

### Metadata

#Author	jzan@uber.com	SEMANTIC
#Date Modified	05/30/2018	SEMANTIC
#DateCreated	03/20/2017	SEMANTIC
#Title	Trip Anomaly Detection Project V3 Review - Final	SEMANTIC
Account	nick.silver@uber.com	SEMANTIC
All Custodians	Silver, Nick	SEMANTIC
All Paths	Silver, Nick: \MassTort_Category2_DRIVE\MassTort_Category2_DRIVE_11.zip; Silver, Nick: \MassTort_Category2_DRIVE\MassTort_Category2_DRIVE_11.zip	SEMANTIC
Application	Microsoft 2007 PowerPoint Presentation	SEMANTIC
Attachment Names	ppt	SEMANTIC
Begin Family	UBER_JCCP_MDL_000188659	SEMANTIC
Collaborators	jsevert@uber.com; ambar@uber.com; nirveek@uber.com; frank@uber.com; akankshu@uber.com; dmitriy@uber.com; nick.silver@uber.com; dpurdy@uber.com; uber.com	SEMANTIC
Confidentiality	Confidential	SEMANTIC
Date Created	03/20/2017 7:52 pm	SEMANTIC
Date Modified	05/30/2018 12:25 am	SEMANTIC
DocID	1A-J5Ek_hUx51Qs3R9XIYx8RNNRUE_U9CT31w7C62Tes	SEMANTIC
End Family	UBER_JCCP_MDL_000188685	SEMANTIC
File Path	\MassTort_Category2_DRIVE\MassTort_Category2_DRIVE_11.zip	SEMANTIC
File Size	5142760	SEMANTIC
Filename	Trip Anomaly Detection Project V3 Review - Fi_1A- J5Ek_hUx51Qs3R9XIYx8RNNRUE_U9CT31w7C62Tes.pptx	SEMANTIC
GoogleDocumentType	PRESENTATION	SEMANTIC
Hash Value	c4b00ea1359e1f8f7cc313500b77f9eb	SEMANTIC
Hidden Content	Yes;	SEMANTIC
ILS All Bates	UBER_JCCP_MDL_000188659;UBER_JCCP_MDL_000188660;UBER_JCCP_MDL_0 00188661;UBER_JCCP_MDL_000188662;UBER_JCCP_MDL_000188663;UBER_JC CP_MDL_000188664;UBER_JCCP_MDL_000188665;UBER_JCCP_MDL_000188666 ;UBER_JCCP_MDL_000188667;UBER_JCCP_MDL_000188668;UBER_JCCP_MDL_ 000188669;UBER_JCCP_MDL_000188670;UBER_JCCP_MDL_000188671;UBER_J CCP_MDL_000188672;UBER_JCCP_MDL_000188673;UBER_JCCP_MDL_0001886 74;UBER_JCCP_MDL_000188675;UBER_JCCP_MDL_000188676;UBER_JCCP_MD L_000188677;UBER_JCCP_MDL_000188678;UBER_JCCP_MDL_000188679;UBER JCCP_MDL_000188680;UBER_JCCP_MDL_000188681;UBER_JCCP_MDL_0001886 82;UBER_JCCP_MDL_000188683;UBER_JCCP_MDL_000188684;UBER_JCCP_MD L_000188685	SEMANTIC
ILS Document Date	05/30/2018	SEMANTIC
ILS Prod Date	8/26/2024	SEMANTIC
ILS Prod Vol	Vol044	SEMANTIC
Other Custodians	Silver, Nick	SEMANTIC
Primary Date	03/20/2017 7:52 pm	DOC_TYP E_ALIAS
Production Volume	Vol044;	SEMANTIC
Redacted	No	SEMANTIC
Sort Date	05/30/2018 12:25 am	SEMANTIC
SourceHash	ea4a7f7adf9a343e7c7bb25389920aed	SEMANTIC

**Trial Exhibit No.**  
**P-00801**



# Trip Anomaly Detection

Telematics Team  
Trust & Safety Data Science

May 2017

#### TO CHANGE THE BACKGROUND PHOTO:

- Go to 'View > Master'
- Right-click on the image and select 'Replace image...'
- Right-click on the photo, go to 'Order > Send to Back'

PHOTOS SHOULD BE AT LEAST 1700 x 866px FOR OPTIMAL RESOLUTION

#### WHEN THE PRESENTATION TITLE EXTENDS TO TWO OR MORE LINES:

- Select the background blur by clicking to the left of the presentation title
- Resize the background blur to ensure sufficient contrast with the background photo.

# Motivation



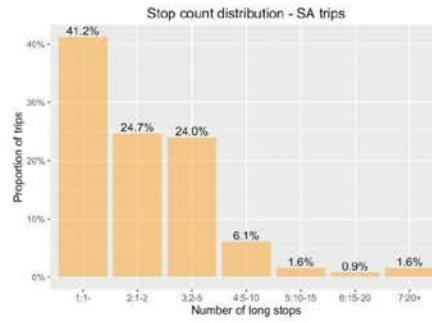
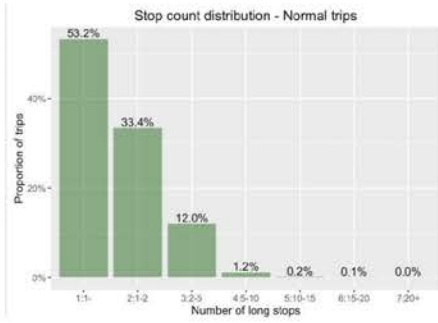
- **No problem in driver profile**
  - Active with Uber for over 6 months
  - Two level 1 tickets
  - Pass BGC
  - Life-time rating: 4.78
- **Risky signals on trip**
  - Mid-night trip
  - At least 6 bars within 1 minute walking distance
  - **Stopped for over an hour near the destination**



## Summary

- Goal:
  - Develop a real-time rule that detects trip anomaly such as
    - On-trip: going in wrong direction, making unnecessary detour, long stops.
    - After-trip: unnecessary contact, long distance between destination and drop-off location, stay close to rider for anomalous long time.
  - Detect the anomalies as early as possible on the trip.
- Status:
  - Multiple well-performed models have been developed
  - Product design jam sessions started
- Rule Performance:
  - Sexual assaults: recall 11.1% / trigger rate 0.5% / lift 22X
  - Rapes: recall 20.3% / trigger rate 0.5% / lift 40X

# SA Trips Have More Long Stops

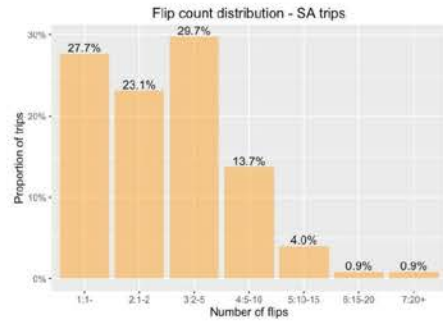
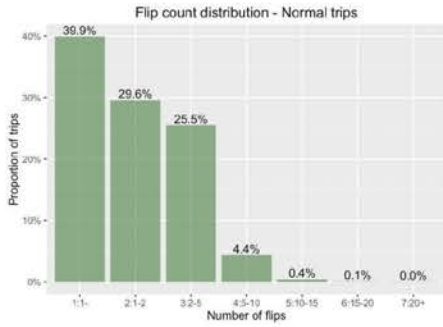


Avg. Number of Long Stops - Normal Trips	Avg. Number of Long Stops - SA Trips	P-value (T-test)	Relative Difference
0.73	1.99	<0.001	173%

Notes: Normal Trips refer to all UberX in the first week of Nov 2016 in LA; SA trips refer to all US UberX SA trips from Oct 1st 2016 to Dec 31st 2016 that have data stored in the ETA table

Stop: the progress made to destination at any checkpoint is less than 10m compared to the previous checkpoint.

# SA Trips Have More Flips



Avg. Number of Flips - Normal Trips	Avg. Number of Flips - SA Trips	P-value (T-test)	Relative Difference
1.29	2.75	<0.001	113%

Notes: Normal Trips refer to all UberX in the first week of Nov 2016 in LA; SA trips refer to all US UberX SA trips from Oct 1st 2016 to Dec 31st 2016 that have data stored in the ETA table

# Trip Anomaly Model

Check UberX trips at every checkpoint (about every 1 minute):

- Real-time features for trip anomaly detection:
  - Number of slow movements in previous 3/5/10 checkpoints
  - Number of moving to wrong directions in previous 3/5/10 checkpoints
  - Relative progress made to destination compare to the previous checkpoint
  - Difference between real-time ETA and original ETA
- Predict sexual assaults using rules with trip anomaly score, trip time, gender, spatial features



# Model Performance - With Gender

*Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Driver-rider different gender,
- Trip Anomaly model score is above threshold  $p^*$ .

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.5%	11.1% (64/575)	22.2X
1%	16.5% (95/575)	16.5X
<b>Non-Consensual Intercourse Only</b>		
0.5%	20.3% (13/64)	40.6X
1%	23.4% (15/64)	23.4X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% (72/647) SA trips lost data in ETA table
  - LB on recall is 9.9% (with 0.5% trigger rate)
- Recall on Rape tickets is calculated from all UberX rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% (30/94) Rape trips lost data in ETA table
  - LB on recall is 13.8% (with 0.5% trigger rate)

# Model Performance - Without Gender

*Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Trip happens on Friday, Saturday or Sunday,
- Trip Anomaly model score is above threshold  $p^*$ .

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.5%	9.6% (55/575)	19.2X
1%	14.3% (82/575)	14.3X
<b>Non-Consensual Intercourse Only</b>		
0.5%	10.9% (7/64)	21.8X
1%	14.1% (9/64)	14.1X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 8.5% (with 0.5% trigger rate)
- Recall on Rape tickets is calculated from all UberX rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 7.4% (with 0.5% trigger rate)

## Case Review


Triggered

- Ticket
  - Driver continued to harass passenger in front seat
  - Made sexual advances both physically and verbally
  - Ignored the correct route
  - Took longer route to destination
- Trip
  - Midnight trip
  - Male driver, female passengers
  - Multiple passengers
  - Went on wrong directions multiple times
  - Made long stops multiple times
- Anomaly detection
  - Model detected anomalous low speed and wrong direction at very early stage of the trip
  - If intervention was sent out right after trip began, we may prevent the situation getting worse




Case 1: triggered


- Ticket
  - Passenger offered sexual favors to driver multiple times
  - Touched driver's private areas
- Trip
  - 9PM trip
  - Male driver male passenger
  - A slight unexpected U-turn on the trip



**Not Triggered**



Sumac Rd  
Broadleaf Dr  
Less than 1 minute  
Poppy Lane  
Sunflower Ave



Heler Ave  
Parkland  
South Parkland  
Wilson Ave  
Bolling Ave  
W Le  
Algonquin Pkwy  
Oregon Ave  
Watson Ct  
Camp Ground Rd  
Cass Run Rd  
Mallory Ln  
Dove Hwy  
Ralph Ave  
Farmley Rd  
Berry Blvd  
Shively  
Crums Ln  
Garrs Ln  
Jacobs/Adr  
Saint Dennis

- Anomaly detection
  - Model did **NOT** detect any anomaly for this trip

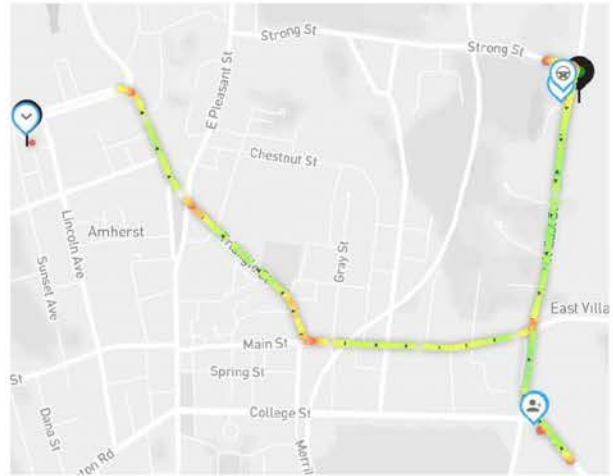
Case 4: not triggered

## Next Steps

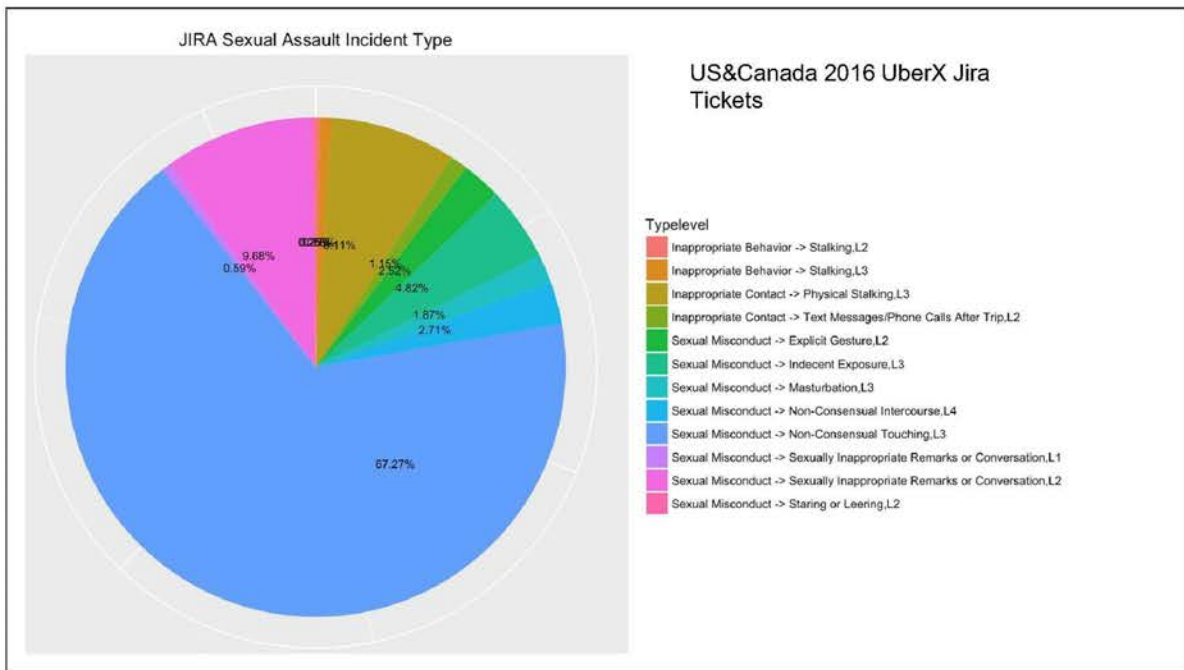
- Proper use cases
- Better intervention

Questions?

- Ticket
  - Driver touched passenger's leg multiple times
- Trip
  - Midnight trip
  - Male driver female passenger
  - Normal route was taken
  - Dropoff location is a few blocks from destination
- Anomaly detection
  - Model did **NOT** detect any anomaly for this trip



Case 3: not triggered



US&Canada 2016 UberX Jira tickets distribution

## Next Steps

- Aggregate the trip anomaly features to account level
  - Check if drivers with SA tickets are more likely to have flips/long stops
  - Potential features for Gwyneth
- Support real-time interventions
  - Prevent opportunistic bad behaviors
  - Offer help and support in real-time

# Naive Model Performance (1) - Without Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Trip happens on Friday, Saturday or Sunday,
- Slow movement in last 3 minutes.

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.5%	7.1% (41/575)	14.2X
<b>Non-Consensual Intercourse Only</b>		
0.5%	20.3% (13/64)	40.6X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 6.3%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 13.8%

## Naive Model Performance (2) - Without Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Trip happens on Friday, Saturday or Sunday,
- Current ETA + time elapsed exceeds 1.8 times initial ETA; and current ETA + time elapsed exceeds initial ETA + 15 minutes

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.54%	8.5% (49/575)	15.7X
<b>Non-Consensual Intercourse Only</b>		
0.54%	9.4% (6/64)	17.4X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 7.6%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 6.4%

# Appendix

# After-trip Rule Performance

*Prediction Rule:* Predict trip will end in sexual assault if:

- Haversine distance between drop-off and destination is longer than  $x^*$  km

	Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>			
$x^*=13$ km (8.08 miles)	0.5%	6.3%	12.6X
$x^*=7$ km (4.35 miles)	1%	11.3%	11.4X
$x^*=3$ km (1.86 miles)	2%	16.9%	8.5X
<b>Non-Consensual Intercourse Only (62% of Rape tickets do not have destination lat/ing data)</b>			
$x^*=13$ km (8.08 miles)	0.5%	3.1%	
$x^*=7$ km (4.35 miles)	1%	6.3%	
$x^*=3$ km (1.86 miles)	2%	12.5%	

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 5.6% (with 0.5% trigger rate)
- Recall on Rape tickets is calculated from all UberX rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 2.1% (with 0.5% trigger rate)

# Model Performance - With Gender

*Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Driver-rider different gender,
- Trip Anomaly model score is above threshold  $p^*$ .

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.07%	1.7% (10/575)	24.3X
0.5%	11.1% (64/575)	22.2X
1%	16.5% (95/575)	16.5X
<b>Non-Consensual Intercourse Only</b>		
0.07%	3.1% (2/64)	44.3X
0.5%	20.3% (13/64)	40.6X
1%	23.4% (15/64)	23.4X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% (72/647) SA trips lost data in ETA table
  - LB on recall is 9.9% (with 0.5% trigger rate)
- Recall on Rape tickets is calculated from all UberX rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% (30/94) Rape trips lost data in ETA table
  - LB on recall is 13.8% (with 0.5% trigger rate)

# Naive Model Performance (1) - Without Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Trip happens on Friday, Saturday or Sunday,
- Slow movement in last 10 minutes.

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.07%	2.1% (12/575)	30X
<b>Non-Consensual Intercourse Only</b>		
0.07%	9.4% (6/64)	134X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 1.9%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 6.4%

# Naive Model Performance (1) - With Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Driver-rider different gender,
- Slow movement in last 10 minutes.

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.07%	2.1% (12/575)	30X
<b>Non-Consensual Intercourse Only</b>		
0.07%	9.4% (6/64)	134X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 1.9%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 6.4%

## Naive Model Performance (2) - With Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Driver-rider different gender,
- Slow movement in last 3 minutes.

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.5%	7.1% (41/575)	14.2X
<b>Non-Consensual Intercourse Only</b>		
0.5%	20.3% (13/64)	40.6X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 6.3%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 13.8%

## Naive Model Performance (3) - With Gender

*Naive Prediction Rule:* Predict trip will end in sexual assault if:

- Trip time between 7pm-5am,
- Driver-rider different gender,
- Current ETA + time elapsed exceeds 1.8 times initial ETA; and current ETA + time elapsed exceeds initial ETA + 15 minutes

Trigger Rate	Recall	Lift in Ticket Rate
<b>All UberX US Sexual Assaults</b>		
0.55%	9.7% (56/575)	17.6X
<b>Non-Consensual Intercourse Only</b>		
0.55%	18.8% (12/64)	34.2X

**Notes:**

- Trigger rate is calculated from all UberX trips in the first week of Nov in LA
- Recall on SA tickets is calculated from all UberX SA trips from Oct 1st to Dec 31st 2016 that have data stored in the ETA table
  - 11% SA trips lost data in ETA table
  - LB on recall is 8.7%
- Recall on Rape tickets is calculated from all UberX Rape ticket trips from Aug 1st to Dec 31st 2016 that have data stored in the ETA table
  - 32% Rape trips lost data in ETA table
  - LB on recall is 12.8%

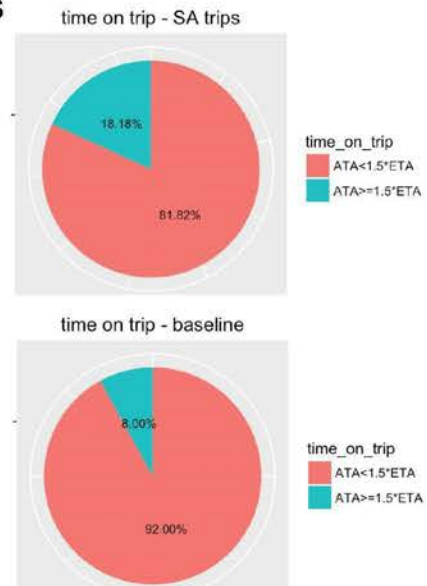
- Ticket
  - Passenger was drunk and made a stop to throw up
  - After driver helped passenger get back he started to ask for a kiss and touch the passenger
  - Hint at prostitution at the end of the trip
  - Mentioned never use Uber again
- Trip
  - Midnight trip
  - Male driver female passenger
  - Passenger was very drunk
  - Picked up from a bar
  - Stopped for very long time in the middle of the trip
- Anomaly detection
  - Model detected anomalous long stop
  - If intervention was sent out right after anomaly was detected, the driver may not have the improper behavior after knowing UBER is 'watching' the trip

Case 2: triggered

## Characteristics of Sexual Assault Trips

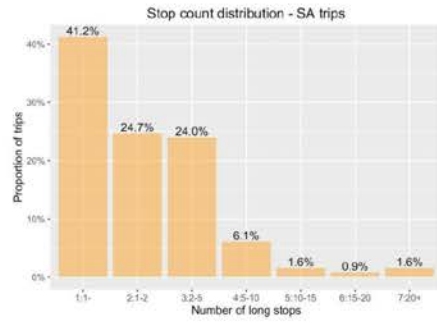
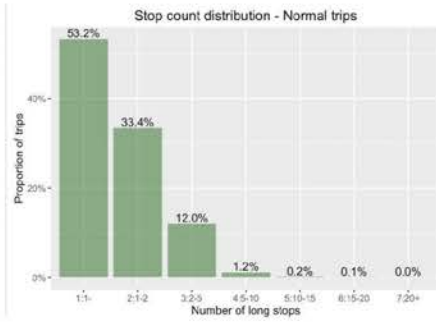
- $\frac{1}{3}$  of the sexual assault trips take 50% longer than initial ETA
- $\frac{1}{3}$  out of those anomalous sexual assault trips are caught by the model

*\*Distribution over SA trips is based on all US&Canada UberX SA trips from 2016/10/01 - 2017-01-01;  
Distribution for the baseline is based on all UberX trips that are longer than 5 miles in Chicago from 2016-10-01 - 2016-10-07*



Focus on sexual assault trips and check the ticket distributions

## SA Trips Have More Long Stops (refreshed data with 2017 Dec)



Avg. Number of Long Stops - Normal Trips	Avg. Number of Long Stops - SA Trips	P-value (T-test)	Relative Difference
0.73	1.99	<0.001	173%

Notes: Normal Trips refer to all UberX in the first week of Nov 2016 in LA; SA trips refer to all US UberX SA trips from Dec 1st 2017 to Dec 31st 2017 that have data stored in the ETA table

Stop: the progress made to destination at any checkpoint is less than 10m compared to the previous checkpoint.