12-10-2023

The Influence of COVID-19 on the Food Delivery Infrastructure

Kushagra Kushagra
Auburn University, kzk0104@auburn.edu

Ashish Gupta
Auburn University, azg0074@auburn.edu

Xiao Qin
Auburn University, xqin@auburn.edu

The Influence of COVID-19 on the Food Delivery Infrastructure

Research-in-Progress

Kushagra Kushagra
Auburn University
kzk0104@auburn.edu

Dr. Ashish Gupta
Auburn University
azg0074@auburn.edu

Dr. Xiao Qin
Auburn University
xqin@auburn.edu

Abstract

COVID-19 has a profound global impact on various industry sectors. It has led to significant changes in societal behavior. Social distancing made public spaces hazardous, shifting consumer habits, including purchasing and spending patterns. People have been driven toward online resources and delivery services, causing disruptions and impacting industries.

The study investigates and identifies new online food delivery patterns that emerged during COVID-19. We focus on the food delivery industry in a University town, integrating 183 restaurants to understand how e-commerce and consumer behavior with respect to restaurant food delivery changed from pre-COVID to the COVID-19 times. We use AI and machine learning techniques to collect and analyze data collected over three years. Findings suggest that new emerging patterns require adaption to the variability in the types of food that consumers are ordering, ordering times, delivery locations, etc. Such insights provide for resource planning and allocation decisions.

Keywords

COVID-19, Delivery, Machine-Learning, Segmentation, Food.

Background

The far-reaching impacts of the COVID-19 pandemic are not limited to the global population but also extend to various sectors such as healthcare, manufacturing, tourism, transportation, and supply chain, among others. COVID-19 has posed numerous challenges not just for the medical community but also for the business sector. The impact on health has also shown demonstrated effect on people's lifestyle. For example, Research has revealed a range of symptoms experienced by those affected, including fatigue, muscle or body aches, shortness of breath, and headaches (Fadaka et al., 2020). This has caused people to alter their lifestyles, which in turn altered customer buying habits. This was further amplified due to social distancing and the mask mandate in crowded areas. This resulted in a significant impact on businesses and required them to take steps to sustain survivability (Campbell et al., 2020).

The COVID-19 era witnessed considerable fluctuations, prompting the need to recalibrate and develop new models in the realm of E-commerce, specifically concerning sales and customer behavior. The E-commerce sector, owing to its vast size and wide variation in a multitude of factors such as food item preferences, order times, restaurant choices, order frequency, and delivery locations, can facilitate our understanding of shifts in consumer demand patterns. While extensive research has been conducted in the broader E-commerce domain, no prior studies have specifically addressed the food delivery industry within a university town or incorporated Machine Learning techniques, particularly within the context of a pandemic. A distinctive demographic makeup in a university town serves as motivation to construct models focused on food delivery services, with a predominant customer base consisting of students, faculty, and university staff. This company encompasses 183 restaurants and offers comprehensive delivery services. Various machine learning models, including heatmaps and K-means clustering, have been employed to analyze data spanning three years within the delivery industry (Kanungo et al., 2002; Marcus, 1998).

Our methodology encompasses the utilization of heatmaps, segmentation analysis, temporal analysis, association rule mining (ARM), and K-means clustering. Heat maps are generated using various metrics such as restaurant-wise daily sales, daily counts versus the company's daily sales and counts, and hourly sales versus the company's daily sales. K-means clustering is employed to identify customer groupings based on their ordering frequency. ARM is applied to menu items across all restaurants to uncover related items commonly ordered together. Subsequently, we delve deeper into the analysis with timeline graphs and boxplots to illustrate the fluctuations and overall statistical impact of COVID-19 on the delivery company.

Two key questions explored in this study are a) What is the impact of COVID-19 on food delivery services? And, b) in what ways such an impact is manifested by external events such as major social events and weather patterns in the area? This could potentially help us gain a deeper understanding of shifts in food delivery patterns during the pandemic, which could aid in effectively managing customer demand and resources. Our investigation begins by scrutinizing the sales and performance of a relatively new food delivery company, commencing on January 1, 2017. The "normal" sales period extends until March 1, 2020, after which the COVID-19 impact is considered. This company, established within the past seven years, has demonstrated a significant sales upsurge during this timeframe.

Methodology

Data Acquisition

The data we gathered originates from a food delivery company and was obtained through a series of HTTP API requests facilitated by Python and the Postman application. We utilized POST requests, leveraging an account endowed with the necessary privileges provided by the company, which granted us access to API endpoints for exporting data in JSON format. Subsequently, we employed GET requests to retrieve data pertaining to three distinct entities: customers, restaurants, and orders. The data extraction period spans from January 1, 2017, to May 14, 2020. The process of acquiring this data unfolds in two phases. In the initial phase, we acquired data encompassing approximately 60,000 orders spanning from January 1, 2017, to May 14, 2020. The API requests yielded information on 2,000 orders per request. After executing each request, a total of 35 times, we compiled the resulting data into a single comprehensive CSV file for further extraction and analysis. Our data acquisition strategy involved multiple iterations of API requests to encompass the entire dataset for the specified time frame of customer orders. This approach was necessitated by server limitations, which restricted each JSON file to a maximum of 2,000 orders. In total, this extraction process yielded a dataset comprising 70,155 orders.

This dataset encompasses a wealth of information, including details about customers, their locations, ordered items, restaurants, and local events. The use of the JSON format greatly facilitated data extraction, as variables were organized hierarchically. For instance, within the "address" variable, a list of sub-variables related to the address, such as street and zip code, were neatly structured. This hierarchical arrangement within the JSON format also simplified the handling of closely related data, aiding in the determination of delivery destinations and any time-location dependencies. Similarly, we had access to information regarding customer addresses, street names. We supplemented this data with an externally compiled dataset of social events and weather conditions within the areas served by the delivery service.

In our study, we concentrate on orders, extracting data related to sales revenue, ordered menu items, delivery zones, and the restaurants fulfilling each order since the company's inception. To safeguard customer data privacy, we exclude other parameters provided by the company's dataset.

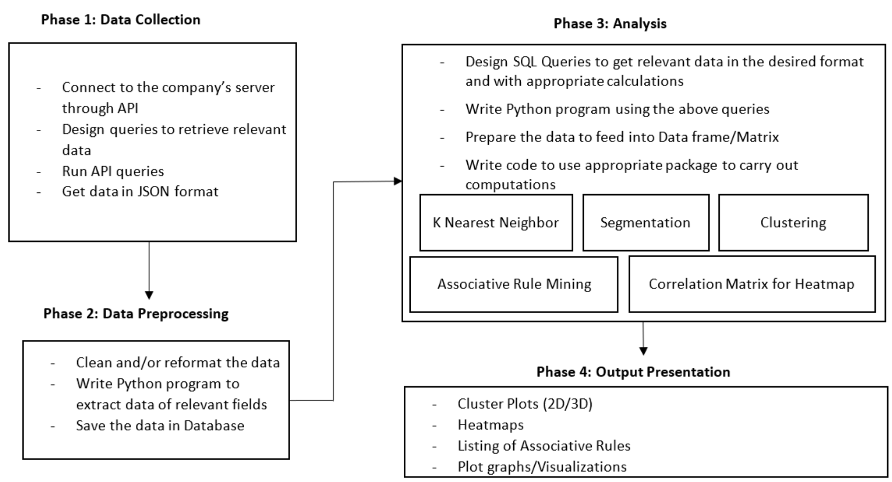

Figure 1. Bird's eye-view of the process used for the report.

Data Preprocessing and Analysis

As shown in Figure 1, we employ heatmaps to visually represent correlation matrices associated with company and restaurant-specific sales that have a particular focus on the topmost cluster of restaurant sales. We also extend this approach to examine the correlation between the company's order count and the order count of individual restaurants. Similarly, we create heatmaps to analyze revenue patterns across different delivery zones and menu item IDs. These heatmaps provide valuable insights, helping the company prioritize its efforts on restaurants, delivery zone IDs, and menu item IDs that significantly impact revenue and order volume. Simultaneously, they shed light on areas where restructuring may be necessary. In each case, the first column/row represents the company's revenue or order count. A high positive correlation indicates the entities that warrant attention for future improvements.

Customer segmentation plays a pivotal role in tailoring interactions with customers, a practice commonly undertaken by sellers but often without a systematic approach. Machine Learning offers a solution to this by addressing customer segmentation through unsupervised learning techniques. Among these techniques, K-Means clustering is the most widely used, and therefore, we employ it in this study to perform customer segmentation. Unsupervised learning, a category of machine learning, deals with data lacking human-labeled annotations. In the context of our dataset, which is subject to privacy considerations, we have access to order-related data that includes sales values, order counts, menu items, menu item values, delivery zones, weather periods, and customer street addresses. None of these parameters come with predefined labels.

Consequently, we can apply various unsupervised learning algorithms such as clustering, anomaly detection, neural networks, among others. In our specific scenario, clustering algorithms are the most fitting because our objective is to segment customers by defining boundaries around clusters based on the aforementioned parameters.

Expected Contributions

Identifying the factors that influence a food delivery company's revenue during both Pre-COVID-19 and COVID-19 periods is crucial for similar businesses seeking to operate efficiently and profitably. We delve into these factors, primarily sourced from the demand side, which hinges on key customers and their ordering behaviors. Our analysis takes into account the socio-economic dynamics within the company's local environment by accounting for major social events and weather patterns as these could influence the delivery patterns. While these factors may vary based on location, the fundamental methodology for identifying these factors remains consistent.

The food delivery company operates a software-based platform that aggregates local restaurants onto a single platform, facilitating food orders for customers. These orders are picked up and delivered by the company's delivery personnel. While a single restaurant may not have the capacity to offer such a service, a food delivery company can efficiently handle ordering and delivery by consolidating multiple restaurants.

One technique employed in this research involves analyzing the temporal variation in the company's sales revenue. One noteworthy finding is that there was no decline in sales revenue during the COVID-19 period, contrary to the common expectation of a decrease. One potential explanation is that during the COVID-19 lockdowns, people relied more heavily on food delivery services, maintaining sales at a level similar to the pre-pandemic period, despite adverse economic conditions.

Another technique we employ is the use of correlation-based heatmaps. We utilize this method to create heatmaps for various metrics, including hourly sales revenue versus the company's daily sales revenue, delivery zone-specific daily sales revenue versus the company's daily sales revenue, and restaurant-specific daily sales revenue versus the company's daily sales revenue. These heatmaps reveal intriguing insights, such as the highest revenue-generating hourly slot being from 6 p.m. to 7 p.m.

Work in Progress- Results are presently undergoing analysis and will be presented at the symposium.

References

Campbell, M. C., Inman, J. J., Kirmani, A., & Price, L. L. (2020). In Times of Trouble: A Framework for Understanding Consumers' Responses to Threats. *J Consum Res*. https://doi.org/10.1093/jcr/ucaa036

Fadaka, A. O., Sibuyi, N. R. S., Adewale, O. B., Bakare, O. O., Akanbi, M. O., Klein, A., Madiehe, A. M., & Meyer, M. (2020). Understanding the epidemiology, pathophysiology, diagnosis and management of SARS-CoV-2. *J Int Med Res*, *48*(8), 300060520949077. https://doi.org/10.1177/0300060520949077

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 881-892. https://doi.org/10.1109/TPAMI.2002.1017616

Marcus, C. (1998). A practical yet meaningful approach to customer segmentation. *Journal of Consumer Marketing*, *15*(5), 494-504. https://doi.org/10.1108/07363769810235974