



Vanguard

INSIGHTS

by

HU-ME

KARL TOTO SAME
IKRAM MOHAMED AGUID
HALYNA ABELCHAKOVA



What's going on?

Vanguard is an investment management company based in US.

For about 3 months we were testing a new UI, which is more intuitive and modern, to find if it could make the online process smoother for clients.

And now we are going to answer the main question:

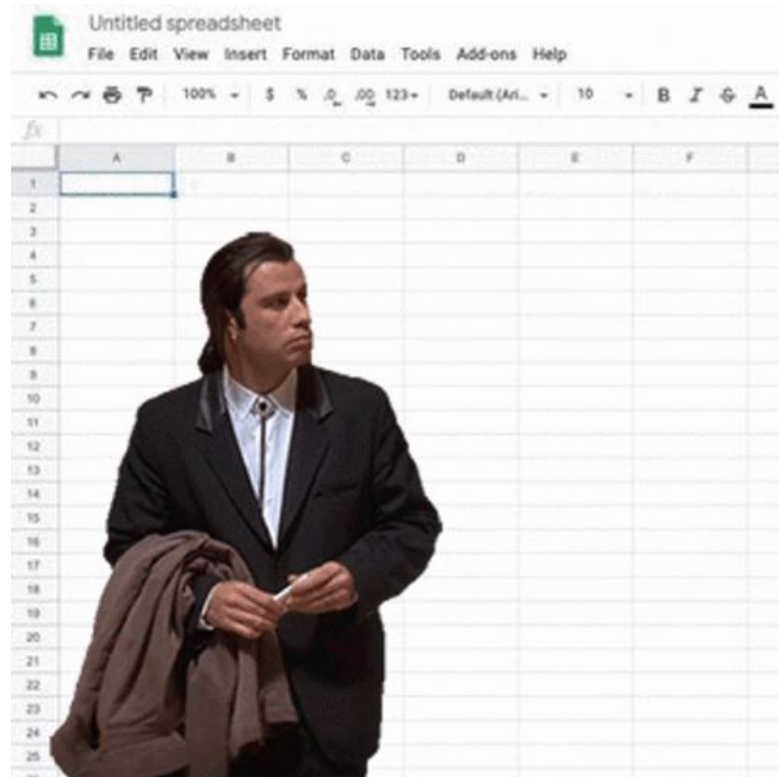
**Would these changes encourage more clients
to complete the process?**



What did we work with?

We used 3 datasets:

- **Client Profiles** (Demographics)
- **Digital Footprints** (online interactions)
- **Experiment Roster:** (list of clients)



What has been done?

EDA and DATA CLEANING



Merging files:

- Concatenation of the two "data footprints" files in Python
- Merge the "data footprint" and "experiment roster" file to have the "variation" column which tells us the group to which the customer belongs

Cleaning data:

- Data standardization (categorical columns)
- Replacing empty values with standard values for all column types

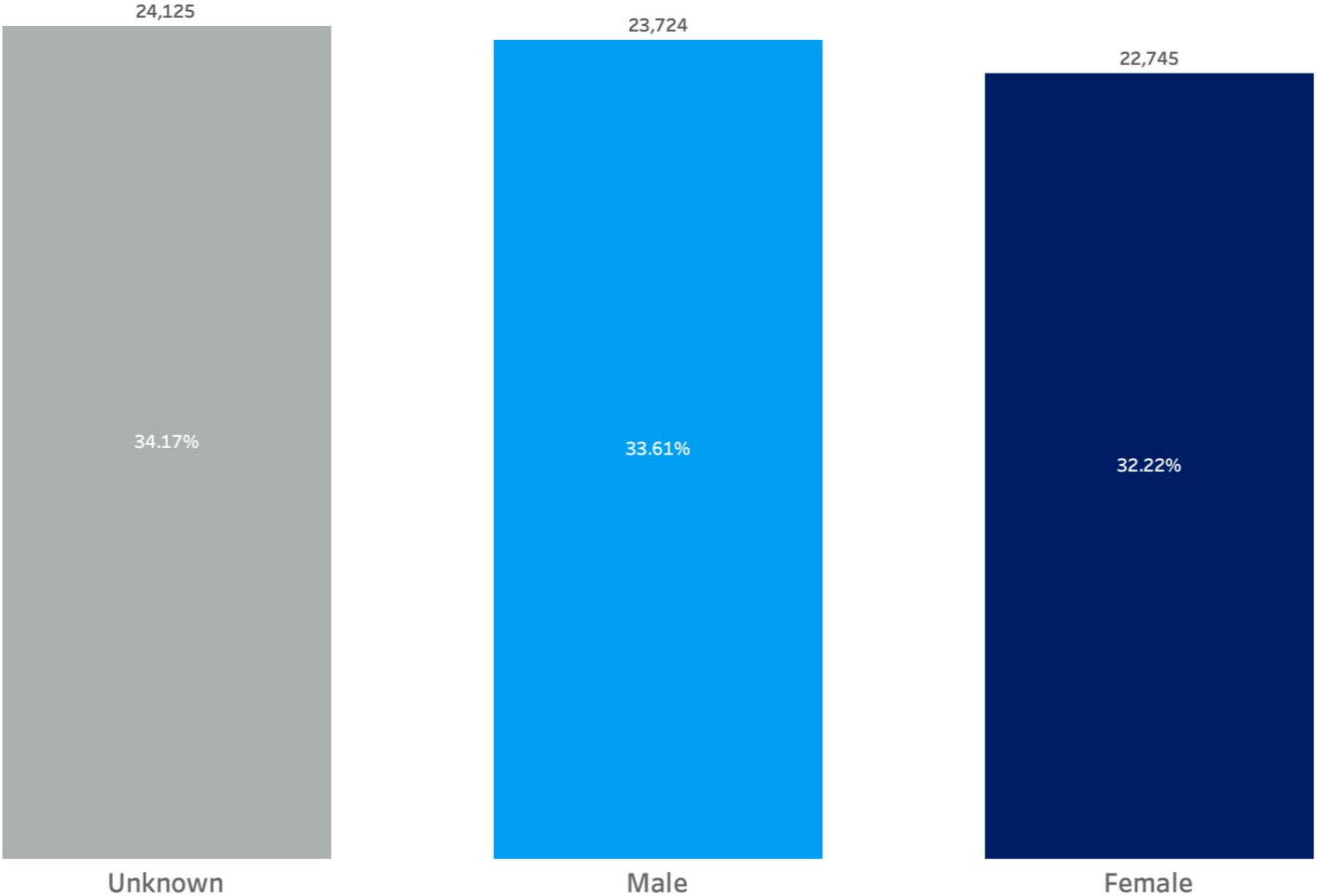
Exploratory Data Analysis

Key demographics of Vanguard's online clients.

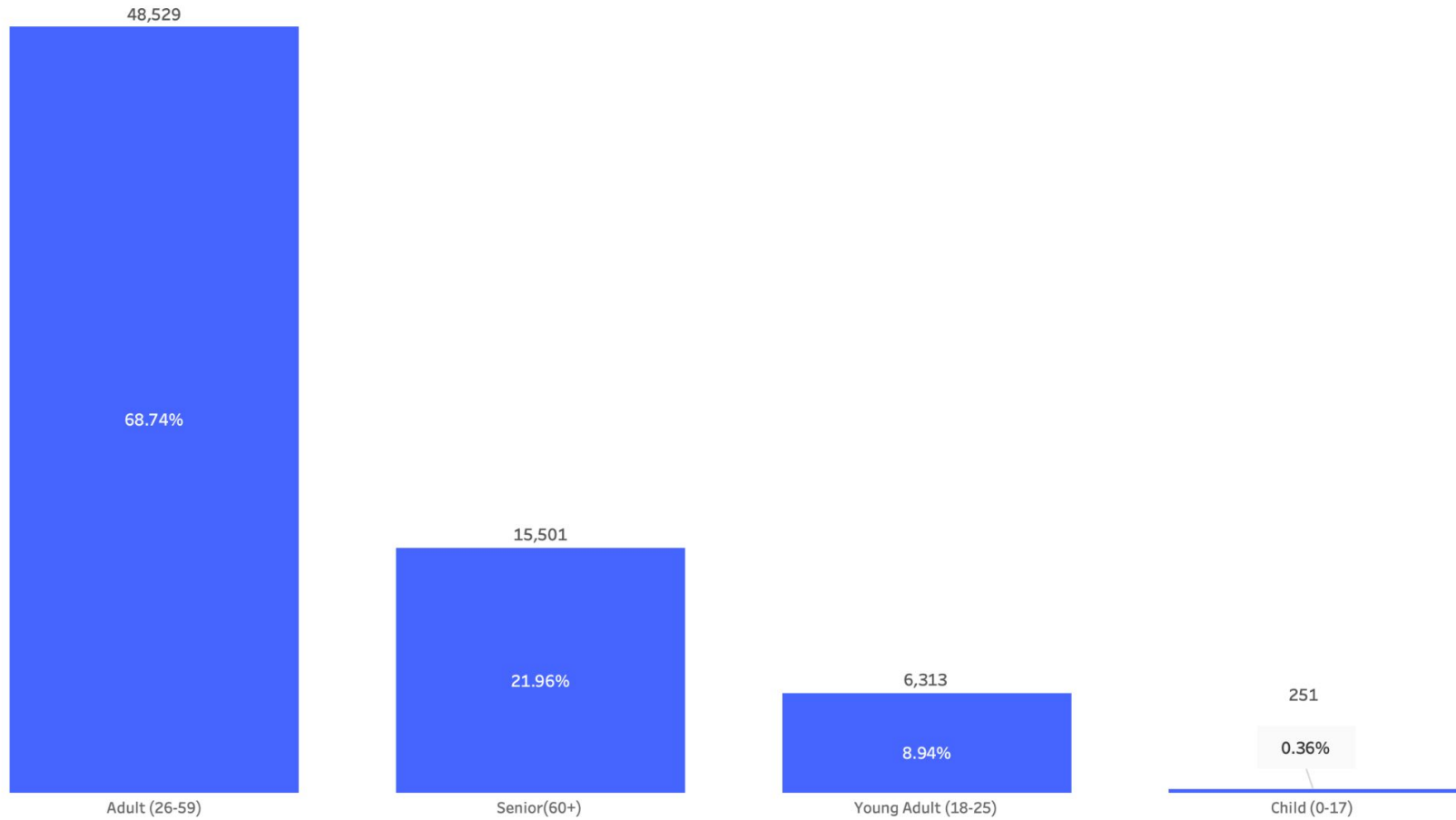
- Gender proportion
- Age distribution
- Tenure distribution



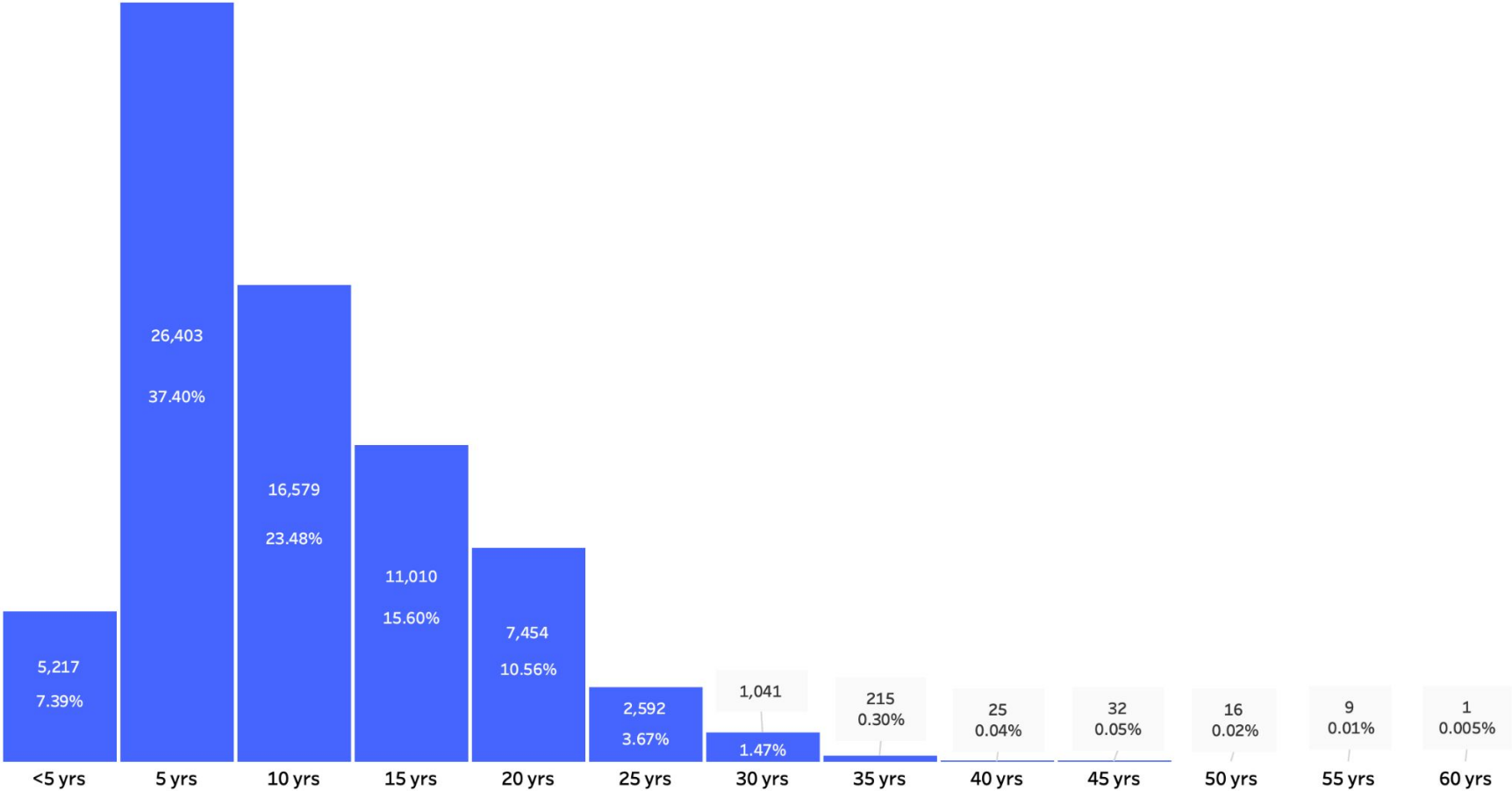
Gender Distribution



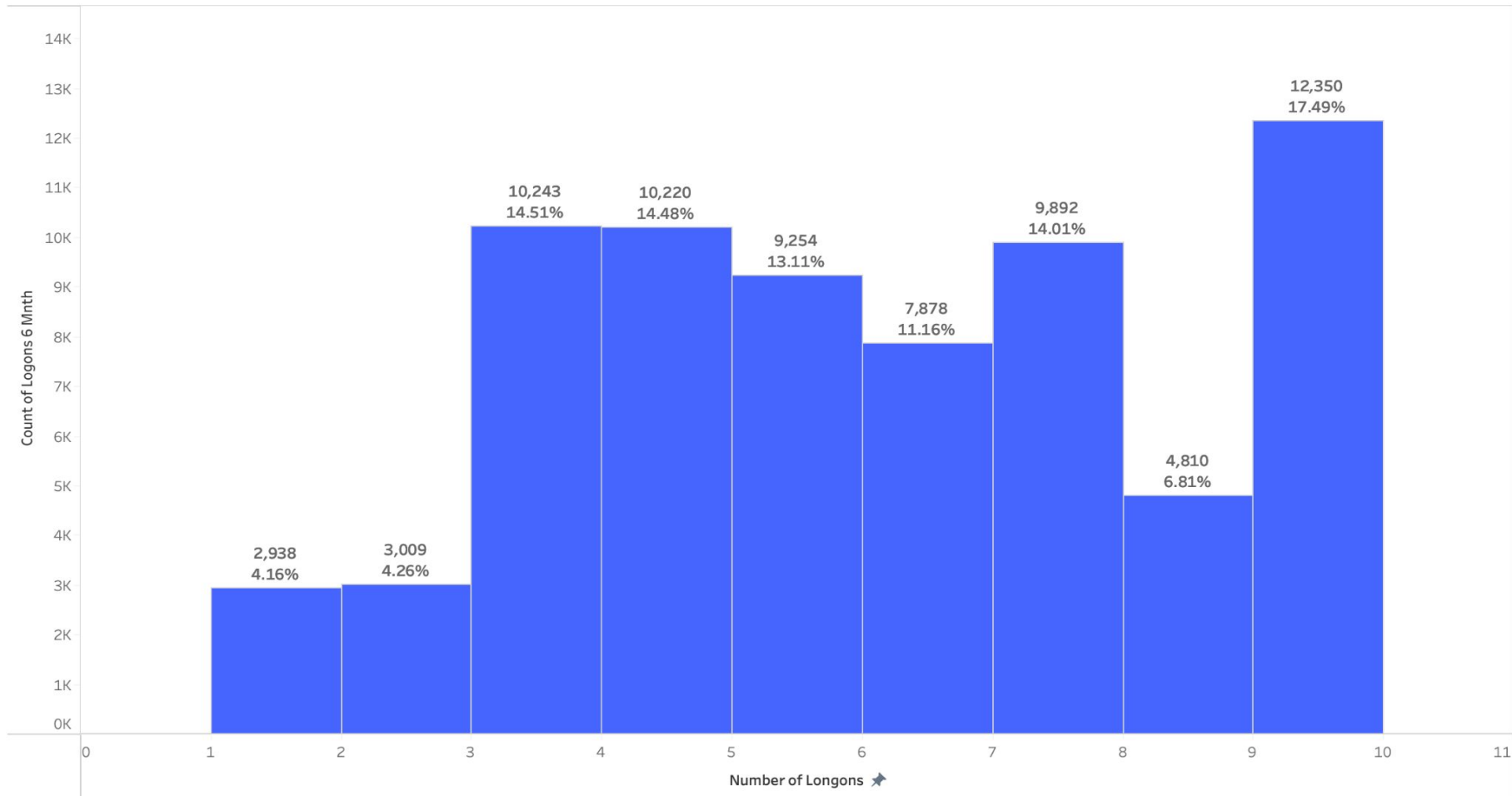
Age Category Distribution



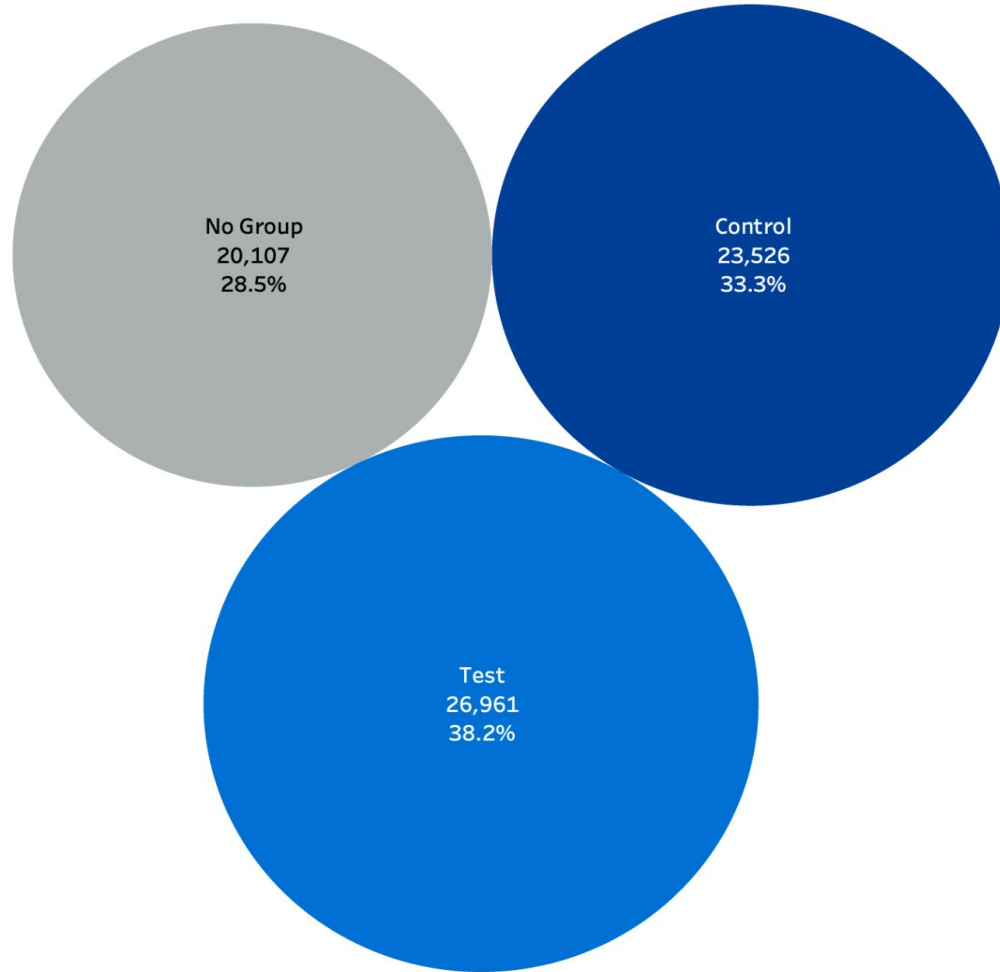
Tenure Distribution



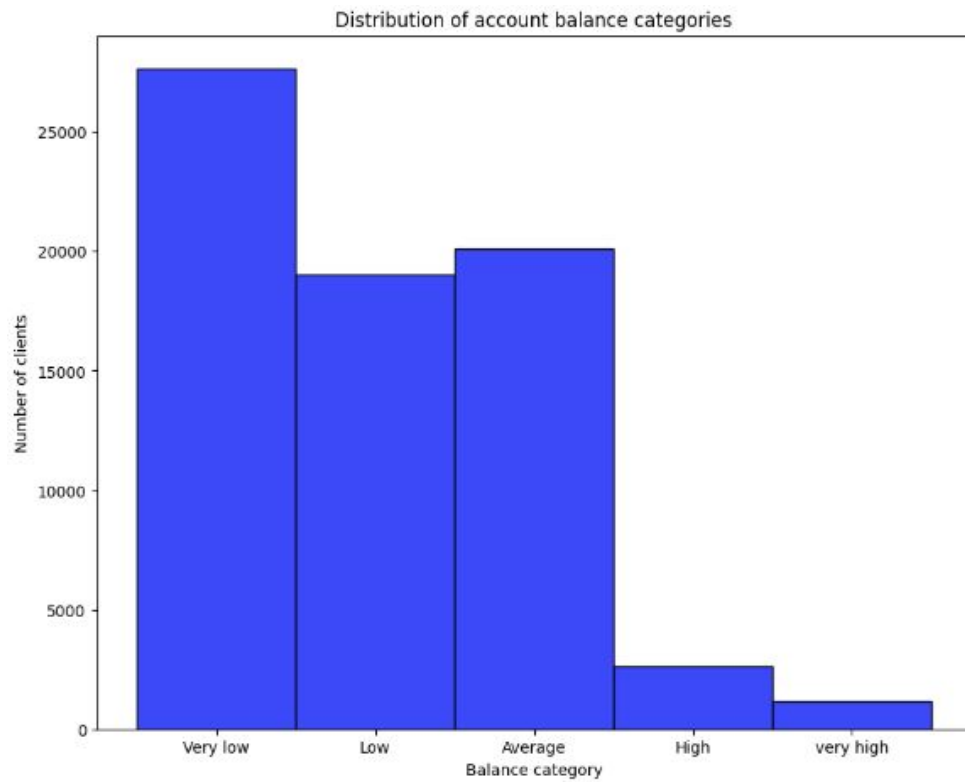
Count of Logons in the last 6 months



Group Distribution



Account balance

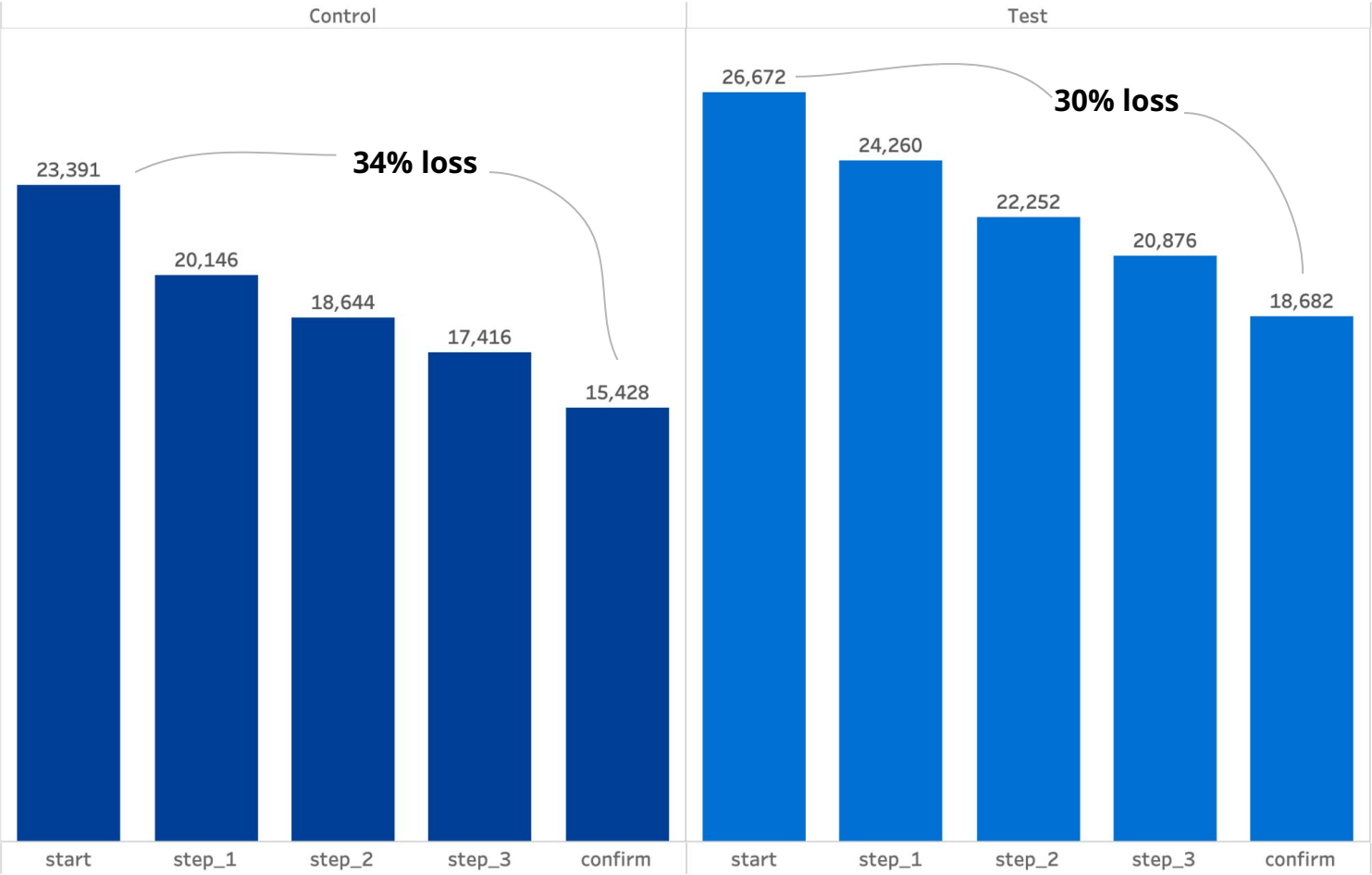


Performance Metrics

- **Completion Rate:** The proportion of users who reach the final 'confirm' step.
- **Time Spent on Each Step:** The average duration users spend on each step.
- **Error Rates.**



Step Completion Count



Completion Rate by Group

Completion Rate TEST

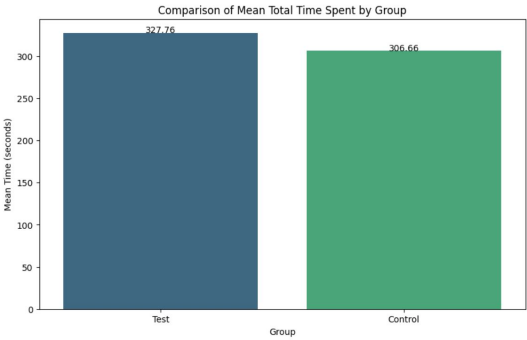
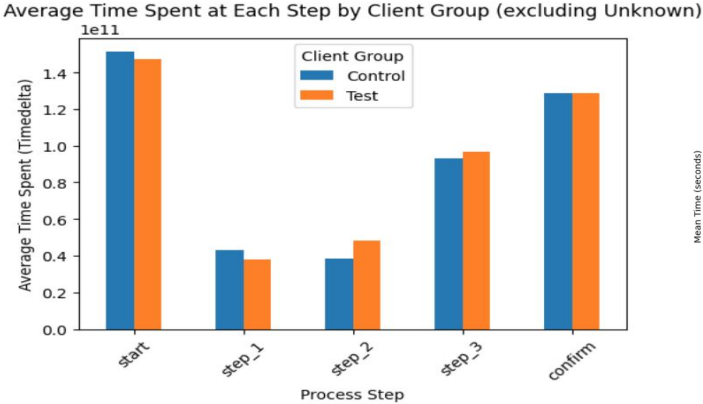
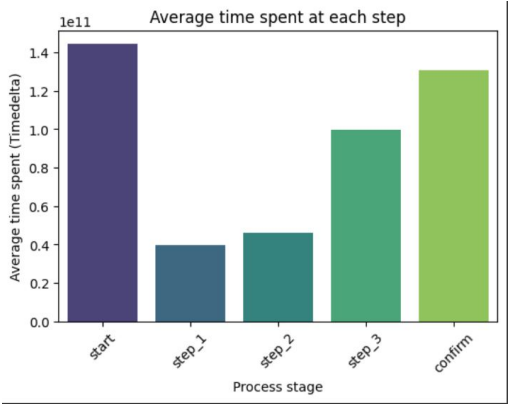
69.29%

Completion Rate
CONTROL

65.59%



Average time spent



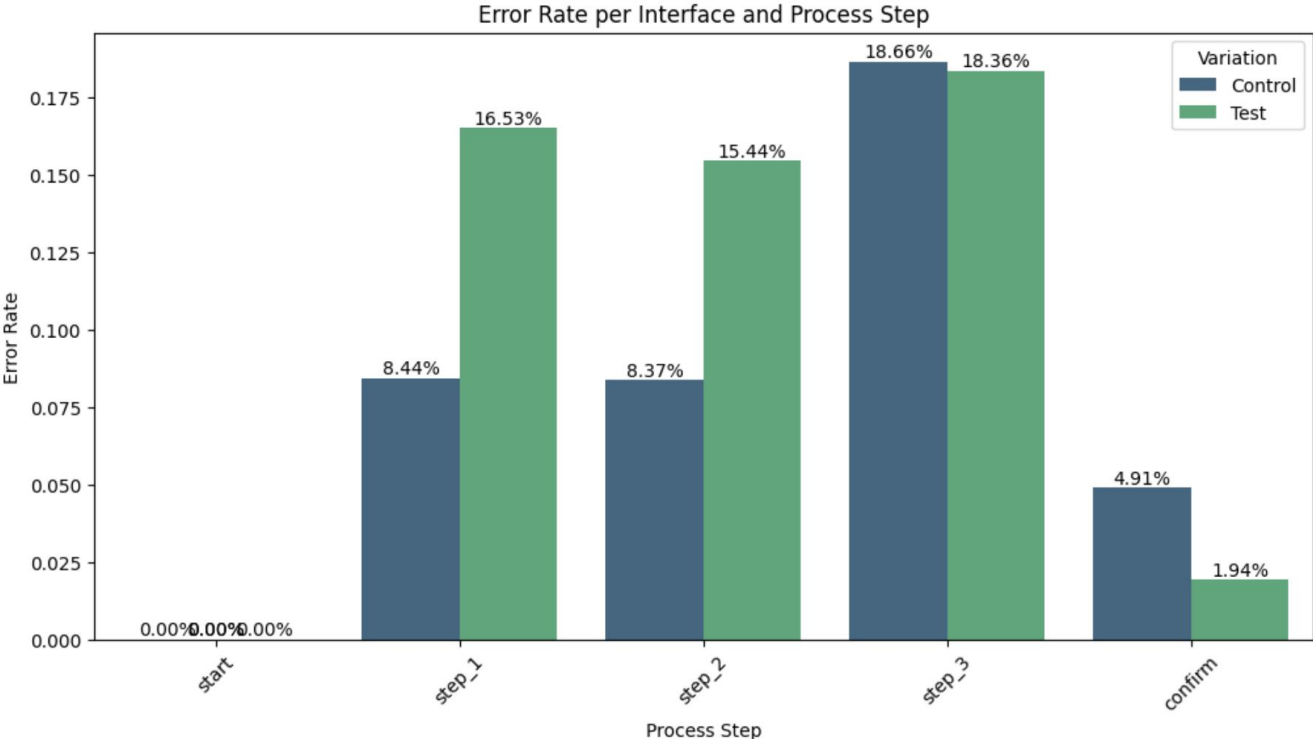
```
process step
start      0 days 00:02:29.762495764
step 1     0 days 00:00:40.272115070
step 2     0 days 00:00:43.590126045
step 3     0 days 00:01:35.397469852
confirm    0 days 00:02:08.396520213
```

```
process step      start      step_1
Variation
Control           0 days 00:02:31.366025986 0 days 00:00:42.998740768
Test              0 days 00:02:27.140325404 0 days 00:00:37.704369310
```

```
process step      step_2      step_3
Variation
Control           0 days 00:00:38.667856588 0 days 00:01:32.944257802
Test              0 days 00:00:48.131611395 0 days 00:01:36.850971669
```

```
process step      confirm
Variation
Control           0 days 00:02:08.664715369
Test              0 days 00:02:08.769657890
```

Error rate



Hypothesis Testing



Completion rate: There is a significant difference between the completion rates of the Test and Control groups.

- Total/ final step '**Test**' group: 26968/ 18687
Proportion: 69.29%
- Total/ final step '**Control**' clients: 23532/ 15434
Proportion: 65.59%

Difference in completion rate: 3.71%

Is the difference in completion rate profitable (threshold at 5%)?
No

Spent time: There is significant difference in total time spent between Test and Control groups.

Mean total time for Test group: 327.7630329878965 seconds
Mean total time for Control group: 306.66207538846135 seconds

Error rate: There is no significant difference between the error rates of the Test and Control groups.

Experiment Evaluation

- Evaluate the experiment design, discussing the **randomization, duration, and potential biases**.
- Suggest what additional data could have been **beneficial**.



Experiment Evaluation

- Age:

T-Statistic: -2.4161, P-Value: 0.0157

Conclusion: Reject the null hypothesis. **There is a significant difference.**

- Balances:

T-Statistic: -0.4408, P-Value: 0.6594

Conclusion: Fail to reject the null hypothesis. There is no significant difference.

- Number of accounts:

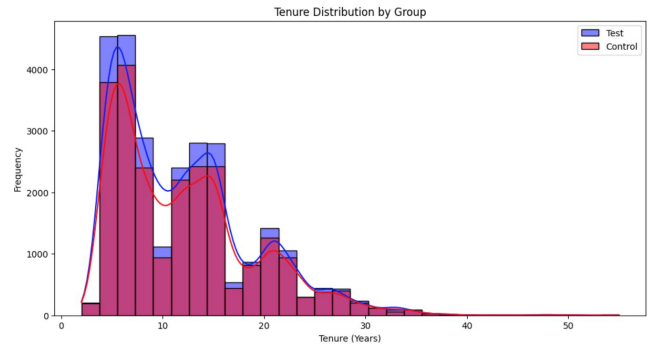
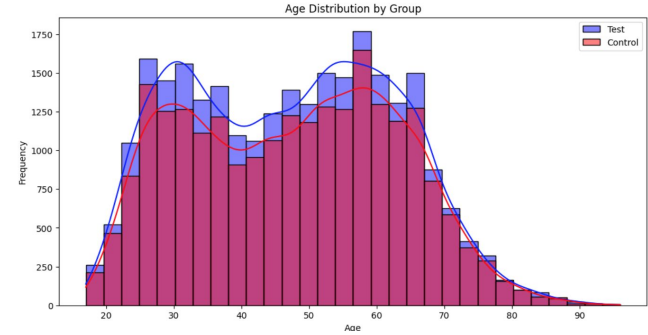
T-Statistic: -2.0974, P-Value: 0.0360

Conclusion: Reject the null hypothesis. **There is a significant difference.**

- Tenure:

T-Statistic: -1.7115, P-Value: 0.0870

Conclusion: Fail to reject the null hypothesis. There is no significant difference.



Teamwork & Project Management

**When you're the only
one working on the
group project**



- We used Trello to divide our tasks;
- To make the process more effective we assigned each task to a specific person with details and deadlines.

Challenges & Learnings

- Lack of time
- Organization



Conclusion

Client behavior: There is a significant difference between Test and Control for the completion rate, the spent time, but not the error rate

Client profile: There is a significant difference between test and control for the age and the number of accounts, but not the balances and tenure

Recommendations:

- focus on improving certain parts of the interface (to improve Start and Confirmation steps, because these steps took the most time for clients)
- further randomize clients for testing

Based on the hypothesis testing, we conclude that the difference of **3.71%** percent doesn't justify the expenses.

