Bank Purchase Classification Case Study

Virtual Internship

20-Nov-2022

Background – Bank Purchase Classification case study

- ABC Bank wants to sell it's term deposit product to customers and before launching the product they want to develop a model which help them in understanding whether a particular customer will buy their product or not (based on customer's past interaction with bank or other Financial Institution).

- Objective: Analyze previous bank customer data to propose an efficient solution for ABC banks upcoming marketing campaign. Identify trends in the data to ultimately create a model to help predict which customers will be most likely to purchase the new product

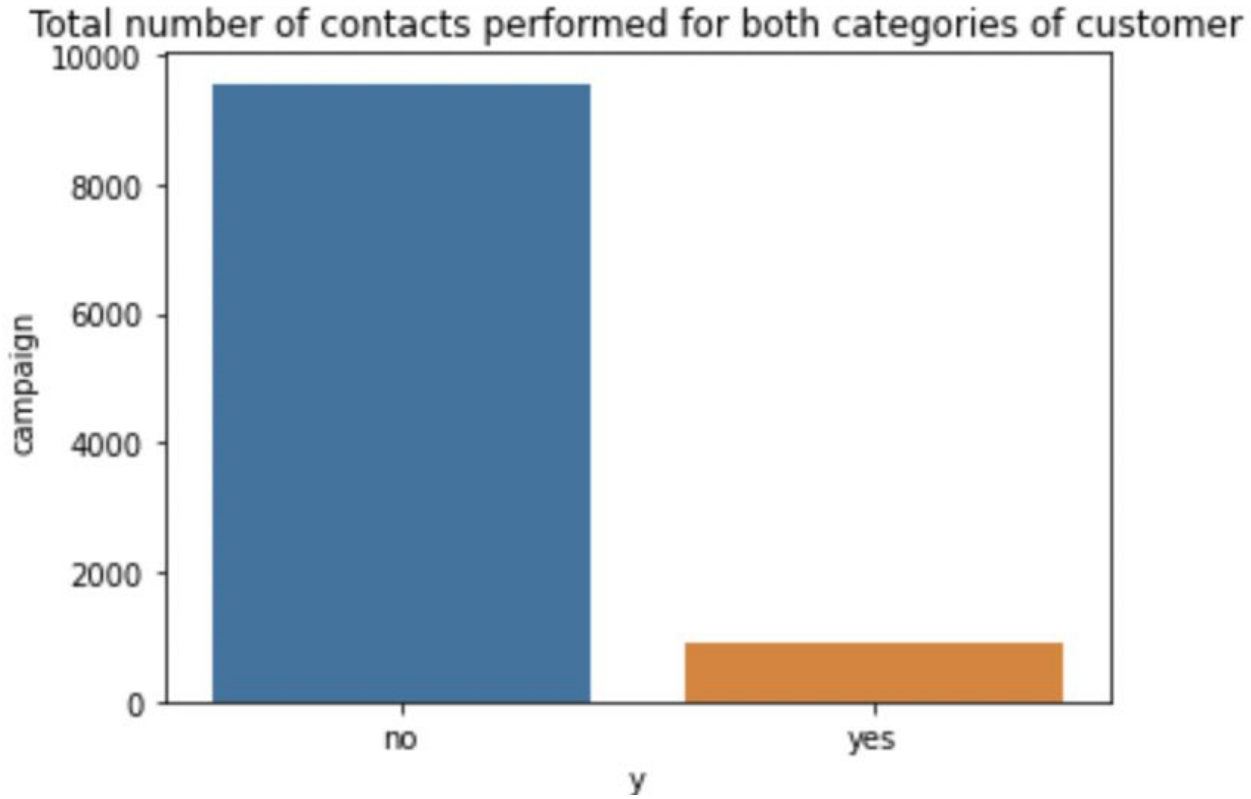
The analysis has been divided into four parts:

- Data Understanding
  - When to run ads, how to run ads
- Finding target groups
  - How we found the target groups
- Recommendations for model building

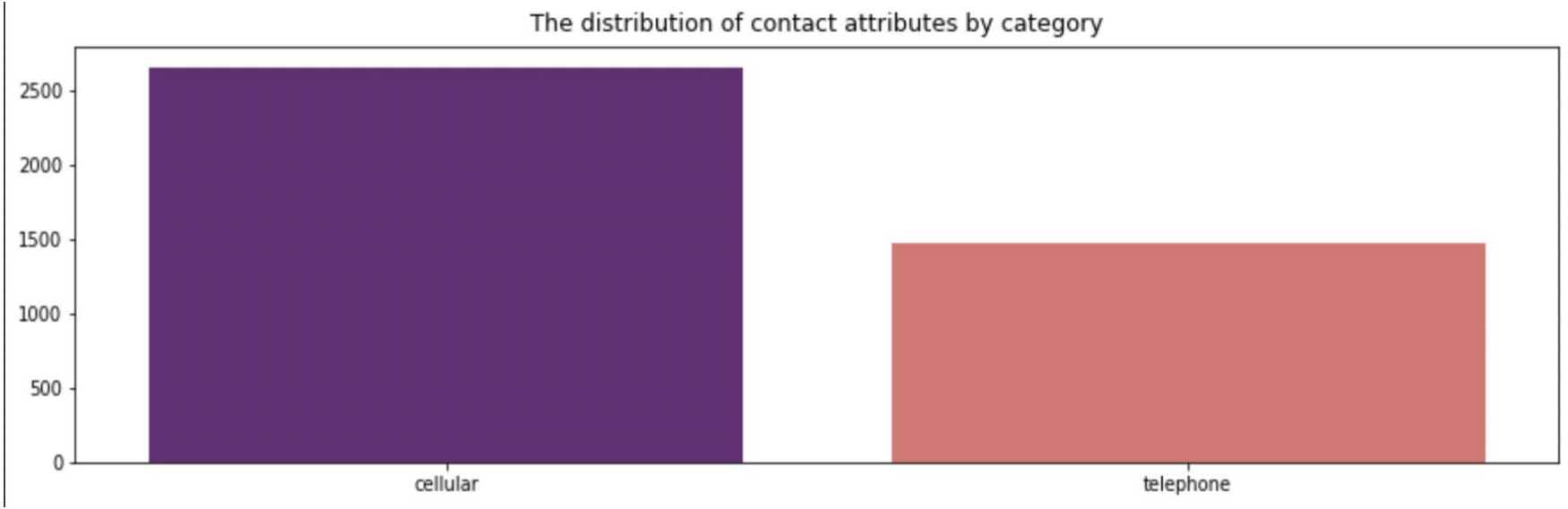
Background – Bank Purchase Classification case study

- It is most effective to advertise during company time (May is the best month to do so).
- In order to improve customer contact, other means of communication should be used.
- Focus on specific categories to increase the chances of your proposal being accepted by students or seniors.
- Socio-economic categories should be used to define target groups. A client's income level (not always high), age, and profession can accurately determine their marketing profile.
- Based on these factors, it is recommended to focus on more promising consumer groups.
- Nearly 25000 married customers and 15000 singles. Admin, blue-collar, and technician are the most popular professions.
- The campaign made nearly 25000 contacts for the three job roles. We can consider these three jobs as target groups.

Data Understanding



Data Understanding - Campaign types

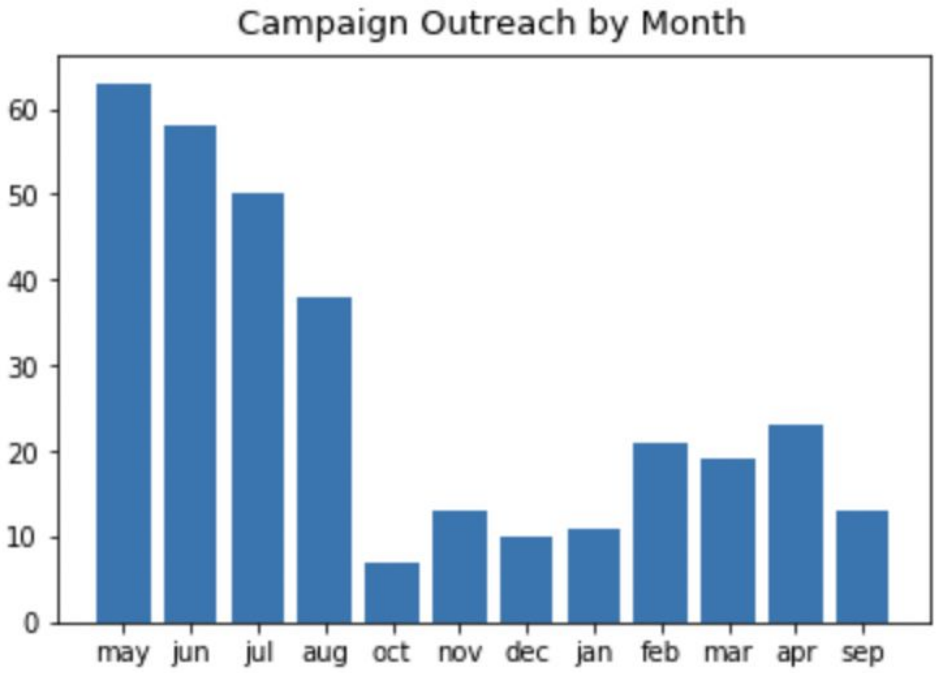


This graph shows the comparison of campaign reach by category and that the campaign reaches more than 50% more customers on a mobile phone compared to a telephone. This will help when determining what the target demographic will be.

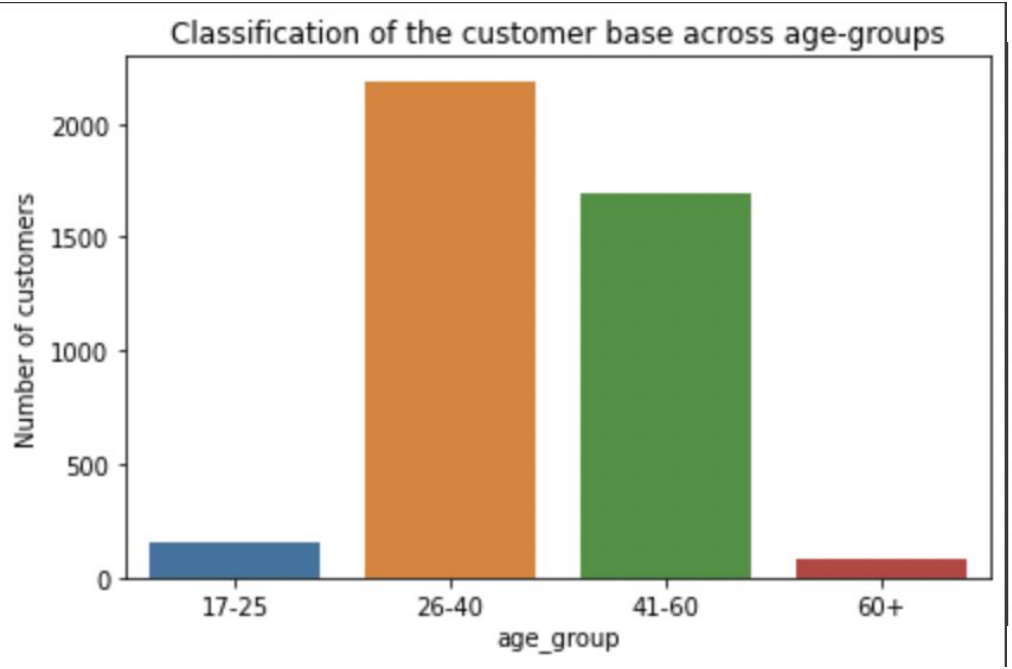
Data Understanding - Campaign types

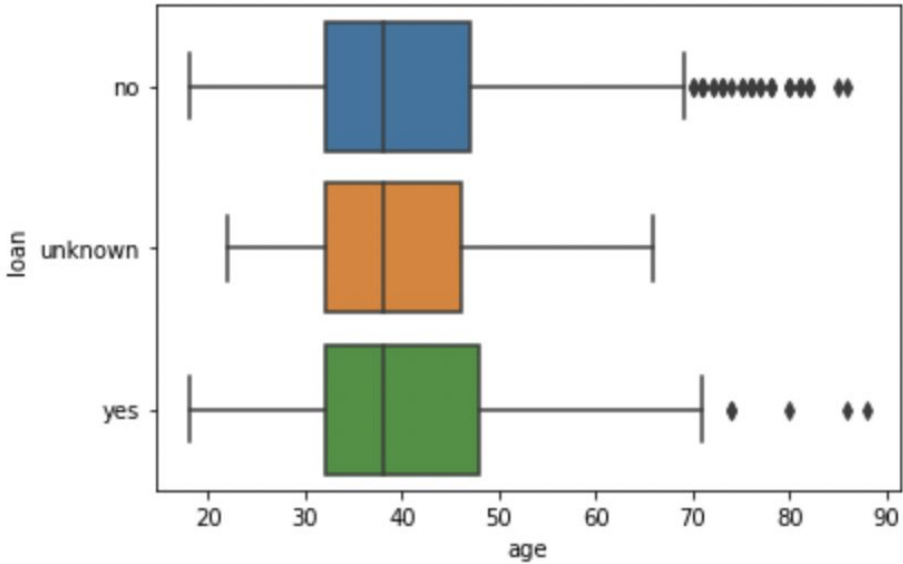
Campaign Outreach by Month:

- Best performing month: May
- Campaign performed the best during the summer months (May-Aug)
- Campaign performed the worst during winter months (Oct-Jan)
- Focus on campaign success early on as it quickly drops in effectiveness



Target Group Identification - Age

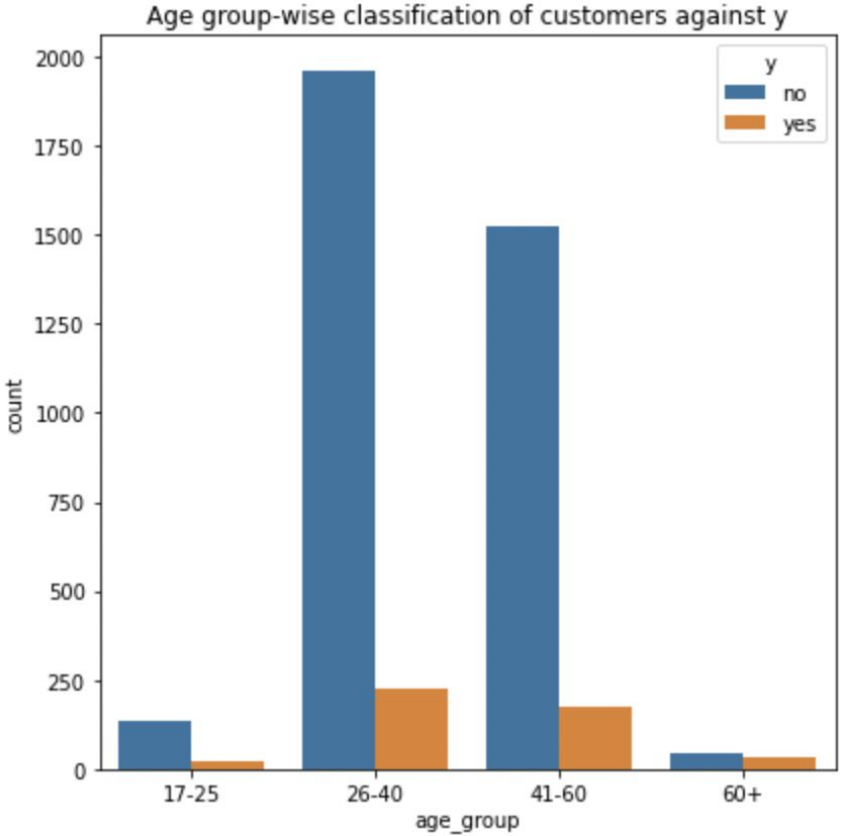




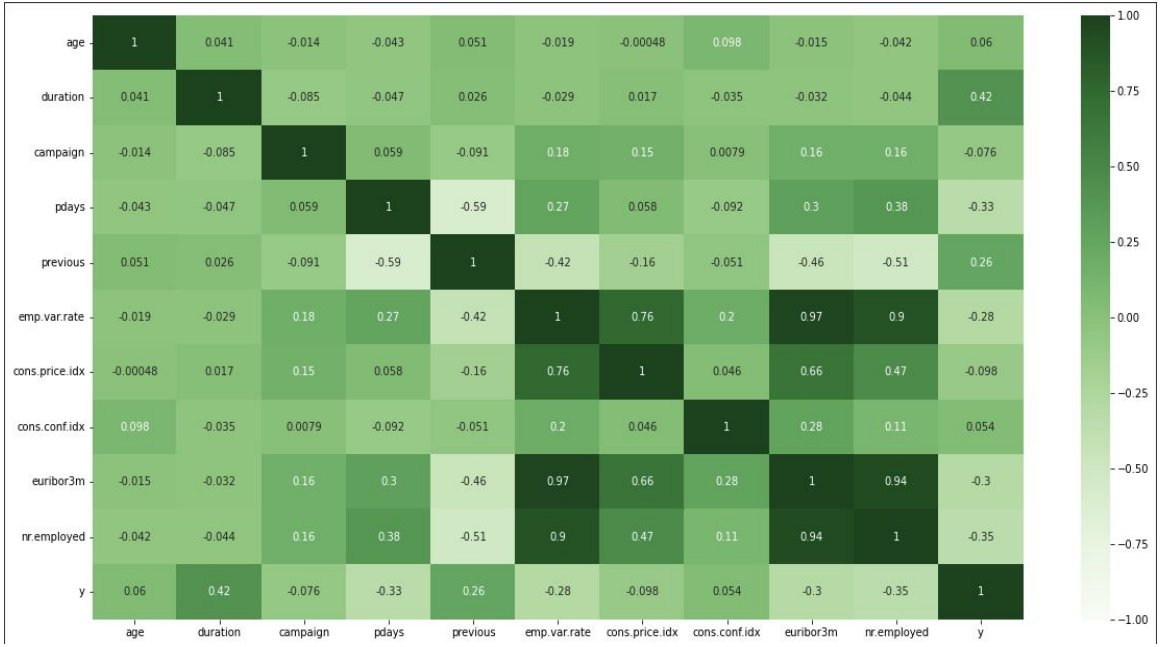
Target Group Identification - Age

Age:

- Most popular age groups:
  - 26-40 y/o
  - 41-60 y/o
- No significant trends between age group and loans
- Highest number of “yes” from the two most popular age groups
  - This may be caused by larger sample size
- We cannot recommend age as a target group on its own



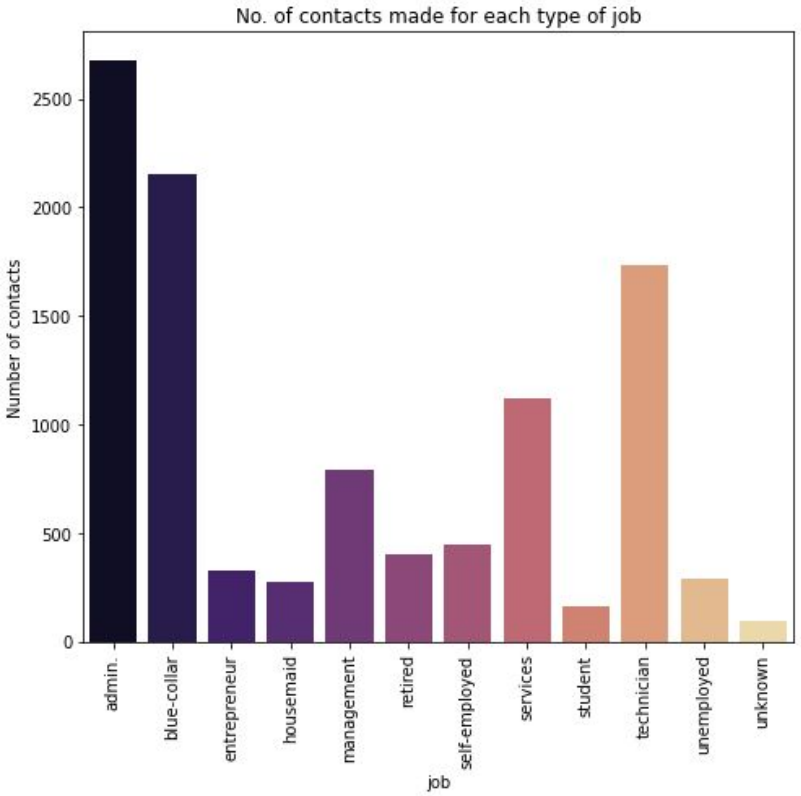
Target Group Identification - Economic Perspectives

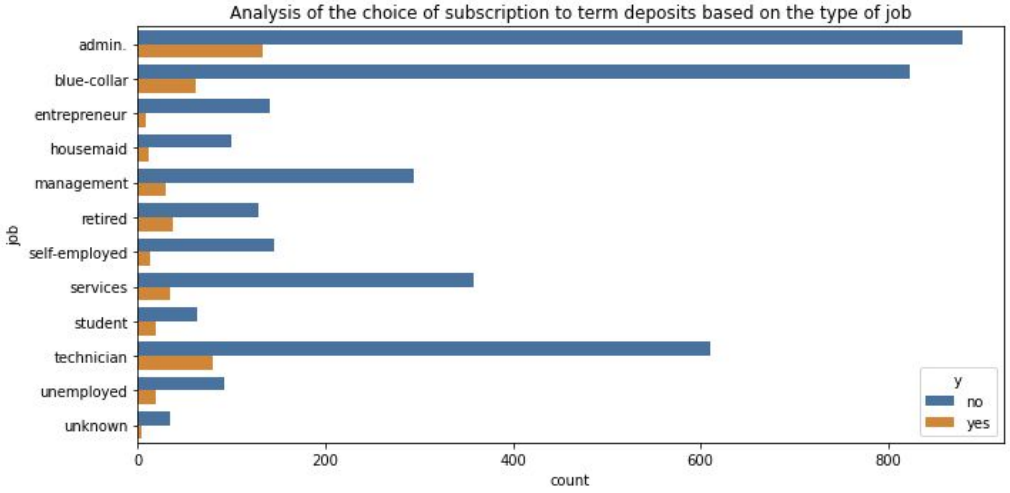


Correlation between attributes:

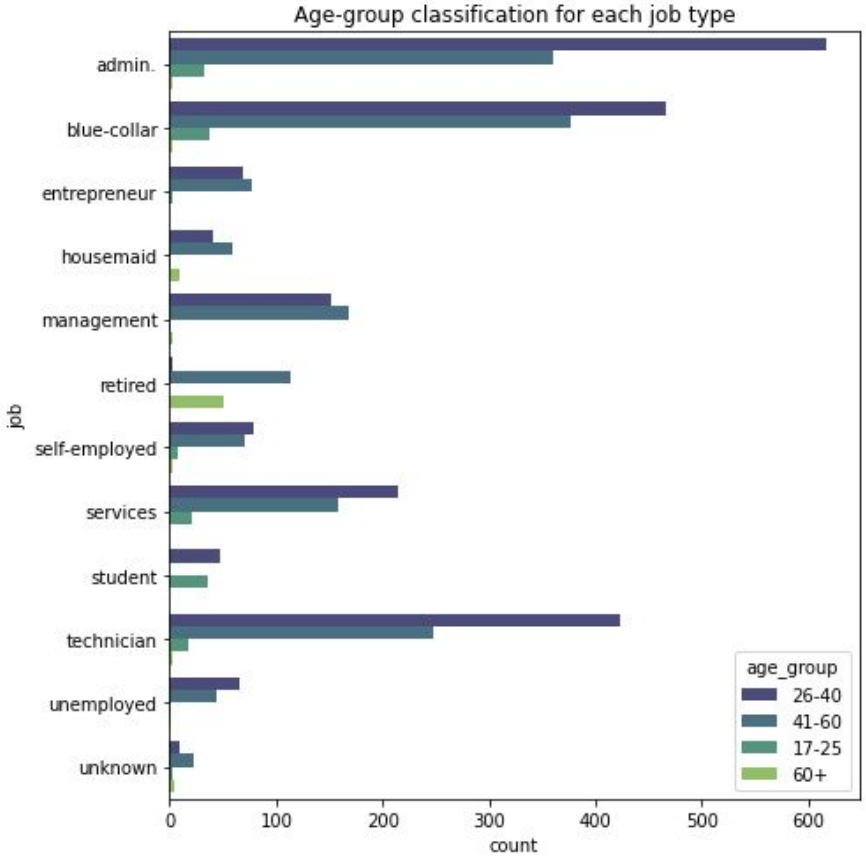
- Employment rate, consumer confidence index, and consumer price index all had high correlations
- These factors may give more insight about target client groups
- We may find that clients who have higher confidence and price index are more likely to purchase the product

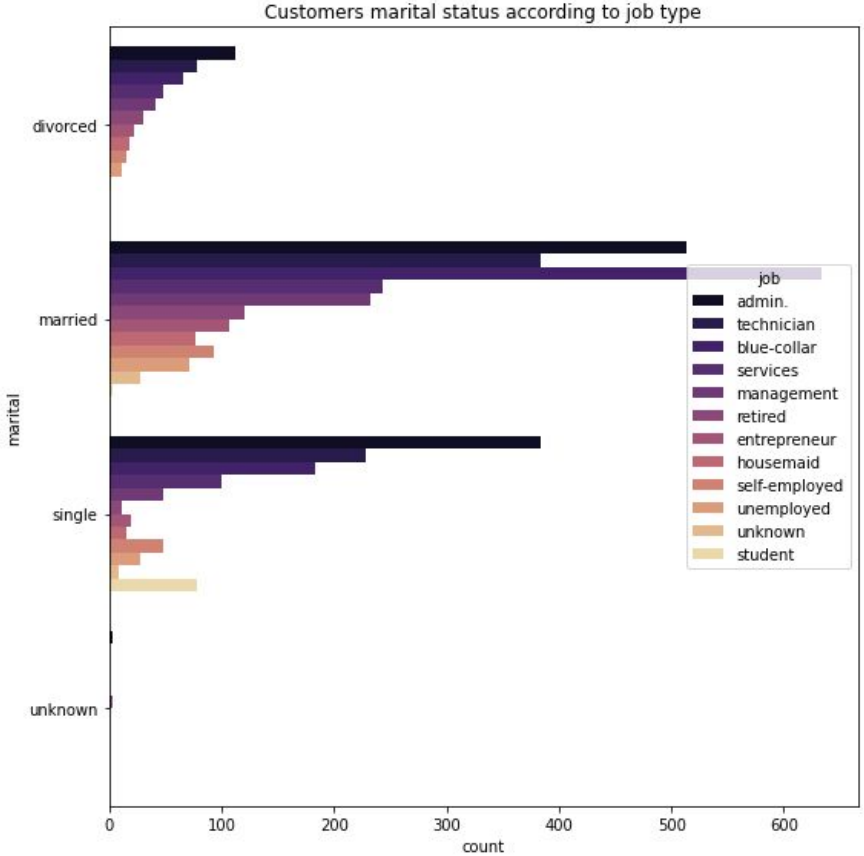
Target Group Identification - Employment and Occupation



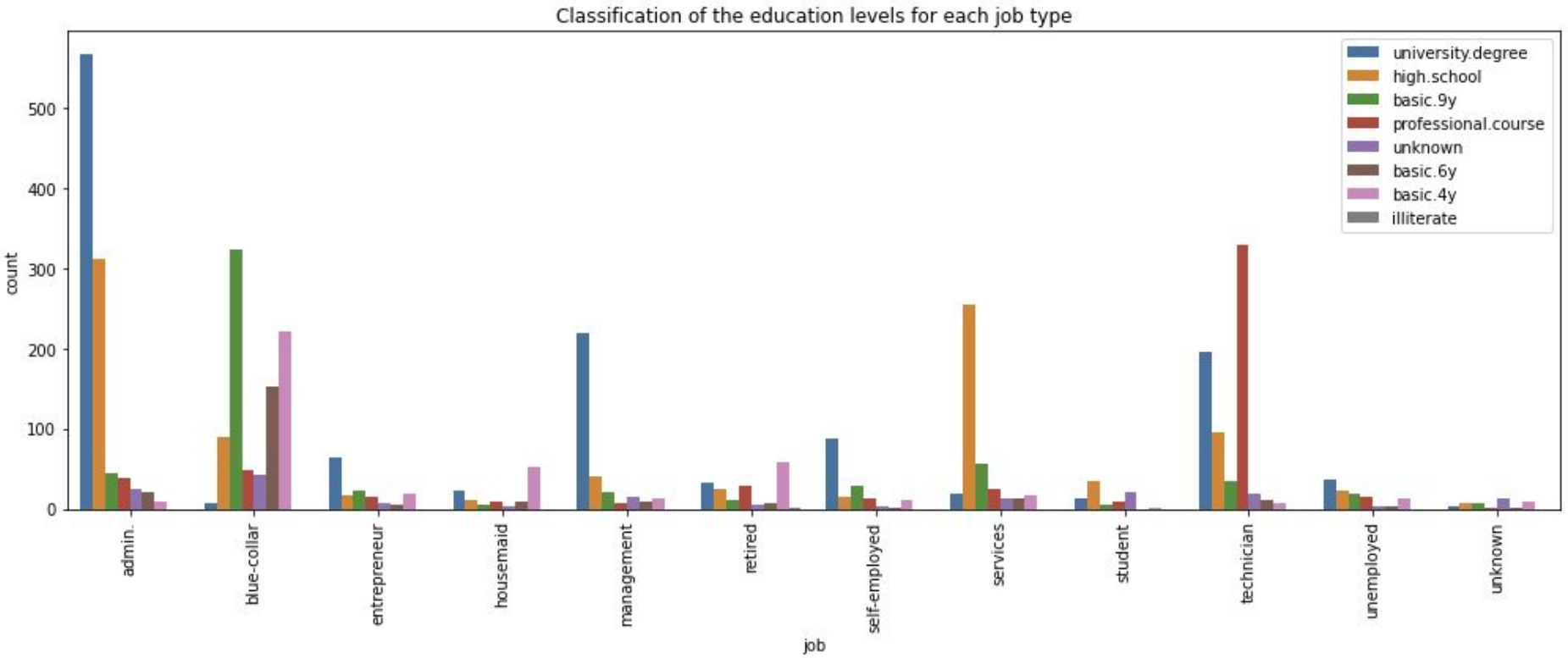


Target Group Identification - Occupation and Other Factors





Target Group Identification - Education and Occupation



Target Group Identification - Final Thoughts

Final Thoughts and Recommendations:

- After exploring many factors and groups, the bank should choose highly efficiency target groups and dates for their ad campaign
  - Suggested date: January - April
  - Suggested groups:
    - Occupation: admin, blue-collar, student, technician
    - Age: 26-40, 17-25
    - Marital Status: married, single (top occupations only)
    - Education: University degree or professional course
- Many useful target groups, but occupation has the largest impact on predicting the purchase rate

Target Group Identification - Model Selection and Execution

Thoughts and Recommendations for ML Model Selection:

- Model should predict whether a client will purchase the new product based on a variety of different data inputs
- We will test 6 different algorithms and choose the best
  - Linear algorithms: logistic regression, linear discriminant analysis
  - Nonlinear algorithms: classification and regression trees, support vector machines, Gaussian Naive Bayes, K-nearest neighbors
- Initial results are shown, but a further analysis of model building will be covered in the final report

```
ScaledLR: 0.860654 (0.034861)
ScaledLDA: 0.857459 (0.038983)
ScaledKNN: 0.715261 (0.037793)
ScaledCART: 0.649699 (0.045427)
ScaledNB: 0.826131 (0.038275)
ScaledSVM: 0.823826 (0.040493)
```