**Q1. How many primary care practices are in the state of Delaware?**

* **Dataset(s):** physician\_compare (with is\_primary\_care flag + site\_key).
* **Logic:**
  + Count **distinct site\_key** values where is\_primary\_care = 1.
  + Assumption: each unique site\_key = one practice location.
* **Query skeleton:**
* SELECT COUNT(DISTINCT site\_key) AS n\_primary\_care\_practices
* FROM physician\_compare
* WHERE is\_primary\_care = 1;

**Q2. Which Delaware primary care practice performed the highest volume of Medicare Annual Wellness Visits (AWVs) in 2019?**

* **Dataset(s):**
  + physician\_supplier\_hcpcs: service-level utilization/payment data (HCPCS codes).
  + physician\_compare: to identify PCPs and join on npi + site\_key.
* **Logic:**
  + AWV = HCPCS **G0438 (initial)** and **G0439 (subsequent)**.
  + Filter physician\_supplier\_hcpcs on those codes.
  + Join to physician\_compare to restrict to PCPs.
  + Aggregate total\_services (or bene\_count) by site\_key.
  + Pick top 1.
* **Query skeleton:**
* SELECT pc.site\_key,
* SUM(h.total\_services) AS total\_awv\_services
* FROM physician\_supplier\_hcpcs h
* JOIN physician\_compare pc ON h.npi = pc.npi
* WHERE pc.is\_primary\_care = 1
* AND h.hcpcs\_code IN ('G0438','G0439')
* GROUP BY pc.site\_key
* ORDER BY total\_awv\_services DESC
* LIMIT 1;

**Q3. Campaign guidance: actionable insights for increasing AWVs**

* **Dataset(s):** physician\_supplier\_agg (population served, demographics, chronic condition prevalence, risk score).
* **Logic:**
  + Use **beneficiary counts, age bands, chronic condition %**, etc., to profile practices.
  + Identify gaps → e.g. practices with **low AWV volumes** but **large panels of eligible 65+ beneficiaries**.
  + Segment by location (city/zip), gender mix, chronic conditions.
  + Actionable recommendations:
    - Target high-volume practices for best-practice sharing.
    - Support low-performing but large-panel practices with outreach tools.
    - Tailor by demographics (e.g. female-heavy, dual-eligible populations).
* **Output:** This is more of a **written/visual summary**, not pure SQL. You’d create a table or chart showing:
  + Practice, AWV count, eligible beneficiaries, AWV penetration rate.
  + Sorted list highlighting underperformers.

**🔑 Why your pipeline works**

* **02\_analysis scripts**: enforce clean datatypes, normalize zip/site\_key, ensure Delaware-only slice.
* **03\_indexes**: allow fast joins on npi, site\_key, city.
* **04\_normalize (compare)**: lets you break out specialties if you want a sensitivity check (is\_primary\_care definition).
* **05\_qa**: validates no bad ZIPs, no orphan NPIs, totals parity across agg vs. hcpcs.

So yes: with these scripts run, your database is in shape to directly query Q1 & Q2, and to mine Q3 insights.

Do you want me to now **draft the starter 02\_analysis for physician\_supplier\_hcpcs** (like we did for compare & agg), so you’ll have all three datasets in analysis-ready shape for the AWV query?

**You said:**

so we don't need to normalize for the physician\_supplier\_agg like we did for compare?

**ChatGPT said:**

Exactly — you **don’t need a separate normalization step for physician\_supplier\_agg** the way you did for physician\_compare.

Here’s why:

* **physician\_compare**:
  + Has *multiple specialty slots* (primary + up to 4 secondaries).
  + To analyze specialties properly, you need a **one-row-per-specialty** table → hence physician\_specialty.
* **physician\_supplier\_agg**:
  + Already **aggregated at the NPI level**.
  + All measures (HCPCS counts, services, charges, demographics, risk, etc.) are **single-valued per provider**.
  + There are no repeating groups or embedded multi-column lists to break out.

So for physician\_supplier\_agg, your **02 (analysis table) + 03 (indexes) + 05 (QA)** is sufficient. No 04 (normalize) needed.

👉 The *only place normalization is essential* is in **physician\_compare** (to break out specialties) and sometimes in **physician\_supplier\_hcpcs** (if you want to roll service-level rows into categories like “AWV” or “Drug vs Non-Drug”).

Do you want me to prepare the **02/03/05 set for physician\_supplier\_hcpcs** next, with the HCPCS normalization logic included (to flag AWV codes and make Q2 easier)?

Top of Form

Bottom of Form

ChatGPT can make mistakes. Check important info.