

# Fast Food Marketing Campaign AB Test

IBM Watson Analytics Marketing Campaign

Urh Peček

September 2022

## Contents

<b>1. Task</b>	<b>2</b>
<b>2. Data</b>	<b>2</b>
<b>3. Exploratory analysis</b>	<b>2</b>
3.1 Weeks . . . . .	2
3.2 Location ID . . . . .	3

# 1. Task

Data and description of the task is part of one of Kaggle's data sets. The data is obtained for (fictional) research in Fast Food Marketing Campaign AB Test.

The task is about a fast-food chain which plans to add a new item to its menu. They are still undecided between three possible marketing campaigns for promoting the new product. So in order to determine which promotion has the greatest effect on sales, the new item has been introduced at locations in several randomly selected markets. A different promotion was used at each location, and the weekly sales of the new item are recorded for the first four weeks.

Our goal was to evaluate the results of AB test and decide which marketing strategy works the best.

# 2. Data

The data available to us for research is 7 variables with 548 recorded observations where there is no missing values. Variables and their short description is as follows:

- MarketID: unique identifier for market
- MarketSize: size of market area by sales
- LocationID: unique identifier for store location
- AgeOfStore: age of store in years
- Promotion: one of three promotions that were tested
- week: one of four weeks when the promotions were run
- SalesInThousands: sales amount for a specific LocationID, Promotion, and week

We want to decide on one (or more) out of three campaigns (1, 2, 3). A variable that determines the performance of a campaign or which we focus on is Salesinthousands. Other variables are a kind of noise in the data that we need to get rid of or explore their influence and how they could affect our final decision and to adjust it accordingly.

# 3. Exploratory analysis

Let's first look at all the variables that could indirectly affect the campaign performance and determine how we will treat them accordingly.

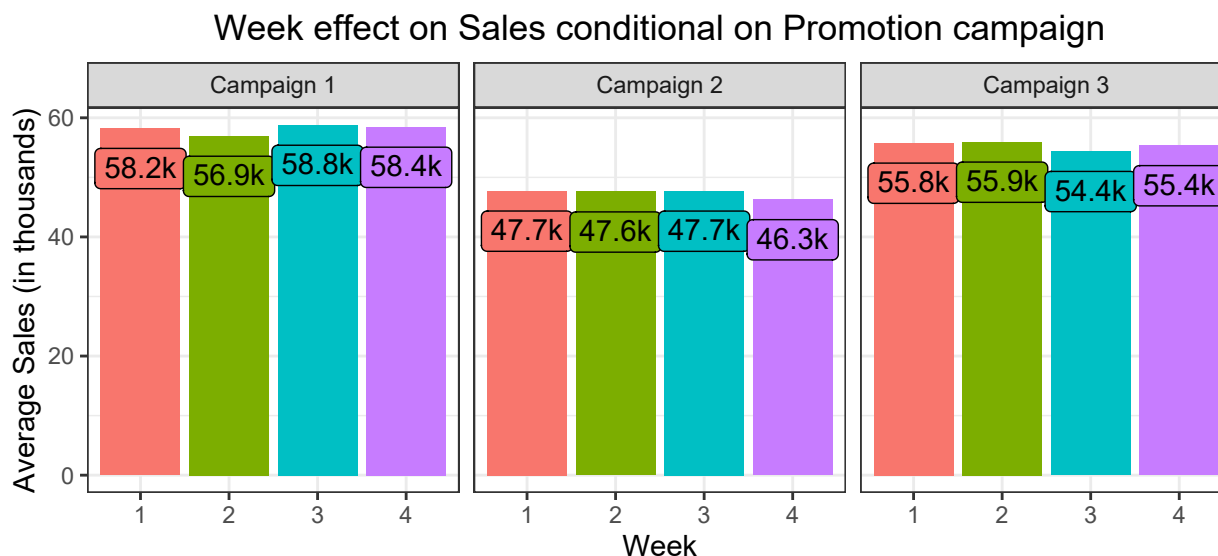
## 3.1 Weeks

For each location, we have recorded sales for each of the first 4 weeks separately. Before anything else, let's look at weeks to see if it needs to be considered within other variables.

First, let's look at what influence the weeks of campaigns have on sale, where we do not consider any other variable.

Week	Average Sales (in thousands)
1	53.79
2	53.39
3	53.47
4	53.21

We see that there are no differences between weeks in terms of Sales. This was also confirmed by the Kruskal-Wallis test. Comparing sales according to weeks for each promotion campaign separately, we see that there are no significant differences between weeks. That is again confirmed by Kruskal-Wallis tests.



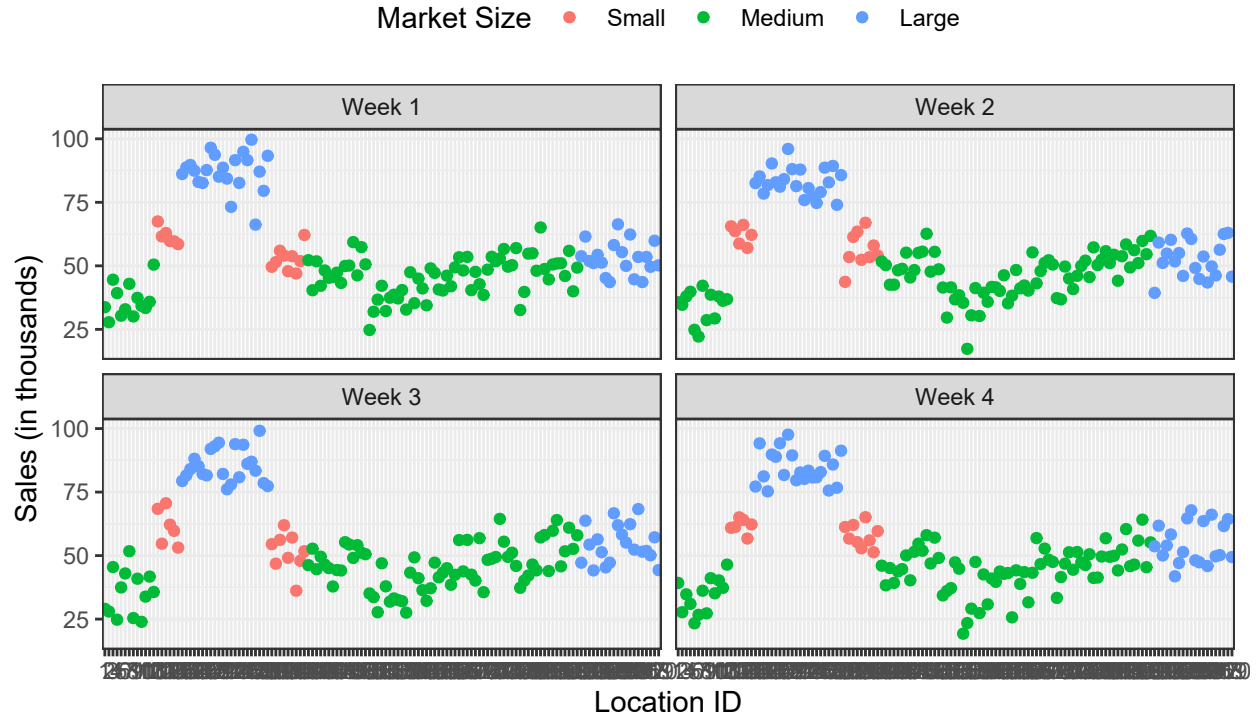
We can conclude that promotional campaigns do not have to be taken into account conditionally for weeks, but we can add up all 4 weeks together and treat it as a joint sale. This also means that one promotional campaign can be done in all four weeks, as there are no differences in their efficiency. If there were differences between weeks and two campaigns were similar efficiency, we could mix campaigns with each other.

We also see that taking into account the cumulative values of sales according to the promotion campaign, campaigns 1 and 3 deviate significantly from the campaign 2. We will investigate on whether there is an impact of any other variable here.

### 3.2 Location ID

There are 137 store locations where each location is represented exactly once and at each location, exactly one campaign was carried out.

Let's look if we have any outliers, i.e. locations where there are visibly higher or lower sales than we would expect from similar locations and could thus affect the final choice of campaign.



We see that there are no visible outliers considering both the week and the market size. So we don't have to pay attention to any location that would be significantly better than the rest and thus distorting the final conclusions.

The location is also associated with its number within each campaign. We see that the number of locations and thus campaigns are different in terms of promotional group. This should be taken into account in sales where it is not appropriate to consider the cumulative value of sales within the campaign.

Campaign	Number of locations
1	43
2	47
3	47

The joint sale within the campaign must thus be taken into account according to the number of locations and thus take the sale of the average location within the campaign (normalized sales). This also allows us the above fact that there are no outliers.

Campaign	Number of locations	Sales (k)	Normalized sales (k)	Sales [%]	Normalized sales [%]
1	43	9993	232	34	36
2	47	8898	189	30	29
3	47	10409	221	36	34