Independent Pharmacy Database

What We Built | What It Contains | What Comes Next

RETAILMYMEDS

Records	**41,775** independent pharmacies
Coverage	All 50 states + District of Columbia
Source	CMS NPPES Data Dissemination (Feb 9, 2026)

PREPARED FOR **Arica Collins, PharmD** & **Kevin McCarron**

PREPARED BY **Matthew Scott**

February 12, 2026

Contents

What This Document Is

This document is a standalone record of one working session. It does not depend on or reference any previously created RetailMyMeds deliverables. Its purpose is threefold:

1. **Document what was built** — the independent pharmacy database, how it was constructed, and exactly what it contains.

2. **Anticipate Arica's questions** — the six questions she will ask when she sees 41,775 records, and why each one matters.

3. **Map the enrichment path** — what public data sources exist to answer each question, ranked by difficulty and value.

Everything in this document is grounded in the data file `qualified_independent_pharmacies_feb2026.csv` and the publicly verifiable sources used to create it. No projections, no strategy, no sales language. Just the data and what it needs to become useful.

The Deliverable

File: `qualified_independent_pharmacies_feb2026.csv`

A single CSV file containing 41,775 records, sorted by estimated operating status (Active first), then by state, city, and name. Each record represents one NPI-registered community/retail pharmacy in the United States that is *not* a chain, *not* a hospital/health system pharmacy, *not* a government pharmacy, and *not* a specialty/mail-order/institutional pharmacy.

Every record includes the pharmacy's estimated operating status and the name, title, and phone number of the Authorized Official (typically the owner or pharmacist-in-charge).

Fields Per Record (16 columns)

Column	Type	Description
npi	String	10-digit National Provider Identifier
display_name	String	Best available pharmacy name (DBA if present, else legal name)
legal_name	String	Legal business name as registered with CMS
dba_name	String	"Doing business as" name (often blank)
address_1	String	Street address line 1 (practice location)
address_2	String	Street address line 2 (suite, unit, etc.)
city	String	City of practice location
state	String	Two-letter state code
zip	String	5-digit ZIP code
phone	String	10-digit phone number (no formatting)
enumeration_date	Date	Date the NPI was first assigned
last_updated	Date	Date the NPI record was last modified
estimated_status	String	Active / Likely Active / Uncertain / Likely Closed
owner_name	String	Authorized Official full name
owner_title	String	Title or position (CEO, Owner, PIC, etc.)
owner_phone	String	Authorized Official phone number

Enrichment Coverage

Column	Coverage	Source
estimated_status	41,775 / 41,775 (100%)	Computed from last_updated
owner_name	41,775 / 41,775 (100%)	NPPES Authorized Official fields
owner_title	41,775 / 41,775 (100%)	NPPES Authorized Official fields
owner_phone	41,775 / 41,775 (100%)	NPPES Authorized Official fields

Supporting Files

File	Size	Purpose
independent_pharmacies_usa_feb2026.csv	5.3 MB	Base extraction (12 columns, no enrichment)
state_summary.csv	436 B	Count of independent pharmacies per state
chain_pharmacies_excluded.csv	5.1 MB	42,336 chain records removed (audit trail)
non_independent_excluded.csv	431 KB	2,864 non-independent records removed (audit trail)
extract_independent_pharmacies.py	20 KB	Base extraction script
enrich_pharmacies.py	8 KB	Enrichment script (status + owner)

How It Was Built

Source Data

The CMS National Plan and Provider Enumeration System (NPPES) Data Dissemination file, released **February 9, 2026**. This is the authoritative federal database of all healthcare providers in the United States.

- **Download URL:** https://download.cms.gov/nppes/NPI_Files.html

- **File:** NPPES_Data_Dissemination_February_2026.zip (1,058 MB compressed)

- **Extracted size:** 11.2 GB (main CSV: npidata_pfile_20050523-20260208.csv)

- **Total records:** 9,368,082 provider NPIs

Filtering Pipeline

Step	Filter	Records	Removed
0	Raw NPPES file	9,368,082	—
1	Entity Type = Organization (Type 2)	—	Individuals
2	NPI Deactivation Reason = blank (active)	—	Deactivated
3	Taxonomy = 3336C0003X (Community/Retail)	88,378	Non-pharmacy
4	State = US 50 states + DC	—	Territories
5	Chain pharmacy name exclusion (80+ patterns)	–42,336	Chains
6	Non-independent exclusion (hospital/govt/etc.)	–2,864	Institutional
	Final output	**41,775**	

Chain Exclusion Patterns (80+)

The extraction script matches organization names against compiled regular expressions. Categories of excluded chains:

- National retail (CVS, Walgreens, Walmart, Rite Aid)
- Grocery chains (Kroger, Safeway, Albertsons, Publix, H-E-B, Meijer, etc.)
- Big box (Costco, Target, Sam's Club, BJ's)
- Historical chains with active NPIs (Genovese, Hook-SupeRx, Pathmark, Shopko, Phar-Mor, Revco, etc.)

- Pharmacy benefit chains (Express Scripts, Optum, Omnicare, PharMerica, Genoa, CenterWell)
- Grocery subsidiaries (Carr-Gottstein, Roundy's, Riser Foods, Giant of Maryland, etc.)
- Regional drug chains (Bartell, Kerr Drug, Discount Drug Mart, Fred's, Snyder's, etc.)

Non-Independent Exclusion Patterns

- Hospital pharmacies
- Health system pharmacies
- VA / military pharmacies
- Indian Health Service / tribal

- Corrections / detention
- University / academic
- Federal / state / county government
- Infusion pharmacies

- Home health / home care
- Long-term care / LTC
- Nursing facility / skilled nursing
- Hospice
- Closed-door pharmacies

- Nuclear pharmacies
- Clinical trial / research
- Mail-order / central fill
- Telepharmacy

The 41,775 vs. 19,000 Question

NCPA's 2025 Digest reports 18,960 independent pharmacy *locations* as of July 2025. The PCMA/NCPDP/Quest Analytics data from February 10, 2026 shows a net increase of 67 pharmacies year-over-year, putting the current operating count at approximately **19,027**.

This file contains **41,775** records — more than double. The gap is not an error. It is a known property of the NPI database:

> **NPI registrations persist after a pharmacy closes.** CMS does not automatically deactivate an NPI when a business ceases operations. The pharmacy (or its successor) must actively request deactivation. Many never do.

Recency Distribution

The `last_updated` field indicates when the NPI record was last modified by the provider. Records modified recently are far more likely to represent operating pharmacies.

Last Updated	Records	% of Total	Interpretation
2024–2026 (Active)	10,359	24.8%	Almost certainly operating
2020–2023 (Likely Active)	11,660	27.9%	Likely operating (updated within 6 years)
2015–2019 (Uncertain)	10,185	24.4%	May be open or closed
Pre-2015 (Likely Closed)	9,571	22.9%	Almost certainly closed
	41,775	100%	

The 10,359 records updated 2024–2026, plus a portion of the 11,660 updated 2020–2023, account for the approximately 19,000 pharmacies that NCPA counts as currently operating. The remaining records are historical — pharmacies that closed but whose NPI was never deactivated.

Why the Full Dataset Is Still Valuable

- **It is a superset.** Every currently operating independent pharmacy is in this file.

- **Recently closed pharmacies have owners who may be opening new locations** or transitioning to consulting — still potential contacts.

- **Adding an `estimated_status` column** based on `last_updated` immediately segments the file into actionable tiers.

- **Cross-referencing with pharmacy dispensing data** (see Section 6) via NCPDP DataQ or CMS PDE can confirm which NPIs are actively dispensing.

State-by-State Inventory

All 50 states and the District of Columbia are represented. Sorted by record count, descending.

#	State	Count	#	State	Count	#	State	Count
1	NY	4,819	18	MD	649	35	OR	282
2	TX	3,977	19	MS	643	36	NV	246
3	CA	3,827	20	VA	638	37	ID	240
4	FL	3,442	21	AR	629	38	MT	215
5	MI	2,110	22	WI	593	39	ND	173
6	PA	1,677	23	WA	585	40	NM	157
7	NJ	1,576	24	SC	534	41	SD	141
8	NC	1,236	25	MN	505	42	HI	137
9	GA	1,204	26	KS	442	43	ME	111
10	IL	1,086	27	IA	430	44	DC	88
11	OH	983	28	MA	401	45	DE	80
12	LA	924	29	WV	369	46	WY	69
13	AL	911	30	CT	359	47	VT	68
14	TN	906	31	IN	358	48	NH	62
15	KY	856	32	AZ	357	49	RI	58
16	MO	847	33	UT	331	50	AK	49
17	OK	783	34	NE	321	51	(DC)	(incl.)

Six Questions Arica Will Ask

When Arica sees 41,775 rows in a spreadsheet, she will ask six questions. Each one represents a gap between what the data currently contains and what she needs to take action. These are numbered 1 through 6 in priority order.

Question 1: "Which of these are actually still open?"

RESOLVED — The `estimated_status` column is now in the qualified file.

Why she asks it: 41,775 is more than double the ~19,000 she knows exist. She needs to know which rows to trust before investing time in any of them.

What the data now provides: The `estimated_status` column, computed from the `last_updated` field:

Status	Criteria	Count	Action
Active	Updated 2024–2026	10,359	Outreach immediately
Likely Active	Updated 2020–2023	11,660	Outreach with verification
Uncertain	Updated 2015–2019	10,185	Verify before outreach
Likely Closed	Updated pre-2015	9,571	Skip or use for market research

How to use it: Filter to "Active" and "Likely Active" for 22,019 pharmacies that are almost certainly operating. This aligns with the ~19,000 NCPA count plus a margin of recently closed businesses whose owners may still be reachable.

Question 2: "How many prescriptions are they filling?"

REQUIRES PAID DATA — Not available through free public sources.

Why she asks it: RetailMyMeds charges \$275/month. A pharmacy doing 1,500 Rx/month won't generate enough savings to justify the cost. She needs volume to qualify prospects.

What the data currently provides: Nothing. Monthly Rx volume is not in the NPI database.

What's needed: A new column — `monthly_rx_volume` or a proxy for it.

Critical finding from this session: The CMS Medicare Part D "Prescriber" Public Use Files are keyed to *prescriber NPIs* (doctors who write prescriptions), *not* pharmacy/dispenser NPIs. We downloaded and tested both the Part D Prescriber PUF (556 MB) and the Part D by Provider and Drug file (3.6 GB). Zero pharmacy NPIs from our database appear in either file. Pharmacy-level dispensing data lives in the Part D Event (PDE) files, which require a formal CMS Data Use Agreement — they are not publicly downloadable.

Available sources:

- **NCPDP DataQ** — Paid subscription. Contains actual dispensing volume, payer mix, and more. This is the only comprehensive source for pharmacy-level Rx counts.

- **CMS PDE Data Use Agreement** — Free data, but requires a formal application and approval process through CMS/ResDAC. Contains pharmacy-level dispensing records for all Medicare Part D claims.

Difficulty: **Hard** — No free public download exists for pharmacy-level dispensing volume. NCPDP DataQ or CMS PDE access is required.

Question 3: "Which ones are filling GLP-1s?"

REQUIRES PAID DATA — Same data access barrier as Question 2.

Why she asks it: The GLP-1 loss problem (\$37–42 loss per fill, 95% of pharmacies underwater) is the highest-urgency pain point. Pharmacies actively dispensing semaglutide (Ozempic/Wegovy) and tirzepatide (Mounjaro/Zepbound) are the hottest prospects.

What the data currently provides: Nothing. Drug-level dispensing is not in the NPI database.

What's needed: A boolean column — `dispenses_glp1` — or a count of GLP-1 claims.

Critical finding: Same as Question 2. The CMS Part D "by Provider and Drug" file (3.6 GB, 26M+ rows) tracks which *doctors* prescribed which drugs, not which *pharmacies* dispensed them. Zero pharmacy NPIs match. Drug-level pharmacy dispensing data requires CMS PDE access or NCPDP DataQ.

Available sources:

- **NCPDP DataQ** — Paid subscription. Can filter by drug category at the pharmacy level.

- **CMS PDE Data Use Agreement** — Contains pharmacy-level drug dispensing records. Can filter to semaglutide/tirzepatide NDCs.

- **Practical workaround:** Nearly every independent pharmacy filling 3,000+ Rx/month is dispensing GLP-1s. The volume question (Q2) effectively answers this one. A pharmacy with meaningful Part D volume is dispensing GLP-1s.

Difficulty: **Hard** — Same paid data sources as Question 2.

Question 4: "What pharmacy management system are they on?"

Why she asks it: RetailMyMeds integrates with pharmacy management systems (PMS). PioneerRx is Priority 1 for integration, Liberty is Priority 2, PrimeRx is Priority 3. Knowing the PMS system lets her prioritize pharmacies where integration will be seamless.

What the data currently provides: Nothing. PMS system is not in the NPI database.

What's needed: A new column — `pms_system`.

Available sources:

- **NCPDP DataQ** — Paid subscription. Contains "software vendor" field that identifies the PMS system per pharmacy. This is the definitive source.

- **PMS vendor customer directories** — Some vendors (PioneerRx, Liberty) publish customer maps or customer counts by state. These are partial but free.

- **State board of pharmacy data** — Some states collect PMS system information as part of licensure. Availability varies widely by state.

Difficulty: <mark>Hard</mark> — No free comprehensive public source. NCPDP DataQ is the only path to full coverage. Vendor directories provide partial data for specific systems.

Question 5: "Who's the owner and what's their email?"

<mark>PARTIALLY RESOLVED</mark> — Owner name, title, and phone are now in the qualified file. Email requires paid data.

Why she asks it: A spreadsheet of pharmacy names and addresses is a starting point. To do outreach, she needs the owner's name and a way to contact them directly.

What the data now provides:

- `owner_name` — Full name of the Authorized Official (100% coverage, 41,775 / 41,775)

- `owner_title` — Title or position: CEO, Owner, President, Pharmacist-in-Charge (PIC), RPH, etc.

- `owner_phone` — Direct phone number of the Authorized Official (separate from the pharmacy phone)

What's still needed:

- **Owner email** — *Not* in NPI data. Available through:

    - State board of pharmacy licensee directories (free, but 50 different sources with different formats and access methods)
    - Commercial data providers (Data Axle, ZoomInfo, etc. — paid)
    - LinkedIn prospecting (manual, one-at-a-time)
    - NCPA member directory (requires NCPA membership)

Difficulty: Owner name/title/phone is <mark>Done</mark>. Owner email is <mark>Hard</mark> (50 different state sources or paid commercial data).

Question 6: "Are any of these already my customers or prospects?"

Why she asks it: Before doing outreach to 10,000+ pharmacies, she needs to exclude pharmacies she's already engaged with — current customers, active prospects in her pipeline, and pharmacies that have already declined.

What the data currently provides: Nothing. Customer/prospect status is internal to RetailMyMeds.

What's needed: A new column — `rmm_status` — with values like: Customer, Prospect, Declined, New. This requires Arica to provide her customer list and CRM export so it can be matched against the database by NPI, pharmacy name, or address.

Difficulty: `Easy` — once Arica provides her internal data. This is a simple merge/match operation.

Enrichment Priority Matrix

The six questions map to seven enrichment tasks. Updated to reflect what was completed in this session and what was discovered about data availability:

#	Enrichment	Source	Cost	Difficulty	Status
1	Active vs. closed status	`last_updated` field (in data)	Free	Easy	**Done**
2	Owner name/title/phone	NPI "Authorized Official" fields	Free	Easy	**Done**
3	Customer/prospect flag	Arica's CRM export	Free	Easy	**Blocked**
4	Monthly Rx volume	NCPDP DataQ or CMS PDE DUA	Paid	Hard	**Paid data**
5	GLP-1 dispensing flag	NCPDP DataQ or CMS PDE DUA	Paid	Hard	**Paid data**
6	PMS system	NCPDP DataQ subscription	Paid	Hard	**Paid data**
7	Owner email	State boards / commercial data	Varies	Hard	**Paid data**

Items 1–2 are complete. Every row in the qualified file has a status and an owner.

Item 3 requires Arica's customer list — a simple merge once she provides it.

Items 4–7 all require paid data. The key discovery in this session: CMS Part D public use files track *prescribers* (doctors), not *dispensers* (pharmacies). There is no free public download of pharmacy-level dispensing volume, drug-specific fills, or PMS system. The three viable paid paths:

- **NCPDP DataQ** (https://dataq.ncpdp.org/) — Subscription. Contains Rx volume, payer mix, PMS system, and more per pharmacy. Covers items 4, 5, and 6 in one purchase.

- **CMS PDE Data Use Agreement** — Free data, formal application through ResDAC (https://resdac.org/). Contains pharmacy-level Medicare dispensing records. Covers items 4 and 5.

- **Commercial data providers** (Data Axle, ZoomInfo) — Paid. Covers item 7 (owner email).

Data Sources and Citations

Every number in this document is traceable to a public source.

1. **CMS NPPES Data Dissemination (February 9, 2026)**
   https://download.cms.gov/nppes/NPI_Files.html
   Source of the 9,368,082 NPI records. Updated monthly by CMS.

2. **NCPA 2025 Digest Report (October 2025)**
   https://ncpa.org/newsroom/news-releases/2025/10/19/ncpa-releases-2025-digest-report
   Reports 18,960 independent pharmacy locations as of July 2025. Published by NCPA in partnership with Cardinal Health.

3. **PCMA / NCPDP / Quest Analytics (February 10, 2026)**
   https://www.pcmanet.org/press-releases/independent-pharmacies-increase-in-2026/02/10/2026
   Reports 67 net new independent pharmacies from 2025 to 2026 and 321 net new over the past 10 years. Based on NCPDP data analyzed by Quest Analytics.

4. **NPI Taxonomy Code 3336C0003X**
   https://npidb.org/taxonomy/3336C0003X/
   Healthcare Provider Taxonomy classification for Community/Retail Pharmacy.

5. **CMS Medicare Part D Prescriber Public Use File**
   https://data.cms.gov/provider-summary-by-type-of-service/medicare-part-d-prescribers
   Annual release of claim counts and costs per *prescriber* NPI (doctors, not pharmacies). Downloaded and tested during this session — contains zero pharmacy NPIs. Not usable for pharmacy-level enrichment.

6. **CMS Medicare Part D Prescribers by Provider and Drug**
   https://data.cms.gov/provider-summary-by-type-of-service/medicare-part-d-prescribers/medicare-part-d-prescribers-by-provider-and-drug
   Drug-level claims per *prescriber* NPI. Downloaded and tested (3.6 GB, 26M+ rows) — zero pharmacy NPI matches. Not usable for GLP-1 dispensing identification.

7. **NCPDP DataQ Pharmacy Database**
   https://dataq.ncpdp.org/
   Paid subscription. Definitive source for PMS system, dispensing volume, payer mix, and other pharmacy-level attributes.

8. **NPPES NPI Registry API (v2.1)**
   https://npiregistry.cms.hhs.gov/api-page
   Public API for querying individual NPI records. Used during this session to validate data structure and test taxonomy filtering before bulk download.

Reproduction Instructions

The database can be regenerated from scratch at any time using the two scripts and the CMS source file. This ensures the data is never stale and the methodology is auditable.

1. Download the latest NPPES monthly file from https://download.cms.gov/nppes/NPI_Files.html

2. Place the ZIP in the `Pharmacy_Database/` directory (rename to `nppes_feb2026.zip` or update the script path)

3. Run: `python3 extract_independent_pharmacies.py` — produces the base 12-column CSV

4. Run: `python3 enrich_pharmacies.py` — adds status and owner columns, produces the qualified 16-column CSV

Runtime: Extract ~2.5 minutes + Enrich ~2 minutes = ~5 minutes total.
Dependencies: Python 3.10+ standard library only. No `pip install` required.

What Happens Next

The qualified CSV is the deliverable. It contains 41,775 independent pharmacies, each with an operating status estimate and the owner's name, title, and phone number. The file is sorted so Active pharmacies appear first.

What Arica Can Do Today

1. Open the CSV in Excel

2. Filter `estimated_status` to "Active" (10,359 rows)

3. Filter `state` to her target state(s)

4. Sort by `city` or `display_name`

5. Each row has the owner's name and a phone number to call

What Requires a Decision

Enrichment items 3–7 are blocked — not by technical difficulty, but by access:

- **Item 3 (customer flag):** Blocked on Arica providing her CRM data. Free and fast once she does.

- **Items 4–6 (Rx volume, GLP-1, PMS):** All behind the same paywall — NCPDP DataQ. One subscription covers all three. The question for Kevin: is the subscription worth it?

- **Item 7 (owner email):** Commercial data providers or 50-state manual aggregation. Separate cost decision.

The file as delivered is actionable without any of these. The enrichment items make it *more* actionable — they let Arica prioritize who to call first, not whether she can call at all.

How This Connects

This database is the foundation layer. On top of it, the **GLP-1 Pharmacy Targeting** analysis (separate report) scores and ranks all 41,775 pharmacies by financial exposure, area health context, and outreach priority — giving Arica a prioritized list of who to call and what to say.

When those pharmacies engage — through Arica's direct outreach or through the RetailMyMeds website — the **qualification form** that Kevin is building in Wix Studio scores them using the same dimensions and delivers a personalized scorecard automatically. The database identifies prospects. The website converts them. Together they give Arica a system, not just a spreadsheet.

> **Repository:** `github.com/guitargnarr/RetailMyMeds` (private)
>
> **Branch:** `main`
>
> **All files version-controlled and reproducible.**