alphabet

Vanguard A/B Test

Francisco, Jugal & Max

Main Question

alphabet

Did the new UI lead to a higher completion rate?

What we did

alphabet

EDA

- Import and inspect the data sets
- Handling missing data
- Analyzing relationships between variables

KPIs & Hypothesis

- Reviewing the KPIs to determine the success of the new design
- Hypothesis testing to make data-driven conclusion

Evaluation

- Get the conclusion rather the experiment was successful or not
- Tableau visualization

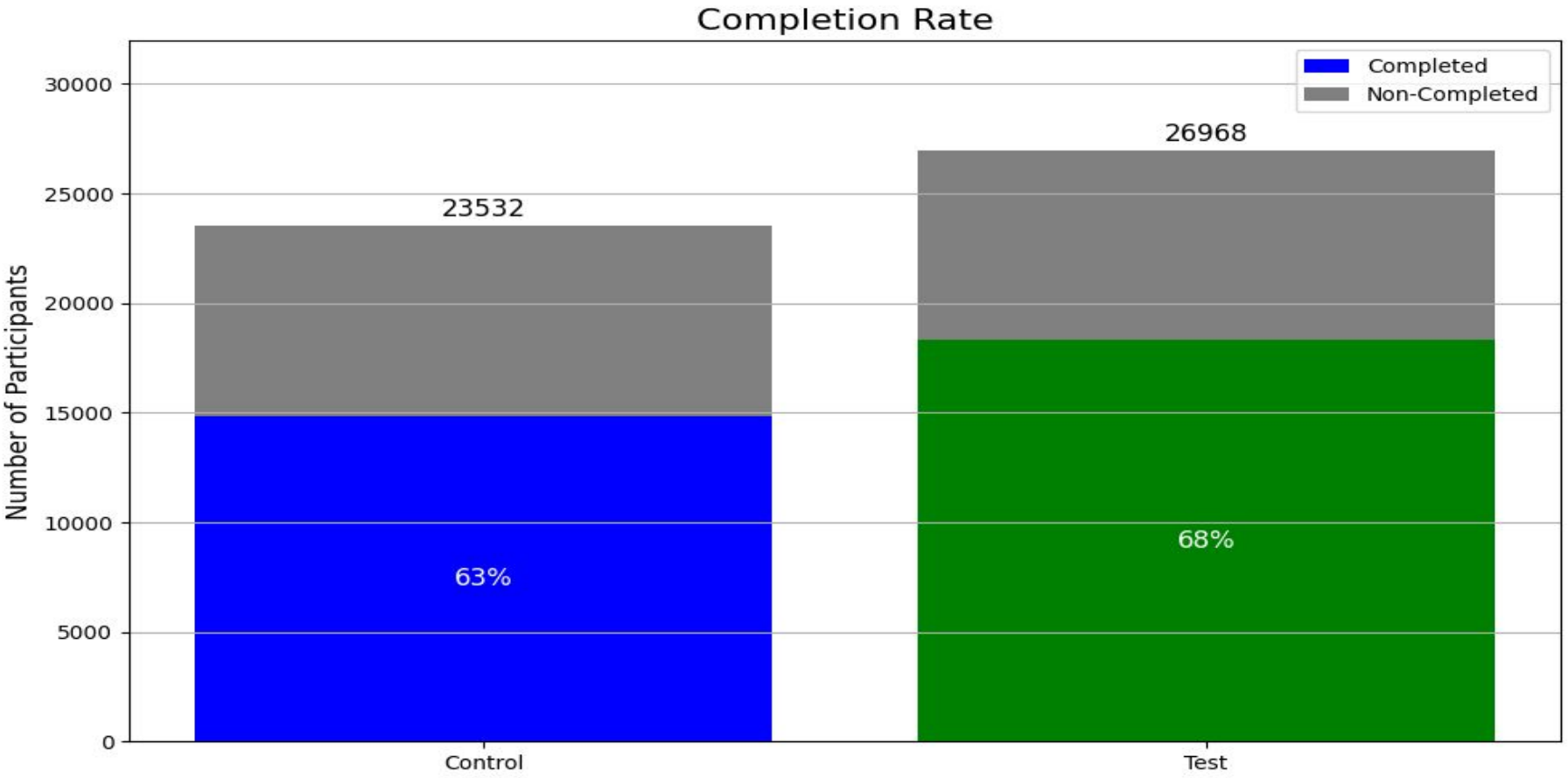
alphabet

KPIs

63% vs 68%

Completion Rate = Participants that reached confirmed / Total Participants

alphabet

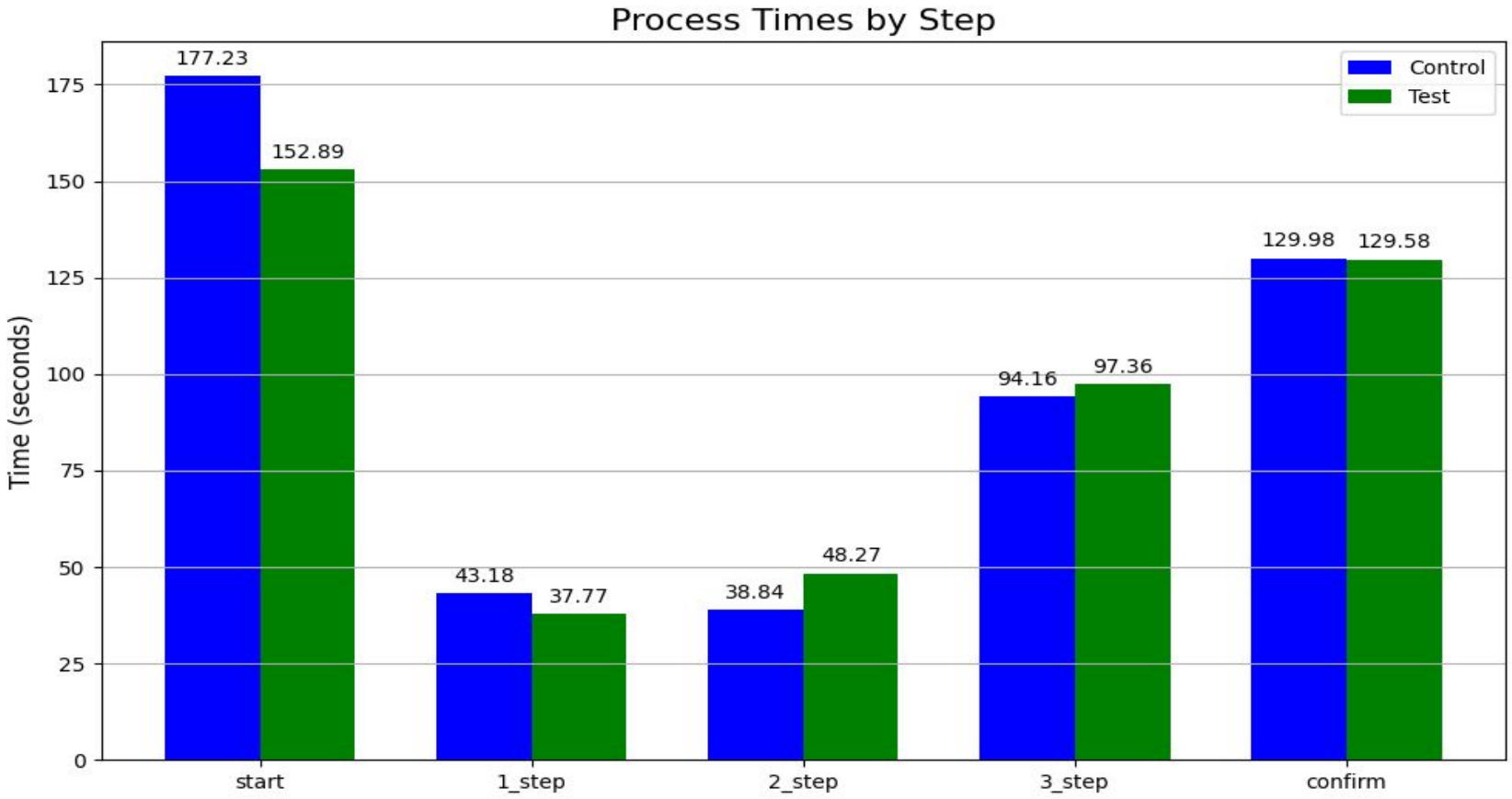

Completion Rate
Completed
Non-Completed
Number of Participants
30000
25000
20000
15000
10000
5000
0
23532
26968
63%
68%
Control
Test

alphabet

96,7 vs 93,2 seconds

Time Spent = Timestamp of next step - Timestamp of current step

alphabet


Process Times by Step
Time (seconds)
0
25
50
75
100
125
150
175
Control
Test
177.23
152.89
43.18
37.77
38.84
48.27
94.16
97.36
129.98
129.58
start
1_step
2_step
3_step
confirm

19,1% vs 24,3%

Error Rate = Backward Steps Taken / Total Steps Taken

alphabet


Error Rate Comparison
Error Rate (%)
30
25
20
15
10
5
0
19.1%
24.3%
Control
Test

alphabet

Hypothesis Testing

Are our hypothesis true?

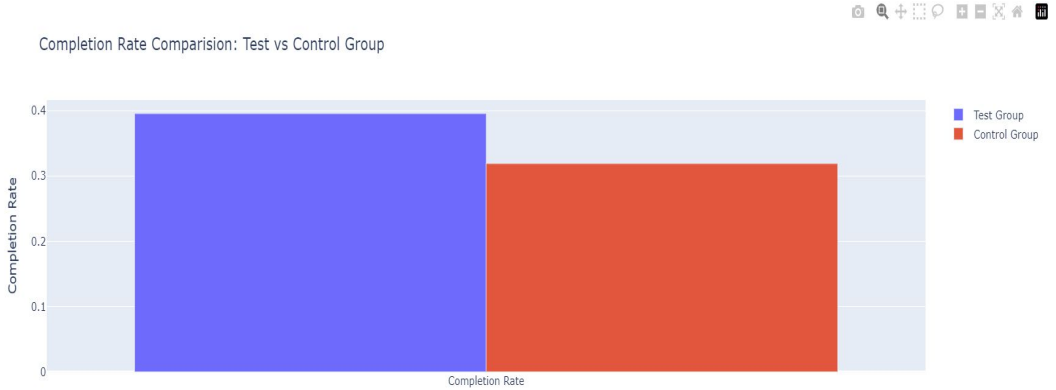
alphabet

1. Completion Rate

Theory:
H0: No significant difference in completion rates between the Test and Control group

H1: Significant difference in completion rates between the Test and Control group

Result:
Test group completion rate: 0.3954
Control group completion rate: 0.3189
Z-statistic: 75.6573
P-value: 0.0000



Conclusion:

P-value is extremely small(much less than our significance level of 0.05) which provides the evidence to reject the null hypothesis. Therefore, we conclude that there is a statistically significant difference in completion rates between the test and control groups. The new UI design (Test group) demonstrates a significantly different completion rate compared to the old design (Control group).

Are our hypothesis true?

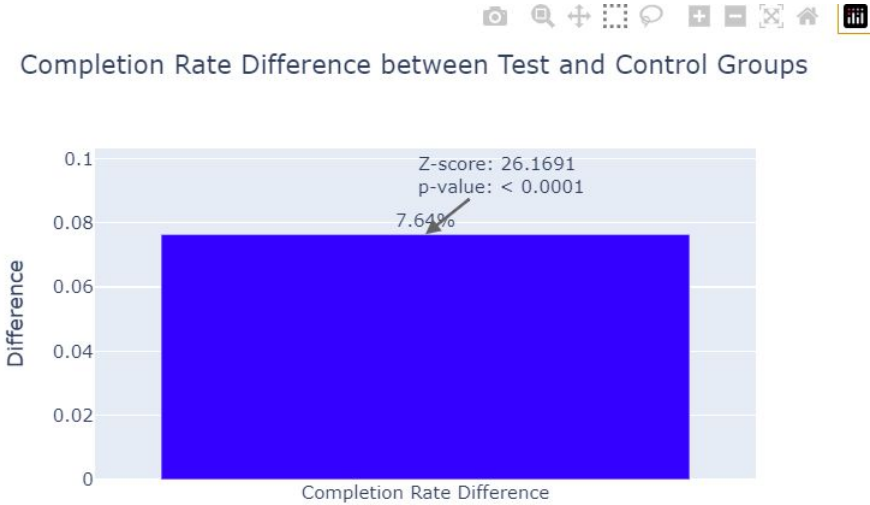
alphabet

2. Completion Rate with a Cost-Effectiveness Threshold

Theory:
H0: Difference in completion rates (Test - Control) <= 5%

H1: Difference in completion rates (Test - Control) >= 5%

Result:
Observed difference in completion rates: 0.0764
Z-score: 26.1691
P-value: 0.0000



Conclusion:

P-value is less than significance level, i.e., 0.05 which suggests that it rejects the null hypothesis. Based on it, we can conclude that the new UI design leads to an increase in completion rate that exceeds 5% threshold, making it cost-effective.

Are our hypothesis true?

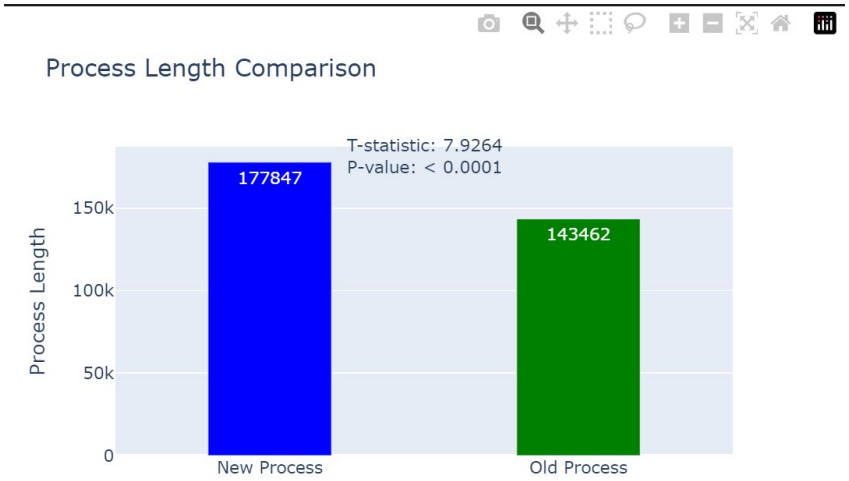
alphabet

3. Average age of clients

Theory:
H0: No difference in the average age of clients engaging with the new process versus the old process

H1: Difference in the avg. age of clients engaging with the new process versus the old process

Result:
New Process Length: 177847
Old Process Length: 143462
T-statistic: 7.926400435419626
P-value: 2.263054043650787e-15



Conclusion:

P-value is less than 0.05 which reject the null hypothesis and conclude that there is significant difference in the average age of clients engaging with new process, i.e., Test group and engaging with old process, i.e., control group

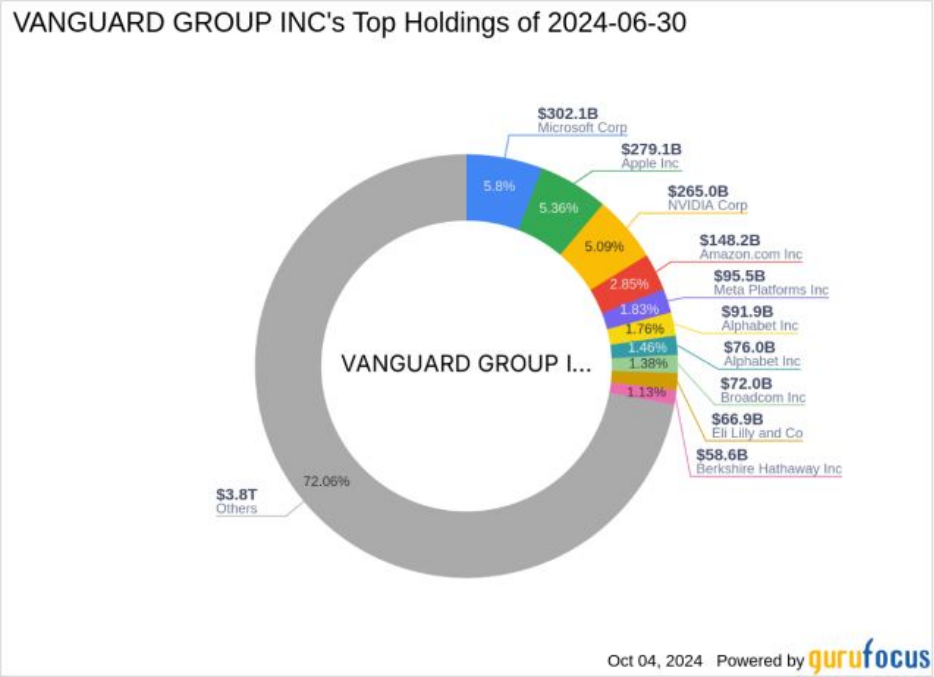
alphabet

Experiment Evaluation

Vanguard Group

alphabet

- Founded in 1975
- Number of clients —> 50M
- Employees —> 20,000
- 10,1$T VS 11,48$T AUM(Blackrock)



Experiment Structure

alphabet

Control

- Participants: 23,532
- Avg Age: 48 y.o
- Avg Tenure: 12 years

VS

Test

- Participants: 26,968 +14%
- Avg Age: 49 y.o
- Avg Tenure: 12 years

Experiment Structure

alphabet

Was the experiment well-structured? Yes, but..

Any biases? No, but..

3/15/2017 - 6/20/2017 adequate? Yes

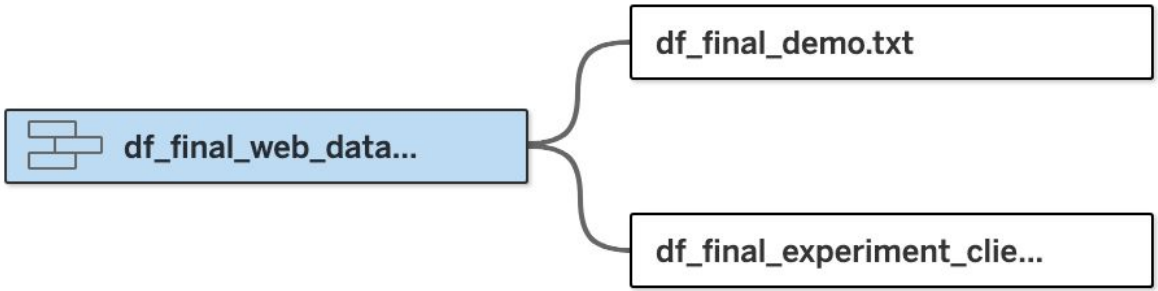
Any suggestions? Having a 50M client base..

Vanguard®

alphabet

Tableau

alphabet

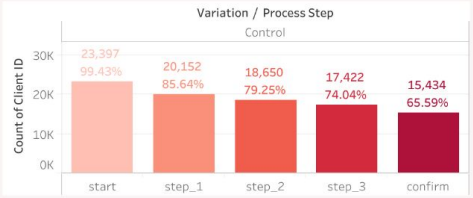


alphabet

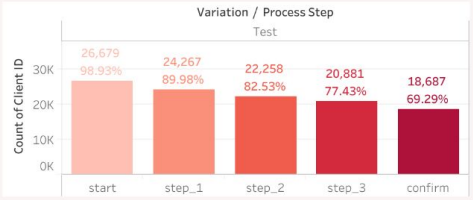
Vanguard

A/B Test Results

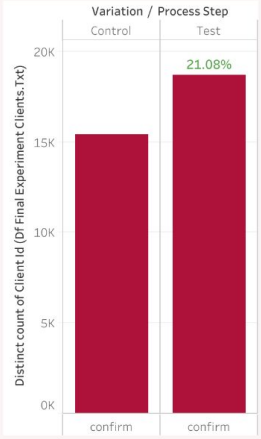
Completion Rate Control



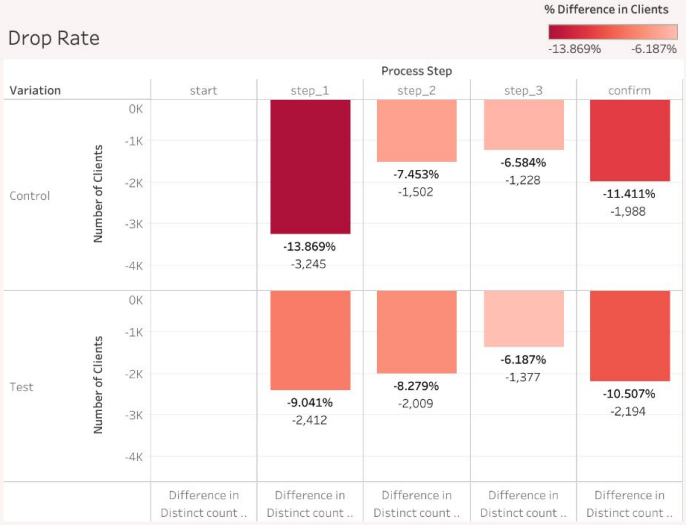
Completion Rate Test



Completion Control vs. Test



Drop Rate



Search Bar

Test Participants

Client Id Str..	Gendr	Variation	Bal	Clnt Age	Logons 6 M..	Time Spent
169	M	NA	501,571	48	4	0
555	U	Test	25,455	30	6	0
647	M	Test	30,526	58	4	0
733	F	NA	22,466	60	1	0

Average time spent per step

alphabet

Conclusion

Key Learnings

alphabet

Merging?

- Combining data sets
- Asking for help
- Still productive

Tableau

- Design vs. Data

alphabet

Thank you!