CIS 671 Final Project Proposal

Title:

Cost-Quality Map for Medicare with MS-DRG Drill-Down

Member: Yu-Ting Liu

Motivation & Problem

Payers, consulting teams, and health-system leaders need to find markets and facilities delivering higher quality at lower standardized spend to guide network strategy, steerage, and utilization management. CMS publishes public data on spending, utilization, quality, and inpatient service-line activity, but these signals are scattered across heterogeneous tables and time periods. We will build an interactive **Geographic Variation and Cost–Quality Explorer** with a **click-through Medicare Severity Diagnosis-Related Group (MS-DRG)**. The explorer provides an at-a-glance map overview (USA) and drill-down to Michigan (and counties), reveals where value is high/low, how trends move over time, which hospitals are outliers, and explains why through each hospital's MS-DRG mix and payment profile.

Datasets

Combining multiple CMS datasets (plus cartographic boundaries)

- **Medicare Geographic Variation — State & County:** standardized per-beneficiary spend, ED visits, preventable hospitalizations; used for map and trend views. (Medicare Geographic Variation - by National, State & County | CMS Data)
- **Care Compare / Provider Data Catalog — Hospital General Information:** hospital roster + **Overall Star Rating**. (Hospital General Information | Provider Data Catalog)
- **Medicare Inpatient Hospitals — by Provider & Service (MS-DRG) for drill-down:** hospital-level inpatient service-line mix and payments; used to explain outliers. (Medicare Inpatient Hospitals by Provider & Service Data | CMS Data)
- **U.S. Census Cartographic Boundary Files** (states, counties) for choropleths. (2023 Cartographic Boundary File (SHP), County and Equivalent for United States, 1:5,000,000 - Catalog)

Using CMS Geographic Variation 2023 (latest complete year), MS-DRG FY2023 (latest available vintage), and Hospital Star Ratings as of the 2025 Care Compare release. Choropleths use U.S. Census Cartographic Boundary Files (2023).

Dataset Types & Attribute Types

- **Types:** tabular CSV (CMS PUFs), geospatial GeoJSON/Shapefile (Census)
- **Attributes:** quantitative (spend, rates, star scores, payments), ordinal (stars), categorical (state, ownership, MS-DRG), geographic (FIPS, polygons)

Questions (Tasks)

1. **Where are high-value markets?** (↑quality, ↓spend) at US/state/county levels.
2. **Which hospitals in a selected market are outliers** (efficient frontier vs laggards)?
3. **Where does Michigan rank nationally in 2023** on market spend and market quality?
4. **Why is a hospital an outlier?** *Explainability* via MS-DRG drill-down (service-line concentration, top DRGs, weighted payments, medical vs surgical mix).

Visualization Design

A) Bivariate Choropleth (Michigan counties)

- **Lightness** = z_spend (z-score of standardized spend per beneficiary by year)
- **Hue** = z_quality (z-score of market quality)
- **Interaction:** Click a county to filter the scatter

B) Quadrant legend + live counts (interaction idiom)

- 2×2 legend for High-Value / Invest / Efficiency / Low-Value; shows live counts of MI counties in each quadrant

C) Hospital Scatter vs Market (Grand Rapids area)

- **X** = market z_spend proxy; **Y** = hospital quality (stars)
- **Color:** ownership/system; **optional size:** beds; efficient-frontier line
- **Tooltip:** hospital name, stars, beds, county, plus MS-DRG summary *(see below)*

D) Service-Line Drill-Down

- **Top-5 MS-DRGs by discharge share** (bar chart)
- **Service-mix concentration** (**HHI = Σ share²**) to show specialization vs breadth
- **Weighted average Medicare payment** across MS-DRGs (complexity/resource proxy)
- **Medical vs Surgical mix** (stacked bar)

E) State ranking bar

- Sortable bar chart of states by z_spend (toggle to z_quality), with Michigan highlighted and a percentile badge

F) Global filters toolbar (interaction idiom)

- Ownership/system toggles, hospital search; Year fixed to 2023

Computation Methods

Quality:

- hospital **Overall Star Rating** from Hospital General Info.
- quality_county= $\frac{\sum_{h \in c} stars_{sh} \times bed_{sh}}{\sum_{h \in c} bed_{sh}}$

Z-scores: compute within-year: z_quality = $\frac{quality\_county - \mu_{quality}}{\sigma_{quality}}$

Market assignment: map hospital ZIP → county FIPS; join to Geographic Variation.

Quadrants (value map): High-Value (zQ>0, zS<0); Invest (zQ>0, zS≥0); Efficiency (zQ≤0, zS<0); Low-Value (zQ≤0, zS≥0).

MS-DRG Aggregation

For each hospital (**provider_id / CCN**) :

- **Share per DRG:** `share_drg = discharges_drg / total_discharges_hospital`
- **Service concentration:** `hhi = Σ share_drg²` (0–1; higher → more concentrated)
- **Weighted average payment (Medicare):** `wavg_payment = Σ(avg_medicare_payment_drg × discharges_drg) / total_discharges_hospital`
- **Medical vs Surgical mix:** map DRGs to Medical/Surgical; compute category shares

Integration to visuals:

- Append MS-DRG profile fields (HHI, wavg_payment, med_share, surg_share, top-5 list) to the hospital table used by the scatter; load the detailed top-5 table into the drill-down drawer on click.

Limitations: hospital-level spend is proxied by market; component vintages may differ; some MS-DRG series may not align on years; handle missingness via exclusion or flags.

Task Plan

Pipeline: Python (pandas) for ingestion/cleaning → compute indices → write tidy CSVs → **Observable (D3/Plot) single-page dashboard** with linked interactions.

- Download & normalize datasets; compute **z_spend** (GV 2023) and **z_quality** (bed-weighted Stars → county → z-score); export tidy CSVs.
- Map: Build **Michigan choropleth (A)** and **Quadrant legend with live counts (B)**.
- Scatter: Build **hospital scatter (C)** linked to county selection; efficient frontier + tooltips.
- Drill-down: Implement **drawer (D)** with Top-5 bar, Med/Surg stacked bar, HHI & Weighted Payment KPIs.
- Trends: Build **State ranking bar (E)**; highlight Michigan.
- Filters: Add **global toolbar (F)**; wire to (A)–(E); polish accessibility and performance.

Deliverables

- **Observable single-page dashboard** (unlisted) with ~6 components: (A) Michigan bivariate choropleth, (B) quadrant legend + counts, (C) hospital scatter, (D) MS-DRG drill-down drawer (Top-5 bar, Med/Surg bar, KPIs), (E) State ranking bar**,** (F) global filters toolbar.
- **Presentation video** (walkthrough of interactions and findings)
- **Short report (3-5 pages):** executive summary, methods, findings (include a GR hospital case explaining an outlier via MS-DRGs), limitations, and future work.