A CHANGING ENVIRONMENT FOR THE CHECKOUT LANE

SARAH BRABEC, SARAH FIRESTONE, BRENNAN MADIAR, LUCAS MILEWSKI
DUSTIN RABIN, GAYTHRI SRIRAM, EMILY WEBER, SCOTT WINNER

PURDUE UNIVERSITY | 84.51°

Our Project

Who is 84.51°?
- Data analytics and marketing company based in Cincinnati, Ohio
- Provide strategy and insights to The Kroger Company
- The Kroger Company is a retail company also based in Cincinnati.

What are our project's goals?
- Increase the efficiency of the grocery store checkout lane
- Make a list of the products that will generate the most profit

What are the limitations of the project?
- Data provided to us only has products that were already in the checkout lane, so we don't have the ability to add new products to the checkout lane
- No specific dimensions for the area that products can take up

How will we accomplish our goal?
- Splitting into three teams to analyze the data based off of Household, Transactions, and Regional data sets
- Integrated our findings to take a deeper look into pricing, basket penetration, and household penetration.
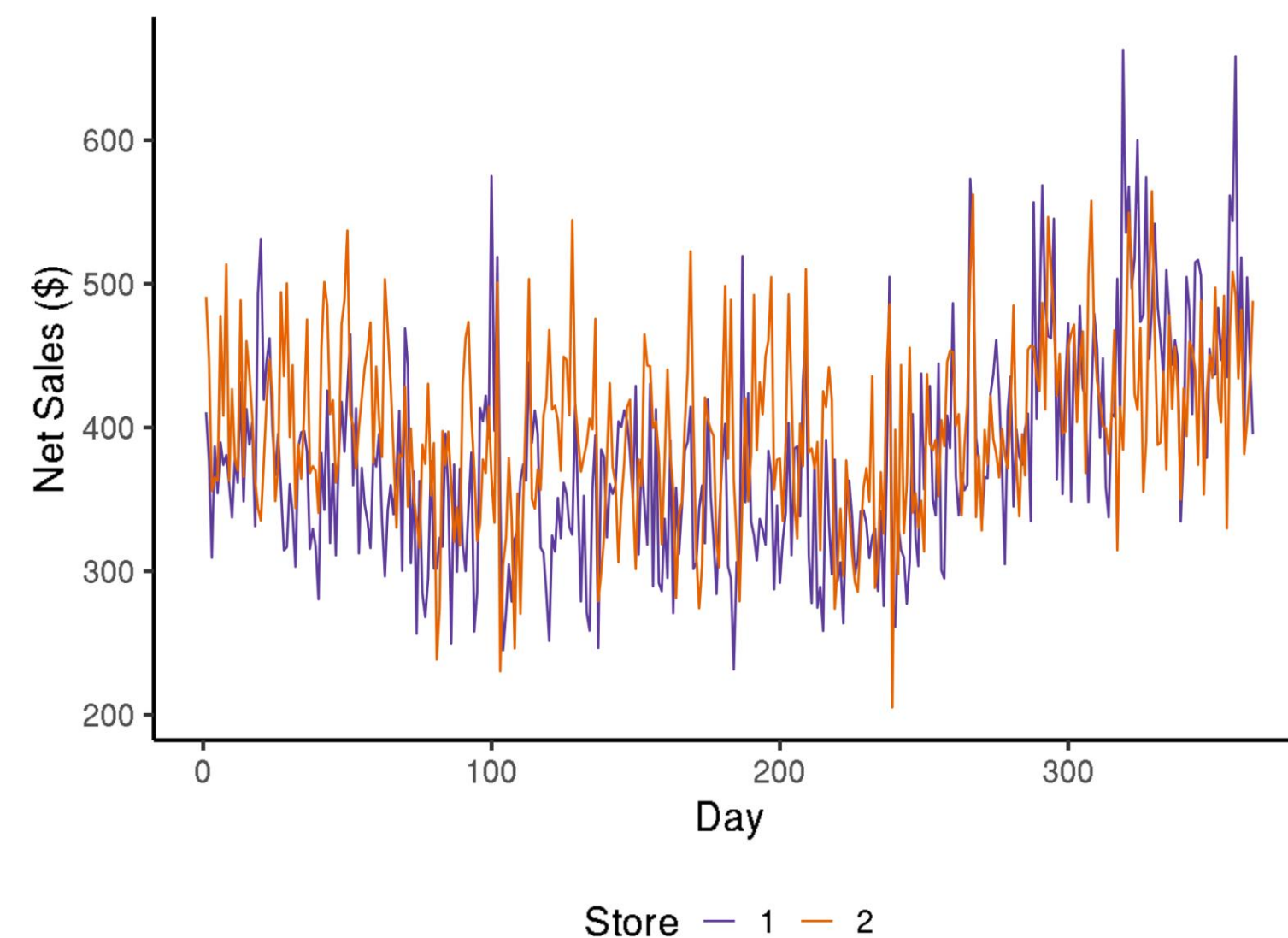- See how those three statistics vary based off Household, Transactions, and Regional data sets

Figure 1 Net sales per day for Stores 1 and 2

References:
1: Wang, Y., Qi, X., & Davison, B.D. (2007). Standing on the Shoulders of Giants Ranking by Combining Multiple Sources.
2: Optimizing Front End Checkout Merchandising. *FMI2012*.
http://www.ipda.org/DownLoadFiles/Final-EDU-slides-FMI-project%204_19_2012%20compressed.pdf

Figure 2
Store 1
Values are ranked UPC numbers based on each metric

HH PEN	TOP HH HH PEN	BASKET PEN	TOP HH BASKET PEN	NET SALES	TOP HH NET SALES
72	13	52	13	13	13
13	138	23	52	72	72
52	120	120	72	52	52
87	90	72	87	87	43
19	87	87	90	43	87

Figure 3
Store 2
Values are ranked UPC numbers based on each metric

HH PEN	TOP HH HH PEN	BASKET PEN	TOP HH BASKET PEN	NET SALES	TOP HH NET SALES
72	13	52	13	13	13
13	138	23	52	72	72
52	120	120	72	52	52
87	90	72	87	87	43
19	87	87	90	43	87

Figure 4
Store 1
Final Rank

UPC	NORM SUM	DESCRIPTION
72	2.900543	Energy Drink
13	2.862745	Energy Drink
43	2.652020	Energy Drink
120	2.582329	Soda
52	2.476730	Soda

Figure 5
Store 2
Final Rank

UPC	NORM SUM	DESCRIPTION
13	2.726983	Energy Drink
120	2.636281	Soda
72	2.549443	Energy Drink
40	2.510571	Soda
52	2.509434	Soda

Our Data

- Evaluated store, product, household, and transaction datasets
- Store Dataset: Store ID, Zip Code, and Region
- Product Dataset: Product ID, Description, Size, Manufacturer, and what category it falls under
- Household Dataset: Loyalty, Household's Demographics, How many people, Their income bracket, and Information about their main shopper.
- Transaction Dataset: Information about each transaction, Which products were bought, How many units were bought, Which household and how much it cost.

Datasets we created:
- Unique Transactions: Dataset to show each unique transaction since there was overlap sometimes in the Transactions Dataset
- Basket Penetration: Dataset to show the basket penetration of every product at every store
- Price Dataset: Shows the average price of every product since price can vary depending on the store
- Top Households Dataset: Show our households with the most activity

R Studio

Personalizing and Ranking

Personalizing to Stores vs Online:
We personalized on a store level. We would like to also have looked at personalization at an online level, but the data we were given was better suited for a store personalization.

Store Personalization:
- Allows different region buying patterns to be considered
- Meets customers from that store needs based on past buying habits
- Don't know who is walking through the checkout lane when a purchase is made

Online Personalization:
- If a user is logged in, we know who is checking out
- Allows for more accurate personalization to an individual
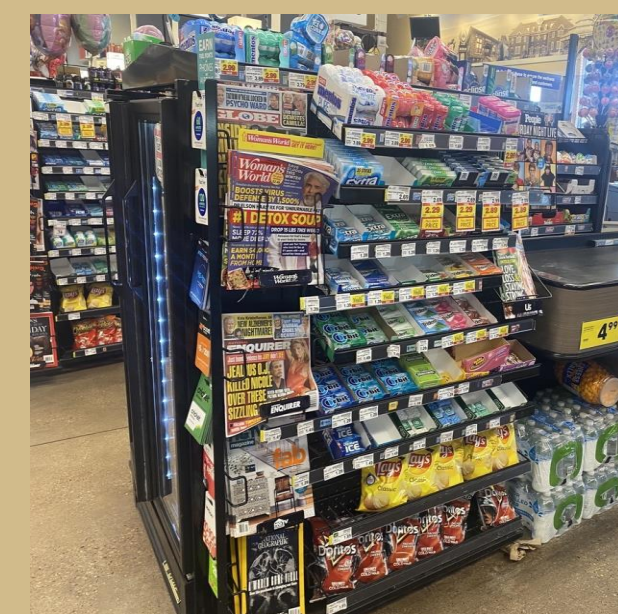- Online platform can change and recommend more options based on purchases
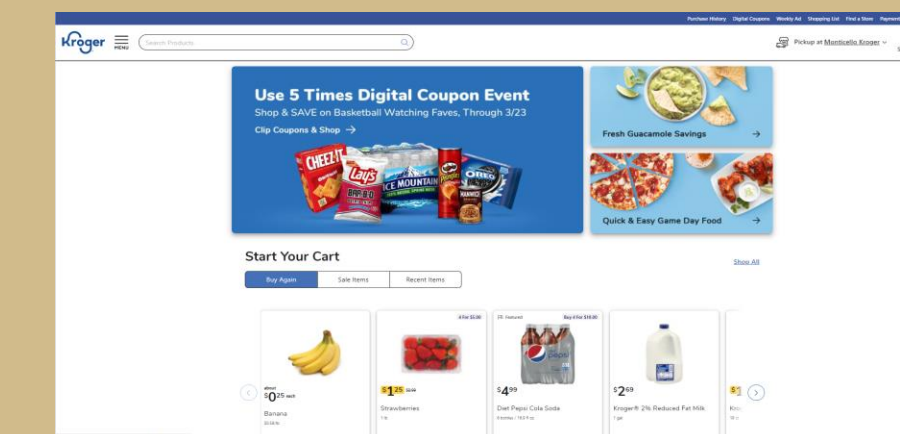

Figure 6 shows a store checkout lane

Figure 7 shows an online checkout experience

Ranking Products:
We ranked products a couple different ways. We then merged these rankings to find the top products that had the highest scores. The NormRank ranking formula was used to merge the rankings.

Net Sales:
- Net Sales were calculated for each product
- A Higher Net Sales made the products rank higher

Household Penetration (HH PEN):
- Compute the percentage of households that purchased each item
- A higher percentage is better, and ranked the item higher

Basket Penetration (BASKET PEN):
- Looked at the products most common in each basket
- The greater number of times a product showed up the higher the score

Top Households (TOP HH)
- Find the top 20% of households for each store that contribute 80% of the visits
- Recalculate the above rankings looking only at transactions from the top households

NormRank[1]:
- NormRank Formula: $(\text{Max rank} - \text{Product rank}) / \text{Max rank}$
- Each product after being ranked by the ways above was put into this formula
- Gave each product a final score
- Each store was given a unique set of Products (upc). See **Figure 4** and **Figure 5**

Conclusion

How to Place Products on the Shelves:
- We ranked the products using "Norm Rank"
  - Combined ranking: Constraints (size, ranking, space...)
- Figure 8: Start placing items where people look the most (eye tracking). Swap products based on aesthetics after that initial placement.
  - Place top ranked products on the middle shelves
  - Place lower ranked products on all the corners/edges
  - Place similarly colored items together
  - Place top ranked drinks on the center shelf
  - Place drinks that are popular with kids near the bottom shelf
  - Overall, consider each product's size to create a visually pleasing layout which will generate the most sales.
- If there is a tie:
 - Prioritize product aesthetics (constraints, colors, etc.)

Figure 8: Main side eye tracking[2]
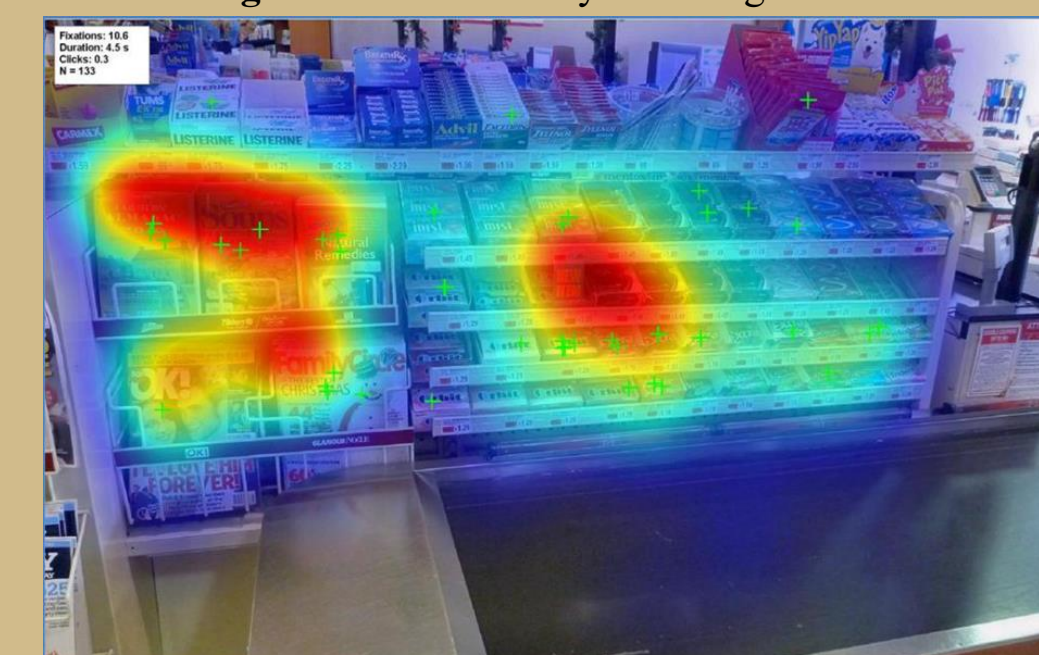
Next Steps:
- Split our data into two sections to see if there is a difference between:
  - Pre-COVID-19
  - Once COVID-19 started
- Test and implement our predicting ranks and see how successful they are in a real 84.51° store
- Compare stores with the new vs old method to test functionality and efficiency
- Compare loyal vs non-loyal purchases in the top 20% of households

Figure 9: Our hypothetical sample shelf
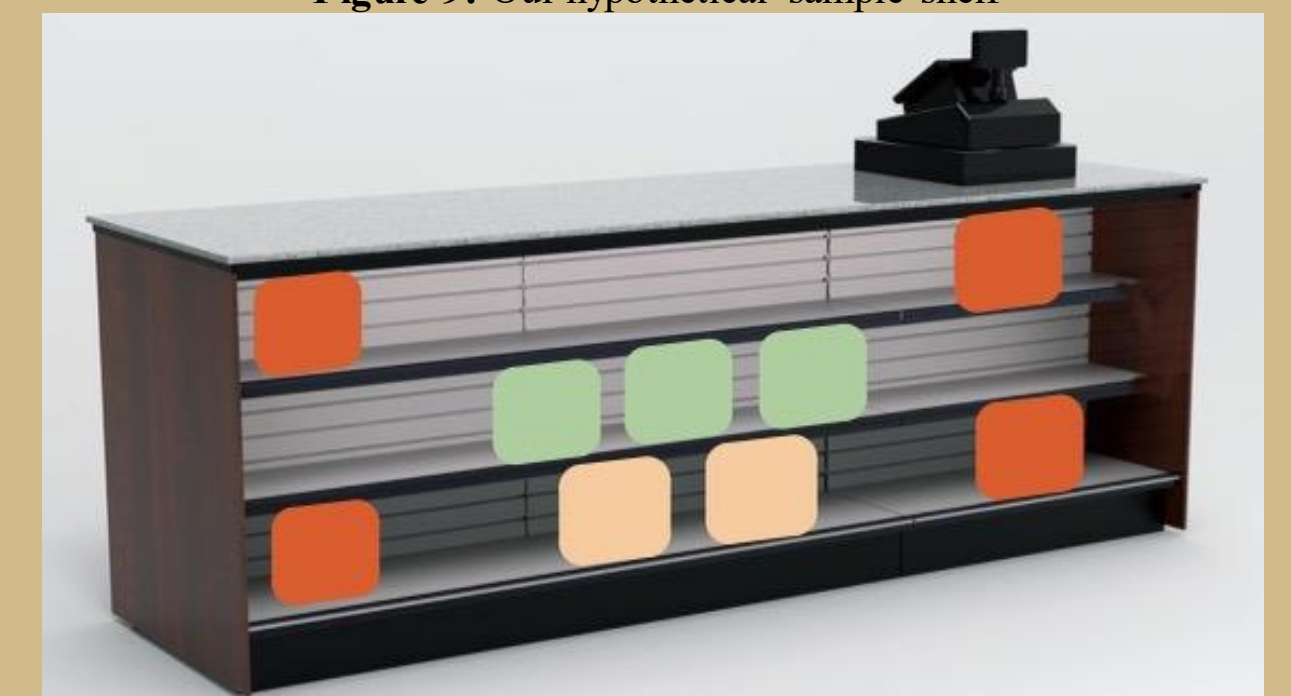
- Low Ranked Products
- High Ranked Products
- Popular Kid Products