Impact of new candy in grocery stores

Soyoung An

December 23, 2021

Outline of Experiment

- **Situation**: Client X owns stores that sells a variety of snacks and beverages. He has a number of stores all around the US and believes introducing new Brand Z candies will increase overall sales across all the stores.
- **Complication**: This is a high risk move (time and money) in case if Client X's intuition is incorrect.
- **Resolution**: Need to test with a small subset of stores to validate what the incremental revenue lift is going to be.
- **Approach**: We need to select a set of "treatment" stores where the brand Z will be introduced. We will then do a one tail t-test if the incremental revenue is statistically significant or not. We will base line it with the same store sales before the brand Z was introduced.
- If the test stores show statistically significant higher incremental revenue (lift) compared to the control stores, we will recommend introducing brand Z across all of the stores.

Objectives

Understand whether the Brand Z's candies will create an increase in total revenue before fully implementing the strategy across all stores

Datasets

Transactions.csv

- 1,048,575 rows, 5 features
- Features: date, store_id, product_id, currency, revenue

Store_attributes.csv

- 16,538 rows, 6 features
- Features: store_id, attribute_id, attribute_type, attribute values (integer, float, string)

Products_of_interest.csv

- 502 rows, 1 feature

control_stores.csv

- 25 rows, 1 feature

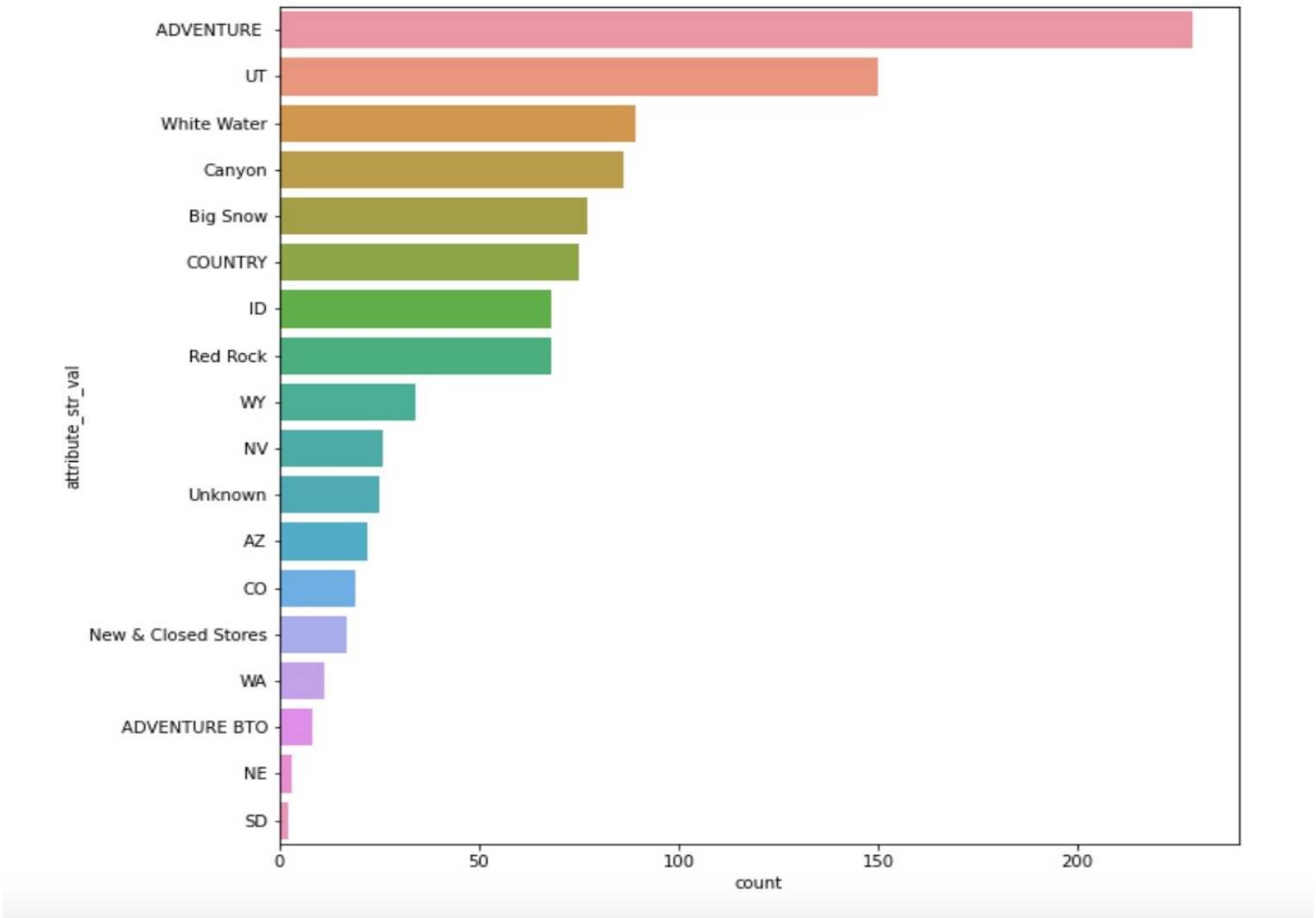
treatment_stores.csv

- 25 rows, 1 feature

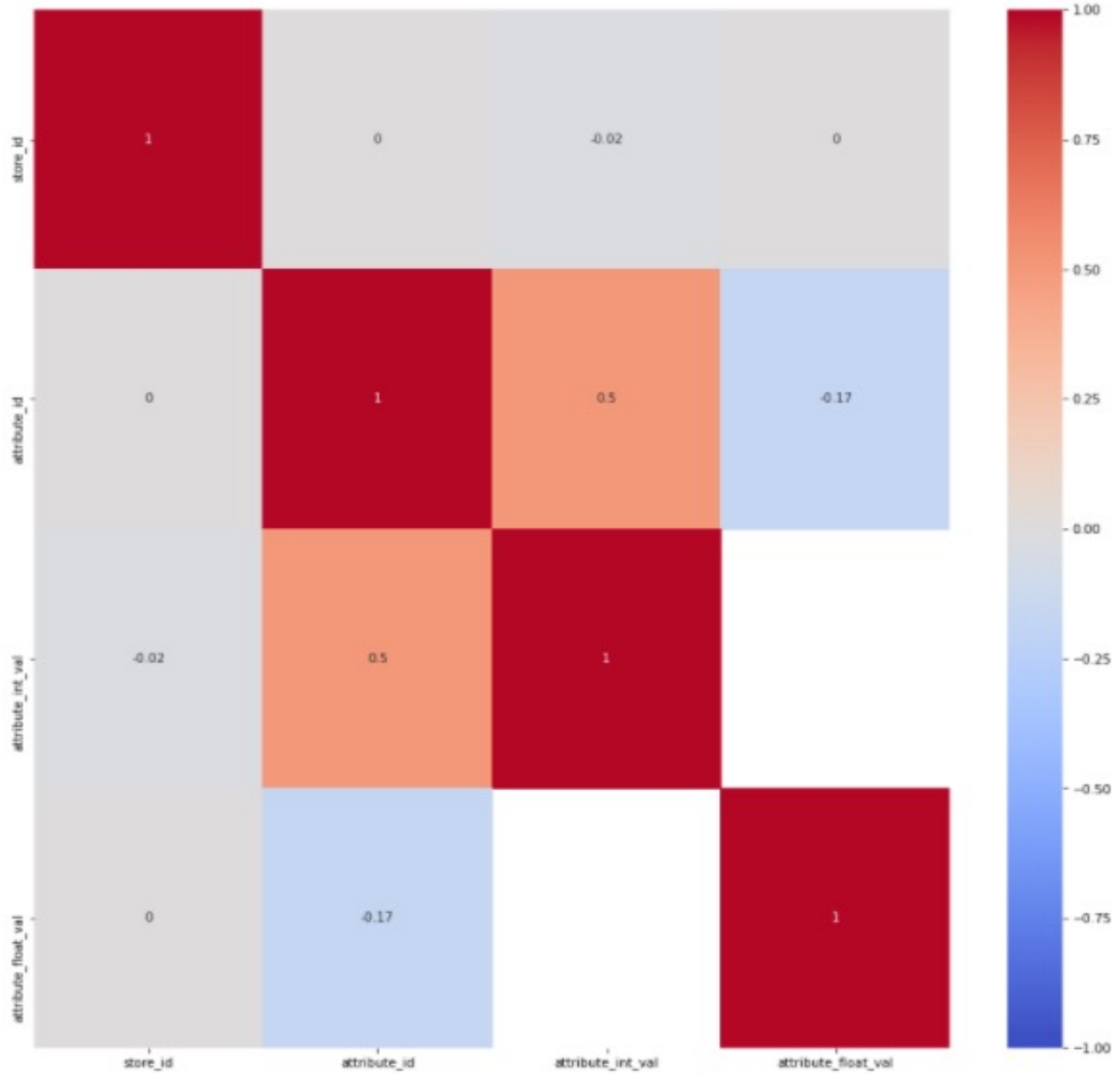
Explore the data

- Total of 337 unique stores in the dataset
- Each store consists of many different attributes (combination of float, integer, string attribute types)
- Total of 53 different attribute_id
  - 4 string type attributes (9, 15, 51, 36)
  - 2 integer type attributes (28, 35)
  - 47 float type attributes
- Stores are located in 11 different states
  - Utah, Idaho, Wyoming, Nevada, Arizona, Colorado, Washington, Nebraska, South Dakota, Montana, Oregon

Count plot of string type attribute



Heatmap


store_id
attribute_id
attribute_int_val
attribute_float_val
1
0
-0.02
0
0
1
0.5
-0.17
-0.02
0.5
1
0
-0.17
1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

Sampling Introduction

Overview

Experiment design is critical for the success of AB Testing. The distribution of the treatment and control has to be similar. The optimal approach is to use stratified sampling. This ensures the two segments are evenly distributed along the critical feature.

Challenges in our data

In the data, we had two challenges to do stratified sampling in the straightforward way:

- a) Features are unknown
- b) Features are mixed

How to draw samples?

Stratified Sampling Approaches:

1. Store store_id, attribute_id, and attribute values by different attribute types (string, integer, float)
2. Using pivot table, convert row and columns
3. Impute any null values with mean of each features
4. Merge 3 pivot tables
5. Convert string values using one hot encoding method
6. Cluster to different group using K-Mean Clustering algorithm
    1. Use Elbow Sum of Square Method to find optimal K value
7. Assign the treatment group from each cluster with a same distribution and rest on the control group

K mean Clustering

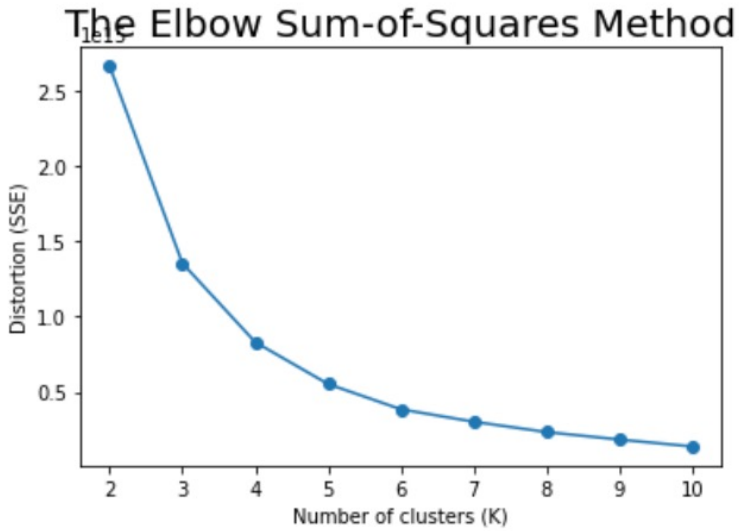


Figure 1

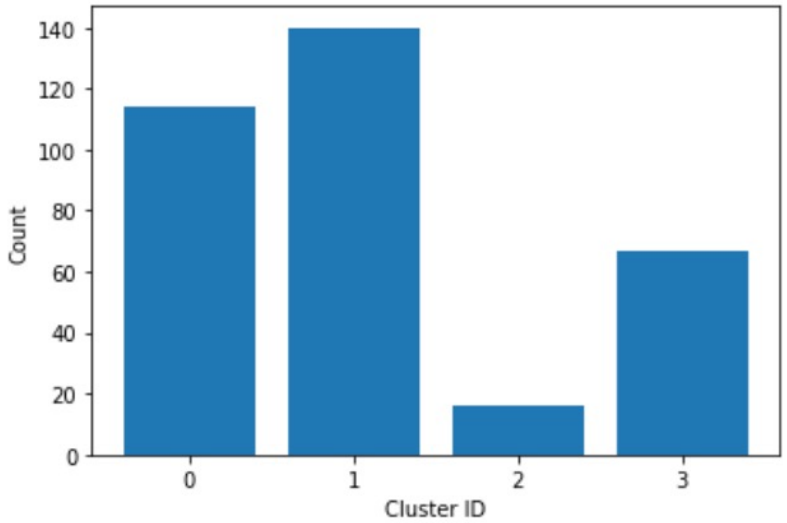


Figure 2

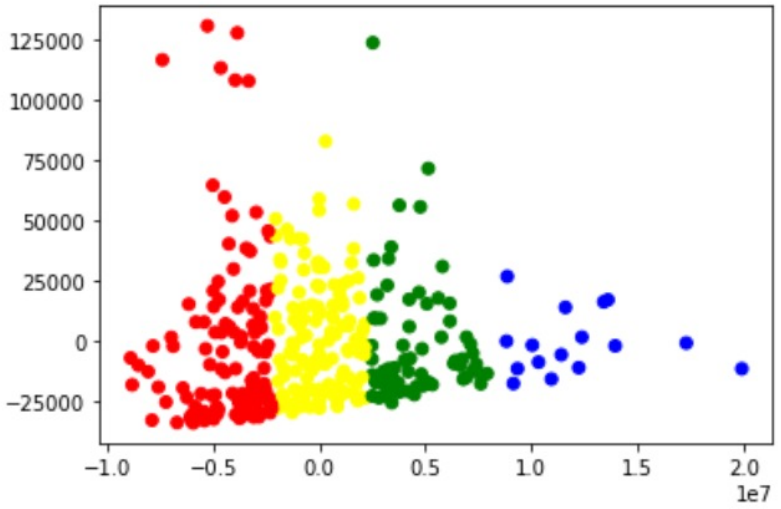


Figure 3

Sample Results

- Treatment group: 34 stores
- Control group: 303 stores

- Total number of stores: 337 stores

Treatment Group Store ID
386
397
401
441
445
510
630
369
380
422
477
505
543
561
565
607
613
654
592
595
345
365
385
390
406
443
472
509
526
532
559
594
640
655

A/B Testing Methods

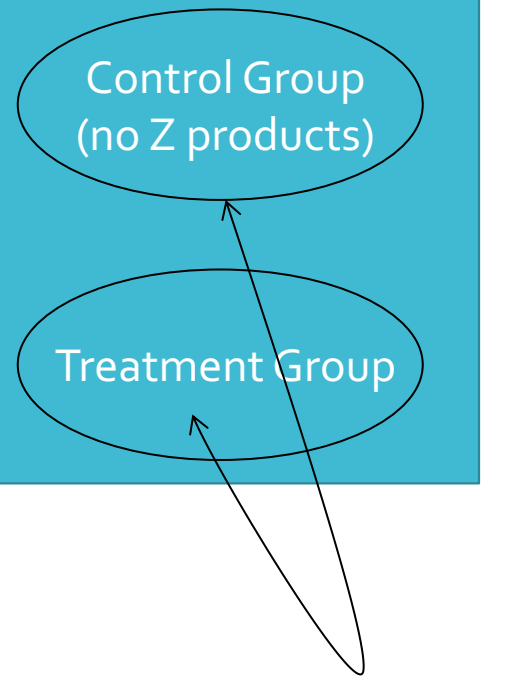
1. Split transactions for control groups and treatment groups
2. Split each control and treatment groups as different period (pre-test, implementation, test)
3. Removed Product Z for all control groups and treatment (pre-test, implementation) groups
4. Conduct one-sided t-test to see significance for each test cases

Experimental Test Design

Pre-Test Period 07/17/2016 - 10/15/2016	Implementation Period 10/16/2016 – 11/12/2016	Test Period 11/13/2016 – 02/12/2017
Control Group (no Z products)	Control Group (no Z products)	Control Group (no Z products)
Treatment Group (no Z products)	Treatment Group (no Z products)	Treatment Group

Test Design 1

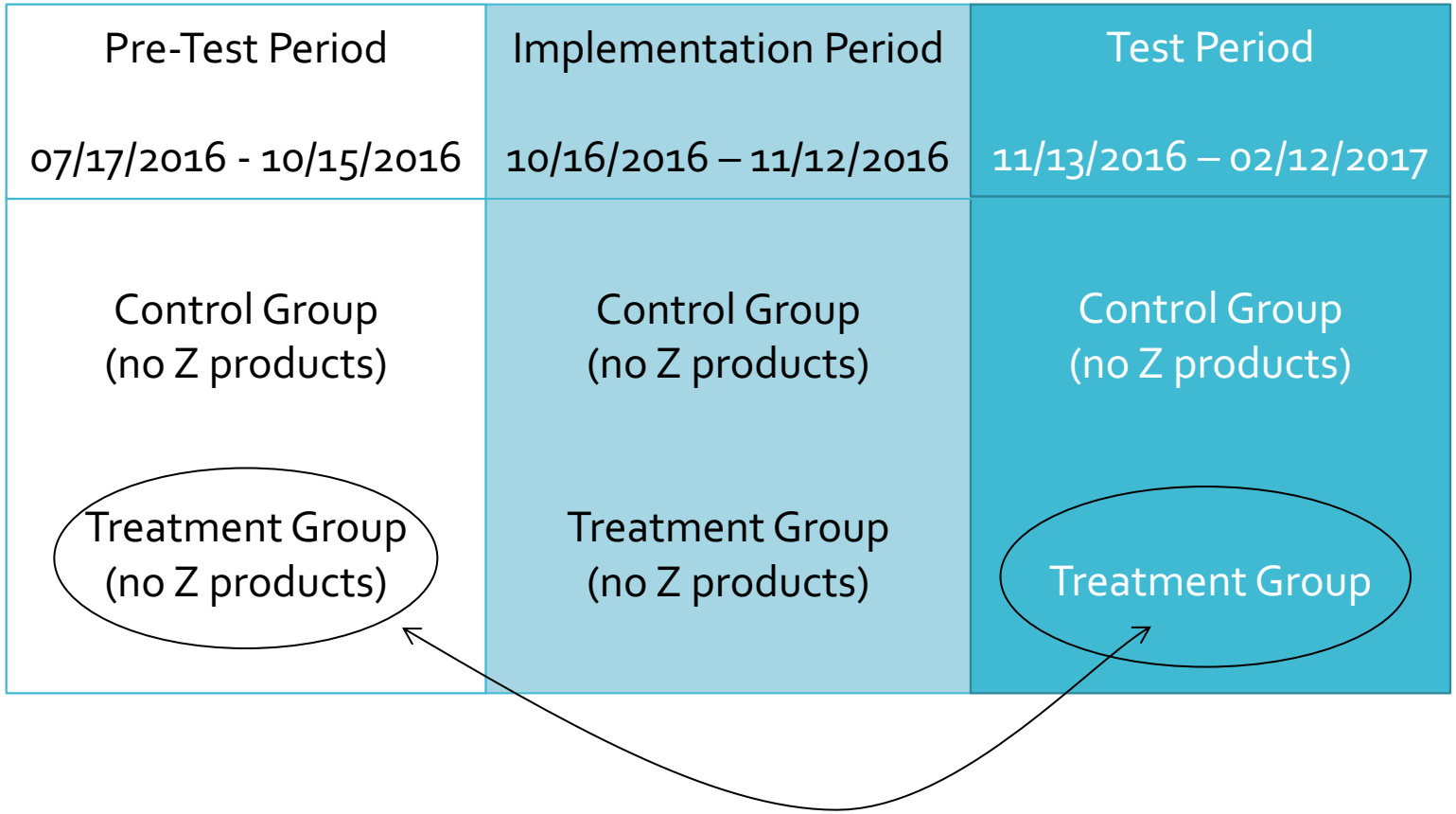
Pre-Test Period 07/17/2016 - 10/15/2016	Implementation Period 10/16/2016 – 11/12/2016	Test Period 11/13/2016 – 02/12/2017
Control Group (no Z products)	Control Group (no Z products)	Control Group (no Z products)
Treatment Group (no Z products)	Treatment Group (no Z products)	Treatment Group



Did Treatment stores have a higher incremental revenue during Test period compared to Control stores?

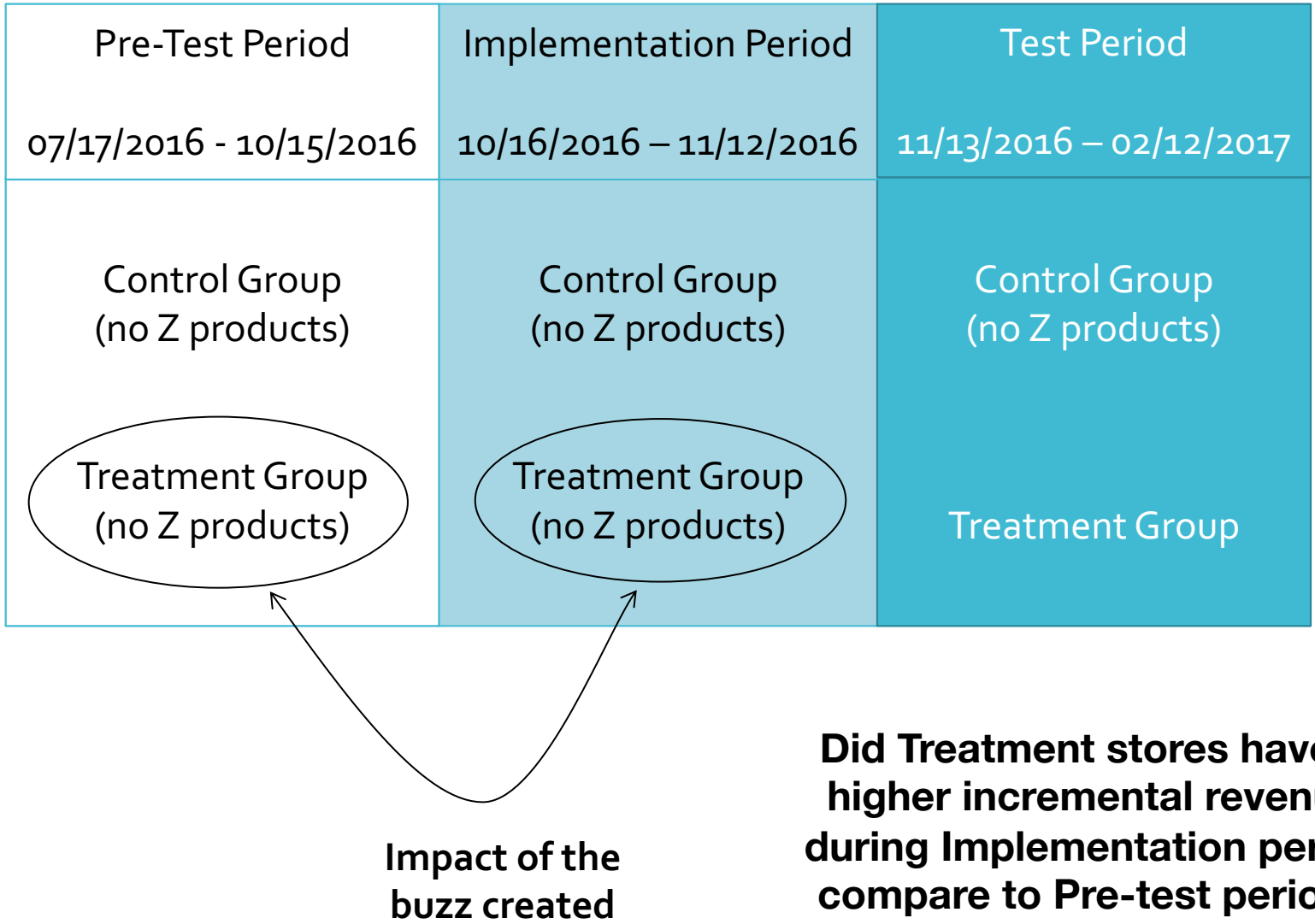
Impact of Brand Z

Test Design 2



Did Treatment stores have a higher incremental revenue during Test period compared to Pre-Test period?

Test Design 3



Did Treatment stores have a higher incremental revenue during Implementation period compare to Pre-test period?

Results

- Observed a strong statistical significance for both Test Design 1 (Brand Z impact) and 2 (Same Store Impact of Brand Z)
- No significant difference was observed in the impact of the buzz created during the implementation period

- Effect Size Analysis
  - $$\frac{\text{Increase with treatment} - \text{Increase without treatment}}{\text{Increase without treatment}} * 100\%$$
  - $\text{Increase without treatment} = \text{Avg}(\text{Control Test Revenue}) - \text{Avg}(\text{Control Pretest Revenue})$
  - $\text{Increase with treatment} = \text{Avg}(\text{Treatment Test Revenue}) - \text{Avg}(\text{Treatment Pretest Revenue})$
  - **503.34% increase** in revenue due to Brand Z candies

Conclusions

Implement new Brand Z's candies will increase revenue for the business

Thank you!