## Prescribers Database

For this exericse, you'll be working with a database derived from the [Medicare Part D Prescriber Public Use File](https://www.hhs.gov/guidance/document/medicare-provider-utilization-and-payment-data-part-d-prescriber-0). More information about the data is contained in the Methodology PDF file. See also the included entity-relationship diagram.

1.

a. Which prescriber had the highest total number of claims (totaled over all drugs)? Report the npi and the total number of claims.

b. Repeat the above, but this time report the nppes\_provider\_first\_name, nppes\_provider\_last\_org\_name, specialty\_description, and the total number of claims.

2.

a. Which specialty had the most total number of claims (totaled over all drugs)?

b. Which specialty had the most total number of claims for opioids?

c. \*\*Challenge Question:\*\* Are there any specialties that appear in the prescriber table that have no associated prescriptions in the prescription table?

d. \*\*Difficult Bonus:\*\* \*Do not attempt until you have solved all other problems!\* For each specialty, report the percentage of total claims by that specialty which are for opioids. Which specialties have a high percentage of opioids?

3.

a. Which drug (generic\_name) had the highest total drug cost?

b. Which drug (generic\_name) has the hightest total cost per day? \*\*Bonus: Round your cost per day column to 2 decimal places. Google ROUND to see how this works.\*\*

4.

a. For each drug in the drug table, return the drug name and then a column named 'drug\_type' which says 'opioid' for drugs which have opioid\_drug\_flag = 'Y', says 'antibiotic' for those drugs which have antibiotic\_drug\_flag = 'Y', and says 'neither' for all other drugs. \*\*Hint:\*\* You may want to use a CASE expression for this. See https://www.postgresqltutorial.com/postgresql-tutorial/postgresql-case/

b. Building off of the query you wrote for part a, determine whether more was spent (total\_drug\_cost) on opioids or on antibiotics. Hint: Format the total costs as MONEY for easier comparision.

5.

a. How many CBSAs are in Tennessee? \*\*Warning:\*\* The cbsa table contains information for all states, not just Tennessee.

b. Which cbsa has the largest combined population? Which has the smallest? Report the CBSA name and total population.

c. What is the largest (in terms of population) county which is not included in a CBSA? Report the county name and population.

6.

a. Find all rows in the prescription table where total\_claims is at least 3000. Report the drug\_name and the total\_claim\_count.

b. For each instance that you found in part a, add a column that indicates whether the drug is an opioid.

c. Add another column to you answer from the previous part which gives the prescriber first and last name associated with each row.

7. The goal of this exercise is to generate a full list of all pain management specialists in Nashville and the number of claims they had for each opioid. \*\*Hint:\*\* The results from all 3 parts will have 637 rows.

a. First, create a list of all npi/drug\_name combinations for pain management specialists (specialty\_description = 'Pain Management) in the city of Nashville (nppes\_provider\_city = 'NASHVILLE'), where the drug is an opioid (opiod\_drug\_flag = 'Y'). \*\*Warning:\*\* Double-check your query before running it. You will only need to use the prescriber and drug tables since you don't need the claims numbers yet.

b. Next, report the number of claims per drug per prescriber. Be sure to include all combinations, whether or not the prescriber had any claims. You should report the npi, the drug name, and the number of claims (total\_claim\_count).

c. Finally, if you have not done so already, fill in any missing values for total\_claim\_count with 0. Hint - Google the COALESCE function.