Vanguard

Customer Experience Analysis

Lu, Sheryll, Ngoc, Rao

Agenda

- ❑ Introduction
- ❑ Data Overview
- ❑ Exploratory Data Analysis (EDA)
- ❑ Performance Metrics
- ❑ Hypothesis Testing
- ❑ Experiment Evaluation
- ❑ Tableau Visualizations
- ❑ Project Management, Challenges & Learnings
- ❑ Conclusion

Introduction

- We are **Vanguard**, a US-based investment management company, launching our newly designed User interface with timely in-text prompts (cues, messages, hints and instructions) because we strive to make the online process smoother for our customers.

- With these new functions, we would love to learn whether these changes would encourage more clients to complete the process. In other words, did the new UI lead to higher completion rates?

Data Overview

Datasets Used:

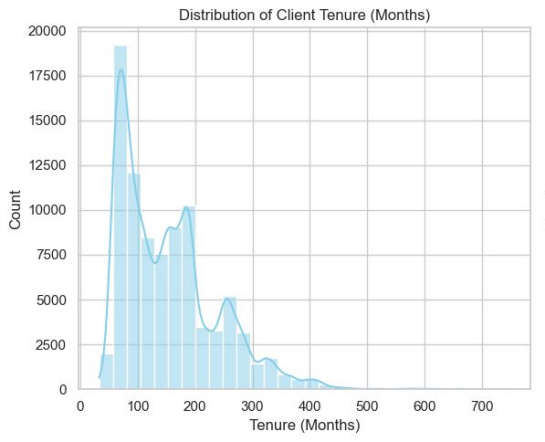
- **Client Profiles** contains demographic information of our customers (age, gender, account details);
- **Digital Footprints** details our clients' interactions online (ids, process step,date, time);
- **Experiment Roster** reveals the clients invited to the test are in the test group or control group.

Data cleaning & Merging Process:

- **Merging:**
  - Two Digital Footprints dataframes into one.
  - Experiment client & Digital Footprints dataframe
- **Cleaning:**
  - Removing duplicates
  - Fill NAs with "unknown" for experiment_client.
  - Saved final versions of dataframe for Tableau.

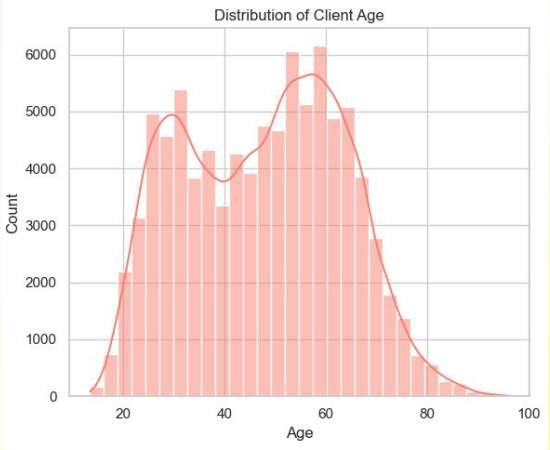
Exploratory Data Analysis

Key demographics and behaviors of Vanguard's online clients: (n=70.609)



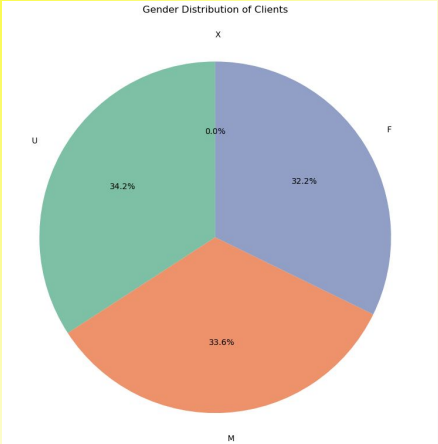
Client Tenure Distribution (in months)

Most of the customers have been using us for 8 years, we also have some customers who have been with us for longer time, centered around 15 years and 22 years. (mean tenure= 12.6 years)



Client Age Distribution (in years)

The majority of our customers is centered around 25 or 60 years old; (mean age=46.4 years old)

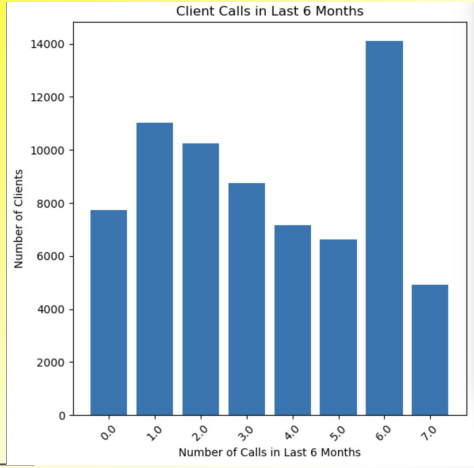


Gender Distribution (male, female, unknown)

There is an even gender distribution between male, female and unknown among our customer base.

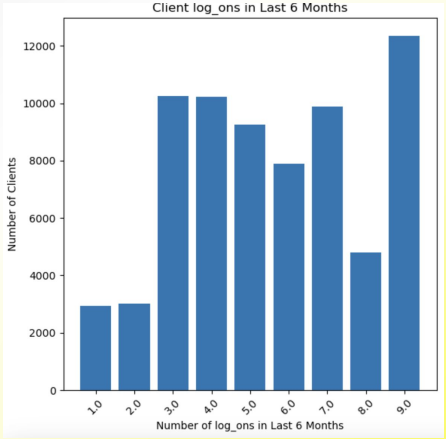
Exploratory Data Analysis

Initial findings about client engagement before diving into the A/B test results:(n=70.609)



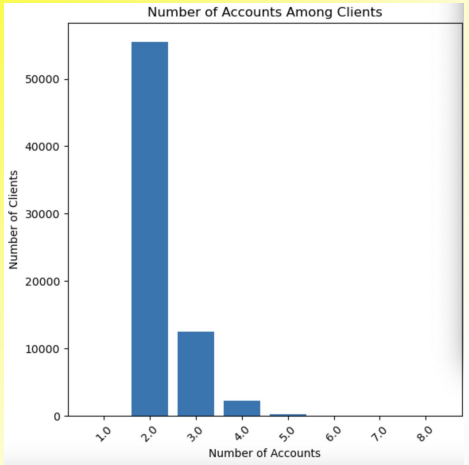
Client Calls in Last 6 Months

Most clients call 6 times in the past 6 months. The average client calls received in the last 6 months are 3.4.



Client log_ons in Last 6 months

Most clients log in to accounts 9 times in the past 6 months, the average log-ins in the past 6 months is 5.6.



Number of Accounts Among Customers

Most clients have 2 accounts (mean=2.3 accounts)

Performance Metrics

We evaluated the new design using the following KPIs:

- **Completion Rate**

  → Completion Rate for all participants: **68.1%**
  → Test group: **69.9%** ; Control Group: **65.8%**

- **Time Spent on Each Step (in hours)**

Start to Step 1	24.13
Step 1 to Step 2	9.76
Step 2 to Step 3	11.75
Step 3 to Confirm	23.93
Start to Confirm	67.66

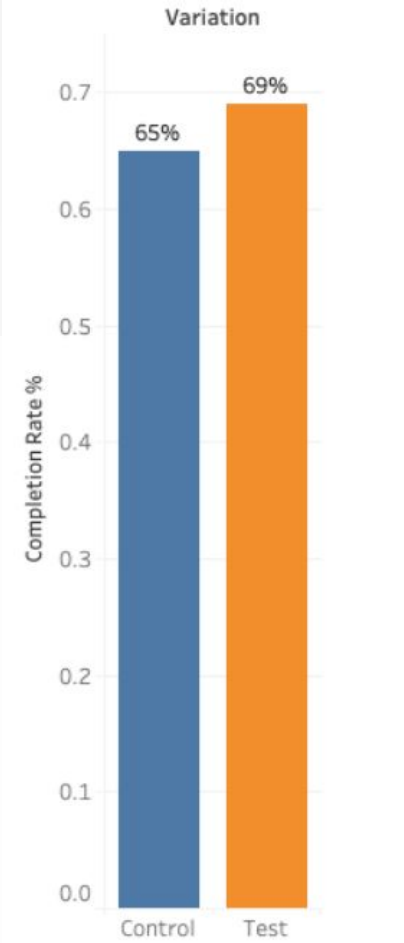
- **Error Rate**

Start to Step 1	11.23%
Step 1 to Step 2	7.94%
Step 2 to Step 3	6.3%
Step 3 to Confirm	10.8%

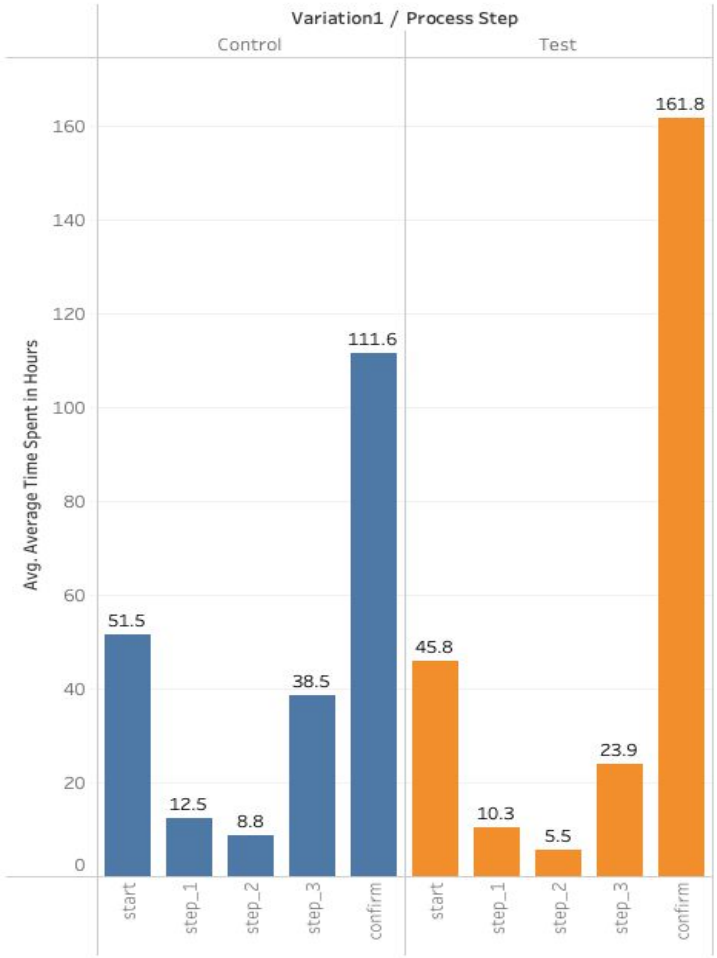
- **Drop-outs**

Start	70014
Step 1	62153
Step 2	57218
Step 3	53613

Completion Rate %



Average Time Spent in Hours



Error Rate %

Hypotheses Testing

Hypothesis 1: Completion Rate between Test and Control Group	**H0:**The completion rates are the same in the Test and Control groups.	Z-test	z-statistics=9.812 *p-value*=0.000	✅ **Reject the null hypothesis** that there is no difference between Test and Control Group in completion rate.
Hypothesis 2: Completion Rate with a Cost-effective Threshold(5%)	**H0:**The increase in completion rate from the new UI (Test group) is less than 5% compared to the old UI (Control group). p-test-p-control<5%	Z-test (one-tailed)	Z-statistics = -2.120 *p-value*=0.0983	❌ **Cannot reject the null hypothesis** that the new design meets the 5% cost-effectiveness threshold
Hypothesis 3: Average Tenure between Test and Control Group	**H0:** The average tenure is the same between Control and Test groups.	T-statistics (two-tailed)	T-statistic: 0.0526 *P-value*= 0.9580	❌ **Cannot reject the null hypothesis** that the average tenure is the same between Test and Control Groups.
Hypothesis 4: Gender Distribution between Test and Control Group	**H0:** The gender distribution is balanced between Test and Control Group.	chi-square	Chi-square: 3.6439 P-value: 0.3026	❌ **Cannot reject the null hypothesis** that gender distribution is balanced between Test and Control group

Hypotheses Testing

Hypothesis 1: Completion Rate between Testing & Control Group

We reject null hypothesis that there is no difference on completion rate between Testing & Control Group.
(Z-statistics= 9.8129, *p-value*=0.000)

Hypothesis 2: Completion Rate with a Cost-Effectiveness Threshold

We cannot reject the null hypothesis that the increase in completion rate from the new UI (Test group) is less than 5% compared to the old UI (Control group).
(Z-statistics= -2.1203, *p-value*=0.983)

Hypotheses Testing

Hypothesis 3: Average Tenure between Control & Testing Group

We fail to reject the null hypothesis that there is significant difference in average client tenure between the Test and Control groups.
(T-statistics= 0.0526, *p-value*=0.9580)

Hypothesis 4:Gender Distribution between Control & Testing Group

We reject the null hypothesis that the gender distribution is balanced between Test and Control Group.
(Z-statistics= -2.1203, *p-value*=0.983)

Hypothesis Testing 1 : Client Age by Process Group

Objective:

To test whether the average age of clients differs between the old (Control) and new (Test) process groups.

Null Hypothesis (H_o):

$\mu_1 = \mu_2$
The average age of clients is **the same** across the two groups.

Alternative Hypothesis (H_1):

$\mu_1 \neq \mu_2$
The average age of clients is **different** between the two groups.

Significance level (α) : 5%

Results:

P-Value: n/a

Since the p-value is significantly less than 0.05, we **reject the null hypothesis**.

Conclusion:
There is a **statistically significant difference** in the average age of clients between the Test and Control groups.

Hypothesis Testing 2: Client Tenure by Process Group

Objective:

To evaluate whether there is a difference in **average client tenure** between users engaging with the old (Control) and new (Test) processes.

H_0 (Null Hypothesis):

$\mu_1 = \mu_2$

The average tenure is the same across both groups (no difference)

H_1 (Alternative Hypothesis):

$\mu_1 \neq \mu_2$
The average tenure is **not** the same (there is a difference)

Significance level (α) : 5%

Results:

T-statistic = -1.62, P-value = 0.1051

Since the p-value is greater than 0.05, we fail to reject the null hypothesis.

Conclusion: There is **no statistically significant difference** in average client tenure between the Test and Control groups.

Hypothesis Testing 3: Gender and Process Engagement

Objective:
To determine whether gender distribution differs between clients engaging with the old (Control) and new (Test) processes.

H_0 (Null Hypothesis):

Gender distribution is the same across both the Test and Control groups.

H_1 (Alternative Hypothesis):

Gender distribution is different between the Test and Control groups.

Significance level (α) : 5%

Results:

- **Chi-Squared Statistic:** 3.81
- **Degrees of Freedom:** 6
- **P-Value:** 0.7023

Conclusion:

Since the p-value (0.7023) is **greater than** the significance level (0.05), we **fail to reject the null hypothesis**.

There is **no statistically significant** association between gender and process variation assignment.

Experiment Evaluation

Experiment Design:

Was the experiment well-structured?

"Yes, the experiment was well-structured. The gender distribution is proportionally balanced across groups, and client tenure is comparable, indicating similar levels of platform familiarity among participants."

Time Frame:

Was the timeframe of the experiment adequate to gather meaningful data and insights?

"Yes, the timeframe from March 15 to June 20, 2017, was **adequate** for evaluating web page completion and error rates. The cost-effective duration allowed enough user traffic to observe consistent interaction patterns, detect usability issues, and collect sufficient data."

Experiment Evaluation

Further Data Exploration

What data would we like to explore further or can help us?

- **Landing Page Changes**
  - Identify specific design, layout, content, and CTA differences between Control and Test pages
  - Understand how these changes may influence user behavior
- **User Interaction (via Hotjar)**
  - Click maps: Which elements are used or ignored
  - Heatmaps: Scrolling and engagement patterns
  - Form & CTA interactions (e.g., “Buy Now”, “Subscribe”)
  - Session recordings: Navigation paths and friction points
  - On-page surveys for user intent and feedback

Experiment Evaluation

Further Data Exploration

What data would we like to explore further or can help us?

- **Funnel Insights (via GA4)**
  - Page performance: Load speed, error rates, broken elements
  - User behavior by device and browser (desktop vs. mobile)
  - Traffic source analysis: Do sources behave differently on each version?
  - Bounce rates and exit points
  - Geographic distribution of visitors
- **Business Context**
  - Product details: Type, pricing, and presentation on each page version
  - Company’s market segment and key competitors

Tableau Visualization: see Application

Project Management,Challenges & Learnings

Project Management:

Divided the tasks and worked individually on different sections. Ran into challenges....

Challenges:

- Collaboration on Git branches, merging and pulling.
- Understand the other analysts' thinking process (syntax coding).
- Time-restricted.

Learnings:

- Importance of standardized work process: variable names, dataframe names.
- Clear communication is key.

Conclusion

Final Verdict: The new User Interface in Vanguard fails to make the online process smoother for clients as we couldn’t find significant completion rate improvement in the Test Group.

Key Findings: The experiment is well-designed (duration, age distribution, gender distribution); A more intuitive and modern User Interface coupled with timely in-text prompts may not be enough to significantly improve user experience.

Recommendations: We suggest adding additional functions in landing page changes and more user interactions designs to test if this will impact on user experience; Making steps easier and clearer with pre-trainings to follow with less confusion.

Any questions? Ask away!

Thank you!