

GLOBAL FEATURES DEFINITIONS

Variable name	Description	Motivation	30(b)(6) Notes
city_sexual_assault_l1234_rate_6mo	The city specific sexual assault (all levels) incident rate	This variable helps to establish a baseline for the expected risk profile of a pairing; helps to allow user features represent more general risk trends	
trip_dow_hour_sexual_assault_l1234_rate_6mo	The city and hour of week specific sexual assault (all levels) incident rate	Works with other trip context features to establish a baseline risk profile of a pairing; helps to allow user features represent more general risk trends	
client_1star_rate_total	The percentage of a client's (i.e. rider's) star ratings received that are 1 star	Low feedback received may be a measure of the elevated riskiness of a user's trips	Using all historical ratings
client_2star_rate_total	The percentage of a client's star ratings received that are 2 star	Low feedback received may be a measure of the elevated riskiness of a user's trips	Using all historical ratings
client_1star_given_rate_total	The percentage of a client's star ratings given that are 1 star	Low feedback given may be a measure of the elevated riskiness of a user's trips	Using all historical ratings
client_2star_given_rate_total	The percentage of a client's star ratings given that are 2 star	Low feedback given may be a measure of the elevated riskiness of a user's trips	Using all historical ratings
client_bliss_l1234_count_all_ao_total	The number of client bliss contacts across all incident levels	Higher bliss contact counts may indicate previous risky behavior	
client_bliss_l1234_rate_all_ao_total	The bliss contact rate of a client across all incident levels	Higher bliss contact may indicate previous risky behavior, but long client tenure may be a mitigating factor	
client_canceled_on_rate_total	The percentage of a client's trips with a driver cancellation	High cancelled on rates may be a measure of the elevated riskiness of a user's trips	
client_cancellation_rate_total	The percentage of a client's trips with a rider cancellation	A possible indication of driver screening behavior from the rider	
client_trip_completion_rate	The percentage of a client's trips that end completed	Low trip completion rates may be a measure of the elevated riskiness of a user's trips	
client_weekend_request_rate_total	The percentage of a client's trips that occurred on a weekend	Higher weekend request rates correspond to a high percentage of a user's trips occurring during periods of elevated risk	"Weekend" is from Friday at 12:00 am to Sunday at 11:59pm
client_latenight_request_rate_total	The percentage of a client's trips that occurred late at night	Higher late night request rates correspond to a high percentage of a user's trips occurring during periods of elevated risk	Latenight is defined as 12:00am to 4:59am
client_cash_rate	The percentage of a client's trips with cash as the payment method	Cash is an anonymous payment method; higher rates may indicate ill intent	
client_request_to_begin_distance_miles_avg	The average distance in miles between a client's trip request and begin locations	Longer average distances between trip request and begin may be linked to elevated trip risk	
client_avg_ata	The average actual time of arrive of a client's trips	Longer trip wait times may be linked to elevated trip risk	
client_first_token_type	The client's first payment token type (cash, Apple Pay, PayPal, etc.)	Different payment methods have different risk profiles	
client_confirmed_mobile	Is the client's mobile confirmed	Having a confirmed mobile device corresponds to lower conditional risk of a trip	
client_ipc_incident_count_driver_ao	The number of interpersonal conflict (IPC) incidents reported by drivers on the client's trips	Higher IPC incident counts may indicate previous risky behavior	
client_ipc_incident_count_rider_ao	The number of IPC incidents reported by the client on the client's trips	Higher IPC incident counts may indicate previous risky behavior	
client_sexual_count_driver_ao	The number of sexual incidents reported by drivers on the client's trips	Higher sexual incident counts may indicate previous risky behavior	
client_sexual_count_rider_ao	The number of sexual incidents reported by the client on the client's trips	Higher sexual incident counts may indicate previous risky behavior	
client_signup_dow	The day of week that a client signed up	May help to differentiate between otherwise similar riders; for example, assessing the risk of a rider's first trip during the week versus the weekend	
client_signup_hour	The hour of the day that a client signed up	May help to differentiate between otherwise similar riders; for example, assessing the risk of a rider's first trip during the day versus the night	
client_days_since_signup	The number of days since a client signed up	A long rider tenure may indicate higher confidence that a trip with this rider will not result in a serious sexual assault or sexual misconduct (SASM) incident compared to short tenure riders	

driver_bliss_l1234_rate_all_ao	The bliss contact rate of a driver across all incident levels	Higher bliss contact may indicate previous risky behavior, but long driver tenure may be a mitigating factor	
driver_ipc_incident_rate_driver_ao	The IPC incident rate of a driver across all incident levels	Higher IPC incident rate indicate previous risky behavior, but long driver tenure may be a mitigating factor	
driver_latenight_request_rate_total	The percentage of a driver's trips that occurred late at night	Higher late night request rates correspond to a high percentage of a user's trips occurring during periods of elevated risk	Latenight is defined as 12:00am to 4:59am
driver_sexual_rate_driver_ao_total	The sexual incident rate of a driver across all incident levels	Higher sexual incident rate indicate previous risky behavior, but long driver tenure may be a mitigating factor	Using lifetime rides
driver_weekend_request_rate_total	The percentage of a driver's trips that occurred on a weekend	Higher weekend request rates correspond to a high percentage of a user's trips occurring during periods of elevated risk	"Weekend" is from Friday at 12:00 am to Sunday at 11:59pm
driver_avg_trip_time	The average time of a driver's trips	Longer average trip times may be related to a higher probability that trips with this driver will successfully end without incident	
geohash6_cash_trip_rate_6mo	The overall cash trip rate in the neighborhood around the begin trip location over the past 6 months	Very high percentages may indicate higher trip risk; this variable allows the model to learn any marginal characteristics of user behavior	Geohashes are roughly rectangular. geohash6 represents an area approximately 1.22km by 0.61km, while geohash7 represents a smaller area, roughly 152.9m by 152.4m.
geohash6_flack_ipc_rate_all_ao_6mo	The overall IPC incident rate in the neighborhood around the begin trip location over the past 6 months	Higher localized IPC rates mean higher conditional risk of the trip; this variable allows the model to learn any marginal characteristics of user behavior	
geohash6_latenight_trip_rate_6mo	The percentage of trips occurring late at night in the neighborhood around the begin trip location over the past 6 months	High percentages may indicate higher trip risk; this variable allows the model to learn any marginal characteristics of user behavior	
geohash6_log_trip_count_6mo	The number of trips that occurred in the neighborhood around the begin trip location over the past 6 months	Areas with few trips or many trips have different risk profiles; this variable allows the model to learn any marginal characteristics of user behavior	
log_client_ltd_completed_count	The number of lifetime client trips	Long rider tenure may indicate lower conditional risk of the trip	
log_driver_lifetime_trip_count	The number of lifetime driver trips	Long driver tenure may indicate lower conditional risk of the trip	
log_driver_trip_duration_seconds_sum_total	The total amount of a driver's on trip time	Long driver tenure or long history of being successfully on trip may indicate lower conditional risk of the trip	
trip_dow_local	The day of week of the trip	Different days of the week have different conditional risks; this variable allows the model to learn any marginal characteristics of user behavior	
trip_global_product_name	The flow type of the trip (UberX, Black, etc.)	Different product types have different conditional risks; this variable allows the model to learn any marginal characteristics of user behavior	
trip_hour_local	The hour of the day of the trip	Different times of the day have different conditional risks; this variable allows the model to learn any marginal characteristics of user behavior	
trip_sub_region	The sub region of the trip	Different trip regions have different conditional risks; this variable allows the model to learn any marginal characteristics of user behavior	
bar_count_geohash7	The number of identified bars in the neighborhood around the trip request location	Serious SASM rates tend to be correlated with the number of bars in the area around the trip request	Bar area definition includes 'Bar', 'Liquor & Wine', 'Night Club'
restaurant_count_geohash7	The number of identified restaurants in the neighborhood around the trip request location	Serious SASM rates tend to be correlated with the number of restaurants in the area around the trip request	