ifood case

Marketing study of Gadget campaign

Robson Müller

Deliverables

presentation index

As a result of our study on the dataset provided by the marketing team. We achieved these final products:

1. Exploration of the 2240 customers in the beta Gadget campaign.
2. Found an outlier buying behaviour
3. Tree classification model

Let’s dive in each one

ifood

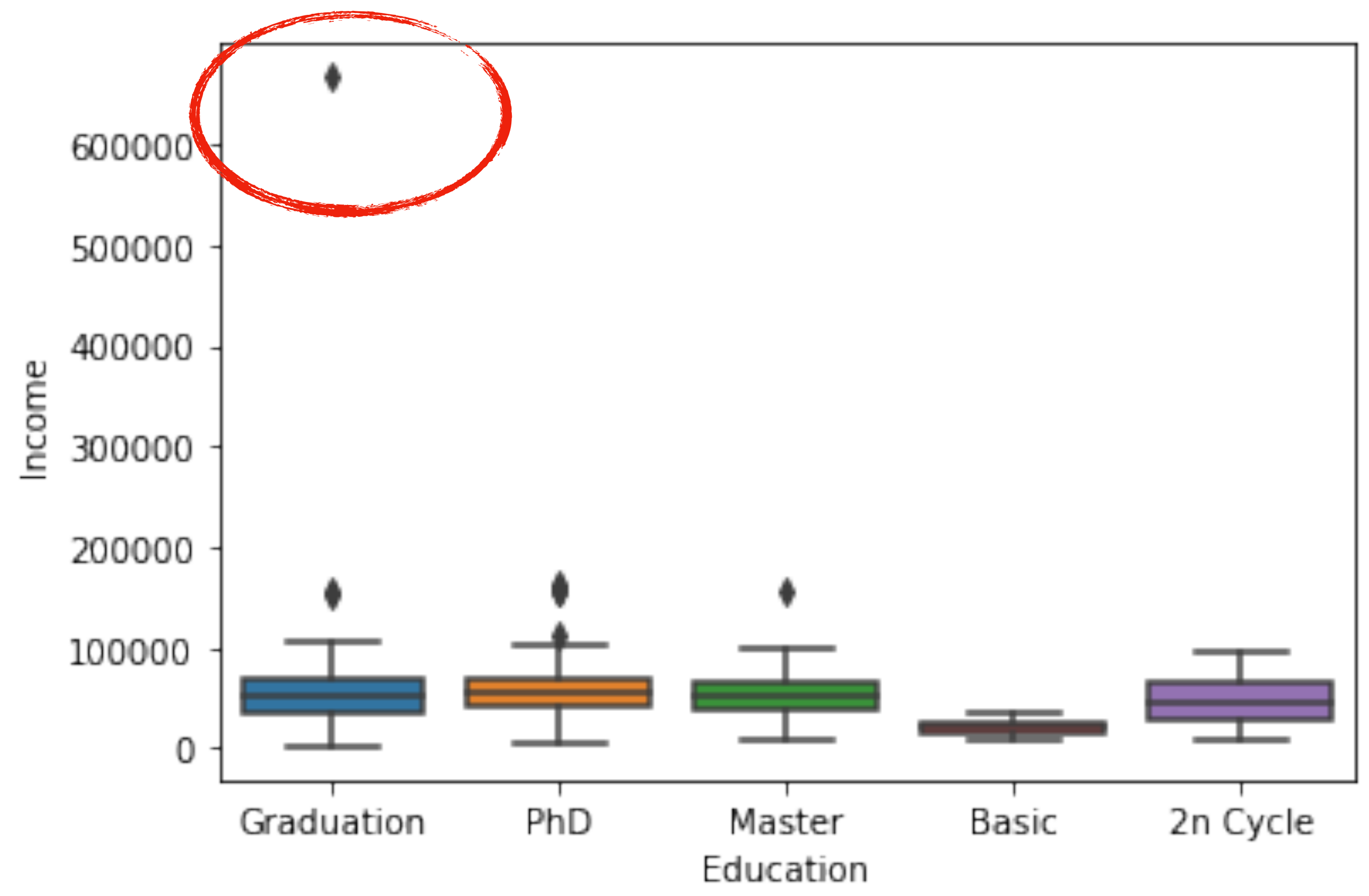
Data exploration

Data Exploration

Cleaning and Understanding

- As a first approach we checked the general statistics for the dataset, and found a very large standard deviation for some features.

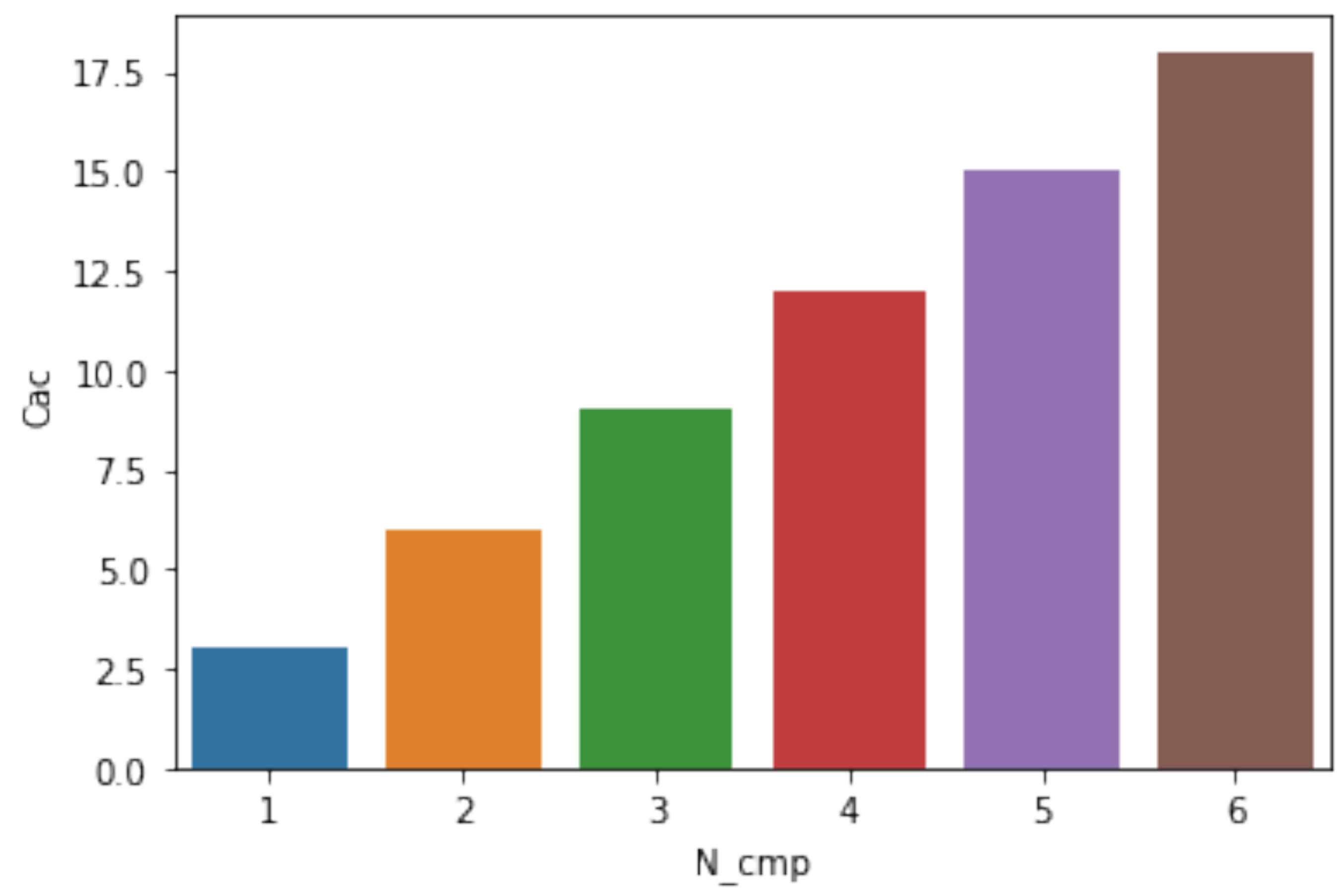
- One of them was Income, and we found an outlier with an Income of 1.176% bigger than the average. We cut him down in order to do not bias our analysis.



Data Exploration

Cleaning and Understanding

- We found two columns not described in the documentation.
  - Z_CostContact: Which we considered as being the cost in each marketing attempt/ campaign
  - Z_Revenue: Which we considered as being the revenue generated in each buy of the offer.



Having both these numbers, we estimated the CAC for each costumer, considering how many attempts were necessary until him/her convert to the offer, and the revenue generated.

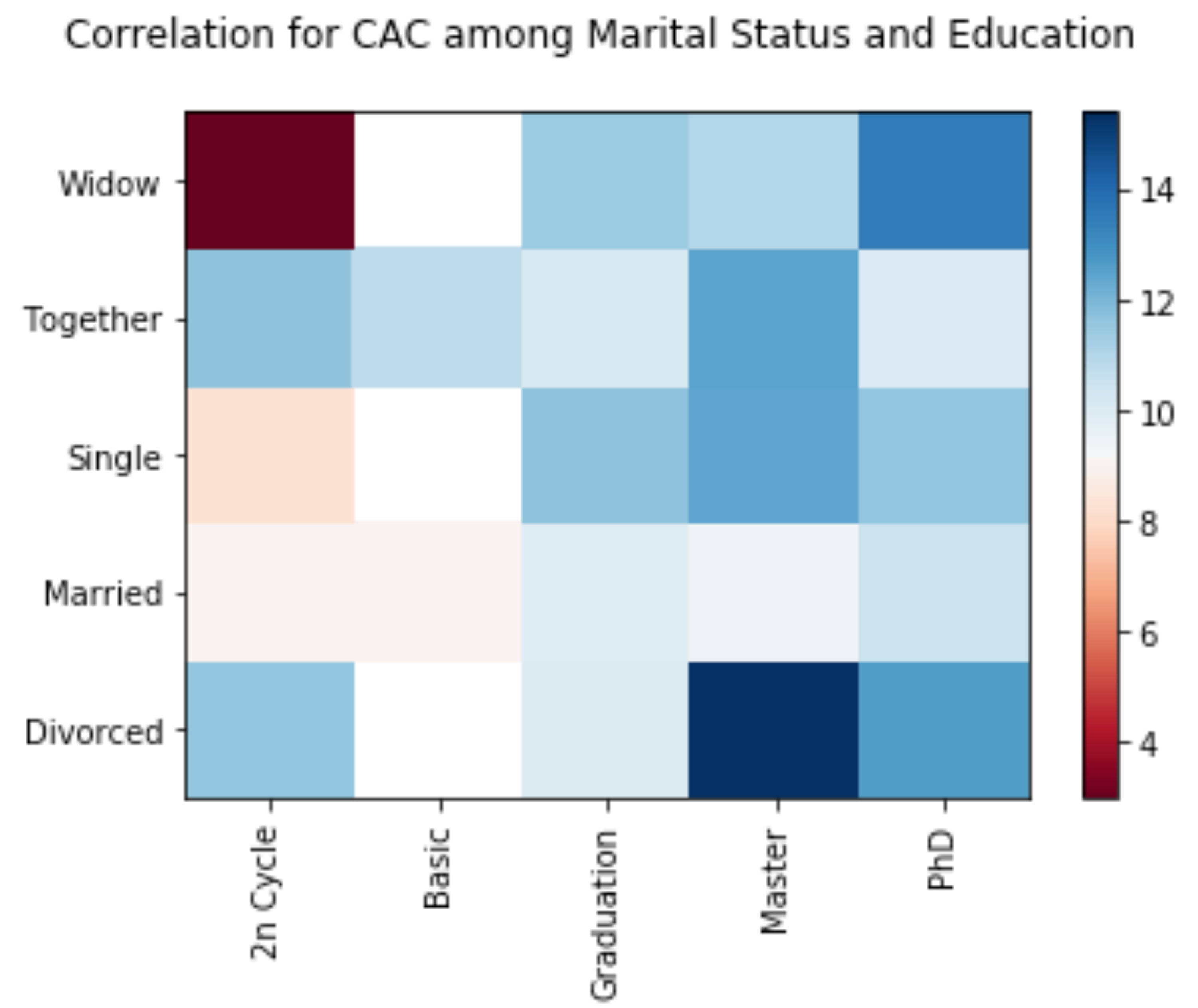
```
Conversion per Campaign

        23% in the 1st campaign
         3% in the 2nd campaign
        22% in the 3rd campaign
        18% in the 4th campaign
        10% in the 5th campaign
        24% in the 6th campaign
```

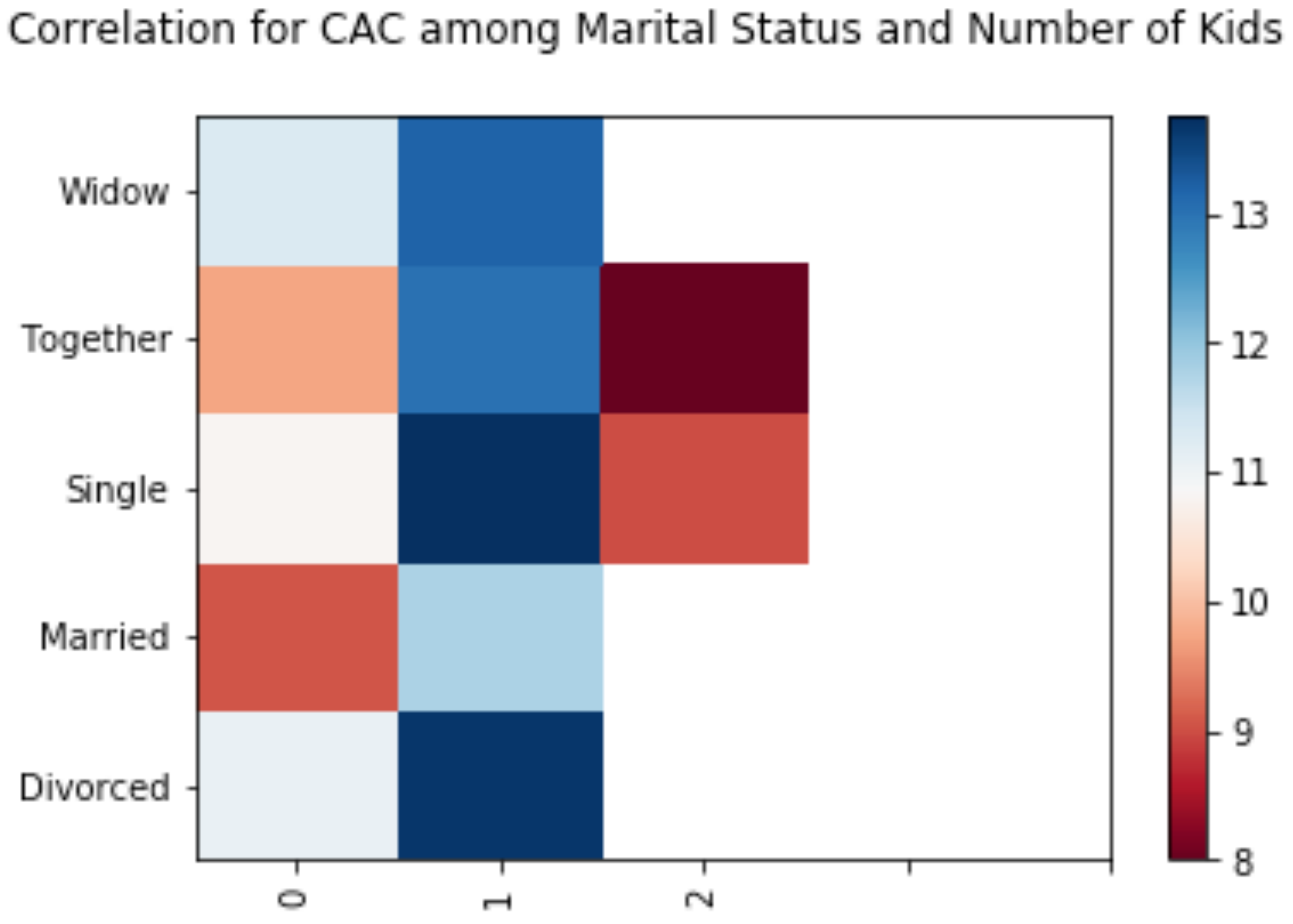
Data Exploration

Offer Acceptance

- Another important factor observer is that people with lower Education level and Income did not accepted the offer when they live alone (the white box means there is no CAC for these groups, in other words they are labeled as “no_conversion”)

- As a curiosity, customers with 1 kid at home, presented Higher CAC. Which means they were more resilient accepting the offers, the single and the divorced with 1 child, converted in the last campaing, summing the max CAC.

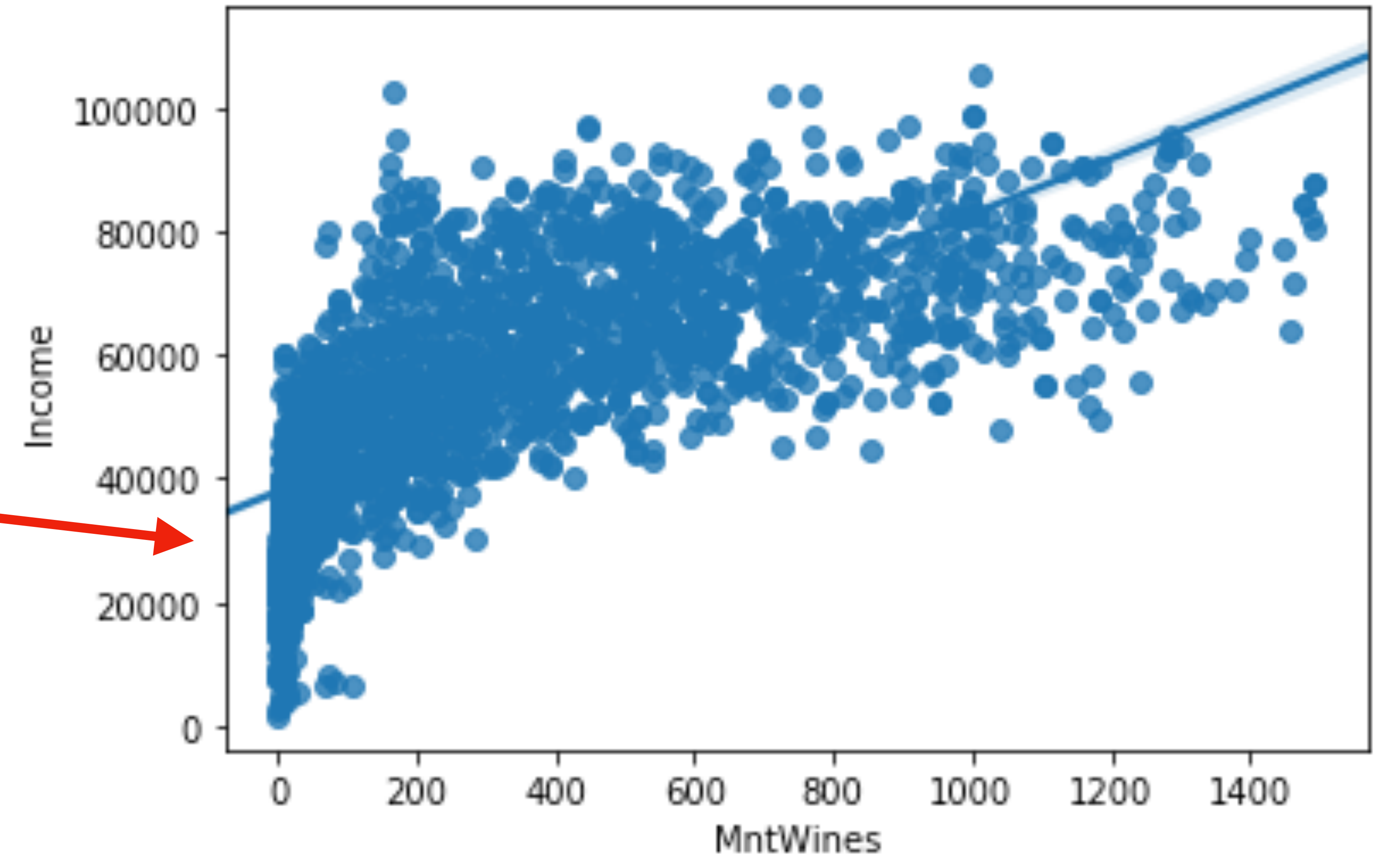




Data Exploration

Buying behaviour

- When customers have an Income higher than 30K their Wine consumption increases drasticaly.



- When customers accumulate more than 4 visits in stores their Wine consumption triple, and the meat consumption double.

Avg amount of Wine bougth by Income Group

	Income_binned	MntWines
0	Low Income	23.147651
1	Medium Income	311.590035
2	High Income	657.138229

Avg amount of Meat bougth by Income Group

	Income_binned	MntMeat
0	Low Income	23.187919
1	Medium Income	123.078671
2	High Income	452.997840

Avg amount of Going to Store by Income Group

	Income_binned	N_purchases_store
0	Low Income	2.979866
1	Medium Income	6.257867
2	High Income	8.399568

In a short way. That's it about the findings!

Feel free to ask me more detailed question in order to build some action plans to the next campaigns.

Let's move to the **PREDICTION MODEL**

Prediction Model

Prediction Model

Model Development

- For this specific approach, we choose the Tree Classifier.
- During the data cleaning, understating and preparation, we made a reverse hot encoding in the cmp feature, and realised that some customers did not buy the offer. So, for those we labeled them as “no_conversion”
- The classifier had the best result with a test set of 30% based on entropy criterion with max depth of 5
- It were evaluated with “metrics” from sk.learn. with an accuracy of 100%
- The final product of the classification where 7 possible labels, classifying in which campaing the customer would convert, or even if it would not.
- This can help the marketing department to estimate the exact return for each customer group, and even do not apply a similar campaign to customers which did not converted to it.