Carnegie Mellon University
Tepper School of Business

INCREASING TRANSACTION SIZE THROUGH DATA ANALYTICS

MS Business Analytic Program

Christian Meyer,
Haider Aly Reza,
Nico Roussel,
Paul Lim

ABSTRACT

How can a traditional retailer motivate customers to increase their basket size? The answer to this question may be hiding in Busy Beaver's database of historical customer transactions. Included in this vast dataset is information about how customers responded to prior promotional campaigns, which primarily take the form of biweekly advertisements that promote various discounted products. The key question is how effective these promotions are at driving incremental business. Moreover, are there opportunities for improvement by adjusting product selection and/or discount levels? Finally, how should seasonality influence the design of the promotions?

The approach chosen to address these questions was to apply modern machine learning techniques to understand the relationship between the various characteristics of a promotion and their corresponding impact on customer purchase behaviors. This approach enabled the identification of key insights that will likely improve management's ability to design their biweekly advertisements. Ultimately, these insights formed the basis for a proof-of-concept application that illustrates how advanced data analytics can be used to help improve the effectiveness of key business decisions.

INTRODUCTION

Busy Beaver is a Pittsburgh-based retailer that operates 25 full-line home improvement centers across Pennsylvania, Ohio, and West Virginia. Despite its recent successes and continued expansion, Busy Beaver believes there is significant potential to increase the average "basket size" of its existing customer base. However, the highly competitive nature of the home improvement marketplace makes this a daunting challenge.

In contrast to online operators, traditional "store" retailers have limited ability to personalize their marketing activities for individual customers. Most of Busy Beavers marketing efforts consists of bi-weekly advertisements. Nearly every week, management determines which products to put on promotion and their corresponding discount levels. Each advertisement promotes approximately 20 items with discounts that range from 10% to 50% off regular retail price. Busy Beaver controls many other aspects of the business that arguably influence customer purchases (e.g. product selection, product placement, retail prices) – however, these are more "long term" decisions that have broader business implications (store layout constraints, inventory levels, competitive positioning, etc.)

The biweekly advertisements play an important role in Busy Beaver's overall business strategy. They are the company's primary "short-term" mechanism for driving business and motivating customer purchases. It is critical for Busy Beaver to ensure that the products and discounts selected in the biweekly advertisement deliver the optimal returns. Data-driven decision support tools could help in the design of the optimal promotions.
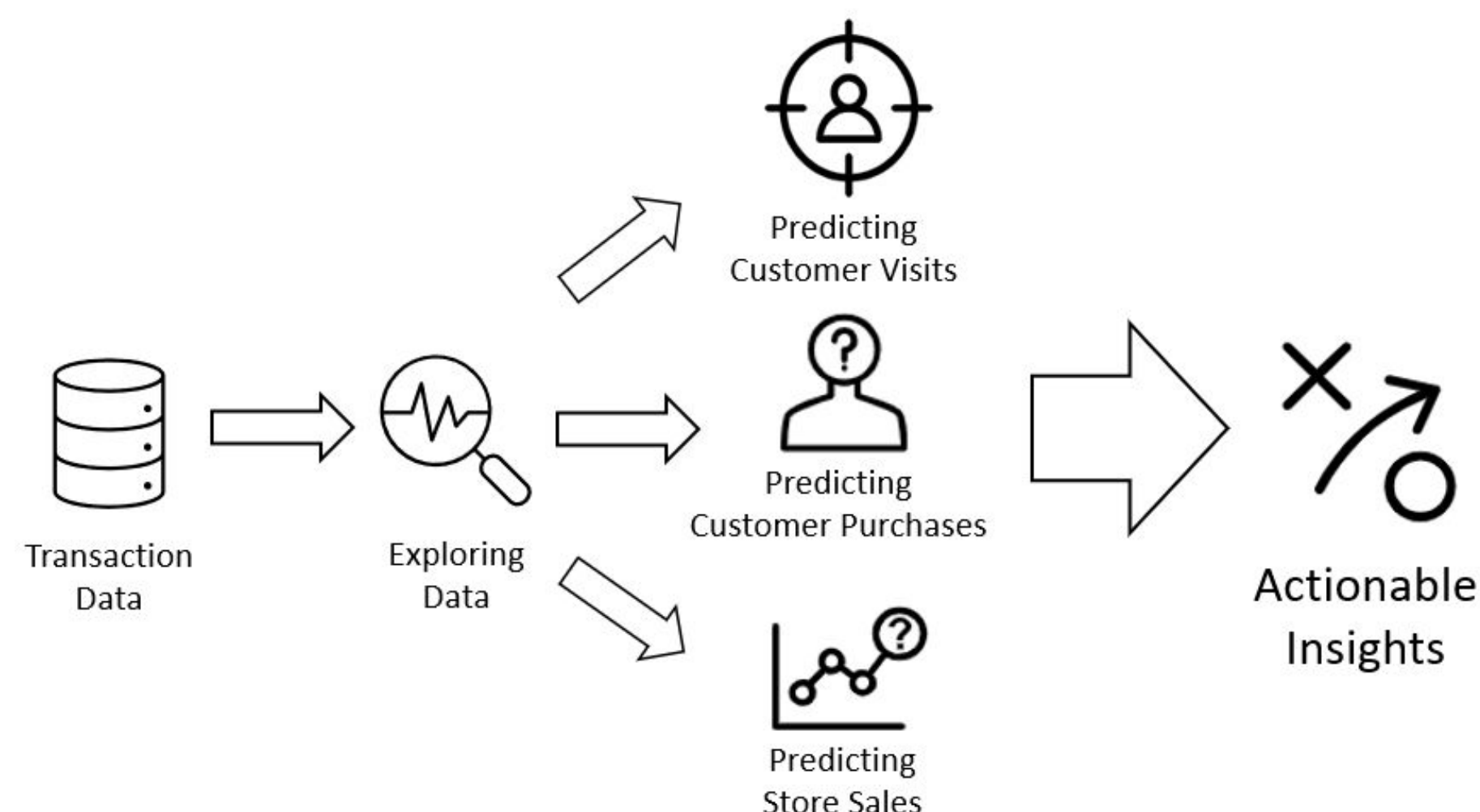
METHODS AND MATERIALS

Busy Beaver's historical transactions were used to develop and train three prediction models that accurately forecast 1) customers' visit decision, 2) customers' purchase decision, and 3) overall sales of any item at a store level.

The first two models make prediction at the individual customer level, which is useful in understanding the factors impacting customers' decision to visit and purchase. The two customer-level models can be combined to forecast sales by aggregating the individual decisions. The last model provides another perspective by approaching the problem from the store level and provides more generalized insights about product sales and promotion performance. Each model considers a variety of factors to inform visit and purchase predictions, including product purchase recency, customer shopping trends (frequency of visits, number of products purchased, basket size), season, weather, store location, product price, and discount level.

Predicting Number of Weekly Visits at All Busy Beaver Stores
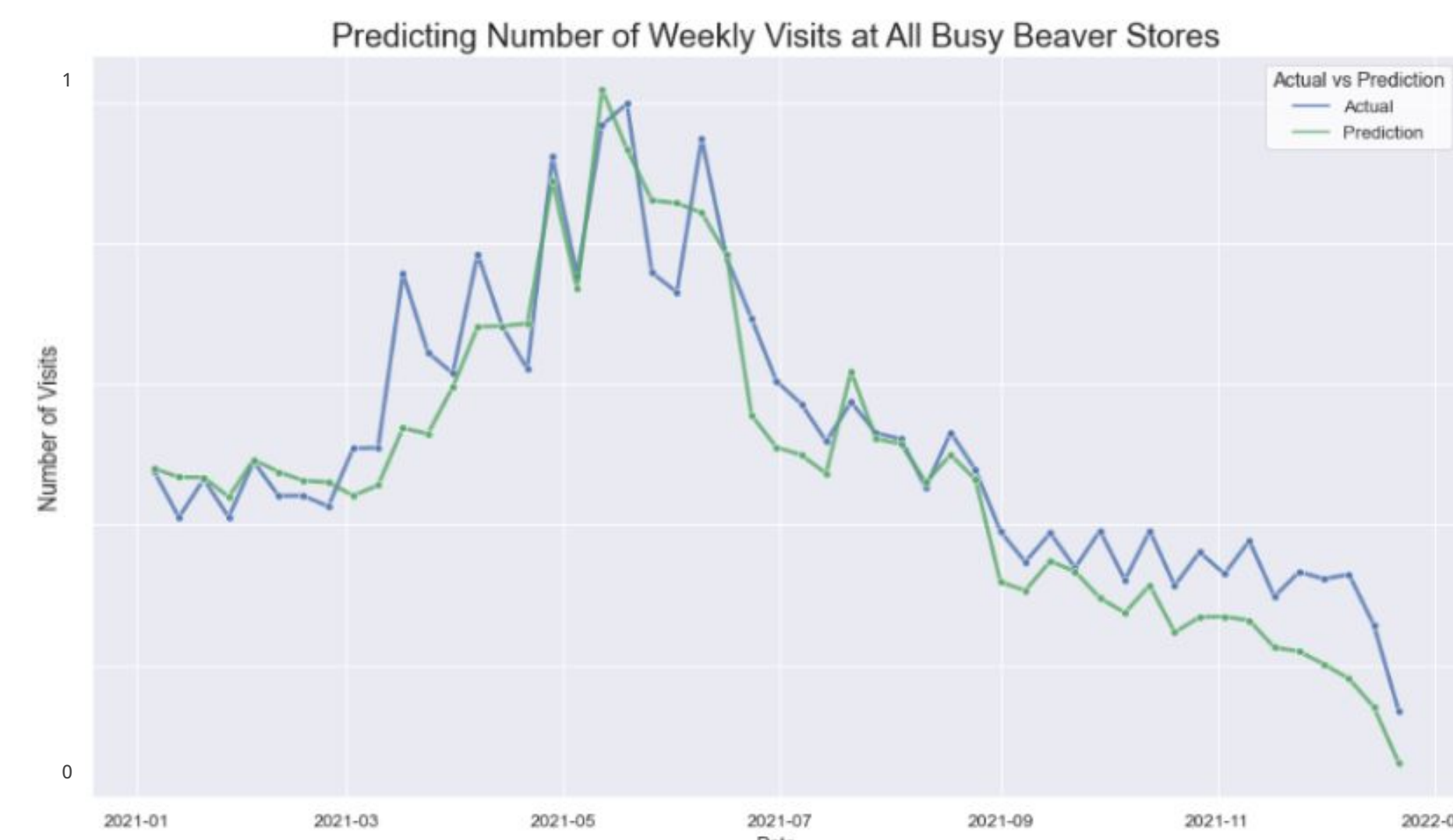
RESULTS

The application of these predictive models can help isolate the impacts of the specific design elements of the bi-weekly advertisements. A key consideration when designing a promotion is setting a discount level such that unit sales are increased beyond the lost revenues associated with the discounted price. The sensitivity of A/C unit sales to discount level was tested in the context of historical promotions. The impact on unit and dollar sales was simulated for several discount scenarios: 1) No discount, 2) 12.5% discount, and 3) 25% discount. The results showed that doubling the discount level had a limited impact on unit sales, with only 5 additional units sold over a two-week period. When factoring in the reduced sales price, total dollar sales was negatively affected by the higher discount. In contrast, the no discount scenario dramatically underperformed the 12.5% discount with significantly less units sold and a sizable decline in net sales.

In fact, the results generally showed that higher discount levels (in various product categories) tended to underperform midrange discounts. The rationale was basically the same as the A/C example in that the incremental units sold generated from the higher discount was not enough to offset the loss in revenue from the higher discounts.

Weekly Sales of A/C Units
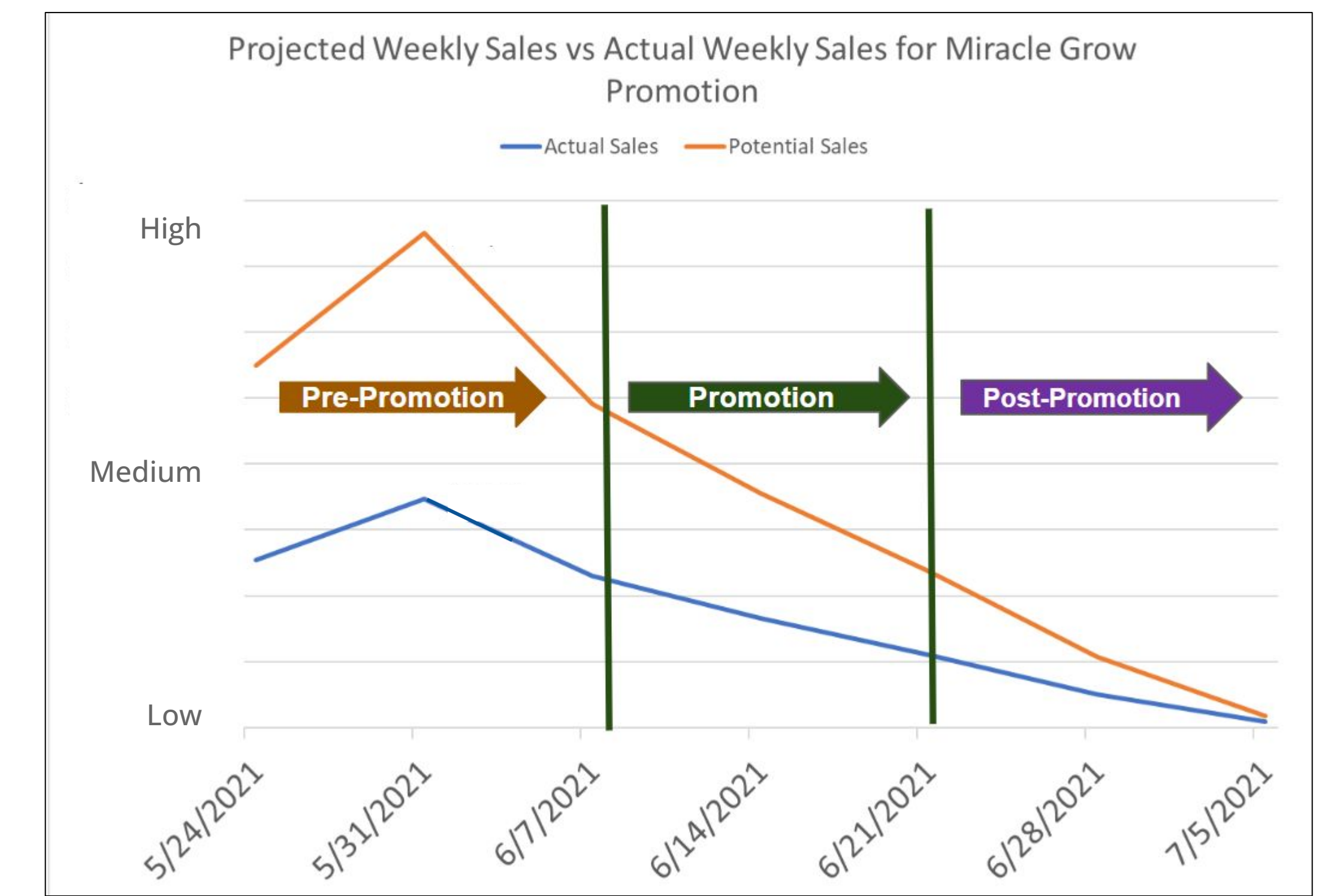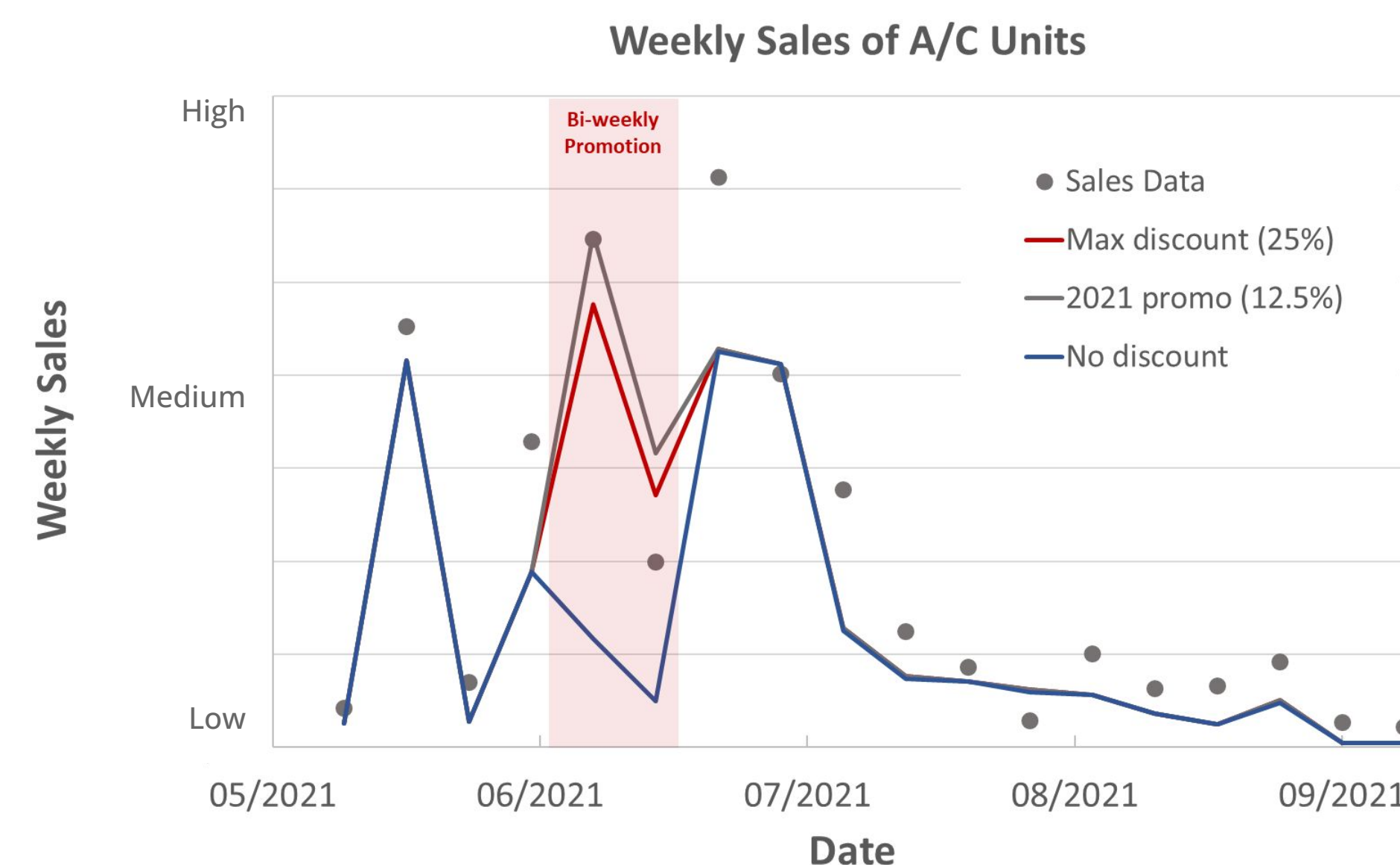
DISCUSSION

When is the best time to promote a product?

The effectiveness of promotions is heavily influenced by the seasonality of specific products. Based on our analysis, sales of seasonal items increase dramatically when promoted with a discount. For home improvement retailers such as Busy Beaver, Spring and early Summer represent peak selling season as homeowners and contractors tend to tackle more projects during warmer weather months.

The quintessential marketing challenge is to determine when in the seasonal cycle an item should be put on discount. There are two primary marketing philosophies: "peak the peek" or "extend the season". The former approach takes advantage of the inherent interest customers have in the early stages of the season whereas the later tries to motivate additional sales as interest wanes. Simulation of promotional timing generally indicated that "peaking the peak" was more effective than "extending the season". For example, Potting mix and AC units generated a higher increase in sales when promoting ahead of a peak and much lower sales from discounts later in the season.

Projected Weekly Sales vs Actual Weekly Sales for Miracle Grow Promotion

CONCLUSIONS

Analysis of a historical promotional campaign using modern machine learning has revealed important insights into customer tendencies and their reaction to promotional campaigns:

• Promotion of big-ticket items (e.g. AC unit, patio furniture) drive incremental trips, but rarely result in the purchase of additional items (e.g. basket is primarily limited to the big ticket items on promotion)

• Biweekly promotions drive store traffic, especially during the first week of a promotion and mainly attract discount-driven customers as opposed to more frequent/loyal customers

• Promotional status, seasonality, and visit frequency are the largest factors in predicting customer visit and purchase decisions

• The effectiveness of promoting top-selling, low-ticket products (e.g. garbage bags) is less driven by the level of discount than the actual inclusion in the advertisement

The "proof of concept" developed for Busy Beaver was effective at identifying insights that could prove helpful in improving the effectiveness of the advertising strategy. That said, the insights and methods need to be generalized to inform key business actions or decisions at an enterprise level.

CONTACT

[Alan Montgomery]
[Tepper School of Business]
[5000 Forbes Avenue, Pittsburgh, PA 15213]
[alm3@andrew.cmu.edu]
[412-268-4562]