Quality and Cost:

Home Health Care Agency Data from the Centers for Medicare and Medicaid

Lia Gaetano

The dataset

United States Centers for Medicare & Medicaid Services (CMS) is the federal government entity that oversees the Medicare and Medicaid insurance policies provided by the United States federal government.
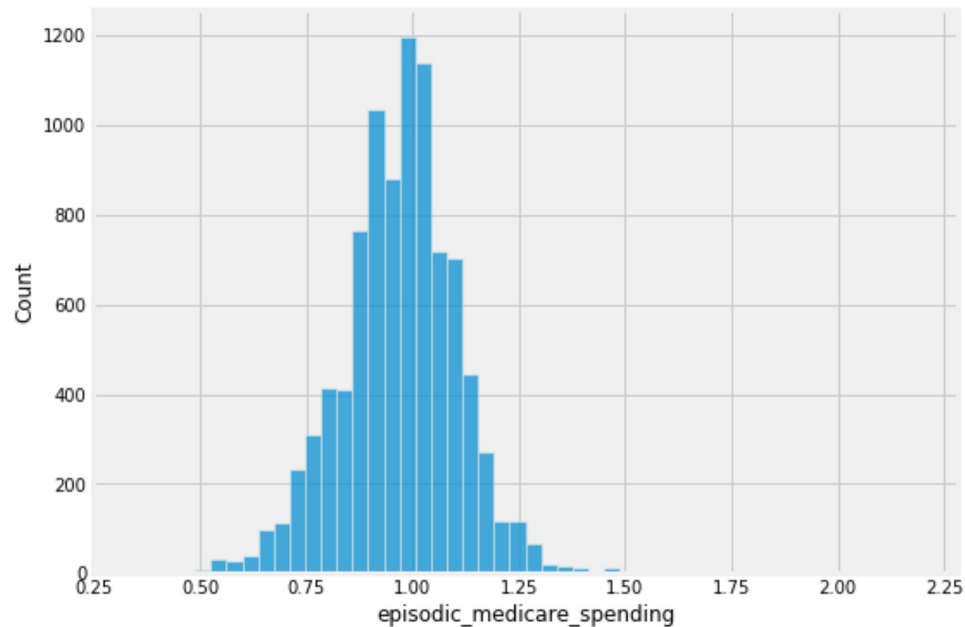
CMS tracks a myriad of metrics regarding access, quality assurance and cost of healthcare in the United States.

The CMS "Home Health Care Agencies" dataset compiles state averages of Home Health Agency demographic and quality measurements, including metrics such as area-wide frequency of hospital admission for home health patients and frequency of improvement in patient ambulation.
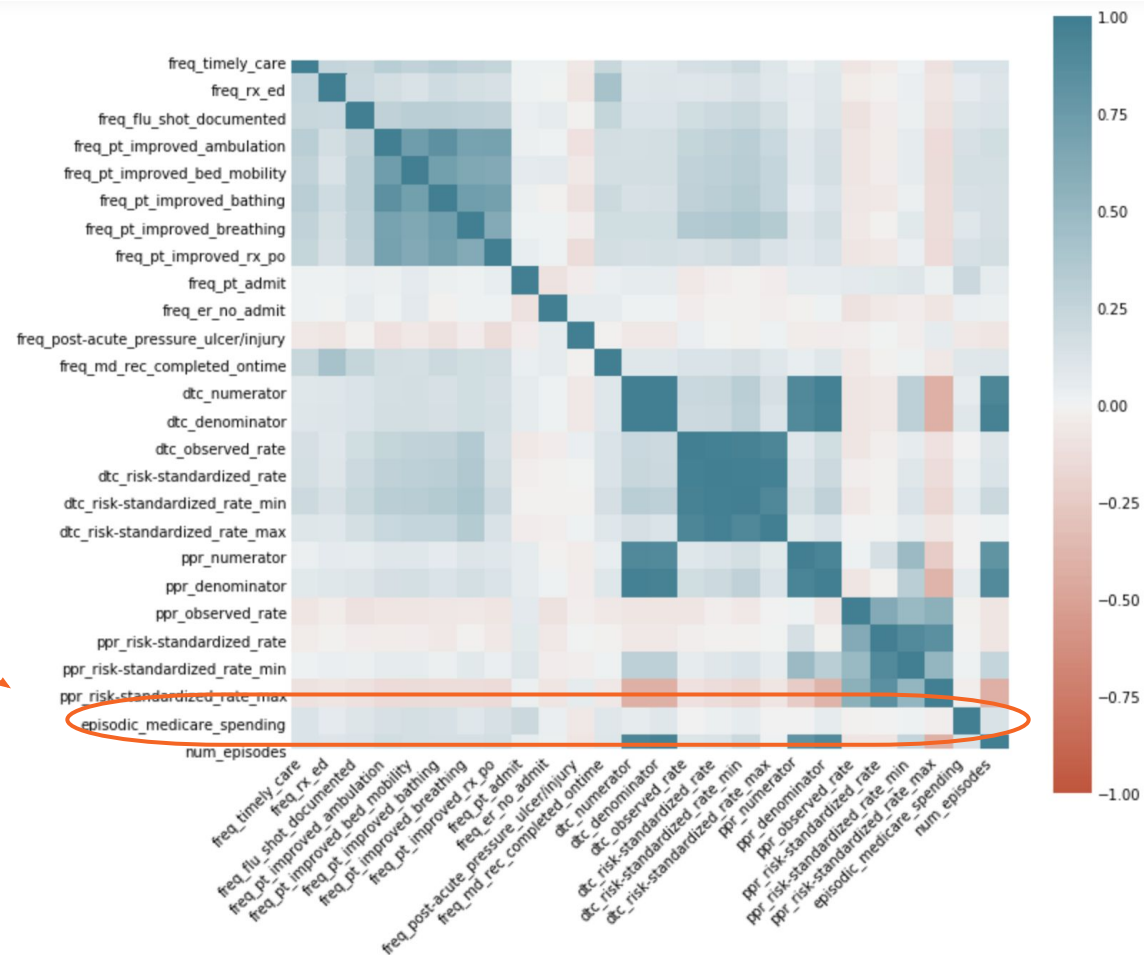
EDA

1. **State** (Text) – The two character postal code for the state or territory in which the home health agency is located.
2. **CMS Certification Number (CCN)** (Text) – The six character identification number assigned to the home health agency by CMS.
3. **Provider Name** (Text) – The name of the home health agency as it appears in the CMS certification system.
4. **Address** (Text) – The street address of the home health agency.
5. **City** (Text) – The city in which the home health agency is located.
6. **ZIP** (Text) – The five-digit ZIP code of the home health agency.
7. **Phone** (Text) – The ten-digit telephone number of the home health agency.
8. **Type of Ownership** (Text) – The general control type of the home health agency. Categories include: • Voluntary Non-Profit – Religious Affiliation • Voluntary Non-Profit – Private • Voluntary Non-Profit – Other • Proprietary • Government – State/County • Government – Combination Government and Voluntary • Government – Local
9. **Offers Nursing Care Services** (Text) – Yes or No
10. **Offers Physical Therapy Services** (Text) – Yes or No
11. **Offers Occupational Therapy Services** (Text) – Yes or No
12. **Offers Speech Pathology Services** (Text) – Yes or No
13. **Offers Medical Social Services** (Text) – Yes or No
14. **Offers Home Health Aide Services** (Text) – Yes or No
15. **Date Certified** (Date) – The original date the home health agency was certified to participate in the Medicare program.
16. **Quality of patient care star rating** (Numeric) – A numeric rating from 1 through 5, in increments of 0.5.
17. **Footnote for quality of patient care star rating** (Text) –If the rating value is missing (blank), the reason the rating was not calculated is explained here.
18. **How often the home health team began their patients' care in a timely manner** (Numeric) – The measure percentage as reported on Care Compare.
19. **Footnote for how often the home health team began their patients' care in a timely manner** (Text) – If the measure value is missing (blank), the reason the measure was not calculated is explained here.
20. **How often the home health team taught patients (or their family caregivers) about their drugs** (Numeric) – The measure percentage as reported on Care Compare.
21. **Footnote for how often the home health team taught patients (or their family caregivers) about their drugs** (Text) –If the measure value is missing (blank), the reason the measure was not calculated is explained here.
22. **How often the home health team determined whether patients received a flu shot for the current flu season** (Numeric) – The measure percentage as reported on Care Compare.
23. **Footnote for how often the home health team determined whether patients received a flu shot for the current flu season (Text) – If the measure value is missing** (blank), the reason the measure was not calculated is explained here.
24. **How often patients got better at walking or moving around** (Numeric) – The measure percentage as reported on Care Compare.
25. **Footnote for how often patients got better at walking or moving around** (Text) – If the measure value is missing (blank), the reason the measure was not calculated is explained here.

Target: Episodic Medicare Spending

How much Medicare spends on an episode of care at this agency, compared to Medicare spending across all agencies nationally

Not much correlation with the target variable.

Linear Regression

- Baseline RMSE: 0.139904

- Ridge Regression (alphas: 0.0-0.9) using normalized features
  - Best performing result -- Alpha=0.0, no normalization, basic linear regression
  - RMSE: 0.122297, R2:0.219594
- Cross Validation using standardized features
  - RMSE: 0.125640, R2: 0.166286
- Cross Validation using selected features
  - RMSE: 0.125640, R2: 0.166286

KNN

- KNN using standardized features and n=3
    - RMSE: 16.762279
    - R2: -0.477399