Part -2: Experiment and Metric Design
Toll Reimbursement to Encourage Cross-City Driving

Problem Statement

Gotham and Metropolis have complementary demand patterns: Gotham is busier at night on weekdays, while Metropolis is busier during the day. Although the cities are connected by a bridge, a two-way toll discourages drivers from crossing, causing drivers to remain in a single city and leading to supply mismatches during peak hours.

The proposed solution is to reimburse toll costs to reduce this friction and encourage drivers to operate in both cities.

Key Measure of Success

Primary Metric: Weekly Cross-City Driver Percentage

The percentage of active drivers in a given week who complete at least one trip in both Gotham and Metropolis.

Why this metric

- It directly measures the desired behavior change: drivers serving both cities.
- A weekly window smooths daily volatility and reflects sustained behavior rather than one-off trips.
- It aligns with operational decision-making, which focuses on long-term policy impact.

Secondary Metrics (Guardrails)

- Driver availability during peak hours (weekday daytime in Metropolis, weekday nighttime in Gotham)
- Rider pickup times and trip completion rates
- Driver earnings per hour and utilization
- Total toll reimbursement cost and cost per incremental cross-city driver

Experiment Design

a) Implementation

- Unit of randomization: Driver
- Treatment group: Drivers receive toll reimbursement during a defined daily window (e.g., a 6-hour weekday period aligned with demand imbalance).
- Control group: Drivers receive no toll reimbursement.
- Eligibility: Active drivers with recent trip history and proximity to the bridge.
- Duration: 3–4 weeks to capture weekday and weekend patterns.
- Reimbursement mechanism: Automatic reimbursement based on verified bridge crossings associated with trips.

b) Statistical Evaluation

We evaluate the treatment effect using a two-proportion z-test for transparency and a logistic regression model for robustness, combining interpretability with statistical rigor.

- Two-Proportion Z-Test
  - Objective: Test whether the proportion of drivers serving both cities differs between treatment and control groups.
  - Setup: Each driver is classified weekly as having served both cities or not. Proportions are computed separately for treatment and control groups.
  - Example: If 65 of 500 treated drivers (13.0%) and 40 of 500 control drivers (8.0%) served both cities, the two-proportion z-test evaluates whether this difference is statistically significant.
  - Hypotheses:
    - H_0: The cross-city driving rate is equal across groups.
    - H_1: The cross-city driving rate is higher for treated drivers.
  - Inference: A p-value below a predefined significance level (e.g., 0.05) indicates a statistically significant treatment effect.
  - Limitation: Does not account for heterogeneity in baseline driver behavior or repeated observations over time.
- Logistic Regression
  - Objective: Estimate the causal effect of toll reimbursement on the likelihood of cross-city driving while controlling for confounders.
  - Outcome Variable: Binary indicator equal to 1 if a driver served both cities in a given week, and 0 otherwise.
  - Primary Predictor: Treatment assignment (toll reimbursement indicator).
  - Controls: Baseline trip volume, typical city of operation, weekday share, and week fixed effects.
  - Example Result: An estimated odds ratio of 1.6 with a 95% confidence interval of [1.15, 2.20] implies that treated drivers are approximately 60% more likely to serve both cities than control drivers.
  - Inference: Statistical significance is established if the confidence interval excludes 1, providing stronger evidence of a treatment effect than the unadjusted proportion test.

c) Interpretation and Recommendations

If results show a meaningful and statistically significant increase:

- Recommend expanding toll reimbursement, potentially limited to specific high-impact time windows.
- Monitor long-term cost efficiency and marketplace balance.

If results show little or no impact:

- Conclude that toll cost alone is not the main barrier.
- Consider complementary incentives such as guaranteed earnings or improved demand visibility.

If negative side effects occur:

- Re-scope the program to avoid supply imbalances or excessive incentive costs.

Caveats and Risks

- Drivers may change behavior temporarily without sustained adoption.
- Information spillover between drivers could reduce experimental separation.
- External factors (weather, events, seasonality) may influence results, reinforcing the need for a control group.