

## Bank Marketing Campaign Analysis

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Load dataset
df = pd.read_csv('bankmarketing.csv')
df.head()
```

↗

	age	job	marital	education	default	housing	loan	contact	month	day_of_week	...	campaign	pdays	previous	poutcome	e
0	56	housemaid	married	basic.4y	no	no	no	telephone	may	mon	...	1	999	0	nonexistent	
1	57	services	married	high.school	unknown	no	no	telephone	may	mon	...	1	999	0	nonexistent	
2	37	services	married	high.school	no	yes	no	telephone	may	mon	...	1	999	0	nonexistent	
3	40	admin.	married	basic.6y	no	no	no	telephone	may	mon	...	1	999	0	nonexistent	
4	56	services	married	high.school	no	no	yes	telephone	may	mon	...	1	999	0	nonexistent	

5 rows × 21 columns

## Data Overview

```
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
```

```
df.isnull().sum()
```



	0
age	0
job	0
marital	0
education	0
default	0
housing	0
loan	0
contact	0
month	0
day_of_week	0
duration	0
campaign	0
pdays	0
previous	0
poutcome	0
emp.var.rate	0
cons.price.idx	0
cons.conf.idx	0
euribor3m	0
nr.employed	0
y	0

dtype: int64

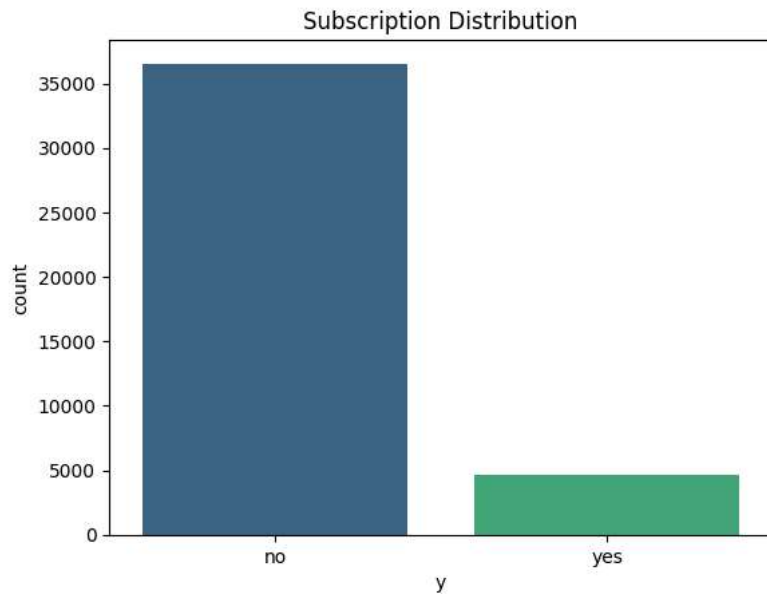
## ▼ Target Variable Distribution

```
sns.countplot(x=df['y'], palette='viridis')  
plt.title('Subscription Distribution')  
plt.show()
```

```
<ipython-input-5-09ad0e52b1ca>:1: FutureWarning:
```

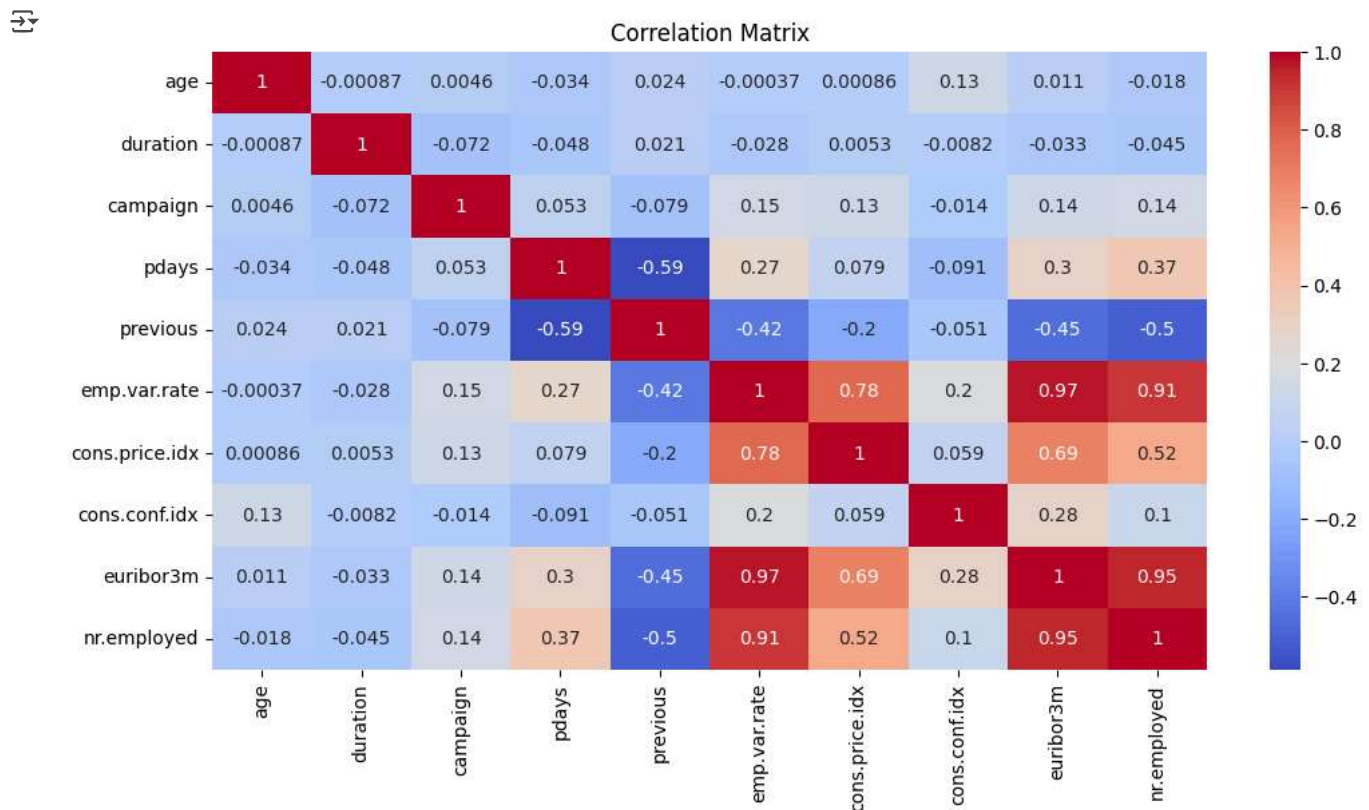
Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend`

```
sns.countplot(x=df['y'], palette='viridis')
```



## Correlation Analysis

```
plt.figure(figsize=(12,6))
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```



## Job Type Impact on Subscription

```
plt.figure(figsize=(12,6))
sns.countplot(x='job', hue='y', data=df, palette='viridis', order=df['job'].value_counts().index)
plt.xticks(rotation=45)
plt.title('Subscription by Job Type')
plt.show()
```

