Understanding the Impact of the Nightlife Industry on Urban Mobility in NYC and Charting COVID-19 Related Disruption

Progress Report 1

Kaifu Ren (kr2516), Nicholas LiCalzi (nl949)

Yingyuan Zhang (yz6378), Yutong Zhu (yz6080)

New York University CUSP

Apr. 1, 2020

Sponsor: VibeLab

CUSP Mentor: Kim Mahler

Abstract:

The nightlife industry (bars, restaurants, clubs, music venues, and more) is a major economic and cultural engine for New York City. In collaboration with VibeLab, an international consultancy specializing in nightlife and cultural vibrancy, our goal is to quantify the specific effects of NYC's night time economy on its modes of transportation. In doing so, we might be able to offer targeted recommendations for ensuring that nightlife activities are adequately serviced by transit and help stakeholders both in the industry and in government collaborate in a productive, informed manner. The unprecedented disruption to the economic activity and movement of New Yorkers unleashed by the COVID-19 pandemic, and the resulting rolling shutdown of bars, clubs, and music venues, offers urban scientists a unique and interesting opportunity to quantify how much transportation activity is uniquely due to specific activities, given that those activities were halted before the rest of the city was shuttered or paused.

Introduction:

New York City has been affectionately known as the "City That Never Sleeps" ever since Liza Minelli performed '*Theme from "New York, New York"'* for Martin Scorsese's film. This nickname alludes to the plethora of nightlife options on offer in the city: dining, dancing, performing, socializing, building a career, and more. Nightlife makes up a significant portion of the city's economy and culture, with more than 25,000 nightlife establishments citywide (NYC Nightlife Report, 2019). This is made possible in no small part because of the City's 24 hour public transportation systems, with subways, buses, cabs, and bikes available around the clock for public use, unlike cities like London or Tokyo where the subway might close at midnight and bus routes and service increase to compensate. NYC has not only excellent provision of transit, but excellent provision of transit-related open data as well; with granular data about Citibike pickups and dropoffs, MTA subway turnstile entrance and exit counts, and taxi pickups and

dropoffs available for exploration by urban scientists. Armed with this data, we can build a baseline understanding about how people move across the city at a given time and to/from a given location.

One unexpected twist we've encountered on the project after embarking upon our research is the arrival and rapid dissemination of COVID-19 across NYC, resulting in the de facto shuttering of the entire city, and the de jure shuttering of its nightlife industry (at least for now, and very likely with permanent repercussions). "As the coronavirus continues its march around the world, governments have turned to proven public health measures, such as social distancing, to physically disrupt the contagion" (Carlsson-Szlezak, Philipp, et al, 2020). While NYC will certainly suffer a huge economic loss due to the forced shutdown and bans on public gatherings and venues closed citywide; the nightlife industry in NYC, largely fueled by freelancers, is experiencing an existentially threatening financial fallout (Lost My GIG NYC).

Having established a snapshot of the urban transportation network and its night time utilization during "normal" times, the research team will then inspect the impact of COVID-19 on the city's nightlife economy in two parts. Firstly, from a social/mobility perspective: with thousands of artists, nightlife workers, and audiences traveling to and from venues nightly, changes to their mobility patterns can have a significant impact over the entire network. The team will leverage the baseline model of all nightly travel under normal conditions to understand the changes in mobility and activities that occurred once nightlife businesses and venues were due to COVID-19. Secondly, from an economic perspective, the team will try to capture and depict the big picture of the scale of economic loss felt by the transportation system and nightlife businesses in NYC. This research is conducted with a goal of helping policymakers and city managers to allocate resources for financial aid and hold deliberations with industry stakeholders.

Literature Review:

We are lucky to have a project sponsor in VibeLab that has carried out a comprehensive initial study of the nightlife industry in NYC (Creative Footprint Report), covering 495 individual music venues. Manhattan has the largest number of venues and has the city's largest and oldest venues, while Brooklyn is dominated by younger venues. Their research shows that venues, young adults, and subway transit are generally correlated in location, and that in both Manhattan and Brooklyn, neighborhoods with venues have higher transit densities than those areas without.

There is a large corpus of academic work about the time-series analysis of multi-modal mobility/trajectory data. One research team proposed a predictive model to collectively forecast the inflow and outflow of crowds in each and every region of a city based on historical taxi and bike data, weather and events (Zhang, Junbo, et al, 2017). In this paper, various mobility trips were aggregated into self-defined city grids to present inflow and outflow analysis. We have been able to find few research works concerning the analysis of traffic flows during major citywide disruptions to transit service, but we hope to be able to source more studies soon.

Problem Statement:

- Questions

1. Understanding the impact of COVID-19 over New Yorkers' access to music venues at night:
    - How does rideshare/ TLC/ CitiBike/ Subway/ Bus usage look like around music venues in various parts of NYC at night (Pre COVID-19)?
    - How does this change during special events (Pre COVID-19)?
    - How does the COVID-19 outbreak change these transportation patterns around music venues?

2. Understanding the economic loss of transportation sector due to the cancellation of nightlife events:
    - What is the economic benefit for the transportation system brought by music events (Pre COVID-19)?
    - What's the economic loss under the situations?

- Hypothesis: expected outcomes
    - To access different nightlife hotspots, New Yorkers tend to prefer different means of transportation.
    - Larger events like concerts may have higher impact over various transportation networks around them.
    - Due to the shut down of music venues and cancellation of events during COVID-19 outbreaks, the usage for various modes of transportation tends to reduce significantly.

Data:

We will integrate all mobility datasets together and grouped them based on some of the following geographies: arbitrarily sized hexagonal tessellation, NYC NTAs/other city geographies, and possibly the NYC taxi zones (Refer to Appendix I for data sources).

1. TLC monthly data: contains region to region commercial vehicles' trip data from 2009 to 2019.
2. Citibike data: primary datasets provide each trip's start time and date, stop time and date, start and end station name, station latitude and longitude as well as the user's gender and type.
3. MTA real-time data: primary datasets provide the weekly turnstile usage, daily MTA traffic situation, NYC transit subway schedules and stations.

4. New York City taxi zone shape file: contains geographic data. We will analyze mobility patterns based on transitions among taxi zones in NYC.
5. NYC Music Venue data: collects all listed venues's type, events' per month, operation time, location's latitude and longitude, as well as number of participants and operating situations. Those data need to be scrapped via Ticketmaster and Eventbrite APIs.
6. COVID-19 data: contains forecast data from John Hopkins CSSE. It will provide evident data for our team to examine the COVID-19 impact on mobility.
7. NYC Regulation data: contains all nighlife-related policy established dates and effective dates. It will be interpreted with COVID-19 data together.

Methods:

Exploratory data analysis and data engineering:

1. Mapping music venues' locations among 5 boroughs in NYC;
2. Understanding music venues' operation situations (open, closed, partially open, online or delivery services) before and after COVID-19 outbreaks;
3. Collecting types, time, number of participants of events in NYC via ticketing APIs;
4. Mapping subway lines, bus services, Citibike stations, taxi pick-up and drop-off points adjacent to music venues
5. Overlaying economic contribution with music venues' locations since 2019

Time Series analysis and comparison study:

1. Monitoring hourly changes of inflow and outflow for different mobility usage (MTA, Citibikes, buses, TLC, rideshares) in each music venues' hotspots and overlaying them with event information to understand how New Yorkers access these music venues and how does certain type of events affect the network in pre COVID-19 period;

2. Aggregating trip's destinations to music venues hotspots and using their origins to understand which neighborhoods in NYC mostly engaged with nightlife events in pre COVID-19 period;
3. Aggregating per day's different mobility usage (MTA, Citibikes, buses, TLC, rideshares) in each music venues' hotspots and overlaying them with key government regulation milestones and daily report of COVID-19 cases in NYC to observe and summarize the trend; and
4. Make comparison study of mobility usage rate with the same month but last year to show differences in mobility demands caused by COVID-19

Analysis and comparison of economic impact:

1. Draw heatmap with music venue's high activity locations in 2020 Q1
2. Combine with time series to estimate economic loss of the music events and transportation system
3. Find relationships within transportation system and nightlife events
4. Compare numbers of music events with March 2019 and analysis the differences due to COVID-19

Research Plan:

- Gantt Chart: steps, milestones, and timelines

Project Risks and Mitigation Strategies

1. Risks: Cities are complex systems. Passengers take different means of transportation for different purposes. Assumptions to construct the causal relationship between traffic

volumes and nearby music venues tend to be unjustified. Specifically, it's hard to define how much change in traffic volume is caused by nightlife events.

Mitigation strategies: Traffic pattern analysis will be conducted to depict the influences.

2. Risks: Data limitation and data unavailability. The 2020 first quarter's TLC data has not been published yet, our study on the mobility pattern for NYC nightlife industry will be unconcret for the immediate Pre COVID-19 period.

   Mitigation strategy: We all actively look for other datasets to fill the infection of our data. Alternatively, we will use predicted data based on past year TLC data.

3. Risks: Due to data privacy, we could not get access to point-to-point level trajectory data in TLC dataset.

   Mitigation strategy: Our analysis has to be aggregated over NYC taxi zones. Further studies need to be carried on to provide finer resolutions and keep privacy issues at the same time.

4. Risks: As the monthly revenue of Citibike and MTA stay unclear, the prediction on economic loss for the transportation system might be inaccurate and biased.

Team Roles

- Nicholas LiCalzi: acquisition, cleaning, and parsing of scraped- and open-datasets; geo-spatial data and spatial auto-correlation analysis; data visualization and presentation; website design and development
- Kaifu Ren: Data mining with APIs to prepare events related information; data engineering and cleaning; building data visualization tools to display the result
- Yingyuan Zhang: Cleaning monthly data of Citibike; building data visualization to show relationship of nightlife events and citibike; building Time Series model and visualize economic trends of music events and transportation system

- Yutong Zhu: Data engineering and cleaning MTA and TLC data; data visualization to display the transportation pattern of MTA and TLC for NYC nightlife events；building EDA models of transportation systems and music venues.

Reference:

1. Carlsson-Szlezak, Philipp, et al. “Understanding the Economic Shock of Coronavirus.” *Harvard Business Review*, 27 Mar. 2020, hbr.org/2020/03/understanding-the-economic-shock-of-coronavirus.
2. “Lost My GIG NYC.” *Helping Freelancers in the Events Industry during the Coronavirus Crisis*, lostmygigny.com/.
3. “Creative Footprint Music, New York” *www.creative-footprint.org*, VibeLab
4. “NYC’s Nightlife Economy Impact, Assets, and Opportunities .” *Https://www1.Nyc.gov/Assets/Mome/Pdf/NYC_Nightlife_Economic_Impact_Report_2019_digital.Pdf*, Nyc.gov, 2019.
5. Zhang, Junbo, et al. “Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction.” *AAAI Conference on Artificial Intelligence*, Feb. 2017, www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14501/13964.
6. Zraick, Karen, and Sandra E. Garcia. “A List of What's Been Canceled Because of the Coronavirus.” *The New York Times*, The New York Times, 9 Mar. 2020, www.nytimes.com/article/cancelled-events-coronavirus.html.

Appendix I: List of Data Sources

Name	Data Sources	Notes（latest updates)
TLC Monthly Data	https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page	Dec 2019
Citibike Monthly Data	https://s3.amazonaws.com/tripdata/index.html	Feb 2020
MTA Real- time Data	http://web.mta.info/developers/developer-data-terms.html#data	March 30, 2020
NYC Taxi zone shape file	https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page	2019
NYC nightlife report	https://www1.nyc.gov/assets/mome/pdf/NYC_Nightlife_Economic_Impact_Report_2019_digital.pdf	2019
Creative Footprint NYC（music）	https://drive.google.com/drive/u/1/folders/1fKMt3mrJVU0vn5YVUeq10W88FufhwYfH	2019
Impact of Nightlife Business Closures	https://drive.google.com/file/d/19SzvU_WbyfM5nKhYgu0-l6ethZ-Gwj7Y/view?usp=sharing	Mar 2020
COVID-Arts/Events/Nightlife Survey	https://docs.google.com/spreadsheets/d/1UodAyPjxbDqPbCcxfU-FmU1NpP2xsH_m1PaF5UJ-nIk/edit#gid=0	Mar 2020
COVID-19 Daily Data	https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset#COVID19_line_list_data.csv	Daily

	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18
	Mar. 30	**Apr. 6**	**Apr. 13**	**Apr. 20**	**Apr. 27**	**May 4**	**May 11**	**May 18**	**May 25**	**Jun. 1**	**Jun. 8**	**Jun. 15**	**Jun. 22**	**Jun. 29**	**Jul. 6**	**Jul. 13**	**Jul. 20**	**Jul. 23**
Preparation & Data Acquisition																		
Structuring Weekly Working Log w/ Mentor																		
Literature Review																		
Proprietary datasets (Creative Footprint, NYC Mayor's Office of Nightlife)																		
Open data sources (TLC, MTA, Citibike, Census)																		
Scraped data sources (Eventbrite, Resident Advisor, Google Maps)																		
Outreach / Interviews (Govt. & Industry Stakeholders)																		
Data Analysis																		
Exploratory Data Analysis (Jupyter Notebooks)																		
Time Series Analysis for Taxis (Jupyter Notebook)																		
Time Series Analysis for MTA Turnstiles (Jupyter Notebook)																		
Time Series Analysis for CitiBike (Jupyter Notebook)																		
Spatial Analysis (Autocorrelation & other methods)																		
Data Visualization																		
Mapping nightlife hotspots (Venue Density Isopleth)																		
Mapping night time transit activity near nightlife hotspots (pre-COVID)																		
Mapping night time transit activity near nightlife hotspots (early COVID)																		
Mapping night time transit activity near nightlife hotspots (after closures)																		
Visualization of spatial autocorrelation analysis																		
Discussion of Findings & Implications																		
DUE: Second Report & Presentation					x (Apr. 29)													
DUE: Third Report & Presentation													x (Jun. 22)					
Finalization of Project GitHub Repo (with descriptive README)																		
Development of Project Website																		
Design Final Presentation																		
DUE: Finalize Technical Project Report																		x (Jul. 23)