

Optimizing CTA Ad Placements Through Foot Traffic and Exposure Analysis

A Datathon project prepared for:

Chicago Transit Authority

April 26, 2025

Yuxin (Coral) Zhai,
Dylan Xia,
Yuhao (Jack) Wang,
Marcus Cao



Today's Agenda

I. Welcome & Introductions

II. Problem Statement

III. Key Findings

IV. Examples of Application

V. Q&A

Problem Statement

Advertising is a major non-fare revenue source for CTA.

The Chicago Transit Authority's current advertising strategy lacks a dynamic, data-driven framework to align ads with real-time commuter behavior and localized audience relevance. Key gaps include:

- 1. Under-captured Exposure from Dwell Time Variability**
- 2. Mismatch Between Traffic Volume and Ad Quality**
- 3. Inflexible Quality Score Integration**
- 4. Inefficient Resource Allocation**

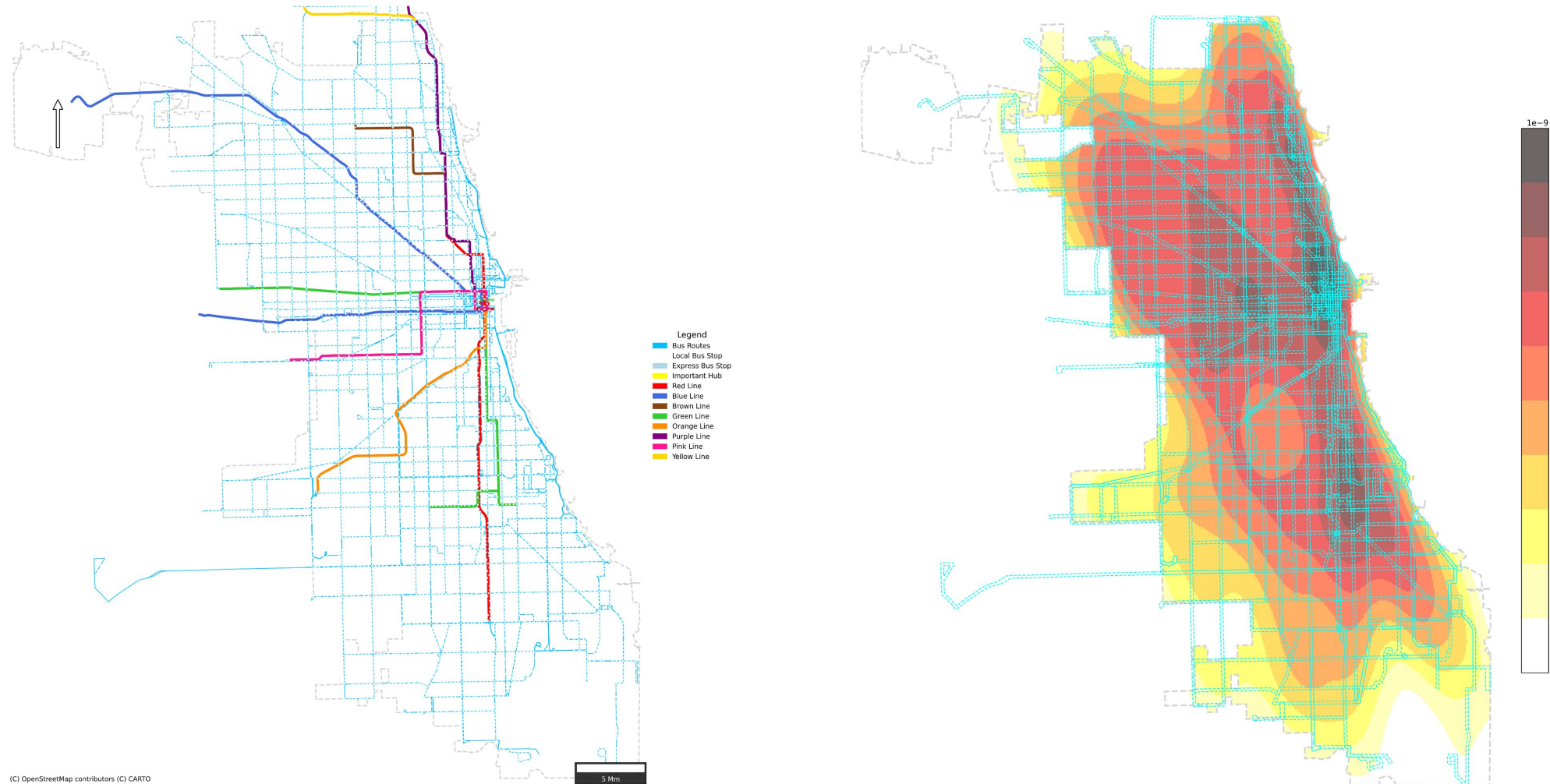
Result: Underutilized advertising spaces and lost potential revenue.

Objective:

Maximize advertising revenue by identifying high-exposure areas and recommending optimal reallocation of public transit ad placements.

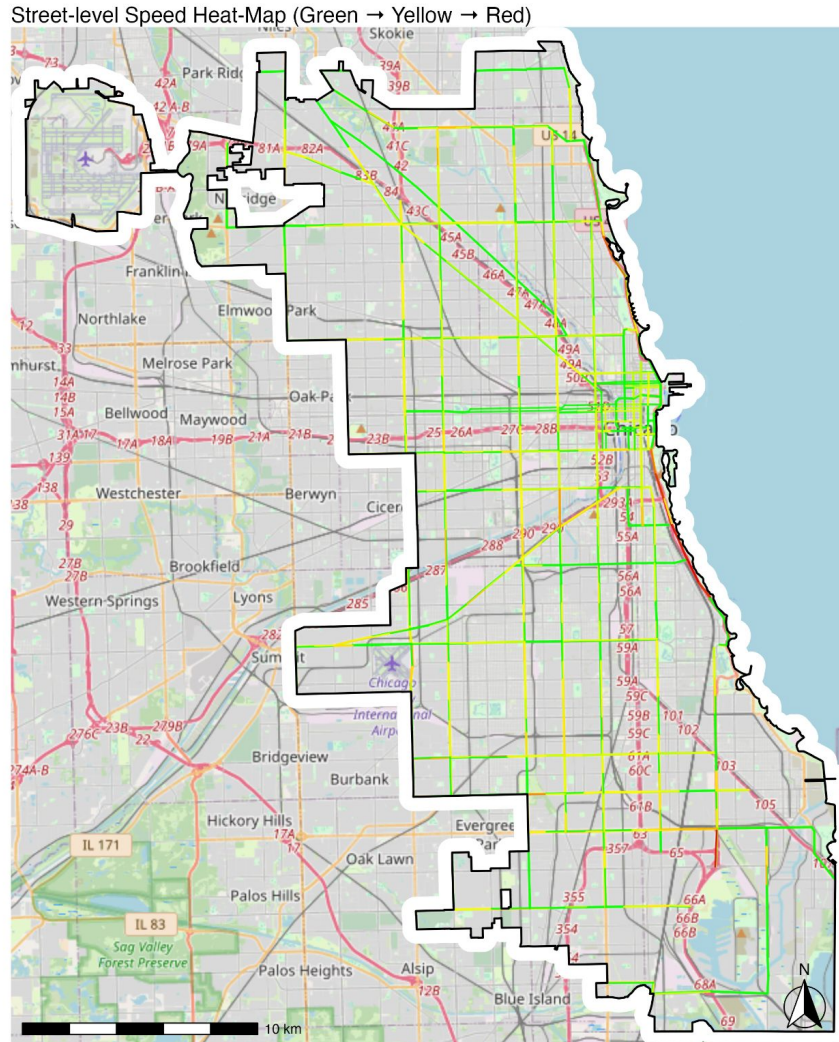


Overview - Chicago Public Transit Map and Heat Map



(C) OpenStreetMap contributors (C) CARTO

Dwell Time Analysis



Objective:

Identify areas with longer dwell times to pinpoint higher advertising exposure opportunities.

Method:

Analyzed street-level speed data; slower speeds indicate longer stops and greater rider attention.

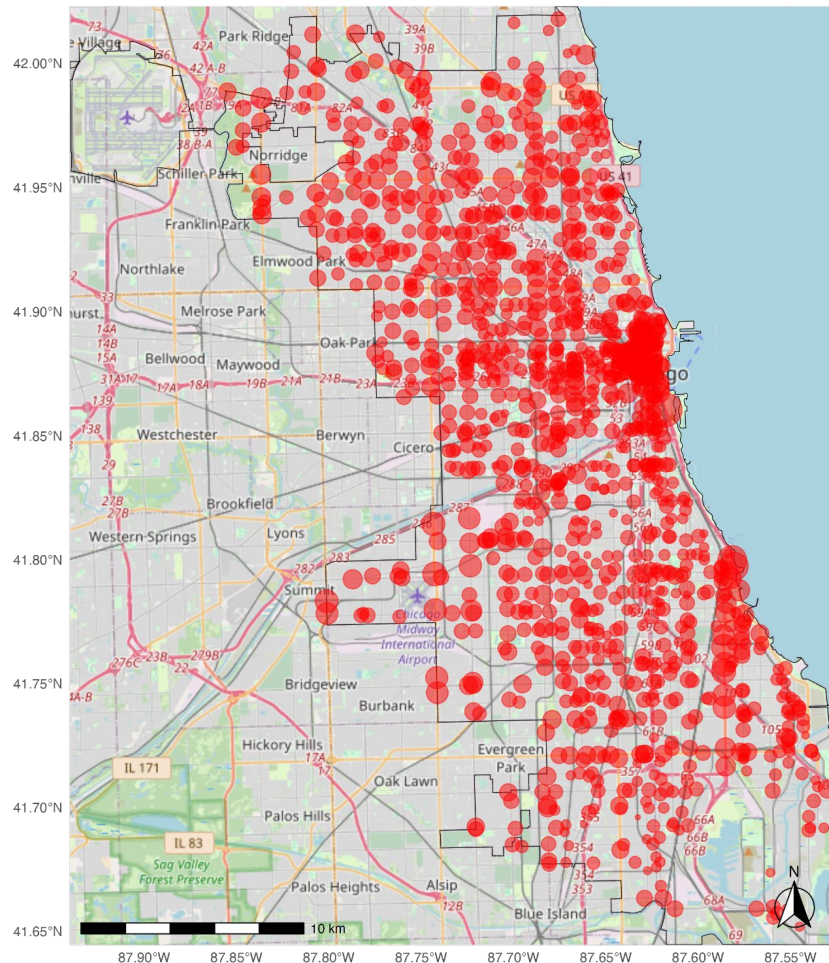
Key Insight:

- Downtown core and major transfer corridors show the longest dwell times
- High-priority areas for advertising placements

*Data Source: Chicago Data Portal

Foot Traffic Analysis

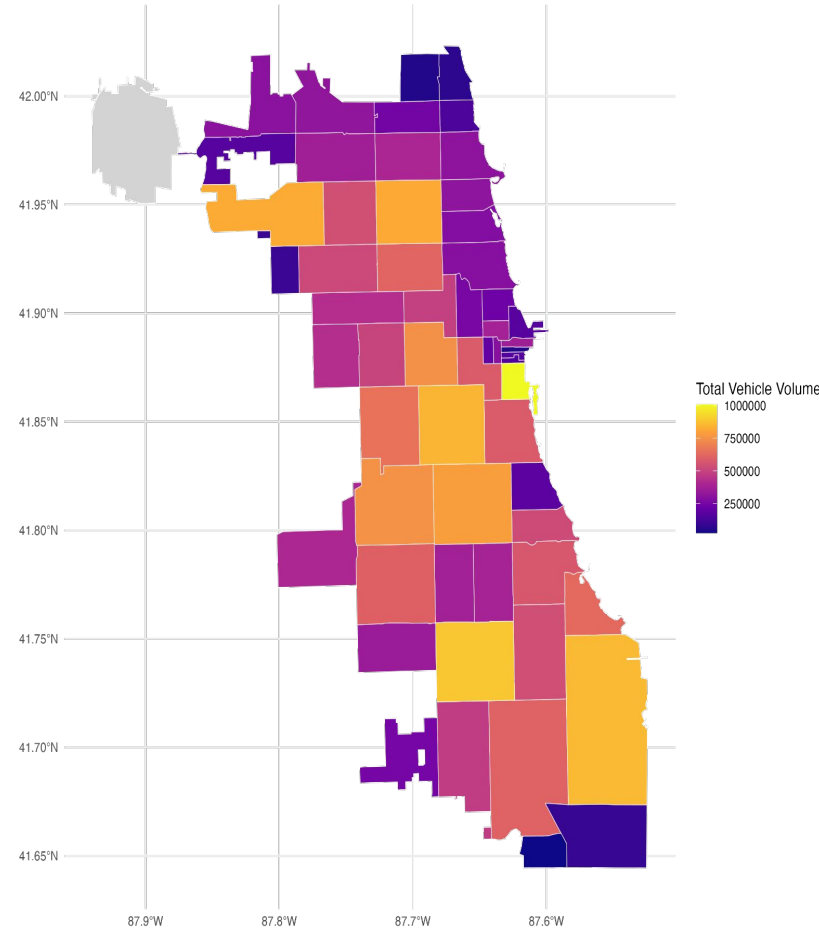
Traffic Volume Bubble Map



Vehicle Volume



Total Traffic Volume Aggregated by ZIP Code



Objective:

Analyze traffic volumes to identify areas with the highest potential rider exposure.

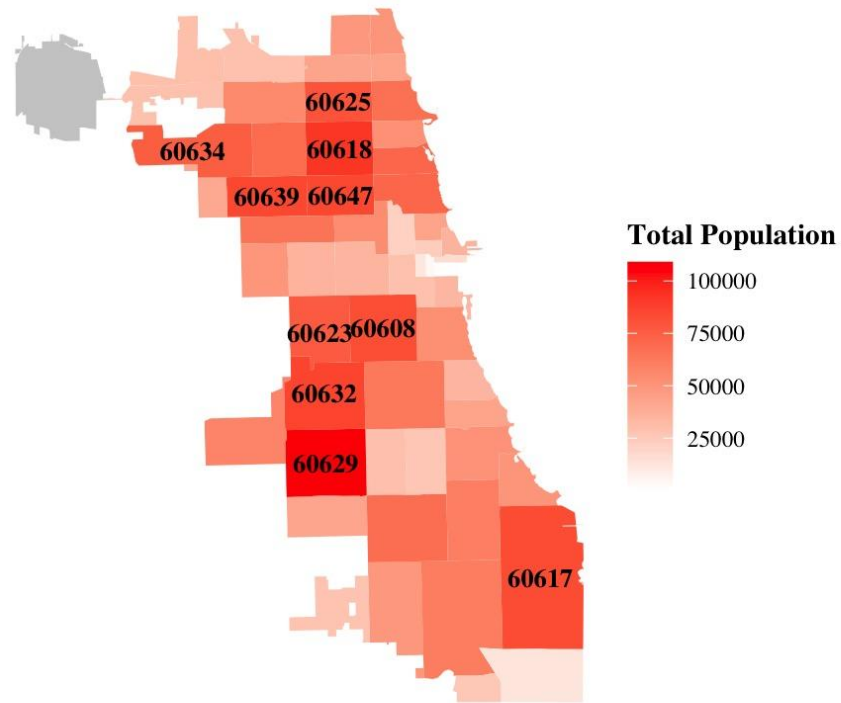
Key Insights:

- High traffic volumes are concentrated in downtown, near major highways, and along key north-south corridors.
- Areas with heavier traffic are prime locations for advertising visibility and engagement.

*Data Source: Chicago Data Portal

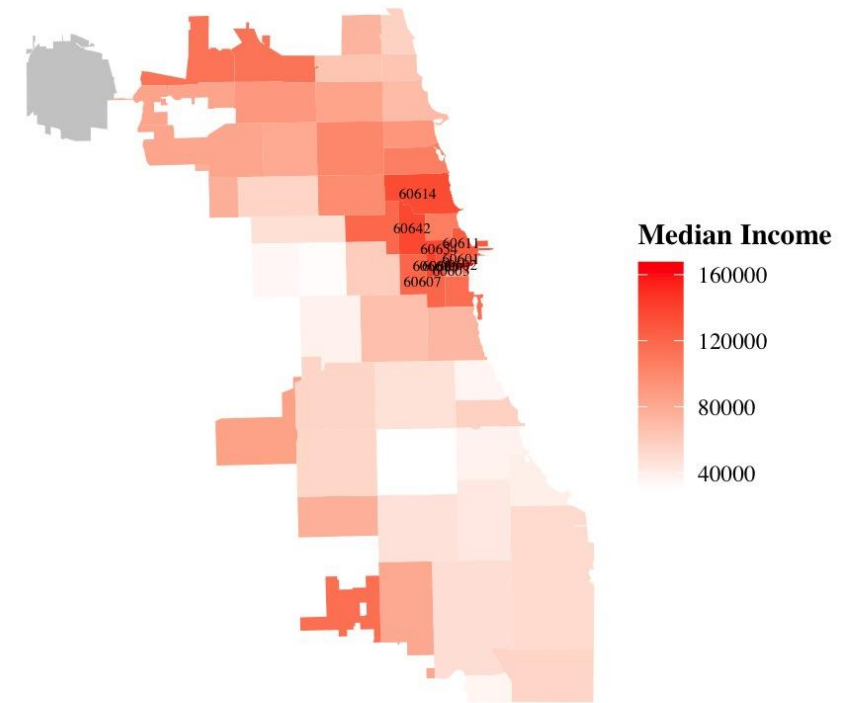
Sociodemographic

Distribution of Total Population by ZIP Code



Top 10 ZIP codes labelled

Distribution of Median Income by ZIP Code

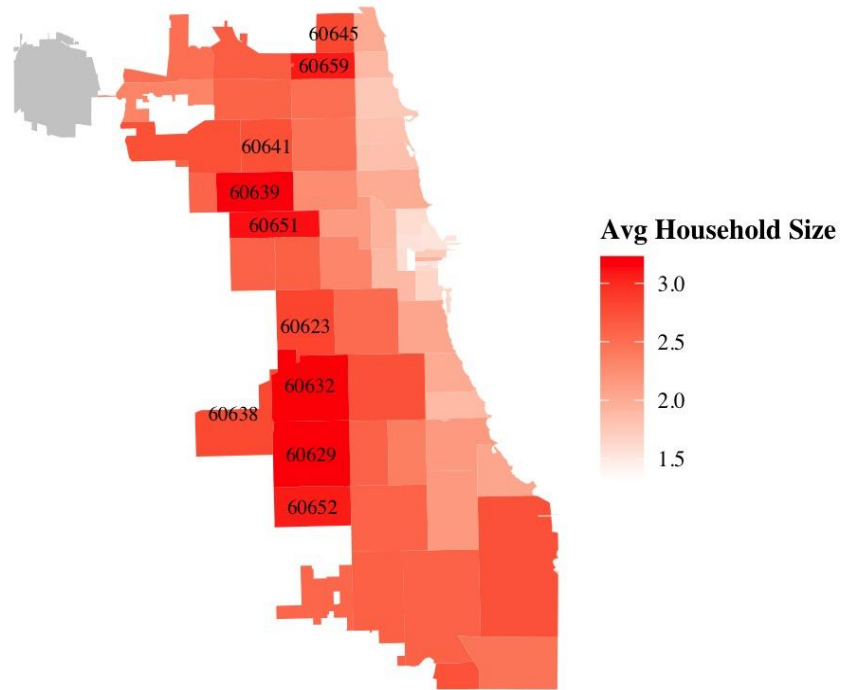


Top 10 ZIP codes labelled

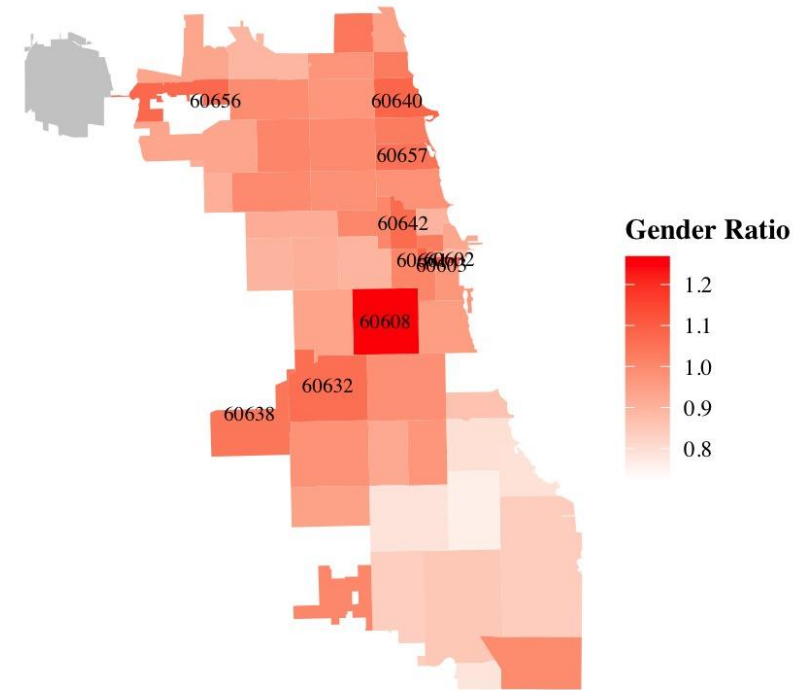
*Data Source: US Census & Chicago Data Portal

Sociodemographic

Distribution of Average Household Size by ZIP Code



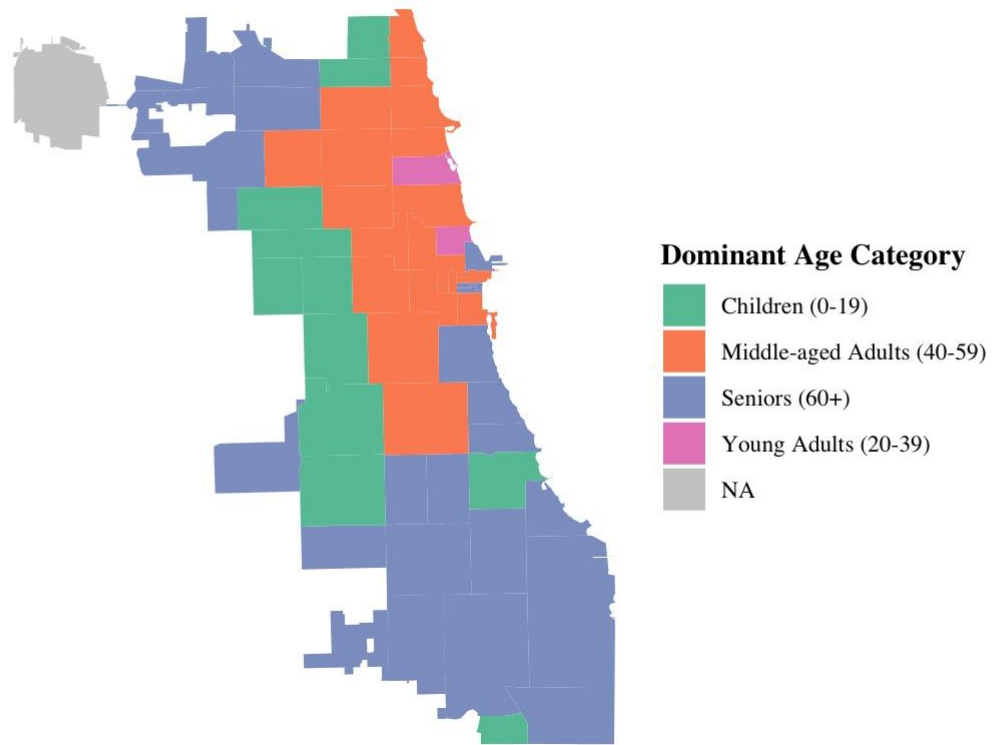
Distribution of Gender Ratio by ZIP Code



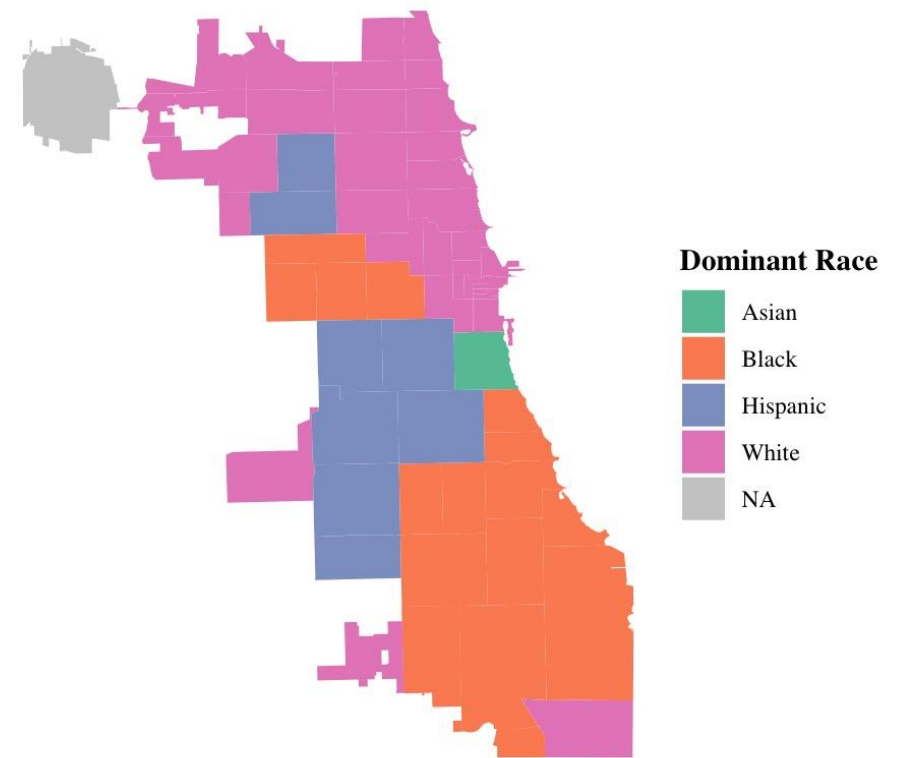
*Data Source: US Census & Chicago Data Portal

Sociodemographic

Distribution of Dominant Age Category by ZIP Code



Distribution of Dominant Race by ZIP Code



*Data Source: US Census & Chicago Data Portal

Calculation Formula for Quality Score

$$\text{Quality Score} = \alpha \times \text{Income Match Score} + \beta \times \text{Race Match Score} + \gamma \times \text{Household Match Score} + \delta \times \text{Gender Match Score} + \epsilon \times \text{Age}$$

$$\text{Income Score} = \min \left(1, \frac{\text{Median Income}}{\text{Target Income}} \right)$$

$$\text{Race Score} = \begin{cases} 1, & \text{if Dominant Race is in Preferred Race} \\ 0.5, & \text{otherwise} \end{cases}$$

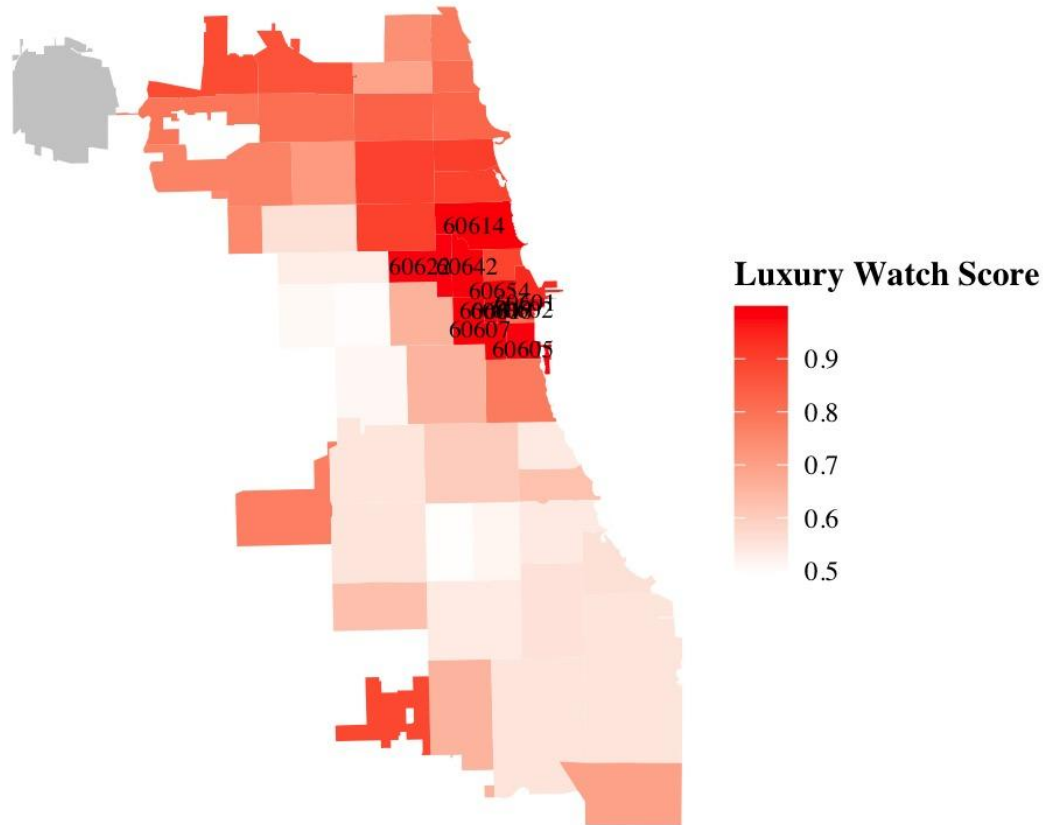
$$\text{Household Score} = \begin{cases} \min \left(1, \frac{\text{Max Household Size}}{\text{Average Household Size}} \right), & \text{if Max Household Size is specified} \\ \min \left(1, \frac{\text{Average Household Size}}{\text{Min Household Size}} \right), & \text{if Min Household Size is specified} \\ 1, & \text{otherwise} \end{cases}$$

$$\text{Gender Score} = 1 - |\text{Gender Ratio} - \text{Target Gender Ratio}| \quad \text{Age Score} = \begin{cases} 1, & \text{if Dominant Age Category matches Target Age Category} \\ 0.5, & \text{otherwise} \end{cases}$$

Application Example: Luxury Watch

$$[\text{Quality Score} = 0.4 \times \text{Income Match Score} + 0.2 \times \text{Race Match Score} + 0.2 \times \text{Household Match Score} + 0.1 \times \text{Gender Match Score} + 0.1 \times \text{Age}]$$

Distribution of Luxury Watch Score by ZIP Code



Top 10 ZIP codes labelled

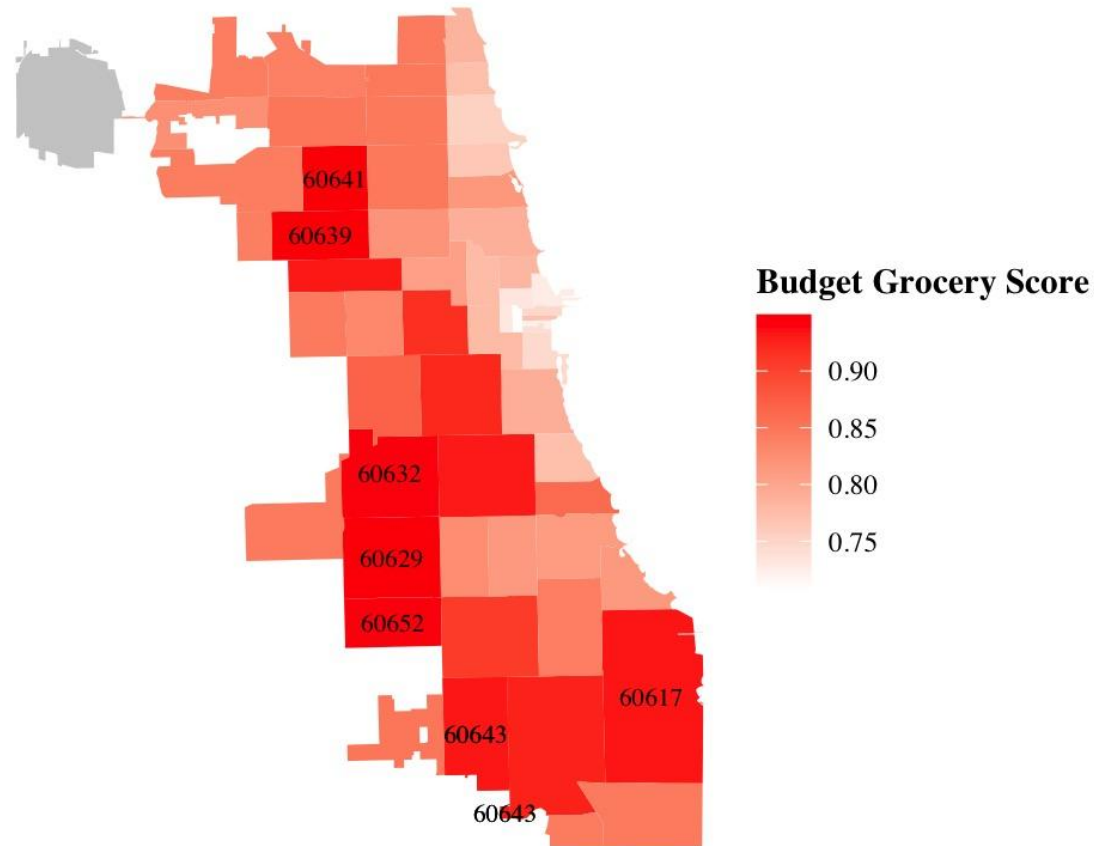
Weights explanation:

- **Income (0.4):** High disposable income is critical for luxury products.
- **Race (0.2):** Targeted marketing based on race demographics.
- **Household (0.2):** Smaller households often correlate with higher disposable income.
- **Gender (0.1):** Slight male preference, but gender balance is still important.
- **Age (0.1):** Middle-aged adults are primary buyers of luxury watches.

Application Example: Budget Grocery

Quality Score = $0.3 \times \text{Income Match Score} + 0.2 \times \text{Race Match Score} + 0.3 \times \text{Household Match Score} + 0.1 \times \text{Gender Match Score} + 0.1 \times \text{Age}$

Distribution of Budget Grocery Score by ZIP Code



Top 10 ZIP codes labelled

Weights explanation:

- **Income (0.3):** Affordability is key for budget products.
- **Race (0.2):** Cultural targeting for community-based marketing.
- **Household (0.3):** Larger households tend to be more price-sensitive.
- **Gender (0.1):** Balanced gender targeting for grocery products.
- **Age (0.1):** Young adults tend to be frequent grocery shoppers.

Policy Recommendation

- Reallocate advertising placements based on quality scores — combining foot traffic, dwell time, and demographics — to maximize the marginal revenue product (MRP) of available advertising spaces.
- This strategy employs principles of constrained resource allocation and profit maximization, ensuring that scarce advertising inventory is deployed where it yields the greatest economic return.
- By aligning advertising inventory with its true marginal value, this new proposed method could assist CTA to unlock previously unrealized revenue streams without raising operational costs or fares.

Questions?

Thank you

