



Vanguard Investment

A/B Test

19/01/2024

Salim & Pollyana

Team: *We don't have a name...* 🕶️

Summary

1. **Introduction**
2. **Data Overview**
3. **Exploratory Data Analysis (EDA)**
4. **Performance Metrics**
5. **Hypothesis Testing**
6. **Experiment Evaluation**
7. **Tableau Visualizations**
8. **Conclusion**

Vanguard New Design AB TEST

Will the New Design lead to a better user experience and higher process completion rate?

- The Experiment Conducted An A/B test was set into motion from 3/15/2017 to 6/20/2017 by the team.
- Control Group: Clients interacted with Vanguard's traditional online process.
- Test Group: Clients experienced the new, spruced-up digital interface.

Data Overview

	client_id	clnt_tenure_yr	clnt_tenure_mnth	clnt_age	gendr	num_accts	bal	calls_6_mnth	logons_6_mnth	age_category
0	836976	6.0	73.0	60.5	U	2.0	45105.30	6.0	9.0	60+
1	2304905	7.0	94.0	58.0	U	2.0	110860.30	6.0	9.0	30-60
2	1439522	5.0	64.0	32.0	U	2.0	52467.79	6.0	9.0	30-60

shape
(70609, 9)

	client_id	visitor_id		visit_id	process_step	date_time
0	9988021	580560515_7732621733	781255054_21935453173_531117		step_3	2017-04-17 15:27:07
1	9988021	580560515_7732621733	781255054_21935453173_531117		step_2	2017-04-17 15:26:51
2	9988021	580560515_7732621733	781255054_21935453173_531117		step_3	2017-04-17 15:19:22

shape
(343141, 5), (412264, 5)

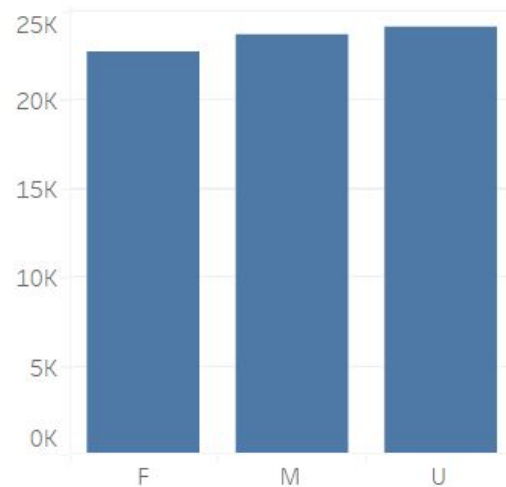
	client_id	variation
0	9988021	Test
1	8320017	Test
2	4033851	Control

shape
(70609, 2)

Exploratory Data Analysis (EDA)

- Almost even number of M/F clients
- We observed a huge number of Unknown gender

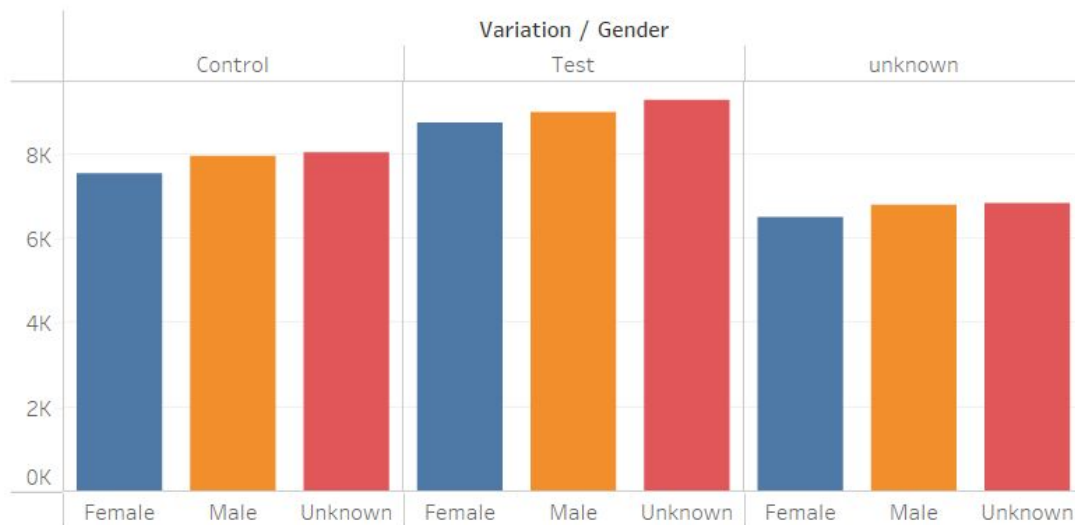
Total Clients Per Gender



Exploratory Data Analysis (EDA)

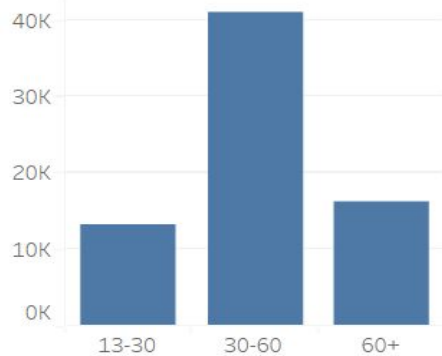
- Count of Male on both Control and Test are slightly higher than Female
- Male and Female distribution on Test and Control seems symmetric

Count of Gender by Variation

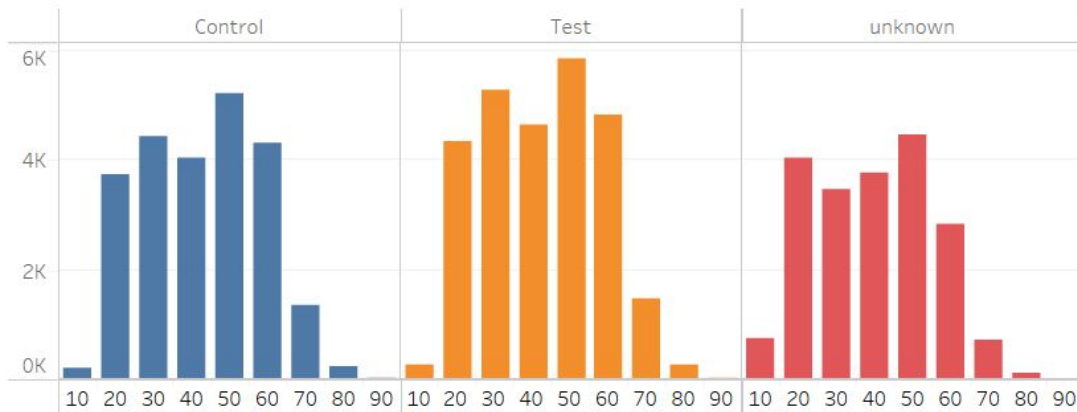


Exploratory Data Analysis (EDA)

Count of Clients Per Age Category

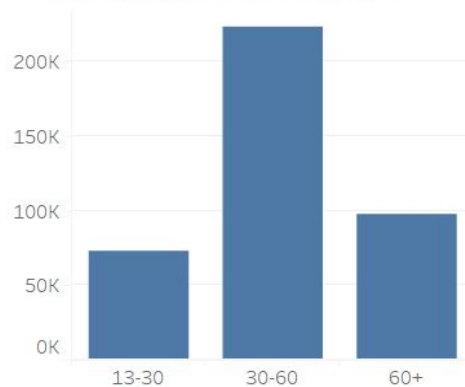


Client Distribution Per Ages Group

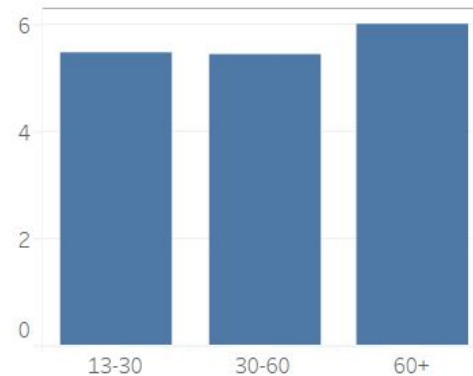


Exploratory Data Analysis (EDA)

Count of Login Per Age Category

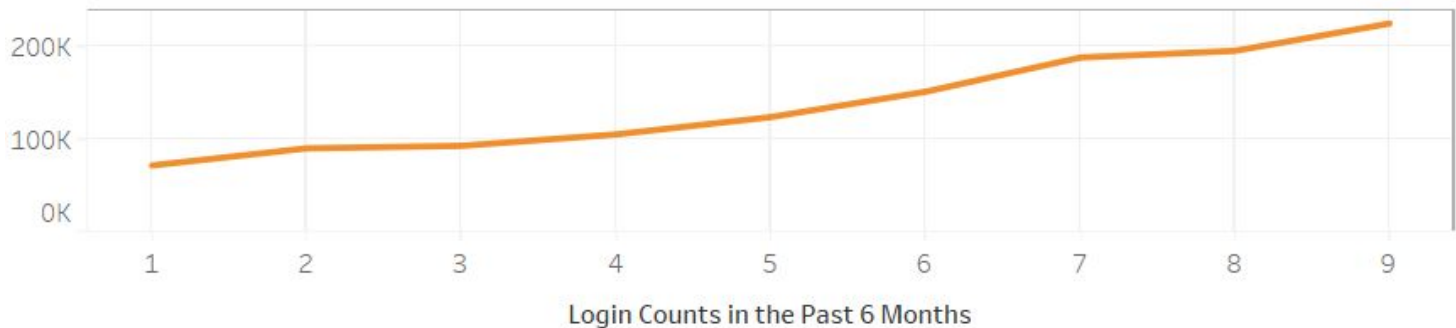


Avg. Logins Per Person (Past 6 Months)

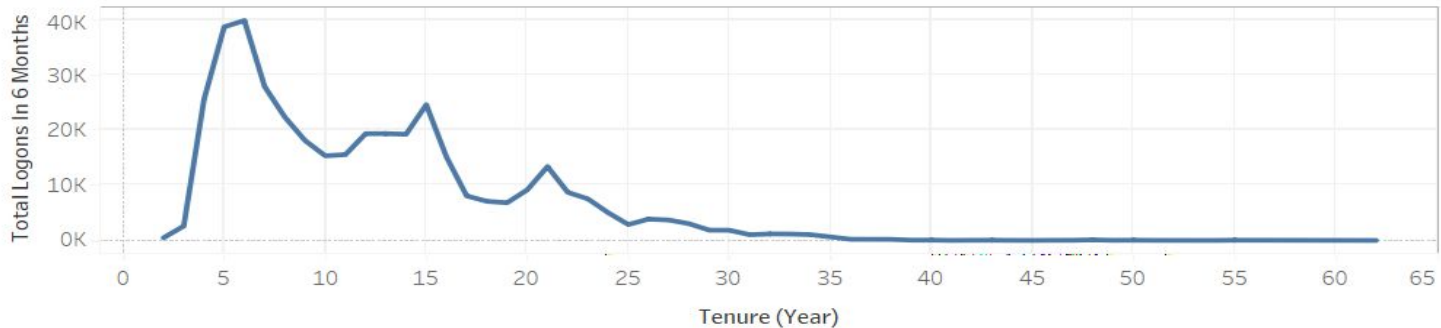


Exploratory Data Analysis (EDA)

Average Balance Per Logins



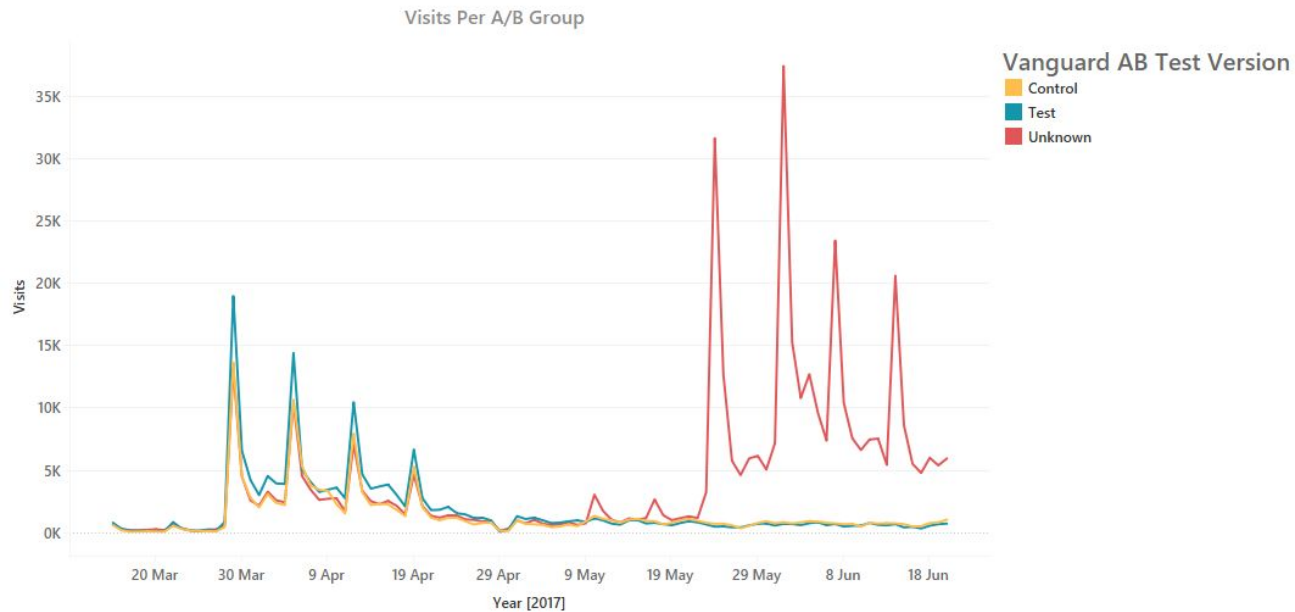
Total Logons by Client's Tenure



Performance Metrics



Visits and users during A/B Test





KPIs used to evaluate Vanguard new design's performance:

Time Spent on Each Step

The average duration users spend on each step

Error Rates

Each time a client go back to a previous step

Number of steps

number of actions (steps) taken by users

Completion Rate

The proportion of users who reach the final 'confirm' step

Error Rates by Step

The error rates by step, how each step is performing

Site Version	Completion rate	Error rate	Avg Number of Steps	Avg Step Duration
Test	0,69	0,38	4,18	01:14
Control	0,66	0,34	4,04	01:12
Unknown	0,68	0,35	4,10	01:13

Performance Metrics



Completion Rate (Percentage of clients who completed the step “confirm” in each group):

69.2 %

Test

65.5 %

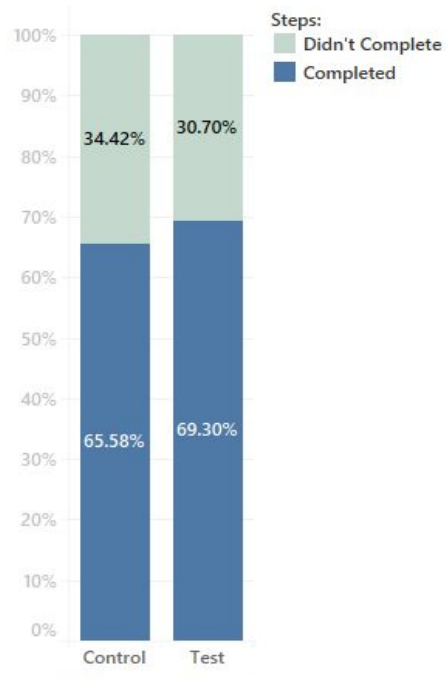
Control

+ 3.7

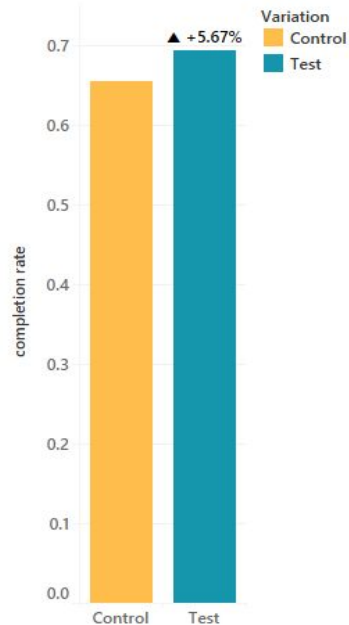
**Completion rate
points
increase**

+ 5.7 %

Lift percentage



**Completion Rate
Variation**

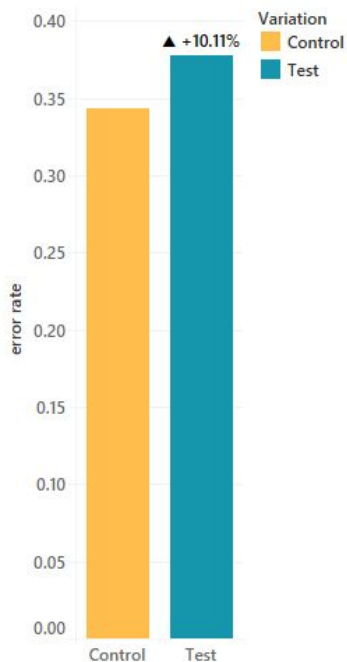


Performance Metrics

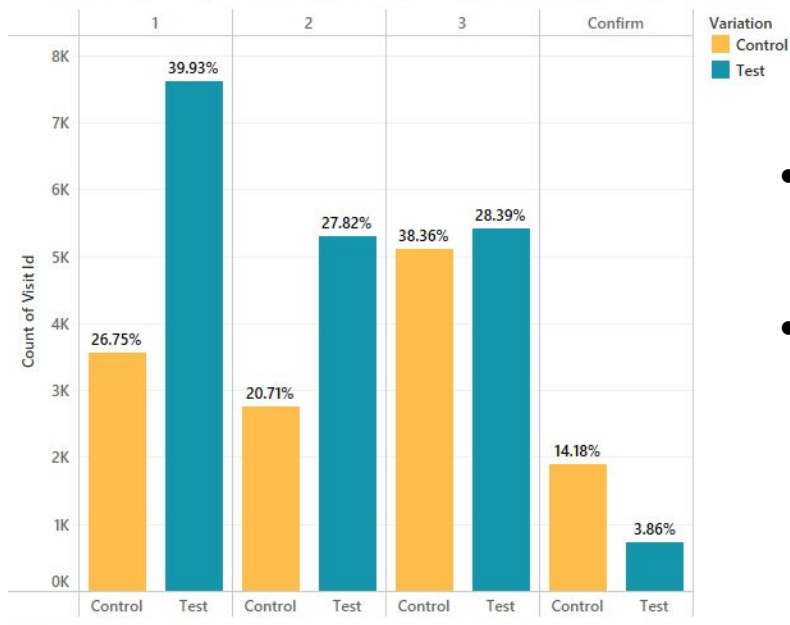


Error Rate (Percentage of clients who went to a previous the step “confirm” in each group):

Error Rate Variation



Process Steps: Steps Generating Errors (% of A/B Test Group Errors)



- Step 1 is the step generating more errors in the Test Version
- Step 3 is the one generating more errors in the Control Version



Steps performance Which previous steps are the users returning to?

Site Version	Avg Number of Steps	Avg Step Duration
Test	4,18	01:14
Control	4,04	01:12





Completion rate differences using Two Proportions Z-test

5 percent level of significance. $\alpha = 0,05$ `confidence_interval(confidence_level=0.95)`

Size Group 'Clients Test Version': 26968

Size Group 'Clients Control Version' 23532

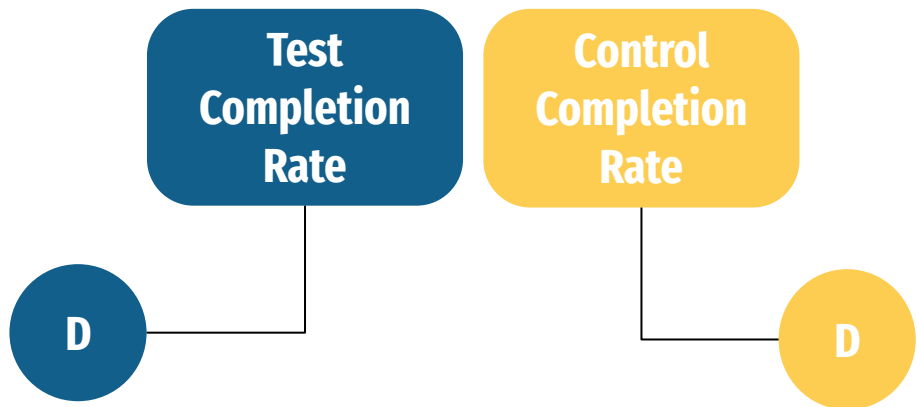
Null Hypothesis H_0 :

Alternative Hypothesis H_1

Z Test or T Test

Z or T-statistic:

P-value:





Completion rate differences using Two Proportions Z-test

Two proportion z-test allows comparing two proportions to see if they are the same.

Null Hypothesis H_0 :

The new design (Test group) had a higher completion rate compared to the old design (Control group), differences observed are not statistically significant

Alternative Hypothesis H_1

The new design (Test group) had a higher completion rate compared to the old design (Control group), and this difference is statistically significant

We reject the null hypothesis

There is a significant difference between groups.

Z-statistic:

-8.8745141890702

Z-test **P-value**:

7.023933247581432e-19

**Test
Completion
Rate**

0.69

**Control
Completion
Rate**

0.66



Cost-effectiveness threshold evaluation using Z-test

Null Hypothesis H0:

The completion rate for the Test group (new design) is equal to or less than the completion rate for the Control group (old design) increased by 5%, and this difference is statistically significant.

Alternative Hypothesis H1:

The completion rate for the Test group (new design) is greater than the completion rate for the Control group (old design) increased by 5% and this difference is statistically significant.

We fail to reject the null hypothesis

not confident we will see a 5% lift should we implement the changes

Z-statistic: 1.5186905692212727

Z-test P-value: 0.06442020077768941

**Test
Completion
Rate**

0.69

**Control
Completion
Rate**

0.66



Other Hypothesis - Average Age using T-Test

Null Hypothesis H0:

The average age of clients engaging with the new process is the same as those engaging with the old process (the samples are drawn from populations with the same population means)

Alternative Hypothesis H1:

The average age of clients engaging with the new process is not the same as those engaging with the old process (the samples are drawn from populations with different population means)

We reject the null hypothesis

There is a significant difference between groups.

TtestResult statistic :

-2.416068061200627

TtestResult **P-value:**

0.015692719461388675

**Test
Average
Age**

47.5

**Control
Average
Age**

47.2

Experiment Evaluation

It has a clear and defined **purpose**: whether a new UI can improve customer experience

The experiment was conducted over a reasonable **timeframe** of three months

clients were **equally divided** between old and new design.

Unknown gender type in dataset which can affect the test result.



Thank You!

Questions?

