamic visualization capabilities.

The application includes two key visualization tabs. The first tab, the dot distribution map, presents hospital-level financial metrics for a sample of 4,000 hospitals. Each hospital is represented by a dot, where the color gradient indicates revenue per bed, with darker colors reflecting higher values, and the dot size scales with the cost-to-revenue ratio. Users can interact with the map through pop-ups that display details such as the hospital name, location, revenue per bed, and cost-to-revenue ratio. A temporal slider allows users to explore trends across the dataset's time range from 2011 to 2022. This feature is particularly useful for identifying temporal patterns and evaluating changes in financial performance over time.

The second tab aggregates financial metrics at the state level, providing a broader view of regional financial trends. The state-level map uses color gradients to highlight variations in metrics such as average revenue per bed and average cost-to-revenue ratios. Interactive pop-ups further enrich the experience by displaying detailed summaries for each state, including aggregated metrics and comparisons across time periods. This functionality allows users to quickly identify states that are outperforming or underperforming financially, offering actionable insights for resource allocation and policy planning.

This application not only empowers stakeholders to make evidence-based decisions but also showcases the versatility of R Shiny in handling large and complex datasets. By combining statistical analysis with user-friendly visualization tools, the application serves as a valuable resource for understanding and improving hospital financial performance at multiple levels. Future enhancements could include the integration of additional metrics, predictive model outputs, and scenario analysis features to further enrich the insights provided to users.

Exploratory Data Description

Table 1 provides a summary of the key financial and operational metrics used in this study, offering insights into the dataset's composition and variability across hospitals. The dataset, which spans over 50,000 observations from 2011 to 2022, captures critical financial indicators such as total revenue, operating costs, and bad debt expenses, alongside operational factors like the number of beds and facility location.

Regression and Geospatial Analysis

Linear Regression Model

We start with a basic linear regression model and the model can be listed as the follows:

$$Y_{it} = \beta_0 + \beta \mathbf{X_{it}} + \mu_{\mathbf{i}} + \tau_{\mathbf{t}} + \epsilon_{\mathbf{it}}$$

where Y_{it}: Target variable (e.g., total revenue or operating costs). β_i: Coefficients for predictors. $\mathbf{X_{it}}$ are the key independent variables that are reported in the statistics summary table, μ_i is the state level fixed effect and the τ_t is the time fixed effect (year). ϵ_{it}: Error term.

```
source("./sources/linear_model.R")
```

```
##
## Call:
## lm(formula = baseline_formula2, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -54.882  -1.403  -0.405   0.777  84.217
##
## Coefficients:
##                                                                  Estimate
## (Intercept)                                                    -3.181e+02
## Total.Discharges..V...XVIII...XIX...Unknown.                    1.334e-04
## Hospital.Total.Days..V...XVIII...XIX...Unknown..For.Adults...Peds -6.074e-05
## Total.Salaries.From.Worksheet.A                                -2.256e-03
## Inpatient.Total.Charges                                        -2.945e-03
## Outpatient.Total.Charges                                        1.241e-04
## Total.Income                                                    7.507e-04
## Total.Other.Income                                             -1.581e-03
## Total.Liabilities.and.Fund.Balances                            -2.800e-04
## Accounts.Payable                                                3.117e-04
## Total.Current.Assets                                            1.461e-04
## Total.Fixed.Assets                                              2.566e-04
## General.Fund.Balance                                           -5.266e-05
## Inventory                                                       6.084e-10
## Total.Patient.Revenue                                           3.587e-03
## Number.of.Beds                                                 -3.798e-06
## year                                                            1.597e-01
## State.CodeAL                                                   -1.490e+00
## State.CodeAR                                                   -1.838e+00
## State.CodeAS                                                   -2.963e+00
## State.CodeAZ                                                   -2.462e-01
## State.CodeCA                                                    7.559e-01
## State.CodeCO                                                   -1.179e-01
## State.CodeCT                                                   -1.771e-01
## State.CodeDC                                                   -1.823e-01
## State.CodeDE                                                   -7.705e-01
## State.CodeFL                                                   -2.467e-01
## State.CodeGA                                                   -9.387e-01
## State.CodeGU                                                   -1.509e+00
## State.CodeHI                                                   -7.634e-01
## State.CodeIA                                                   -1.801e+00
## State.CodeID                                                   -1.234e+00
## State.CodeIL                                                   -8.386e-01
## State.CodeIN                                                   -3.518e-01
## State.CodeKS                                                   -2.128e+00
## State.CodeKY                                                   -1.388e+00
## State.CodeLA                                                   -1.414e+00
## State.CodeMA                                                   -4.303e-01
## State.CodeMD                                                   -1.315e+00
## State.CodeME                                                   -7.148e-01
## State.CodeMI                                                   -8.573e-01
## State.CodeMN                                                   -9.402e-01
## State.CodeMO                                                   -8.621e-01
## State.CodeMS                                                   -1.971e+00
## State.CodeMT                                                   -2.166e+00
```

```
## State.CodeNC                                                    -7.982e-01
## State.CodeND                                                    -2.499e+00
## State.CodeNE                                                    -1.910e+00
## State.CodeNH                                                     1.264e-01
## State.CodeNJ                                                     7.269e-01
## State.CodeNM                                                    -7.709e-01
## State.CodeNV                                                    -2.204e-03
## State.CodeNY                                                    -3.462e-01
## State.CodeOH                                                    -6.826e-01
## State.CodeOK                                                    -1.242e+00
## State.CodeOR                                                     2.667e-02
## State.CodePA                                                    -4.624e-01
## State.CodePR                                                    -2.364e+00
## State.CodeRI                                                    -1.001e+00
## State.CodeSC                                                    -9.094e-01
## State.CodeSD                                                    -1.802e+00
## State.CodeTN                                                    -1.140e+00
## State.CodeTX                                                    -5.858e-01
## State.CodeUT                                                    -1.617e+00
## State.CodeVA                                                    -5.557e-01
## State.CodeVI                                                    -2.286e+00
## State.CodeVT                                                     5.919e-02
## State.CodeWA                                                     7.131e-01
## State.CodeWI                                                    -3.799e-02
## State.CodeWV                                                    -1.633e+00
## State.CodeWY                                                    -1.618e+00
##                                                                Std. Error
## (Intercept)                                                      7.336e+00
## Total.Discharges..V...XVIII...XIX...Unknown.                     4.683e-06
## Hospital.Total.Days..V...XVIII...XIX...Unknown..For.Adults...Peds 1.216e-06
## Total.Salaries.From.Worksheet.A                                  2.796e-04
## Inpatient.Total.Charges                                          1.126e-04
## Outpatient.Total.Charges                                         1.174e-04
## Total.Income                                                     1.861e-04
## Total.Other.Income                                               2.284e-04
## Total.Liabilities.and.Fund.Balances                              3.529e-05
## Accounts.Payable                                                 2.333e-04
## Total.Current.Assets                                             5.898e-05
## Total.Fixed.Assets                                               1.235e-04
## General.Fund.Balance                                             5.303e-05
## Inventory                                                        1.257e-09
## Total.Patient.Revenue                                            1.082e-04
## Number.of.Beds                                                   1.903e-06
## year                                                             3.637e-03
## State.CodeAL                                                     2.392e-01
## State.CodeAR                                                     2.415e-01
## State.CodeAS                                                     2.167e+00
## State.CodeAZ                                                     2.406e-01
## State.CodeCA                                                     2.272e-01
## State.CodeCO                                                     2.391e-01
## State.CodeCT                                                     2.717e-01
## State.CodeDC                                                     3.407e-01
## State.CodeDE                                                     3.489e-01
## State.CodeFL                                                     2.304e-01
## State.CodeGA                                                     2.331e-01
## State.CodeGU                                                     7.342e-01
## State.CodeHI                                                     2.856e-01
## State.CodeIA                                                     2.364e-01
## State.CodeID                                                     2.569e-01
## State.CodeIL                                                     2.314e-01
```

```
## State.CodeIN                                                    2.342e-01
## State.CodeKS                                                    2.340e-01
## State.CodeKY                                                    2.377e-01
## State.CodeLA                                                    2.329e-01
## State.CodeMA                                                    2.451e-01
## State.CodeMD                                                    2.557e-01
## State.CodeME                                                    2.655e-01
## State.CodeMI                                                    2.342e-01
## State.CodeMN                                                    2.351e-01
## State.CodeMO                                                    2.361e-01
## State.CodeMS                                                    2.399e-01
## State.CodeMT                                                    2.491e-01
## State.CodeNC                                                    2.371e-01
## State.CodeND                                                    2.576e-01
## State.CodeNE                                                    2.400e-01
## State.CodeNH                                                    2.767e-01
## State.CodeNJ                                                    2.437e-01
## State.CodeNM                                                    2.584e-01
## State.CodeNV                                                    2.557e-01
## State.CodeNY                                                    2.334e-01
## State.CodeOH                                                    2.314e-01
## State.CodeOK                                                    2.349e-01
## State.CodeOR                                                    2.507e-01
## State.CodePA                                                    2.312e-01
## State.CodePR                                                    2.547e-01
## State.CodeRI                                                    3.326e-01
## State.CodeSC                                                    2.448e-01
## State.CodeSD                                                    2.516e-01
## State.CodeTN                                                    2.364e-01
## State.CodeTX                                                    2.253e-01
## State.CodeUT                                                    2.524e-01
## State.CodeVA                                                    2.387e-01
## State.CodeVI                                                    6.609e-01
## State.CodeVT                                                    3.177e-01
## State.CodeWA                                                    2.394e-01
## State.CodeWI                                                    2.347e-01
## State.CodeWV                                                    2.520e-01
## State.CodeWY                                                    2.783e-01
##                                                                 t value
## (Intercept)                                                    -43.362
## Total.Discharges..V...XVIII...XIX...Unknown.                    28.480
## Hospital.Total.Days..V...XVIII...XIX...Unknown..For.Adults...Peds -49.945
## Total.Salaries.From.Worksheet.A                                 -8.068
## Inpatient.Total.Charges                                        -26.150
## Outpatient.Total.Charges                                         1.057
## Total.Income                                                     4.034
## Total.Other.Income                                              -6.922
## Total.Liabilities.and.Fund.Balances                             -7.934
## Accounts.Payable                                                 1.336
## Total.Current.Assets                                             2.476
## Total.Fixed.Assets                                               2.077
## General.Fund.Balance                                            -0.993
## Inventory                                                        0.484
## Total.Patient.Revenue                                           33.152
## Number.of.Beds                                                  -1.997
## year                                                            43.918
## State.CodeAL                                                    -6.231
## State.CodeAR                                                    -7.610
## State.CodeAS                                                    -1.367
## State.CodeAZ                                                    -1.023
```

```
## State.CodeCA                                                    3.327
## State.CodeCO                                                   -0.493
## State.CodeCT                                                   -0.652
## State.CodeDC                                                   -0.535
## State.CodeDE                                                   -2.208
## State.CodeFL                                                   -1.070
## State.CodeGA                                                   -4.028
## State.CodeGU                                                   -2.056
## State.CodeHI                                                   -2.673
## State.CodeIA                                                   -7.618
## State.CodeID                                                   -4.803
## State.CodeIL                                                   -3.624
## State.CodeIN                                                   -1.502
## State.CodeKS                                                   -9.096
## State.CodeKY                                                   -5.838
## State.CodeLA                                                   -6.074
## State.CodeMA                                                   -1.756
## State.CodeMD                                                   -5.144
## State.CodeME                                                   -2.692
## State.CodeMI                                                   -3.661
## State.CodeMN                                                   -3.999
## State.CodeMO                                                   -3.651
## State.CodeMS                                                   -8.218
## State.CodeMT                                                   -8.693
## State.CodeNC                                                   -3.367
## State.CodeND                                                   -9.701
## State.CodeNE                                                   -7.961
## State.CodeNH                                                    0.457
## State.CodeNJ                                                    2.983
## State.CodeNM                                                   -2.984
## State.CodeNV                                                   -0.009
## State.CodeNY                                                   -1.483
## State.CodeOH                                                   -2.950
## State.CodeOK                                                   -5.289
## State.CodeOR                                                    0.106
## State.CodePA                                                   -2.000
## State.CodePR                                                   -9.281
## State.CodeRI                                                   -3.008
## State.CodeSC                                                   -3.715
## State.CodeSD                                                   -7.160
## State.CodeTN                                                   -4.820
## State.CodeTX                                                   -2.600
## State.CodeUT                                                   -6.405
## State.CodeVA                                                   -2.328
## State.CodeVI                                                   -3.459
## State.CodeVT                                                    0.186
## State.CodeWA                                                    2.978
## State.CodeWI                                                   -0.162
## State.CodeWV                                                   -6.479
## State.CodeWY                                                   -5.812
##                                                              Pr(>|t|)
## (Intercept)                                                   < 2e-16 ***
## Total.Discharges..V...XVIII...XIX...Unknown.                  < 2e-16 ***
## Hospital.Total.Days..V...XVIII...XIX...Unknown..For.Adults...Peds  < 2e-16 ***
## Total.Salaries.From.Worksheet.A                              7.27e-16 ***
## Inpatient.Total.Charges                                       < 2e-16 ***
## Outpatient.Total.Charges                                     0.290343
## Total.Income                                                 5.49e-05 ***
## Total.Other.Income                                           4.51e-12 ***
## Total.Liabilities.and.Fund.Balances                          2.15e-15 ***
```

```
## Accounts.Payable                                    0.181585
## Total.Current.Assets                                0.013278 *
## Total.Fixed.Assets                                  0.037797 *
## General.Fund.Balance                                0.320704
## Inventory                                           0.628349
## Total.Patient.Revenue                                < 2e-16 ***
## Number.of.Beds                                      0.045883 *
## year                                                 < 2e-16 ***
## State.CodeAL                                        4.65e-10 ***
## State.CodeAR                                        2.78e-14 ***
## State.CodeAS                                        0.171635
## State.CodeAZ                                        0.306123
## State.CodeCA                                        0.000877 ***
## State.CodeCO                                        0.621954
## State.CodeCT                                        0.514553
## State.CodeDC                                        0.592552
## State.CodeDE                                        0.027215 *
## State.CodeFL                                        0.284455
## State.CodeGA                                        5.64e-05 ***
## State.CodeGU                                        0.039826 *
## State.CodeHI                                        0.007512 **
## State.CodeIA                                        2.61e-14 ***
## State.CodeID                                        1.57e-06 ***
## State.CodeIL                                        0.000290 ***
## State.CodeIN                                        0.132982
## State.CodeKS                                         < 2e-16 ***
## State.CodeKY                                        5.32e-09 ***
## State.CodeLA                                        1.26e-09 ***
## State.CodeMA                                        0.079158 .
## State.CodeMD                                        2.70e-07 ***
## State.CodeME                                        0.007095 **
## State.CodeMI                                        0.000251 ***
## State.CodeMN                                        6.37e-05 ***
## State.CodeMO                                        0.000261 ***
## State.CodeMS                                         < 2e-16 ***
## State.CodeMT                                         < 2e-16 ***
## State.CodeNC                                        0.000760 ***
## State.CodeND                                         < 2e-16 ***
## State.CodeNE                                        1.74e-15 ***
## State.CodeNH                                        0.647699
## State.CodeNJ                                        0.002856 **
## State.CodeNM                                        0.002849 **
## State.CodeNV                                        0.993121
## State.CodeNY                                        0.138061
## State.CodeOH                                        0.003179 **
## State.CodeOK                                        1.23e-07 ***
## State.CodeOR                                        0.915294
## State.CodePA                                        0.045464 *
## State.CodePR                                         < 2e-16 ***
## State.CodeRI                                        0.002631 **
## State.CodeSC                                        0.000203 ***
## State.CodeSD                                        8.15e-13 ***
## State.CodeTN                                        1.44e-06 ***
## State.CodeTX                                        0.009326 **
## State.CodeUT                                        1.52e-10 ***
## State.CodeVA                                        0.019893 *
## State.CodeVI                                        0.000542 ***
## State.CodeVT                                        0.852207
## State.CodeWA                                        0.002902 **
## State.CodeWI                                        0.871398
```

```
## State.CodeWV                                                  9.28e-11 ***
## State.CodeWY                                                  6.19e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.049 on 63277 degrees of freedom
## Multiple R-squared:  0.3495, Adjusted R-squared:  0.3488
## F-statistic: 485.7 on 70 and 63277 DF,  p-value: < 2.2e-16
```

Machine Learning Implementation and Analysis

References

- Chandawarkar, R., Nadkarni, P., Barmash, E., Thomas, S., Capek, A., Casey, K., & Carradero, F. (2024). Revenue Cycle Management: The Art and the Science. Journal of Health Care Finance. PMCID: PMC11219169.
- Nevola, A., Pace, C., Karim, S. A., & Morris, M. E. (2016). Revisiting 'The Determinants of Hospital Profitability' in Florida. Journal of Health Care Finance.
- Ly, D. P., & Cutler, D. M. (2018). Factors of U.S. Hospitals Associated with Improved Profit Margins: An Observational Study. Journal of General Internal Medicine, 33(7), 1020–1027. https://doi.org/10.1007/s11606-018-4347-4
- Payerchin, R. (2024). Health care bankruptcies in 2023 reach highest level in five years. Medical Economics. Retrieved from https://www.medicaleconomics.com/view/health-care-bankruptcies-in-2023-reach-highest-level-in-five-years