

# USER EXPERIENCE

Analysis of an AB test from **Vanguard®**

Carlota Gordillo | Celia Manzano | Laura Sánchez

# PROJECT INDEX

Data Exploration & Cleaning

Hypothesis Testing

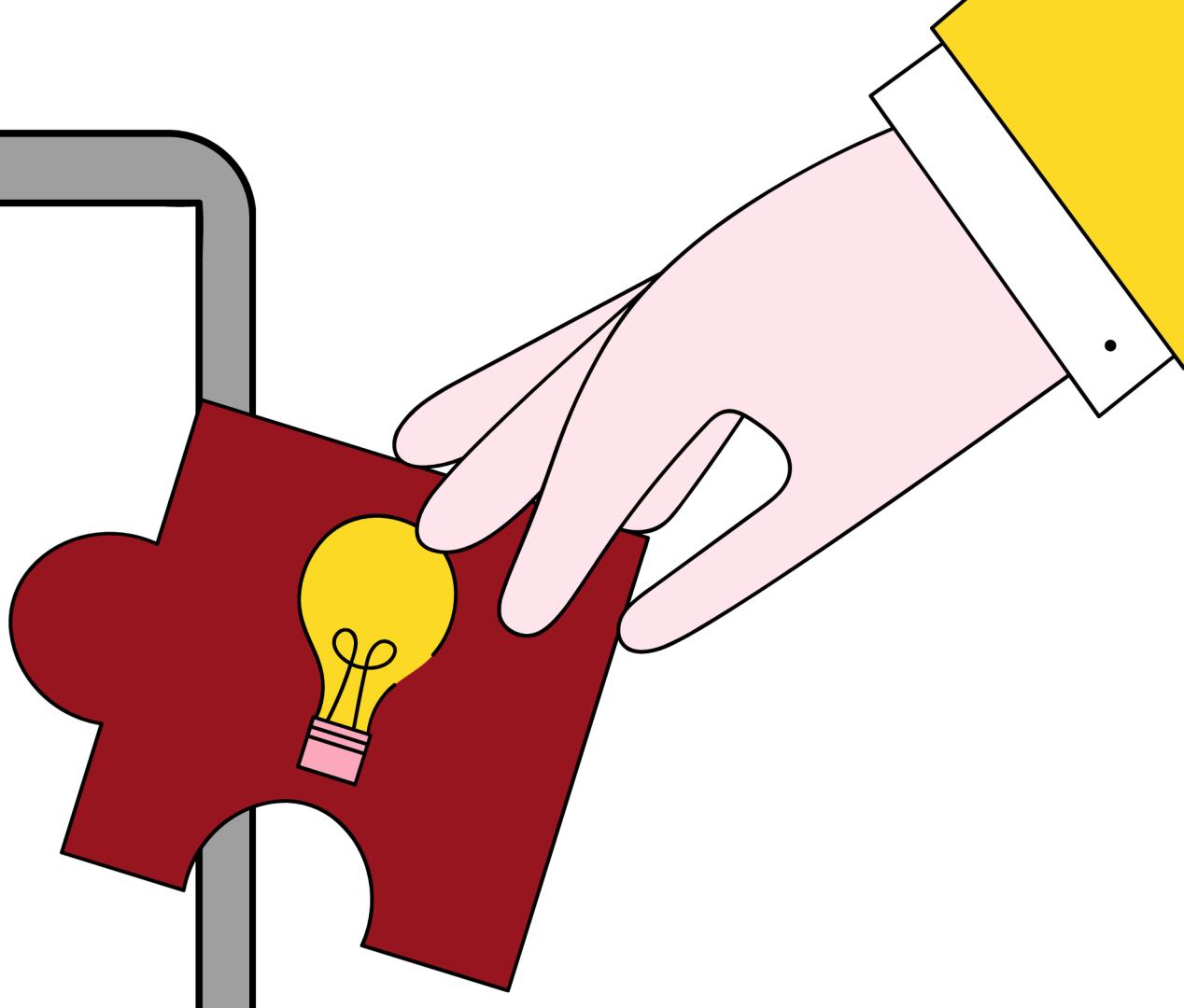
Exploratory Data Analysis

Experiment Evaluation

Performance Metrics

Conclusion

Power BI Visualizations

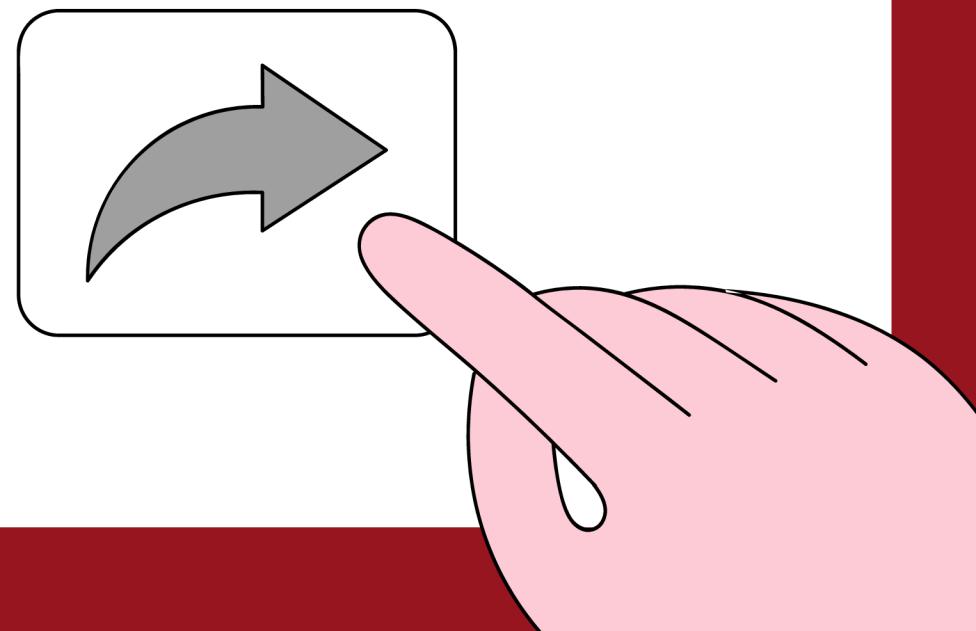


# Introduction-Vanguard Context

## Vanguard Context

About Vanguard: A trusted investment management company focused on innovation and client satisfaction.

The Challenge: Vanguard introduced a modernized User Interface (UI) with intuitive design and in-context prompts to streamline the online client process

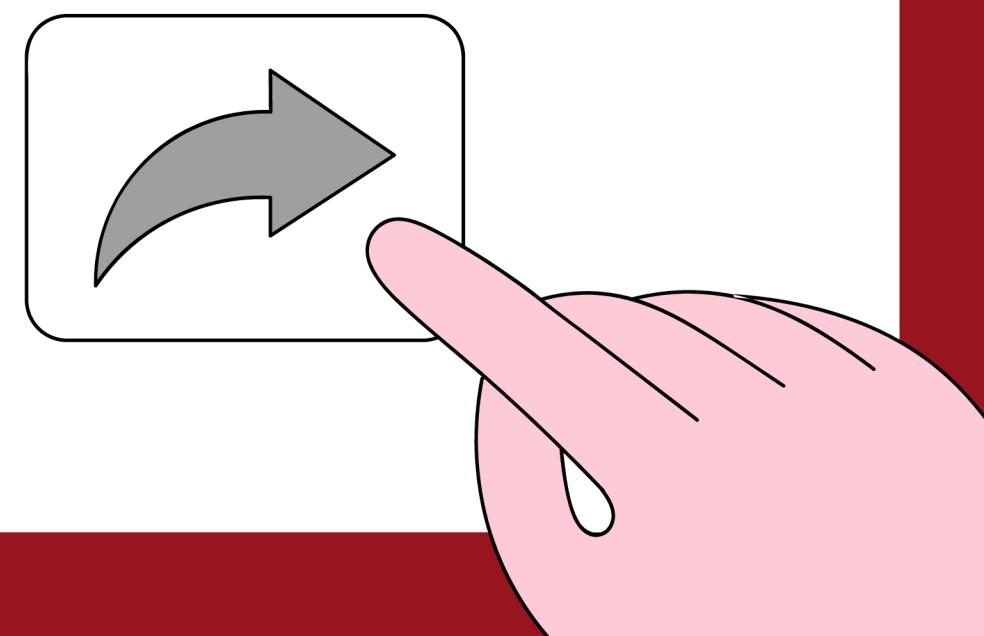


# Introduction-Vanguard Context

## Key Details of the experiment design

- Experiment Period: March 15, 2017 – June 20, 2017.
- Groups:
  - Control Group: Used the traditional process.
  - Test Group: Used the improved UI.
- Customer journey:

**Start → 3 steps → Confirm**



# Introduction-Vanguard Context

Key question:

**Did the new UI improve  
previous design?**

# Data Exploration & Cleaning



**Client Profiles**

**Digital Footprints**

**Experiment Roster**

Merged all info



**Final Customer**

Discretize numerical columns into  
categorical bins and clean the  
dataset by handling duplicates  
and null values.



# Clients Demographics

## Insights

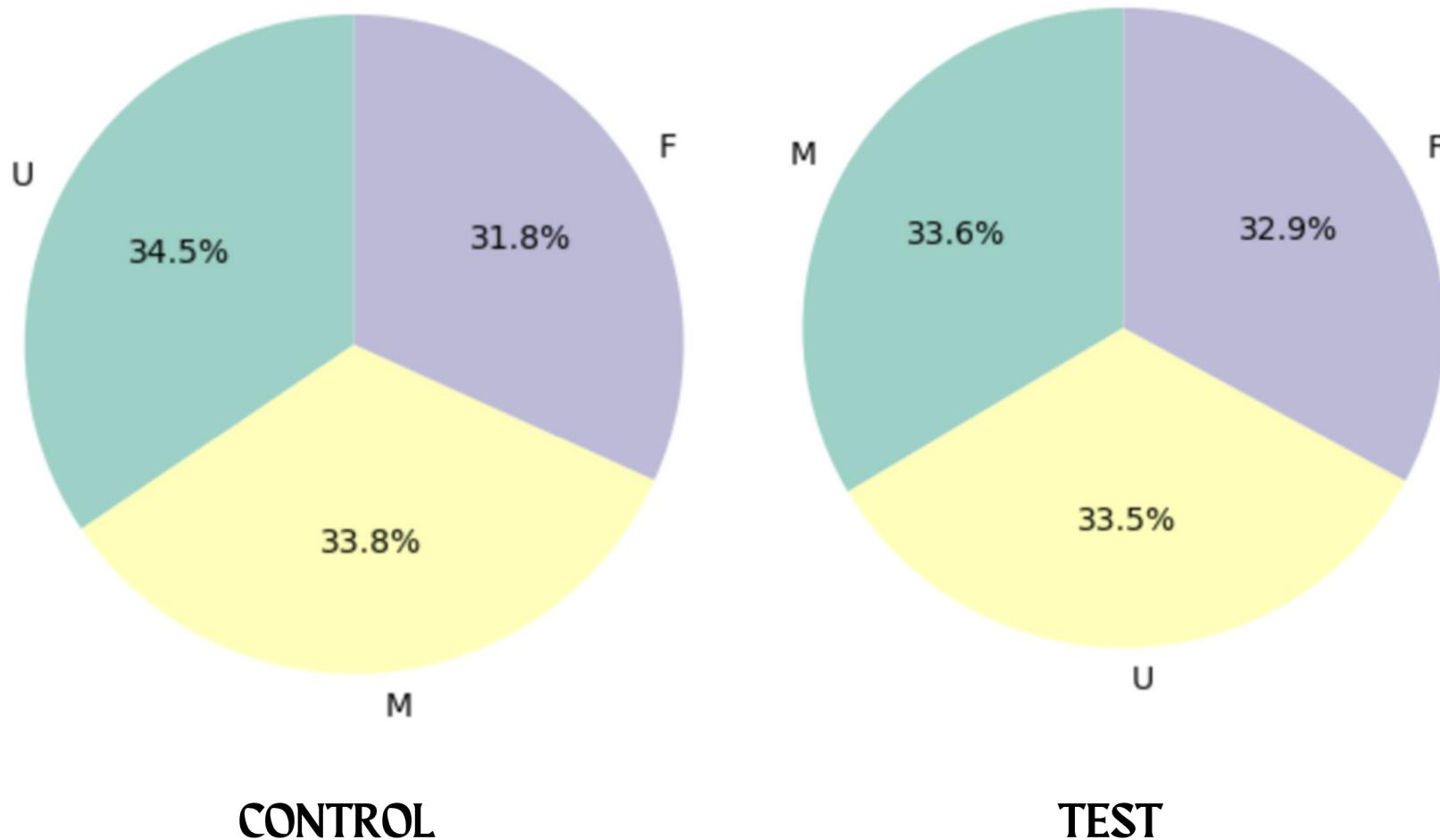
- Control sample size: **23,526**
- Test sample size: **26,961**
- The most common number of interactions is between **2 and 6** for both groups.
- The groups appear to be **fairly proportional**, except for the difference in the number of participants.



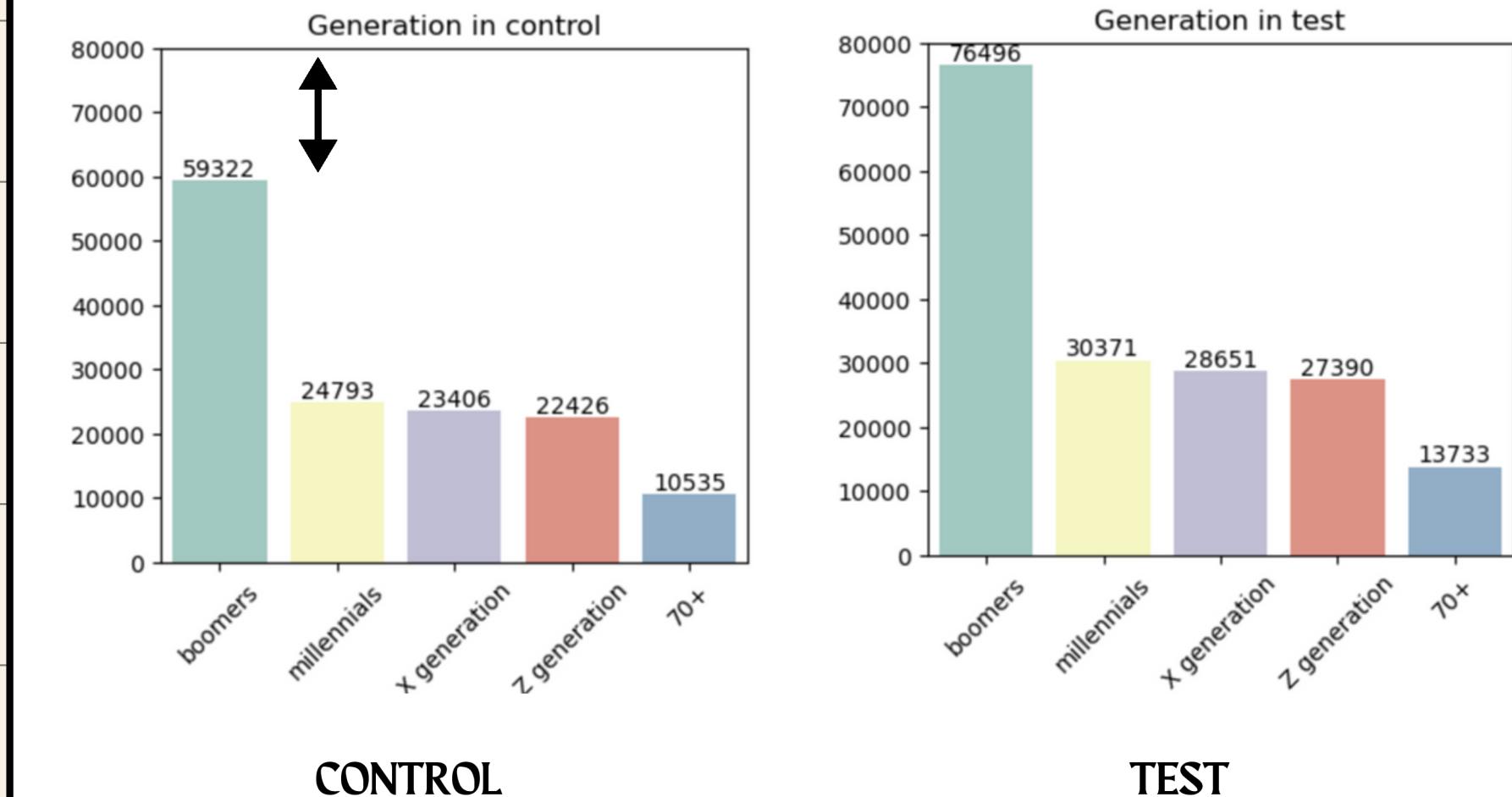
Sample size should be okay to detect a significance difference

# Clients Demographics

## Gender



## Age

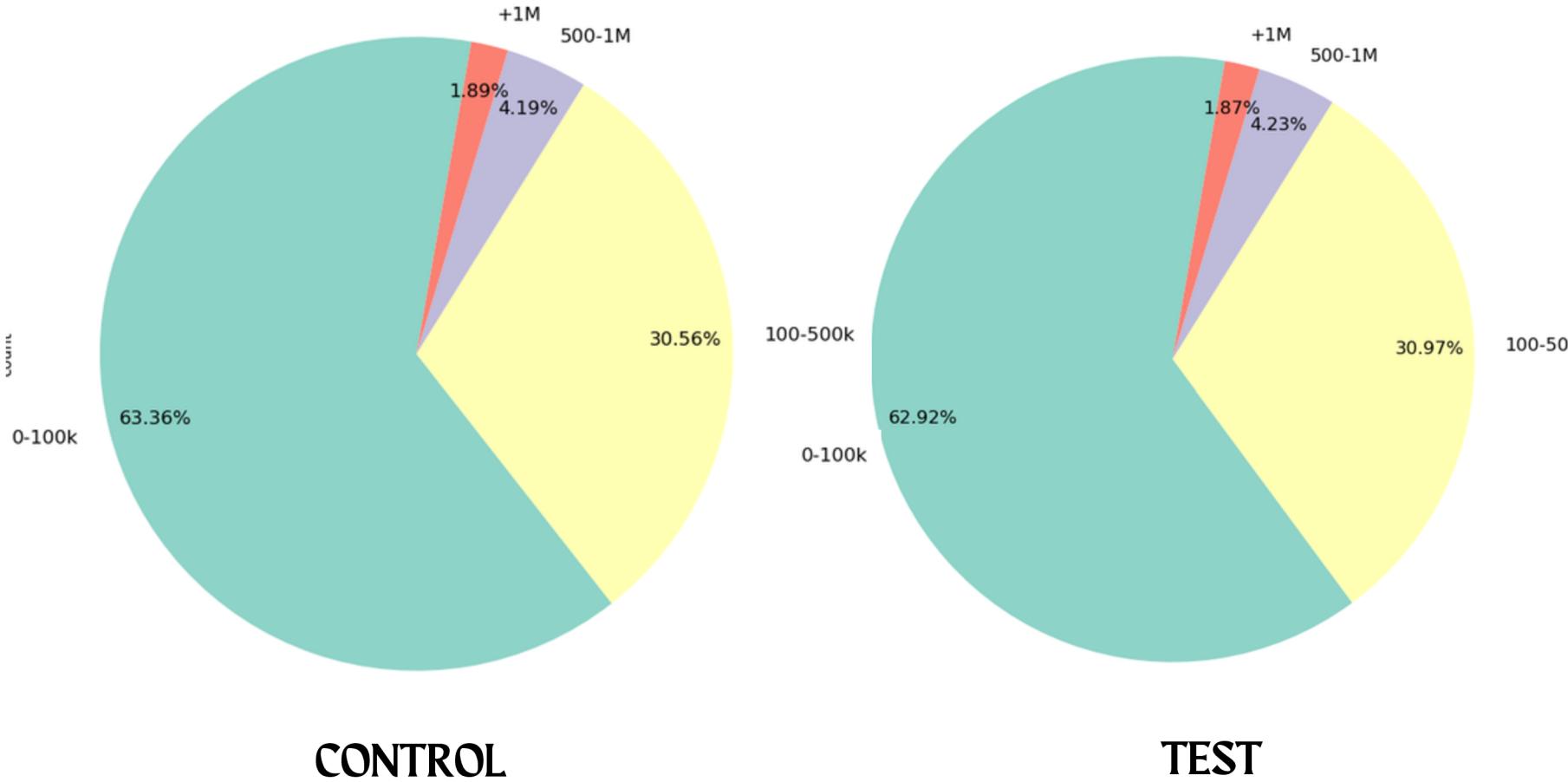


There are no differences between the gender in both groups.

There are differences between the generation in both groups.

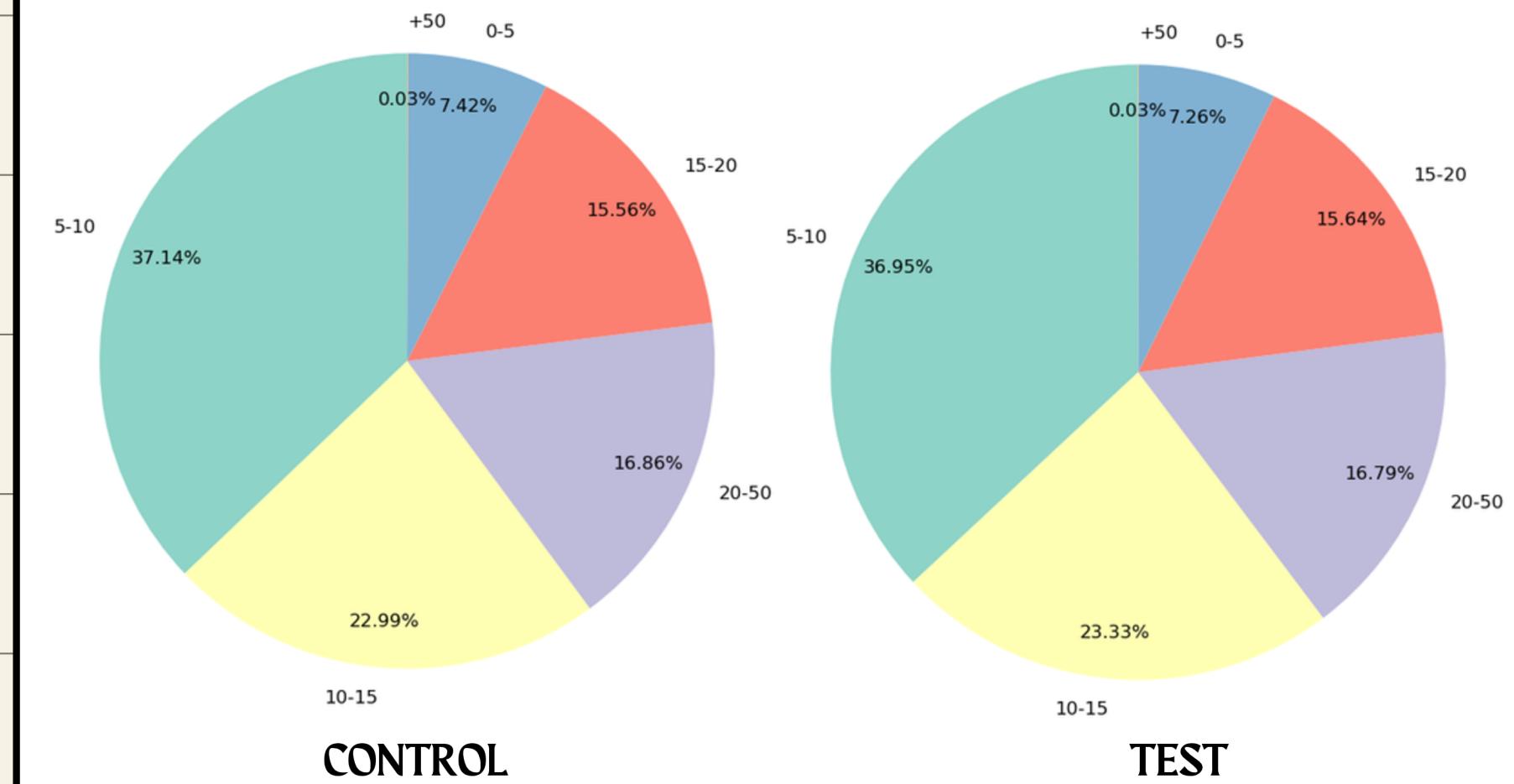
# Clients Demographics

## Balance



There are no differences between the balance in both groups.

## Tenure



There are no differences between the tenure in both groups.

# KPI'S

01

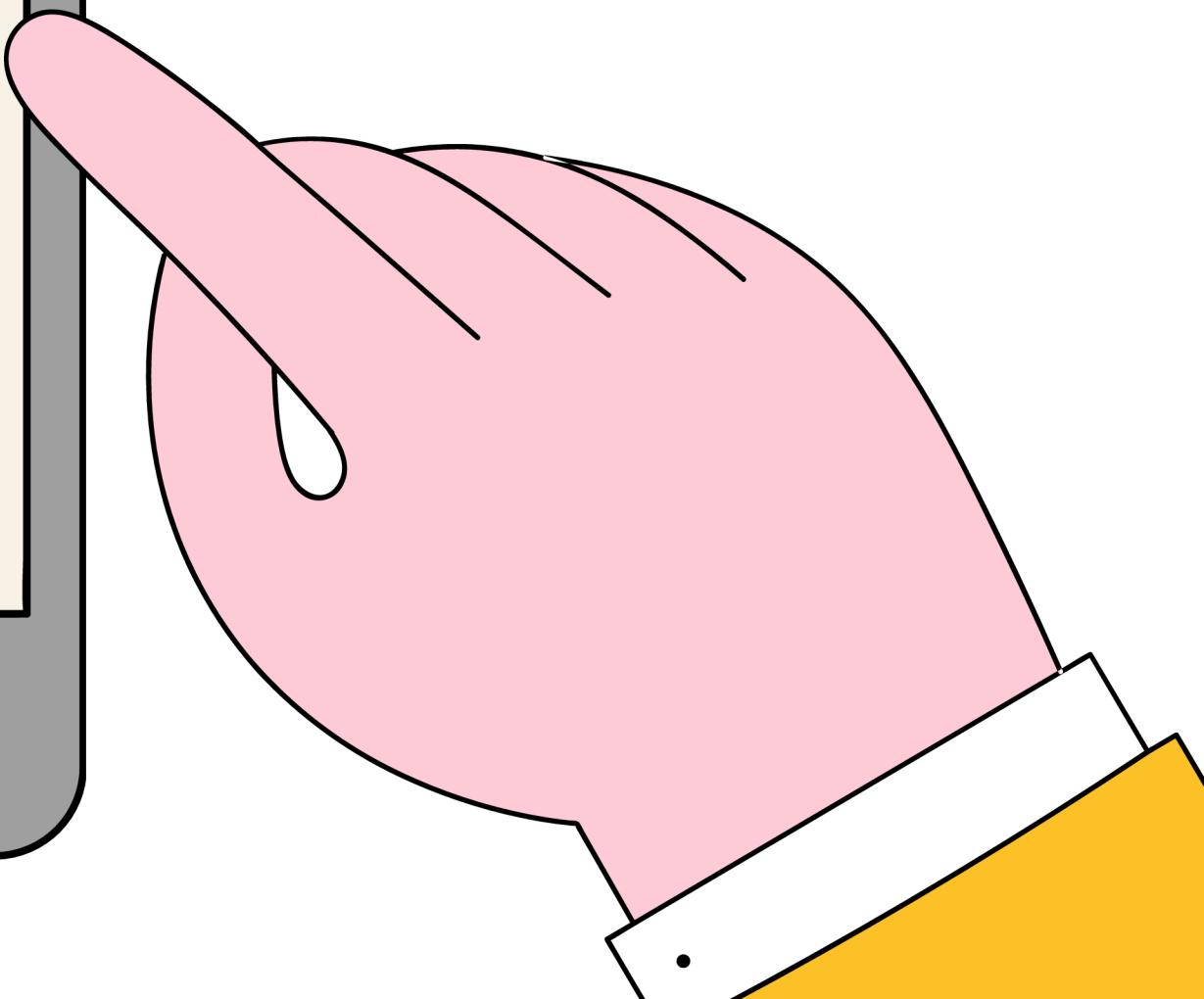
**Completion Rate:** The proportion of users who reach the final *confirm* step

03

**Time Spent between steps:** The average duration users spend on each step.

02

**Error Rates:** Consider moving from a later step to an earlier one as an error.



# Completion Rate

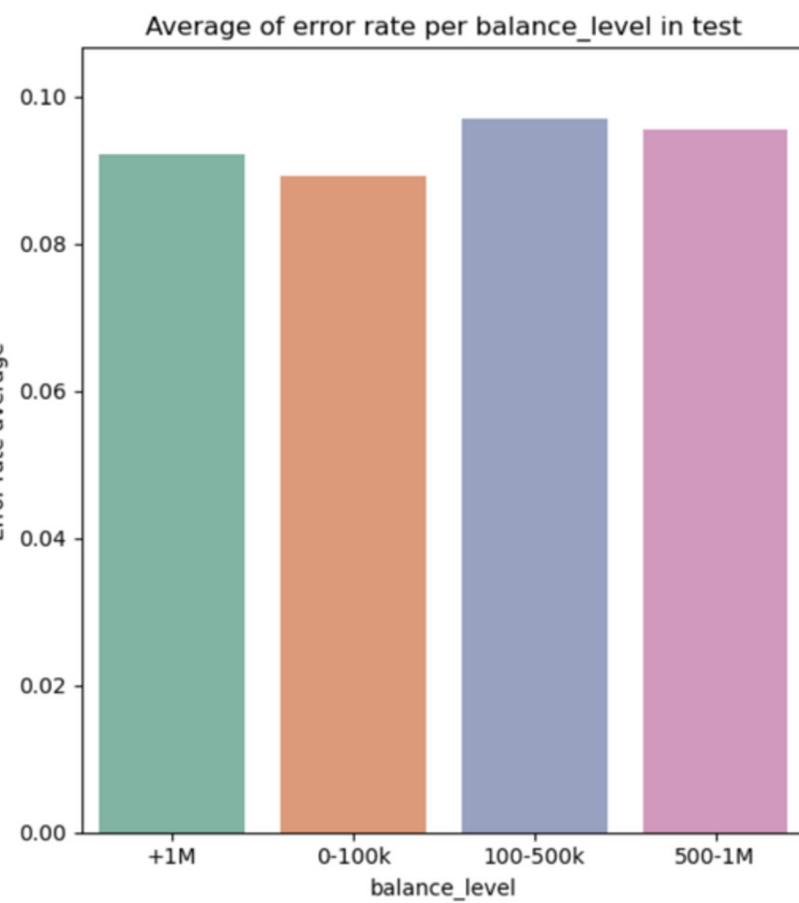
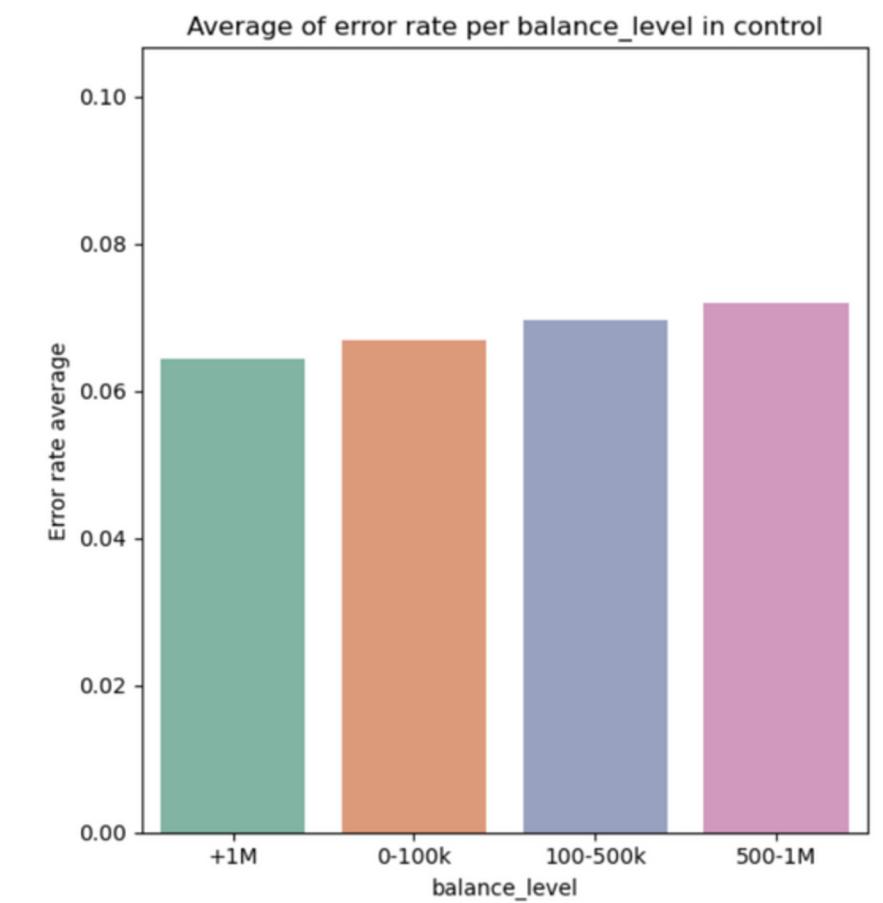
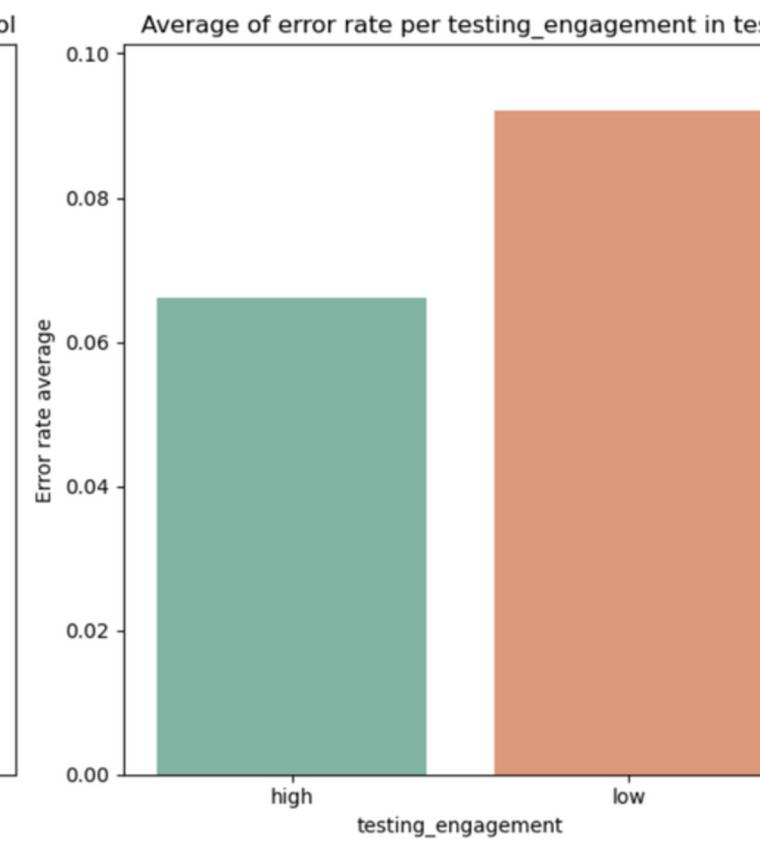
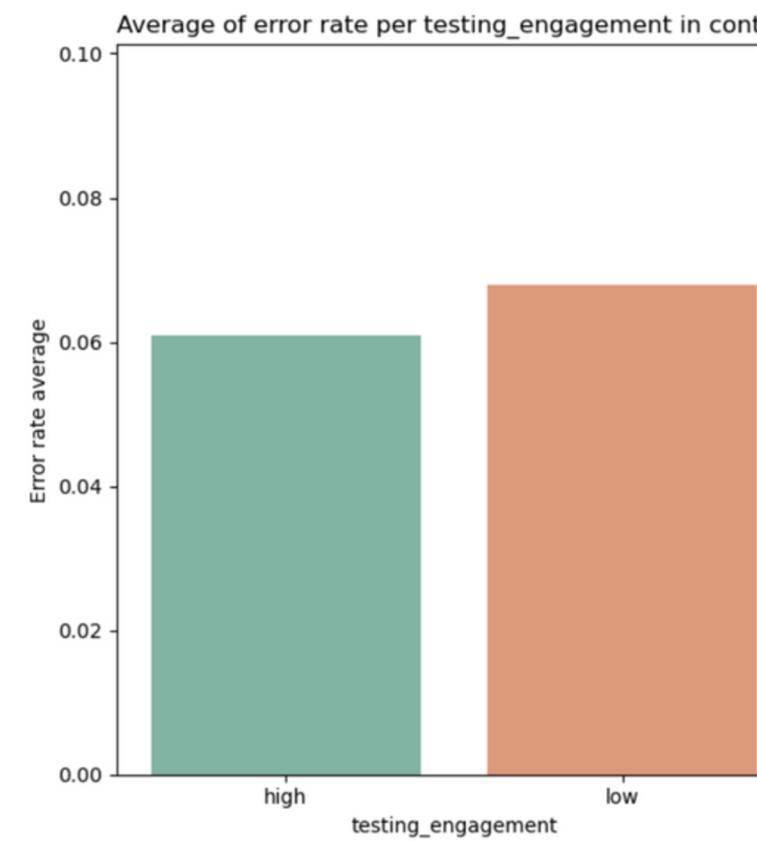
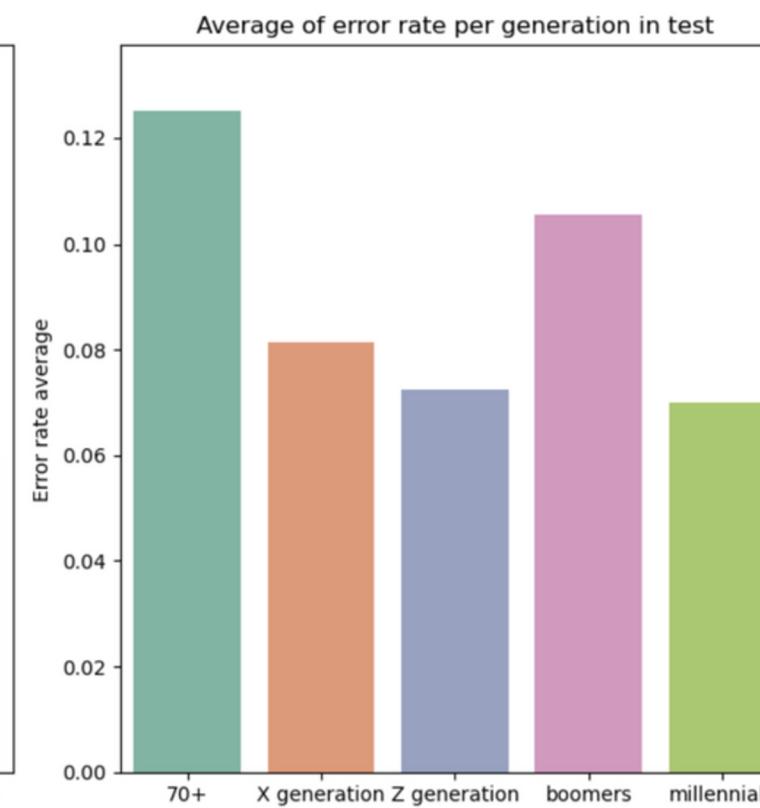
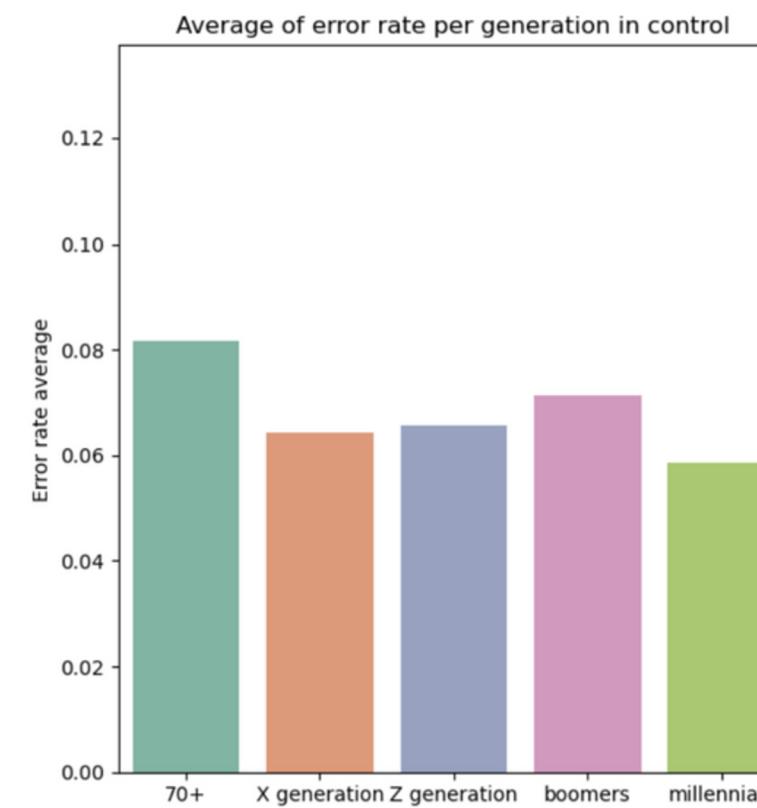
49.82%

CONTROL

67.53%

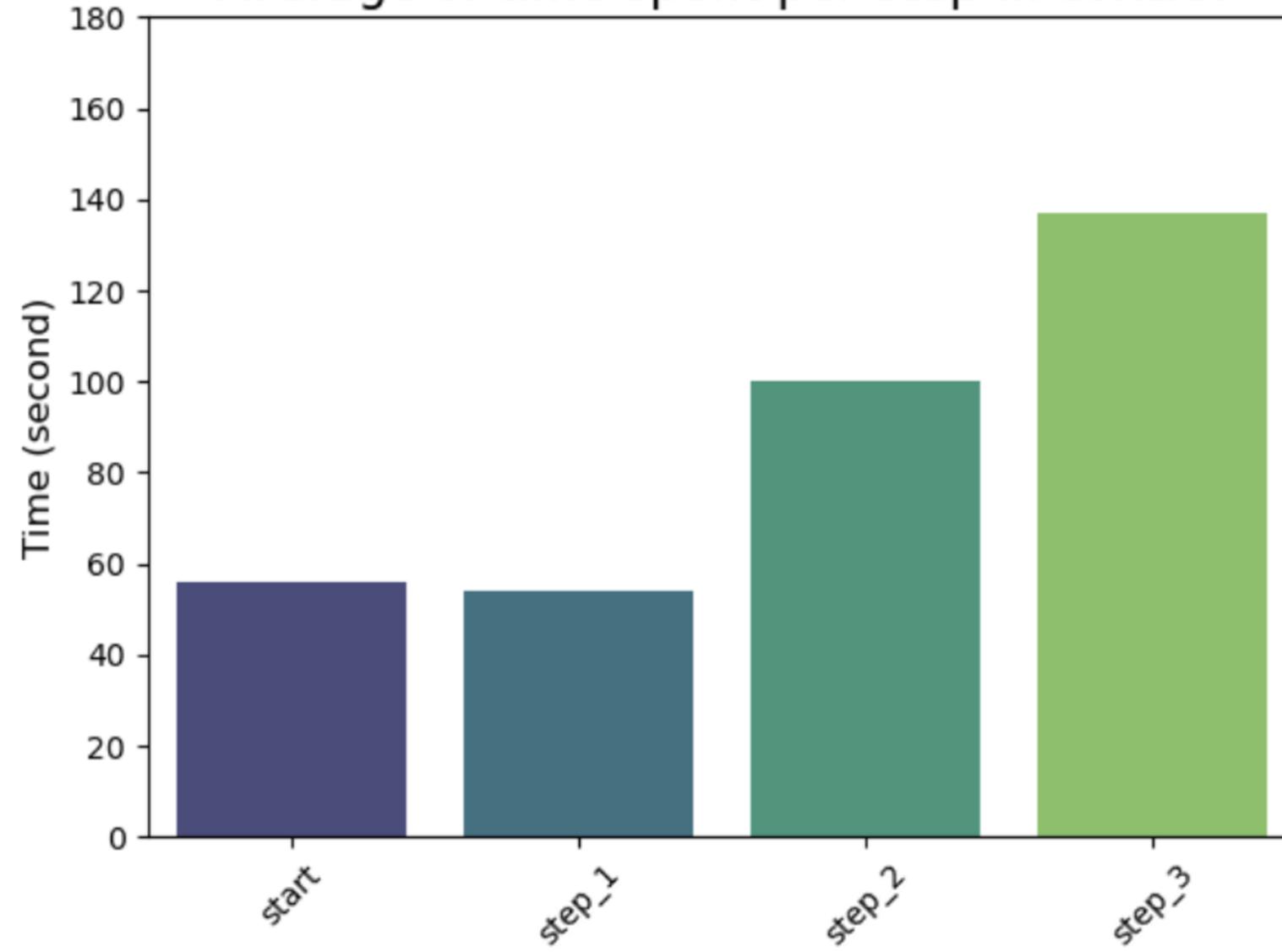
TEST

# Error Rates



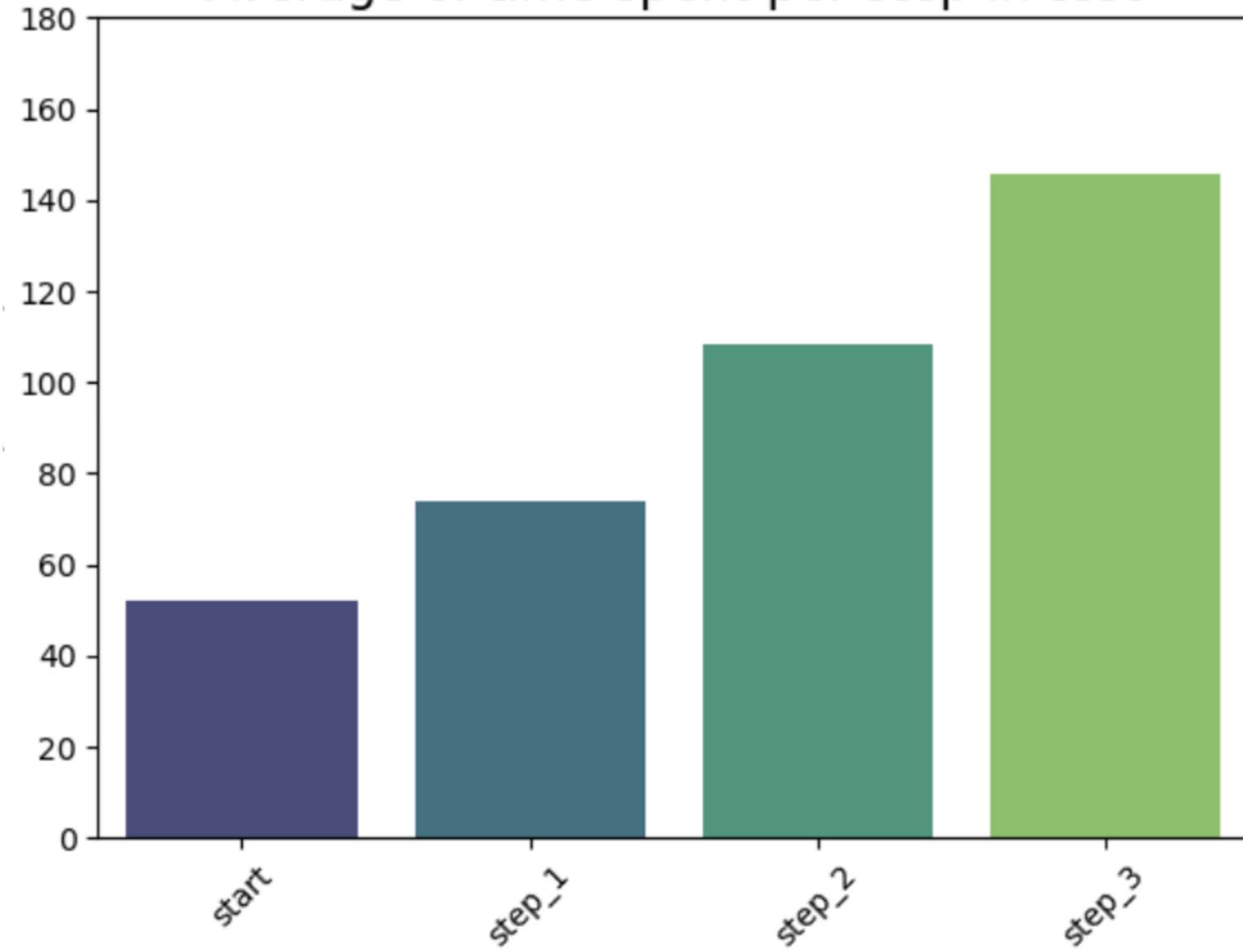
# Time spent between steps

Average of time spent per step in control



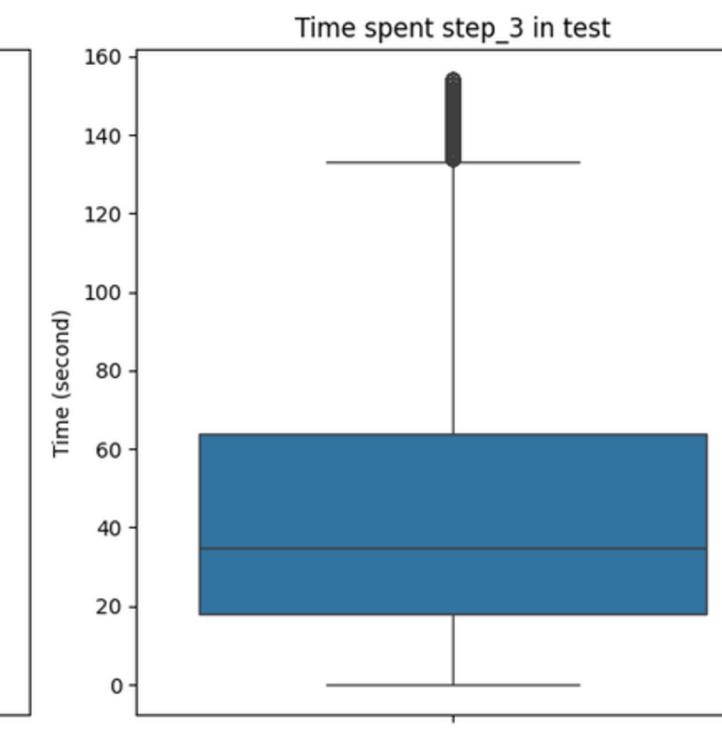
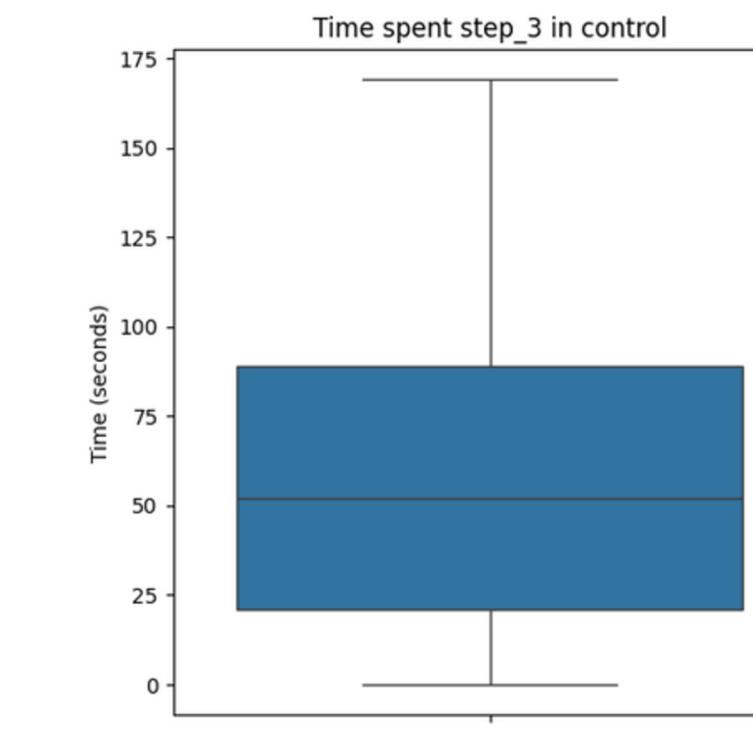
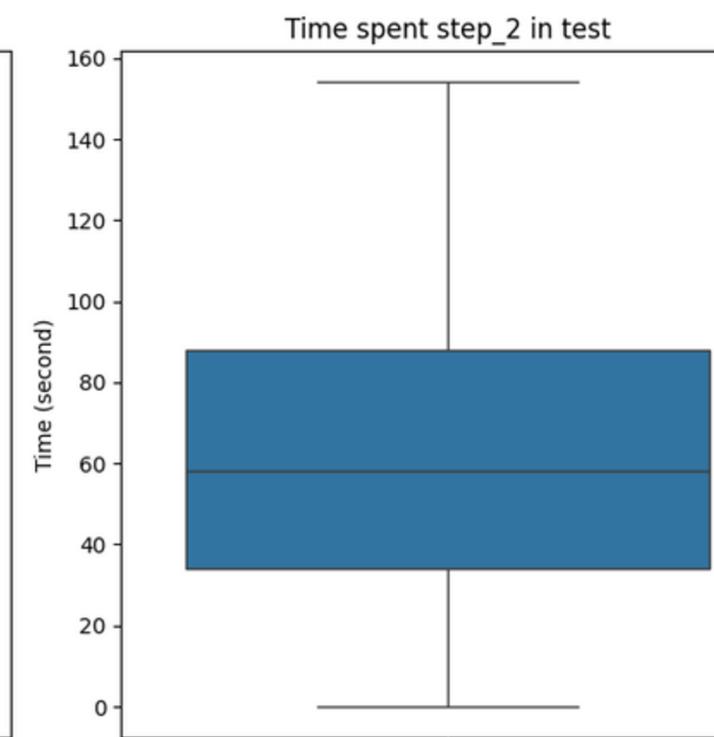
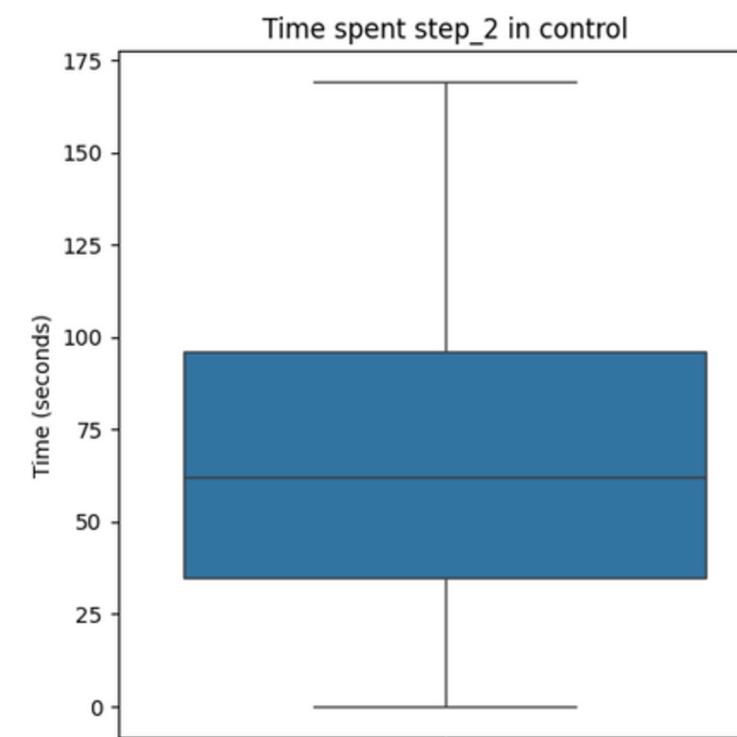
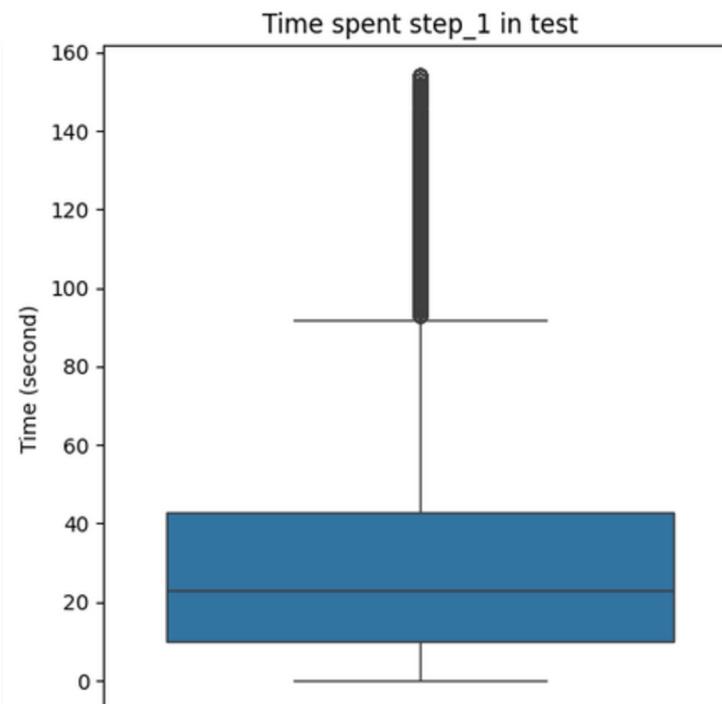
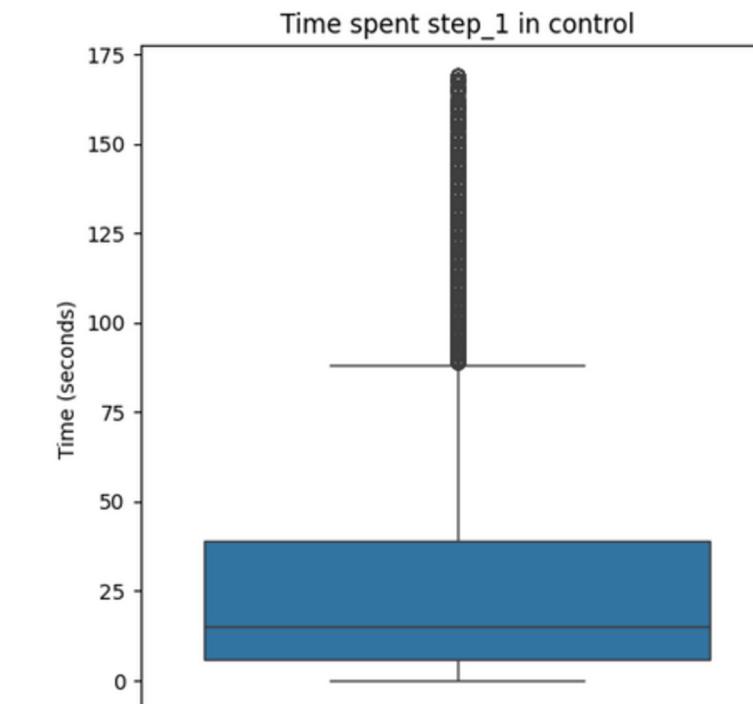
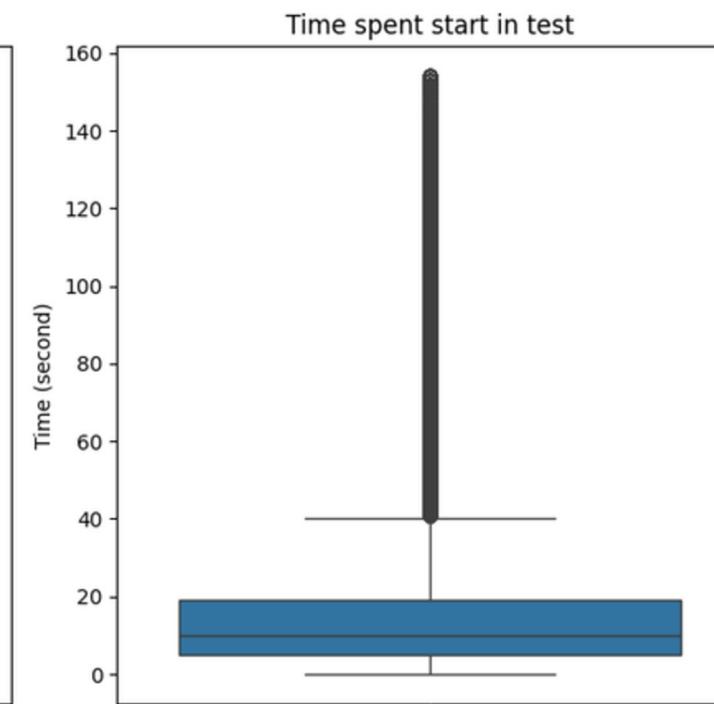
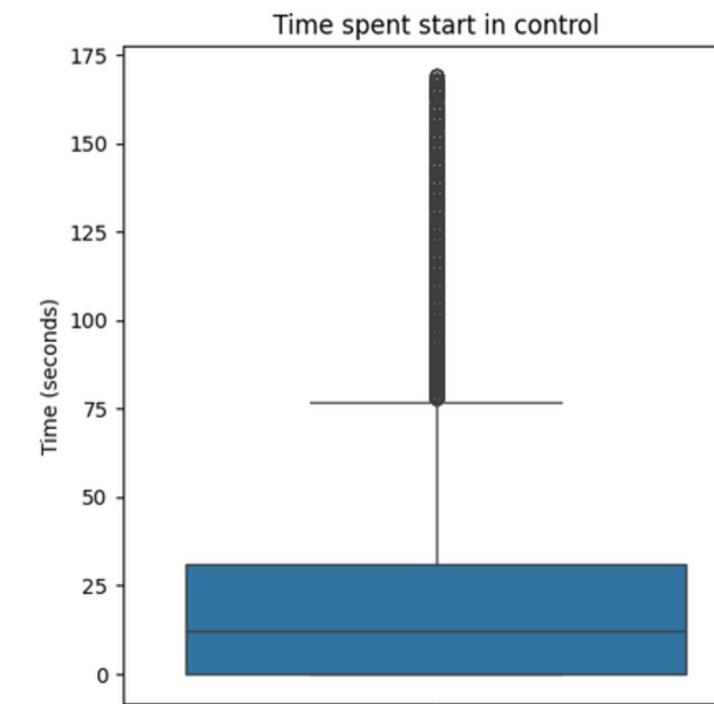
87" average

Average of time spent per step in test



95" average

# Time distribution per step





# Vanguard®

## A/B Testing

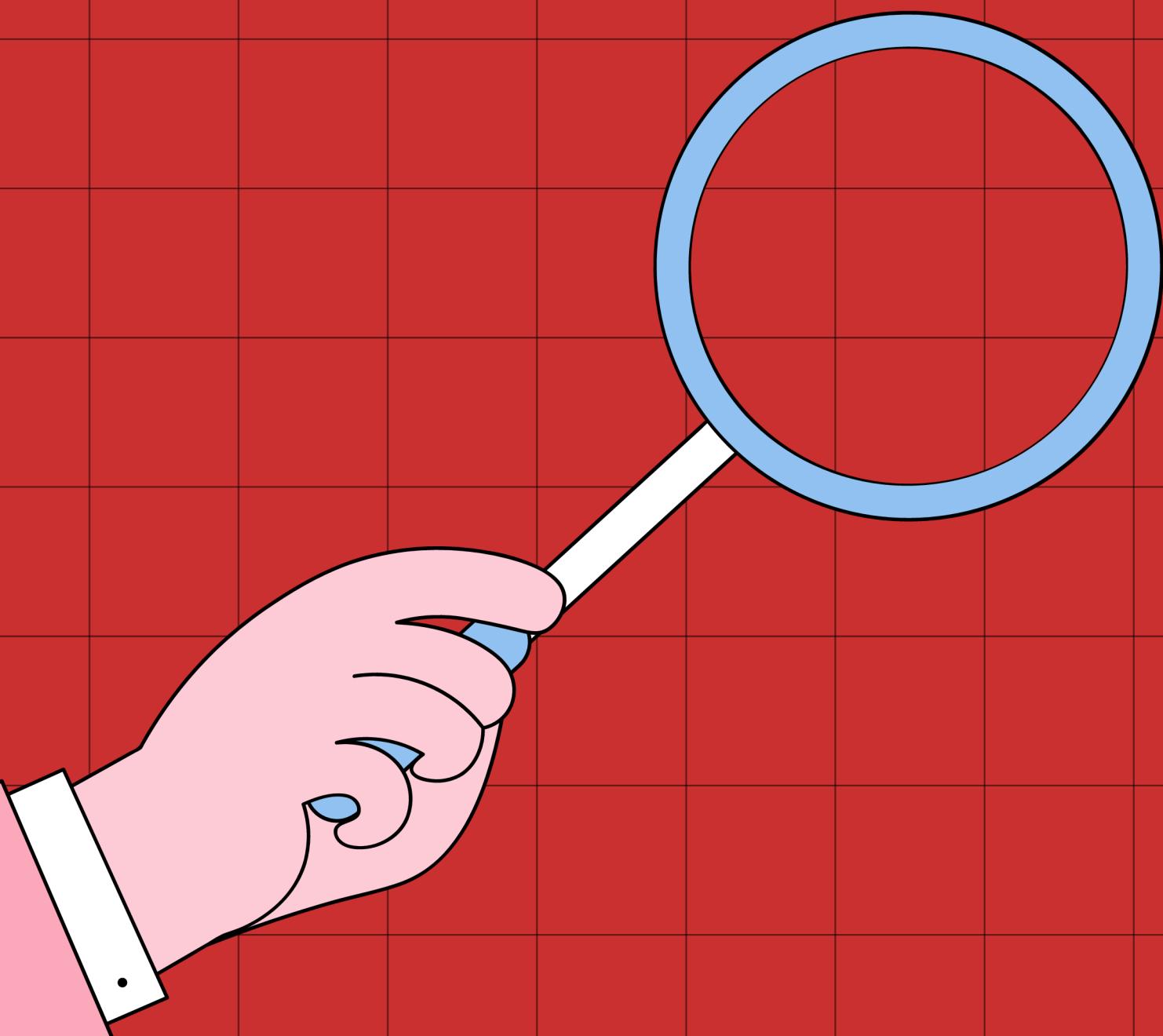
[Client analysis](#)[Time Spent](#)

### Conversion Rates

[Control](#)[Test](#)**49,82 %****67,53 %**

### Error rates

[Cont](#) [Test](#)**20,15 %****26,93 %**



# HYPOTHESIS TESTING

**What result do you hypothesize from  
the experiment?**

# Difference in completion rate

H0: The completion rate is independent of the design.

H1: The completion rate depends on the design.

**1.546e-16**

**p-value**

**80.244**

**Chi-Squared Statistic**

**5%**

**significance levels**

# Difference in completion rate

H0: There is no significant difference in completion rates between the Test and Control groups.

H1: The completion rate is significantly higher in the Test group compared to the Control group.

**3.3087e-19**

**p-value**

**8.9579**

**Statistic**

**5%**

**significance levels**

# Average of time in test and control

H0: The average of time spent in all complete is the same in both designs.

H1: The average of time spent in all complete is the same in both designs.

0.7823

p-value

-0.2889

t - Statistic

5%

significance levels

# Other hypothesis

**Group are balanced?**

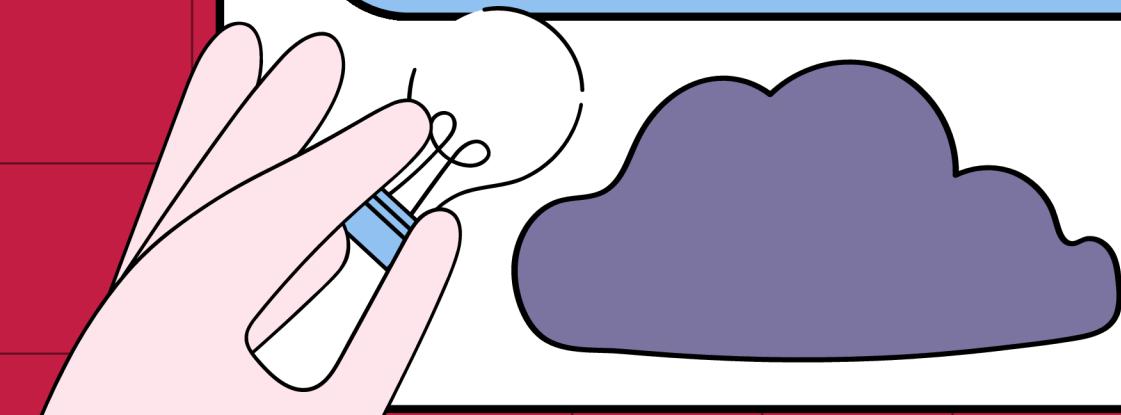
The groups aren't balanced. The difference between them are 30816 users

**Are any differences in num\_accts?**

There are significance differences in num\_accts between test and control.

**Have clients the same age?**

There are significance difference in age between both groups.



# Conclusions

- The completion rate is higher in the test group.
- The test group has a larger sample size, which could introduce potential measurement errors.
- The mean and median times are nearly identical, indicating little to no difference in time spent across both groups.
- Users tend to make more errors in the new interface, suggesting that the design may not be clear.
- A small sample of users follow the perfect customer journey: Start → 3 steps → Confirm

# Veredict

Redesign. Improve UI.

Choose other metrics

# THANK YOU!

Any questions, feel free to ask!