

# **BINARY.COM**

Data Scientist – Recruitment Exercise

## **REPORT**

Name: Syed Arham Ali Rizvi

**Question:** The management of sell-your-stuff.com wants to set up marketing strategies for the next fiscal year based on the summary tables. **Your job is to analyse the performance of past campaigns, coupons, profit by countries, etc. and make a report showing information to support your insights.**

## Analysis:

**NOTE:** We are looking at data from 2017-2018. Please refer to the appendix (at the end of the report) for more details.

### **Past Campaigns**

To analyse the performance of past campaigns, we must question what type of a campaign it is. The campaign:

1. Maybe aimed to increase the number of clients on the website (awareness)
2. Or it may aim to increase the profit generated
3. Or maybe both of the above

**NOTE:** Before we move on to how ‘successful’ a campaign was, it is important to note that a considerable number of factors (which might not be exhibited in the data) might affect how a campaign went. Factors such as:

1. The **timing** of the campaign. Consider a campaign done during another big event in the region. For example, if an important football match is taking place during that time, the campaign can choose to capitalize on it. With that being said, it could possibly go the other way too such as an unexpected tragic event in the region in which case the campaign might go in loss.
2. The willingness/advancement of technology (**acceptance**) of the audience to use the website properly and take advantage of the campaign.
3. Political **stability/instability** in the region.
4. Various government **rules**.
5. Degree to **freedom** of trade (import/export of the products).
6. **Marketing** strategy used to attract the audience.
7. **Buying and selling power** of the population.
8. Currency conversion rate and level of **inflation** in various countries.
9. The **population** of a country.
10. Other **competitors**.
11. The list goes on...

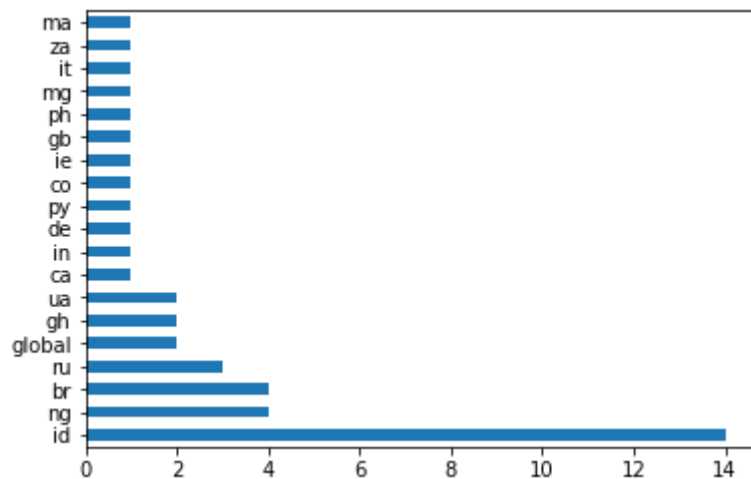
To better understand the data, the ‘Number of days’ a campaign lasts was calculated (Figure 1.0):

	<i>start_date</i>	<i>end_date</i>	<i>total_spend</i>	<i>country</i>	<i>no_of_days</i>
0	2017-01-03	2017-01-04	9935.67	Global	<b>1 days</b>
1	2017-01-04	2017-01-06	17696.41	id	<b>2 days</b>
2	2017-01-08	2017-01-12	22541.26	in	<b>4 days</b>
3	2017-02-13	2017-02-18	745.89	ng	<b>5 days</b>
4	2017-05-02	2017-05-03	15338.78	id	<b>1 days</b>

**Figure 1.0: Campaign table (head)**

<b>No. of days</b>	<b>Frequency</b>
<b>2 days</b>	15
<b>4 days</b>	9
<b>3 days</b>	7
<b>1 days</b>	6
<b>0 days</b>	5
<b>5 days</b>	1

**Figure 1.1: No. of days campaign lasts & their frequency**



**Figure 1.2: No. of campaigns per country**

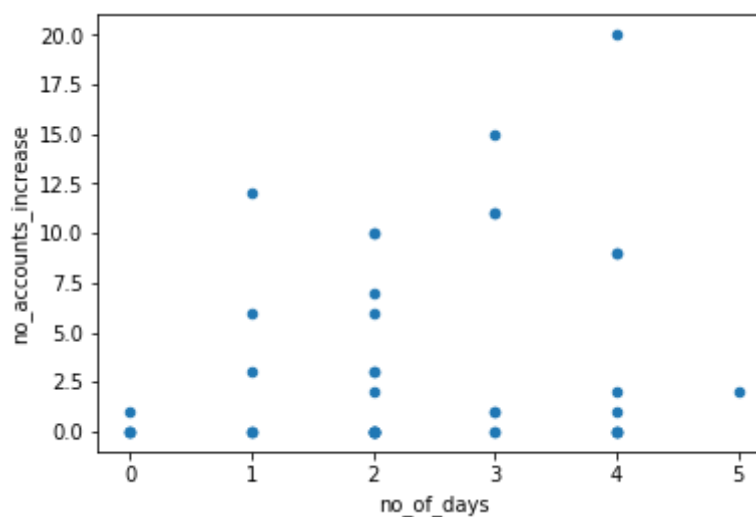
From the figures above, we can see that:

1. Indonesia had the most campaigns, approximately around 14 times.
2. Following it were Nigeria, Brazil and Russia.
3. Most campaigns were for 2 days.
4. The average campaign lasted for 2 days.

Country	no_accounts_increase
<b>Id</b>	124
<b>Br</b>	12
<b>Ru</b>	5
<b>Ng</b>	4

**Figure 1.3: No. of accounts increase per country**

Does the number of days a campaign last, affect the number of clients that joined?



**Figure 1.4: No. of days & Acc. increase**

Figure 1.4 shows that there is no such correlation between the increasing number of days and the increase in account numbers.

Does the frequency of campaigns, affect the number of clients that joined?

According to figures 1.2 & 1.3 the data presents that the increasing frequency of campaigns did in fact have a **positive** impact on the number of the clients joining the website from their respective countries.

Through the campaigns, only **145** new clients joined the website which is only about **2.9%** of the 5000 accounts registered on the website.

From my intuition, there can be 2 possibilities:

1. The campaigns might not have particularly aimed towards increasing the number of clients on the website although any added number of clients was a bonus.
2. The company is a start-up (since it only has 5000 accounts registered) and might be trying to test the market in several countries.

If the aim was not on increasing the number of clients, it might be possible that the company was trying to generate a profit.

### **Profit Analysis**

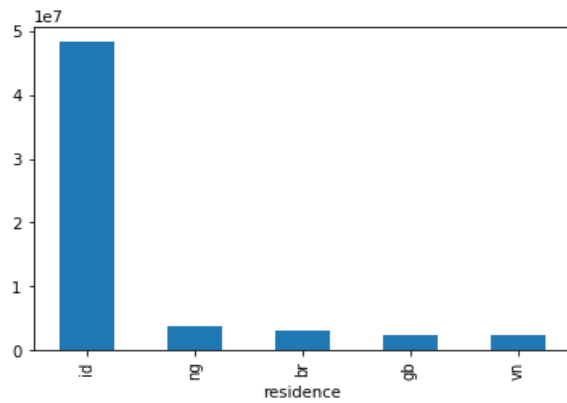
To calculate the profit (general steps):

1. The total revenue was calculated
2. 1% of the total revenue was only considered since 'the company sells the data by roughly 1% of the total clients have bought'
3. From that amount, the expenditure (amount spent on campaigns) was deducted and the final profit presented.

Profit Analysis of the company based on the whole year:

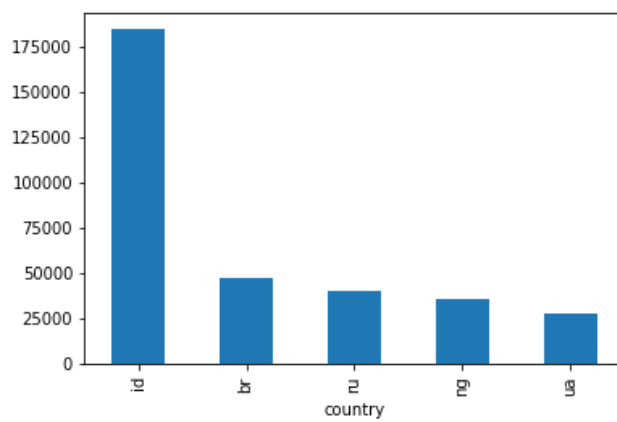
1. The total **Revenue** the company earned the whole year was **77189298.38**
2. The total **Amount spent on campaigns** was **516208.68**
3. The total **Profit** the company earned the whole year was **255684.30**
4. Only **18** countries had **43** campaigns in them while the database contains accounts from **137** countries.

### Profit Analysis per country:



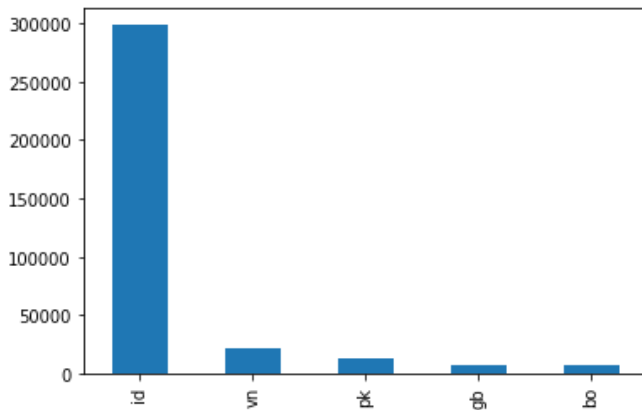
Country	Revenue
Id	48342021.31
Ng	3640249.68
Br	3074899.40
Gb	2440897.59
Vn	2234566.03

**Figure 1.5: Total revenue per country (Top 5)**

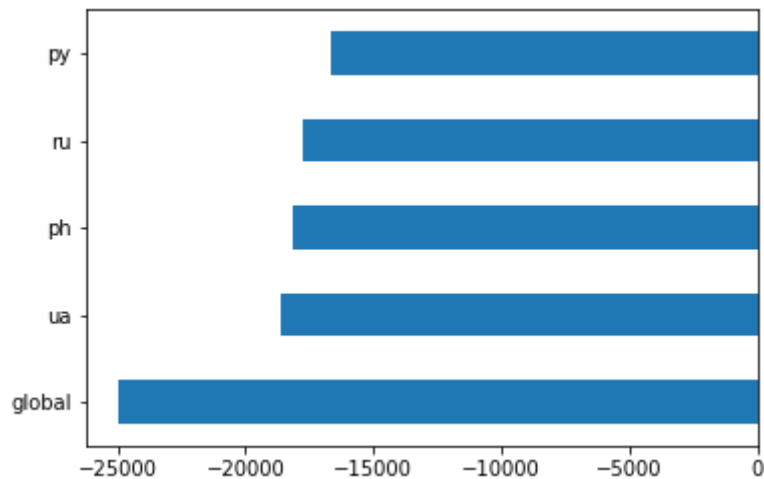


Country	Campaign spent
id	184919.92
br	47229.13
ru	40109.58
ng	35351.00
ua	27302.04

**Figure 1.6: Total spent on campaigns per country (Top 5)**



**Figure 1.7: Profit per country (Top 5)**



**Figure 1.8: Loss per country (Top 6)**

From the figures above, we can conclude the following:

1. Indonesia has generated the most revenue, spent the most on its campaigns and earned the most profit compared to any other country.
2. Even though the company has made an overall profit of 255684.30 (as seen earlier), it has incurred losses in a few countries.
3. Brazil, Russia, Nigeria and UAE were amongst the countries that spent the most on campaigns (as well as conducted the most number of campaigns). Even though Nigeria and Brazil were able to generate some considerable amount of revenue, UAE, Russia and Brazil were still in the top 6 countries that went into loss.



4. Amongst the top 10 countries that made a profit, most were in the Asian region (5/10), 2 of them in Europe, 2 of them in Africa and 1 in South America
5. From the top 5 countries that generated a profit, only United Kingdom had a campaign in their country which was just for a day.

## Conclusion

From the statistics, it seems like the company:

1. May be a start-up, as it only has 5000 accounts, which is scarce and unsubstantial for an ecommerce website aiming to capture the markets all around the globe.
2. Maybe an Indonesian company or concentrated in Indonesia as it invested there considerably and received the most desirable response.
3. The Asian region made the most profit.
4. The campaigns may be used to create awareness in countries such as UAE, Russia, Brazil and Nigeria since significant amount of money was spent there while for Indonesia the campaigns might have been aimed to generate profit since the company has already captured the market.

## FUTURE WORK

1. The company can determine what the goal for their campaign is for example if the campaign's goal is to raise awareness, make profit or both.
2. Once the goal is known, classify for all the previous campaigns as successful or unsuccessful.
3. A decision tree or a random forest model can be trained to make a prediction whether based on various variables, a campaign in some country would be a success or a failure.
4. To go further than that, a regression model (e.g. random forest regressor model) can be used to predict 'how much money should be spent', 'how much time spent', 'how frequent a campaign' etc., should be done to achieve a level of success without exhausting resources.

## Appendix

### Client table:

	account	indication_coupon	first_deposit_amount	first_transaction_amount	balance_amount
count	5000.000000	5000.000000	3959.000000	2606.000000	3.964000e+03
mean	3499.500000	280.059000	7972.617336	62.299371	-3.614863e+05
std	1443.520003	72.958655	5898.371677	296.155525	1.812420e+06
min	1000.000000	0.000000	1.390000	0.070000	-2.905925e+07
25%	2249.750000	308.000000	9090.910000	0.520000	-9.090909e+04
50%	3499.500000	308.000000	9090.910000	1.550000	1.964535e+03
75%	4749.250000	308.000000	9090.910000	9.089630	9.090910e+03
max	5999.000000	308.000000	100000.000000	6887.185000	3.551842e+05

1. first\_deposit\_amount mean: 7972.617336
2. first\_transaction\_amount mean: 62.299371
3. There are clients from 136 'known' countries
4. Most clients were from Indonesia (1667) followed by Brazil (438) and Nigeria (284)

From the data given we already know that there are 5000 unique accounts currently registered on the website. However, out of those 5000 accounts, 1041 accounts have never made a deposit (20.82% of the total number of clients) and only 3964 accounts have ever made a transaction. This may be the case because:

1. The accounts are new, and they are yet to make a transaction
2. The accounts are inactive

Indication coupon	Frequency
308	4229
41	61
116	48
36	20
181	18

### Campaign table:

total_spend	
count	43.000000
mean	12004.853023
std	6862.693494
min	745.890000
25%	7499.460000
50%	12463.880000
75%	16249.620000
max	29866.080000

1. Average amount spent on campaigns: 12004.85
2. Maximum amount spent on a campaign: 29866.08 (spent on russia)
3. Minimum amount spent on a campaign: 745.89 (spent on nigeria)
4. 2 global campaigns
5. Campaigns in 18 different countries

## Transaction table:

	account	total_buy	total_sell	total_dep osits	total_withdr awals	count_cont racts	count_dep osits	count_withdr awals
count	36747.00 0000	33169.000 000	32921.000 000	8242.000 00	16685.00000 0	36747.0000 00	36747.000 000	36747.000000
mean	3098.360 927	2327.1518 10	1807.7898 38	6860.939 06	88211.23585 9	255.985468	0.292786	0.084850
std	1361.246 105	10855.111 743	8532.6869 75	9210.638 11	31763.76768 0	4529.95017 6	0.707798	0.524256
min	1001.000 000	0.070000	0.030000	1.30000	0.200000	0.000000	0.000000	0.000000
25 %	1873.500 000	4.500000	3.860000	350.0000 0	100000.0000 00	3.000000	0.000000	0.000000
50 %	3153.000 000	28.190000	23.920000	9090.910 00	100000.0000 00	15.000000	0.000000	0.000000
75 %	4052.500 000	259.17000 0	217.18000 0	9090.910 00	100000.0000 00	68.000000	0.000000	0.000000
max	5999.000 000	100000.00 0000	100000.00 0000	100000.0 0000	100000.0000 00	248789.000 000	17.000000	33.000000

1. The average buyer spent: \$2327.15
2. The average seller sold: \$1807.79
3. The average number of items sold in a day: 255.98
4. Total clients that have ever made a transaction: 3964