**Lee’s Pizza – NYC Market Entry**

By: Li Sun

**Abstract**

The goal of this project is use linear regression to identify relationships between Cryptocurrency Price and its related factors that could contribute in fundamental analysis of Crypto and result in investment opportunities. I worked with 3 datasets focused on Circulating Supply, Market Capitalization, Social Media volumes, and OnChain Transactions. After running baseline regression models, applying

The goal of this project was to use exploratory data analysis to create market entry strategies into New York City for Lee’s Pizza, a brick & mortar pizza parlor that serves late-night pizzas. I worked primarily with the MTA turnstile dataset and complemented it with existing Yelp studies/data. After analyzing the foot traffic of potential pizza buyers in 6 targeted Manhattan neighborhoods and evaluating the pattern of foot traffic throughout the day, I recommended two neighborhoods for potential entry as well as expanding beyond Lee’s Pizza’s normal hours of operation to capture more pizza buyers*.*

**Design**

New York City (NYC) is historically the mecca for authentic, traditionally made pizza in the United States. It is therefore not uncommon for Pizza stores like Lee’s Pizza to desire opening a store in New York City. Because of Lee’s unfamiliarity with entering the NYC pizza market, the problem definition is two-fold:

* **Question #1 -** What Manhattan neighborhood(s) should Lee’s Pizza consider opening a store?
* **Question #2** - Should Lee’s Pizza expand beyond their normal hours of operation (8 pm - 4 am)?

Lee’s Pizza’s customers are typically nightlife party people who are hungry for grab-and-go pizza on the way to or back from Nightlife Establishments (e.g. bars, clubs). Nightlife Pizza buyers are driven by foot traffic. We can reasonably make two assumptions: 1) Healthy conversion rate of foot traffic to Pizza buyers, 2) Majority of late-night traffic of pizza buyers (NightLife partiers) takes the subway to or from their Nightlife Establishment.

**Data**

The dataset contains 15 weeks of 2019 MTA Turnstile Data (July 20 – Oct 26, 2019). Additionally, I manually recorded the number of late-night pizza stores/competitors per targeted neighborhood on Yelp.

**Algorithms**

Methodology & Steps

1. Used an existing Yelp study[[1]](#footnote-1) to first identify 6 Target Neighborhoods that had a high density of Nightlife Establishments (e.g. bars, clubs) in Manhattan.
   * Target Neighborhoods: East Village, West Village, Hell’s Kitchen, NoMad, Murray Hill, Lower East Side
2. Imported, cleaned, grouped, and aggregated 15 weeks of MTA turnpike data. Filtered out the 13 stations located in the 6 Neighborhoods. Added “Neighborhood” column to isolate the 6 target neighborhoods.
3. After cleaning the ENTRIES & EXITS data for negatives, outliers, and using shift(1) to create incremental 4HR entries & exits, I combined ENTRIES & EXITS to derive **“Foot Traffic”.**
4. Filtered by two 4HR TIME blocks to match Lee’s pizza’s hours of operation (8pm-4am).

Result

1. Grouped the dataset by the 6 Neighborhoods and took the mean of “Foot Traffic”. Converted “Average Foot Traffic” to daily average to derive **“Average Daily Foot Traffic (8pm-4am) per Neighborhood”.** Plotted this data in Bar Chart.
2. Paired **“Average Daily Foot Traffic (8pm-4am) per Neighborhood”** data with # of competitor Pizza stores per Neighborhood (Yelp) to reach conclusion to Question #1.
3. Un-filter dataset to include all 4HR TIME blocks. Cleaned TIME data and created a times series line chart measuring **“Average Foot Traffic” per 4HR TIME block** to reach conclusion to Question #2

**Tools**

Pandas for data manipulation, Matplotlib for visualization, SQL & Excel to lookup data, Google Map, Yelp

**Communication**

Slides and visuals are in PPT, submitted PDF.

1. <https://www1.nyc.gov/assets/mome/pdf/ESI-NYCEDC-Nightlife-Report-2018.pdf> (page 37, Exhibit 3.21) [↑](#footnote-ref-1)