The Ride-Sharing Analytics Game: December 4, 2018

IE1086/IE2086, University of Pittsburgh, Fall 2018

Instructor: Daniel R. Jiang, Designer: Nick Kelly

November 27, 2018

Imagine that you work in the analytics department for a ride-sharing company that features both contracted drivers and autonomous vehicles (AVs), which need to be integrated into a system of 10 U.S. cities. In each month, you will make decisions related to (1) pricing to manage supply and demand, (2) manufacturing AVs, (3) repositioning of AVs, and (4) advertising to build a customer base. There are two manufacturing facilities where you can build AVs that become available in the next period; once they are available, you will be allowed to transport them to any city of your choice. Even after they are operational, you are allowed to reposition them between cities in order to satisfy seasonal demand. The company would like you to maximize their yearly profits by providing thorough analysis of their performance and then designing a decision-making policy.

This is a "real-time" game: (1) you will submit decisions to the instructors; (2) our simulation system will run your decisions for a single period (one month) and provide feedback to you in the form of a new state of the system, revenues/costs, and other useful information; (3) you will then submit decisions for the next month. Communications will occur via Slack and Dropbox (instructions below).

Information Provided: To meet these objectives, you will be provided with a year of historical data from the 10 cities in the U.S. serviced by the company, a snapshot of the current state of the company in each city, monthly reports, and any relevant static information.

Historical Data. Several `.csv` files are provided to you that contain data relevant for analysis:

- `Advertising_Data.csv`: the amount spent on advertising, the riders and drivers gained, and the population of the cities. Advertising adds to the numbers of riders and drivers who sign up for the platform (but proportional to the population – e.g., advertising will bring in larger quantities of users in bigger cities). Some further observations:
  - We have learned that no driver on the application is also a rider, so the population can be split into two groups, but we are unsure of the size of each.
  - We have noticed that when we spend a specific amount of money on advertising, it does not always have the same effect and this may be due to our rider and driver populations nearing the maximum of what is available.
- `Lost_Users_Data.csv`: the number of riders and drivers who chose to discontinue service with our company each month. Since our contracts are month to month, we have noticed that some of our riders and drivers will only use our service for a month or 2 before deleting their accounts.

- `Fare_Data.csv`: the number of riders and drivers that sought our services based on the ride price we chose. Each day of the month (assumed 30 days each month), every rider or driver on the platform has some probability of choosing to ride or drive based on the fare price. The formula is

$$\begin{aligned}\text{monthly revenue} = 20\% \cdot \text{fare price} \cdot \sum_i \min(&\text{drivers on day } i, \text{riders on day } i) \\ + 100\% \cdot \text{fare price} \cdot \sum_i \min(&\text{AVs on day } i, \text{remaining riders on day } i).\end{aligned}$$

  Some further observations:

  – We have noticed that when we charge more per ride in a city, our riders may turn to public transportation or other means. We have also noticed that more drivers are inclined to devote their time to driving when they have the opportunity to make more money. Therefore, a higher fare price boosts the per-ride revenue and leads to more drivers but fewer riders.

- `Surplus_Data.csv`: the number of surplus riders that appear each month compared to our projections. These are riders that come into town temporarily for vacation and/or other events and could be a significant source of revenue.

  – Our projections for the number of surplus riders in each city are overly cautious to ensure that we do not underestimate the actual demand in any particular city.

State of the System. The state of the system is provided in `State_Team_TeamName_MX.csv` and contains the following information.

- The number of drivers contracted in each city.
- The number of riders signed-up with the company in each city.
- The number of AVs located in each city and in each of the two factories.

Decisions. You will be expected to create a policy that observes the state of the system and makes the following decisions.

- The fare to charge in each city.
- The amount to advertise in each city to attract riders and drivers.
- The positioning of the available AVs from the current cities and factories.
- The amount of AVs that our factories should build in a given month.

All decisions will be input into `Decisions_Team_TeamName_MX.csv` and submitted to the instructor.

Monthly Reports. At the beginning of each month, you will receive a report, also within `State_Team_TeamName.csv`, regarding the 'transitions' that occurred throughout the previous month.

- Riders served over the course of the previous month.

- Drivers that ended their contracts with the company.
- Riders that deleted their accounts.
- Drivers that saw the ads and started contracts with the company.
- Riders that saw the ads and decided to create accounts with the company.
- Surplus riders that appeared in the city and were served in that month.
- The percentage of drivers that appeared and served riders.
- The percentage of riders/surplus riders that appeared and were served.
- The percentage of AVs that were used to serve riders.
- Rides that were provided by the AVs in each city.

Static Data. Here are the static pieces of data in the system.

- Upkeep cost of each autonomous vehicle: $1.5/month.
- Travel cost per autonomous vehicle $0.002/mile.
- Build cost of one autonomous vehicle $250 (2 factories available).
- Monthly build capacity of each factory (100,000 in Factory 1 and 200,000 in Factory 2).
- Distance (in miles) between cities and factories.
  - See `City_Distances.csv` and `Factory_Distances.csv`.
- Projected amount of surplus riders that will appear in each city.
  - See `Surplus_Data.csv`.

Rooms Available in Benedum. 1041, 1025E, 1021 on 11/29 and 12/4 from 9-11am.

Teams. Teams are assigned in the spreadsheet below. They were chosen based on the project teams, except organized into groups of four or five.

`https://goo.gl/kuX7Pd`

Communication on Game Day. We will use Slack. A private channel will be created for each team and communication with the instructors will occur there. During the game, you will let the instructors know of any problems and confirm when your decisions have been submitted. Everybody, please join Slack via this invitation link using your Pitt email:

`https://goo.gl/cEK3ma`

Dropbox. Please identify one team member who will create a Dropbox account (or use an existing account) and send the account email to the instructors (drjiang@pitt.edu and njk48@pitt.edu). Nick will help with the Dropbox set up.

Grading. This will count as the final homework. Participation will give you 75% credit and the remaining 25% will be based on your performance. Hint: build a simulator to test ideas!