



# uber

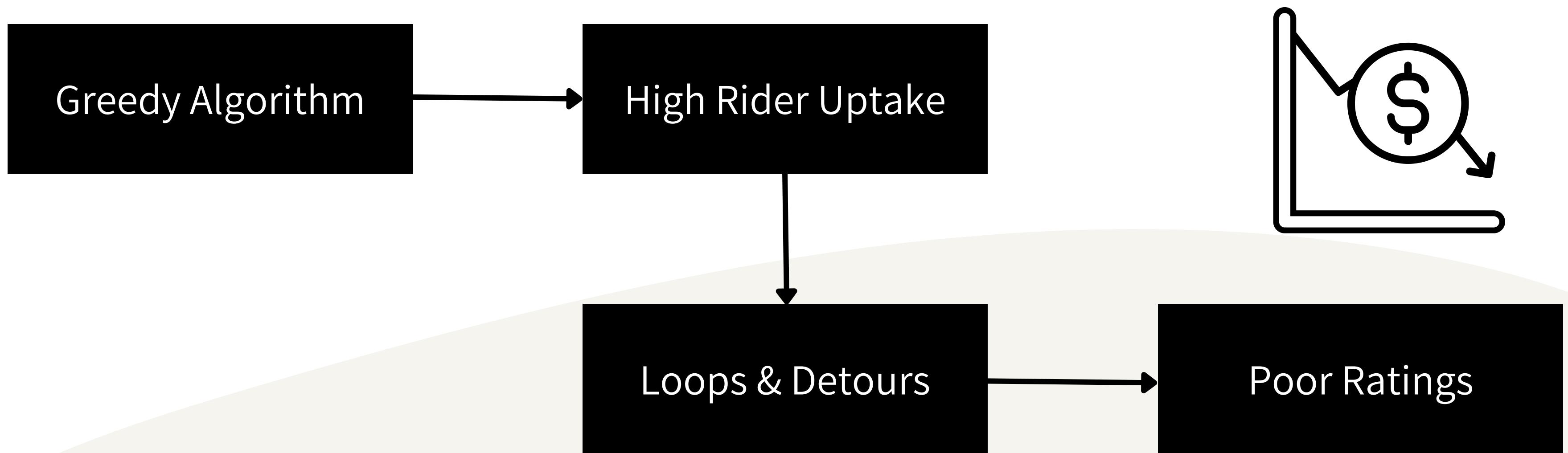
+

Purple Team 4  
Kelly, Peifen, Rish, Shalaka & Tony

# Year: 2017

# Uber Pool

Pooling in Peril: Profits Plunge as 2016 Draws to a Close

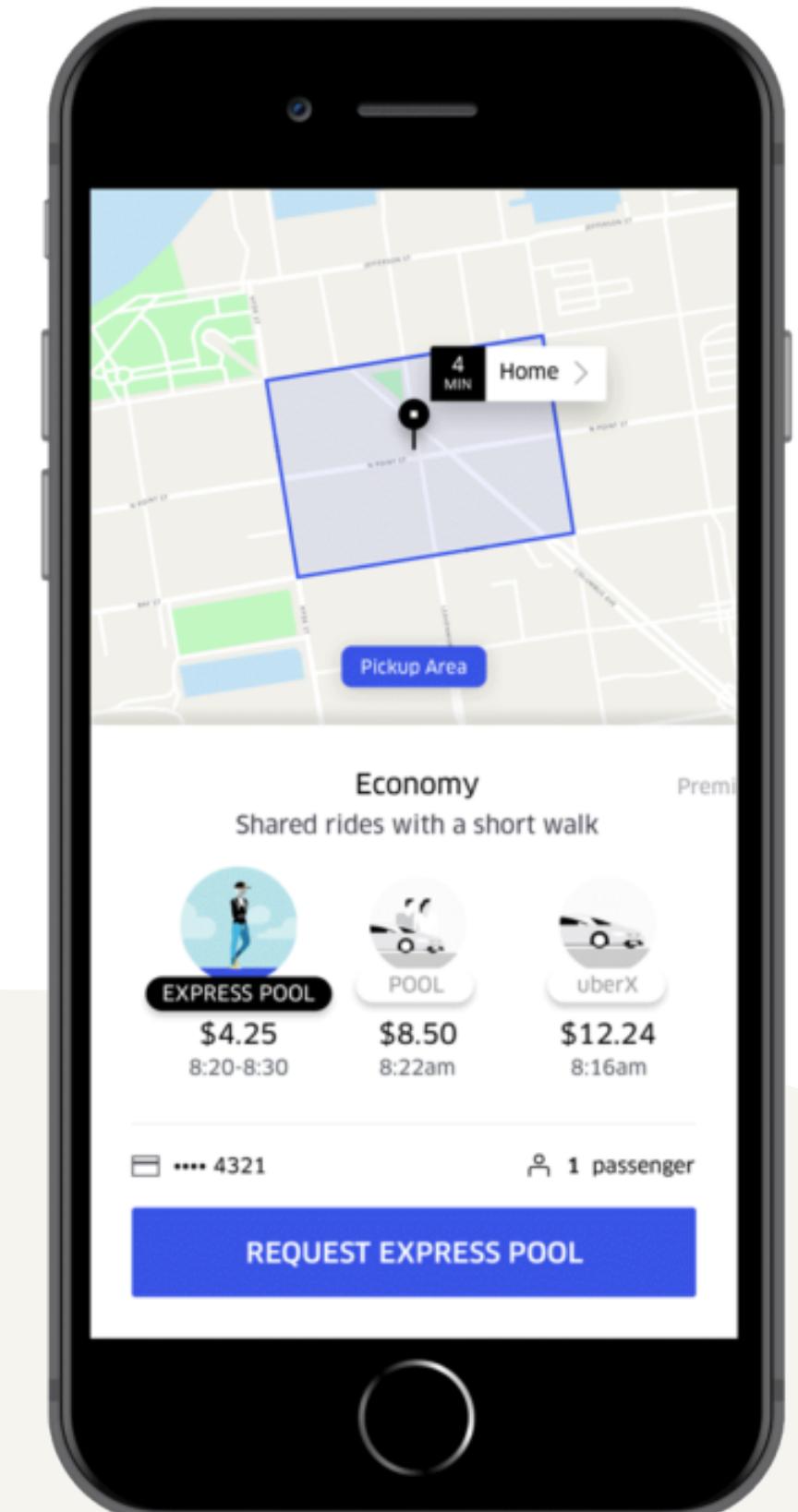


# Solution

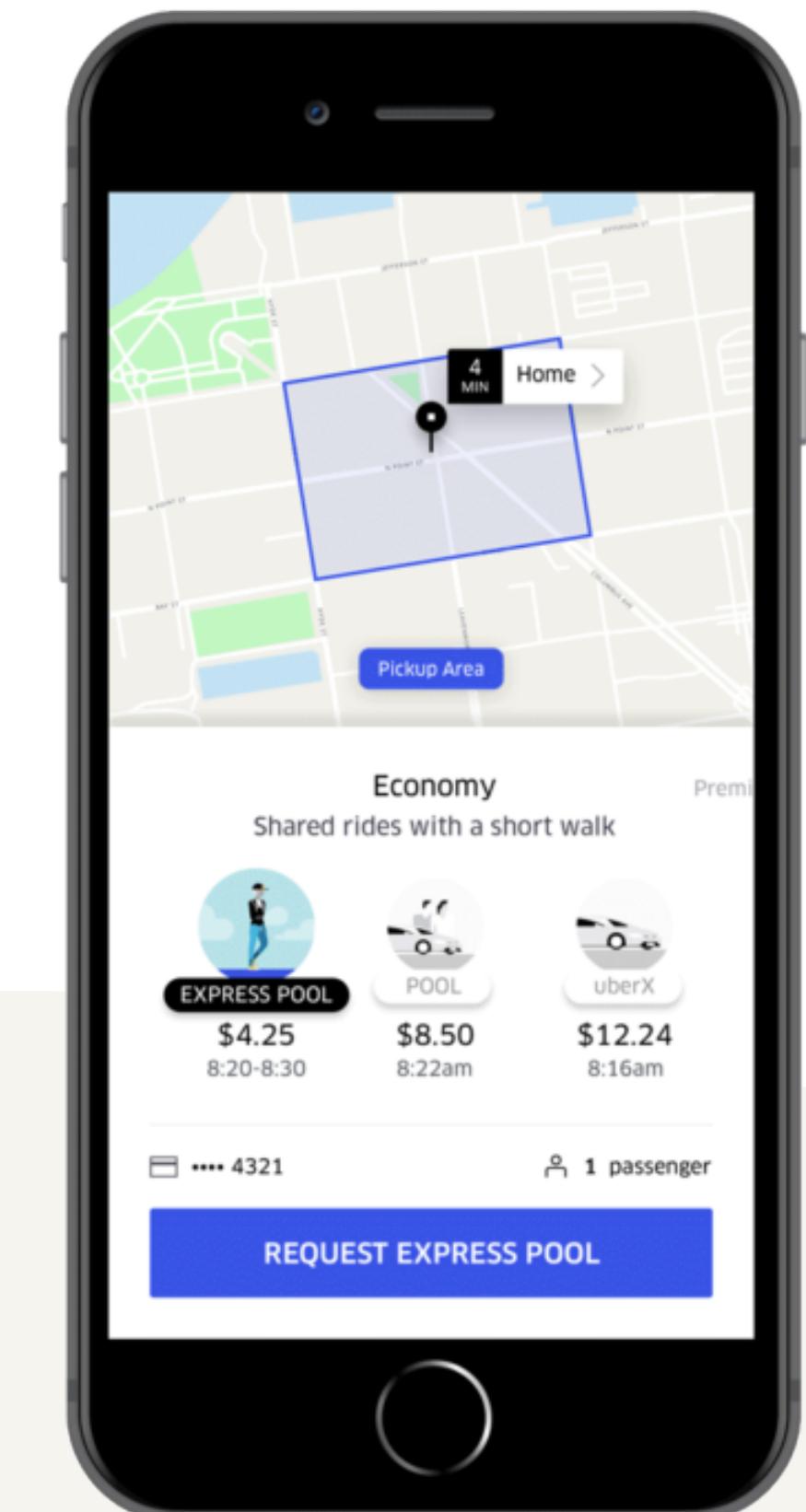
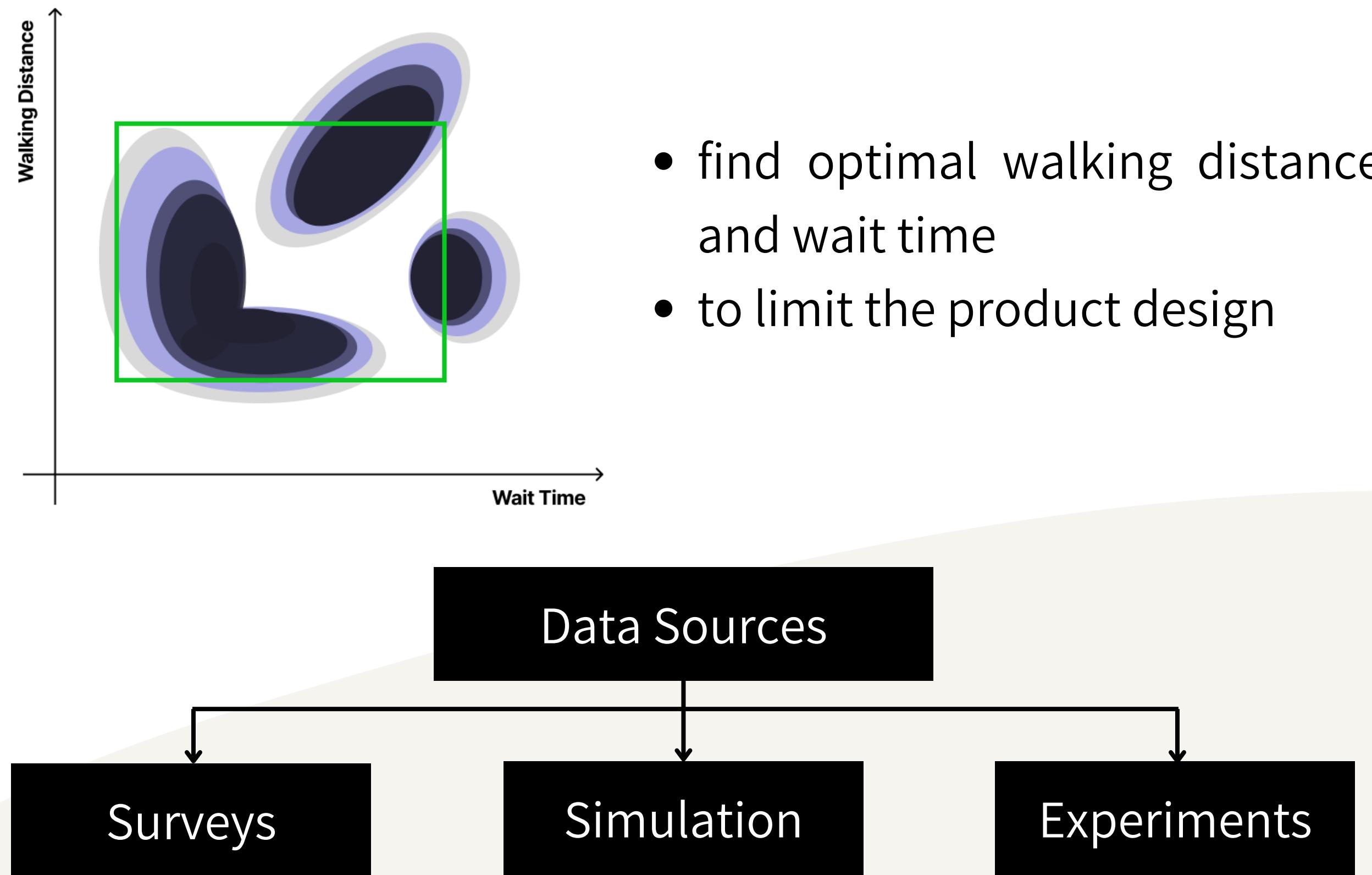
- Introduce a **two-minute waiting period** for batch matching of co-riders.
- Encourage riders to **walk short distances** to and from their pick-up and drop-off points.

New Product

Uber Express Pool = CarPool + Walk & Wait



# Our Approach



# Assumptions

Rush Hour

- Express over Pool
- Lower trip cancellations
- Lower Cost Per Trip Expected
- Increased matches

Not Rush Hour

- Pool over Express
- Lower trip cancellations
- Higher Cost Per Trip Expected
- Reduced matches

Lower Wait Time

- Express over Pool
- Higher trip cancellations
- Lower Cost Per Trip Expected
- Increased matches

Higher Wait Time

- Pool over Express
- Higher trip cancellations
- Higher Cost Per Trip Expected
- Reduced matches

# Metrics

## Customer Experience Metrics



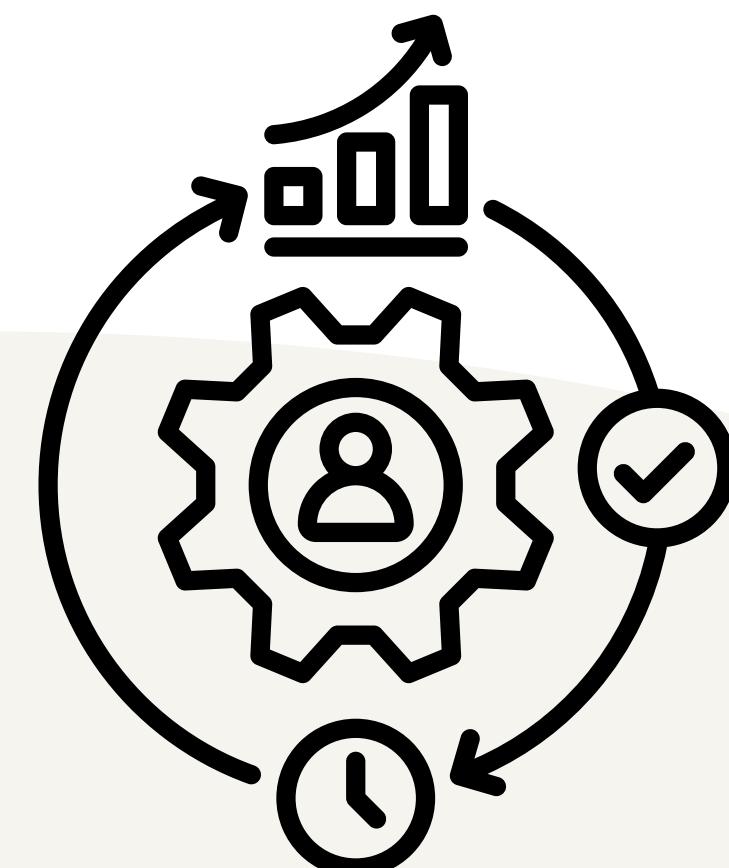
Count of Rider Cancellations

Count of Express Trips

Cost Per Trip

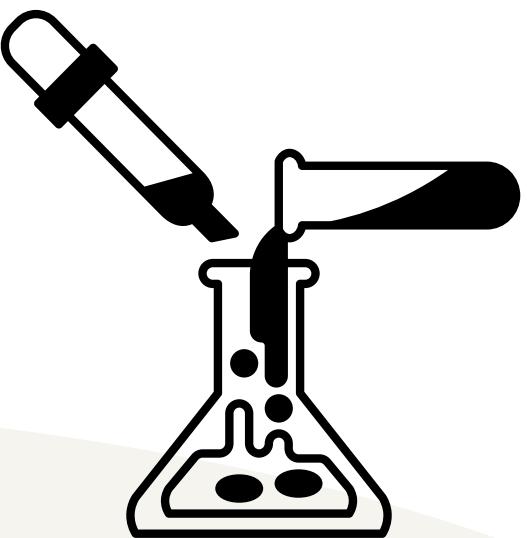
Count of Matches

## Product Efficiency Metrics



# The Experiment

- **Switchback** Experiment to assess the impact of wait time variations on user experience and efficiency.
- The experiment involved altering wait times after a set period:
  - **Control group:** Riders waited up to two minutes.
  - **Treatment group:** Riders waited up to five minutes.
- Data was collected after two weeks to assess experiment outcomes.



**Analysis Focus:** How the increase in wait times from 2 to 5 mins affected user satisfaction, trip completion rates, and overall efficiency of the ride-sharing service.

# The Dataset

	A	B	C	D	E	F	G	H	I	J	K
1	city_id	period_start	wait_time	treat	commute	trips_pool	trips_express	rider_cancellations	total_driver_payout	total_matches	total_double_matches
2	Boston	2/19/18 07:00	2	FALSE	TRUE	1415	3245	256	34458.41163	3372	1476
3	Boston	2/19/18 09:40	5	TRUE	FALSE	1461	2363	203	29764.34982	2288	1275
4	Boston	2/19/18 12:20	2	FALSE	FALSE	1362	2184	118	27437.36736	2283	962
5	Boston	2/19/18 15:00	5	TRUE	TRUE	1984	3584	355	44995.45299	4035	2021
6	Boston	2/19/18 17:40	2	FALSE	FALSE	1371	2580	181	27583.9553	2200	979
7	Boston	2/19/18 20:20	5	TRUE	FALSE	1401	2022	135	23888.11085	2066	1062

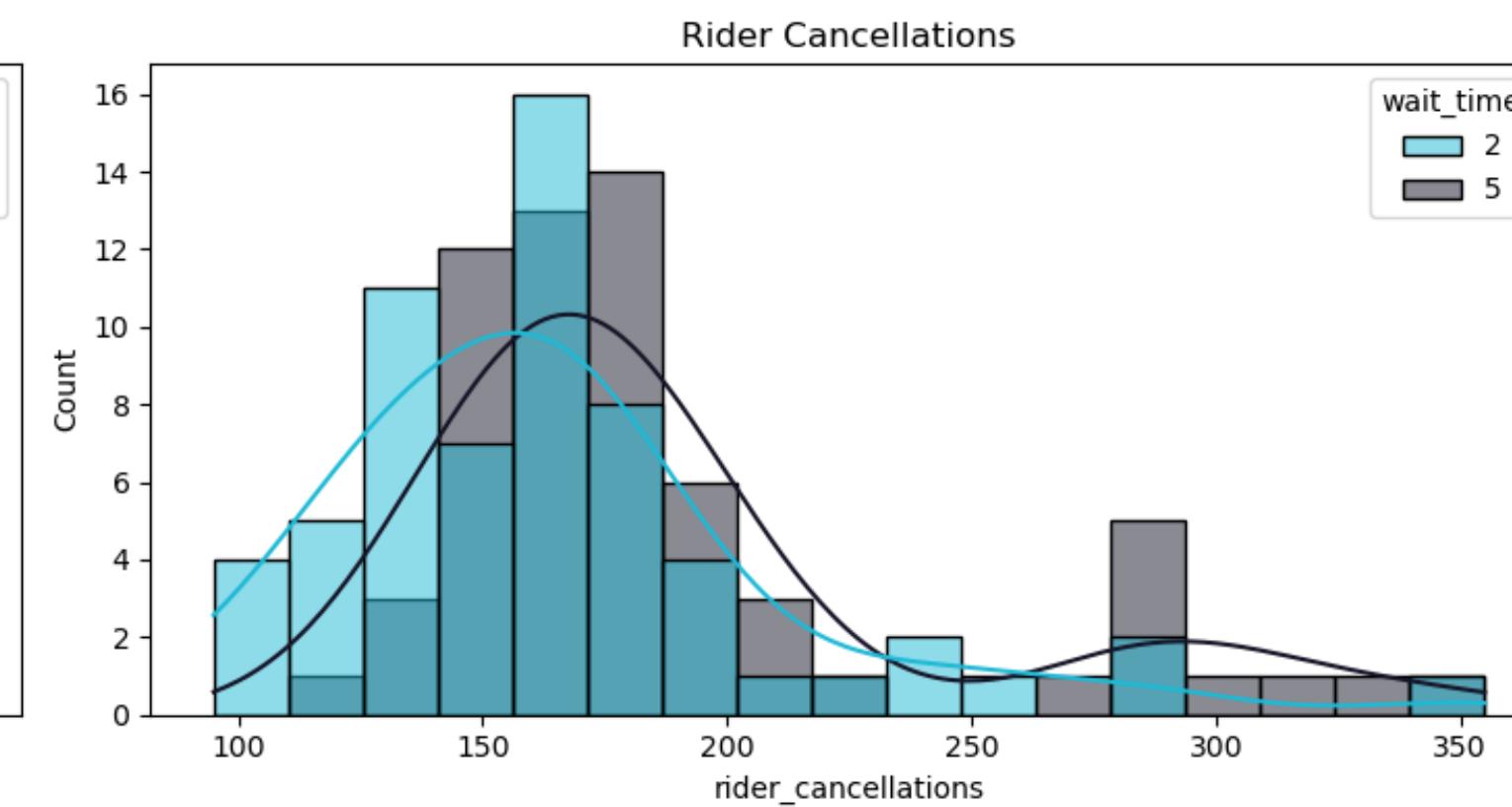
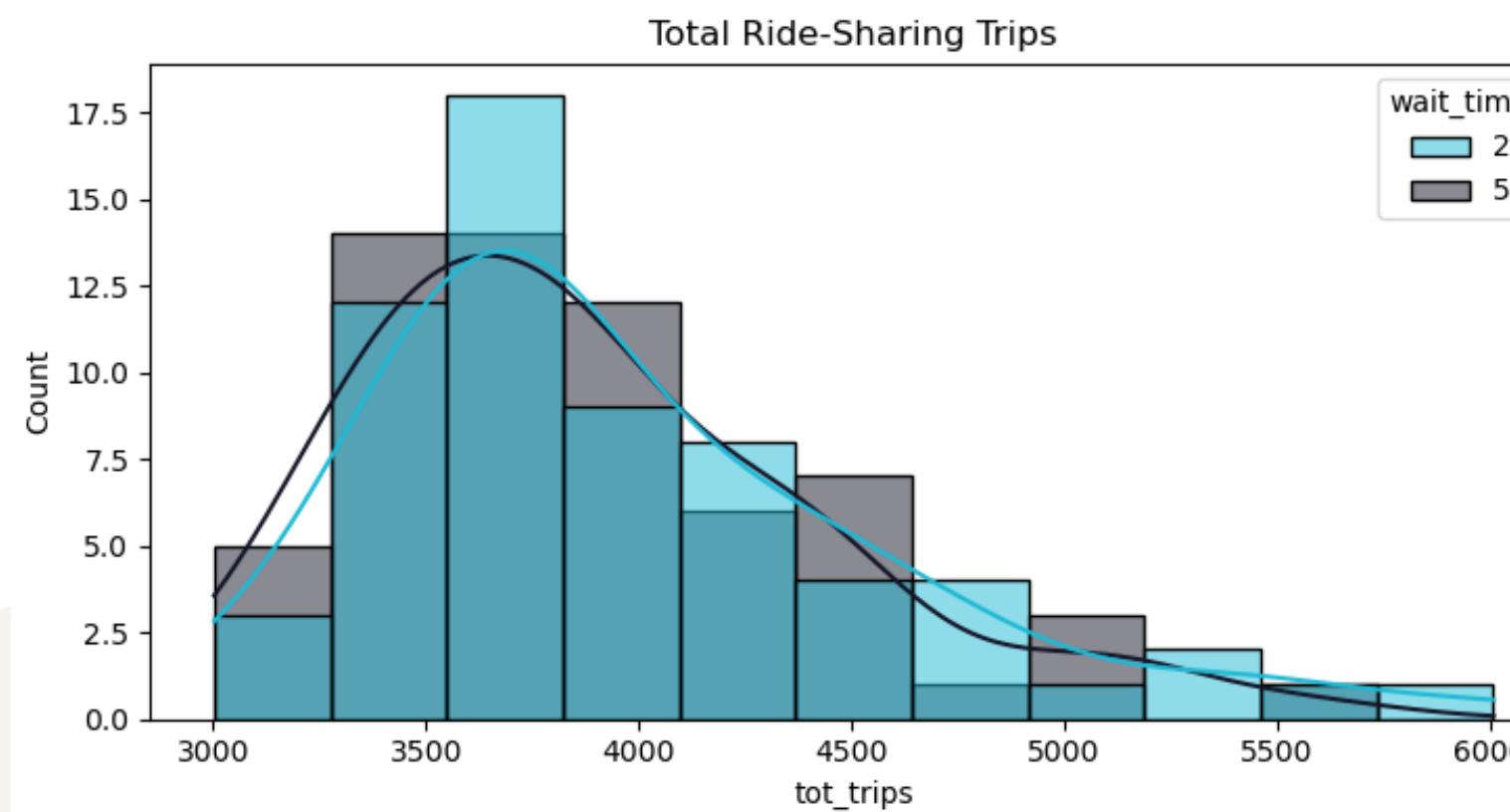
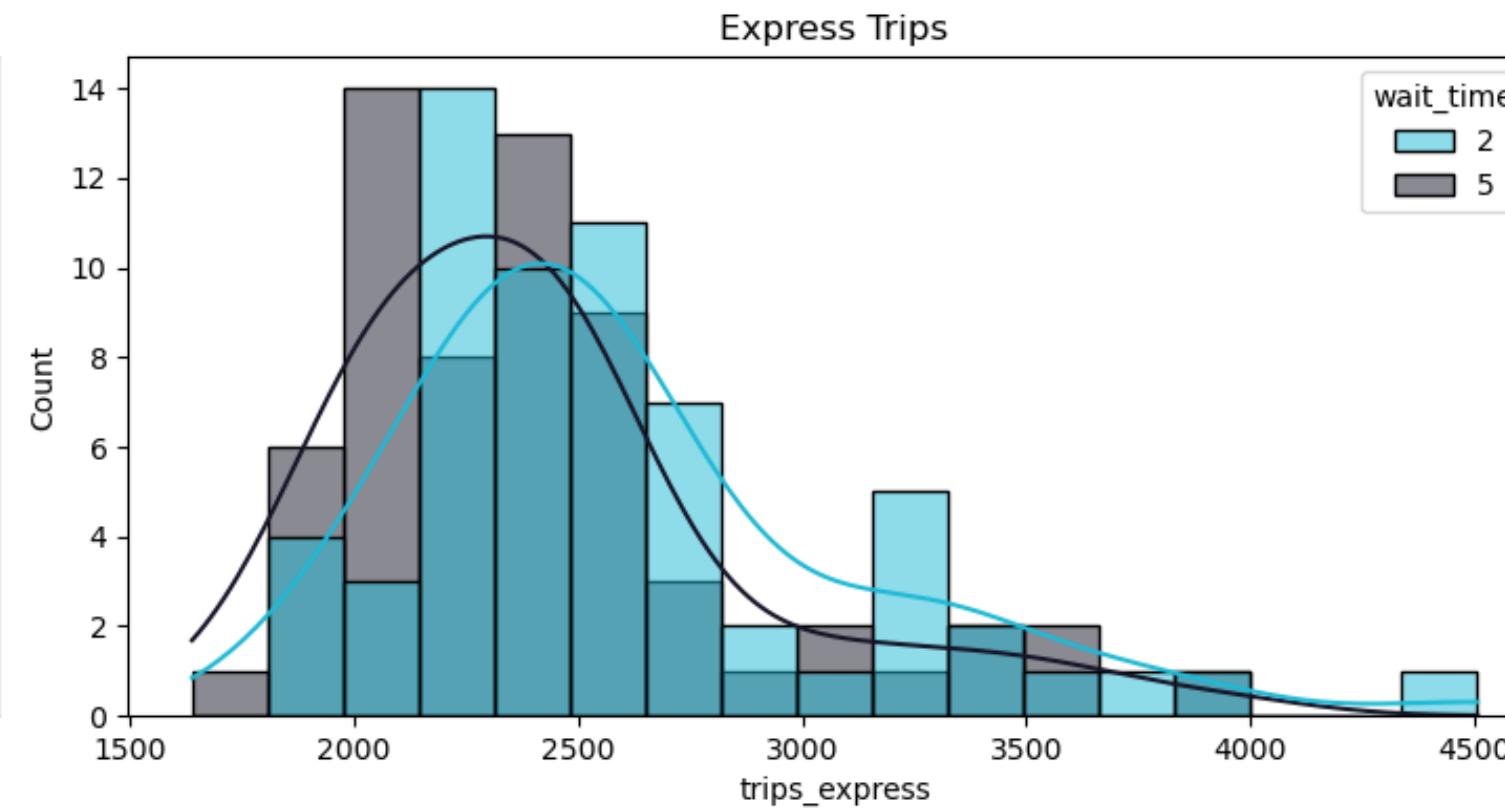
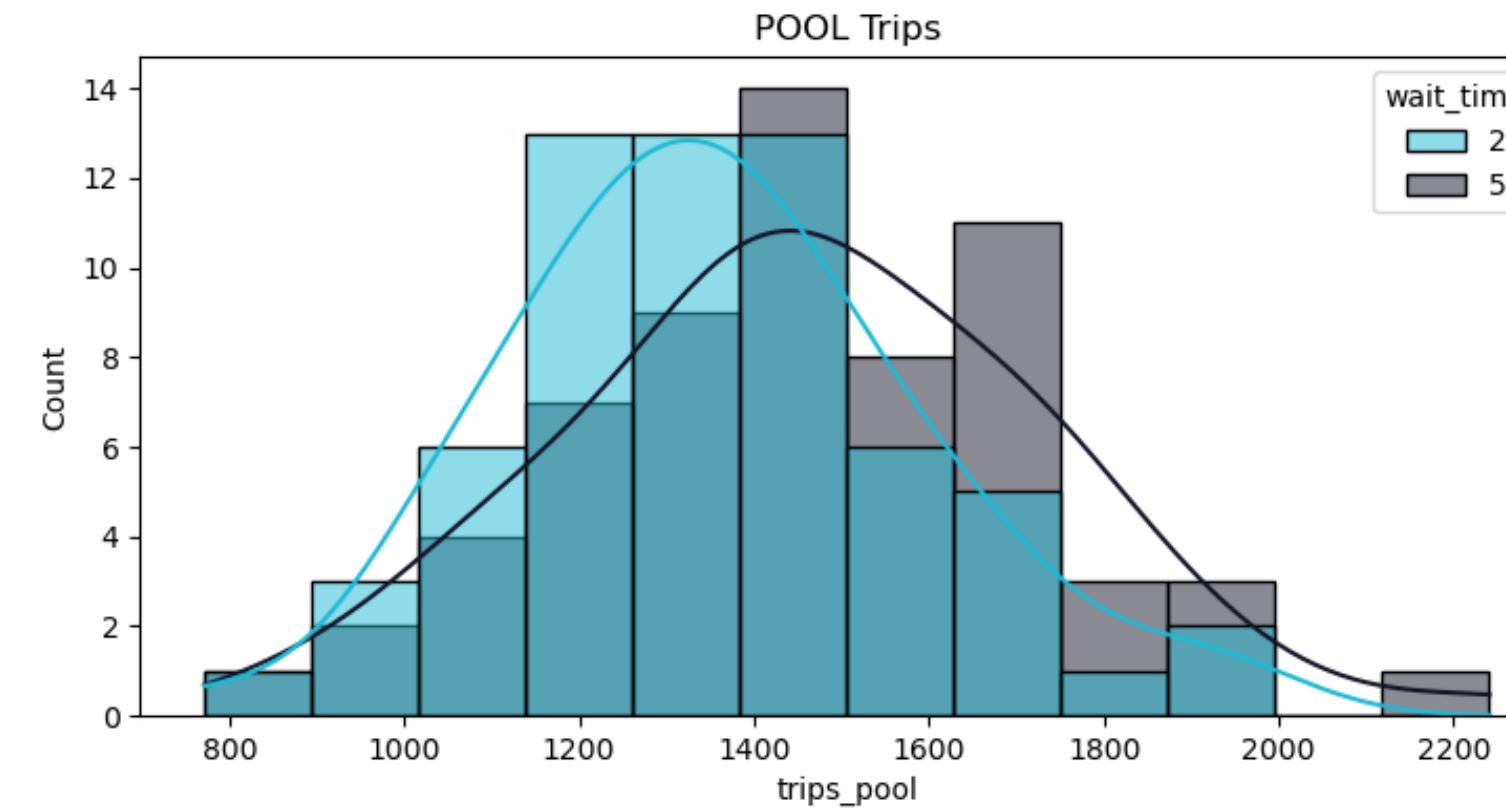
- **trips\_pool**: Total completed POOL trips in the current time period.
- **trips\_express**: Total completed Express POOL trips in the current time period.
- **rider\_cancellations**: Total rider-cancelled trip requests in the current time period.
- **total\_driver\_payout**: Total dollars paid to drivers for completed trips in the current time period.
- **total\_matches**: Total completed trips paired with at least one other rider during the current time period.
- **total\_double\_matches**: Total completed trips paired with at least two other riders during the current time period.

# Data Analysis

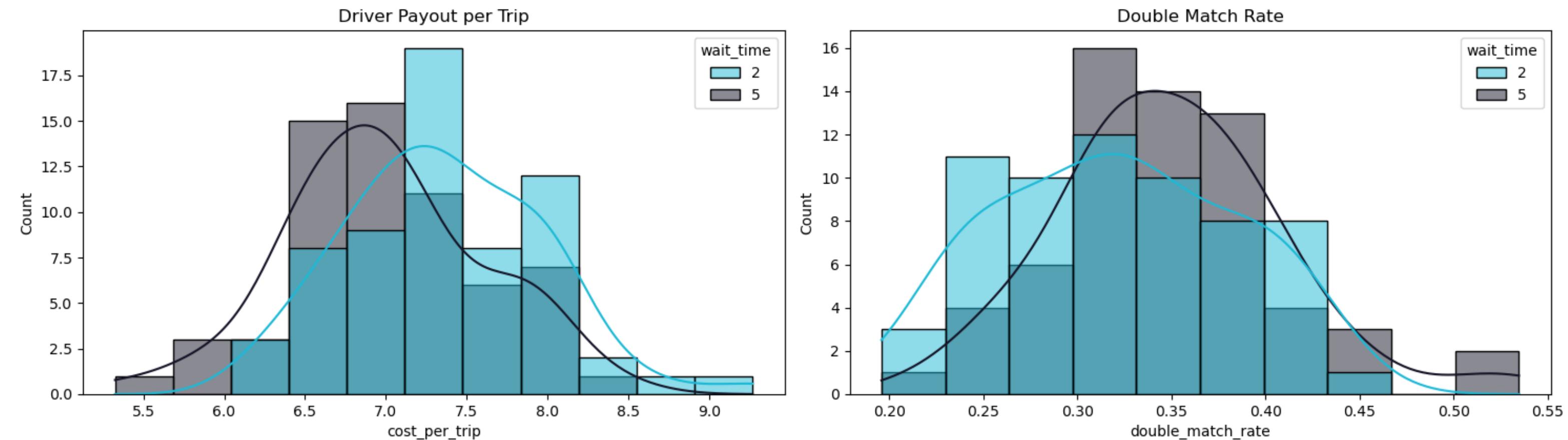
- Investigate The Distribution Of Data Points
- Verify The Underlying Assumptions
- Assess The Impact Of The Treatment On Metrics



# Data Distribution



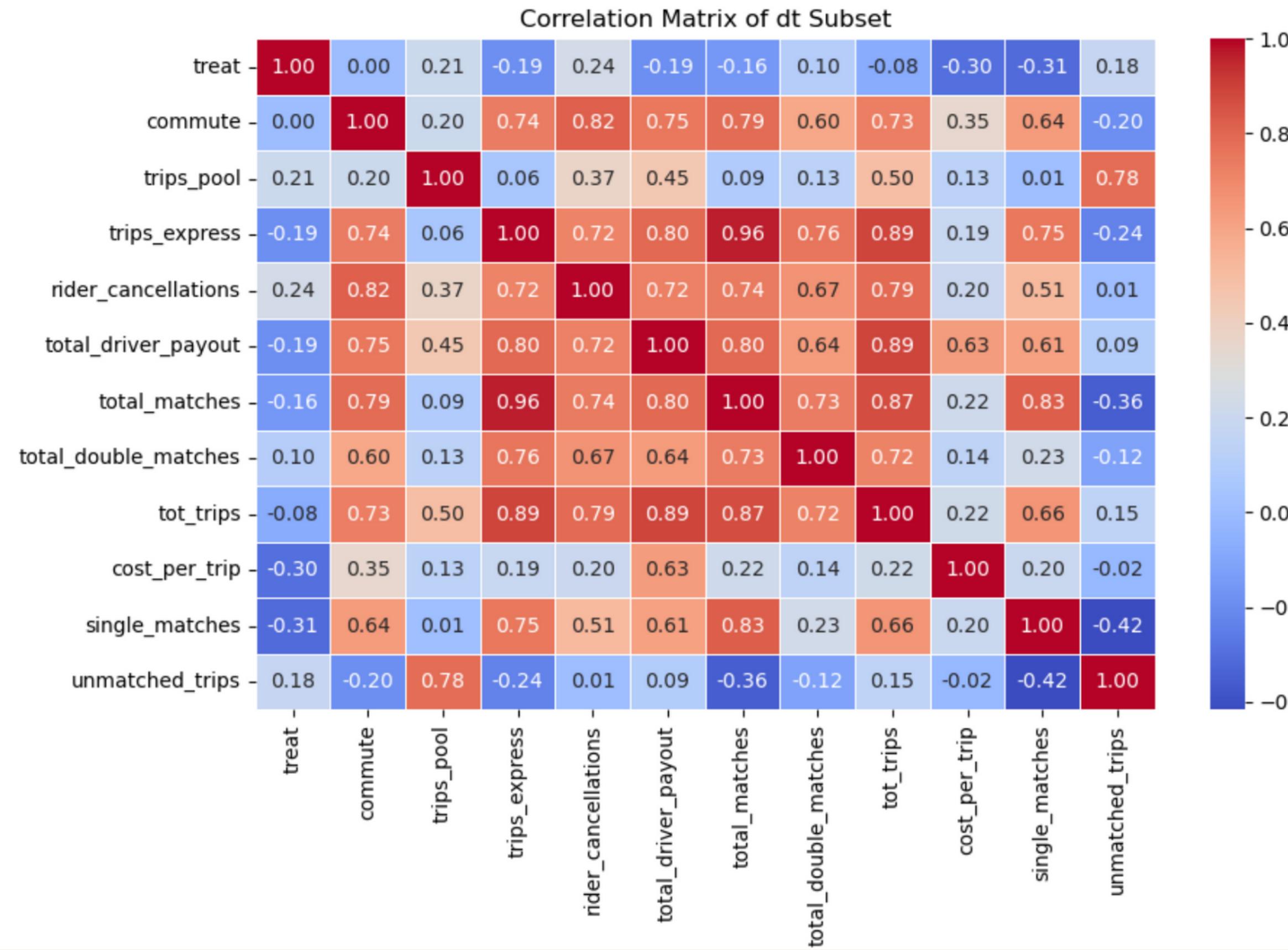
# Data Distribution



**High-level observation:** Distinct peaks and density plots highlight the differences between groups\* across all observed variables.

\*categorized by the treatment variable (wait time)

# Data Distribution



## Interesting observations:

- ‘commute’ and ‘treat’ variables have 0 correlation.
- ‘commute’ is heavily correlated with cancellations, driver payout and ride matches.

commute: TRUE Indicates rush hour periods (7-9:40AM or 3-5:40PM).

# t-tests

T-test results:

```
trips_pool: TtestResult(statistic=-2.3334843926729363, pvalue=0.021232396600015767, df=124.0)
trips_express: TtestResult(statistic=2.2015160137355108, pvalue=0.02955001543931055, df=124.0)
rider_cancellations: TtestResult(statistic=-2.7640050528783227, pvalue=0.0065808982457568384, df=124.0)
cost_per_trip: TtestResult(statistic=3.4621410571047204, pvalue=0.0007359364864662961, df=124.0)
match_rate: TtestResult(statistic=2.516928763195693, pvalue=0.013114044629773199, df=124.0)
```

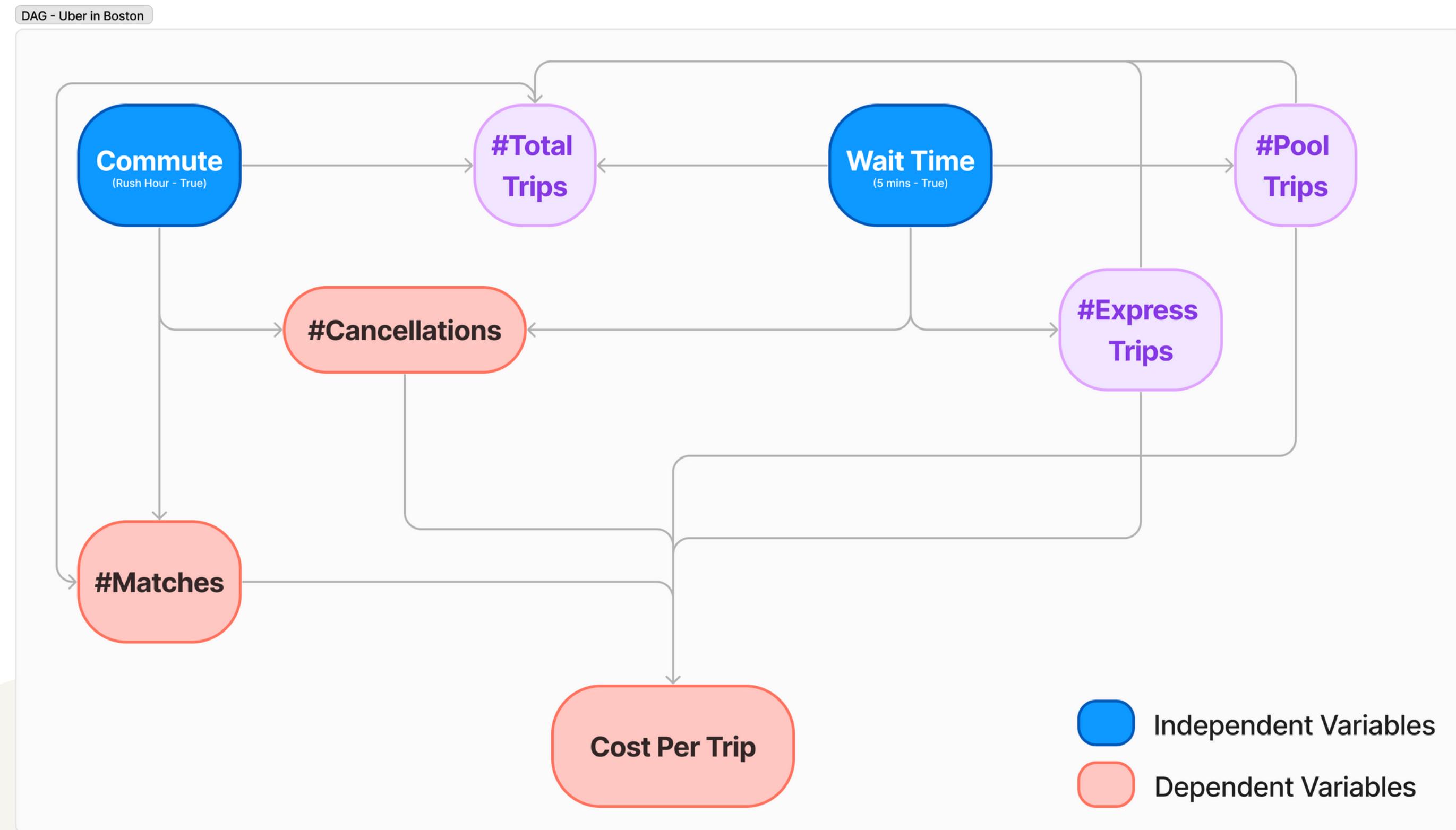
The low p-values indicate a statistically significant difference in the count of POOL trips, Express trips, rider cancellations, cost per trip, and match rate between the two groups\*.

## Importance:

- Distinct, identifiable effects help determine the effectiveness of the treatment.
- Analysis of observed effects helps inform business decisions.

\*Group 'TRUE' has observations where the treatment condition is met; (wait time = 5 mins). Group 'FALSE' has observations where the treatment condition is not met; (wait time is 2 mins).

# Impact Of The Treatment - DAG



# Impact Of The Treatment On Metrics

	trips_pool	trips_express	rider_cancellations	cost_per_trip	match_rate
Intercept	1355.317*** (31.933)	2611.667*** (61.760)	165.349*** (6.339)	7.365*** (0.077)	0.659*** (0.010)
treat[T.True]	105.381** (45.160)	-192.286** (87.342)	24.778*** (8.964)	-0.377*** (0.109)	-0.035** (0.014)
R-squared	0.042	0.038	0.058	0.088	0.049
R-squared Adj.	0.034	0.030	0.050	0.081	0.041

Standard errors in parentheses.

\* p<.1, \*\* p<.05, \*\*\*p<.01

## Key Observations:

- Wait time alone is not a strong determinant of cancellations.
- Wait time only explains ~9% of the variation in the reduction of cost per trip.

# Impact Of The Treatment On Metrics

	trips_pool	trips_express	rider_cancellations	cost_per_trip	match_rate
Intercept	1332.451*** (32.798)	2453.115*** (42.864)	146.988*** (3.541)	7.268*** (0.075)	0.640*** (0.009)
treat[T.True]	105.381** (44.339)	-192.286*** (57.947)	24.778*** (4.786)	-0.377*** (0.102)	-0.035*** (0.012)
commute[T.True]	144.058** (60.667)	998.878*** (79.287)	115.673*** (6.549)	0.612*** (0.139)	0.118*** (0.016)
R-squared	0.084	0.580	0.734	0.212	0.348
R-squared Adj.	0.069	0.573	0.729	0.199	0.337

Standard errors in parentheses.

\* p<.1, \*\* p<.05, \*\*\*p<.01

## Key Observations:

- The coefficients of the treat variable do not change.
- Treat and commute variables explain ~75% of the variance in cancellations.

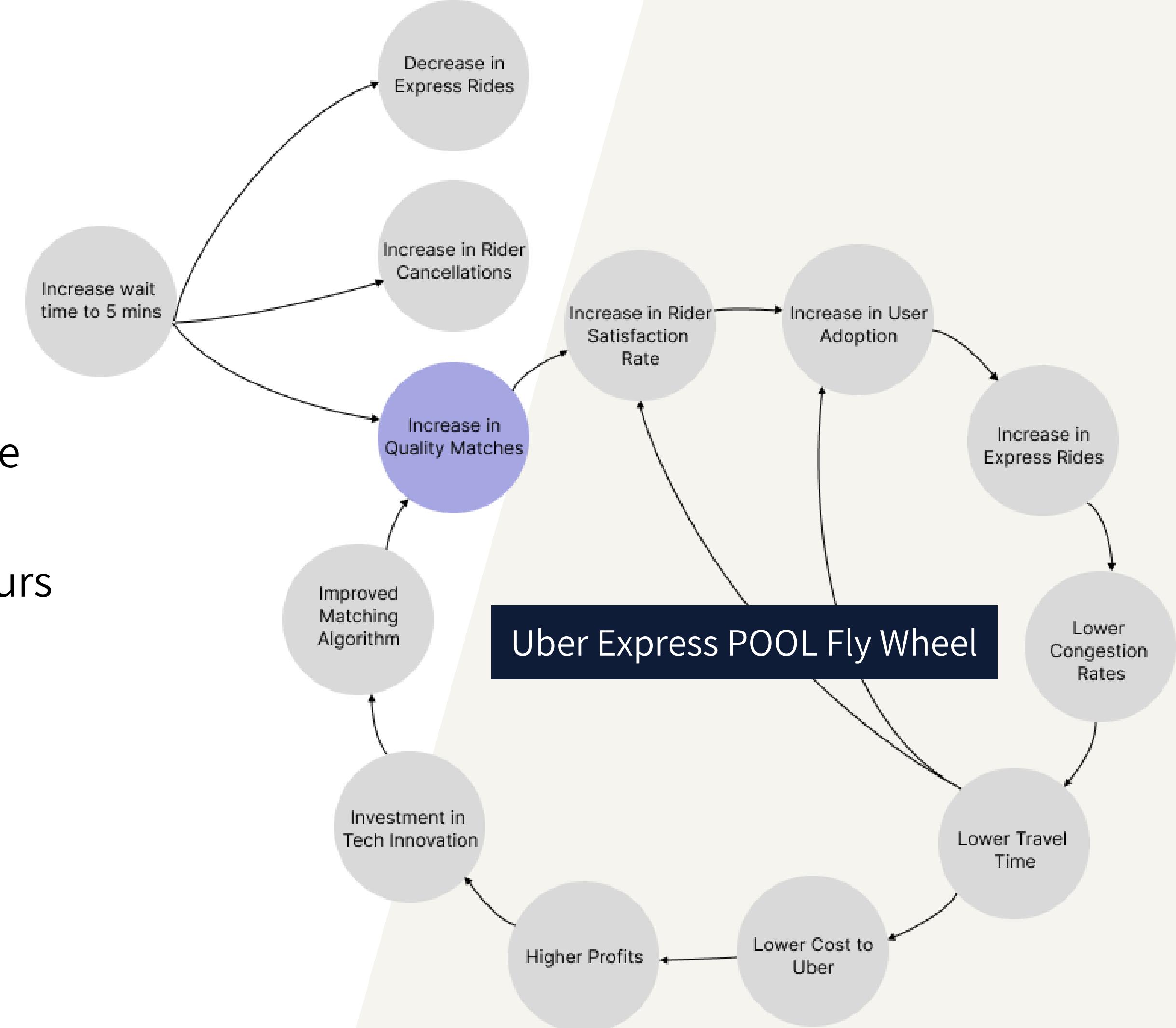
# **Insights & Business Outcomes**



# Action Items

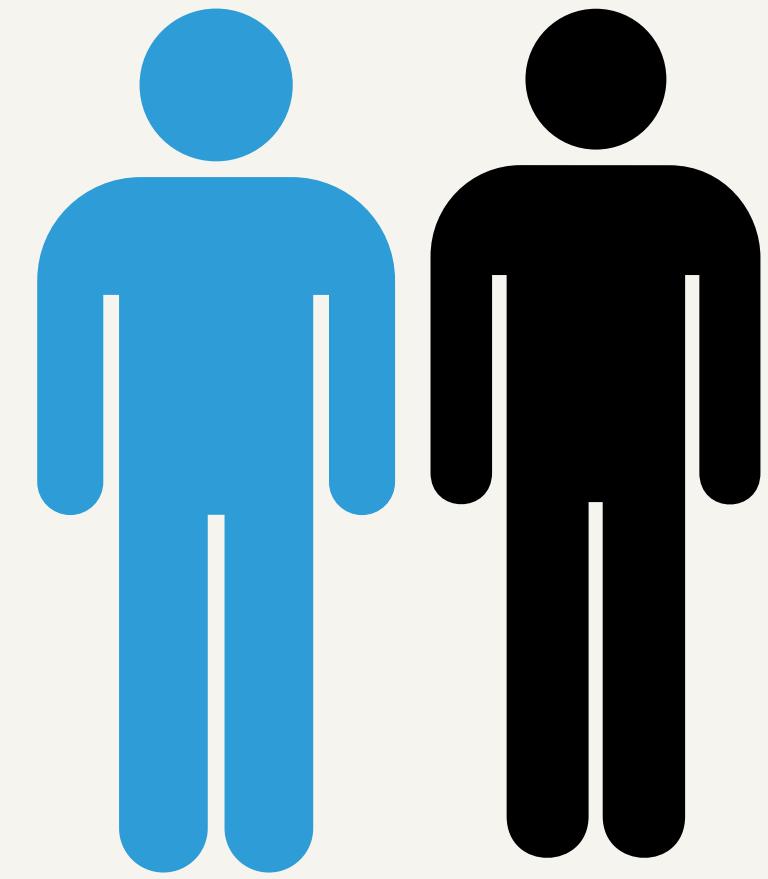
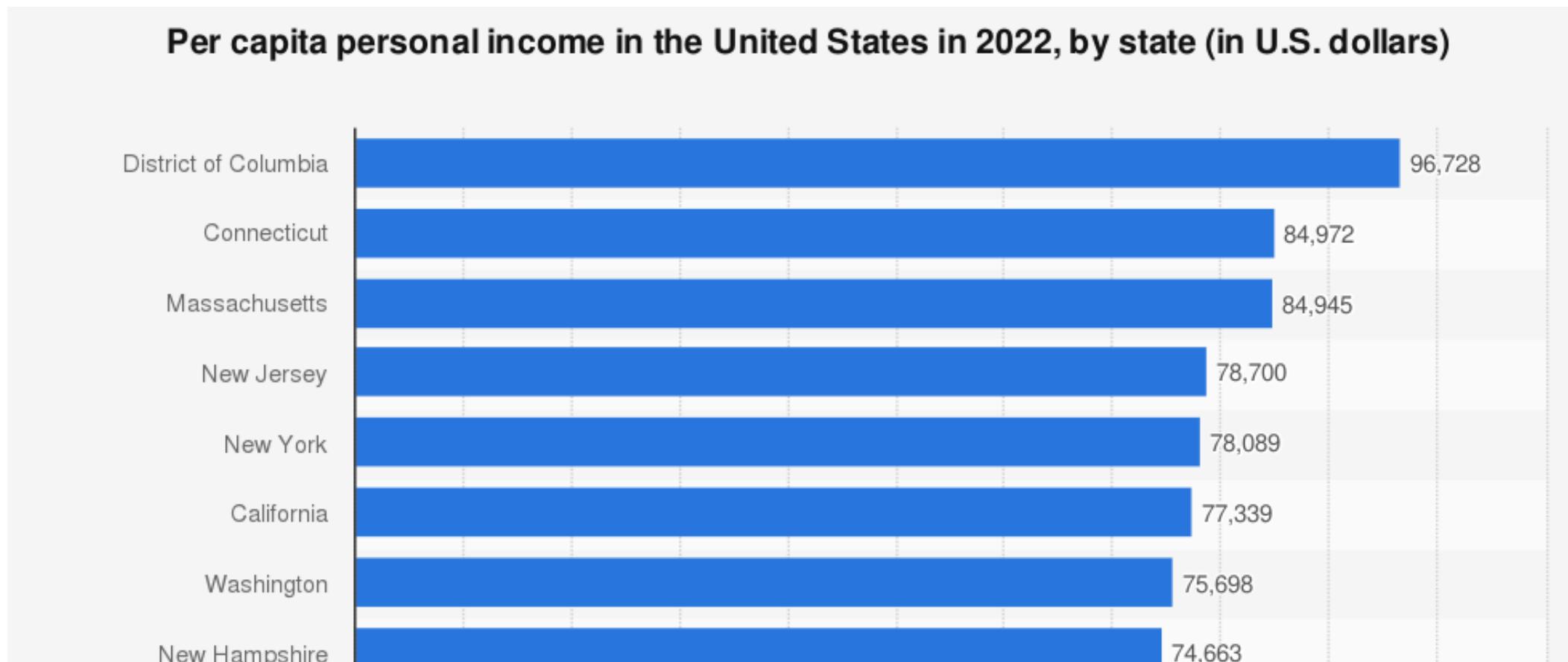
Introduce dynamic waiting time

- 2 mins during rush hours
- 5 mins during non-rush hours



# Limitations

- Data on revenue generated from trips for POOL and Express Pool
- Match rates specific to trip type
- Historic rider cancellation rates
- Selection Bias



Male  
High income  
Higher education level

# Business Value

in Billions USD	2018
Projected Revenue for Uber Mobility	8.9
Proportion of Pool Rides	20%
Uber Pool Revenue	1.78
Feature Adoption Rate	35%
Estimated Increase in Uber Express Pool Revenue	0.623



# Questions

Express POOL

