

Bank-Term-Deposit-Analysis-Report

Problem Statement

The bank conducts direct marketing campaigns to promote term deposits to customers. These campaigns involve contacting customers via phone calls, which incurs operational costs. The objective of this analysis is to understand customer characteristics and campaign factors that influence whether a customer subscribes to a term deposit. By analyzing historical campaign data, the bank aims to improve targeting strategies and increase campaign efficiency.

Business Objective

The primary business objective is to analyze customer demographic, financial, and campaign-related factors to identify patterns associated with successful term deposit subscriptions. The insights from this analysis will help the bank optimize marketing efforts, reduce unnecessary customer contact, and improve overall campaign success rates.

Success Metrics

The success of the analysis will be measured by identifying key variables that have a strong relationship with term deposit subscription outcomes, improving understanding of customer segments more likely to subscribe, and providing actionable insights that can support better targeting decisions for future campaigns.

Data Overview

- The data is related to a bank marketing campaign
- The dataset contains 41,188 rows and 21 columns
- The data includes customer demographic details, socio-economic indicators, and campaign-related attributes.
- . No missing values were found in the dataset but several categorical variables contain the value “unknown”

Columns	Data Types	Unique Values
Job	Object	housemaid, services, admin., blue-collar,technician, retired, management, unemployed, self-employed, unknown, entrepreneur, student
Marital	Object	married, single, divorced, unknown
Education	Object	basic.4y, high.school, basic.6y, basic.9y,professional.course, unknown, university.degree, illiterate
Default	Object	no, unknown, yes
Housing	Object	no, yes, unknown
Loan	Object	no, yes, unknown
Contact	Object	telephone, cellular
Month	Object	may, jun, jul, aug, oct, nov, dec, mar, apr, sep
Day_of_week	Object	mon, tue ,wed, thu, fri
Poutcome	Object	nonexistent, failure, success
y	Object	no, yes

- No major inconsistencies or irregular category labels were observed across the categorical features
- The dataset also contains 12 duplicate records

```
[28]: # some basic information
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 41188 entries, 0 to 41187
Data columns (total 21 columns):
 #   Column            Non-Null Count  Dtype  
--- 
 0   age              41188 non-null    int64  
 1   job              41188 non-null    object  
 2   marital          41188 non-null    object  
 3   education        41188 non-null    object  
 4   default          41188 non-null    object  
 5   housing          41188 non-null    object  
 6   loan              41188 non-null    object  
 7   contact          41188 non-null    object  
 8   month             41188 non-null    object  
 9   day_of_week      41188 non-null    object  
 10  duration         41188 non-null    int64  
 11  campaign         41188 non-null    int64  
 12  pdays            41188 non-null    int64  
 13  previous         41188 non-null    int64  
 14  poutcome         41188 non-null    object  
 15  emp.var.rate     41188 non-null    float64 
 16  cons.price.idx  41188 non-null    float64 
 17  cons.conf.idx   41188 non-null    float64 
 18  euribor3m        41188 non-null    float64 
 19  nr.employed      41188 non-null    float64 
 20  y                41188 non-null    object  
dtypes: float64(5), int64(5), object(11)
memory usage: 6.6+ MB
```

Data Cleaning Strategy

1. Categorical values labeled as “unknown” were retained, as they represent undisclosed or unavailable information rather than missing data
2. Duplicate records were not remove because it can be contact attempt
3. The value `pdays = 999`, indicating customers who were never previously contacted, was preserved due to its strong business
4. No automatic outlier removal was performed, as extreme values may reflect genuine customer behavior

Univariate Analysis

Univariate analysis was conducted **only on the target variable (y)** to understand the overall subscription distribution in the dataset.

Subscription Status Count Percentage

	Count	Percentage
No	36,548	88.73%
Yes	4,640	11.27%

```
df.y.value_counts()
```

```
y
no      36548
yes     4640
Name: count, dtype: int64
```

```
df['y'].value_counts(normalize=True) * 100
```

```
y
no      88.734583
yes     11.265417
Name: proportion, dtype: float64
```

Key Insight

The dataset is **highly imbalanced**, with significantly fewer positive subscription cases

It's important because

- It affects campaign performance interpretation
- It must be considered in further analysis and modeling

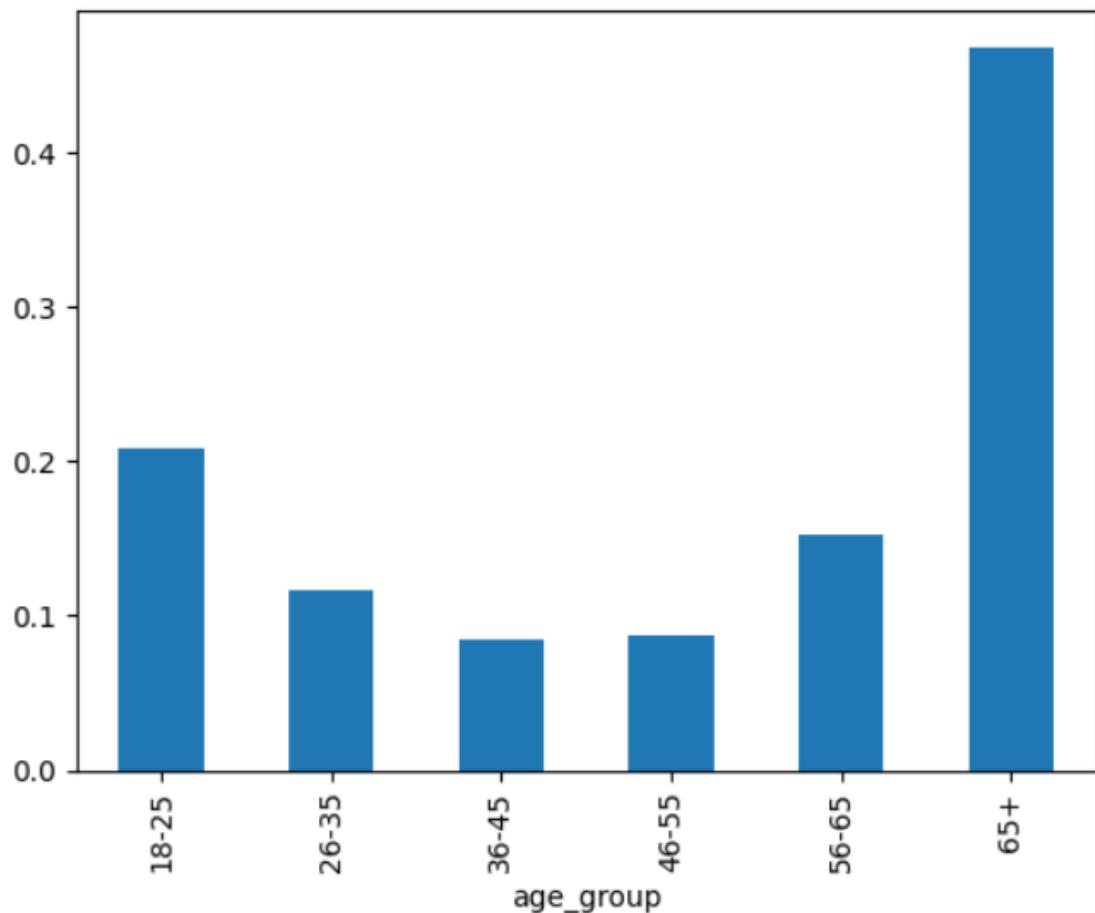
Bivariate Analysis

Bivariate analysis was performed to understand the relationship between **customer attributes** and the **term deposit subscription outcome (y)**.

The analysis focuses on identifying variables that significantly influence subscription behavior.

Age Group vs Subscription

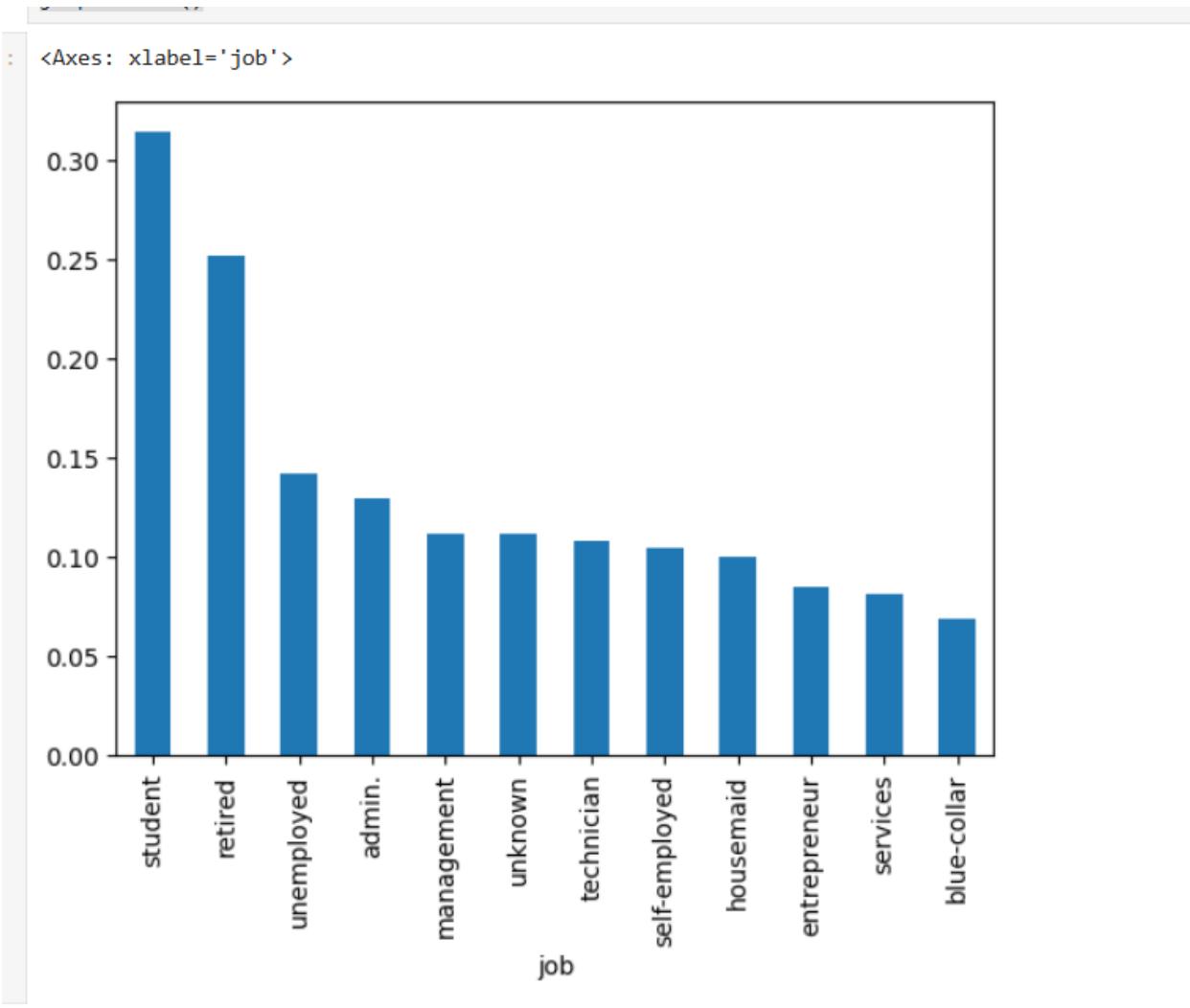
```
<Axes: xlabel='age_group'>
```



Observation:-

- Older age groups show relatively higher subscription rates compared to younger customers
- Younger age groups (18–25 and 26–35) have **lower subscription rates** compared to older customers.

Job vs Subscription



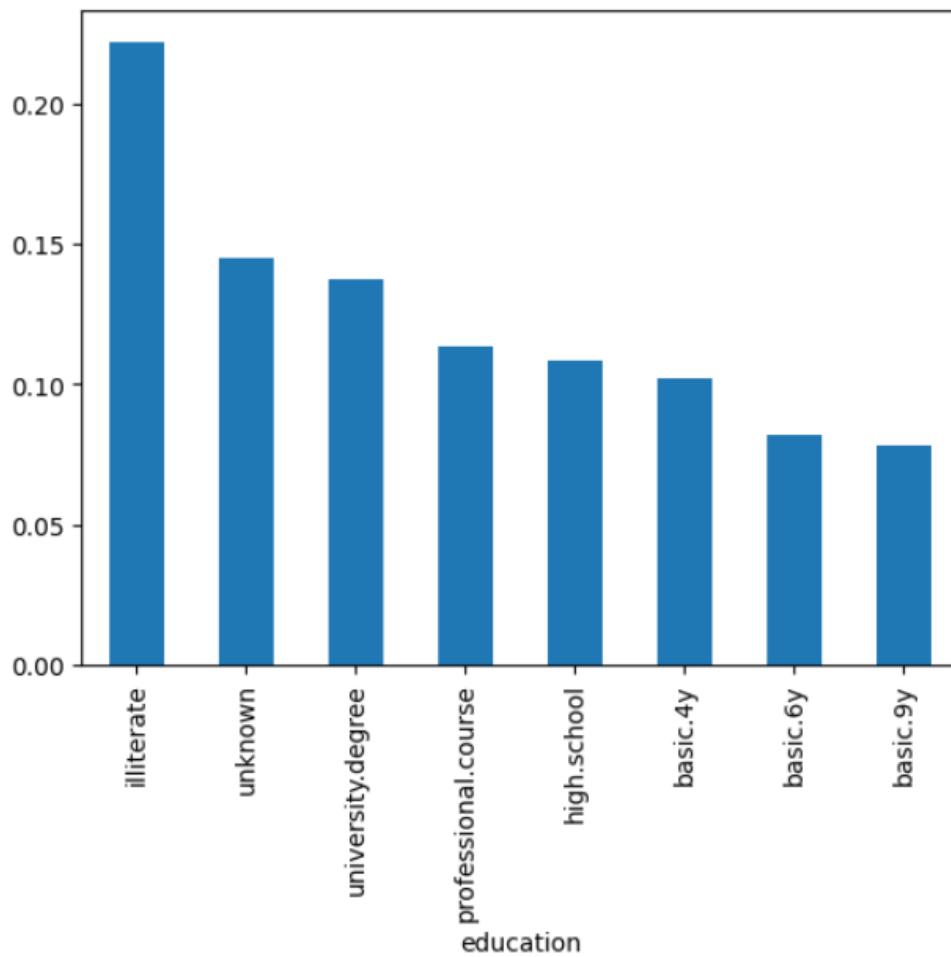
Observation:-

- Students and retired customers exhibit the highest subscription rates
- blue-collar and service-sector customers show the lowest conversion rates.

Insight:- suggests that marketing campaigns may achieve better efficiency by prioritizing job segments with higher historical response rates.

Education vs Subscription

<Axes: xlabel='education'>

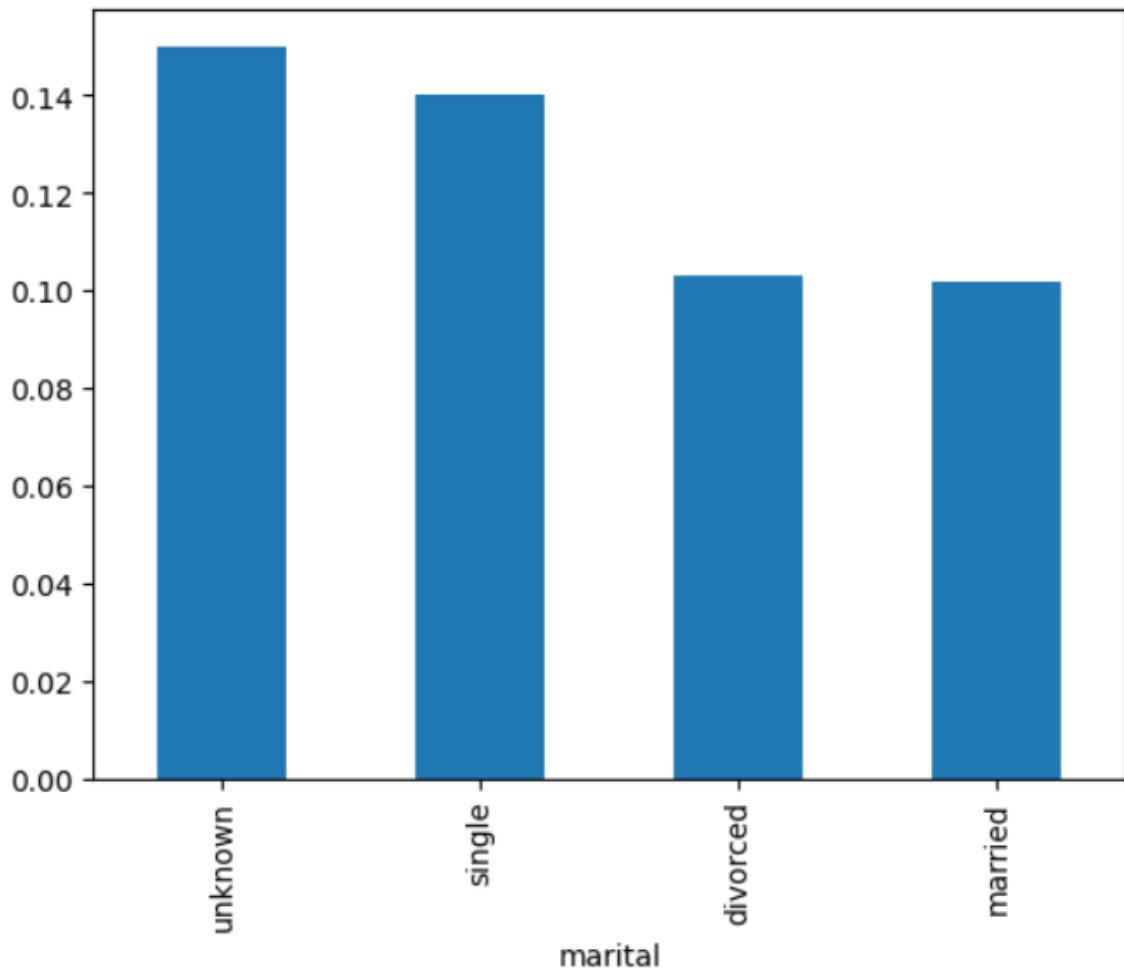


Observation:-

- Education-wise analysis indicates that customers with university degrees and professional courses demonstrate higher subscription rates compared to those with basic education levels.
- Although illiterate and unknown categories show higher proportions, these results should be interpreted cautiously due to their relatively small sample sizes.
- Education appears to play a meaningful role in subscription behavior, with more educated customers being more receptive to term deposit offerings.

Marital vs Subscription

```
(Axes: xlabel='marital')>
```

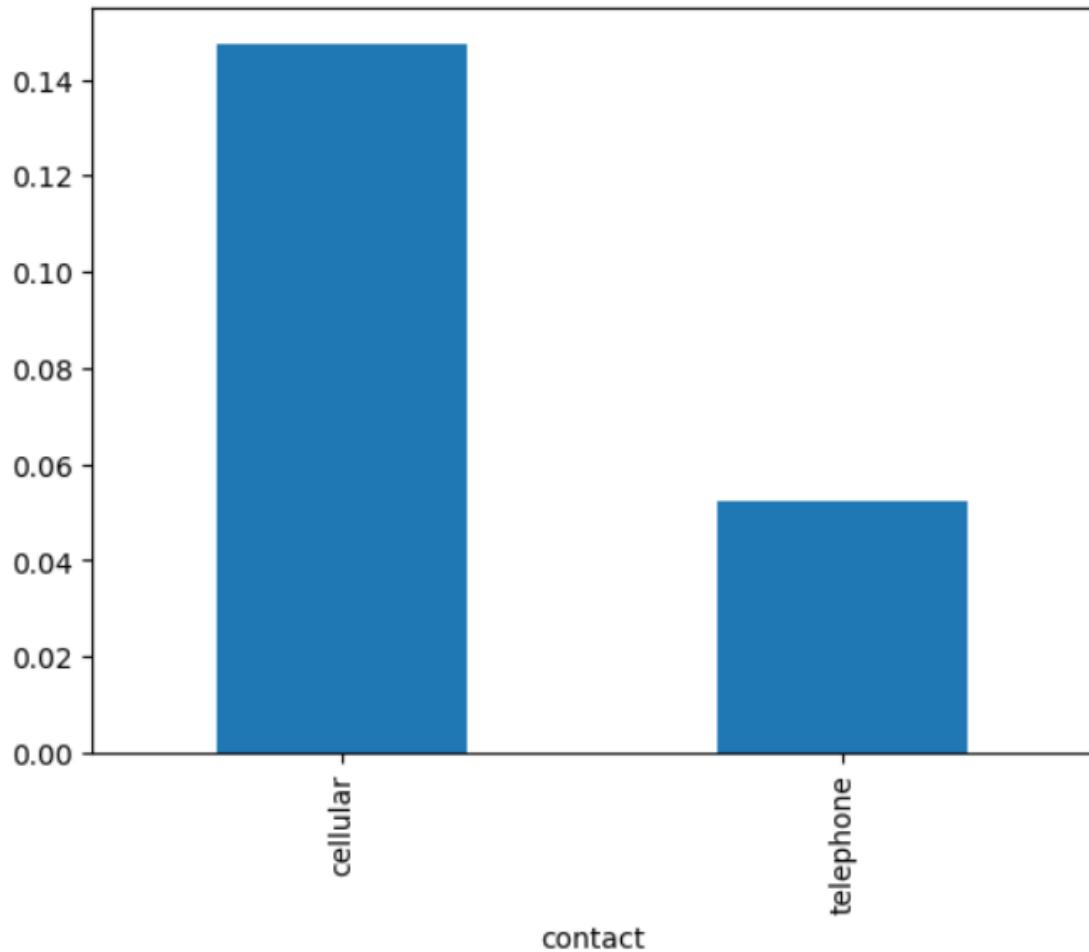


Observation:-

- Single customers exhibit a higher subscription rate compared to married and divorced individuals.
- This suggests that personal financial flexibility may influence a customer's willingness to invest in term deposits.

Contact Type vs Subscription

```
<Axes: xlabel='contact'>
```

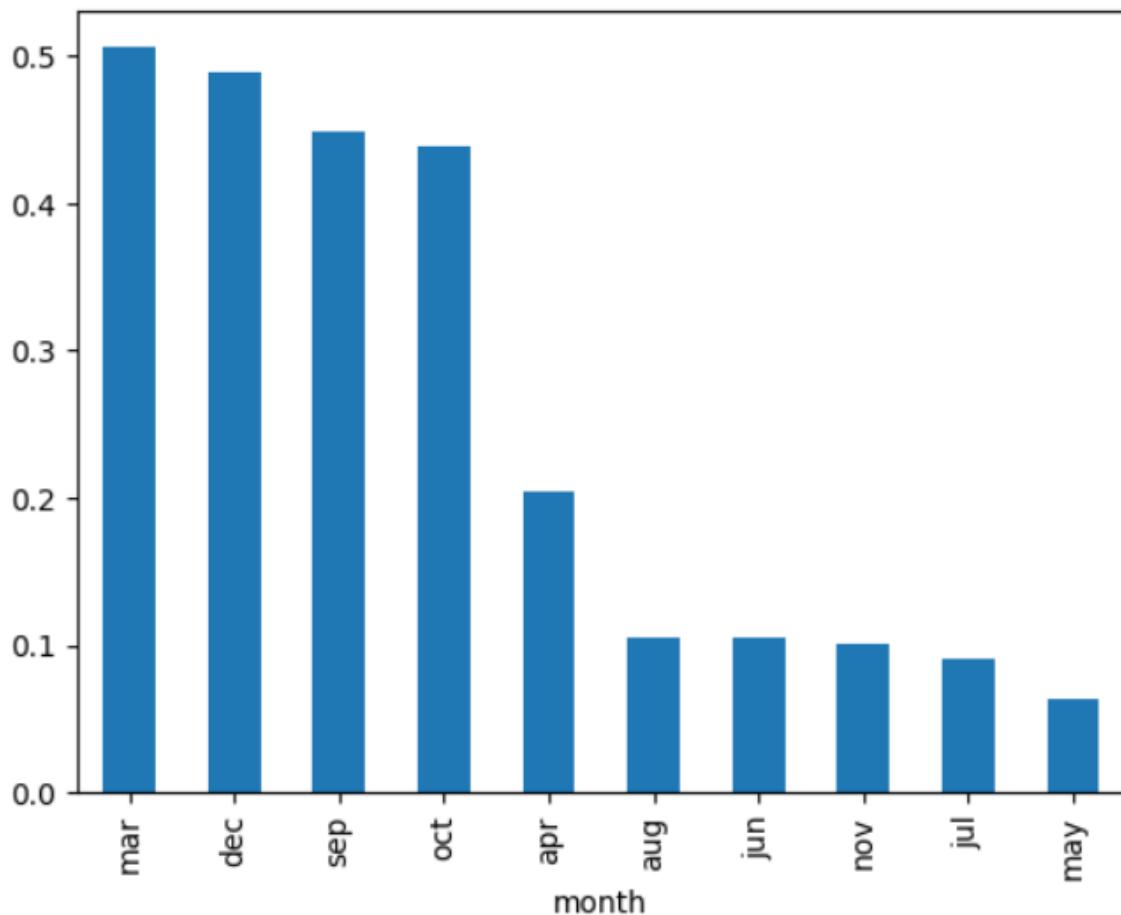


Observation:-

- Analysis of contact type shows that cellular communication results in higher subscription rates compared to telephone-based contact.
- This indicates that customers are more responsive to mobile-based communication channels, suggesting that marketing campaigns should prioritize cellular outreach methods.
- This improves campaign efficiency and maximizes subscription rates while reducing wasted contact attempts.

Month vs Subscription

<Axes: xlabel='month'>

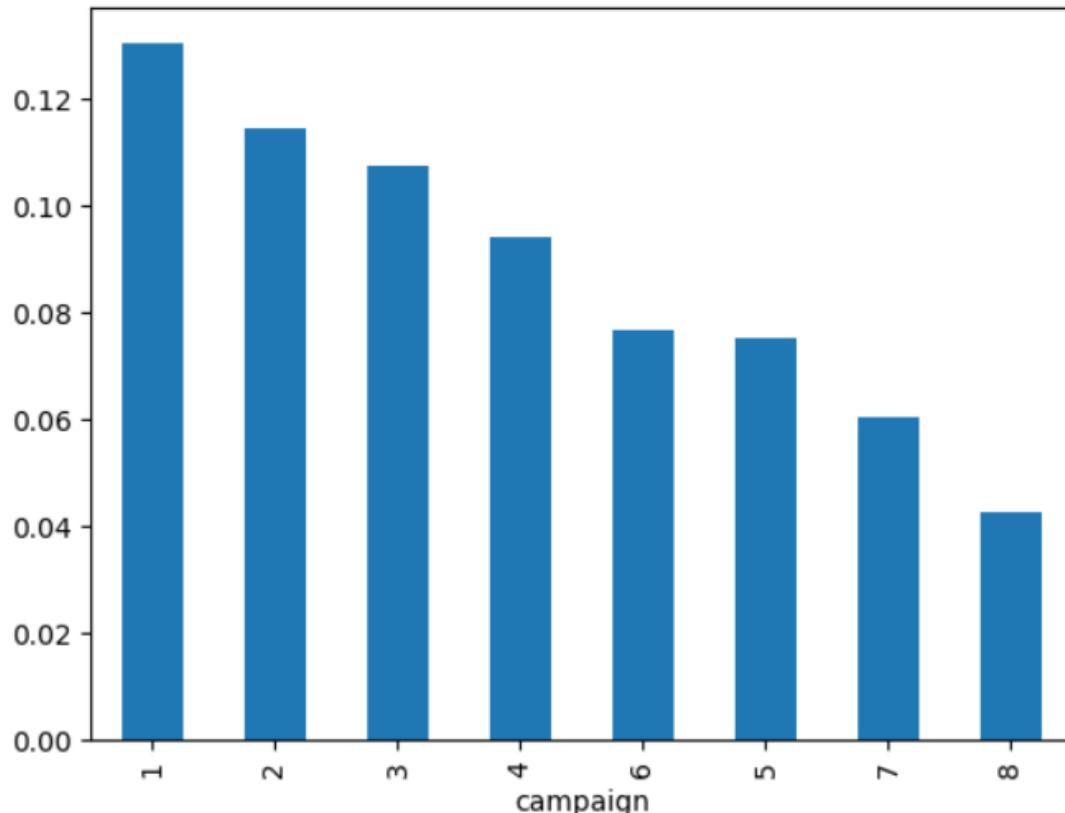


Observation:-

- Campaign timing has a strong impact on subscription success, with certain months delivering significantly higher conversion rates than others
- Campaigns conducted during March, September, October, and December exhibit significantly higher conversion rates, whereas months such as May and July show notably lower effectiveness despite higher campaign activity.
- Campaign timing appears to have a stronger impact than campaign volume.

Campaign Calls vs Subscription

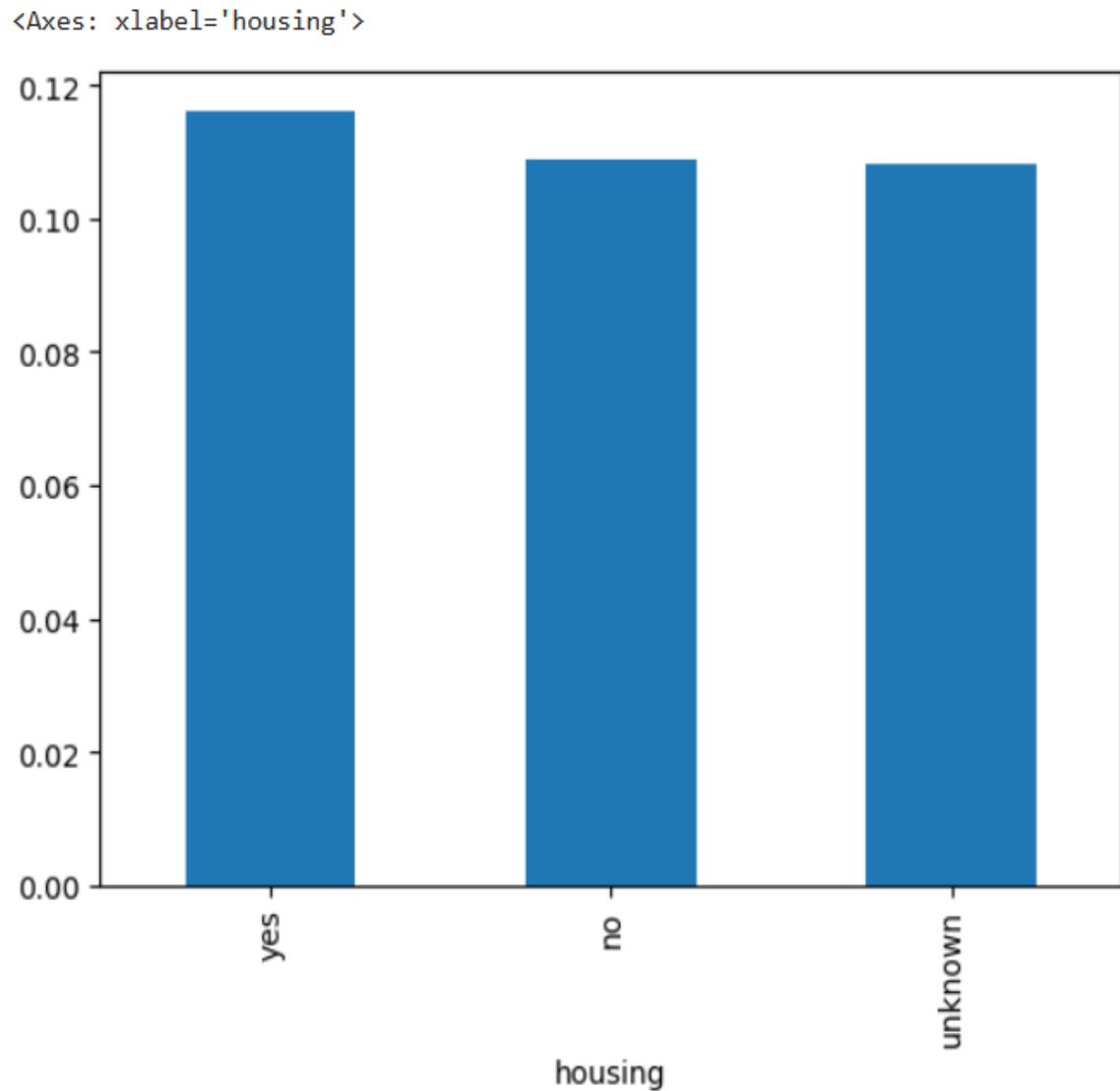
```
<Axes: xlabel='campaign'>
```



Observation:-

- Customers contacted only once have the highest probability of subscribing. As the number of calls increases, the subscription rate consistently decreases, indicating that repeated contact reduces customer interest. [¶](#)
- Excessive campaign calls negatively impact subscription success; therefore, fewer and more targeted contacts are suggested

Housing vs Subscription

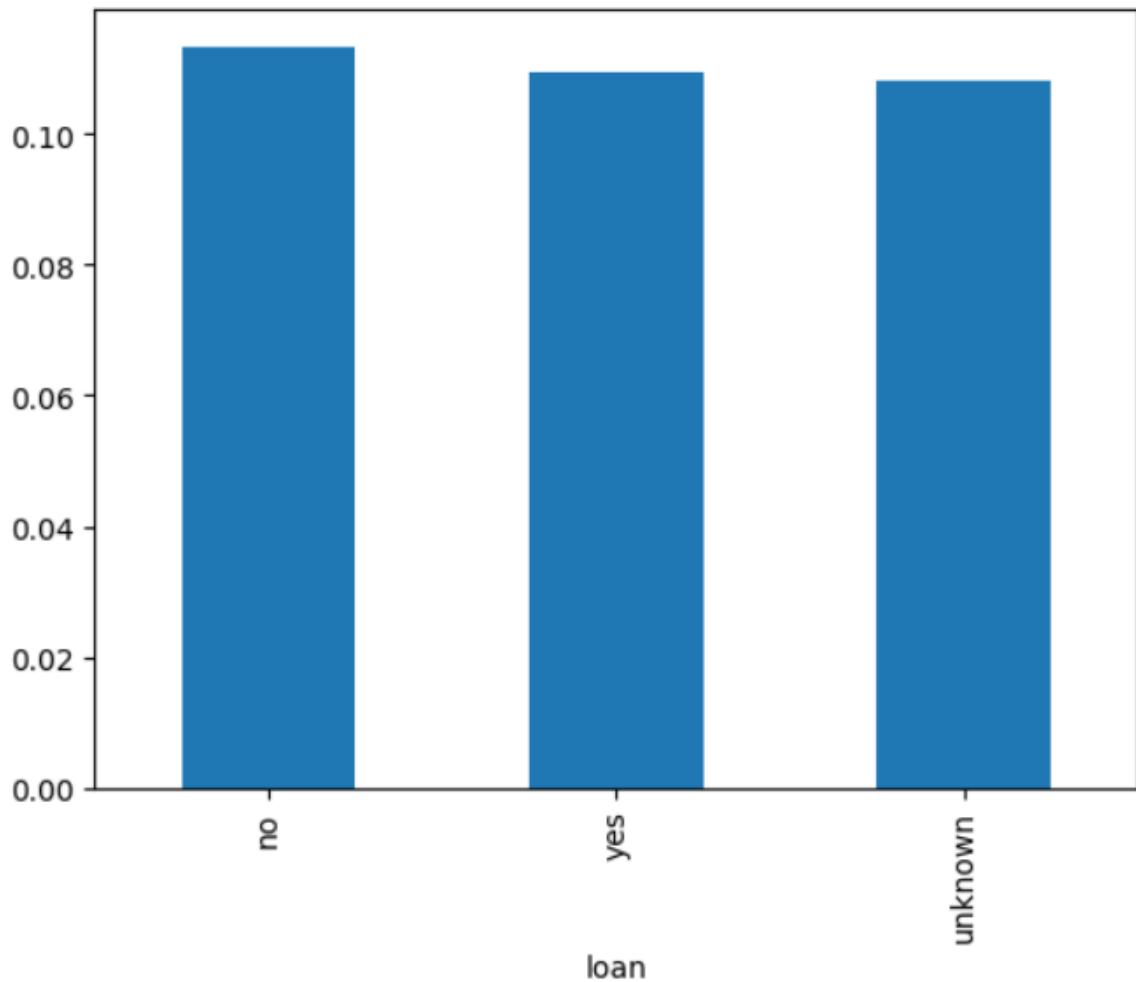


Observation:-

- People with a housing loan subscribed slightly more than people without a housing loan.
- Housing loan status does affect subscription, but [The difference is small, not very strong.](#)
- suggesting housing loan is not a strong deciding factor for subscription

loan status vs Subscription

<Axes: xlabel='loan'>

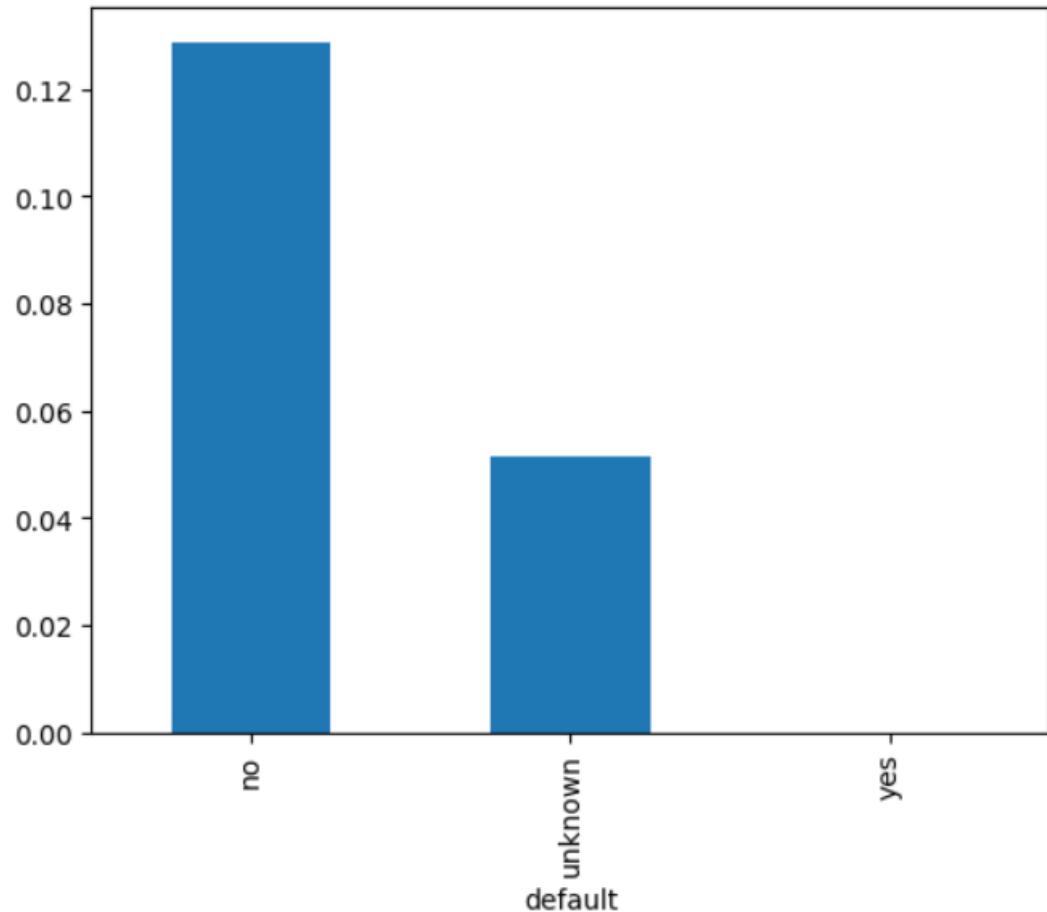


Observation:-

- Customers without a personal loan subscribed the most
- Customers with a loan subscribed slightly less
- Difference is very small So loan status does not strongly influence subscription.

Default vs subscription

<Axes: xlabel='default'>

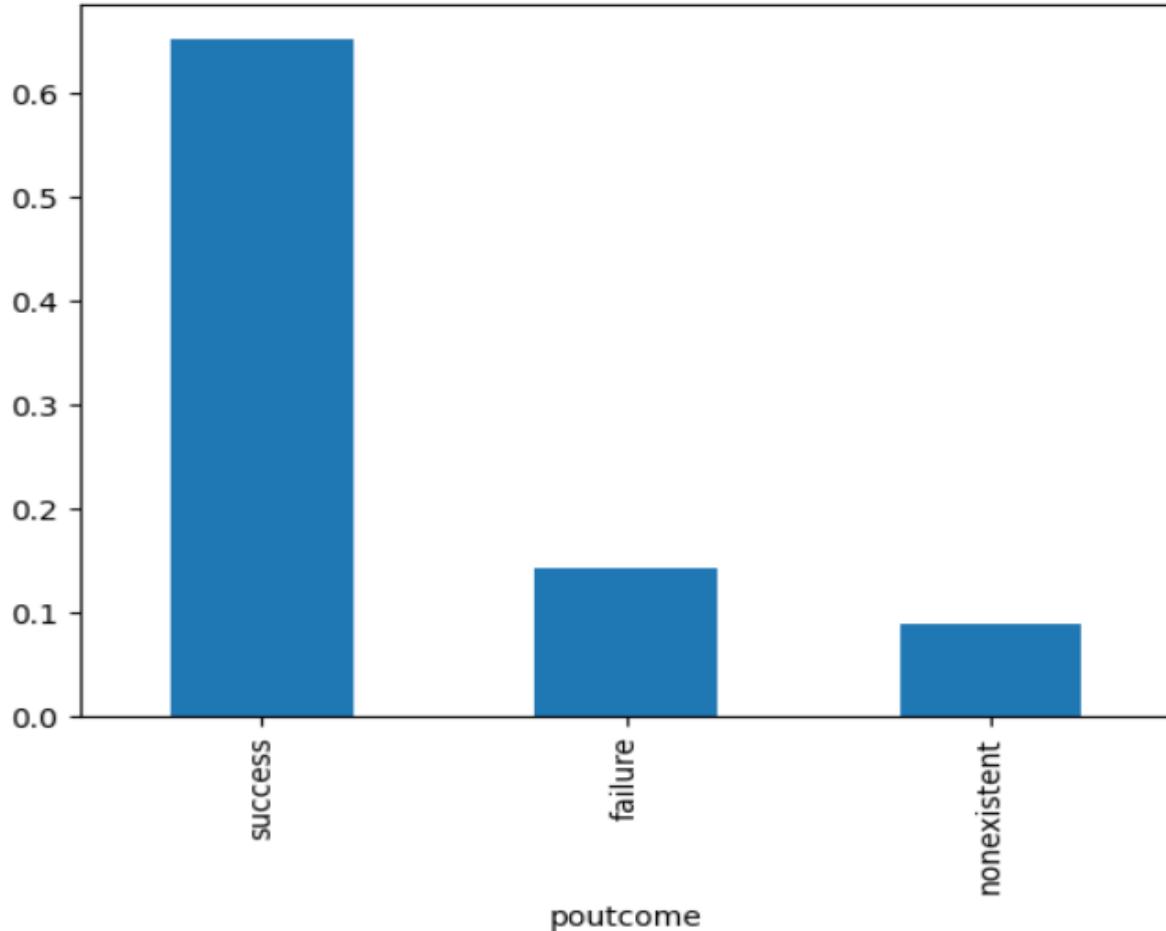


Observation:-

Customers without default are more likely to subscribe. Customers with default are less likely to subscribe, but this group is very small in the dataset. So conclusions should be interpreted cautiously.[¶](#)

poutcome vs subscription

<Axes: xlabel='poutcome'>



Observation :-

- If a customer subscribed in a previous campaign,
- ↗ there is a 65% chance they will subscribe again.
- If a customer rejected earlier, ↗ only 14% chance they subscribe now.
- If customer was never contacted, ↗ conversion is very low (~9%).

Recommendations :-

Customers who previously subscribed are highly likely to subscribe again. Marketing efforts should prioritize such customers to improve campaign efficiency.

Multivariate Analysis

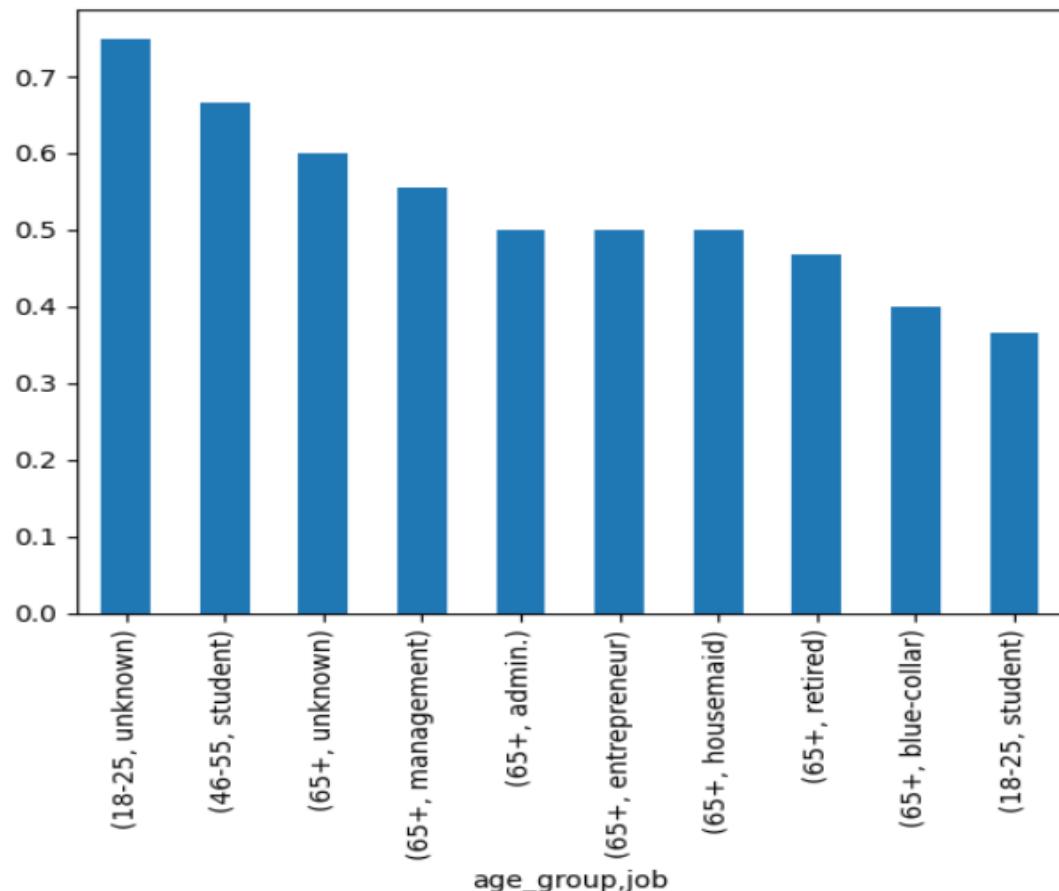
Multivariate analysis was performed to understand how combinations of multiple customer attributes influence term deposit subscription behavior.

age vs job vs subscription

```
agejob_sub
```

```
age_group    job
18-25        unknown      0.750000
46-55        student      0.666667
65+          unknown      0.600000
                  management 0.555556
                  admin.       0.500000
                  entrepreneur 0.500000
                  housemaid   0.500000
                  retired     0.467167
                  blue-collar  0.400000
18-25        student      0.365256
Name: yes, dtype: float64
```

```
<Axes: xlabel='age_group,job'>
```



Insight :-

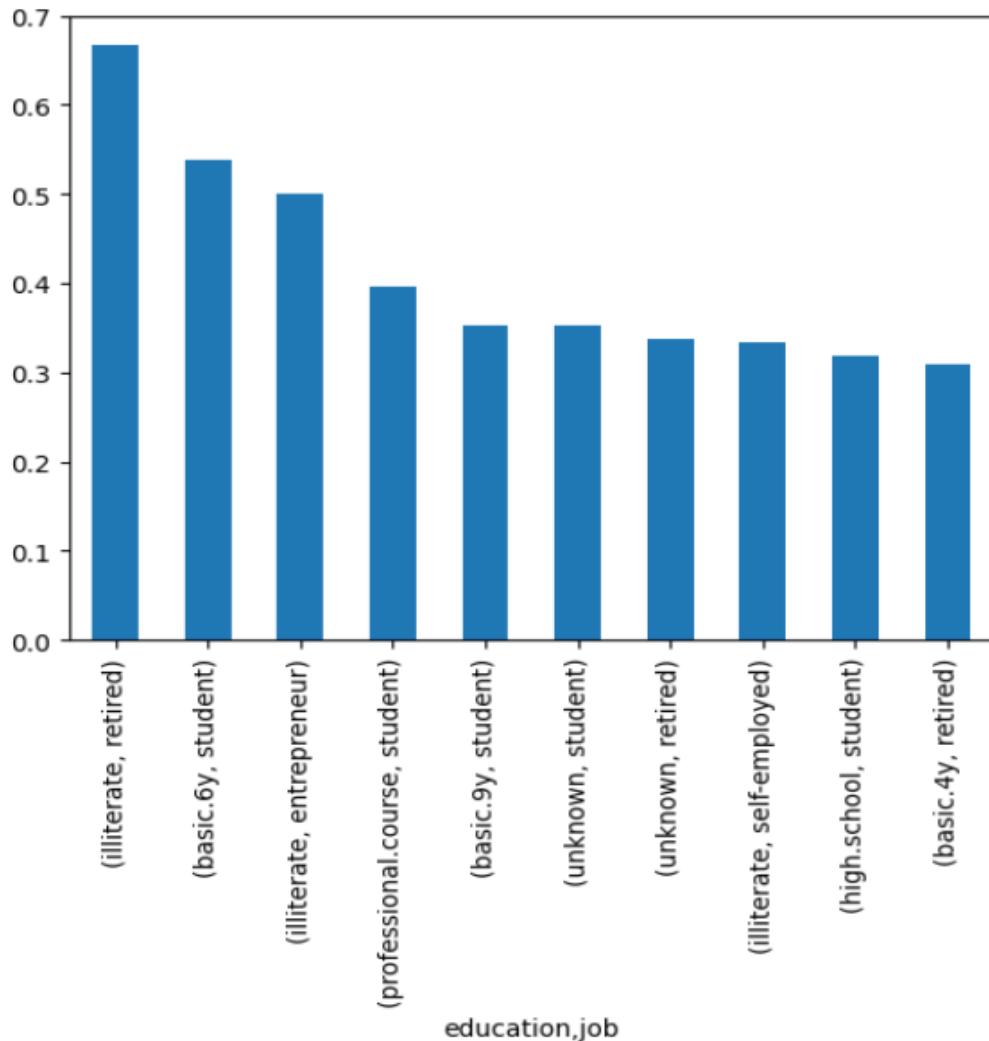
The combined analysis of age group and job role reveals that older customers in professional or retired roles exhibit higher subscription rates. while blue-collar and service employees across most age groups exhibit lower conversion Young students can subscribe well, but only in specific cases

This confirms that subscription behavior is influenced by the interaction of demographic and occupational factors rather than age or job independently.

Education vs job vs subscription

```
: education          job
illiterate        retired      0.666667
basic.6y         student     0.538462
illiterate        entrepreneur 0.500000
professional.course student   0.395349
basic.9y         student     0.353535
unknown           student     0.353293
                  retired     0.336735
illiterate        self-employed 0.333333
high.school       student     0.319328
basic.4y         retired     0.309883
Name: yes, dtype: float64
```

```
: educjob_sub.plot.bar()
: <Axes: xlabel='education,job'>
```

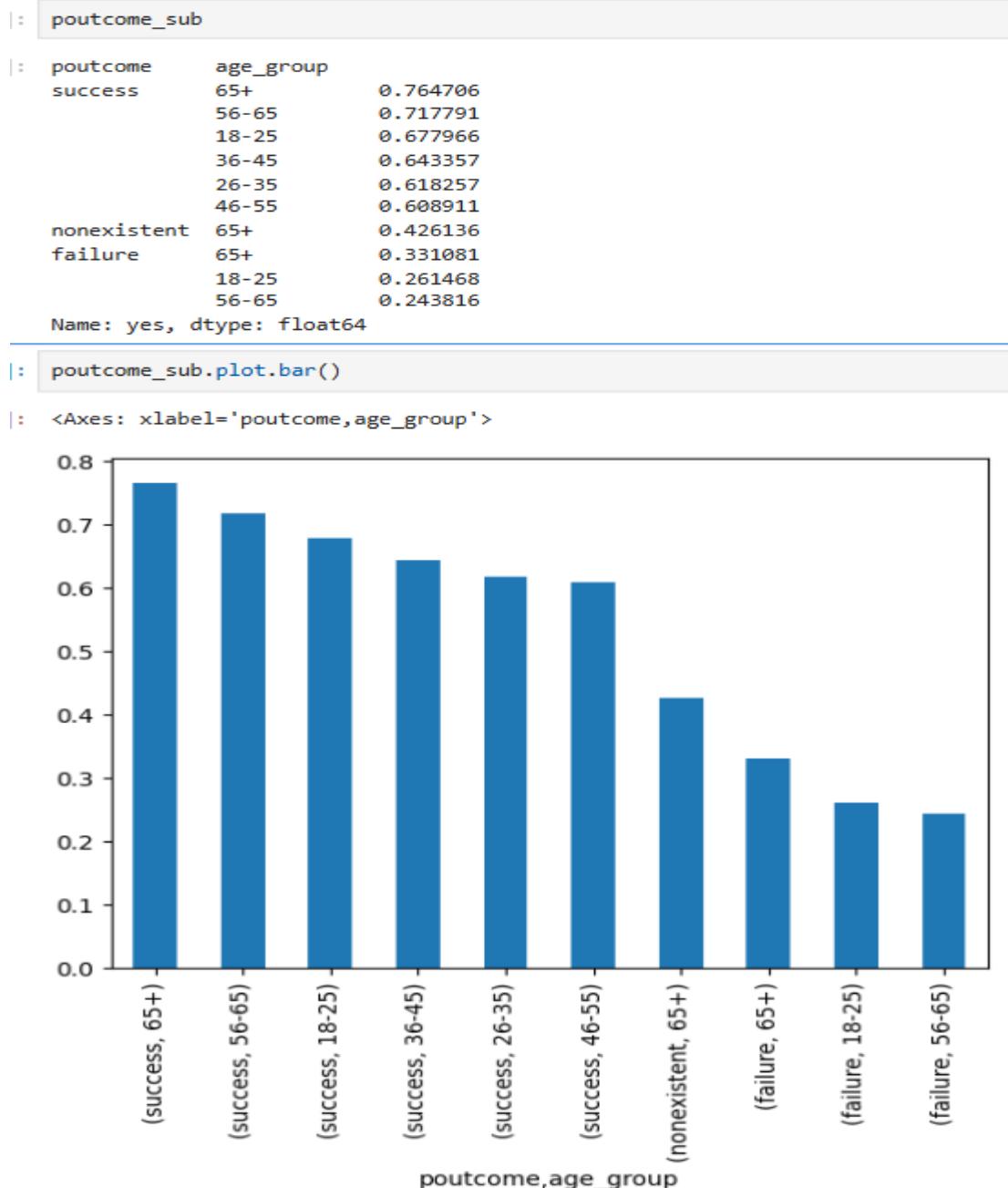


Insight:-

- Multivariate analysis of education and job shows that job role plays a stronger influence than education alone.
- Retired customers demonstrate higher subscription rates across education levels, while students show varying behavior based on educational background.
- Self-employed and blue-collar segments show comparatively lower conversion rates.

Previous Campaign Outcome × Age Group × Subscription

This analysis assesses how **previous campaign outcomes combined with age groups** affect subscription likelihood.



Observation:

- Customers with a **successful previous campaign outcome** show significantly higher subscription rates across all age groups.
- The **impact of success** is especially strong in **middle-age groups (26–45 years)**.
- Customers with failed or non-existent previous outcomes show much lower conversion rates regardless of age group.

Key Insights

Low overall subscription rate Only around 11% of customers subscribed to the term deposit, indicating that customer targeting is critical for improving campaign effectiveness.

Age influences subscription behavior Middle-aged and older customers show a higher likelihood of subscription compared to younger customers, suggesting age-based segmentation can improve results.

Job role is a strong indicator Students and retired customers have significantly higher subscription rates, while blue-collar and service sector customers show lower conversion.

Education level impacts customer decisions Customers with higher education levels or professional backgrounds are more likely to subscribe, reflecting better financial awareness.

Previous campaign outcome is the strongest predictor Customers who subscribed in a previous campaign show a very high conversion rate (~65%), whereas customers with failed or no previous contact have much lower subscription rates.

Campaign intensity affects success Subscription probability decreases as the number of contacts increases, indicating that excessive follow-ups reduce customer interest.

Seasonal patterns exist Higher subscription rates are observed during specific months such as March, September, and December, while some months show consistently low performance.

Multivariate analysis improves targeting precision Combining factors like age group, job, and education reveals specific customer segments with much higher conversion rates than single-variable analysis.

Business Recommendations

Prioritize customers with previous successful outcomes Re-target customers who subscribed in earlier campaigns, as they offer the highest return on investment.

Adopt segmented marketing strategies Focus marketing efforts on high-performing segments such as students, retired customers, and professionally educated individuals.

Optimize campaign contact frequency Limit the number of calls per customer to avoid over-contacting and reduce customer fatigue.

Plan campaigns around high-performing months Allocate more marketing resources during months with higher historical subscription rates and reduce effort during low-performing periods.

Improve engagement for first-time contacts Enhance call scripts, offers, and personalization for customers with no previous campaign exposure to increase initial conversion.