# Exploratory Data Analysis

## Bank Marketing(Campaign)

**11-Dec-2021**

# Bank Marketing


- **I have done final project individually**
- **Name : Mohammad Aminian**
- **Email:** aminanmuhamad24@gmail.com
- **Country : Iran**
- **College/Company : Tehran Kharazmi University**
- **Specialization : Data Science**

# Data Exploration

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).

Bank wants to use ML model to shortlist customer whose chances of buying the product is more so that their marketing channel (tele marketing, SMS/email marketing etc)  can focus only to those customers whose chances of buying the product is more.

This will save resource and their time ( which is directly involved in the cost ( resource billing)).

Develop model with Duration and without duration feature and report the performance of the model.

Duration feature is not recommended as this will be difficult to explain the result to business and also it will

be difficult for business to campaign based on duration.

# Data Exploration

Input variables:
# bank client data:
1 - age (numeric)
2 - job : type of job (categorical: 'admin.','blue-collar','entrepreneur','housemaid','management','retired','self-employed','services','student','technician','unemployed','unknown')
3 - marital : marital status (categorical: 'divorced','married','single','unknown'; note: 'divorced' means divorced or widowed)
4 - education (categorical: 'basic.4y','basic.6y','basic.9y','high.school','illiterate','professional.course','university.degree','unknown')
5 - default: has credit in default? (categorical: 'no','yes','unknown')
6 - housing: has housing loan? (categorical: 'no','yes','unknown')
7 - loan: has personal loan? (categorical: 'no','yes','unknown')
# related with the last contact of the current campaign:
8 - contact: contact communication type (categorical: 'cellular','telephone')
9 - month: last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')
10 - day_of_week: last contact day of the week (categorical: 'mon','tue','wed','thu','fri')c)

# Data Exploration

11 - duration: last contact duration, in seconds (numeric). Important note: this attribute highly affects the output target (e.g., if duration=0 then y='no')

# other attributes:

12 - campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)

13 - 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)

14 - previous: number of contacts performed before this campaign and for this client (numeric)

15 - poutcome: outcome of the previous marketing campaign (categorical: 'failure','nonexistent','success')

# social and economic context attributes

16 - emp.var.rate: employment variation rate - quarterly indicator (numeric)

17 - cons.price.idx: consumer price index - monthly indicator (numeric)

18 - cons.conf.idx: consumer confidence index - monthly indicator (numeric)

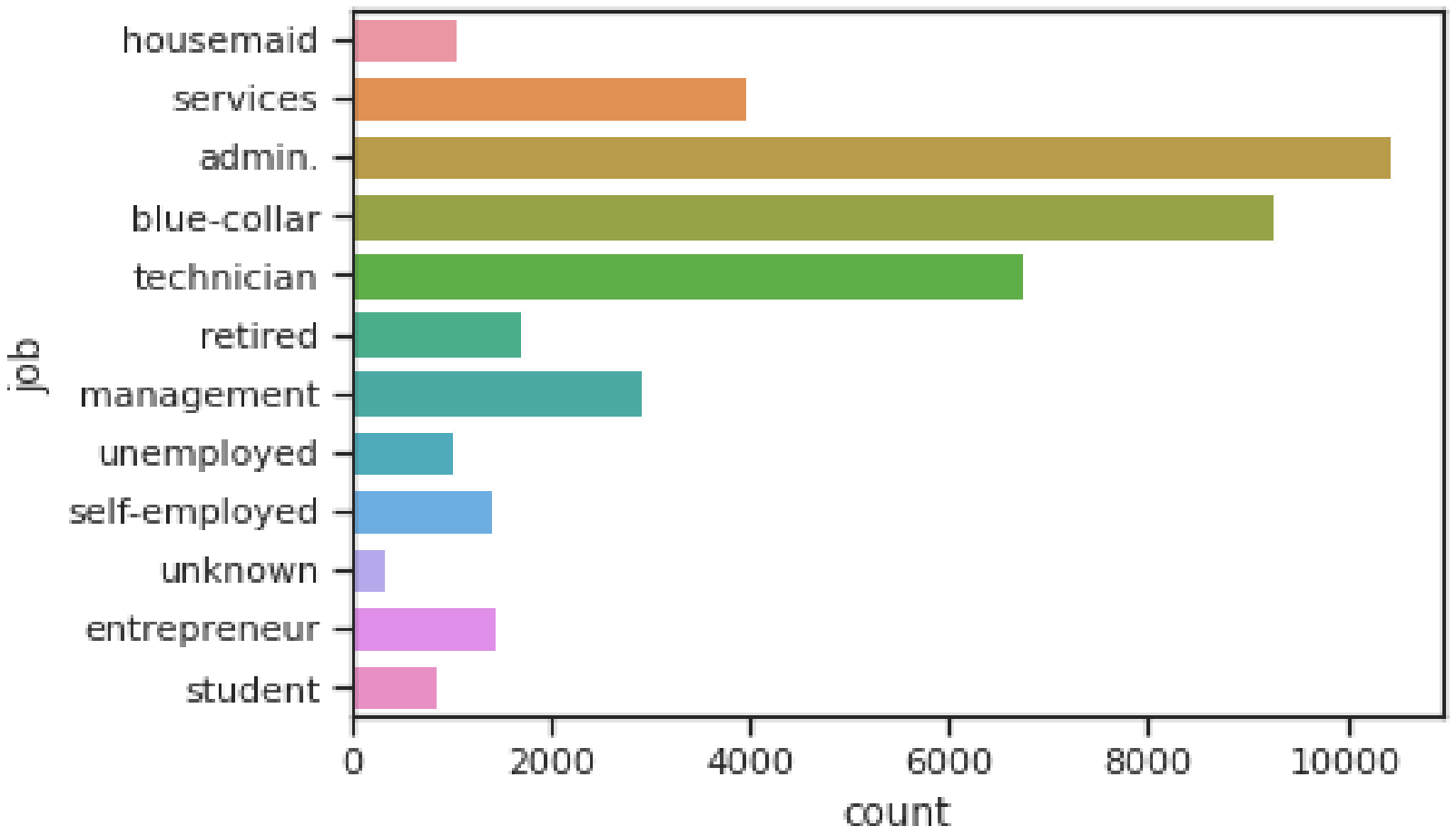
19 - euribor3m: euribor 3 month rate - daily indicator (numeric)

20 - nr.employed: number of employees - quarterly indicator (numeric)

Output variable (desired target):

21 - y - has the client subscribed a term deposit? (binary: 'yes','no')

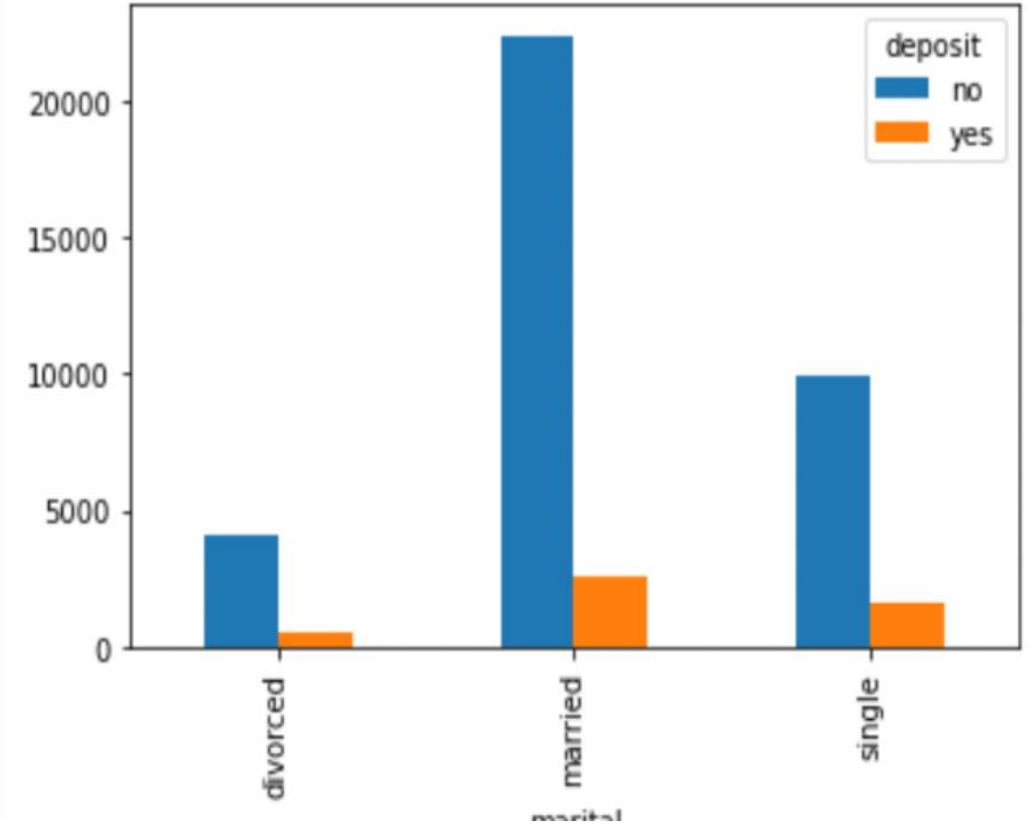
# Categorical EDA

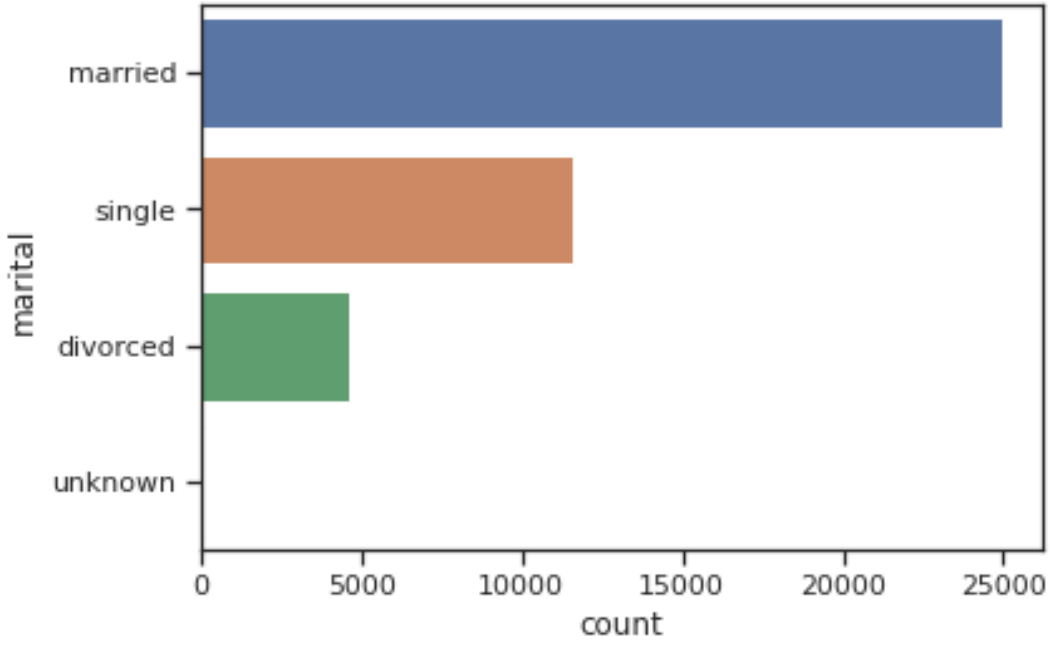



As we see in graph
most of the customers
that have jobs as "admin",
"blue-collar" or "technician"

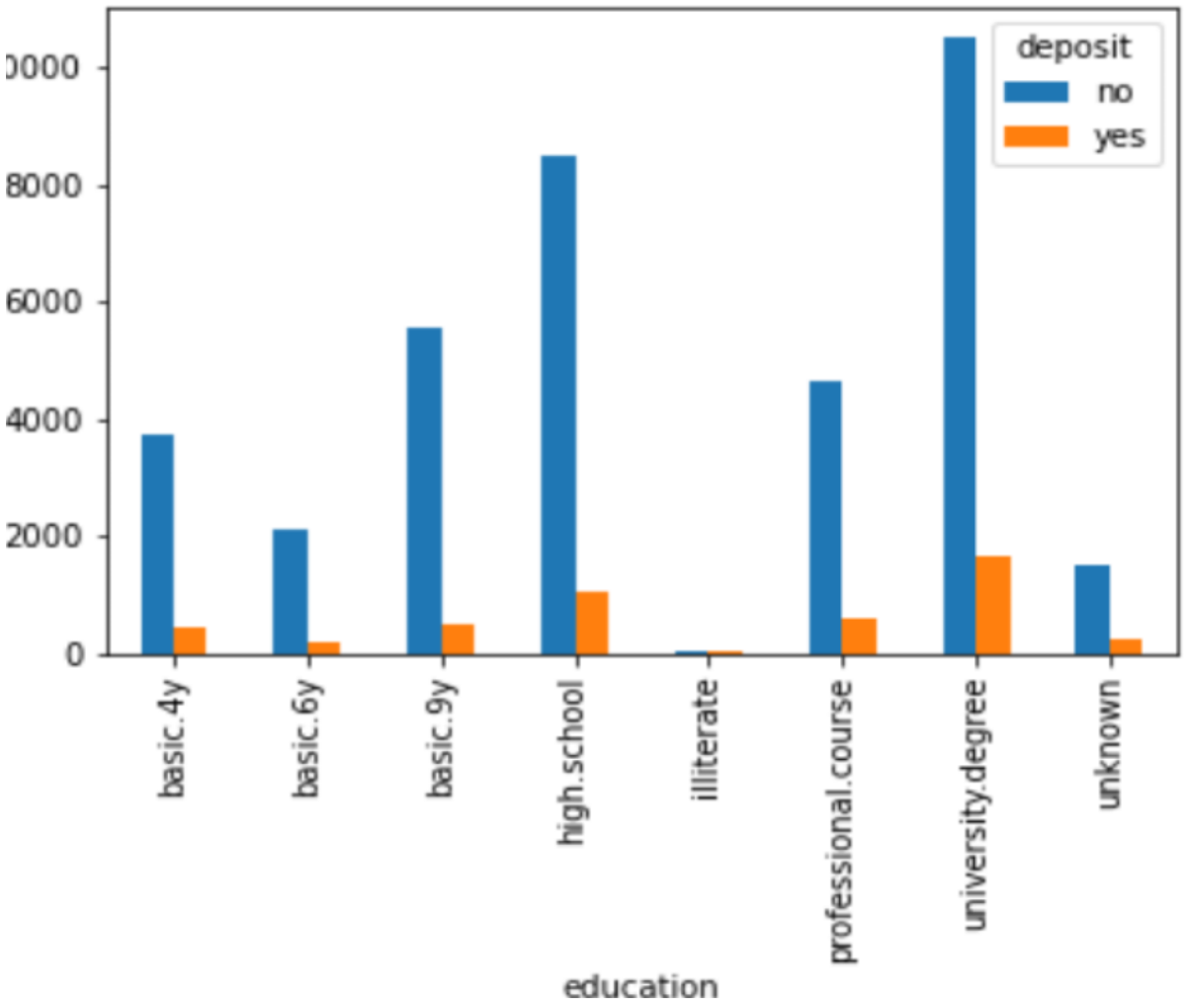
# Categorical EDA

Effect of Marital on deposit Possiblity.

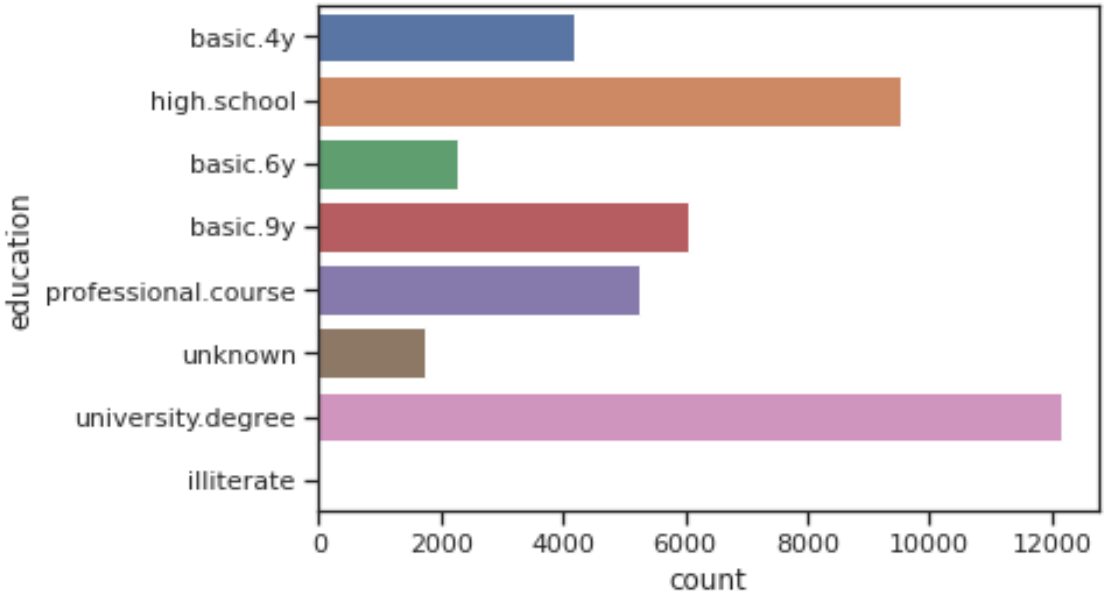


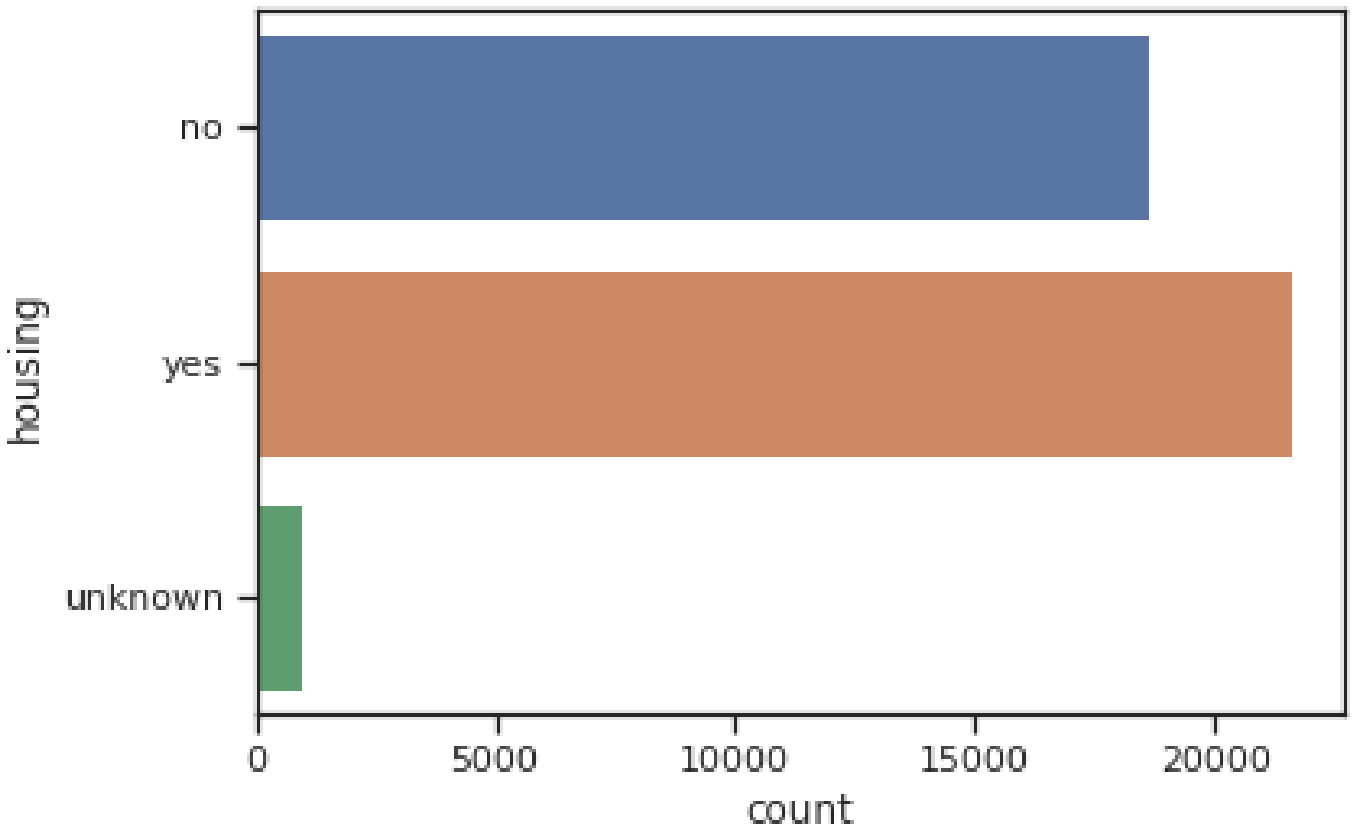
# Categorical EDA




Effect of Education on deposit Possiblity.

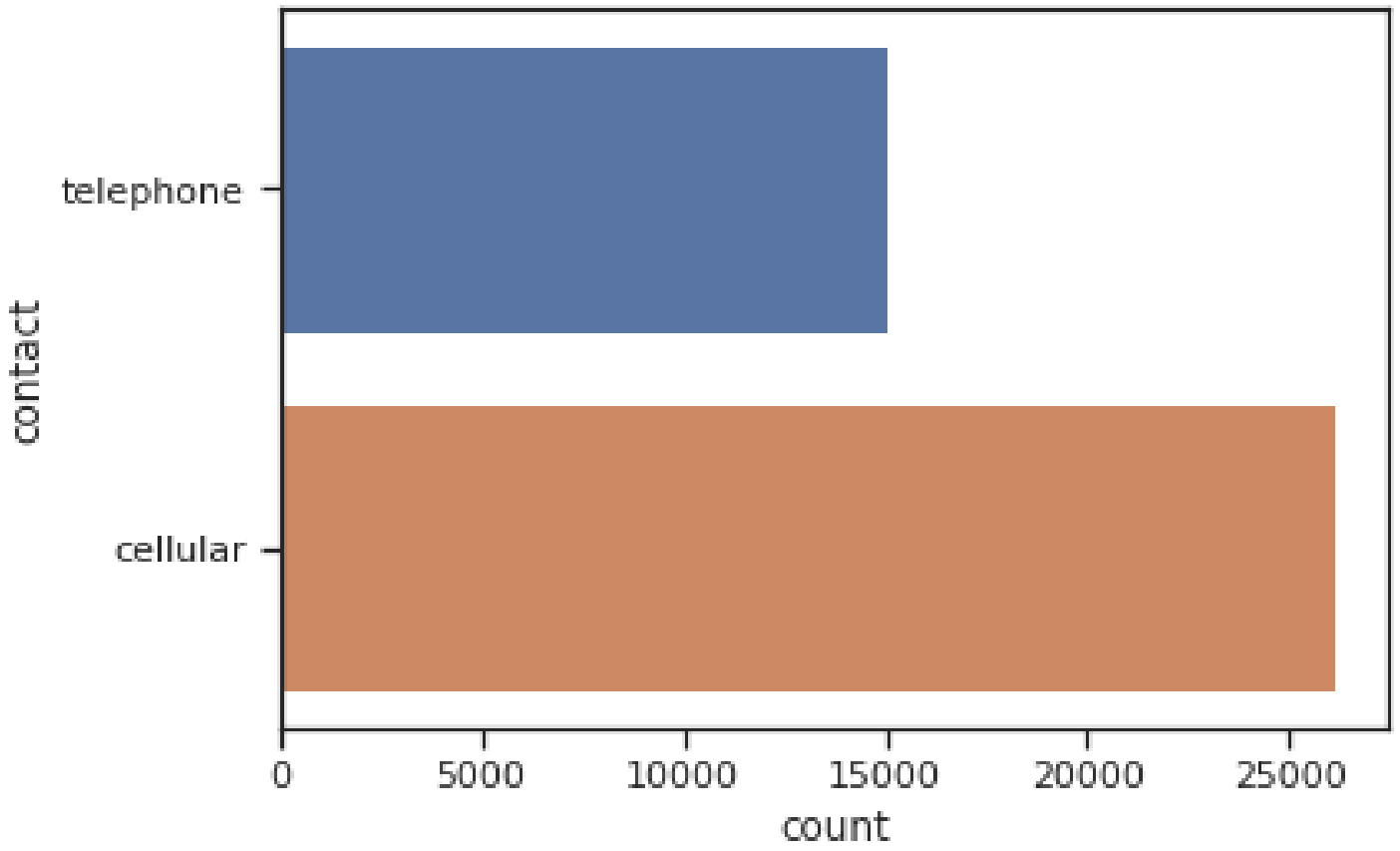


# Categorical EDA



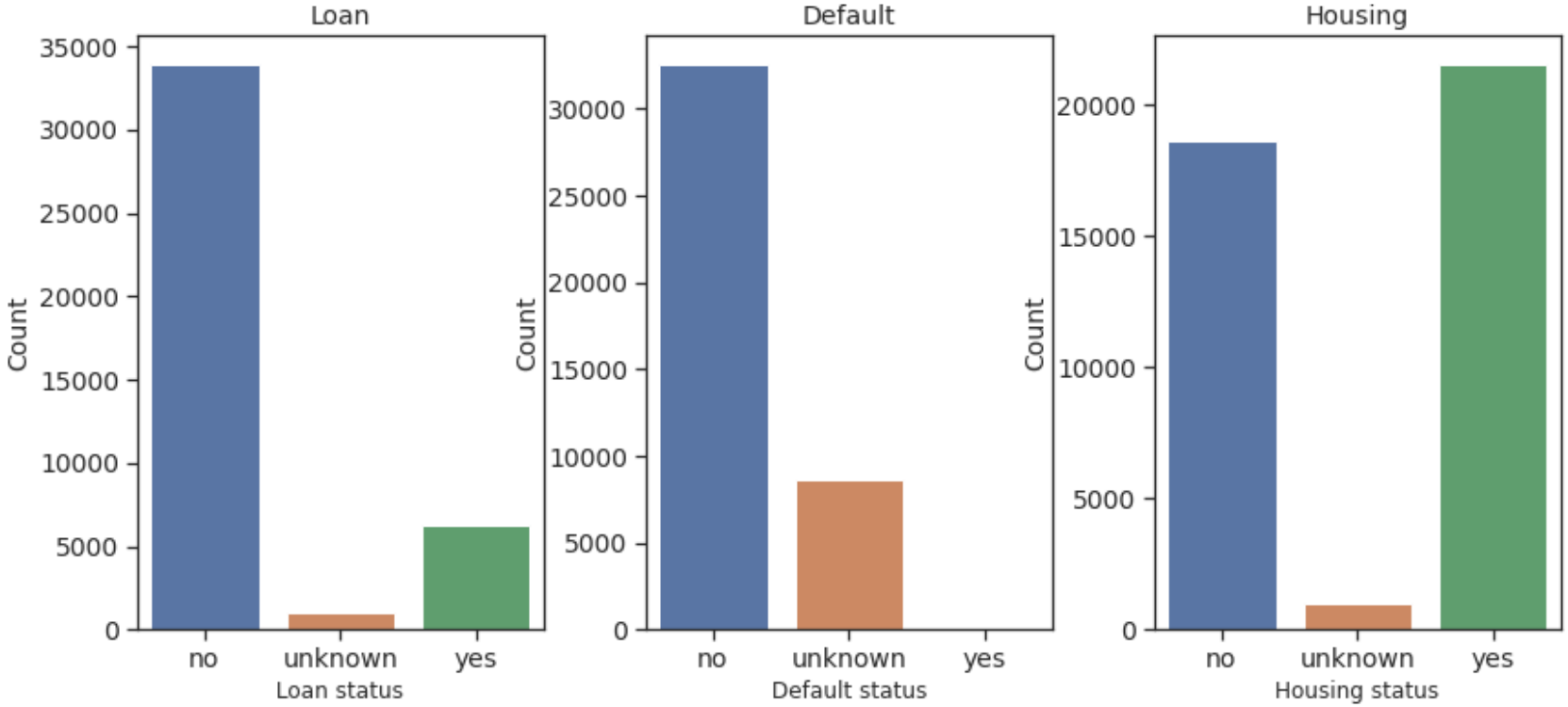

People who have housing loan contacted more.

# Categorical EDA




Most pepole contacted in cellular as it has been shown
In image.

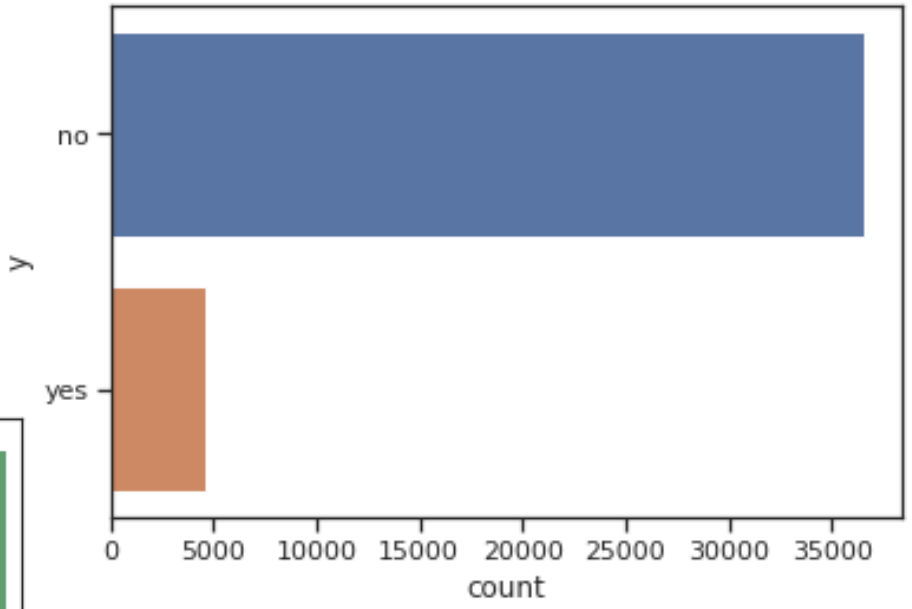
# Categorical EDA



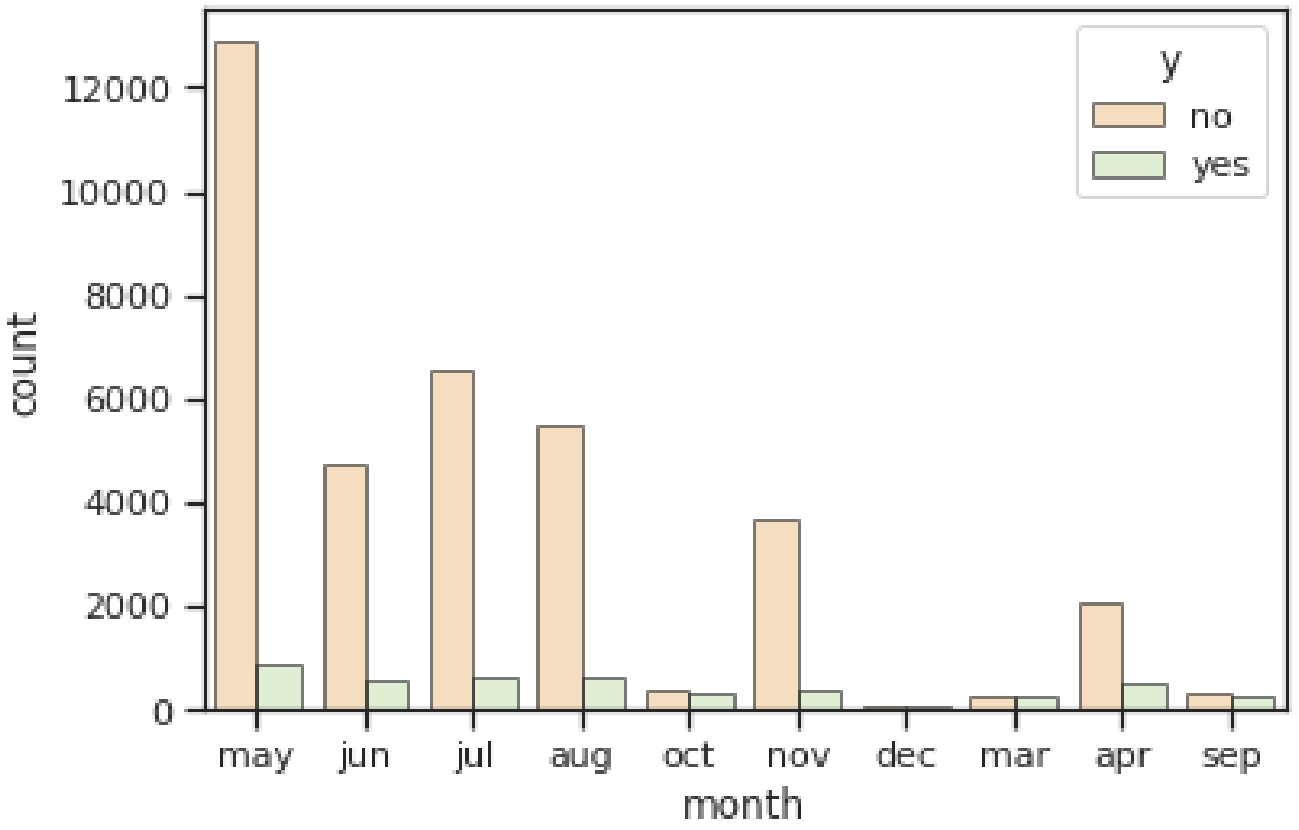






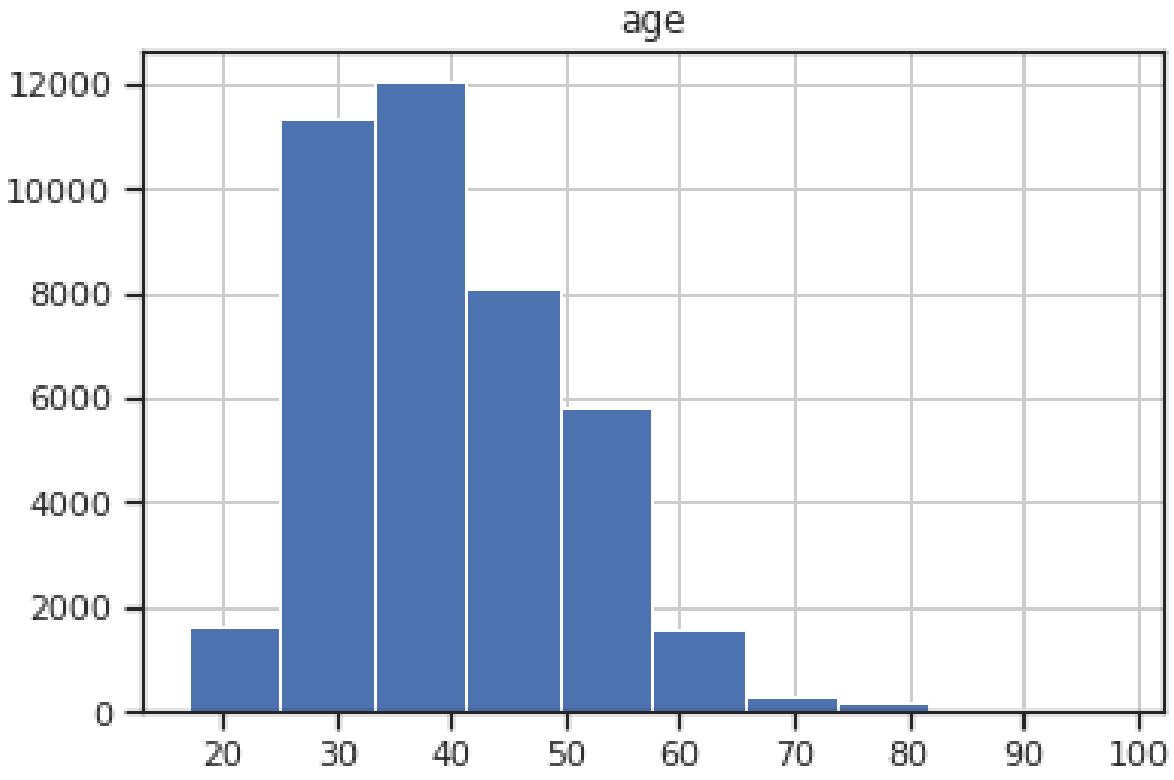



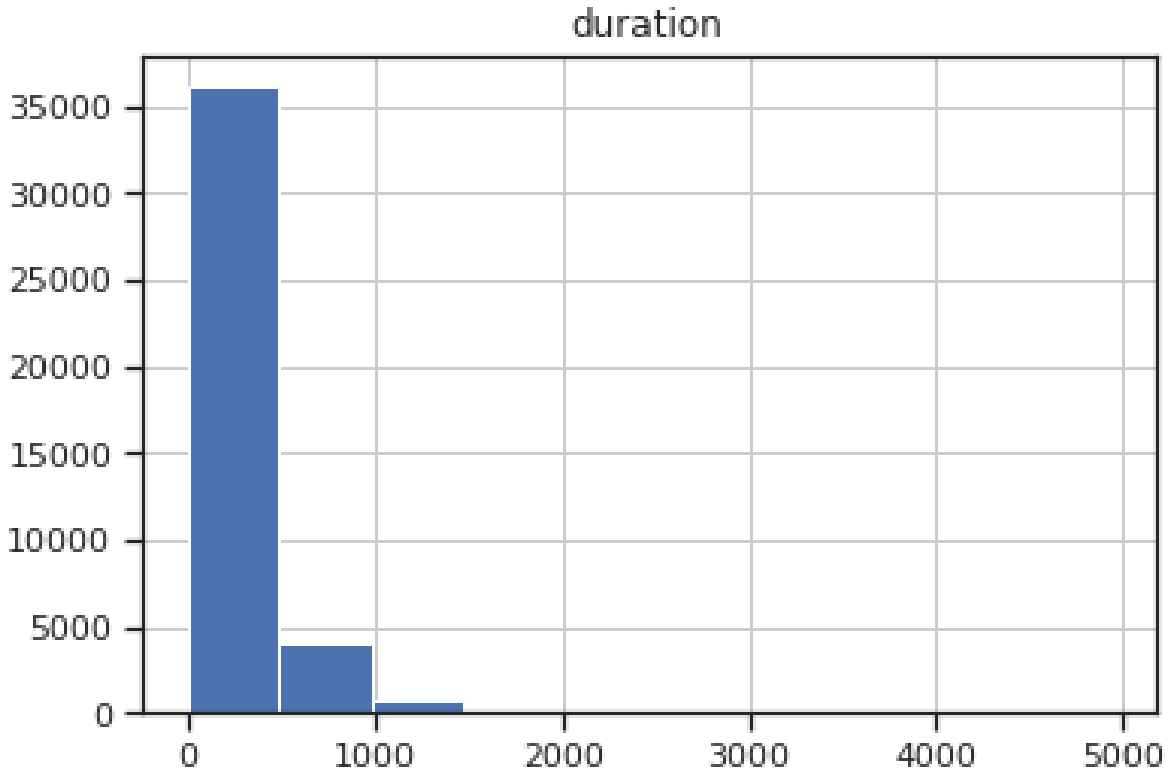
# Categorical EDA



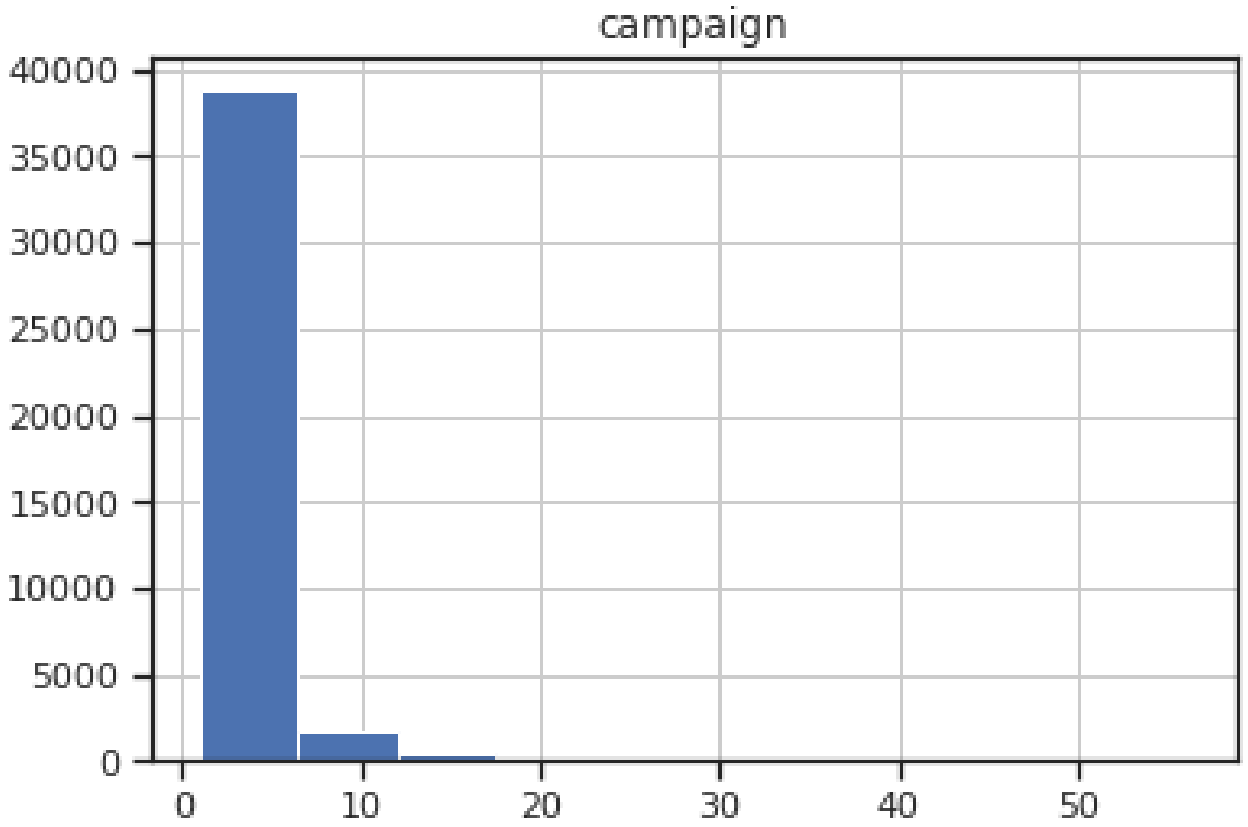

People are being contacted the most in the month of May.more People subscribed in may too.

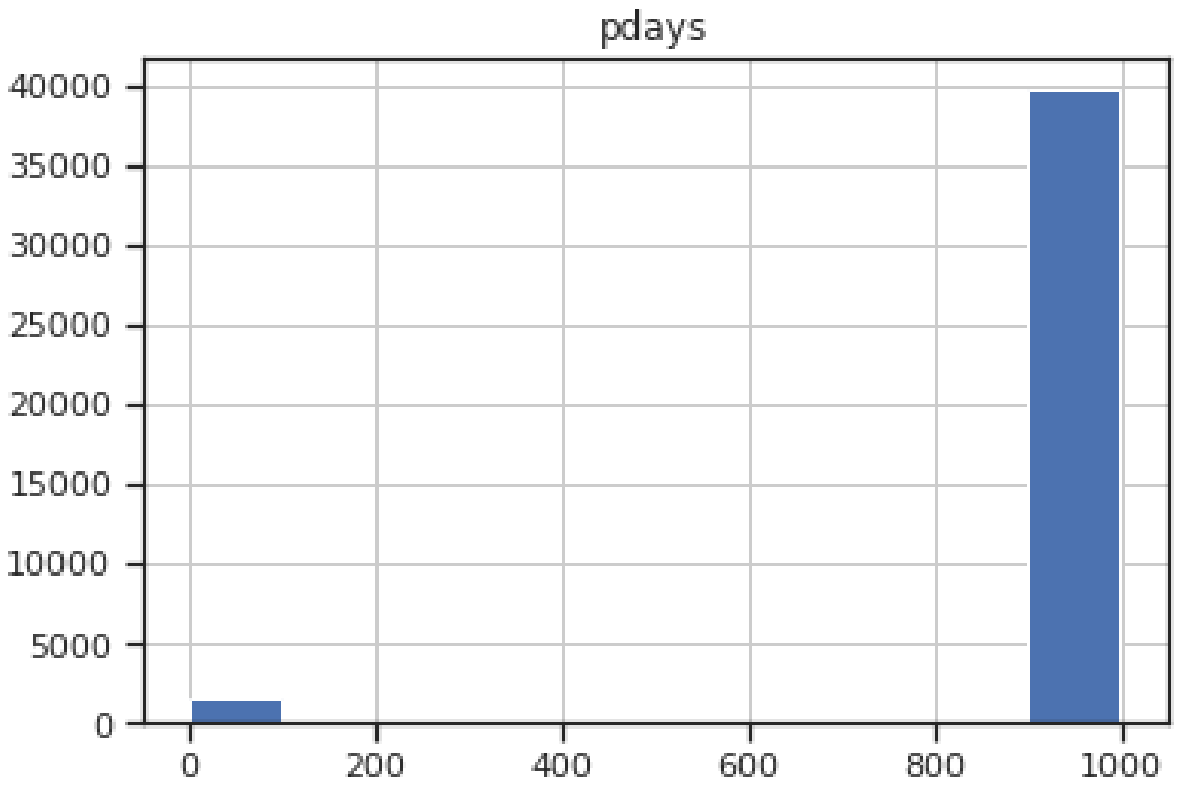
# Numerical EDA



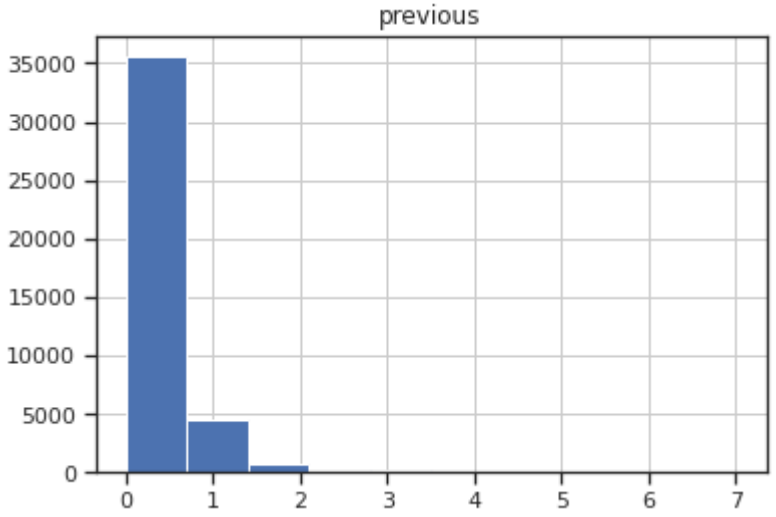



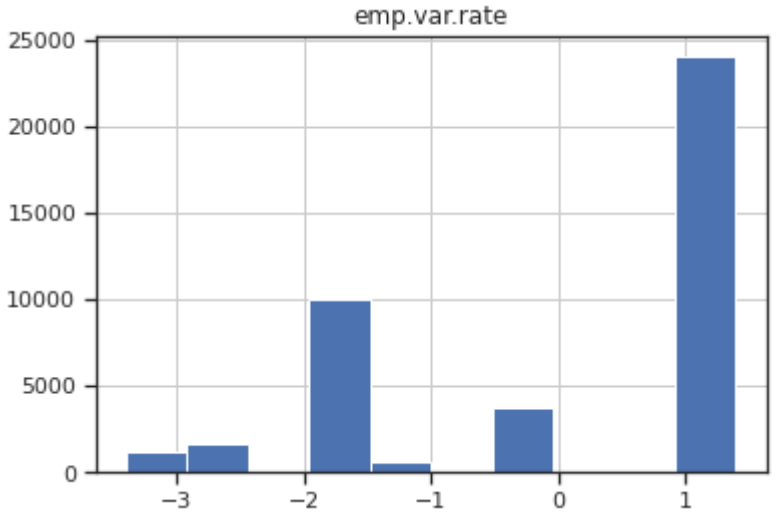
# Numerical EDA

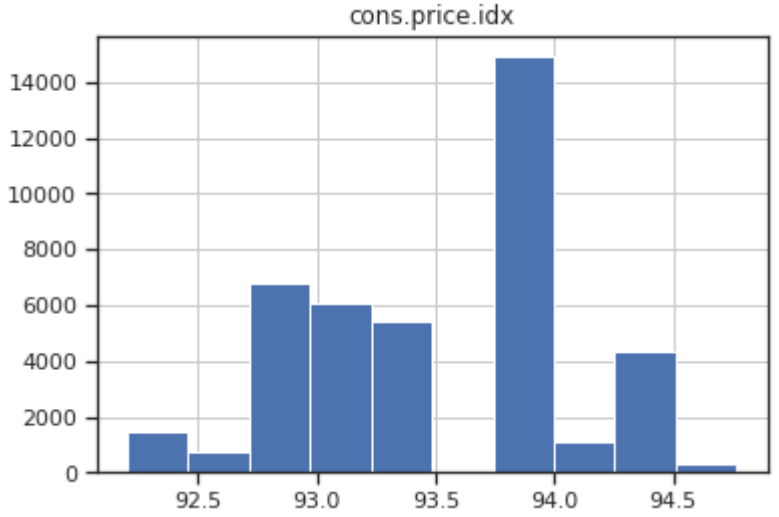


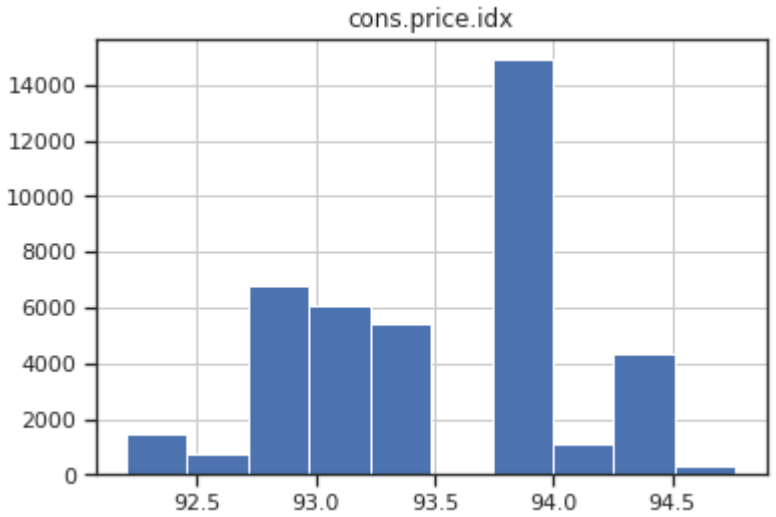



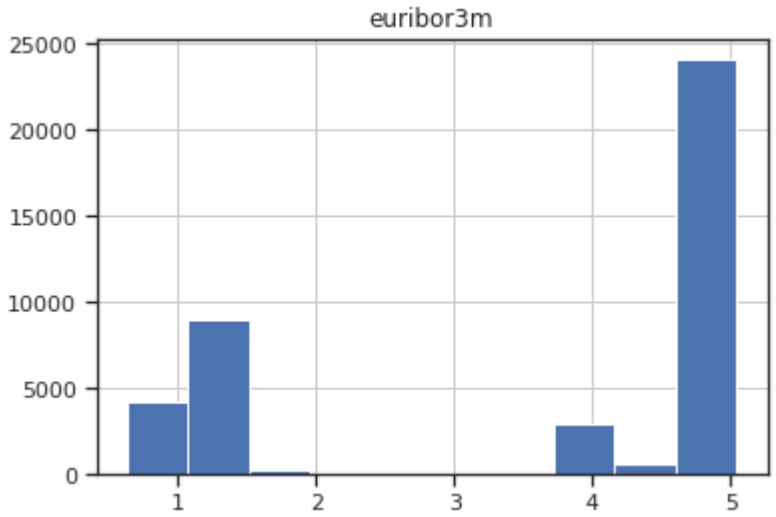
# Numerical EDA

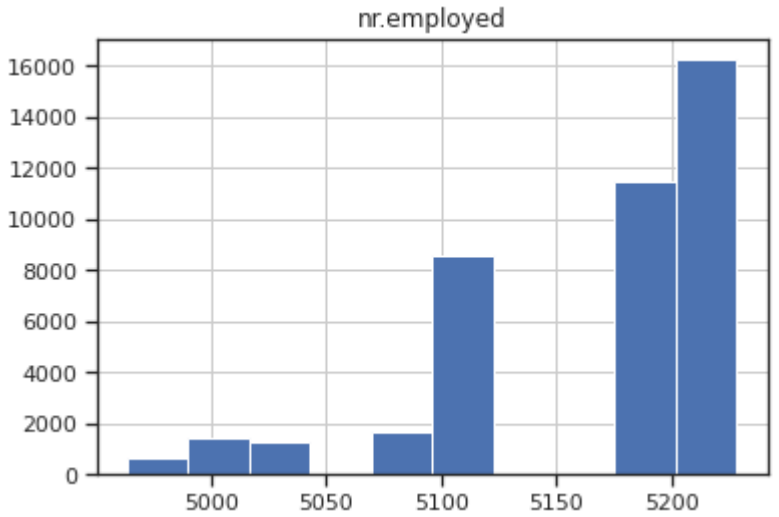


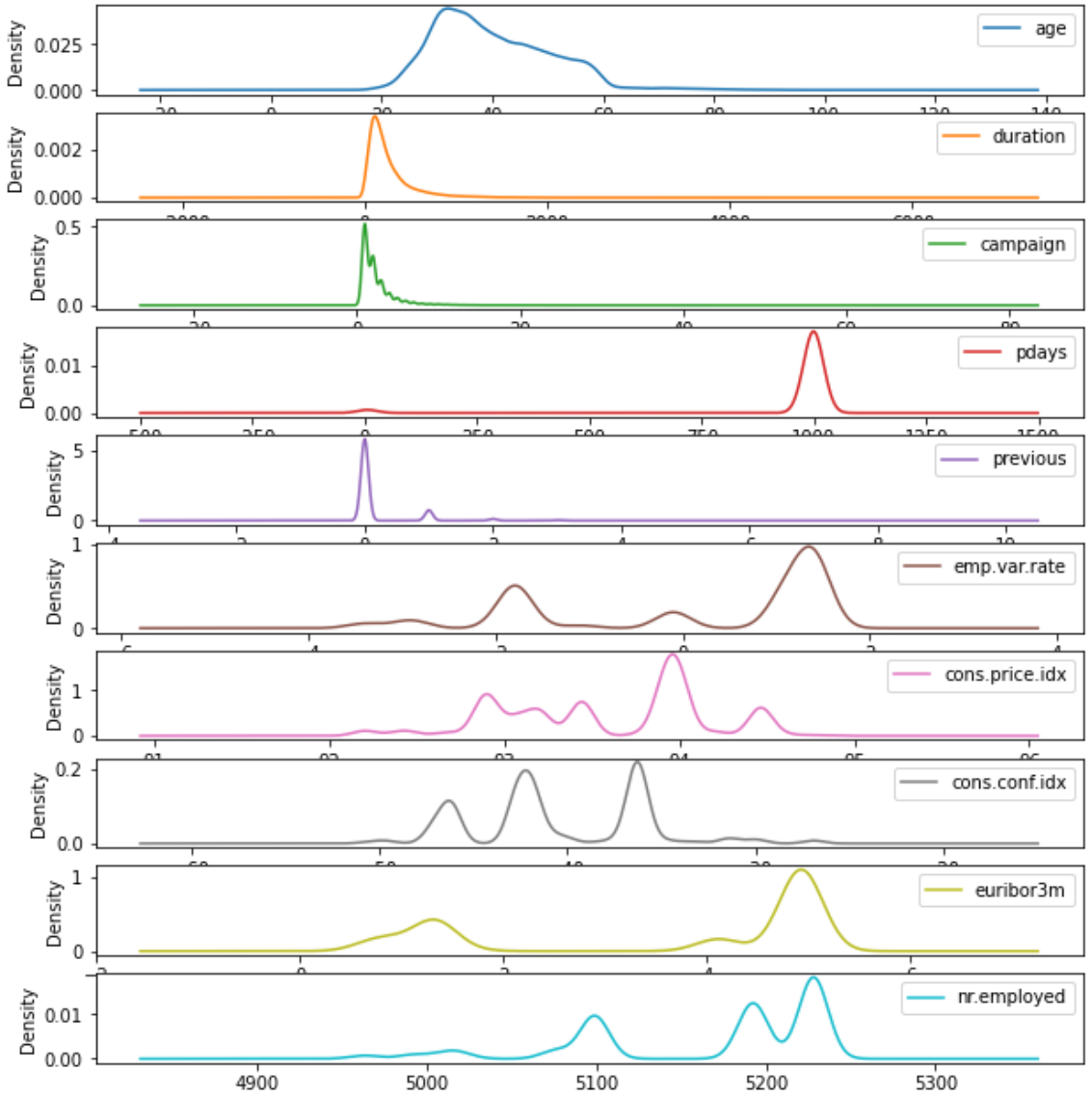



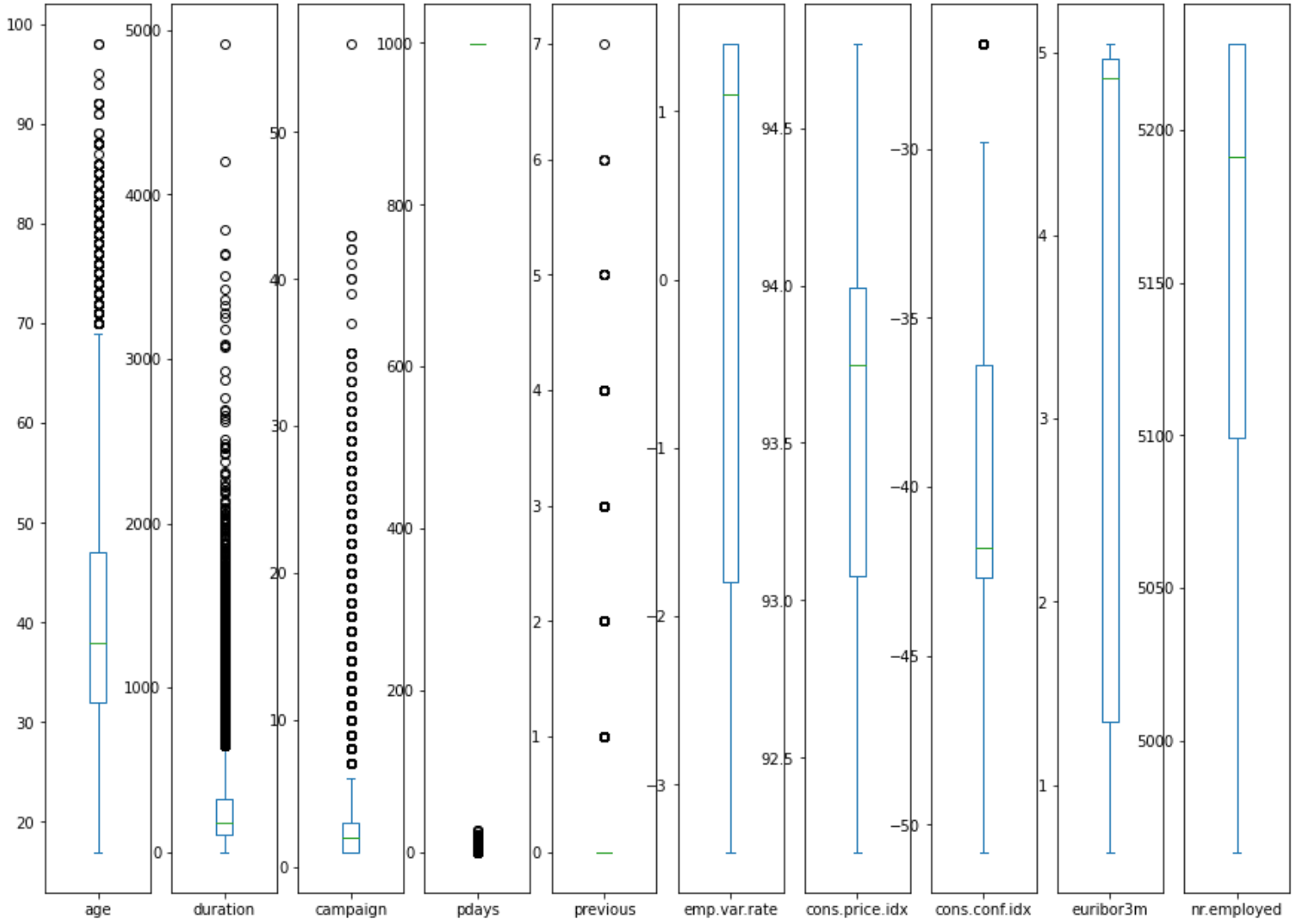



# Numerical EDA



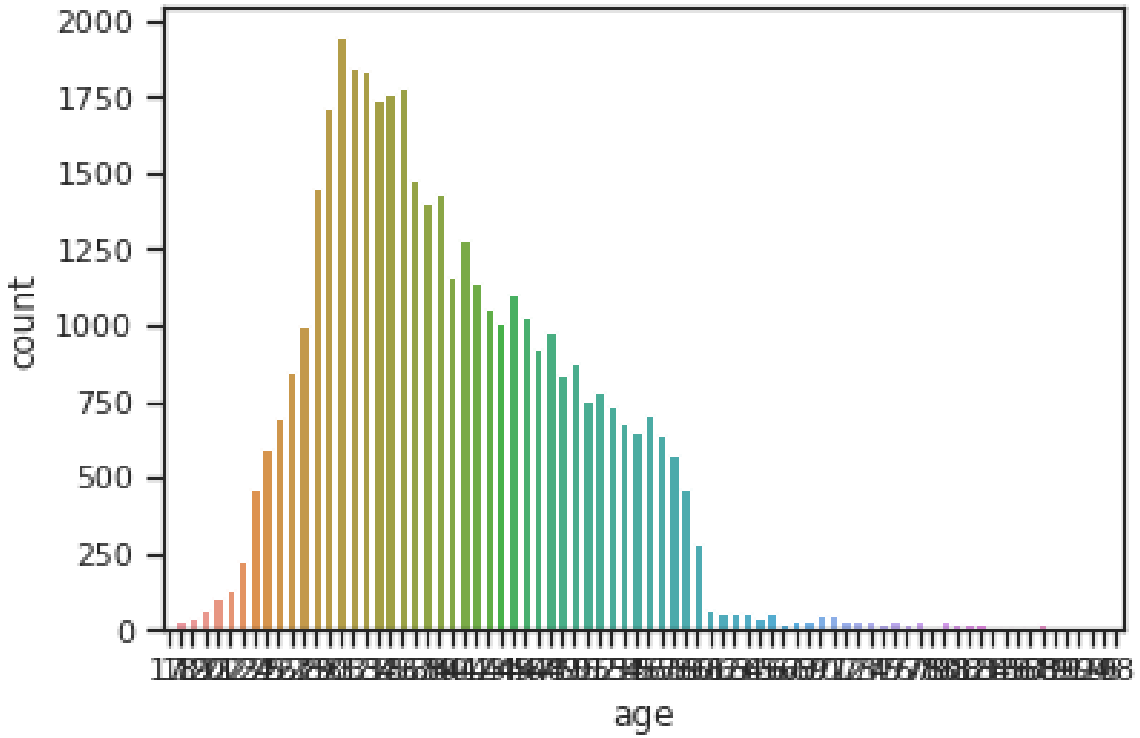

All Numerical Variable density Chart

# Numerical EDA

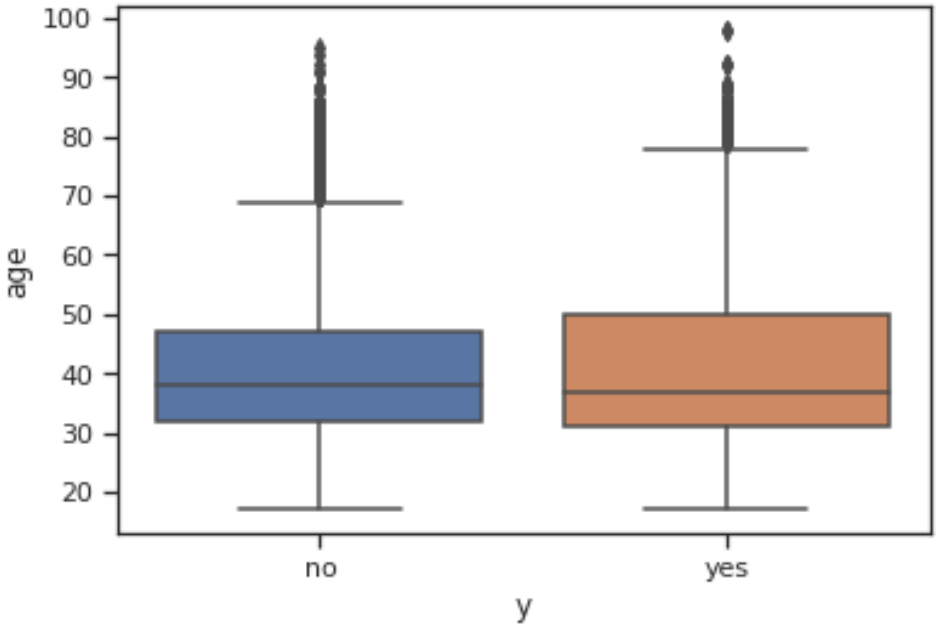



All Numerical Variables
Box plot

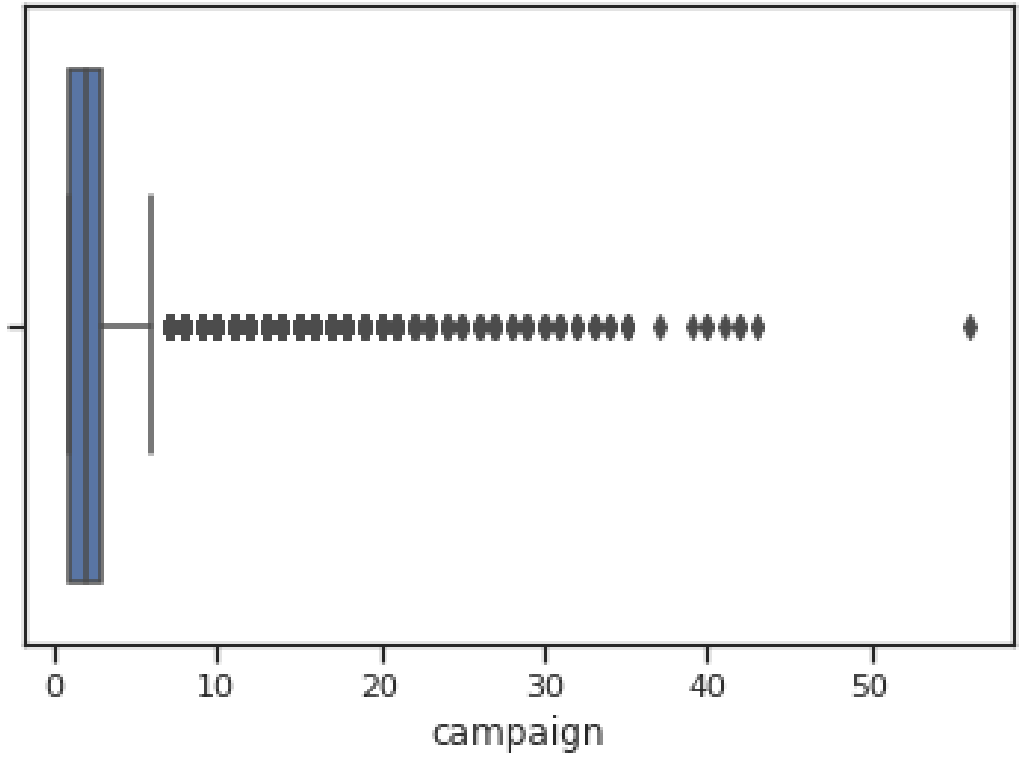
# Numerical EDA

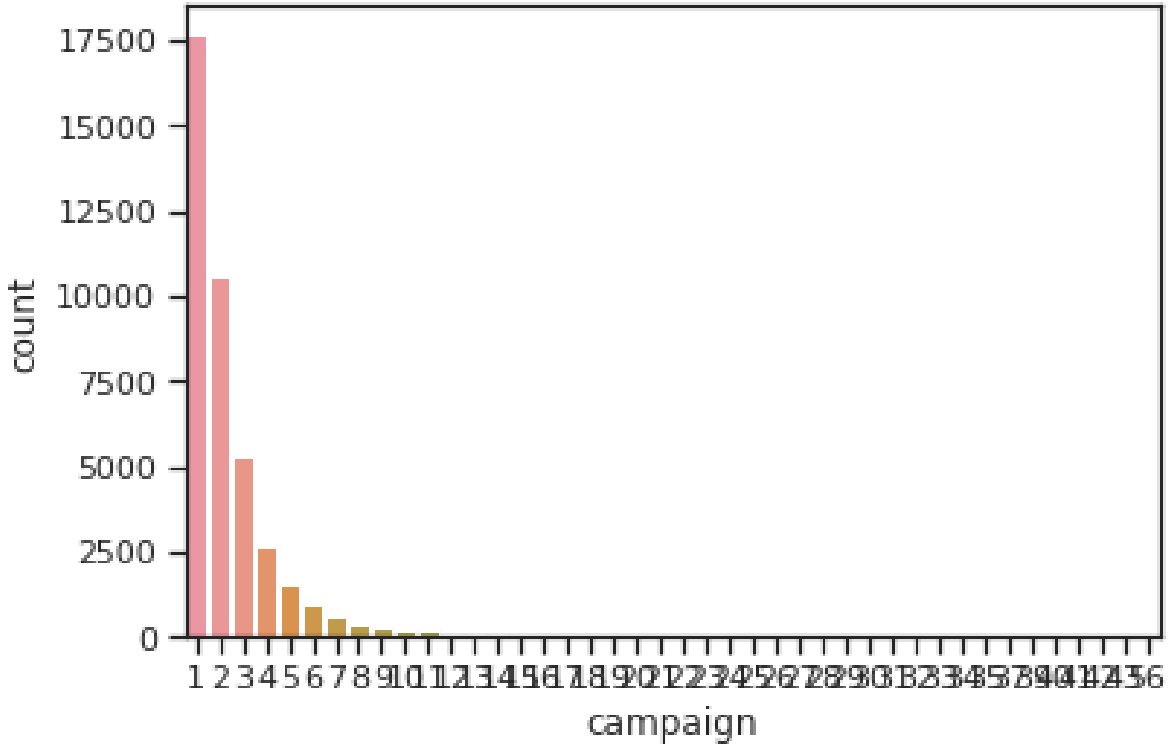



Majority of People are between age 35-40.

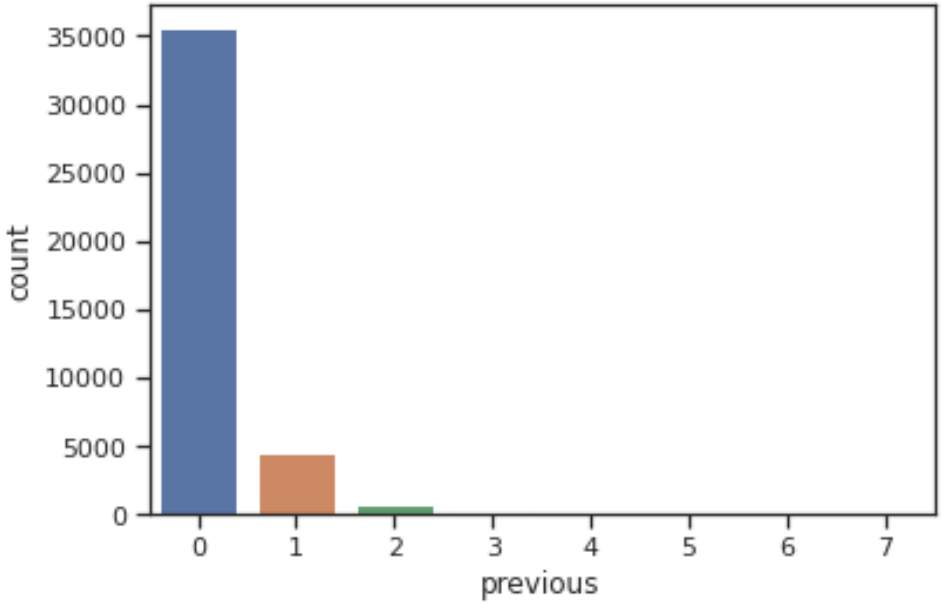


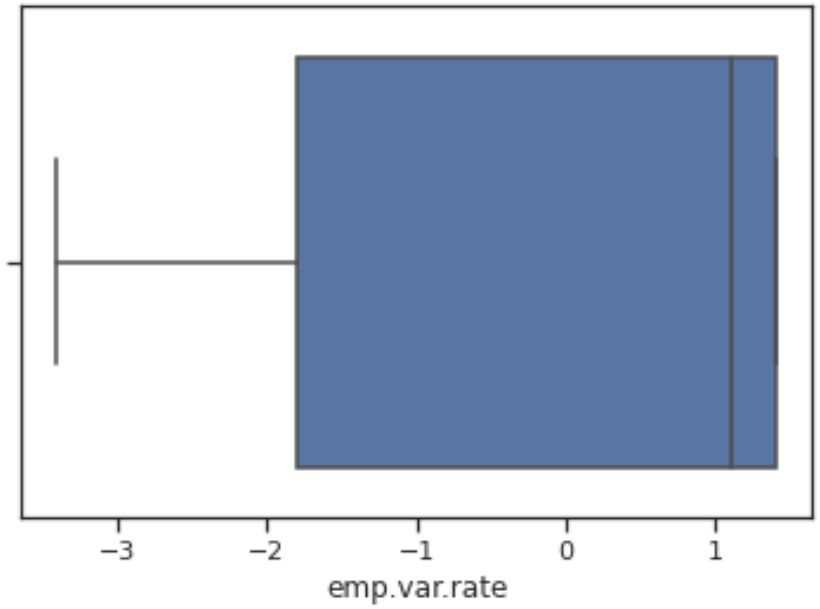
# Numerical EDA

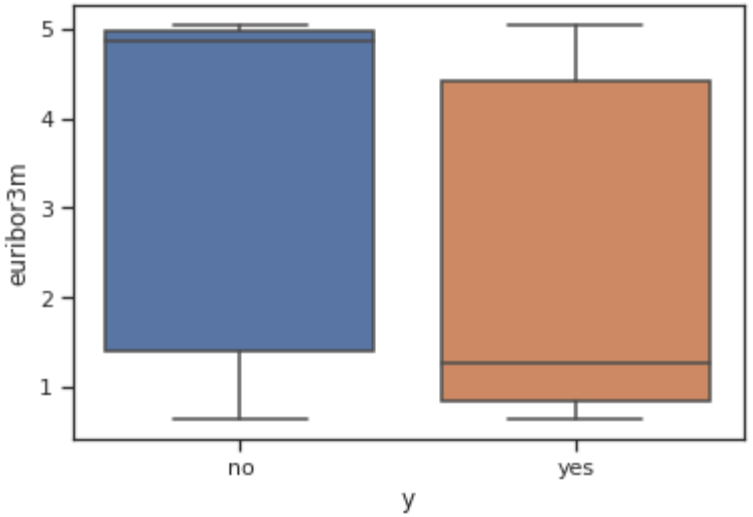


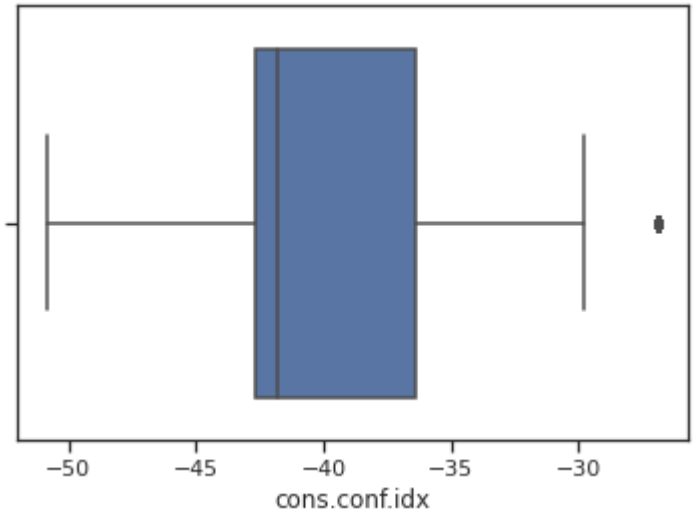



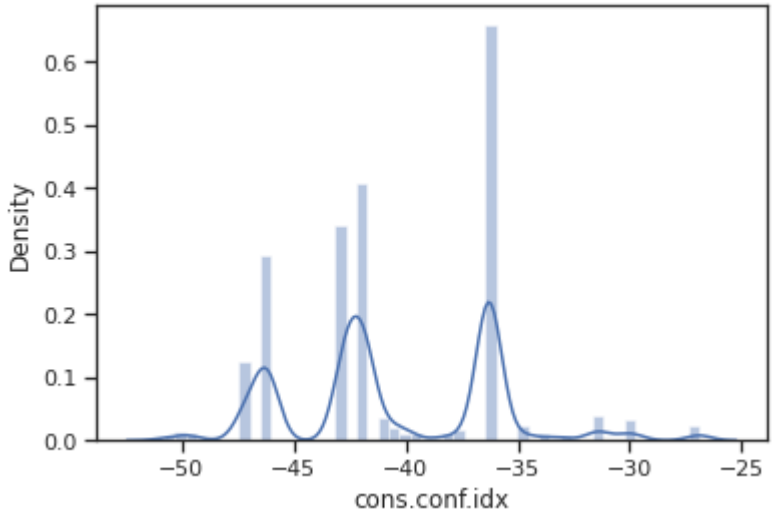
# Numerical EDA



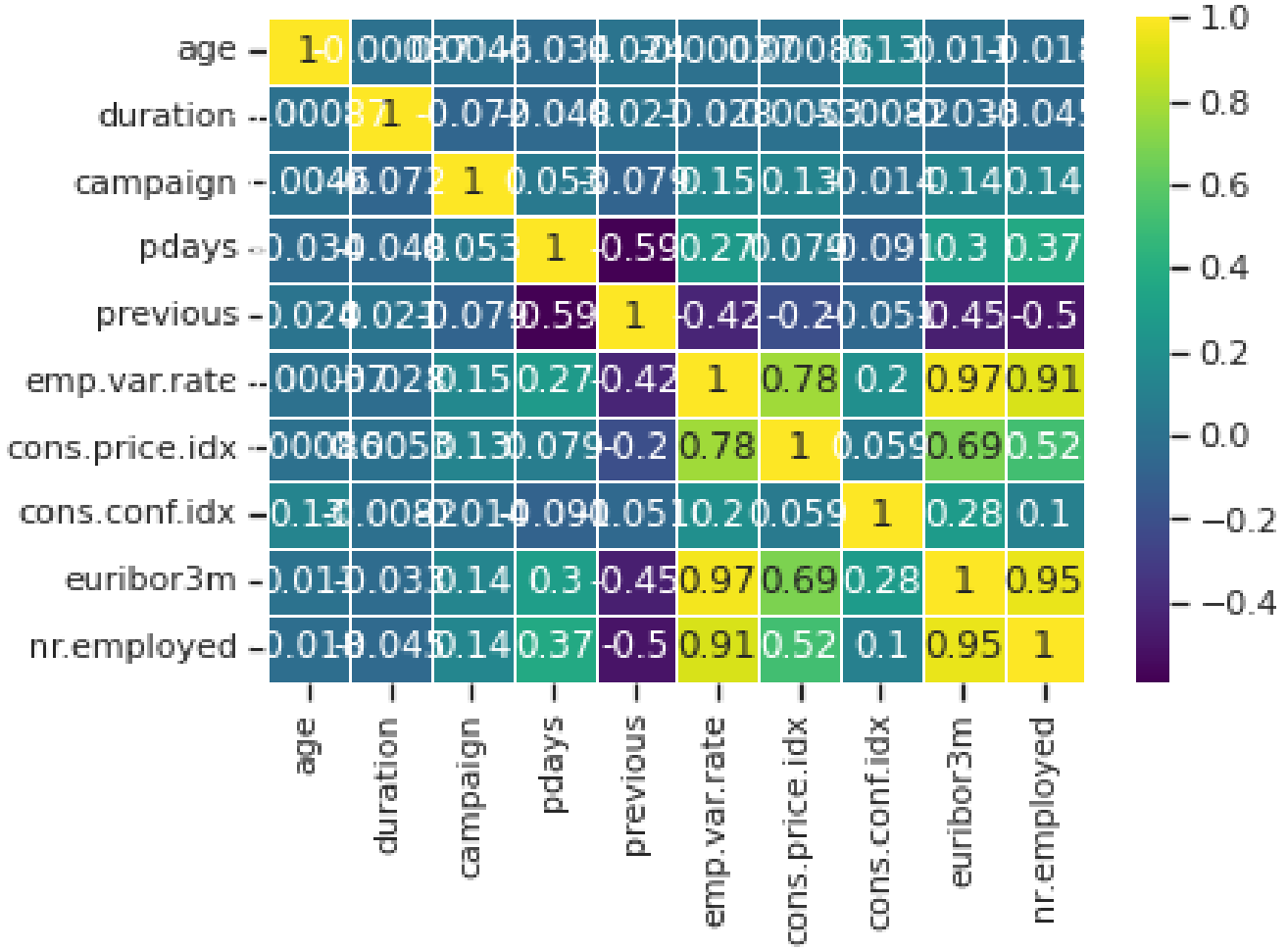



# Numerical EDA




Emp.var. Rate, euribor3m, nr.employedand cons.price.idx have higher correlation

# RECOMMENDED MODELS

**1-Logistic Regression**

**2-XGboos**

**3-Decition Tree**

**4-Random Forest**

**5-SVM**

# Thank You