Sample App Analysis

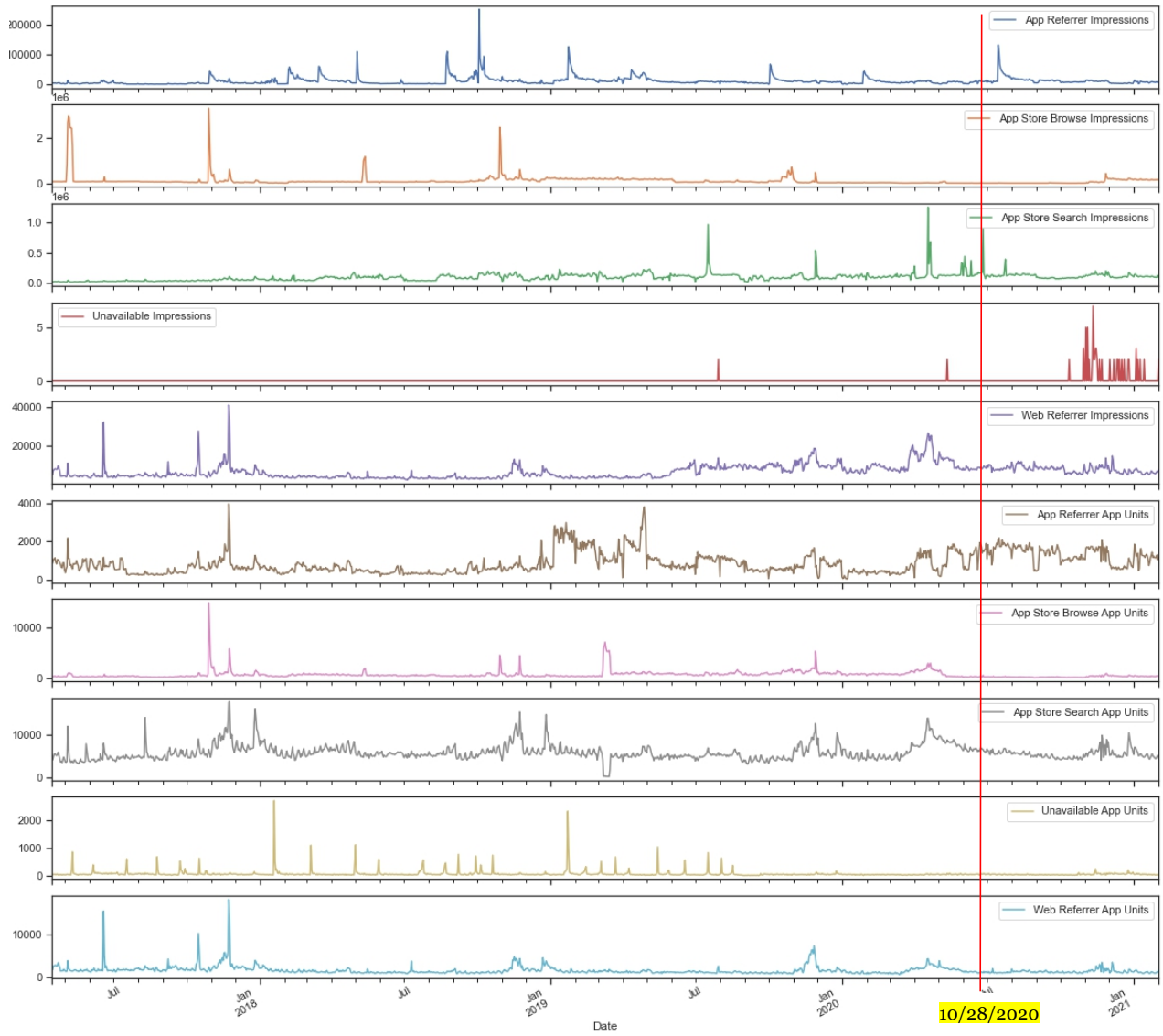
Rachel Ren 2020/03/16

Data Description

This data set consists of impression and app units data from different channels of sample app from apple app store.

Time Period: 4/15/2017 - 2/1/2021
Number of Fields: 11
Channels: App Referrer, App store Browse, App store Search, Unavailable
Number of Observations: 1,389

1. Trend plot from 4/15/2017 - 2/1/2021:



From the trend line plot by time, we can have a general picture about sample app's impressions and downloads information based on different channels:

1. Impressions and app units did not have a significant surge or drop overall. But app referrer impressions, app store browse impressions and app referrer impressions have outliers of a high number.
2. After campaign, unavailable impressions increased.
3. After campaign, app referrer impressions and app referrer app units sightly increased.
4. App store search app units dropped after campaign and then increased.
5. App referrer app units, app store search app units web referrer impressions have a large fluctuations by time.

2. Statistical indicators from 4/15/2017 to 2/1/2021:

Numbers of records: 1292

	App Referrer Impressions	App Store Browse Impressions	App Store Search Impressions	Unavailable Impressions	Web Referrer Impressions	App Referrer App Units	App Store Browse App Units	App Store Search App Units	Unavailable App Units	Web Referrer App Units
mean	11756.9	120477.9	99951.9	0.0	6432.4	831.8	766.8	5813.5	59.2	1675.1
std	14402.2	244103.6	71760.7	0.1	3805.9	552.3	786.5	1782.2	138.0	1063.1
min	1037.0	19308.0	21437.0	0.0	2117.0	32.0	143.0	260.0	0.0	718.0
median	8610.0	76219.0	93915.0	0.0	5216.5	660.0	616.5	5427.5	35.0	1403.5
max	249327.0	3321598.0	1255685.0	2.0	41237.0	3958.0	14833.0	17605.0	2718.0	18323.0

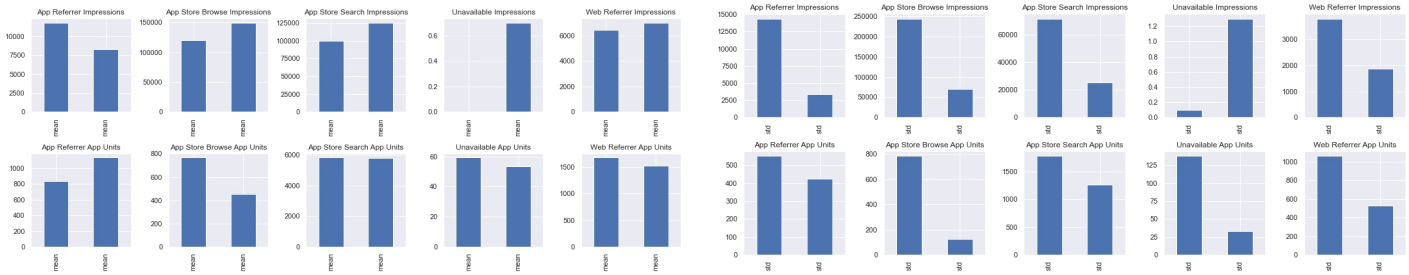
3. Statistical indicators from 4/15/2017 to 2/1/2021:

Numbers of records: 97

	App Referrer Impressions	App Store Browse Impressions	App Store Search Impressions	Unavailable Impressions	Web Referrer Impressions	App Referrer App Units	App Store Browse App Units	App Store Search App Units	Unavailable App Units	Web Referrer App Units
mean	8235.1	149222.9	125021.2	0.7	7017.5	1134.6	453.8	5793.5	53.2	1524.5
std	3410.6	69806.6	25452.1	1.3	1883.6	423.3	125.4	1268.8	32.6	526.0
min	2813.0	21433.0	90740.0	0.0	4805.0	470.0	162.0	4062.0	8.0	958.0
median	8572.0	168387.0	117143.0	0.0	6667.0	1140.0	442.0	5373.0	47.0	1372.0
max	15062.0	446410.0	206392.0	7.0	14617.0	2052.0	977.0	10457.0	237.0	3573.0

4. Mean/Standard Deviation Comparison before and after campaign:

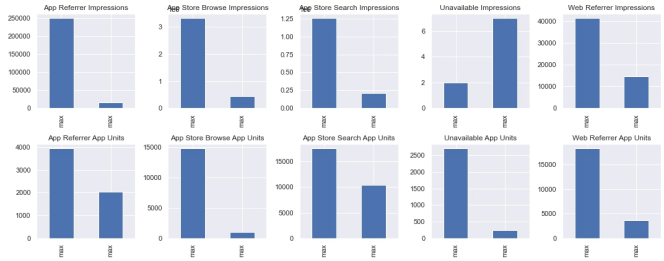
Notes: For each plot, first bar represents the statistical indicator (mean & standard deviation) before campaign change and the second is after campaign change.



From the average plots, we can tell that:

1. After campaign the average of impressions increased from app store browse, app store search, web referrer sources.
2. After campaign the average of app units dropped from most of the sources.
3. After campaign, the average of impression from app referrer dropped but the app unit increased.
4. From the average indicator, we can tell that even though the campaign increased impressions but app units sightly dropped.

5. Min/Max impressions and app units Comparison before and after campaign:



From standard deviation plots:
It's obvious that the numbers of impressions and app units are less spread out after campaign. We have more steady and consistent impressions and app units than before.
From the min/max numbers of impression and app units, we can tell that:
1. we did not have a extreme small or large number of impressions and unit apps after campaign. Max number dropped a lot.
2. But on the other hand, we'd better to think why before campaign we can have reach a higher numbers such as 3321598 of app store browse impression, but after campaign we only have 446410. Why we did not have a high number like before.

Data Cleaning

1. Check missing values: None

2. Check data types: One date time, 10 continuous variables

3. Hypothesis Test:

Indicators	P-Value	Results
App Referrer Impressions	0.0163	Reject null hypothesis, different distributions
App Store Browse Impressions	0.2477	Accept null hypothesis
App Store Search Impressions	0.0006	Reject null hypothesis, different distributions
Unavailable Impressions	0.0000	Reject null hypothesis, different distributions
Web Referrer Impressions	0.1336	Accept null hypothesis
App Referrer App Units	0.0000	Reject null hypothesis, different distributions
App Store Browse App Units	0.0001	Reject null hypothesis, different distributions
App Store Search App Units	0.9133	Accept null hypothesis
Unavailable App Units	0.6657	Accept null hypothesis
Web Referrer App Units	0.1667	Accept null hypothesis

Before campaign and after campaign, market campaign has statistically significant on App Referrer Impressions
Market campaign has statistically significant on App Store Search Impressions, Unavailable Impressions, App Referrer App Units, App Store Browse App Units

4. Check outliers and correlation

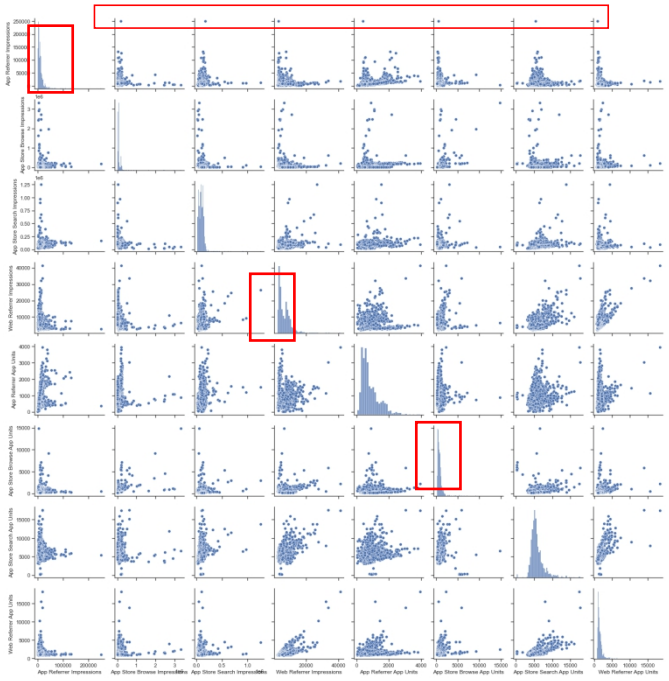
Notes: Table shows the numbers of outliers before campaign and after campaign.
*Outliers calculation: Q1-1.5*Q or Q3+1.5*Q (Q1: first Quantile point, Q3: third quantile point, Q: 1 quantile)*

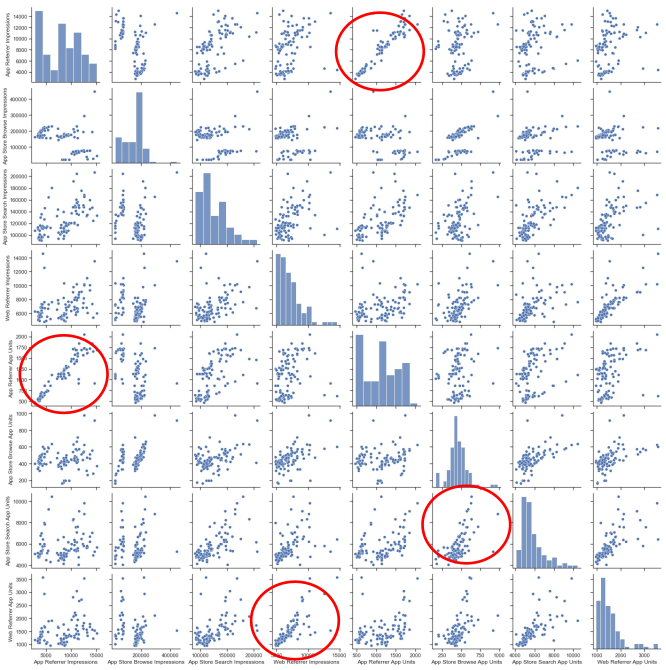
	App Referrer Impressions	App Store Browse Impressions	App Store Search Impressions	Unavailable Impressions	Web Referrer Impressions	App Referrer App Units	App Store Browse App Units	App Store Search App Units	Unavailable App Units	Web Referrer App Units
Before campaign	158	117	26	3	47	82	80	33	128	48
After Campaign	0	11	0	27	1	3	0	0	6	1

App store search channels have less outliers.

4.1 Pair plots:

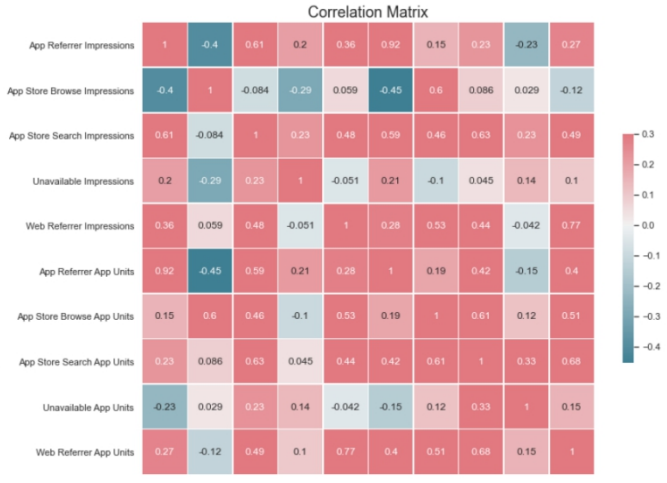
Notes: The left pair plot shows the data distribution before campaign, the right shows after.





From the pair plot, we can tell before the campaign, most of the data points distribution were left-skewed which means there are many outliers. Also, we can tell several strong correlations between different variables.

4.2 Correlation metrics:



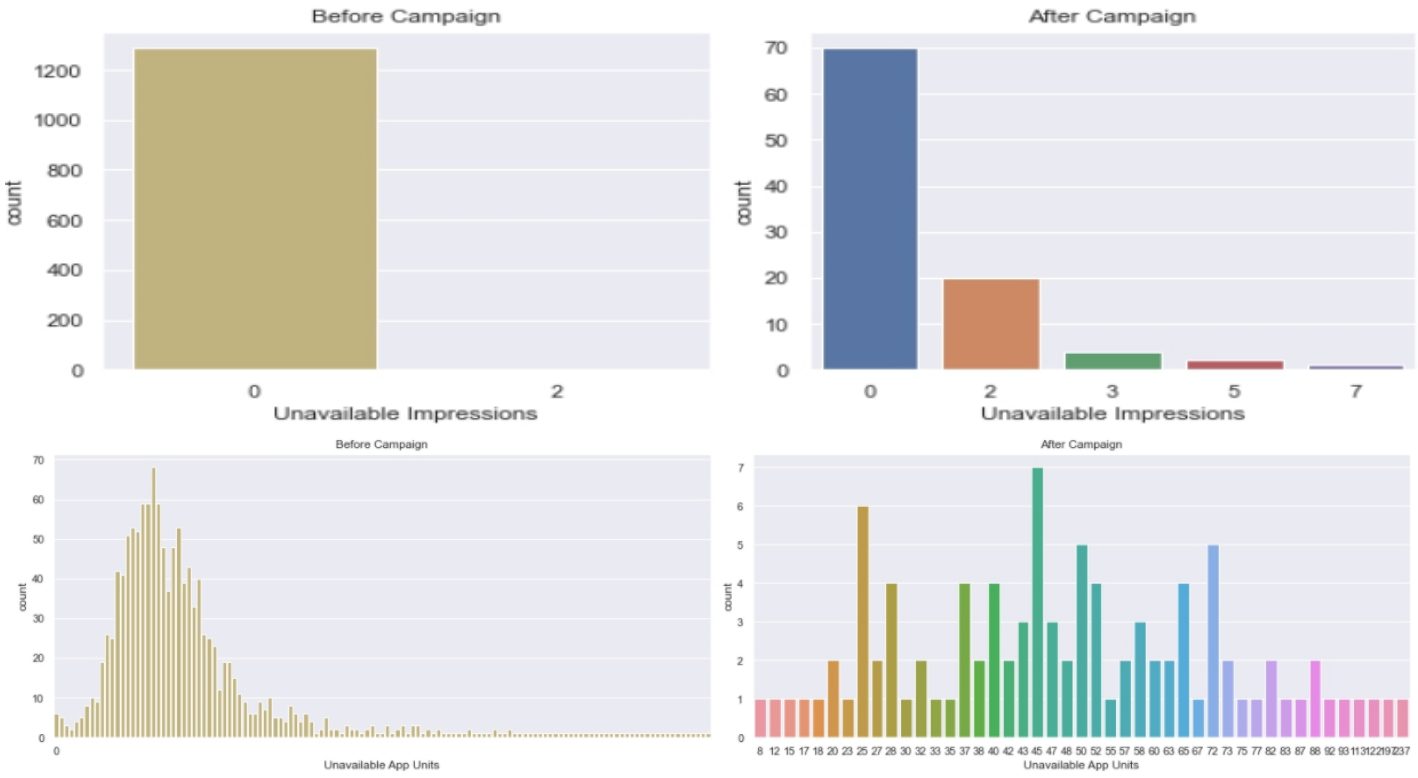
After changed campaign, the correlation between different channels are strong than before.
There are some strong positive correlations:

App Store Browse Impressions	**App Store Browse App Units**	0.603779
App Store Search App Units	**App Store Browse App Units**	0.606043
App Store Search Impressions	**App Referrer Impressions**	0.606909
App Store Search Impressions	**App Store Search App Units**	0.633323
App Store Search App Units	**Web Referrer App Units**	0.684156
Web Referrer Impressions	**Web Referrer App Units**	0.772404
App Referrer Impressions	**App Referrer App Units**	0.917403

From the correlation map, we can tell that after campaign, the impression and app units have stronger positive correlation between each other, which means that once we gain more impressions or downloads, we will gain more impression or downloads from other sources too.
App search channels are influencing web referrer, app referrer, app browse in a positive way.

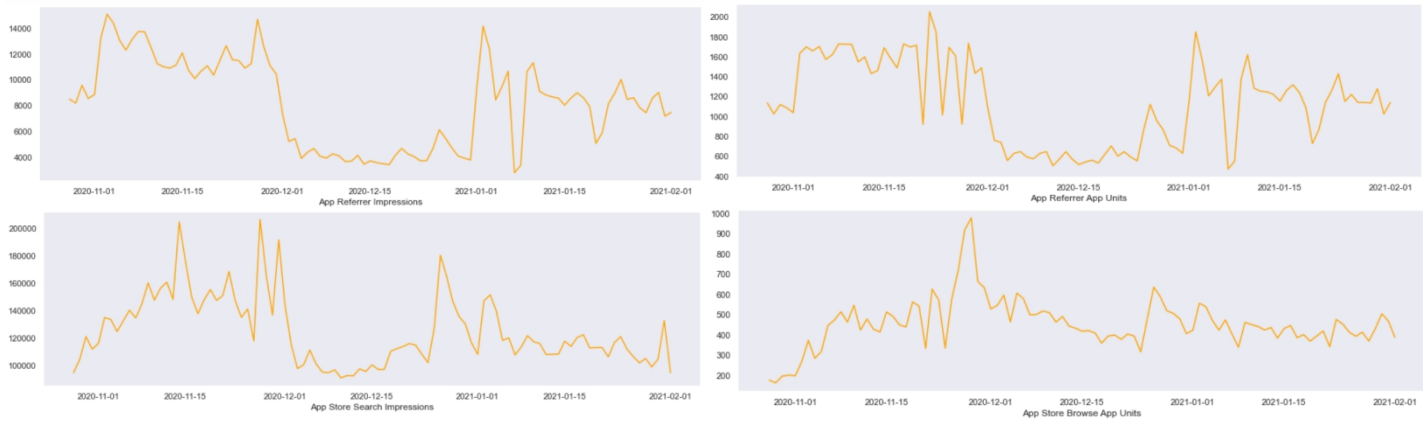
Variables Analysis:

1. Unavailable Impressions and unavailable app units:



There are large numbers of unavailable impressions such as 3, 5, 7 after campaign.
Before campaign, there are more outliers of unavailable app units.

2. Variables that has statistical significant:



App referrer impression, app referrer app units, app store search impression and app store browse app units have the similar trend. Those indicators increased once the market campaign launched, then quickly dropped from 1st Dec until around 25th Dec then sightly increased

3. Compare with 97 days before campaign

	App Referrer Impressions	App Store Browse Impressions	App Store Search Impressions	Unavailable Impressions	Web Referrer Impressions	App Referrer App Units	App Store Browse App Units	App Store Search App Units	Unavailable App Units	Web Referrer App Units	Total impressions	Total app units
Before	1353896	2128563	11177232	2	775489	132618	23464	491723	2554	127067	15435182	28081245
After	798806	14474617	12127058	69	680695	110052	44019	561965	5156	147874	777426	869066
Difference	-69%	85%	8%	97%	-14%	-21%	47%	12%	50%	14%	45%	11%

Compared to the data from 97 days before campaign with 97 days after campaign, we can tell that total impression increased 45%, and total app units increased 11%

Summary

1. Overall, impressions and app units did not have a significant surge, steady increase or significant drop.
2. App referrer impression, app referrer app units, app store search impression and app store browse app units have the similar trend. Those indicators increased once the market campaign changed, then quickly dropped from 1st Dec until around 25th Dec then sightly increased.
3. App search channels have a positive correlation with web referrer, app referrer, app browse.
4. Less outliers of unavailable app units.
5. Based the comparison of the mean of the number between before campaign and after, we can tell that even though the campaign increased impressions but app units dropped.
6. There are five indicators has statistically significant between before campaign and after. In other words, the launch of the campaign has a significant impact on those five indicators. They are App Store Search Impressions, Unavailable Impressions, App Referrer App Units, App Store Browse App Units.
7. Based on the comparison of the 97 days before and after campaign change, we can tell that this campaign change has a positive impact overall. The total impressions and total app units are increased by 45%, 11% respectively. Most of the channels are increased. App referrer decreased.