Fast Food Marketing Campaign AB Test

IBM Watson Analytics Marketing Campaign

Urh Peček

September 2022

Contents

1. About

1.1 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.

1.2 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.

2. 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.

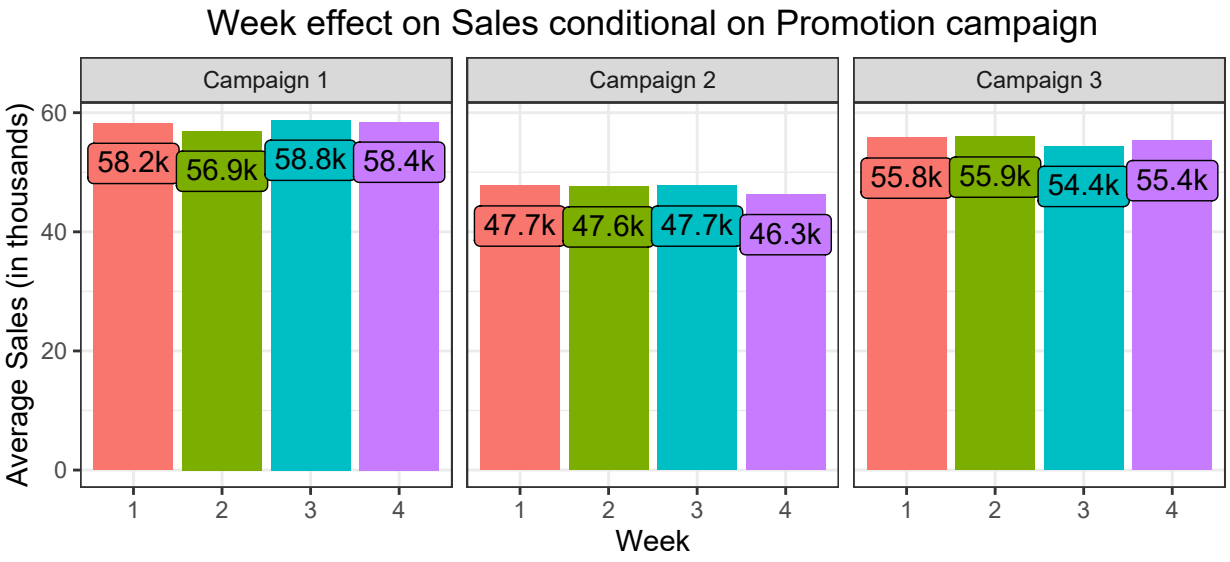
2.1 Weeks

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.



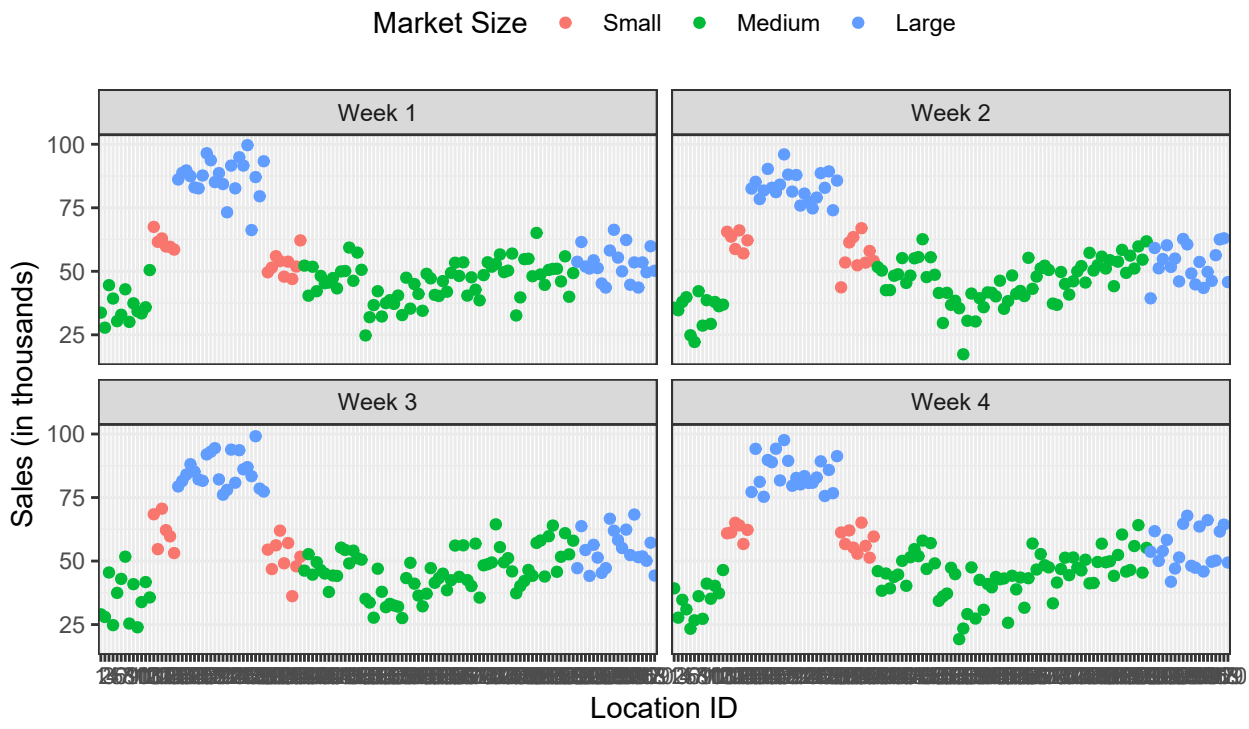
We can conclude that promotional campaigns do not have to be taken into account conditionally for weeks, but we can add up all 4 weeks together and treat it as a joint sale. This also means that one promotional campaign can be done in all four weeks, as there are no differences in their efficiency. If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

2.2 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 campaign. 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.

At the table below, where we compare the cumulative values of sales and the number of locations of normalized sales values and their proportions within total sales, conditionally on the promotion campaign, we see the impact of the number of locations. If we only look at total sales, without taking into account the

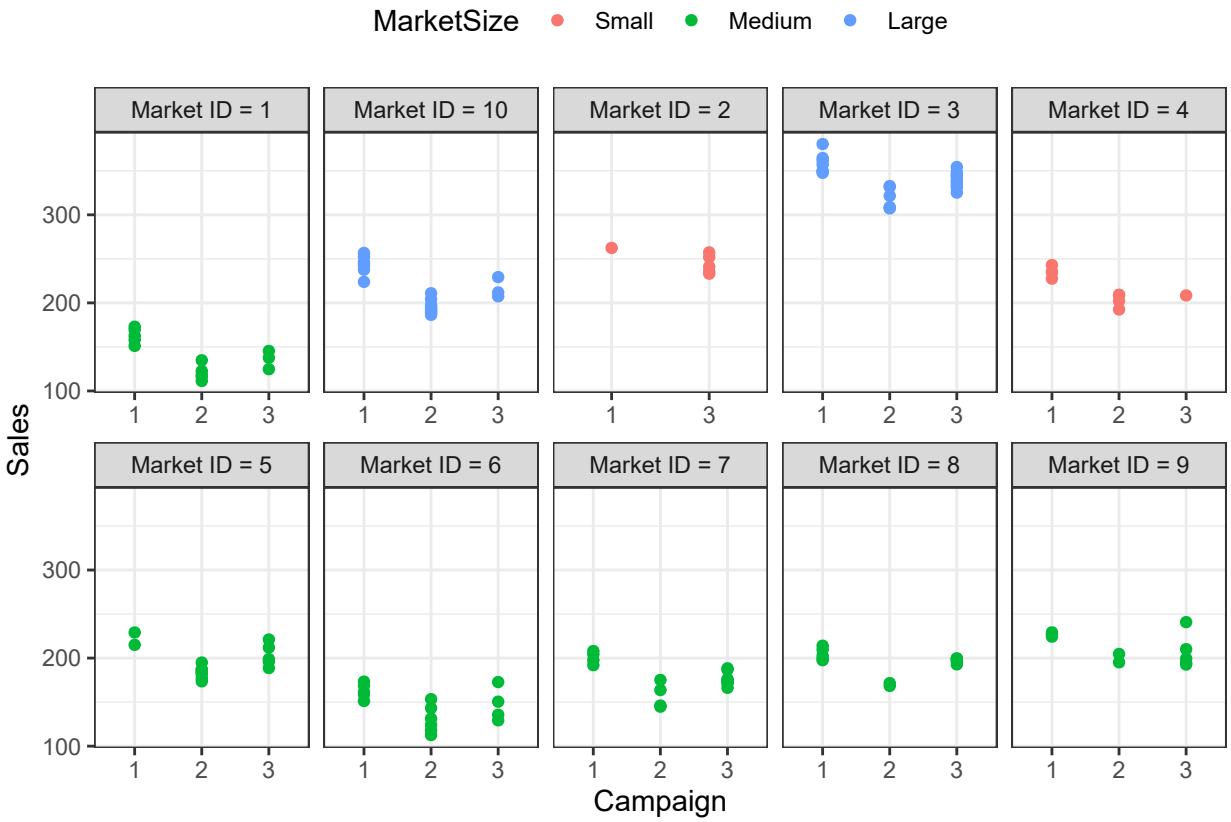
number of locations, the best promotion campaign is 3, but if we normalize for the number of locations, the best promotion campaign is 1.

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

Now the normalized sales values for the number of locations can be compared with the test to see if there are significant differences in sales values between campaigns. The Kruskal-Wallis test tells us that there are significant differences between the campaigns in terms of sales, and the paired Wilcox test with correction for multiple testing tells us that there is no statistically significant difference between campaigns 1 and 3, while campaign 2 is statistically significantly worse.

2.3 Market ID

Similar to Location ID, we can look for outliers in Market ID. 10 different markets were included, which are not equally represented between campaigns. Are customers in a certain market more inclined to one of the promotional campaigns than the other two? If this is true, MarketID may be a variable that affects Sales and is not included in the promotion campaign. Thus, campaigns can also be adjusted to Market ID.

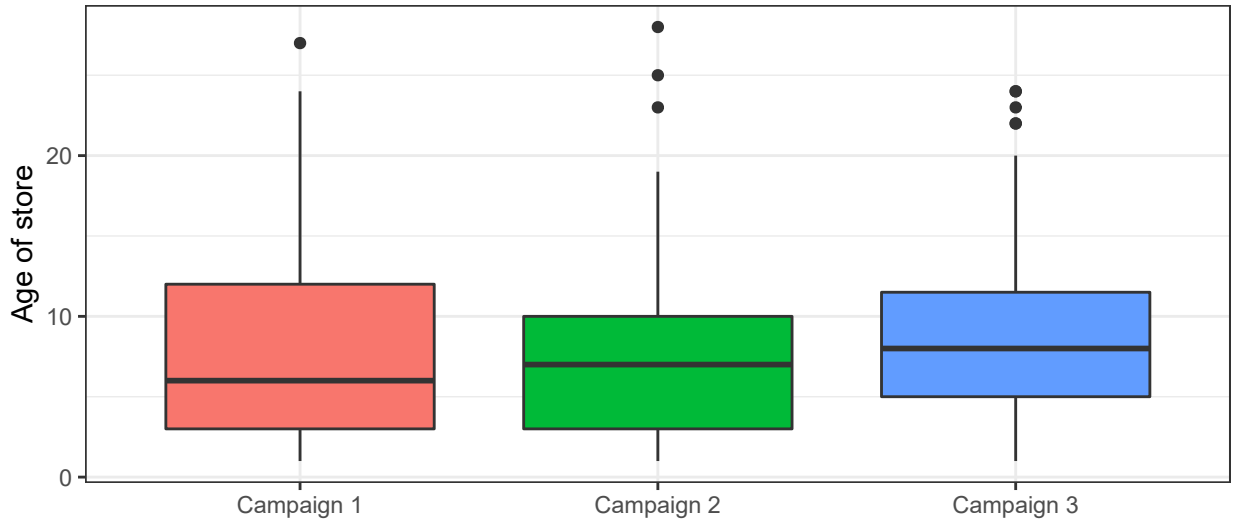


We can see that the sales value conditional on the marketing campaign is similarly distributed among all Markets IDs. They sales differ mostly in terms of market size. Therefore, there are no outliers in the Market ID that could influence the campaign selection or adjust this selection. It is only true that some Market

IDs give on average higher Sales than the others, but there are no differences when comparing promotions. MarketID is thus not a variable that would influence the selection of a promotional campaign.

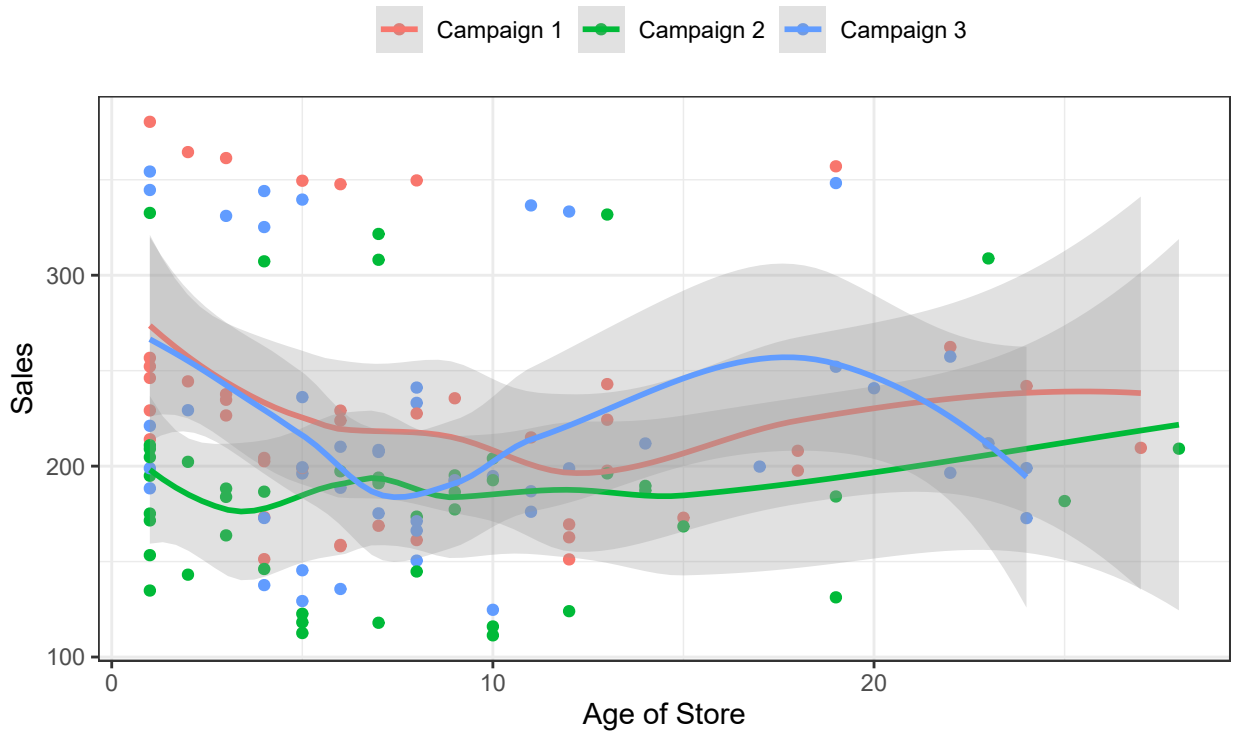
2.4 Age of store

Let's see how the age of stores (Location ID) is distributed between promotional campaigns. Stores age varies from 1 to 28 years. If the distribution was not "equal" between promotional campaigns and there was a relationship between age of stores and sales, this could have an undesirable effect on the choice of campaign.



There are no noticeable differences in the age distribution of stores between the campaigns. This is also "confirmed" or not rejected by the Kruskal-Wallis test. The age of the stores thus does not indirectly affect the choice of promotional campaign.

Although we are not interested in that here, we can also check if the age of the stores is related to sales.



In the figure above, the relationship is also shown conditional on the promotional campaign. We can see that there is a strong overfitting in the data and that there is no visible relationship betweenage of store and sales. This can also be verified with a test, where the non-parametric Spearman correlation test does not reject the hypothesis of zero correlation ($p = 0.43$). Otherwise the value of the correlation coefficient equals $\rho = -0.068$.

2.5 MarketSize

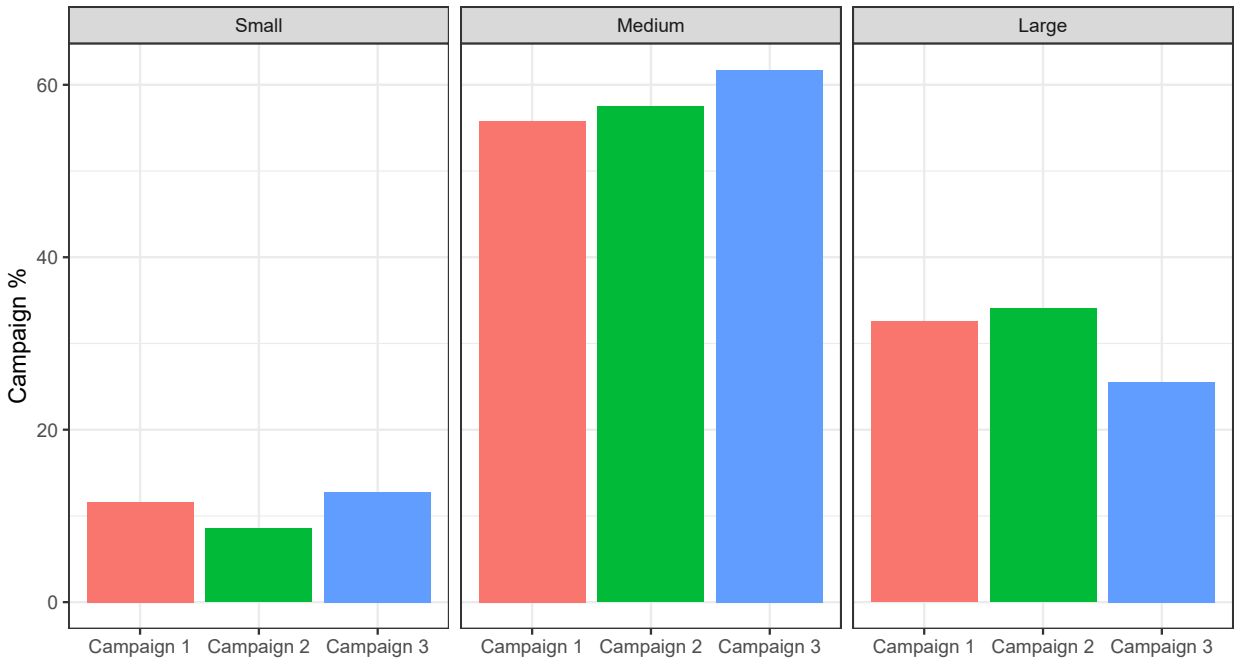
Let's take a look at the market size as the last, perhaps unintentionally influencing variable. According to the table below, we can quickly see that the value of sales varies greatly depending on the market sizes. The difference between average sales conditional to market size can also be confirmed with the Kruskal-Wallis test, where we reject the hypothesis of equality of average sales between market sizes. Paired Wilcox test with corrected p-values for multiple testing confirms that all three market sizes differ significantly in terms of sales.

Market size	Number of stores	Average sales (in thousands)
Small	15	230
Medium	80	176
Large	42	280

As long as the distribution of market sizes between promotional campaigns is the same, there is no problem with this, otherwise market size must be taken into account when choosing the optimal campaign.

So let's see how market size is distributed within promotional campaigns. Perhaps better than absolute values, due to the different number of stores (which we take into account with average sales), we can look at the percentage values of the size of the stores within the promotional campaign.

Market size	Campaign 1 [%]	Campaign 2 [%]	Campaign 3 [%]
Small	12	9	13
Medium	56	57	62
Large	33	34	26



We can see that the distribution of market sizes within promotional campaigns is not exactly the same. This, in combination with a strong statistically significant difference in the value of sales within the size of the markets, could have a significant impact on the selection of the optimal promotional campaign.

Now let's test whether there is a correlation between the the market size and the promotional campaign. We can check this with the chi^2 test based on the table below. It should be noted that one of the assumptions of the χ^2 test is that all frequencies are > 5, which is not completely fulfilled in our case. However, based on the large p-value ($p = 0.88$), we cannot reject the assumption that market size and promotional campaign are not correlated variables.

Market size	Campaign 1	Campaign 2	Campaign 3
Small	5	4	6
Medium	24	27	29
Large	14	16	12

Thus, despite the relationship between market size and sales, market size does not significantly affect the choice of a promotional campaign.

However, what if we don't trust the test completely and we are worried that despite the statistically insignificant correlation between market size and campaign, the correlation is big enough to distort our choice of promotional campaign. For this purpose, we can try to get rid of the influence of market size on sales and leave a different distribution of market size within campaigns. Similarly, we could proceed in the other direction as well.

We will normalize sales so that sales within small market size will be comparable values as within large and medium market size. Thus, we can simply average the sales within each promotional campaign and no longer

need to consider the market size as an influencing variable on sales. We will normalize sales on the basis of market size, so that the average value of sales within each market size will 1 and thus comparable between sizes. So, conditional on market size, we divide the sale value by the average sale within this market size. In order to obtain more easily interpretable values (not around 1), we multiply each obtained sales value by the average value of all sales, disregarding any other variable (market size).

Market size could also be taken into account by comparing promotional campaigns separately, according to market size, but again it would be more confusing, since we would have to somehow weight the 3 values to choose the optimal promotional campaign.

3. Final decisions

3.1 Summary of findings

Our goal is to determine the optimal selection of promotional campaigns, or to choose one (or more) of the three promotional campaigns that brings the highest sales value. For this purpose, before the actual analysis and subsequent selection of the optimal campaign, we had to study additional variables that were included in the campaigns, could affect their effectiveness and could influence their selection. The findings are as follows:

- For each location, we have recorded sales for each of the first 4 weeks separately. It turned out that there are no significant differences in sales between weeks, taking into account all campaigns together as well as individual campaigns. The weeks variable thus does not indirectly affect the selection of the campaign and thus does not need to be taken into account. For each campaign and within that location, all four weeks can be combined and treated as a total value.
- There are 137 store locations where each location is represented exactly once and at each location, exactly one campaign was carried out. There are no outliers regarding the location, 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. We saw that the number of locations and thus campaigns are different in terms of promotional campaign. This means that we have to compare the sales values by campaign according to the average sales value by location and not according to the total value within the promotional campaign. We found that if we compare to the number of locations normalized sales value by campaign, the promotion campaigns 1 and 3 are significantly better than campaign 2 and there is no difference between them. Of course, this is not the final conclusion, as it is necessary to check the possible indirect influence of other variables.
- 10 different markets were included, which were not equally represented between campaigns. We saw that the sales value conditional on the marketing campaign is similarly distributed among all Markets and there were no outliers which could influence the campaign selection. Thus, market does not influence the selection of a promotional campaign.
- Stores (locations) age varies from 1 to 28 years. There are no significant nor noticeable differences in the age distribution of stores between the campaigns. The age of the stores thus does not indirectly affect the choice of promotional campaign. Additionally, there is no relationship between age of store and value of sales.
- Campaigns or locations were implemented in different markets, which were of three possible sizes (small, medium, large). We saw that the value of sales varies greatly depending on the market sizes. Disregarding the campaign, large markets provide statistically significant biggest sales, followed by small markets and the last are medium-sized markets. Considering the statistical test, this does not present a problem, since we do not reject the hypothesis that the size of the markets is equally distributed between the campaigns. However, if we look at the number or shares of market sizes within each campaign, we can see differences. Although these differences are not statistically significant, they could represent a "problem" when choosing a campaign. If we want to get rid of the possible influence of the market size, we can normalize the sales values according to the size of the market and thus put the sales on the same scale and enable an unbiased comparison of promotional campaigns.

3.2 Choosing the optimal promotional campaign

So we have seen that considering only the bare results of the tests and their statistical significance, the only thing that needs to be taken into account when choosing the optimal promotional campaign is that we do not compare the cumulative values of sales but the average values by location.

The average values of sales by campaigns are shown in the table below and the Kruskal-Wallis test and the paired Wilcox test with correction for multiple testing tell us that there is no statistically significant difference between campaigns 1 and 3, while campaign 2 is statistically significantly worse. To choose the optimal campaign, we would thus decide between campaign 1 or 3, favoring campaign 1.

Campaign	Average sales (in thousands)	Average sales [%]
1	232.4	36.1
2	189.3	29.4
3	221.5	34.4

If we do not fully trust the statistical test regarding the correlation between market size and campaign, or if we think that the differences, due to typical differences in sales between market sizes, could be important, we can consider the normalized sales values. We obtain the following normalized average sales values by location according to the campaign.

Results are similar to above (the statistical test was right). Kruskal-Wallis test and the paired Wilcox test with correction for multiple testing tell us that there is no statistically significant difference between campaigns 1 and 3, while campaign 2 is statistically significantly worse. To choose the optimal campaign, we would thus decide between campaign 1 or 3, favoring campaign 1.

Campaign	Normalized average sales (in thousands)	Normalized average sales [%]
1	230.1	35.8
2	188.0	29.2
3	224.8	35.0