

Healthcare and Medicare Data

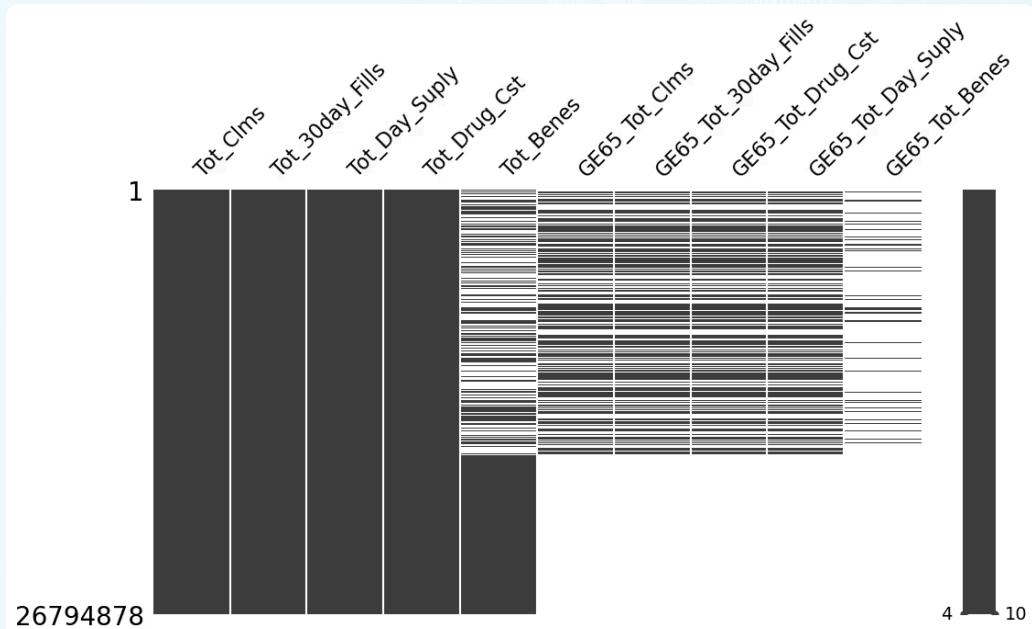
*From Claims to Clarity:
Visualizing Medicare Patterns & Provider Impact*

 New Update / Analysis

1. ANALYSIS OF OPIOID PRESCRIPTIONS
2. BREAKDOWN OF PRESCRIBERS

by Susan Schnitzel

Evaluation and cleaning the data

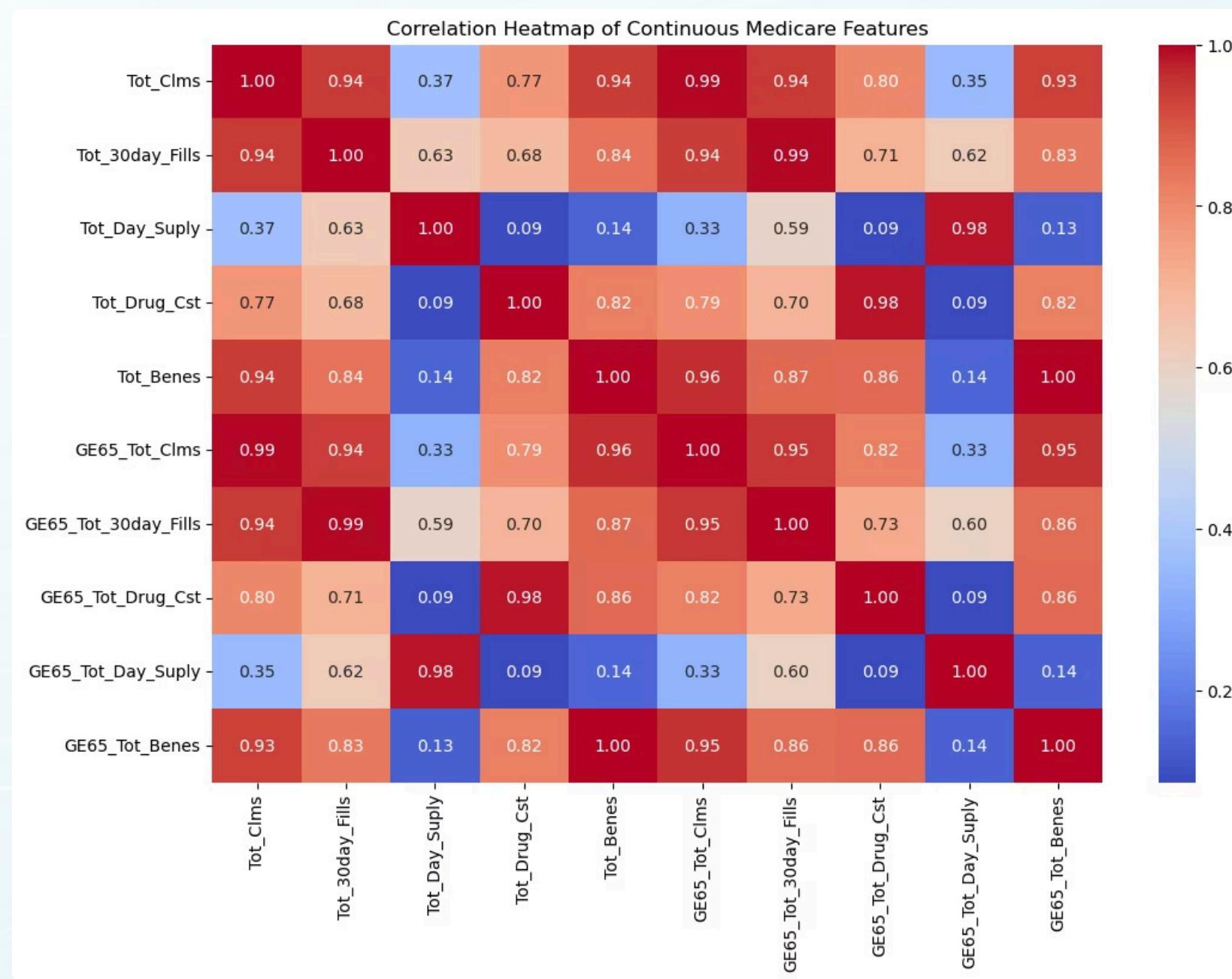


```
!pip install missingno import missingno as msno  
msno.matrix(df[continuous_cols], figsize=(12, 6))
```

I will focus on the First Four Columns but will cleanup the "Total Benefits" to use for some visuals.

Patterns at a Glance:

A Heatmap Exploration of Medicare Data

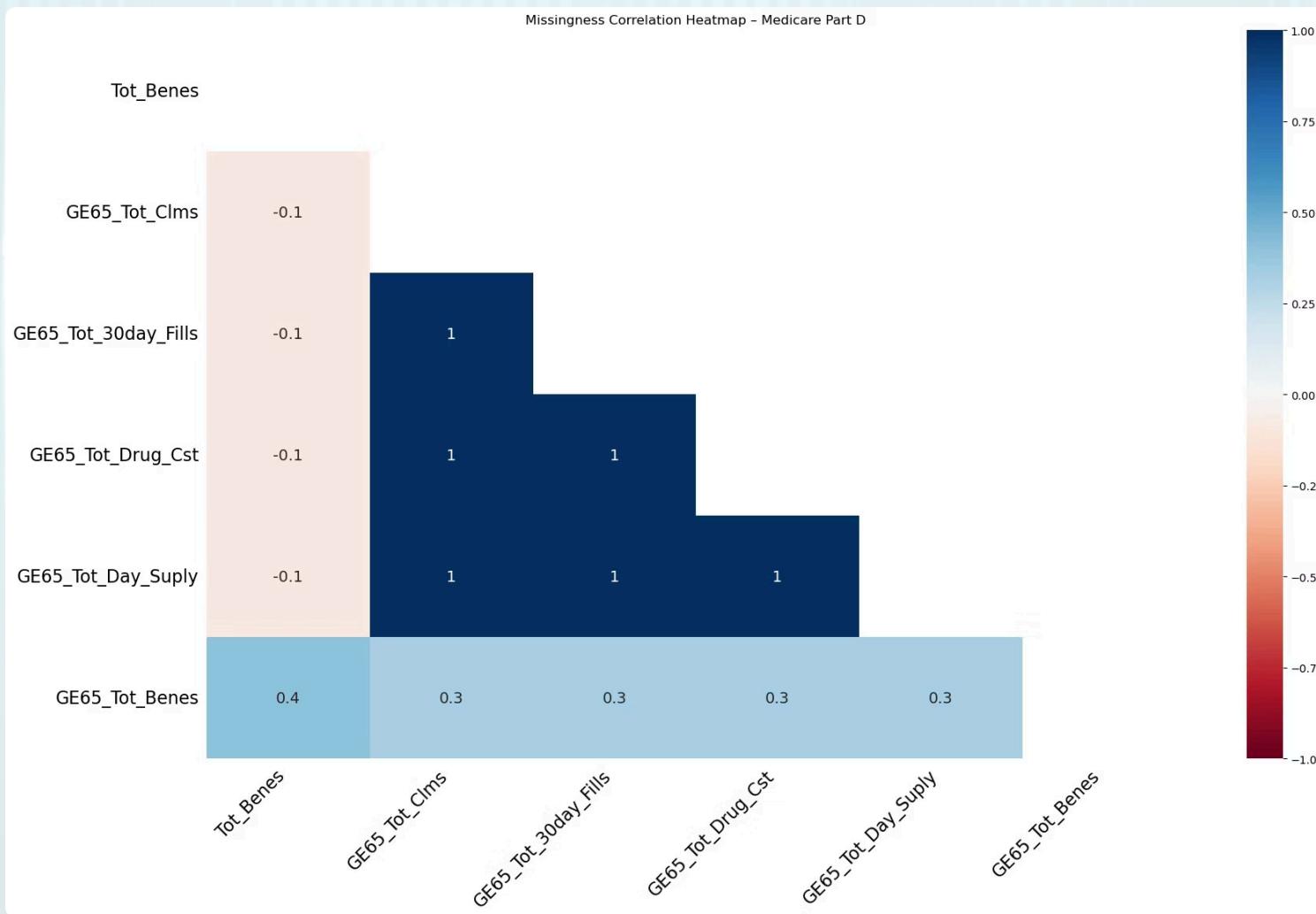


Several trends emerged with some expected, others more revealing.

The GE65-titled columns show a strong correlation with the four key variables I initially planned to focus on. Across those four, the relationships are consistently strong, including with the Benefits column, which demonstrates notable alignment.

Visualizing Data Gaps: Heatmap of Medicare Part D Variables

GE65 variables show perfect alignment, and potentially redundant or policy-driven. GE65 Benefits diverges slightly, suggesting a distinct cost pattern.



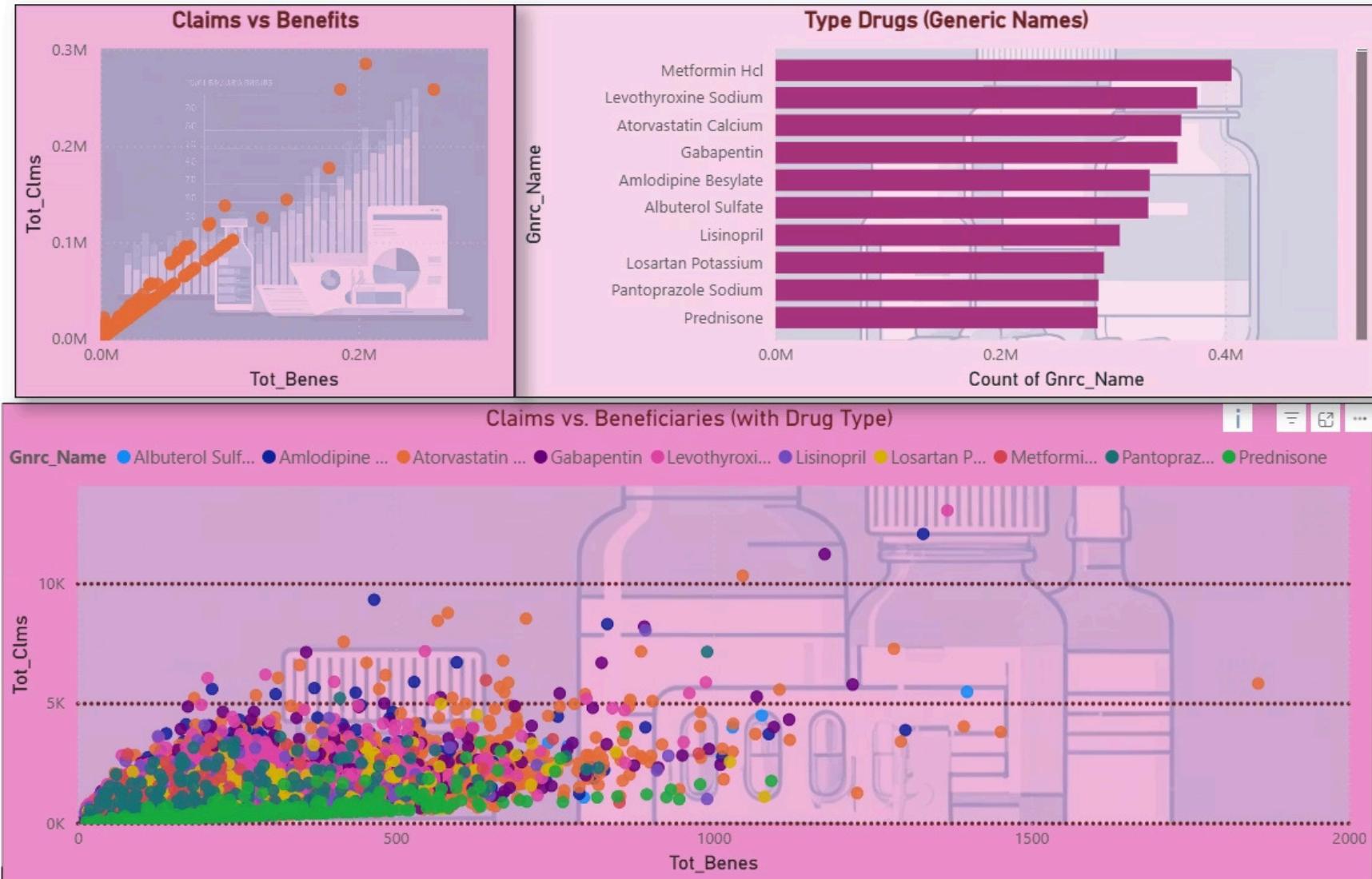


Quick Look of "Claims vs. Benefits" and Top Drugs

Initial findings show a strong, direct correlation between total benefits and total cost.

To explore this further, I filtered out the top 10 drugs; see chart on top right.

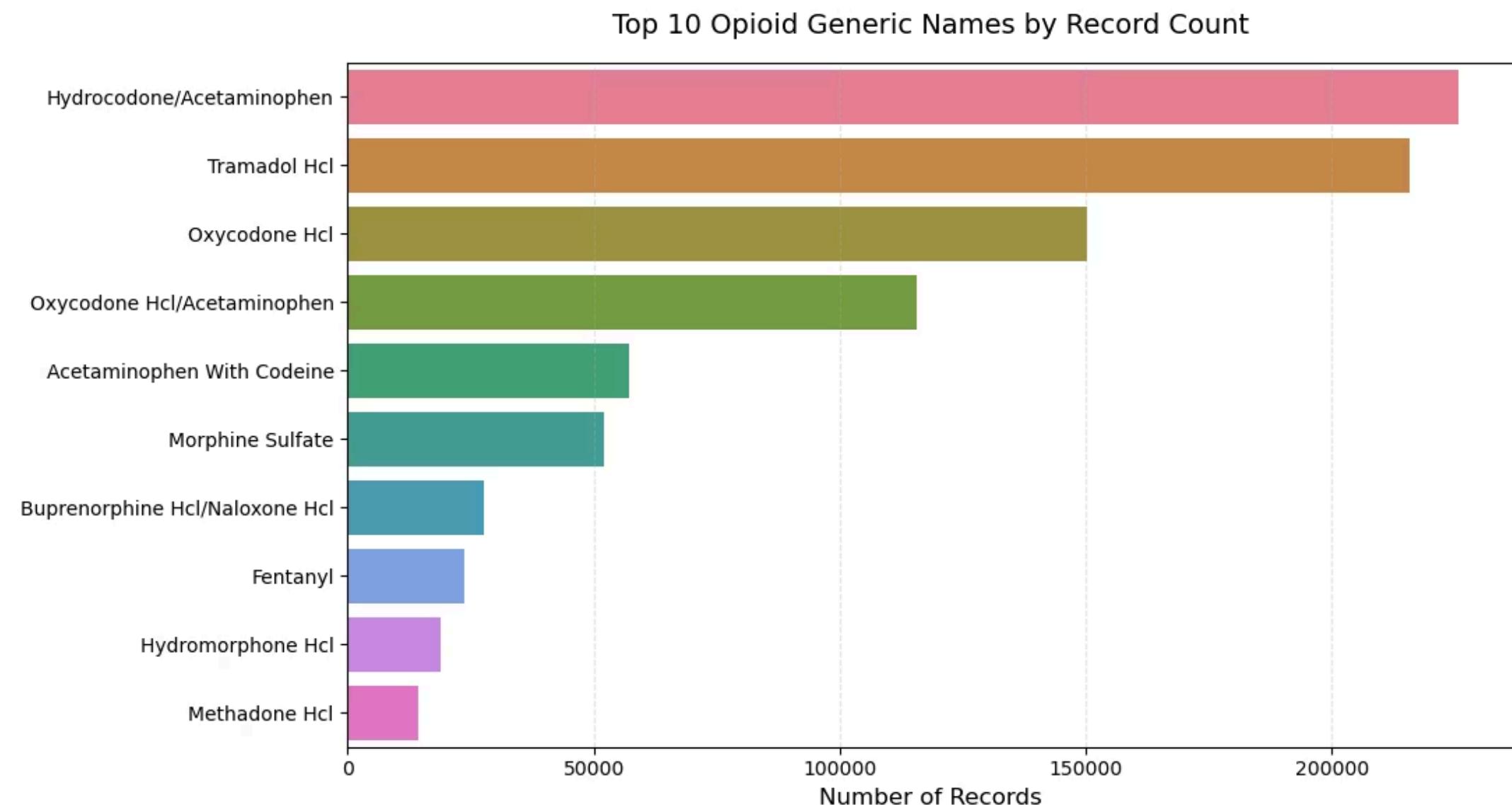
After rerunning the correlation analysis, the relationship becomes more nuanced and slightly less linear. However, a meaningful correlation still persists, suggesting consistent cost dynamics even beyond high-volume prescriptions.



Opioid Exploration in Medicare Part D Data

Filtered the Medicare Part D Prescriber dataset for opioid related drugs by using keywords matching on cleaned drug names.

By Extracting and focusing on specific therapeutic classes from my larger health care dataset,
I was able to Identified 941,497 records.



Hydrocodone/Acetaminophen leads, followed by Tramadol and Oxycodone, consistent with national Medicare opioid trends.

A Snapshot of the Prescribers

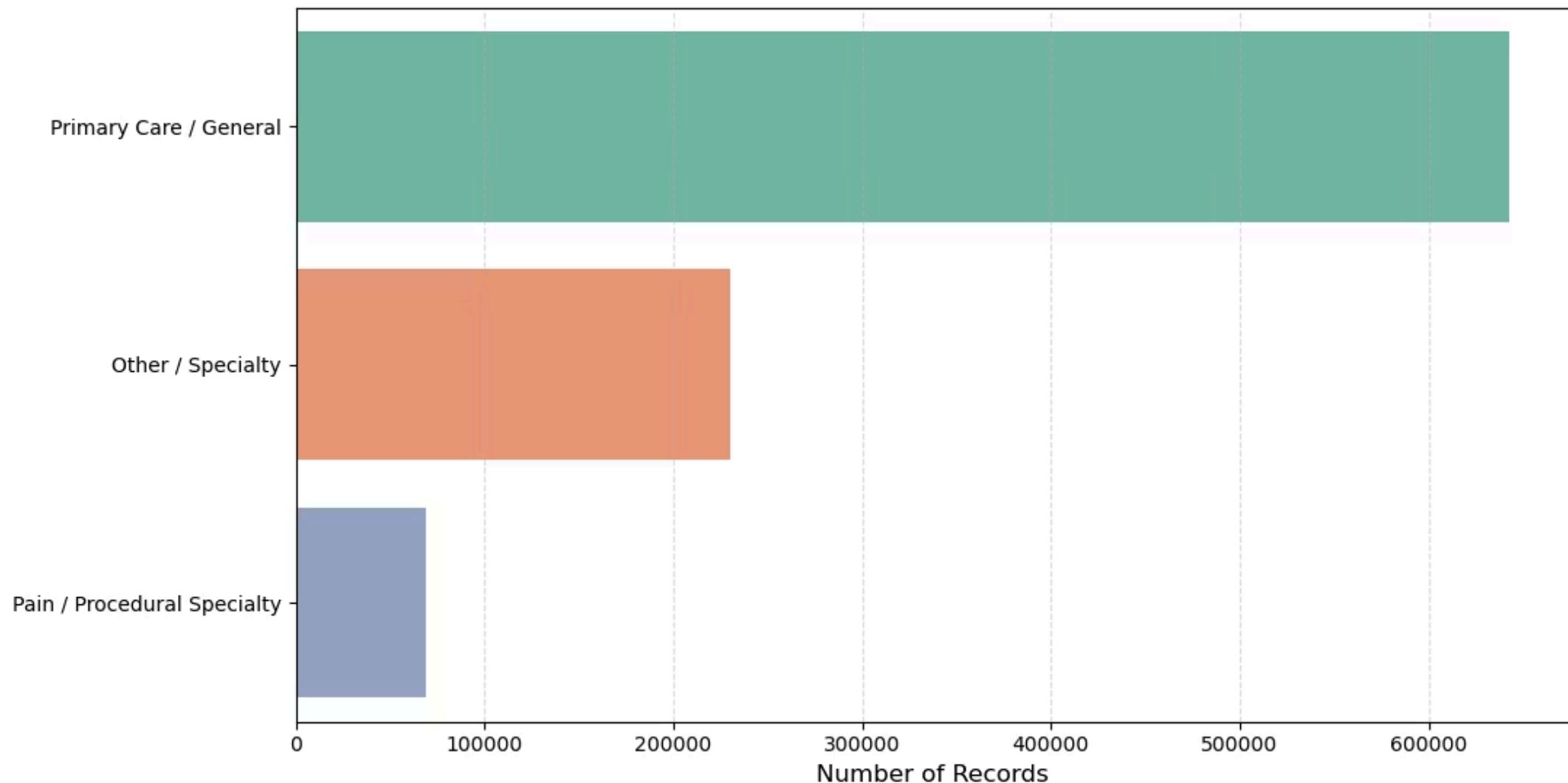
Some of the Prescribers will be combined or removed in future Analysis/Visuals



Prescriber Specialty Grouping *for Analytic Insights of Opioid Drugs*

Using the breakdown of providers in the chart from previous card, the prescribers were binned using my dataset "Prscrbr_Type" column into meaningful groups (e.g., combining 'Family Practice', 'Internal Medicine', 'Nurse Practitioner', 'Physician Assistant' into 'Primary Care / General'). This enables summary reporting on opioid volumes by provider type."

Opioid-Related Records by Prescriber Specialty Group



Primary Care / General providers account for 68.3% of opioid-related records (642,942 out of 941,497), making them a primary focus for prescribing education and quality improvement initiatives

Breakdown of Opioid Prescribing Patterns by Specialty Group

In this Medicare Part D analysis, primary care and general providers (including family medicine, internal medicine, nurse practitioners, and physician assistants) handle the majority of opioid-related prescribing by sheer volume.

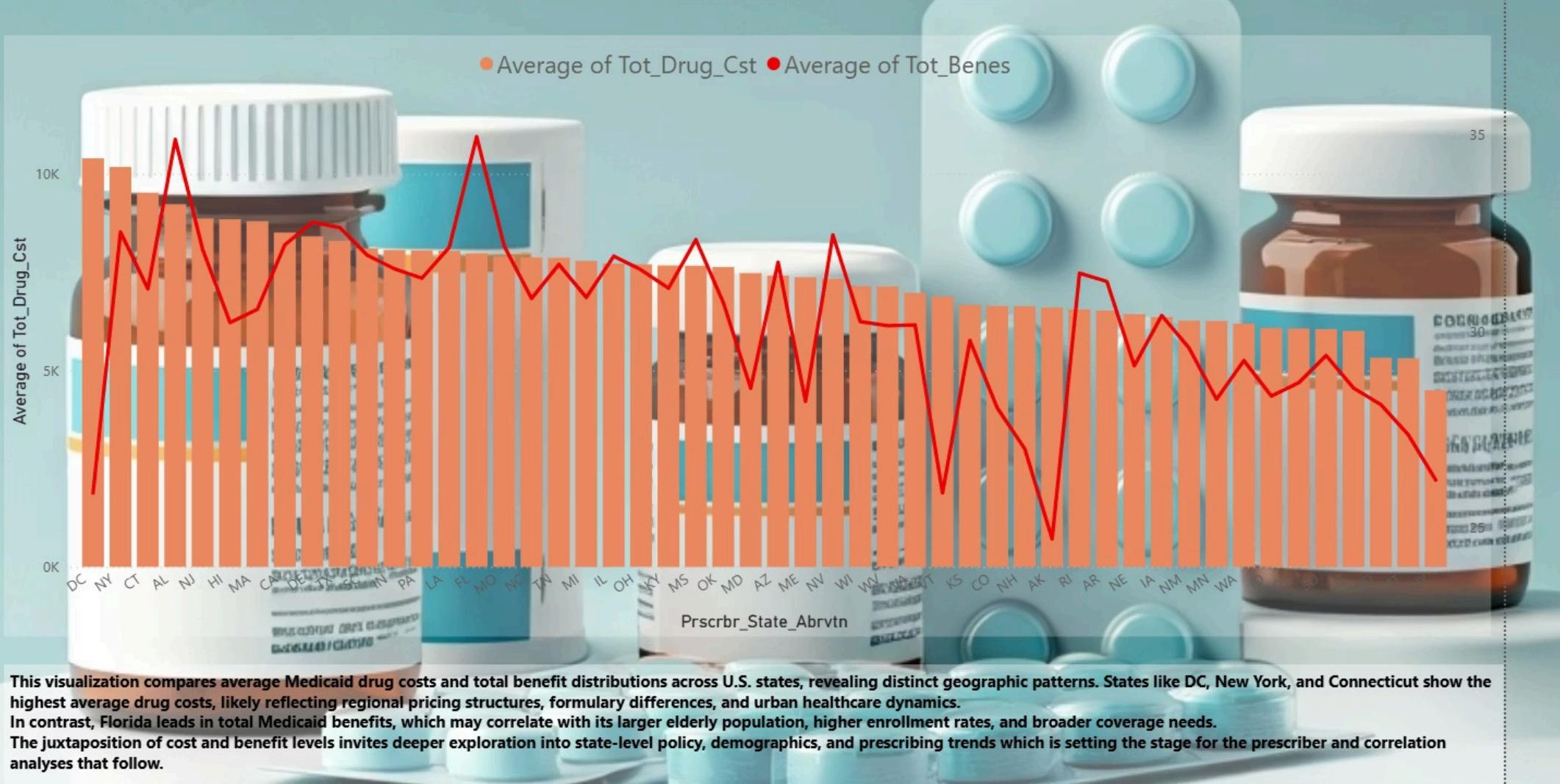
To provide a clearer picture, prescribers were grouped based on their type (Prscrbr_Type), revealing that:

- Primary Care / General accounts for 68.3% of all opioid-related records (642,942 out of 941,497).
- While they drive the bulk of overall opioid claims, specialty groups like Pain / Procedural show higher average claims per individual prescriber-drug record (142 vs. 56 in Primary Care).

Specialty Group	Records	% of Total	Total Claims	Avg Claims per Record
Primary Care / General	642,942	68.3%	36,120,563	56
Other / Specialty	229,927	24.4%	9,565,931	42
Pain / Procedural Specialty	68,628	7.3%	9,738,748	142

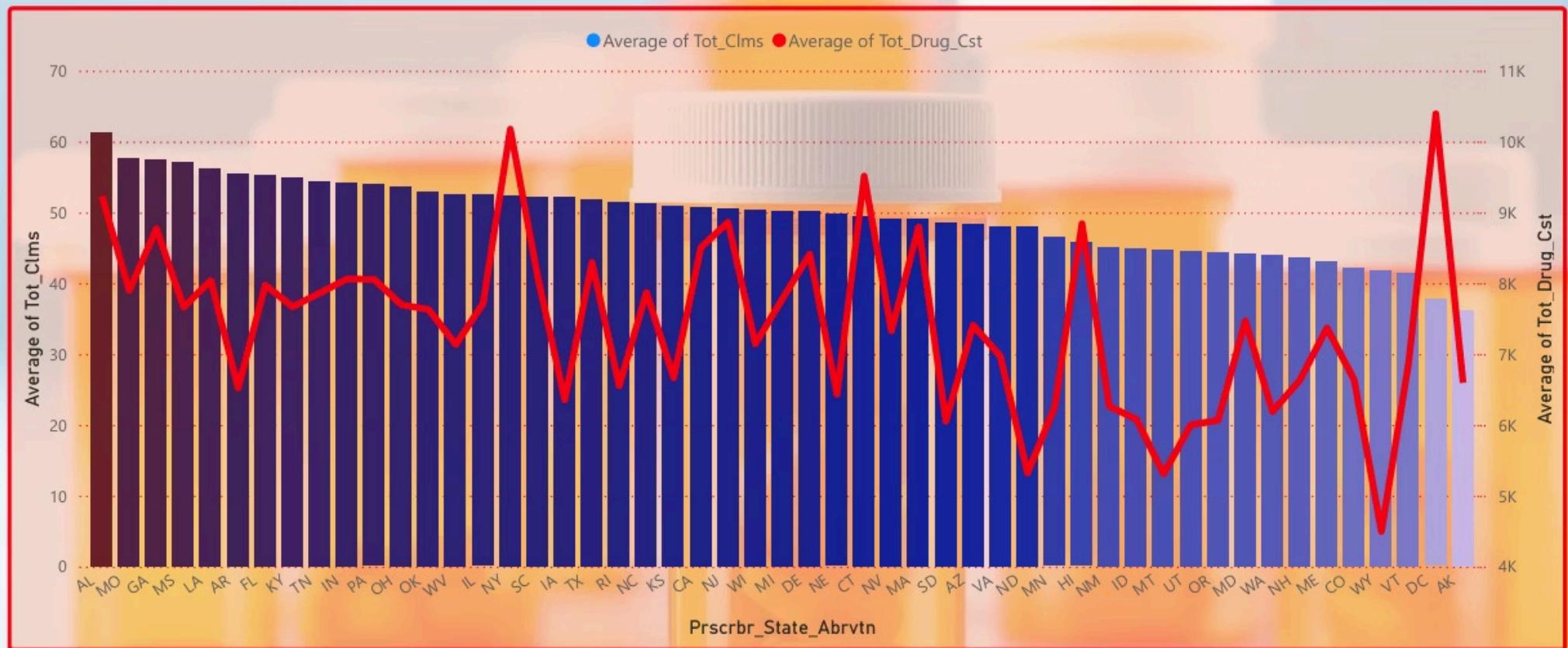
*This suggests more intensive prescribing patterns in specialized settings, which could inform targeted quality reviews or guideline adherence efforts. **code can be found on my github page***

Average Drug Cost and Benefits at US State Levels

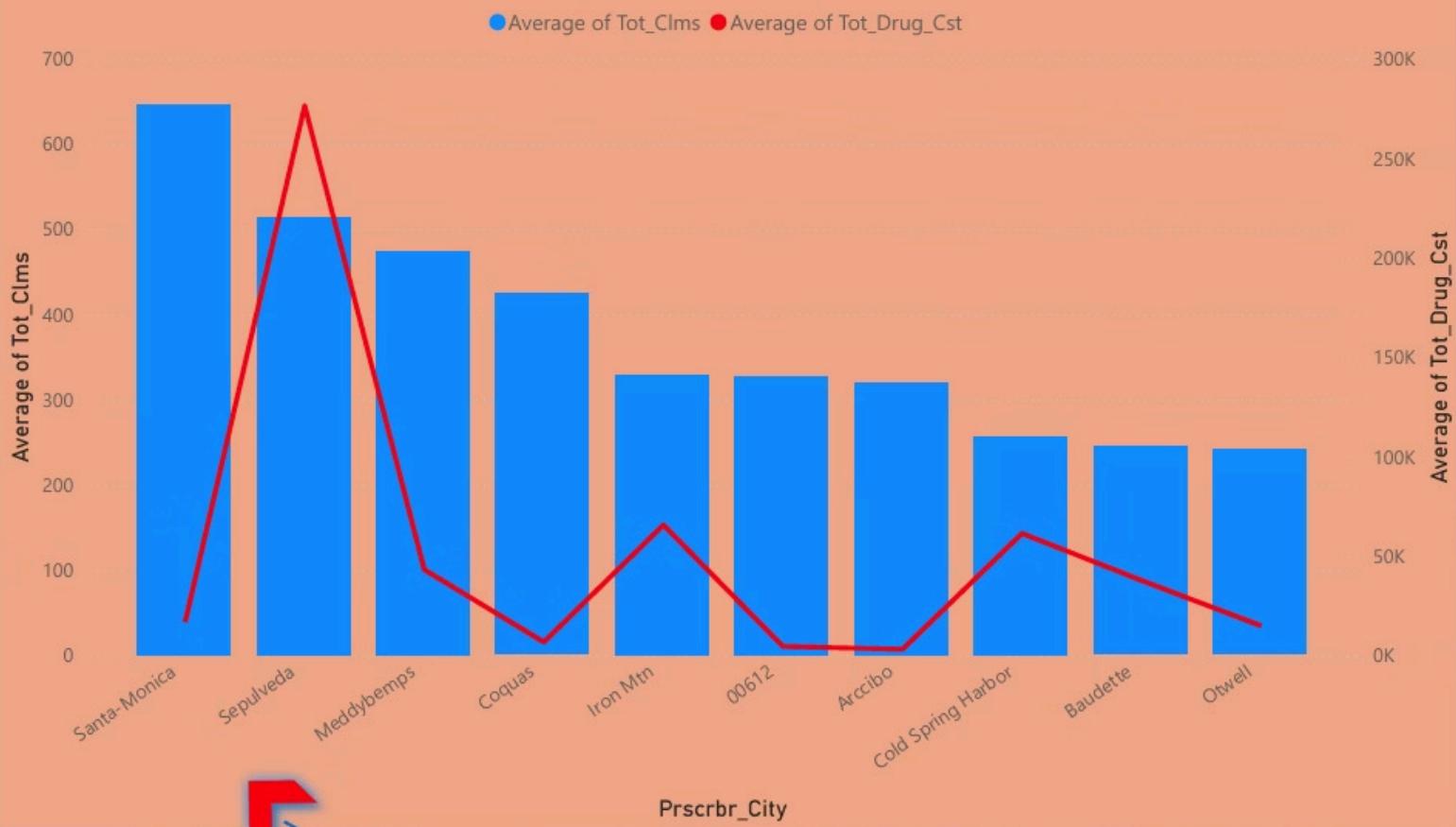


Total Claims and Drug Cost vs. States

This visualization shifts focus to average claims (bars) against drug prices (lines). DC continues to spike in drug costs (as seen in previous visual), but Alabama emerges as the leader in total claims. This reframes the analysis toward utilization patterns and prescribing behavior. Slicers added on the next visuals to dig a little deeper.



Average of Total Claims by CITY (US Territory)



North Hills, formerly known as Sepulveda, is a neighborhood in the San Fernando Valley region of Los Angeles, California. It was originally part of an agricultural community called Mission, and today it is also home to the Sepulveda VA Medical Center.

I'm a little concerned that the Drug Cost numbers here seem unusually high compared to other cities. This could point to a clerical error (perhaps repeated entries inflating the totals) or, on the more dramatic side, we might have stumbled onto something suspicious, maybe even uncovered a hidden ring. Personally, I lean toward the former, but I'll be taking a closer look at the data just to be sure.

Santa Monica is the leader in total claims. The other cities included ones that are all of the US territory, but also some with zip codes too. Messy Data still.

SLICER USED - ALL STATES
ALL CITIES SHOWN

STATE ABV

Deselect...

AK

AL

AR

AZ

CA

CO

CT

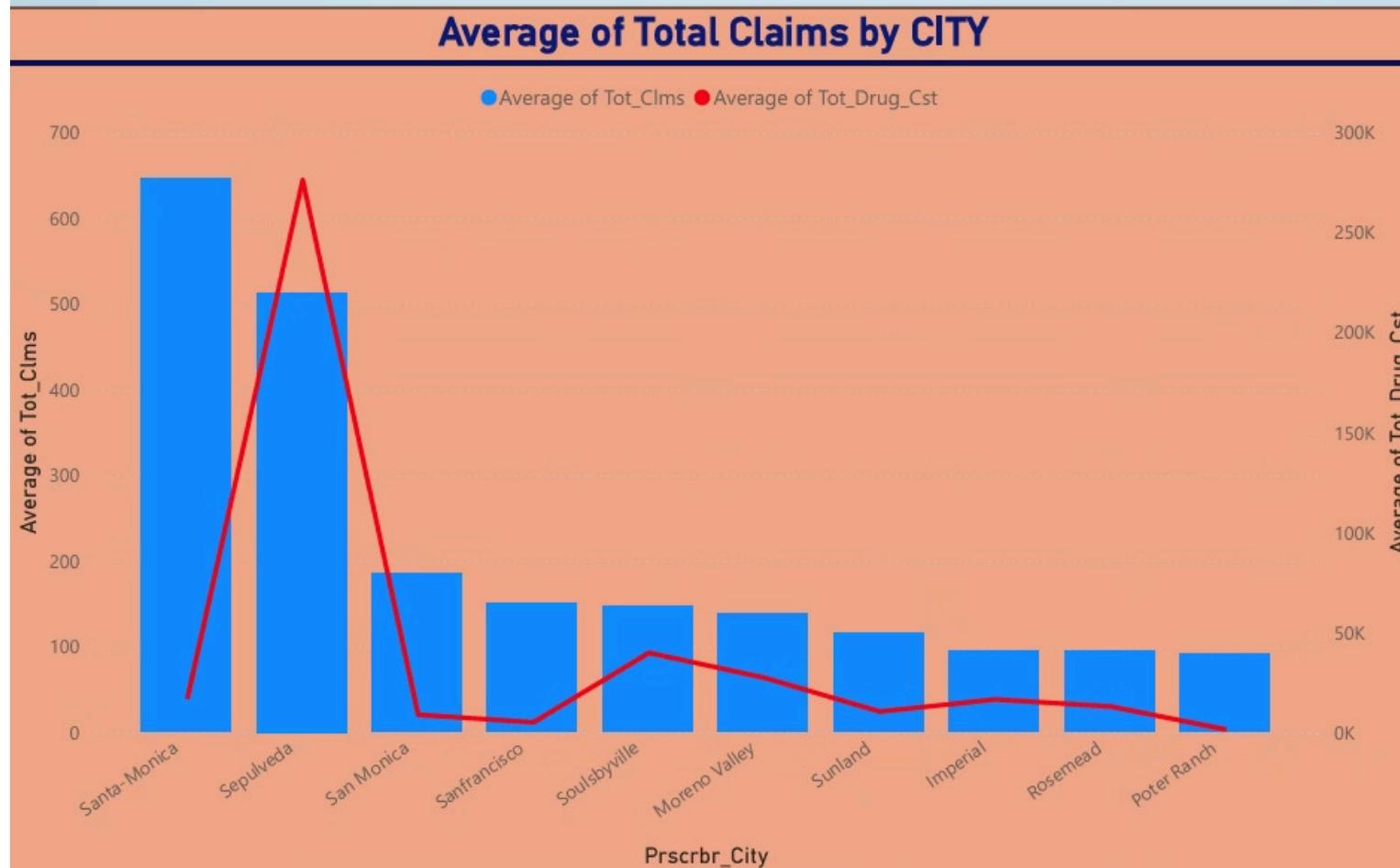


Sliced on California

Top Ten Cities based on Total Claims

As expected based on the previous slide, Santa Monica and Sepulveda are significantly higher in claims and I still believe there might be a clerical error with the drug costs for Sepulveda.

Selected CA only



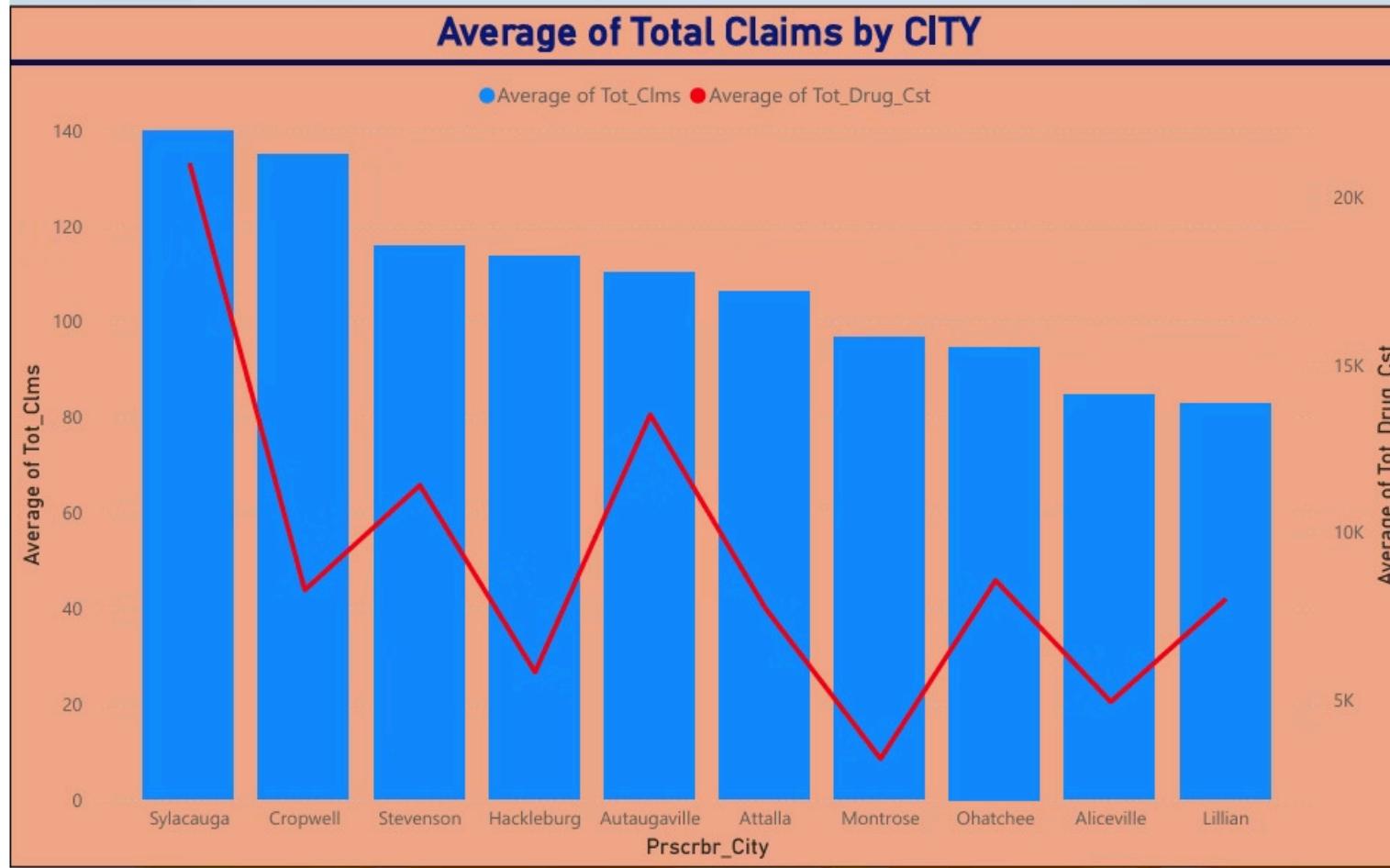


Sliced on ALABAMA

Top Ten Cities based on Total Claims

Alabama, in a previous visual, was the leading state for Claims, so wanted to dig deeper into the cities.

Roll Tide



**Selected AL
only**

STATE ABV

Select all

AK

AL

AR

AZ

CA

CO

CT

Unusual Claim Patterns

Sylacauga, AL vs.

Sepulveda (North Hills), CA

Alabama leads the nation in total healthcare claims, and when I drill down into the city-level data, Sylacauga rises to the top. That's surprising on its own, but even more so when you consider that drug costs in Sylacauga are also notably high, not just the volume of claims.

For context, North Hills, CA (formerly Sepulveda) also shows up as an outlier, particularly in terms of average drug costs. (This can be found in a previous visual.)

Cropwell, AL isn't far behind Sylacauga in total claims, but its cost profile seems more in line with expectations. That contrast makes Sylacauga's numbers stand out even more.

Could this be a case of repeated entries or a data aggregation quirk? Possibly. But when two cities on opposite coasts both show unusually high drug-related claims, it's hard not to wonder if there's a deeper patternor at least a shared reporting anomaly. Personally, I lean toward a clerical or structural explanation, but I'll be taking a closer look at the data.

Still, it's tempting to imagine we've just stumbled onto the opening scene of a healthcare noir.

Overview and Ongoing Exploration

Key highlights include:

- Robust data cleaning and suppression handling
- SQL/Python extraction for opioid patterns
- Prescriber grouping and intensity insights
- Break down by states
- Visual dashboards with PowerBI refinements with filters and slicers

The project remains open and iterative. Potential next steps:

- Merge with Medicaid or state-specific data for multi-source comparison
- Build more interactive Power BI dashboards with filters/slicers based on drug types
- Incorporate advanced stats (e.g., regression on prescribing trends)
- Expand to real-time ETL pipelines or grant-supporting reports

Open for Future Enhancements and Suggestions