**Specialty Drug Cost Analysis in Medicare Advantage Plans**

**Transition to Per-Unit Cost Analysis and Final Data Validation**

**Objective:**

Initially, our analysis focused on total drug spending for specialty drugs in Medicare and Medicaid. However, this approach did not account for price inefficiencies at the dosage level. To provide more precise insights, we transitioned to a per-unit cost analysis, which allowed us to:

* Identify cost disparities between Medicare and Medicaid for the same drugs at the dosage unit level.
* Refine policy recommendations based on drugs where Medicare overpays significantly per unit.
* Provide a stronger data-backed argument for price negotiation policies in Medicare.

**Data Processing for Per-Unit Cost Analysis**

**1. Extracting Per-Unit Cost for the Top 50 Verified Specialty Drugs**

After verifying the Top 50 Specialty Drugs using CVS Specialty Pharmacy data, we extracted per-unit cost data from the following datasets:

* Medicare Part D Spending Dataset (2022)
* Medicaid Drug Spending Dataset (2022)

**The per-unit cost was extracted from the column:**

Avg\_Spnd\_Per\_Dsg\_Unt\_Wghtd\_2022 (Average Spending Per Dosage Unit, Weighted)

This metric allowed us to directly compare the cost per unit between Medicare and Medicaid.

**2. Handling Inconsistencies Using Fuzzy Matching**

One of the biggest challenges in this transition was inconsistencies in drug naming across datasets. To address this, we:

* Used fuzzy string matching to accurately match drugs between Medicare and Medicaid datasets.
* Set a threshold (85% similarity score) to filter only high-confidence matches.
* Manually reviewed any mismatches to ensure correctness.

**3. Identifying and Removing Duplicates**

After merging Medicare and Medicaid per-unit cost data, we:

* Checked for duplicate entries where the same drug appeared multiple times due to different formulations or packaging types.
* Aggregated spending values for duplicate drugs to maintain a single record per drug.
* Ensured all 50 specialty drugs were included and verified.

**Findings:**

Duplicate entries were detected and removed based on Brnd\_Name.

**4. Final Dataset and Per-Unit Cost Discrepancy Findings**

After processing the data, we computed the per-unit cost difference as:

*Per-Unit Cost Difference = Medicare Per-Unit Cost − Medicaid Per-Unit Cost*

**5. Data Integrity Verification**

After all modifications, we validated the dataset by:

* Ensuring only the Top 50 drugs were included.
* Verifying no missing data in the per-unit cost column.
* Checking for outliers in cost differences to ensure reasonable values.

**Conclusion & Next Steps**

* The transition to per-unit cost analysis provided stronger insights into pricing inefficiencies.
* The final dataset is cleaned, verified, and ready for further analysis and reporting.
* Next steps involve updating policy recommendations based on the newly discovered per-unit cost disparities.