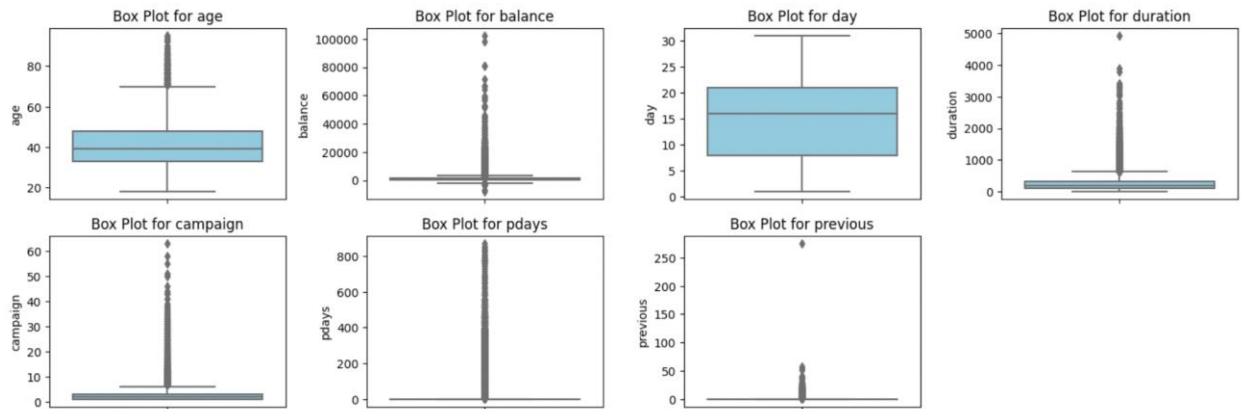


APPENDIX OF FIGURES WITH CAPTIONS

1. Outliers Checking for Numeric Columns

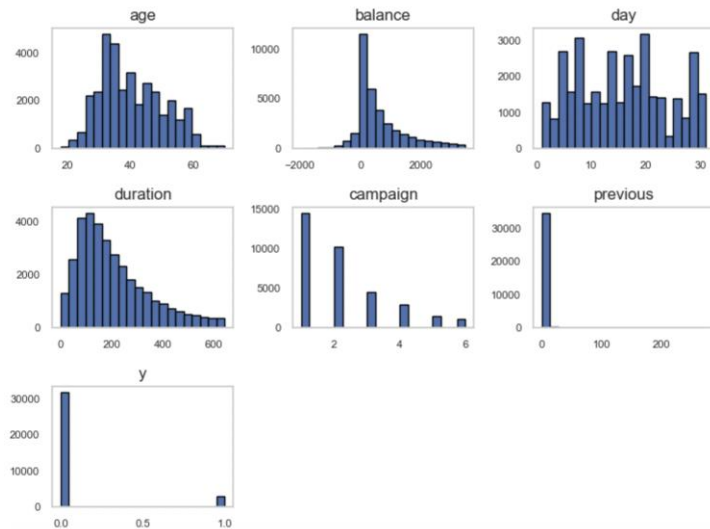


Insight:

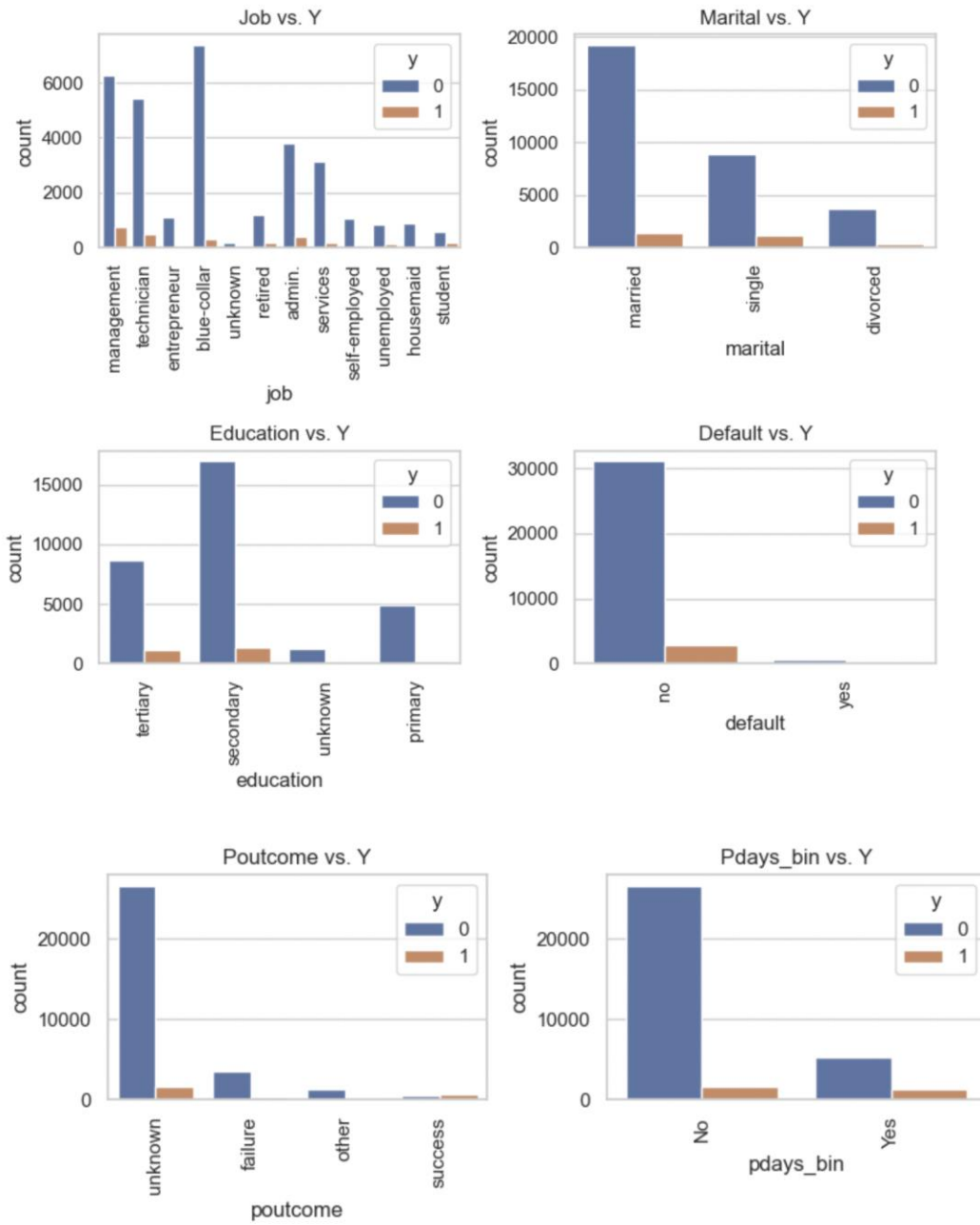
The columns 'age,' 'duration,' and 'campaign' contain a considerable number of outliers, with 'duration' in particular showcasing a prominent distribution of outliers. To remove outliers in these columns, the Interquartile Range (IQR) method was applied.

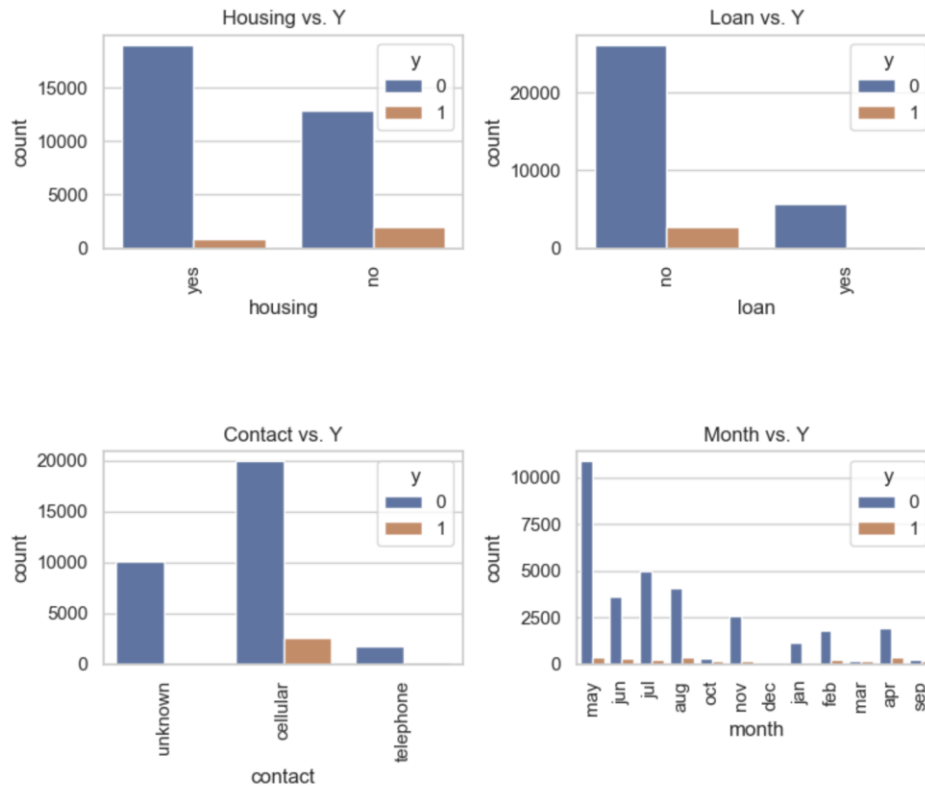
The 'balance' column exhibits a substantial number of outliers. However, in the context of banking and financial data, these values might be normal if there are customers with exceptionally high account balances. Preserving this information could be vital to retain crucial insights about customers.

2. Distribution of numeric features

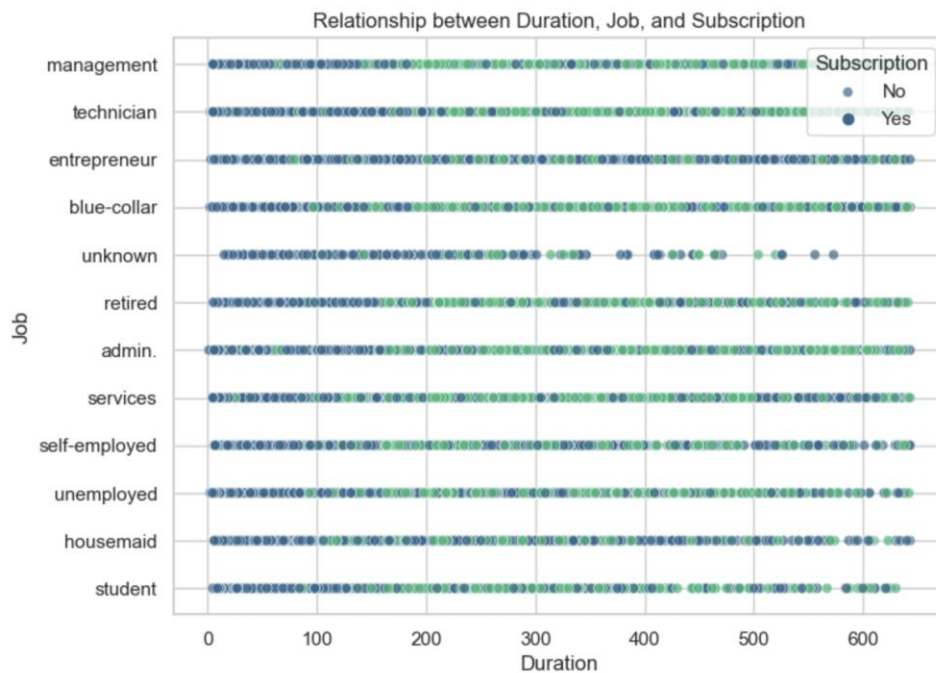


3. Categorical Variables Impact on Subscription Registration





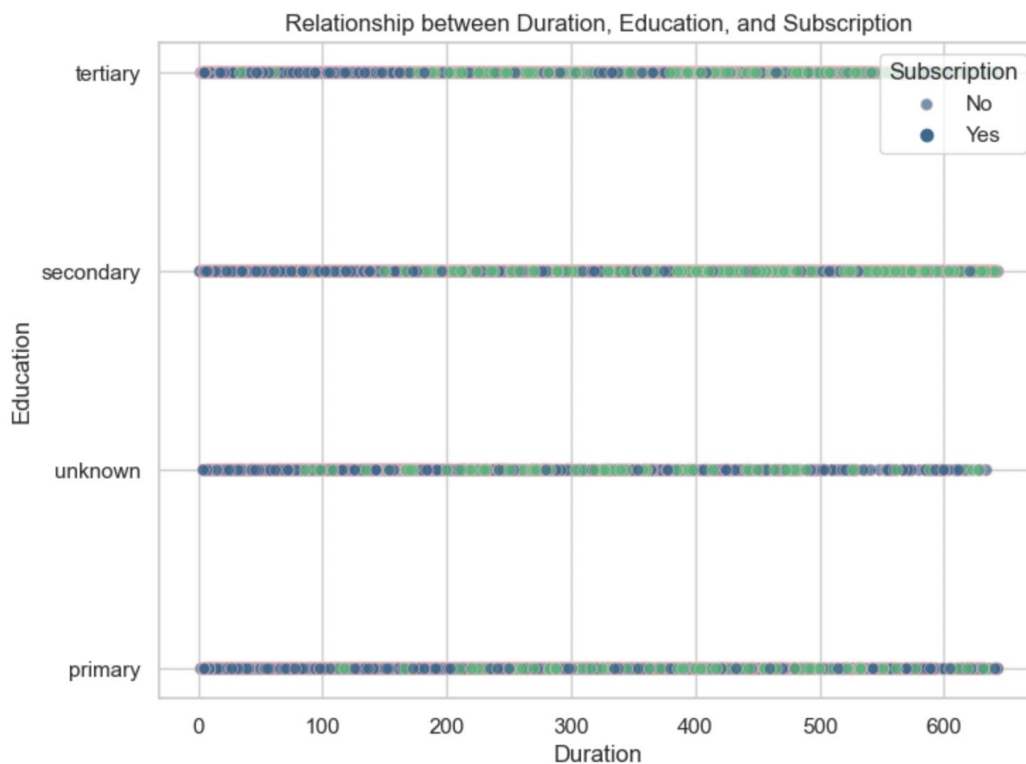
4. Relationship between Duration, Job, and Y Value



Insight:

- There is a significant variation in the duration of calls among different occupational groups when considering subscriptions and non-subscriptions.
- Occupational groups such as 'blue-collar,' 'entrepreneur,' and 'services' tend to have longer call durations when there is a subscription, whereas the 'student' group has the shortest duration.
- The 'unknown' group experiences a substantial decrease in call duration when there is a subscription, possibly due to a small sample size.

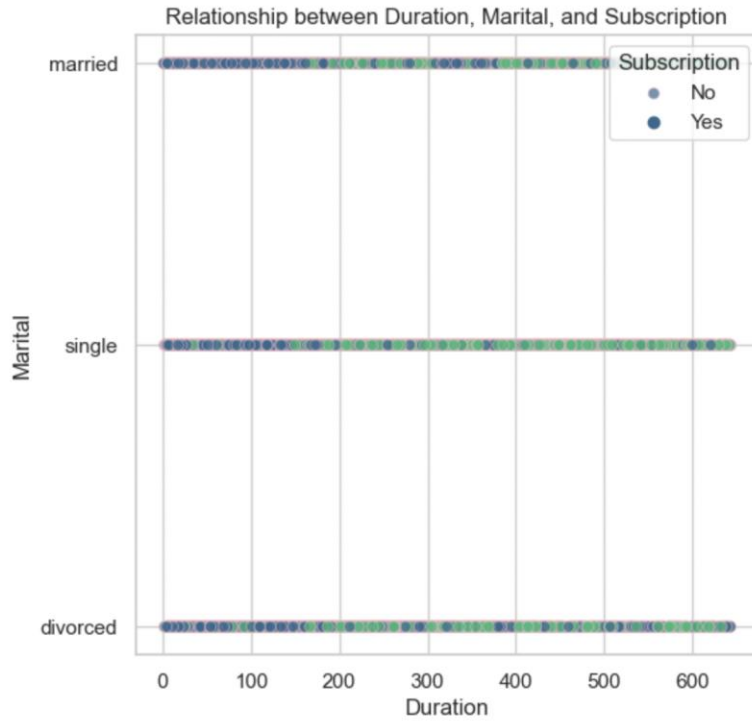
5. Relationship between Duration, Education, and Y Value¶



Insight:

- In general, there seems to be a trend that individuals with higher education levels may have longer call durations when subscribing.

6. Relationship between Duration, Marital, and Y Value¶



Insight

- It seems that the divorced and married groups tend to have longer call durations when subscribing, and the single group also shows a significant increase in call durations.