Data Driven Business Models
Due date: 1:30pm on December 28 (submit on E-class)

1. Backgrounds

• Understanding the Uber system
To begin, it is important to understand the earning systems of Uber. According to the Uber website, the earnings of Uber driver are calculated by: Standard Fare (base fare + amounts for each minute and mile driver drive), Minimum trip earnings, Service fee, Booking fee, Cancellation fee, and most importantly: Surge Pricing.
Surge Pricing is applied when there is a high demand case. Examples of high demand case: Bad weather, rush hour, and special events etc. In these cases of very high demand, prices may increase to help ensure that those who need a ride can get one by attracting more drivers. This is called Surge Pricing. Surge prices are calculated based on a multiplier to standard rates, an additional surge amount, or an upfront fare including the surge amount. Uber's service fee percentage does not change during surge pricing. This is the key of Uber business model to make a balance between passengers and drivers (Uber should control for this balance by time and day, unlike other platforms).

• Competition between traditional businesses and new ride hailing businesses
While Uber rapidly launches its service across many cities in the world, Yellow Cabs are still alive. Uber drivers are getting more number of trips than the traditional taxis, which means Uber drivers are performing twice than the other taxi drivers. However, surprisingly Yellow Cabs are still a big rivalry to on-demand taxi services like Uber and Lyft in New York City .

NYC is the home of traditional Yellow taxis where people can get quick and affordable taxi rides. These days, riding in Yellow Cabs becomes a gloomy experience, and people are fed up with the irregular and messy rides. This is one of the big reasons behind the passengers' migration to on-demand taxi services that promised reliable rides, friendly drivers and seamless payment options to make the customers more convenient. Uber has something called surge pricing, which is pretty higher than the other taxis.

Yellow Cabs are challenging the competition among the on-demand taxi services like Uber and Lyft by placing huge bets on these ride-sharing services. When the cost is certainly a major concern while taking a taxi ride, in NYC, people can get affordable taxi service by hailing Yellow Cabs during peak hours or even during when Uber's surge pricing is in place. On-demand taxi companies like Uber have already raised the expectation towards their potential customers. To survive and stay competitive with the competitors, Yellow Cabs will need to rise to the challenge.

2. Data

•Yellow Cab data

Field Name	Descriptin
VendorID	A code indicating the TPEP provider that provided the record. **1= Creative Mobile Technologies, LLC; 2= VeriFone Inc.**
tpep_pickup_datetime	The date and time when the meter was engaged.
tpep_dropoff_datetime	The date and time when the meter was disengaged.

Passenger_count	The number of passengers in the vehicle. This is a driver-entered value.
Trip_distance	The elapsed trip distance in miles reported by the taximeter.
PULocationID	TLC Taxi Zone in which the taximeter was engaged
DOLocationID	TLC Taxi Zone in which the taximeter was disengaged
RateCodeID	The final rate code in effect at the end of the trip. **1=Standard rate** **2=JFK** **3=Newark** **4=Nassau or Westchester** **5=Negotiated fare** **6=Group ride**
Store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor, aka "store and forward,"because the vehicle did not have a connection to the server. **Y= store and forward trip** **N= not a store and forward trip**
Payment_type	A numeric code signifying how the passenger paid for the trip. **1= Credit card** **2= Cash** **3= No charge** **4= Dispute** **5= Unknown** **6= Voided trip**
Fare_amount	The time-and-distance fare calculated by the meter.
Extra	Miscellaneous extras and surcharges. Currently, this only includes the $0.50 and $1 rush hour and overnight charges.
MTA_tax	$0.50 MTA tax that is automatically triggered based on the metered rate in use.
Improvement_surcharge	$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.
Tip_amount	Tip amount – This field is automatically populated for credit card tips. Cash tips are not included.
Tolls_amount	Total amount of all tolls paid in trip.
Total_amount	The total amount charged to passengers. Does not include cash tips.
Congestion_Surcharge	Total amount collected in trip for NYS congestion surcharge.
Airport_fee	$1.25 for pick up only at LaGuardia and John F. Kennedy Airports

•Ride-hailing service data (Uber and Lyft data)

Field Name	Description
Hvfhs_license_num	The TLC license number of the HVFHS base or business As of September 2019, the HVFHS licensees are the following: • HV0002: Juno • HV0003: Uber • HV0004: Via • HV0005: Lyft
Dispatching_base_num	The TLC Base License Number of the base that dispatched the trip
Pickup_datetime	The date and time of the trip pick-up
DropOff_datetime	The date and time of the trip drop-off
PULocationID	TLC Taxi Zone in which the trip began
DOLocationID	TLC Taxi Zone in which the trip ended
originating_base_num	base number of the base that received the original trip request
request_datetime	date/time when passenger requested to be picked up
on_scene_datetime	date/time when driver arrived at the pick-up location (Accessible Vehicles-only)
trip_miles	total miles for passenger trip
trip_time	total time in seconds for passenger trip
base_passenger_fare	base passenger fare before tolls, tips, taxes, and fees
tolls	total amount of all tolls paid in trip
bcf	total amount collected in trip for Black Car Fund
sales_tax	total amount collected in trip for NYS sales tax
congestion_surcharge	total amount collected in trip for NYS congestion surcharge
airport_fee	$2.50 for both drop off and pick up at LaGuardia, Newark, and John F. Kennedy airports
tips	total amount of tips received from passenger
driver_pay	total driver pay (not including tolls or tips and net of commission, surcharges, or taxes)
shared_request_flag	Did the passenger agree to a shared/pooled ride, regardless of whether they were matched? (Y/N)
shared_match_flag	Did the passenger share the vehicle with another passenger who booked separately at any point during the trip? (Y/N)
access_a_ride_flag	Was the trip administered on behalf of the Metropolitan Transportation Authority (MTA)? (Y/N)
wav_request_flag	Did the passenger request a wheelchair-accessible vehicle (WAV)? (Y/N)
wav_match_flag	Did the trip occur in a wheelchair-accessible vehicle (WAV)? (Y/N)

For the analysis, the following datasets are collected by 10% RANDOMLY sampling from transportation records in March 2024 in New York City. Use the given dataset to answer the following questions.

3. Analysis

1) Identify empirical evidence of network effects in Uber's business model.
Note that while Uber may exhibit certain network effects, it does not encompass all types of network effects. Please focus on finding specific examples of network effects for drivers and riders within Uber's business model.

2) Compare the market shares and pricing models of taxis, Uber, and Lyft.
Additionally, evaluate the pricing strategies employed by these services. Are there any competitive advantages in Uber's pricing model compared to taxis or Lyft?

3) Analyze dynamic pricing in ride-hailing platforms like Uber and Lyft.
As discussed in class, ride-hailing companies act as matchmakers between riders and drivers. The core of their business model lies in maintaining a balance between these two sides. To achieve this, ride-hailing platforms use dynamic pricing (i.e., prices that vary based on time and demand). For example, Uber might increase (surge) prices to attract more drivers when there's high rider demand or decrease prices to attract more riders when drivers outnumber riders.

Based on the comparison in the previous question, do Uber and Lyft need to adjust their prices? If so, explain why and how they should change their pricing (to maximize profits).

NOTE:

1. **Please use double-space and 11 points.**
2. **You report cannot exceed 10 pages.**
3. **Write your report only in plain English.**
4. **Please attach ALL your Python code in Appendix (which does not count for the page limit) e.g., 1) Python code, 1) Python code, etc..**
5. **Please upload "ONE" pdf or doc file on E-class (do not submit multiple files!).**
6. **10% points are deducted for 1 hour late submission (thus please submit your report well ahead of the deadline). Would not accept any excuses such as system failures, etc.**