# IMPACT OF REAL TIME MOBILE MESSAGES

**Shopco Project Report**

**Wilian Delgado**

Chicago, June 7th, 2022

*(table of contents added and updated automatically with page numbers)*

Table of Contents

## 1. EXECUTIVE SUMMARY

In a recent experiment, a mall operator wants to determine the effectiveness of using real-time mobile messaging. The goal of this is to determine what the consumer's behavior is like when being lured to spend more after their initial purchase.

This experiment surveyed a group of 9032 people who visited the mall in the week of October 1, 2019. Within the data, it was found that the majority of people in this pool were women who were, on average, 32 years old. When judging whether the real-time messaging method worked, it was measured that a little over 50% of people receiving a loyalty message went on to shop further. The average amount spent when continuing to shop was $113.

When promising loyalty points through real-time messaging, it's proven that shopping does continue. Using this type of technique when stimulating sales does help in creating a short-term spike in people shopping.

## 2. OBJECTIVES AND QUESTIONS

### 2.1 Project purpose

**Objective of the project**

The objective of this research is to assess the impact of getting a real-time mobile message at a shopping center on a customer's willingness to continue shopping there.

**Questions**

Question 1) How many customers received the real time messages from the experiment data?

Question 2) Did the customer purchase products from the second store after the first purchase?

Question 3) What was the demographics of the customer like Age, Gender ?

Question 4) What percentage of the customers were loyal in their shopping patterns?

**Experiment**

The mall operator ran a field experiment on 9032 visitors who visited the mall during the week of October 1, 2019, in study the influence of real-time mobile messages on customers' shopping decisions. In the trial, 9032 individuals were split evenly between two parts. The first half of customers (4516) received a real-time mobile message guaranteeing 1,000 loyalty points if they made a second purchase on the same day (independent of amount), whereas the second half (4516) kept shopping without getting a real-time mobile message.

A consumer can make many purchases during a shopping excursion to a shopping center. The mall operator uses cutting-edge technology to collect customer and transaction information. In its upscale mall, the mall operator only permits clients to make purchases via the mall's mobile application.

The information is organized into two tables: Consumers table and Purchases table. The customer table records demographic data such as age, gender, and loyalty status. In addition to capturing customer purchases from the first and second retailers, the purchase table also indicates if the consumer received a real-time mobile message. Both the tables have unique customer Id assigned for identification purposes.

## 1.1. Descriptive Analytics

For our data, we analyze the different variables in order to best explain the implications.

For our categorical variable that takes two values: Male (2,129) and Female (6,903)

For our continuous variables of gender consumers on age we get Female with a M (mean) =31 , SD (standard deviation) = 9.85 , min (minimum) = 14 , max (maximum) = 66, while Males received a M (mean) =33 , SD (standard deviation) = 8.38 , min (minimum) = 18 , max (maximum) = 66,

## 1.2. Predictive Analytics

**Regression Analysis**

For the first regression, we trying to predict the impact of age on second store sales where our independent variable was age and dependent variable is second store sales.

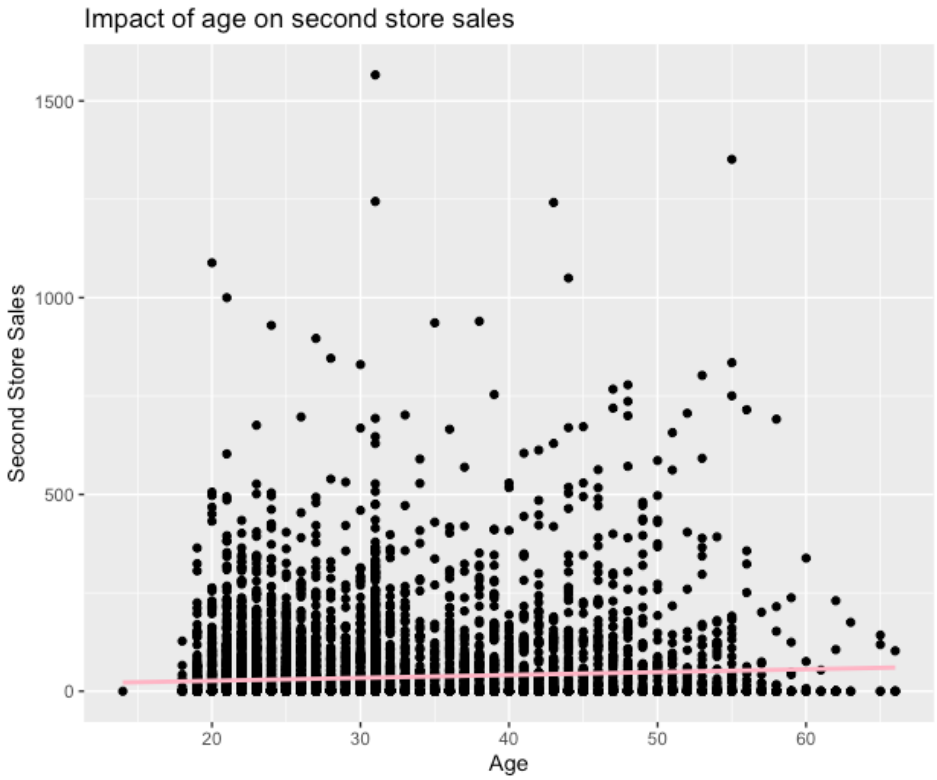
Our regression equation: Y= (second store sales) + (age)X + e,

R code- -lm(from_second_store_sales ~ age, data = experiment)

Here we can see the equation indicates our independent variable (age) is the slope (bX) and will show the rate of change to its dependent variable (second store sales) which is the y intercept(a).

Using the model diagnostics will help us understand the full impact in our independent variable on our dependent variable aswell. P- value generally tells us if there is enough correlation of the independent variable in the dependent variable where we can reject the null hypothesis. Our findings for age of second store sales show a p-value of 1.383E-12 which is < 0.05 indicating there not much correlation from age.

R value- shows what percentage of variation of the independent variable affects our dependent variable. For our regression analysis we found that 0.5435% of age variation of the second store sales. F-value shows how well your model fits into the data. For our regression analysis we found that 50.35 on 1 and 9030 DF for impact of age on sales



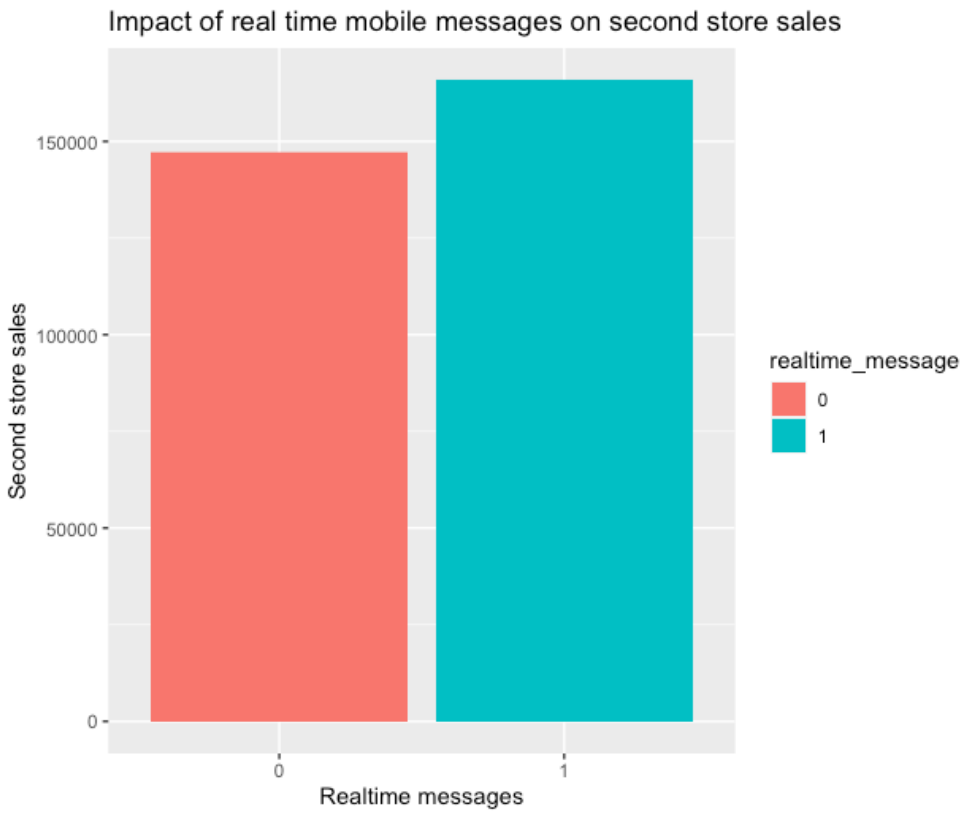

Second regression equation: Y= (second store sales) + (real time message) X + e,

R code- - lm(from_second_store_sales ~ realtime_message, data= experiment)

Here we can see the equation indicates our independent variable (real time message) is the slope (bX) and will show the rate of change to its dependent variable (second store sales) which is the y intercept(a).

Our findings for real time message of second store sales show a p-value of 0.0347 which is >0.05 indicating there is a significant amount of correlation from real time messages.

R value- shows what percentage of variation of the independent variable affects our dependent variable. For our regression analysis we found that 0.04016% of real time message variation of the second store sales. F-value shows how well your model fits into the data. For our regression analysis we found that 4.629 on 1 and 9030 DF for impact of age on sales




**Clustering analysis**

For cluster analysis, we analyzed 3 different cluster groups where we can see the greatest amounts of second store sales related to age. Analyzing the data the cluster was separated in to ranges of second store sales, where it was clear to see that there is less second store sales as age increases. We can also see that clusters for the highest second store sales showed a greater impact to younger age group between 20 and 50 years.

Regression model in the form Y {Cluster 1, Cluster 2, Cluster 3} = f (X1, X2) (1 sentence)

Y is a dependent variable in this case (second store sales) that we wish to predict. Referring to Y as Clusters and subgroups as cluster1, cluster2.

X is an independent variable(age) ; we refer to it as features.


Clusters on Age and Second Store Sales
from_second_store_sales
1500
1000
500
0
20
30
40
50
60
age
Clusters
1
2
3

## 4.RECOMMENDATIONS AND CONCLUSION

From the data that have been collected I conclude that the real time mobile message has lot of impact on the customers intentions to keep them shopping as there is a slight impact of age and real-time message on sales. With one unit increase in the age of the consumer, there was 0.73 increase, and with one unit increase in real-time messages sent, the sales increased by 4.15 times. The cluster with average age of 36.32 has highest sales of 566.53, followed by average age of 31.98 with sales of 176.90 and lastly, average age of 30.84 has sales of 8.53. Therefore, it is recommended to target younger population and send messages to increase sales.

1. I recommend my manager to send messages to larger population.
2. It is highly recommended to send the messages to younger ones to increase sales as we have clearly seen the impact of age on the sales.
3. We can add more products to the men to broaden the market
4. The sale should be sent to the customers more i.e. through the messages in an effective way so that every customer will be more likely to buy the product.