Design: How does a predictive model trained on social media trends perform on grocery store product sales?

Team number: Team C
Team Members: Benjamin He, Kevin Yi, Steven Au
Date: Aug 6, 2024

Overview

The Problem

Due to rapidly changing social media trends, grocery stores have the challenge of adapting to inventory needs outside their traditional retail practices and predictive models. That is, in the modern age, many people use social media such as Instagram where the food and drink category accounts for 43% of all the users' interests.[1] What that entails is people who are exposed to these posts become inspired to learn, emulate, and recreate these dishes on their own by buying those grocery items. Unfortunately, traditional inventory prediction models do not fully account for these shifts in consumer behaviors to determine inventory, leading to situations of overstocking and understocking. These situations cause increases in food waste, lost revenues, and inventory control issues. This problem highlights the need for a new predictive model to better align with consumer demands.

For our research, we will conduct an experimental study comparing our new predictive model that factors in social media data against the existing inventory prediction models currently in use at Safeway in San Francisco's Bay Area. The range of these target products will be determined by the social media trends outside of the specific control of this research. However, all these products measured should be available as part of the grocery stores' inventory. Ultimately, we aim to determine the effectiveness of our new predictive model in predicting grocery store product sales and revenue that traditional inventory prediction models cannot sufficiently capture.

Intended Audience

Our intended audience includes grocery store market researchers, supply chain managers, and executives. We believe that by incorporating social media data into the existing predictive model - effectively an inventory management algorithm - grocery stores can optimize product sales, reduce waste, and improve sales efficiency.

Existing Literature

- Clausen and Li state that research on big data-driven models for inventory management based on multi-variables is in its infancy. They created a dynamic inventory model taking in big data to prove that it was better at predicting orange juice inventory levels than models only relying on historical data. [2]

- Hajli discusses how social media interaction and marketing have a significant impact on building customer trust in ecommerce stores, which then influences customers to be more likely to purchase from the store. [3]
    - The results of the study show that social media influences consumer behavior, suggesting that factoring in social media trends will help grocery stores sell products more efficiently.
- Clark and Lee examine the benefits of the Continuous Replenishment Process (CRP) within the USA grocery industry where those retailers transmit inventory and sales data to suppliers. [4]
    - While their study focuses on traditional data models, it highlights the importance of using data to enhance inventory management while suggesting that we can incorporate social media data into our predictive model to facilitate predicting grocery store inventory product sales.

Anticipated Impact

The direct anticipated impact of this study would be to inform grocery stores about the effect on sales revenue when using a predictive model for inventory management. Research shows that social media trends influence consumer behavior. Extending this, we believe that considering social media as a part of predictive inventory models would improve inventory management efficiency and thus sales revenue for grocery stores.

If this belief is correct, further impacts for grocery stores would include the creation of better predictive models for forecasting grocery product demand. This would bring increased revenue through social media campaigns, greater profit margins due to less product waste, and increased customer satisfaction/brand loyalty from stocking the right products at the right times, all characteristics that enable any grocery store to thrive.

At a greater scale, more efficient inventory management would also encourage optimization of the entire supply chain for these grocery stores, leading to lowered environmental impact from the stores' operations as fewer resources go to waste. With grocery stores being both critical and widespread establishments in numerous nations, the beneficial effects of the knowledge gained from this study can be global.

Research Question

Main Research Question:

Do stores that use a predictive inventory model trained on social media data have higher product sales relative to the stores that do not?

Sub-Questions

- Will stores using the new model see an increase in revenue compared to before?

- How does our model perform differently for stores located in different types of demographic areas?
- How does our new model fare compared to different types of categories of food? Shelf Stable vs Unstable food. Junk food vs healthy food.
- Which grocery product categories benefit the most from using a social media data model?
- How long must a grocery store use the predictive inventory model to begin noticing improvements in sales, if at all? (month, quarter, year, etc.)

Definitions

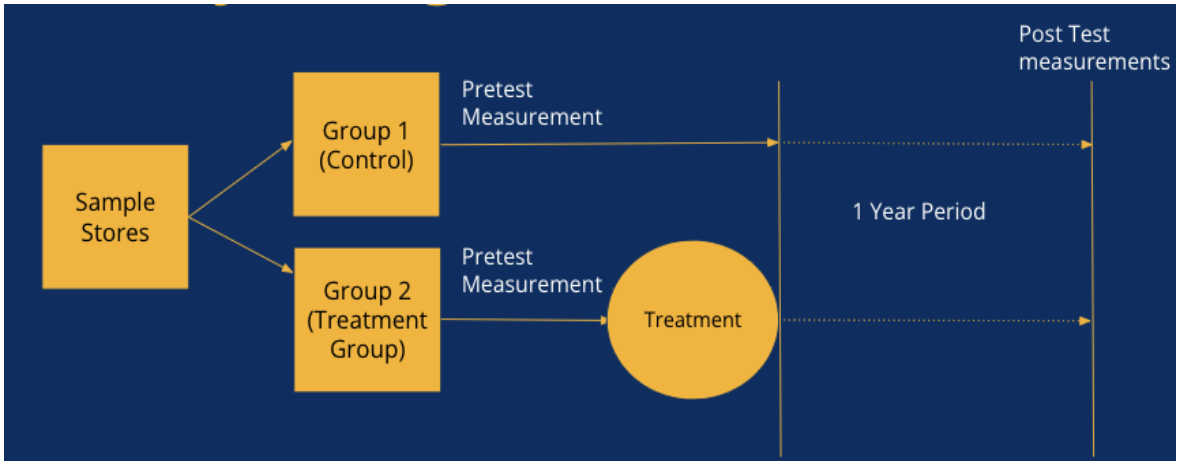
- *Social media trends*: Activities, lifestyle habits, and/or items that are being highly discussed on social media platforms such as TikTok, Instagram, and LinkedIn.
- *Product Sales*: Grocery product items that are sold at the stores.
- *Product/Item*: Note that these two words are used interchangeably throughout. This is effectively what the item/product that is affected and observed in this research.
- *Predictive Model*: A model, algorithm, or procedure that would determine the level of inventory for grocery stores to stock based on how it is implemented/programmed.
- *Demographic Area*: The local region that a store in our sample is located, with cultural/socioeconomic characteristics–probably causing certain buying habits–that are relatively unique to this region.
- *Staple foods/products*: Essential goods that consumers buy regularly regardless of trends. These products have a consistent demand because they are fundamental to daily living.
- *Trend-sensitive foods/products*: Items whose demand fluctuates based on current consumer trends, health fads, seasonal popularity, or social media influence. These products often see spikes in demand driven by trends but may not maintain consistent sales over time.

Study Design

- We will use our model that has already been created that factors in social media data. Our model's performance in predicting historical data has been shown to perform better than existing models in predicting grocery product sales. We will conduct an experimental study to see if our new model fares better than the current model in use by the grocery store by measuring grocery product sales.
- Outcome: Our model will be deemed successful enough to roll out to all Safeways if the increase in grocery product sales results in a 5% increase in revenue. There needs to be a high enough ROI to offset the new model and

operational/maintenance costs to include the cost for engineers and cloud computing for the models.

- Study overview:



- The assignment will include 2 groups: One control, and one Treatment. Random stratified sampling will be taken to assign our grocery stores to our control and treatment groups. Each group will be chosen according to a stratified sampling method to ensure that we have grocery stores in both groups that are representative of the sample. Some factors to consider are the demographics of the area that the grocery store serves such as income level, ethnicity, educational level, and age. Other factors include store proximity to other grocery stores of the same or different company and the store's existing gross revenue and profits. We will conduct a power analysis to determine our control and sample size.
    - The control group will continue to use the existing inventory prediction model that they have been using. All Safeway stores are currently using the same model.
    - The treatment group will use our new prediction model to set product inventory levels. The model will be run weekly to capture recent social media data.
- Pretest Assessment: We will take our store's total grocery product sales for the past year as our baseline, with measurements taken quarterly and at the end of the year. We will also normalize the historical data to exclude events such as civil unrest, extreme weather, and global events that may introduce confounding external variables.
- Post test: After treatment, we will take measures of our sample stores' quarterly grocery product sales for a year (to capture all seasons) and compare how the control and treatment groups fared.

Data

Covariates

The following covariates will be part of our independent variable data used. These include demographic, geographic, store, and time (season) that will be collected on a recurring basis for each store. The data will be matched to the store throughout the research. This serves as the existing data that will be continuously updated for our research as described in the below sections. Other covariates that act as confounding variables such as market power, store loyalty, and product brand quality are covert factors beyond the scope of this study. Weather will be embedded as part of the time (season) instead of using specifics such as temperatures.

Data Collection

All data will continuously be extracted, cleaned, and transformed as appropriate and as described in the study design.

- Grocery Store Data:
    - We will use data as described in the study design and sample sections on a recurring basis.
    - The data will come from the stores as described in the sample section.
    - Data will be collected every quarter for the grocery store item sales.
- Social Media Data
    - The data sources will be from TikTok, Meta's FaceBook, Instagram, and Threads, and X - formerly Twitter.
    - Data will be collected and sent to our new predictive model weekly.
        - The new model will use only regional social media data as described in our sample section's sampling frame.
        - Data will be standardized accordingly per each platform.
    - Sentiment analysis on hashtags, mentions, emojis, and influencers' posts will be conducted. Trending food cooking videos and promotional information will be captured too.
        - All these values will determine the interaction trend of the product(s) to support our model.
        - Data will be sanitized accordingly to account for both internal and external validity. Please see the potential risks section for details.

Intervention

The intervention will be applied on the basis as described in the study design.

- This will simply be determined as the new predictive model's effectiveness with the continually recurring stream of social media data. The specifics of the social media data metrics to determine the item(s) are described in the Intervention variable section.

Sample

- Sampling Frame: Our grocery store chain (Safeway) in San Francisco's Bay Area.
- Sample: Our sample size will be selected from our power analysis and using a stratified sampling method on demographic areas to ensure consumer behavior representation from various demographics. Our selected stores in San Francisco's Bay Area are to be determined to not be affected by significant factors, which include ones that are not scheduled to undergo major renovations or have an existing major personnel shortage.
- Treatment: Stores that use our new inventory prediction model.
- Control: Stores that use their default inventory model.
- Sample groups /randomization: Stratified sampling based on geographic, time, and demographic information the store serves.

Hypotheses

For the below hypotheses, H0 represents the null hypothesis and H1 represents the alternative hypothesis of a given hypothesis set. The subscripts represent each separate set of hypotheses for testing a research insight we aim to understand based on our research questions. The rejection criteria for each of the below null hypotheses is a p-value of under 0.05, which is the conventional standard.

Primary Hypotheses

$H0_p$: The mean difference in product sales between stores that use a predictive inventory model trained on social media data and stores that do not is less than 5%.

$H1_p$: The mean difference in product sales between stores that use a predictive inventory model trained on social media data and stores that do not is 5% or greater.

The primary hypotheses are tested during the post test phase of our study design.

Sub-question Hypotheses

Because the data we collect is product-level, we can use the same set of data to conduct more granular analyses for researching our sub-questions during the post test phase. Our additional subquestion hypotheses are:

$H0_1$: Stores using the new predictive inventory model do not see an increase in revenue compared to before using the model (**$H1_1$**: they do see an increase).

$H0_2$: The model's performance does not differ for stores in different demographic areas (**$H1_2$**: it does differ).

$H0_3$: No specific grocery product categories benefit more from using the model (**$H1_3$**: trend-sensitive products benefit more).

The other sub-questions we will investigate are more exploratory in nature, and we have not formed hypotheses on them. However, they will also be explored using our data from the post test phase.

Variables

Our study will involve several independent variables and an intervention to compare the efficacy of our new predictive inventory model that incorporates social media data against the existing models currently available for the outcome. These variables all cater to the vicinity of each grocery store.

Independent Variables:

- Demography
    - Income level: This is going to be the average income of the people living in the vicinity of the grocery store. This determines the purchasing power of the customers.
    - Ethnicity: This would be the distribution of ethnicities in the grocery store's vicinity. Different ethnicities strongly prefer certain foods such as rice or grain-based foods.
    - Education: This is the average education level of the people in the grocery store’s vicinity. Different education levels can influence dietary preferences.
    - Age: This is the average age range distribution of the people in the grocery store’s vicinity. Different age groups may prefer certain foods such as softer/liquid foods.
- Geography
    - Population Density: The number of people living in the given grocery store vicinity. Density determines the quantity of foot traffic to a certain store.
    - Nearest Neighboring Store (Proximity): The distance between the next closest grocery store (Same chain or competitor), determined through a simple distance formula. Neighboring stores can affect sales due to competition and foot traffic.
- Store
    - Store Size: The size of the grocery store in square feet. This categorizes the amount of inventory and foot traffic that a store can physically accommodate such as small, normal, and large stores.

    - Items: These are all the item listings that a store has for its inventory. This determines what products the store has for sale per the received data.
- Time
    - Season: The types of items that a grocery store can practically offer in their inventory at the time of year. This affects the purchasing of certain available items due to seasonality.
    - Time of Day: This is the 24-hour cycle of time. Different times of day affect the foot traffic at the grocery store.
    - Day of week: The 7-day cycle of a week. Different days of the week change the foot traffic of the grocery store, such as weekdays and weekends or holidays.

Intervention:

- Social Media Engagement: This is the amount of exposure and engagement trends measured as interactions through the form of marketing and promotions per item(s). The specific usages of the intervention variable are described in our data section to determine the affected item's trend. This intervention permits the analysis of the new prediction model to account for its effectiveness on the product(s) sales.

Outcomes (Dependent Variable):

- Product Sales: This is the sale of the grocery item(s) listed per each store. The outcome of each item when measured against the predictive and traditional models.

Statistical Methods

For our statistical analysis, we will use the two-sample t-test. A two-sample t-test works when you have two independent samples (we have our control and treatment group). It also works when you have a continuous variable (we have our measured value of grocery product sales for all stores). [5] Checking the t-values (calculation of t-value will be adjusted when testing $H0_p$) and the corresponding p-value will show us if there is a significant difference between the mean values of the total grocery sales for our treatment and control group.

Potential Risks

- Risk to non-interference: The effect of the treatment or control groups can potentially spill over into each other.
    - As an example, if the stock of a store within the control store is empty, customers may buy from a treatment group store and inflate the inventory turnover rate of the treatment group. The opposite could happen as well.

Customers could also go to another store that is not part of our research or other stores can start promotions/sales that impact the stores in our study.

- Individual stores could see increased foot traffic from nearby grocery stores closing down (reduced competition) and form increased population (A new development/neighborhood could be built).
- Risk to excludability: Individual stores could be impacted by new/improved or poor in-store management led at the individual store level.
- External validity risk factors (interaction of setting and treatment) such as inclement weather, epidemics, and public unrest could negatively impact store performance.
- Risk to internal validity–History: Major events may also artificially increase store performance at certain periods.
- Algorithmic Bias: Inventory items over or under tuned by the new model may favor or disfavor certain items trending for unintended reasons, such as memes or social competition challenges. For example, the “cinnamon challenge” where people consumed cinnamon in one spoonful with the risk of death.

Deliverables

- One year of pretest data gathering (total grocery product sales) for baseline. Measurements of sales will be gathered every quarter and aggregated at the end of the year.
- In the following year, we will conduct our research with the grocery store chain (Safeway) using our new predictive model against the existing predictive models.
    - Throughout the research, assess the intervention with social media data incorporated in the new predictive model.
- Final Deliverable: Presentation and report on the study outcomes. We will also provide a recommendation on further directions for study, or if the new model should be rolled back or implemented across all Safeway stores. Lastly, we will provide an in-depth analysis of the sales outcome and operational/maintenance cost of the newly tested model compared to the existing model.

References

1. The Best Chef. The Power of Social Media Platforms - The Best Chef. The Best Chef. Published July 17, 2023. Accessed August 3, 2024. https://thebestchefawards.com/2023/07/17/the-power-of-social-media-platforms/
2. Johan, Li H. Big data driven order-up-to level model: Application of machine learning. *Computers & Operations Research*. 2022;139:105641-105641. doi:https://doi.org/10.1016/j.cor.2021.105641
3. M. Nick Hajli. A study of the impact of social media on consumers. *International Journal of Market Research*. 2014;56(3):387-404. doi:https://doi.org/10.2501/ijmr-2014-025
4. Clark TH, Ho Guen Lee. Performance, interdependence and coordination in business-to-business electronic commerce and supply chain management. *Information Technology and Management*. 2000;1(1/2):85-105. doi:https://doi.org/10.1023/a:1019108621684
5. Xu M, Fralick D, Zheng JZ, Wang B, Tu XM, Feng C. The Differences and Similarities Between Two-Sample T-Test and Paired T-Test. *PubMed*. 2017;29(3):184-188. doi:https://doi.org/10.11919/j.issn.1002-0829.217070

Statements of Contribution

- Benjamin He:
    - I contributed directly to anticipated impact, hypotheses, and references. However, as a group we essentially discussed all sections, bouncing ideas off of each other for study design and data sections. During our live working sessions, I contributed to existing literature, sub-questions, potential risks, definitions and deliverables in collaboration with my teammates. I had a great time working with Kevin and Steven. Our communication flowed well and resulted in an efficient process of meeting to decide work, working asynchronously, and setting deadlines throughout the project to ensure we stay on track without leaving things to the last minute. We struggled with determining good performance metrics to measure for grocery stores because none of us had a background in the consumer goods space. If I were to do this experiment again, I would've researched grocery/retail performance metrics more beforehand to find additional useful dependent variables apart from product sales.
- Kevin Yi:
    - I mainly contributed to the study design and statistical methods. As part of the team collaboration, I contributed to the research question, sub question, data, sample, variable definitions, and risks. Our group experience was exceptional. We had efficient team meetings to set expectations and discuss main topics before diving into async work and

then coming back together to provide feedback. Would definitely work with Ben and Steven in the future if an opportunity arose. In the future, I would spend more time trying to understand grocery store's KPI and on what benchmarks product predictive models are compared to. This would allow us to make a more informed research question.

- Steven Au:
    - I contributed to revising several sections of the report including the data section based on the feedback and proofread to refocus. The overview problem statement was condensed per the feedback received and I added the section on the variables based on all the information needed for our research. Unchanged with the existing literature of reference 4. Great team! Everyone was motivated to get things moving along knowing that working on this for the last minute may cause significant issues. If we have to do this again, hopefully, we may work together again and perhaps reuse or generate a similar version of our main research question for a future class down the line. Based on what we have learned from this class, the right main research question is the most important component of the overall design.