# Bank Marketing Campaigns

LaChandra Ash

November 10, 2022

## Topic

Attracting and acquiring new clients is the goal of bank marketing, which employs both classic and modern approaches to advertising. These media tactics are used to identify the type of clientele most drawn to particular businesses. This also encompasses the fact that various financial institutions employ varying tactics to draw in the kind of clientele they like (Depino, 2019).

### Business Problem

The Portuguese bank wants to know what steps to take after experiencing a reduction in revenue. The main explanation, according to analysis, is that their consumers are not making deposits as frequently as they once did. Because term deposits allow banks to keep deposits for a certain period of time, banks can invest in financial products with higher profit margins to increase their profits. Additionally, banks have a better chance of convincing term deposit customers to purchase other goods like funds or insurance in order to boost their profits. As a result, the Portuguese bank needs to pinpoint existing customers who are more likely to sign up for a term deposit and concentrate marketing efforts on them.

### Background | History

The banks are challenged with creating the campaigns to bring awareness of their products and services, so they can increase their profits. Before the start of each campaign, the banks must gather important information from each client. The client features they need are age, job type, marital status, education, default on any loan or other credit product, current annual balance amount, do they have a housing loan and or personal loan, does the client have a cellular device for the bank to contact them? The information that was gathered about each campaign includes the last contact month of the year, the last day of contact, what was the duration of the phone call (in seconds). The duration of the call can cause a high effect of the output results target.

### Data Explanation

The data was obtained from a Portuguese bank's direct-marketing efforts (UCI, n.d.). The phone was the central tool of the marketing activities. It usually took multiple interactions with a single customer to determine whether or not they were interested in the product (a bank term deposit) being offered. The dataset is multivariate.

The associated tasks are classification. The dataset has forty-five thousand two hundred and eleven number of instances and seventeen attributes. The y variable is the output and target variable. The explanation of the twenty-one features is below:

- age (numeric)
- job : type of job (categorical: 'admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown')
- marital : marital status (categorical: 'divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)
- education (categorical: 'basic.4y','basic.6y','basic.9y','high.school','illiterate','professional.course','university.degree','unknown')

- default: has credit in default? (categorical: 'no','yes','unknown')
- housing: has housing loan? (categorical: 'no','yes','unknown')
- loan: has personal loan? (categorical: 'no','yes','unknown')
  # related with the last contact of the current campaign:
- contact: contact communication type (categorical: 'cellular','telephone')
- month: last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')

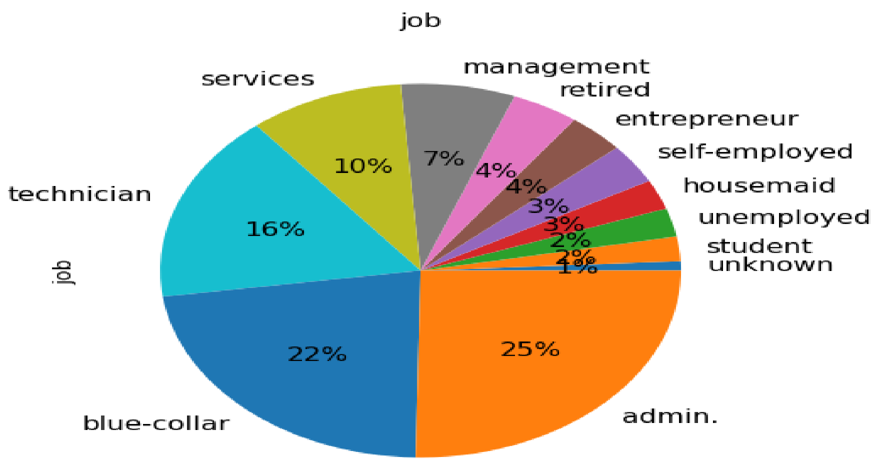
- day_of_week: last contact day of the week (categorical: 'mon','tue','wed','thu','fri'
- duration: last contact duration, in seconds (numeric# other attributes:
- campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
- pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
- previous: number of contacts performed before this campaign and for this client (numeric)
- poutcome: outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')
  # social and economic context attributes
- emp.var.rate: employment variation rate - quarterly indicator (numeric)
- cons.price.idx: consumer price index - monthly indicator (numeric)
- cons.conf.idx: consumer confidence index - monthly indicator (numeric)
- euribor3m: euribor 3-month rate - daily indicator (numeric)
- nr.employed: number of employees - quarterly indicator (numeric)
- y - has the client subscribed a term deposit? (binary: 'yes','no') (UCI, n.d.).

**Methods**

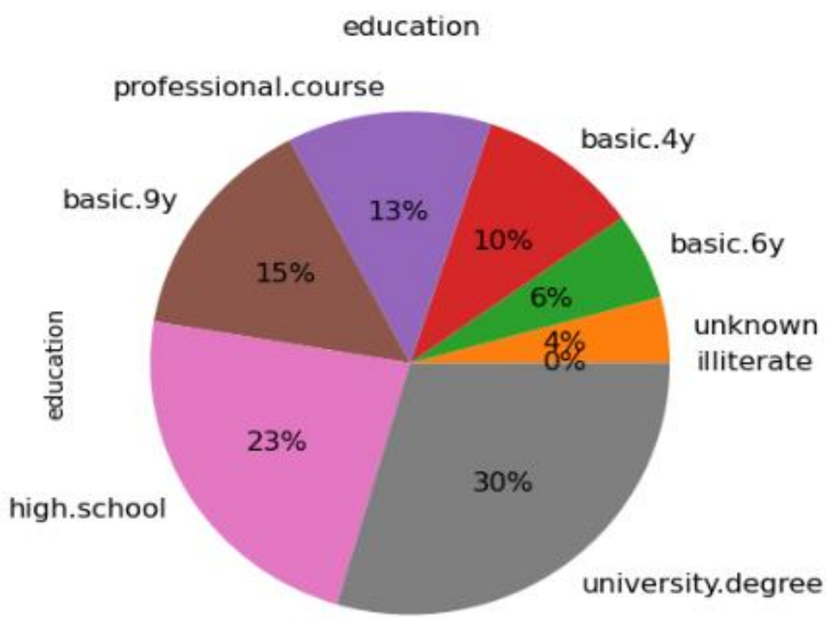
I evaluated the dataframe for missing values, and duplicates. The dataframe was free of missing values. I removed the twelve duplications from the dataframe. I explored the dataframe by checking its shape, description, information, correlation, covariance, counted the features, displayed the columns, viewed the dtypes, revealed the features' unique values, displayed the categories, and displayed the categorical and numerical categories. I selected the Decision Tree Classifier and Naïve Bayes to model my data.

**Analysis**

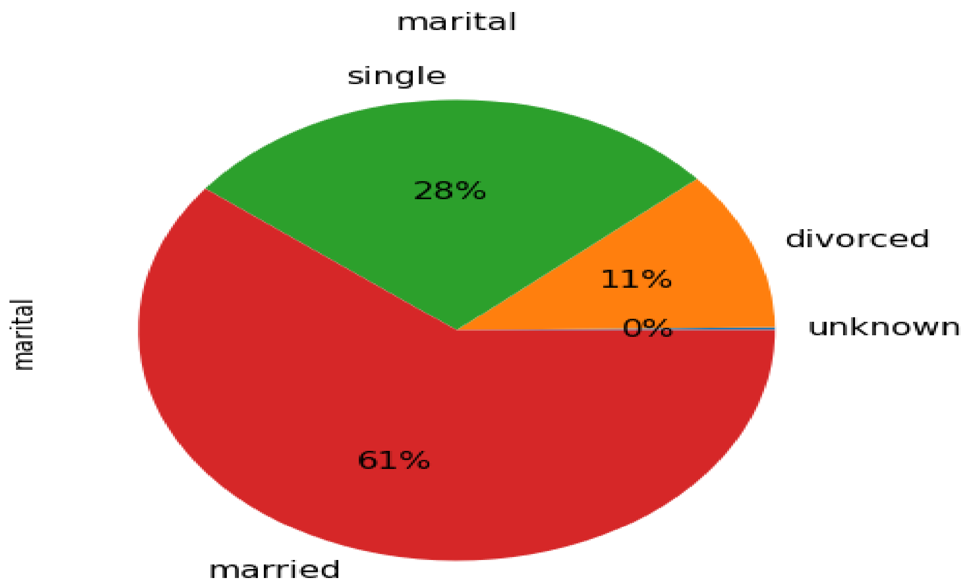
1)Job: Most of the bank's consumers are employed as blue-collar or administration.




2)Education: Most of the bank's consumers who are subscribed to the bank has a high school degree and university degree.

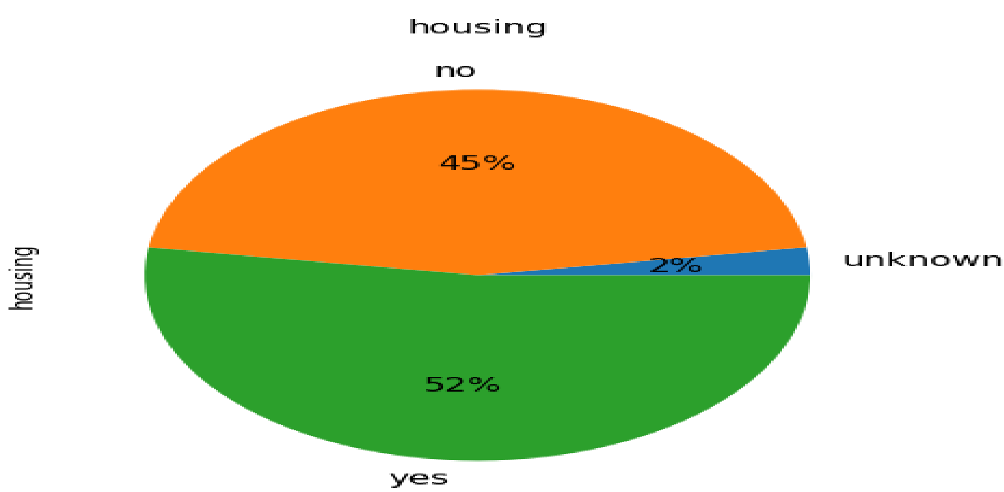


3)Marital: Most of the bank's consumers who are subscribed to the bank are married.

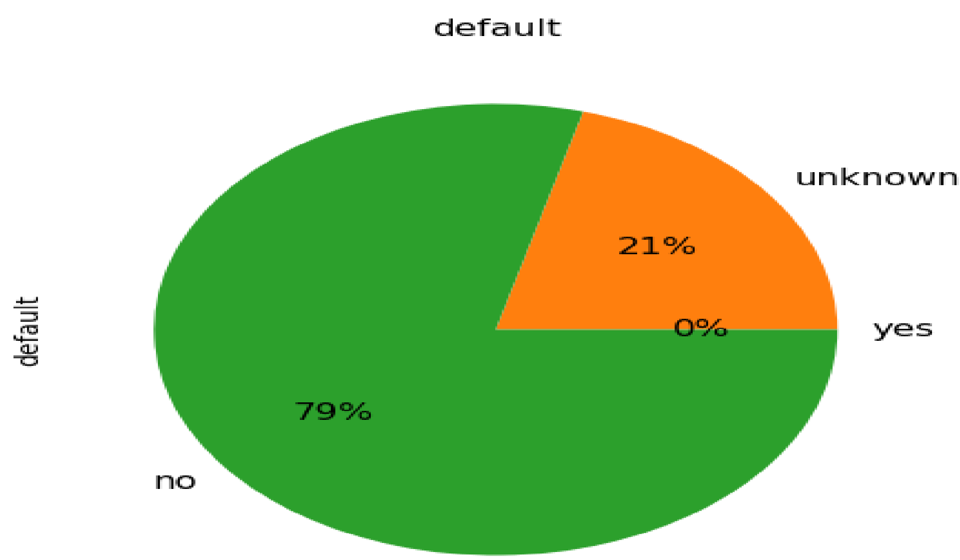



---

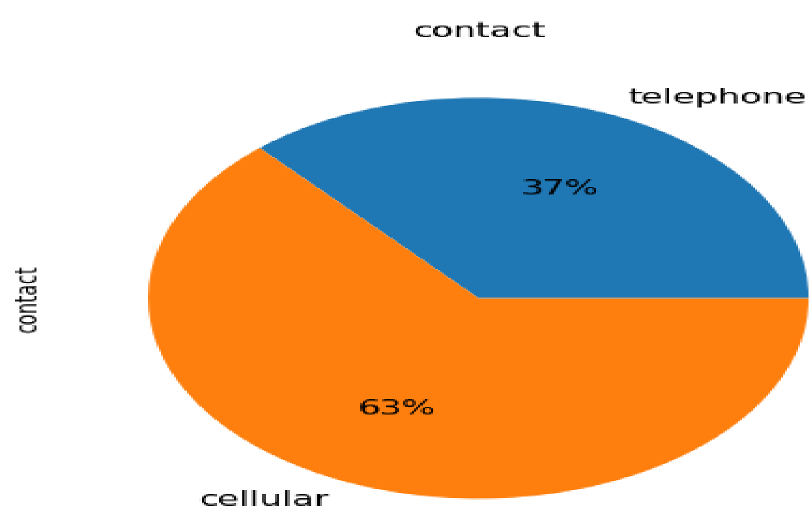
4)Housing: More than fifty percent of the bank's consumers have housing loans at the bank. Forty-five percent of the bank consumers do not have a housing loan.



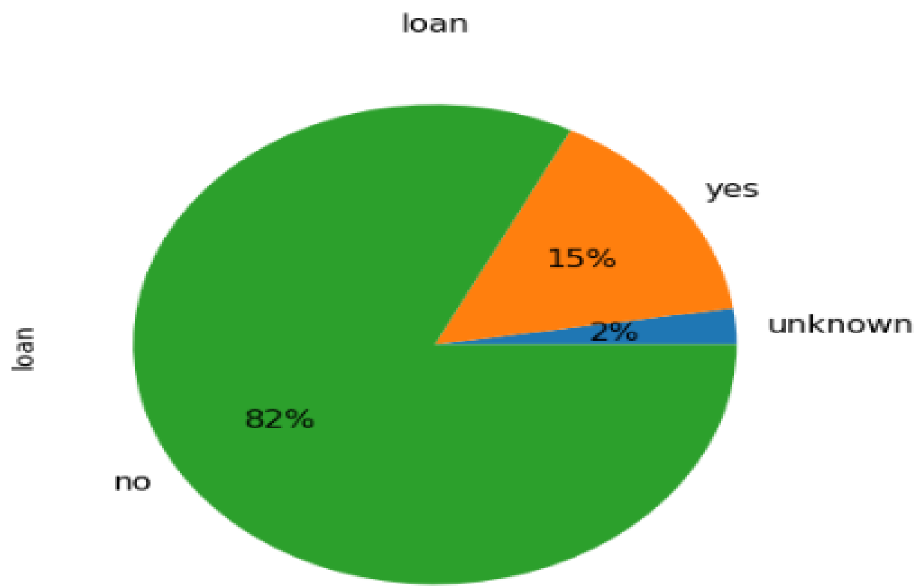
5)Default: Most of the bank's consumers did not default on their loan. The bank has twenty-one percent of consumers who may have or did not default on their loan.



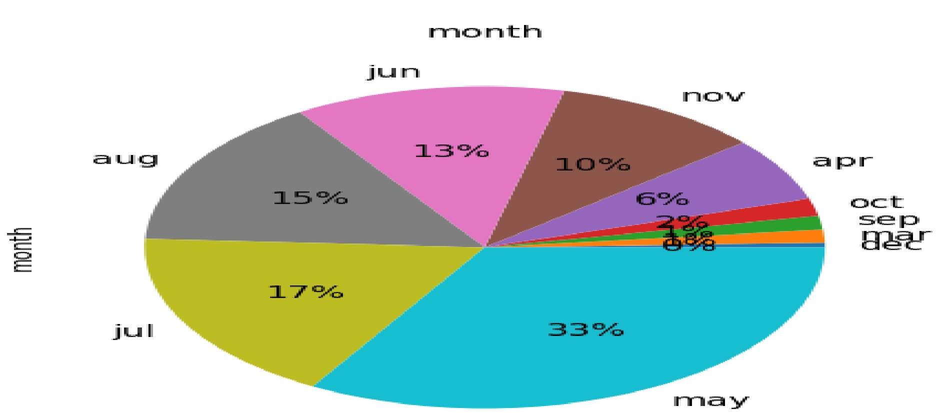

6)Contact: Most of the bank's consumers were contacted by their cellphone number.



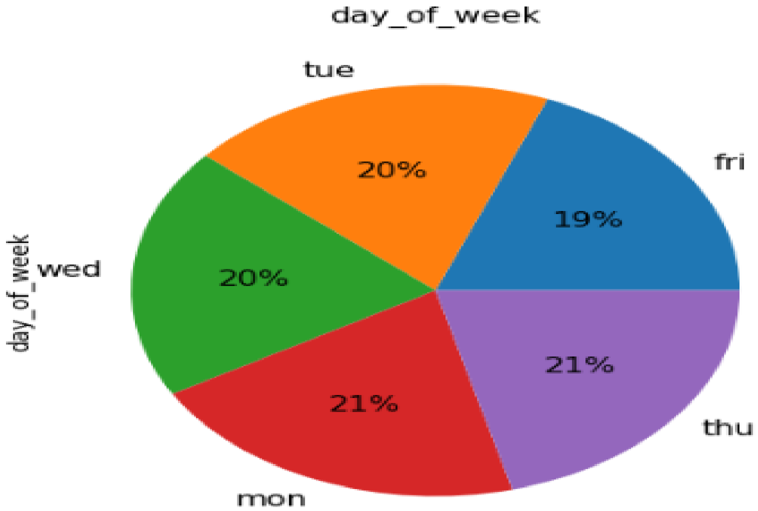
7) Loan: Most of the bank's consumers did not have a personal loan before the start of the campaign.




8)Month: The bank's campaign gained the highest number of consumers during the month of May. The percentage of consumers they gained during the campaign began to decrease each month.

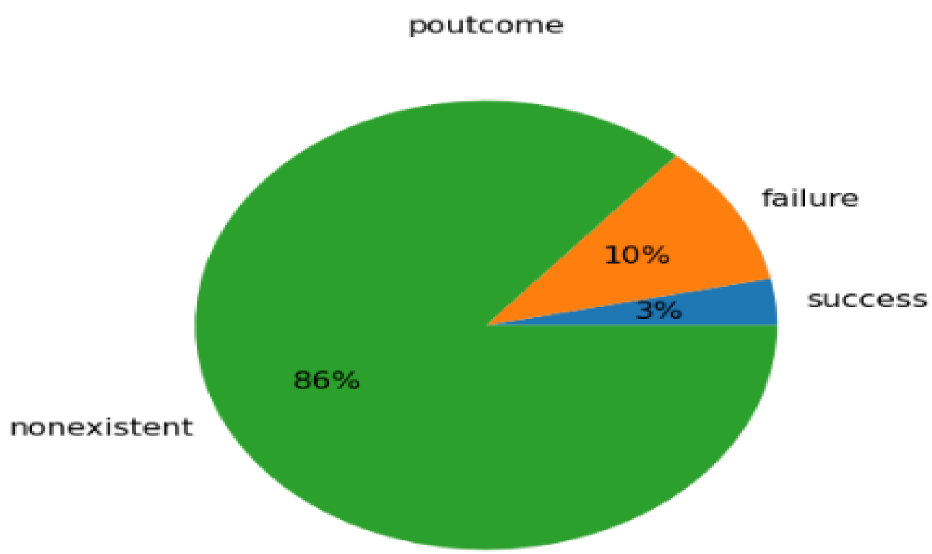


9)Day of Week: The percentage of the bank's consumers who were contacted on the last day of the week was nearly the same each day..

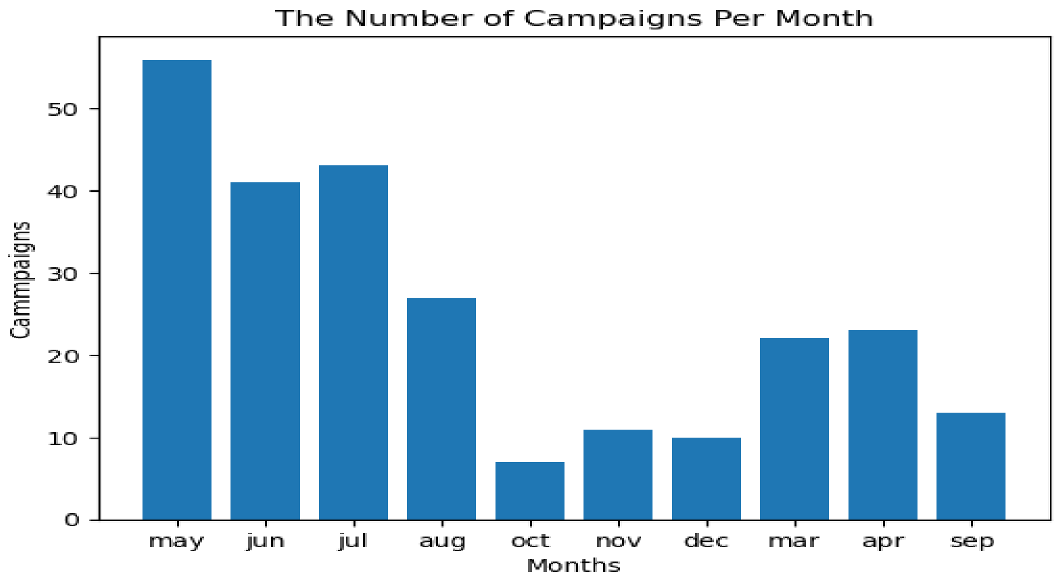



---

10)Poutcome: The highest percentage of outcome results of the previous campaign contacts are non-existent.

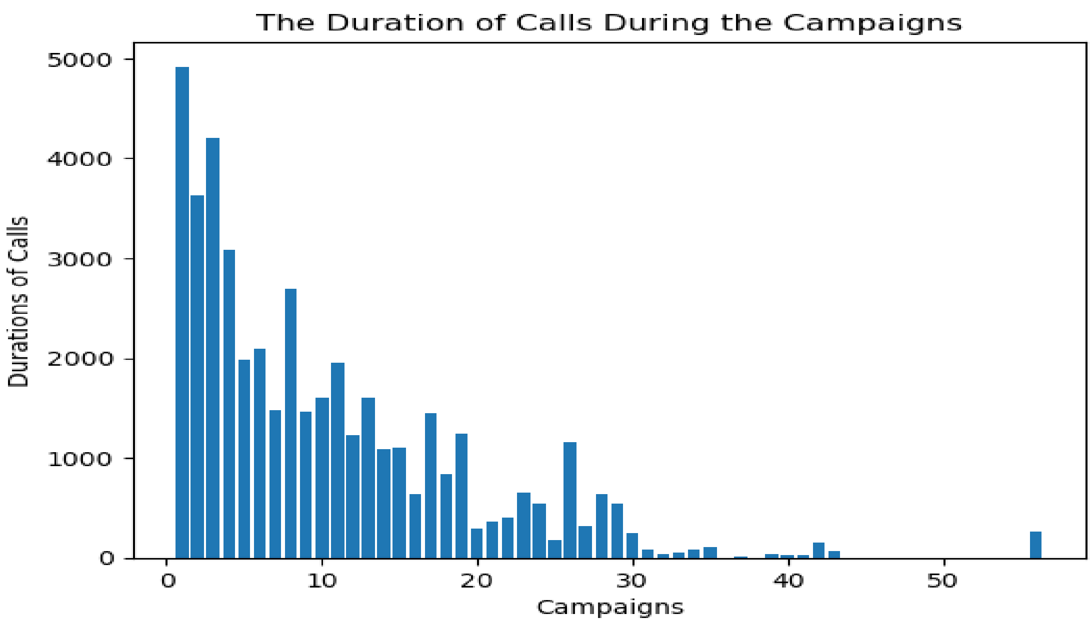


11)The Number of Campaigns Per Month: The bank contacted the greatest number of consumers during the month of May. More consumers were contacted during the summer versus fall and spring months.

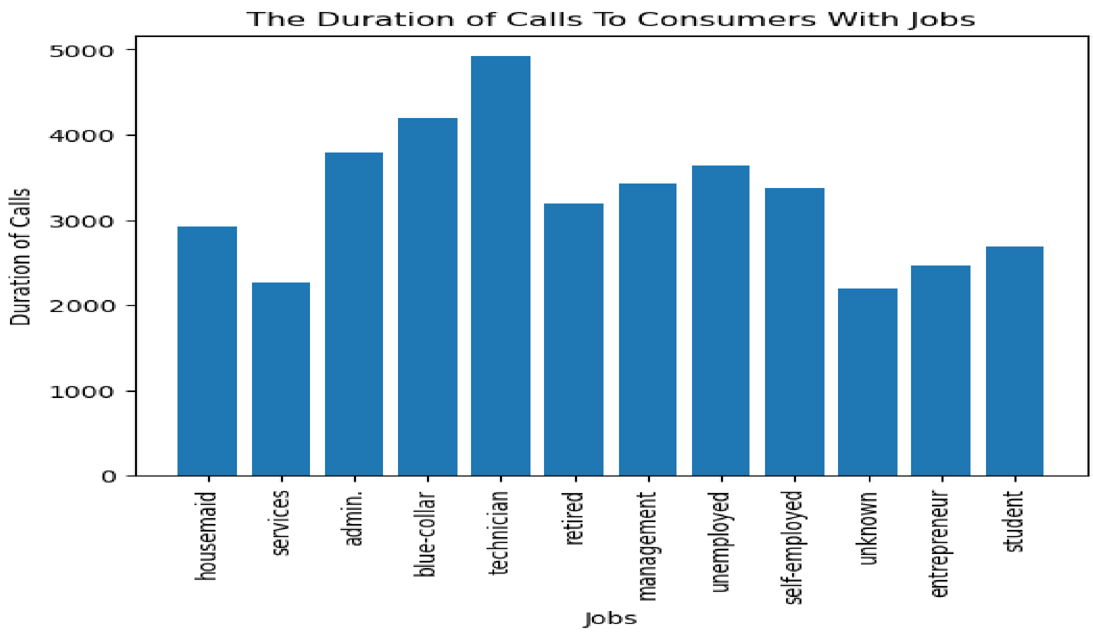



---

12)The Duration of Calls During the Campaigns: The most duration of calls was conducted in the beginning of the campaign contacts, and they decreased when the campaign contacts increased.

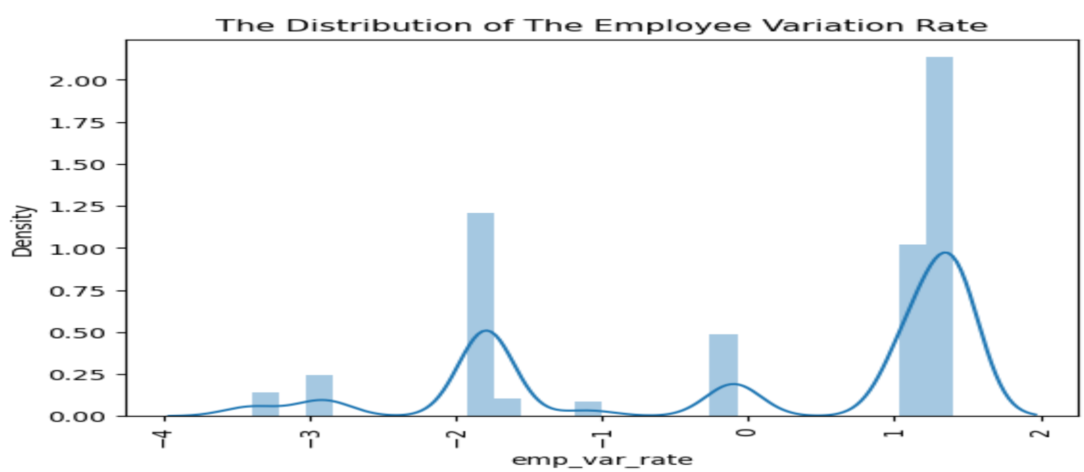


13) The Duration of Calls to Consumers With Jobs: The duration of calls lasted longer with consumers who worked as a technician.

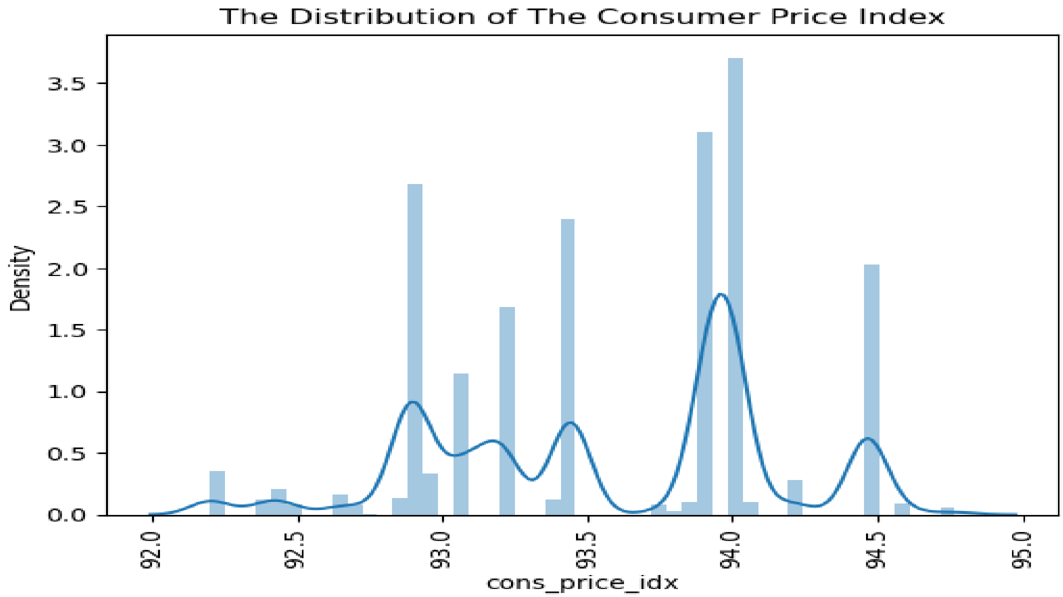



---

14)Employee Variation Rate Distribution: A high turnover rate indicates that the campaign was launched at a time of significant staff mobility caused by economic factors.

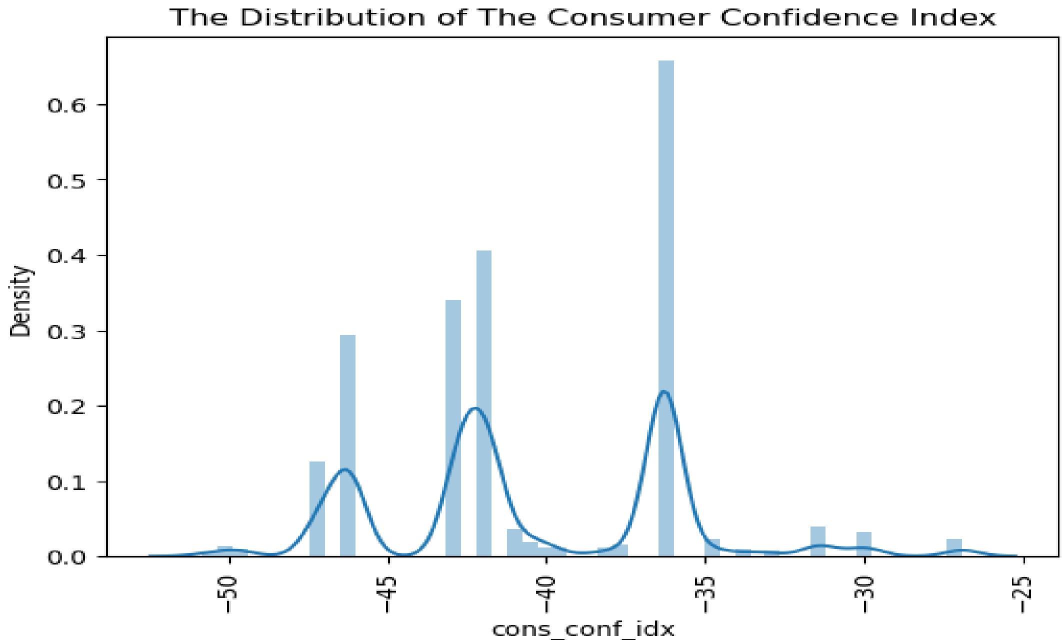


15) The Consumer Price Index: The consumer price index is helpful in indicating to potential savers and depositors which areas of the country have the lowest prices for consumer items.

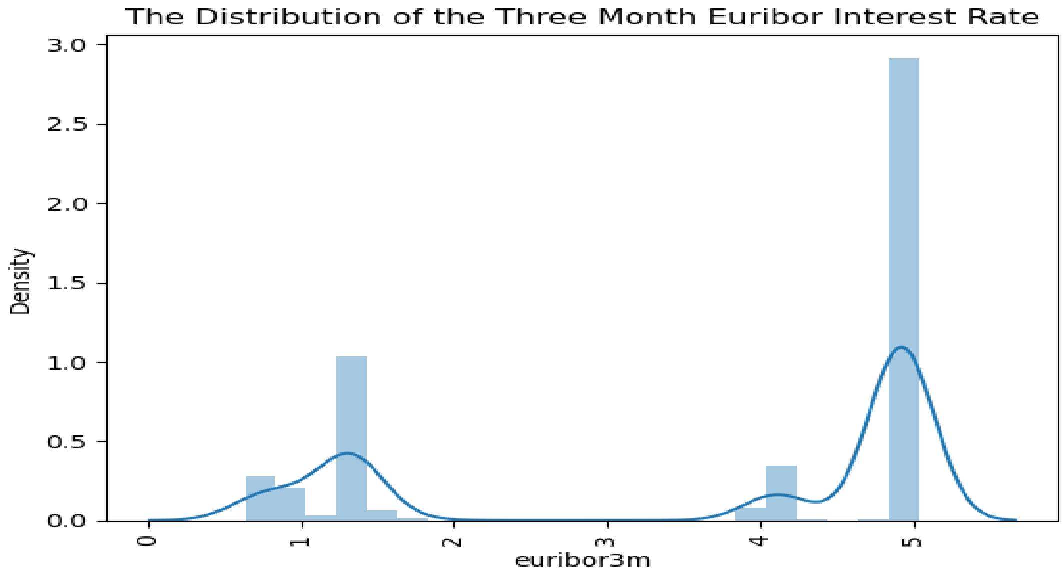



---

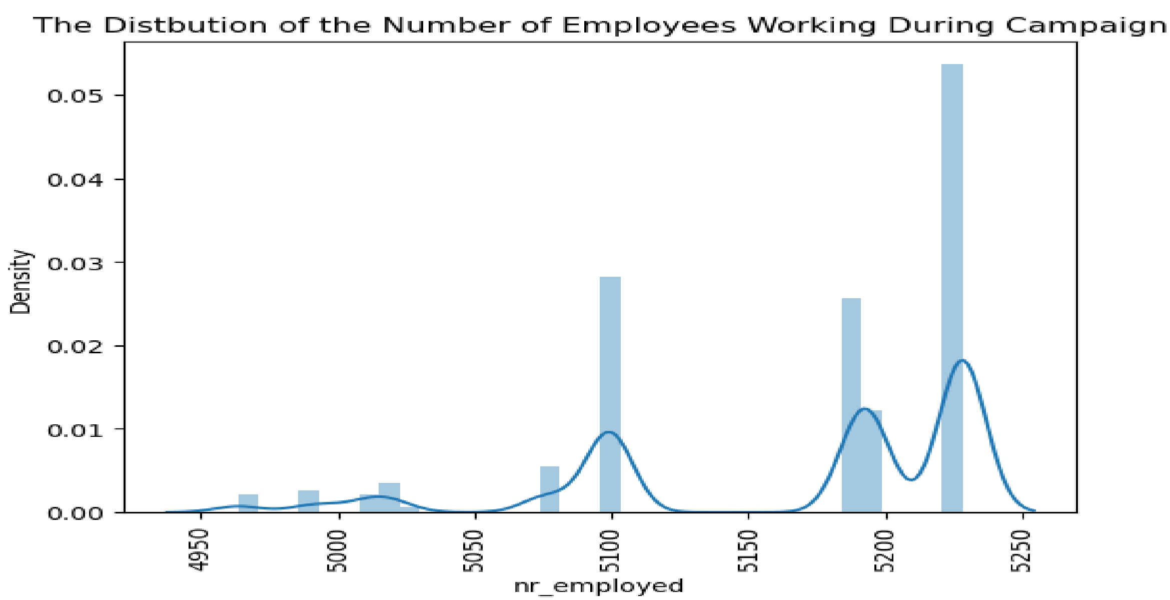
16)The Distribution of The Consumer Confidence Index: The consumer confidence index was somewhat low because people are worried about the uncertain state of the economy.



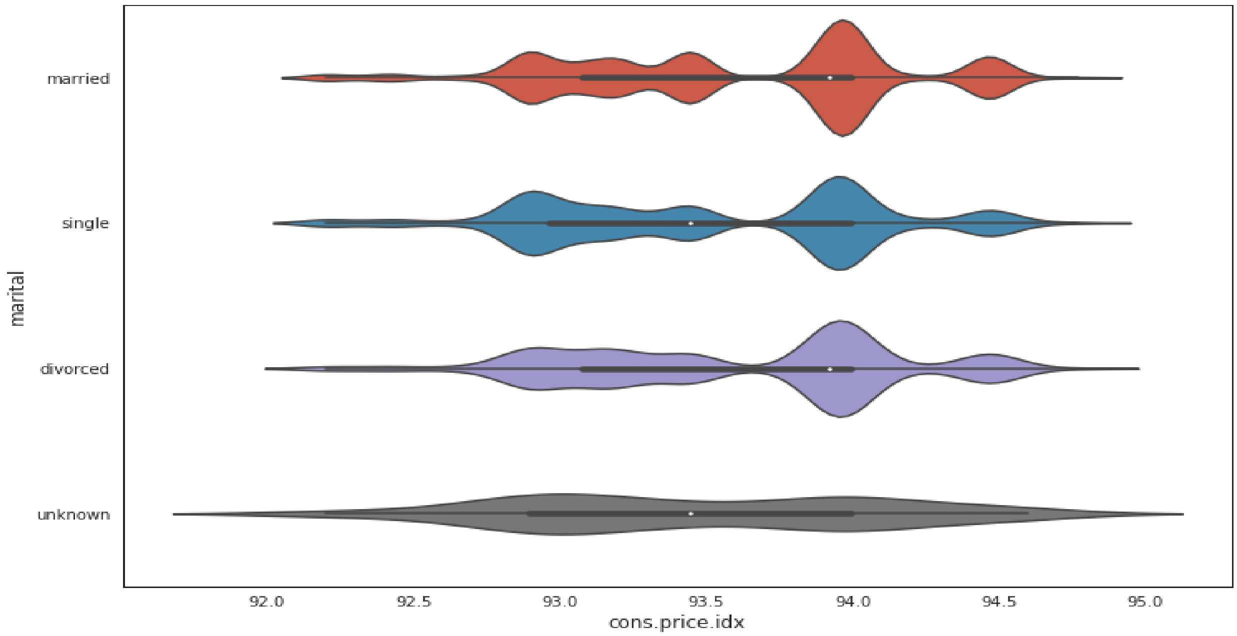
17)The Distribution of the Three Month Euribor Interest Rate: The rates at which they borrow money are quite high.



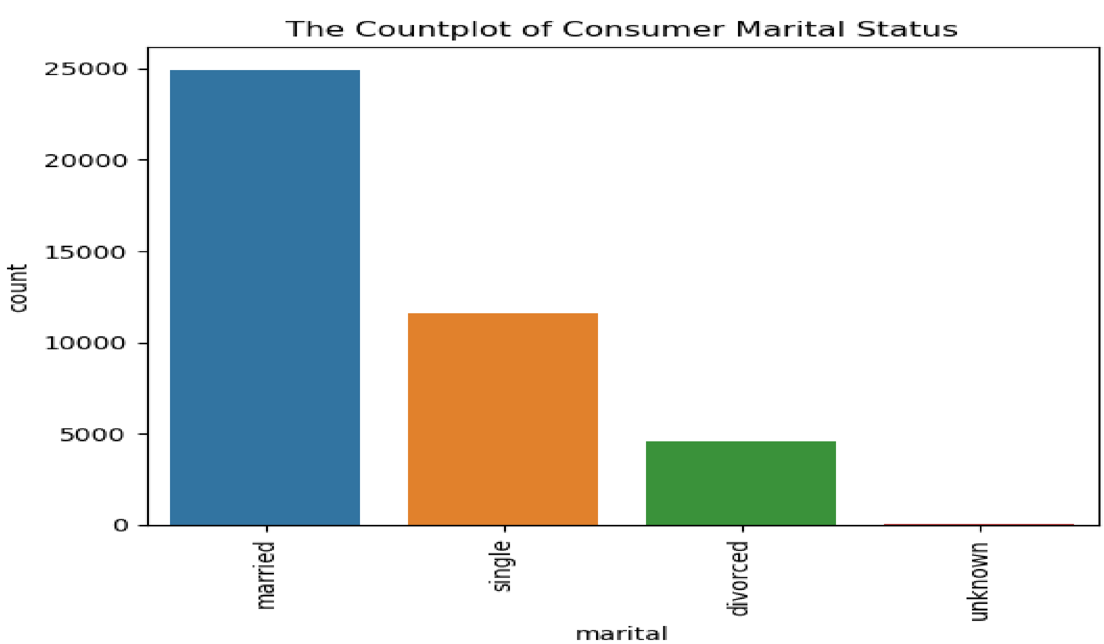

18)The Distribution of the Number of Employees Working During the Campaign: Employee counts were quite high, which is a positive indicator of future financial success.



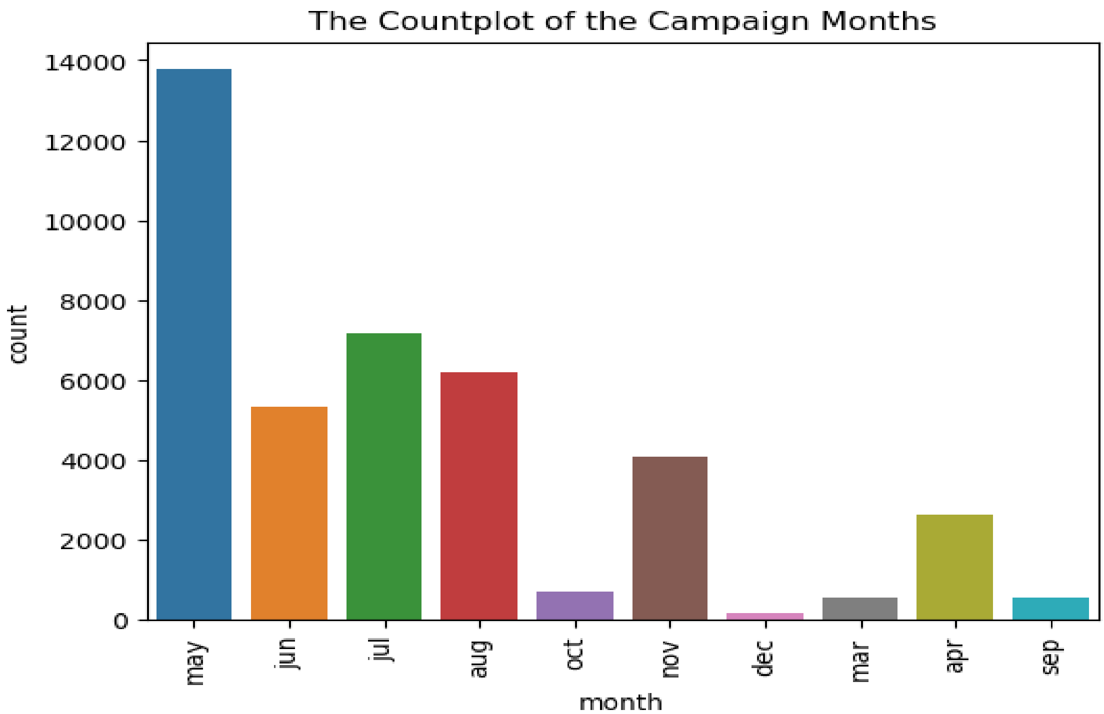
19)Marital Vs. Consumer Price Index: The married consumers are most likely to be considered before other marital statuses because they have higher consumer price index.




20)Most of the consumers were married.

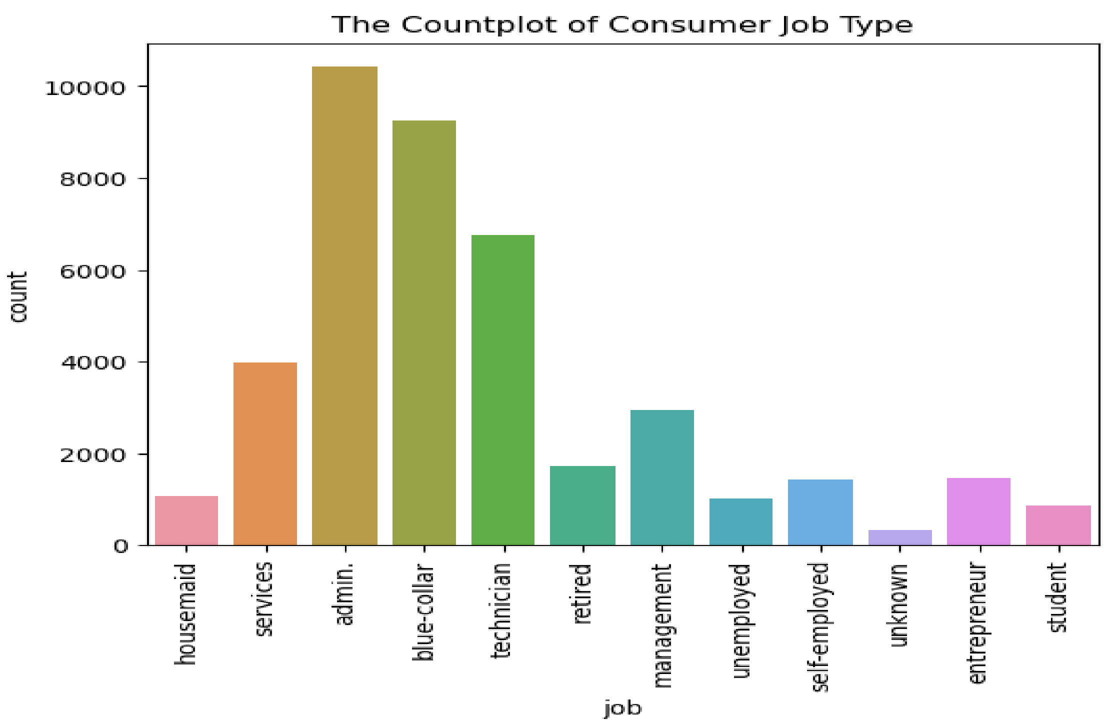



(21) Month: The most campaign contacts were made during the month of May.

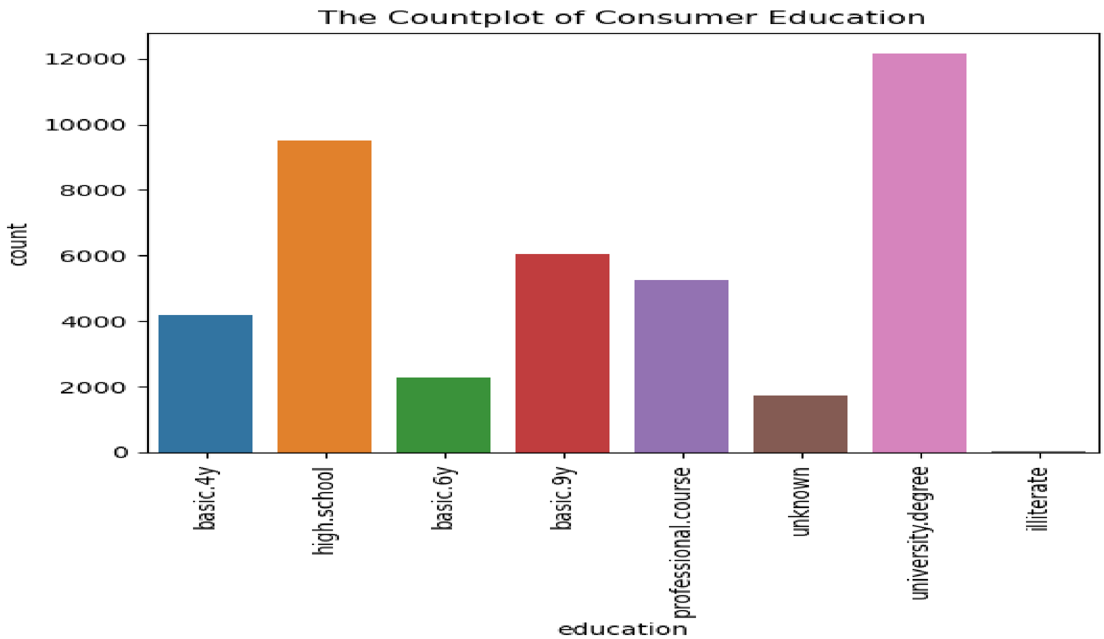



(22)Job: The greatest number of consumers worked in the administration, blue-collar, and technician jobs.

---



(23).Most of the bank's consumers had a high school diploma and university degree.




**The DecisionTreeClassifier and Naïve Bayes Models**

The purpose of the prediction is to predict if the bank's consumers will subscribe to another loan term or not. The DecisionTreeClassifier resulted in a model accuracy score of 89%. The Naïve Bayes model accuracy score was 85.6%. The precision score was 95%. The f1-score revealed a score of 92%..

**View Appendix: Table 1: Model Predictive Metrics**

**Conclusion**

The precision score predicted 95.% of the bank's consumers will subscribe to the loan term. The model can be fine tuned by adding another model. Most of the bank's consumers were married, had a college and high school degree, and worked in a blue-collar, technician, or administration job field.

The bank contacted the highest number of consumers in the month of May. Most bank consumers were contacted during the summer, and the campaign contacts decreased as the year went by. The campaigns did not occur in the months of January and February. Married consumers will more than likely qualify for the loan term because of their consumer price index.

**References**


Data.World (2022). Bank Marketing Dataset. Retrieved from Bank Marketing - dataset by uci | data.world, on November 6, 2022.

DePino, F. (2019). Modern Bank Marketing – A Comprehensive Guide. Retrieved from Modern Bank Marketing - A Comprehensive Guide (2021) (mediaboom.com), on November 6, 2022.

Roshan, B. (n.d.). Bank Marketing Campaign | Predictive Analytics. Retrieved from Bank marketing campaign | Predictive analytics | Kaggle, on November 6, 2022.

UCI (n.d.). Bank Marketing Data Set. Retrieved from UCI Machine Learning Repository: Bank Marketing Data Set, on November 6, 2022.

## Appendix:

Table 1: Model Predictive Metrics

```
#Display the score of the decisiontreeclassifier model.
clf = DecisionTreeClassifier(random_state=0)
clf.fit(xtrain,ytrain, sample_weight=None, check_input=True, X_idx_sorted=None)
clf.get_params(deep=True)
clf.predict(xtest, check_input=True)
clf.predict_log_proba(xtest)
clf.predict(xtest,check_input=True)
print(clf.score(xtest,ytest, sample_weight=None))
```

```
0.8885651857246905
```

```
#Conduct the prediction tests and print the metrics.
y_pred = gb.predict(x_test)
print(confusion_matrix(y_test,y_pred))
print(accuracy_score(y_test,y_pred)*100)
print(classification_report(y_test,y_pred))
```

```
[[6526  807]
 [ 376  529]]
85.63971837824714
              precision    recall  f1-score   support

           0       0.95      0.89      0.92      7333
           1       0.40      0.58      0.47       905

    accuracy                           0.86      8238
   macro avg       0.67      0.74      0.69      8238
weighted avg       0.89      0.86      0.87      8238
```