**Status Report: Analysis of Pharmacy Expansion Opportunities**

Shih Min Lin, Tzu-Yun Huang, Hung-Chen Hsu

Group 5

**Initial Research Questions and Industry of Interest**

Our team focuses on the healthcare industry within the United States, specifically the pharmacy sector. As a real estate investment firm, we aim to identify areas with high demand for pharmacies, defined as regions with a large elderly population and insufficient pharmacy services. This analysis will help guide decisions on where to open new stores for maximum profitability. We are leveraging anonymized cell phone traffic data from Safegraph to identify consumer traffic patterns and underserved areas.

**Specific Questions to Answer from the Data**

1. Identify Popular Pharmacies: Which pharmacy brand in the healthcare industry has had the most traffic in the past five years?

Findings: Our analysis revealed that Walgreens has the highest traffic among pharmacy brands, making it a dominant player in the market. Understanding its traffic patterns provides a benchmark for future expansions.

1. Define Areas of High Demand: Which U.S. regions with a large elderly and high median-income population have fewer than ten pharmacies?

Finding: We then analyzed regions with a large elderly population and a high median-income population. Rogers County represents potential locations where new pharmacies can be opened to meet unmet demand.

1. Determine Optimal Business Hours: What is Walgreens's ideal business day and hour in Rogers City?

Findings: Walgreens in Roger City experiences peak traffic on Fridays during the afternoon and early evening, suggesting the best business hours to ensure profitability. Knowing the busiest times helps determine optimal hours of operation for future stores.

**Reference**

Reuter, D. (n.d.). *Meet the typical CVS shopper: A white, gen X, college-educated city dweller earning a high income*. Business Insider. https://www.businessinsider.com/typical-cvs-shopper-demographic-urban-genx-earning-high-income-2021-9

**A. Brand-Specific Questions (**looking for potential brand**)**

(Calculate the total visit counts for each brand in a specific area divided by the total population of that area (traffic) to understand the most popular brands in the market.)

* **What certain pharmacies in the healthcare industry has the most traffic in the past 5 years?**

-- Finding a pharmacy brand in the healthcare industry has the most traffic in the past 5 years.

SELECT p.location\_name AS pharmacy\_name,

SUM(v.raw\_visit\_counts) AS total\_visits

FROM `elemental-leaf-436616-q5.safegraph.places` AS p

JOIN `elemental-leaf-436616-q5.safegraph.visits` AS v

ON p.safegraph\_place\_id = v.safegraph\_place\_id

WHERE p.top\_category = 'Health and Personal Care Stores'

AND v.date\_range\_start >= TIMESTAMP(DATE\_SUB(CURRENT\_DATE(), INTERVAL 5 YEAR)) -- Convert to TIMESTAMP

GROUP BY p.location\_name

ORDER BY total\_visits DESC

LIMIT 1;

**B. Geographic-Specific Questions**

* **What U.S. County with an above-average and high-income population has the least pharmacy in the country?**

CREATE TEMPORARY TABLE potential\_county\_table AS

SELECT f.county,f.state\_fips,f.county\_fips,

SUM(`pop\_m\_60-61` + `pop\_m\_62-64` + `pop\_m\_65-66` +

`pop\_f\_60-61` + `pop\_f\_62-64` + `pop\_f\_65-66`) AS older\_population,

SUM(`inc\_75-100` + `inc\_100-125` + `inc\_125-150`) AS higher\_income\_population

FROM `elemental-leaf-436616-q5.safegraph.cbg\_demographics` AS d

JOIN `elemental-leaf-436616-q5.safegraph.cbg\_fips` AS f

ON SUBSTRING(d.cbg, 1, 2) = f.state\_fips

AND SUBSTRING(d.cbg, 3, 3) = f.county\_fips

GROUP BY f.county,f.state\_fips,f.county\_fips

HAVING older\_population > (

SELECT AVG(older\_population)

FROM (

SELECT SUM(`pop\_m\_60-61` + `pop\_m\_62-64` + `pop\_m\_65-66` +

`pop\_f\_60-61` + `pop\_f\_62-64` + `pop\_f\_65-66`) AS older\_population,

FROM `elemental-leaf-436616-q5.safegraph.cbg\_demographics` AS d

JOIN `elemental-leaf-436616-q5.safegraph.cbg\_fips` AS f

ON SUBSTRING(d.cbg, 1, 2) = f.state\_fips

AND SUBSTRING(d.cbg, 3, 3) = f.county\_fips

GROUP BY f.county

) AS subquery)

AND higher\_income\_population> (

SELECT AVG(income)

FROM (

SELECT SUM(`inc\_75-100` + `inc\_100-125` + `inc\_125-150`) AS income,

FROM `elemental-leaf-436616-q5.safegraph.cbg\_demographics` AS d

JOIN `elemental-leaf-436616-q5.safegraph.cbg\_fips` AS f

ON SUBSTRING(d.cbg, 1, 2) = f.state\_fips

AND SUBSTRING(d.cbg, 3, 3) = f.county\_fips

GROUP BY f.county

) AS sub);

SELECT county, COUNT(v.safegraph\_place\_id) pharmacy\_num

FROM potential\_county\_table pc

JOIN `elemental-leaf-436616-q5.safegraph.visits` v

ON SUBSTRING(v.poi\_cbg, 1, 2) = pc.state\_fips

AND SUBSTRING(v.poi\_cbg, 3, 3) = pc.county\_fips

JOIN `elemental-leaf-436616-q5.safegraph.places` p

ON v.safegraph\_place\_id = p.safegraph\_place\_id

WHERE p.top\_category = 'Health and Personal Care Stores'

GROUP BY county

ORDER BY pharmacy\_num;

**C. Business hours (days of the week & opening times)**

* **When is the most ideal business day and hour for this store?**

-- Finding the busiest day and hour for Walgreens as the ideal business day and hour.

SELECT p.location\_name,

p.top\_category,

p.city,

p.region,

v.popularity\_by\_day AS busiest\_day, -- Busiest day

v.popularity\_by\_hour AS busiest\_hour -- Busiest hour

FROM `elemental-leaf-436616-q5.safegraph.visits` AS v

JOIN `elemental-leaf-436616-q5.safegraph.places` AS p

ON v.safegraph\_place\_id = p.safegraph\_place\_id

WHERE p.city = 'Rogers'

AND p.top\_category = 'Health and Personal Care Stores'

AND p.location\_name = 'Walgreens'

ORDER BY

v.popularity\_by\_day DESC, -- Order by the day popularity

v.popularity\_by\_hour DESC -- Then by hour popularity

LIMIT 1; -- Get the top result for the busiest day and hour