



New GUI client response

Vanguard CX team - Targaryan House
Benoit-Emmanuel



Business case

- The digital world is evolving, and so are Vanguard's clients
- The trading platform generates fees and is a large contributor to Vanguard's revenue
- Vanguard has developed an improved User Interface for trading operations (buy / sell)
- Before rolling out the new UI to all customers, we need to ascertain that the new platform will:
 - Facilitate the completion of the trading transaction
 - Optimize the trading operations process
 - Duration
 - Repetition of steps
- The CX team has conducted a data analysis study using the A/B testing methodology to
 - Assess the performance of the new platform
 - Provide insights on the Vanguard customer base

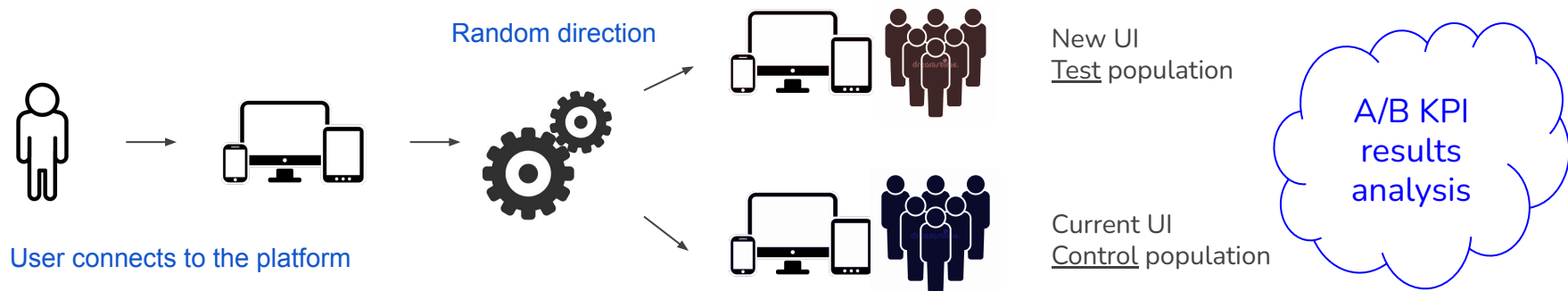


A/B Testing Methodology

- A/B testing is a statistical method used to compare two versions of an application to assess which one performs better
- When the user logs on to the platform, it is randomly directed to the new or old version of the GUI
- The behaviour of the 2 populations (Test: new version / Control: old version) is then analyzed and compared measuring results on performance KPIs:
 - Completion rate (proportion of transactions confirmed)
 - Repetition of transaction steps
 - Duration of the process



An improvement of at least 5% in completion rate is expected to justify the financial cost to roll out the new platform





Data analysis process

Acquisition



shutterstock.com - 12252426 shutterstock.com - 12252427 shutterstock.com - 12252428 shutterstock.com - 12252429

4 txt source files

- 2 web transactions
- Customer info
- A/B list

Data cleaning & Wrangling

- Dropping null values
- Data type casting
- Calculated columns
- Bins for age / tenure

Aggregation

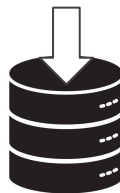
Web transactions files

- Concat 2 files
- Calculated #steps, step duration
- Confirm (True/false)
- Repetition (true/False)
- Summarized web transactions in data frame with one 1 row / visit id

Linking tables on client_id

- Customer base with A/B list
- Web transactions

One single dataset to support Analysis



Analysis

Exploratory Data Analysis (EDA)

- Charts & graphs
- KPI measures
- Hypothesis testing
 - Z-test for A vs B results significance
 - T-test (Anova) for A/B samples homogeneity

Tableau Dashboard for multi-dimensions visualizations





New GUI performance results

Measure the results of A/B testing on the selected KPIs

Completion rate

variation	Total Transactions	Confirmed Transactions	Completion Rate
Control	32026	14688	45.86
Test	36976	17730	47.95

- Improvement in completion rate
- However, the improvement in completion rate is 4.55% **lower** than decision threshold of 5%



Steps duration

	Start	Step 1	Step 2	Step 3	Confirmed	Total
variation						
Control	78.7	39.5	31.1	65.7	66.9	281.9
Test	92.9	39.1	40.2	67.1	77.6	316.9
Difference	14.2	-0.4	9.1	1.4	10.7	35.0

- Globally, there is no improvement on the average duration of the process, overall and by step
- However, Step 1 shows a slight duration improvement



Repetition of steps (for transaction confirmed)

variation	Total	Repetition	Repetition Rate (%)
Control	14688	5818	39.61
Test	17730	7395	41.71

- There is no improvement on the repetition of steps



The new UI does improve the completion rate of the transactions, but does not optimize the process



Test results - Hypothesis testing

Z-test: Method to determine if the differences observed in KPIs between test & control are statistically significant

Completion rate

$H_0: p_{Test} = p_{Control}$

$H_1: p_{Test} \neq p_{Control}$

With $p = \text{Number of confirmed transactions} / \text{Number of Total transactions}$

Z-test result : Z-statistic = 5.48 / p-value 4.28e-08

Since the p-value is so low, we **reject the null hypothesis (H_0)**.

The difference in completion rates between the Test and Control groups is **statistically significant**

Repetition rate =>

$H_0: r_{Test} = r_{Control}$

$H_1: r_{Test} \neq r_{Control}$

With $r = \text{Number of transactions with repeats} / \text{Number of Total transactions completed}$

Z-test result : Z-statistic = 3.82 / p-value 0.000129

Since the p-value is so low, we **reject the null hypothesis (H_0)**.

The difference in repetition rate between the Test and Control groups is **statistically significant**

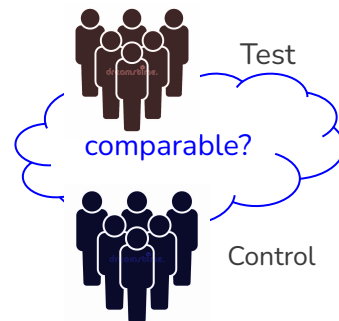
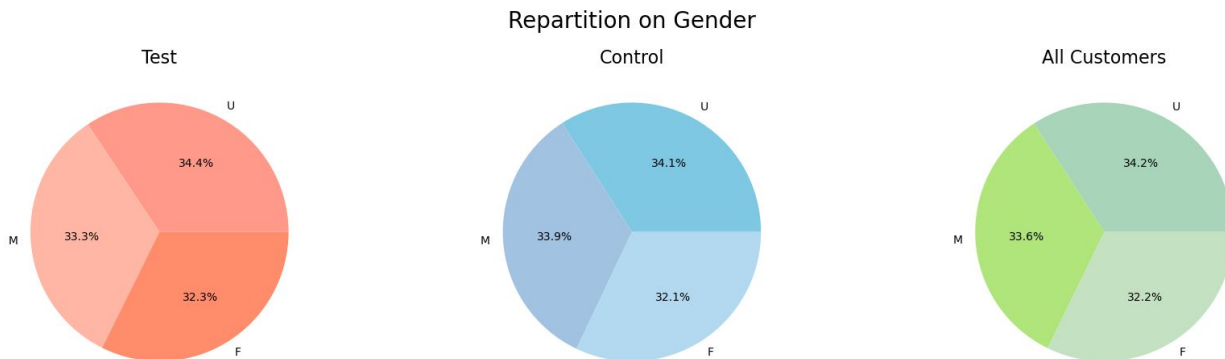
Average duration => Z-test for proportions

	Z-statistic	p-value	Statistically significant difference
Start	4.695688	2.657112e-06	YES
Step 1	-0.395825	6.922339e-01	NO
Step 2	10.243624	1.264108e-24	YES
Step 3	1.234810	2.169012e-01	NO
Confirmed	5.890224	3.856731e-09	YES
Total	8.137079	4.049287e-16	YES

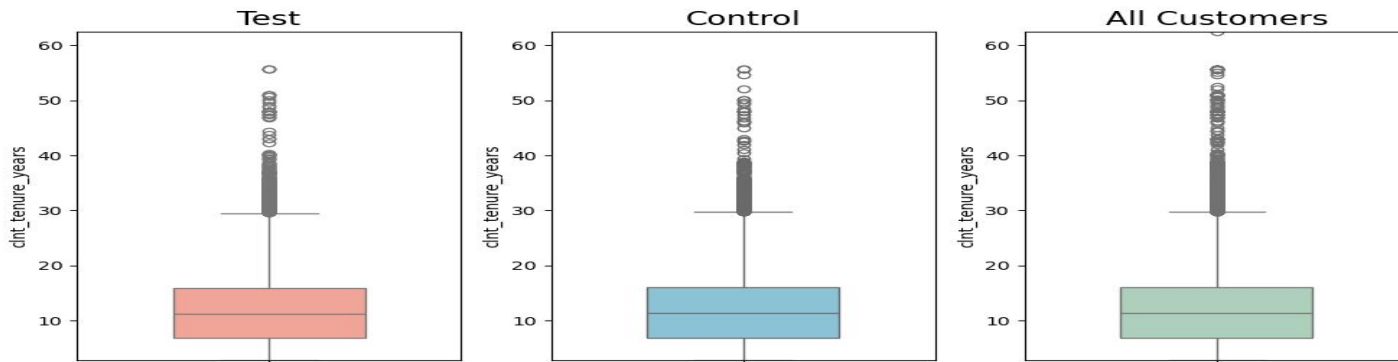


A/B Sample homogeneity - Visualizations

To further ascertain the results, we also need to make sure that the test & control populations are homogenous between themselves and the total customer population => avoid introducing bias on results from the sampling



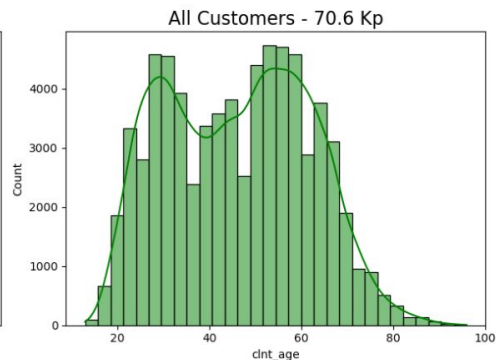
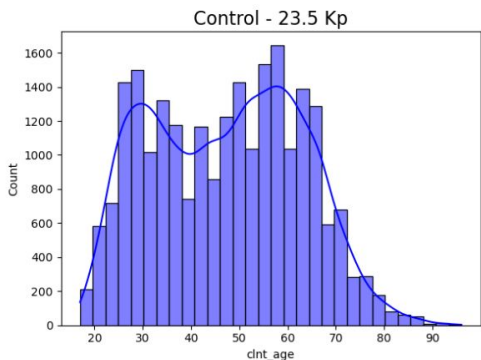
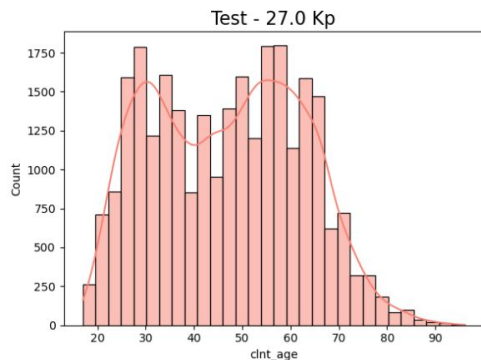
Comparison of Client Tenure in Years



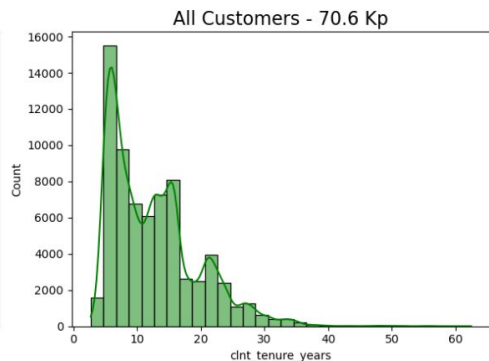
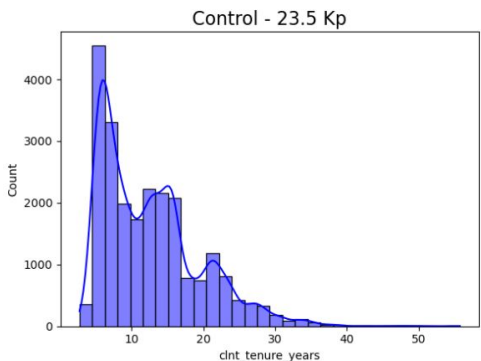
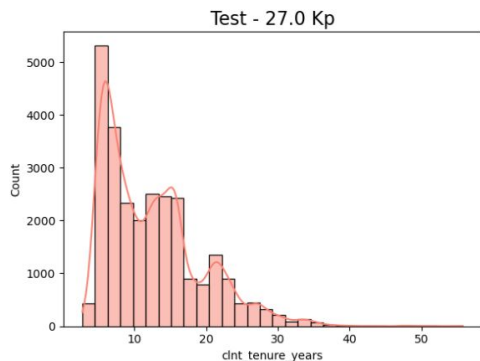


A/B Sample homogeneity - Visualizations

Customer Age distribution



Client tenure distribution





A/B Sample homogeneity - Summary Statistics

	Test						Control						All					
	Mean	Mode	Median	Std Dev	Min	Max	Mean	Mode	Median	Std Dev	Min	Max	Mean	Mode	Median	Std Dev	Min	Max
Age	46.89	31.00	47.00	15.52	17.00	96.00	47.26	58.00	48.00	15.53	17.00	96.00	46.18	57.00	47.00	15.60	13.00	96.00
Number of Accounts	2.25	2.00	2.00	0.53	1.00	7.00	2.26	2.00	2.00	0.54	2.00	7.00	2.26	2.00	2.00	0.53	1.00	8.00
Tenure (Years)	12.49	4.83	11.17	6.82	2.75	55.75	12.59	4.83	11.42	6.84	2.75	55.75	12.55	4.83	11.33	6.84	2.75	62.42
Calls last 6 months	3.06	6.00	3.00	2.19	0.00	6.00	3.13	6.00	3.00	2.18	0.00	6.00	3.38	6.00	3.00	2.24	0.00	7.00
Logons last 6 months	6.10	9.00	6.00	2.18	3.00	9.00	6.17	9.00	6.00	2.17	3.00	9.00	5.57	9.00	5.00	2.35	1.00	9.00

- From visualizations and summary statistics, we tend to conclude that the 2 samples are homogenous
- **Is it enough to full ascertain that they are?**



A/B Sample homogeneity - Hypothesis testing

ANOVA : >2 sample T-test

Method to assess the **coherence** and **homogeneity** between samples

Client Age

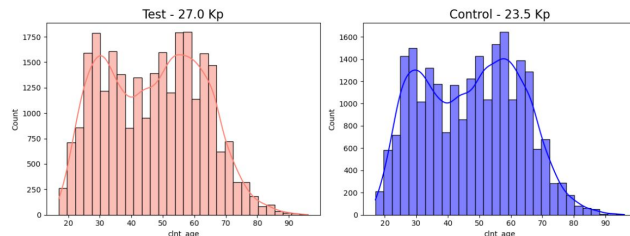
$$H_0 : \mu_{\text{Test}} = \mu_{\text{Control}} = \mu_{\text{All}}$$

$$H_1 : \mu_{\text{Test}} \neq \mu_{\text{Control}} \neq \mu_{\text{All}}$$

Anova test result : F-statistic = 50.38 / p-value = 1.34e-22

Since the p-value is **much smaller than 0.05**, we **reject the null hypothesis**

There is a **statistically significant difference** in the mean client age between the Test, Control, and All groups.



Client Tenure

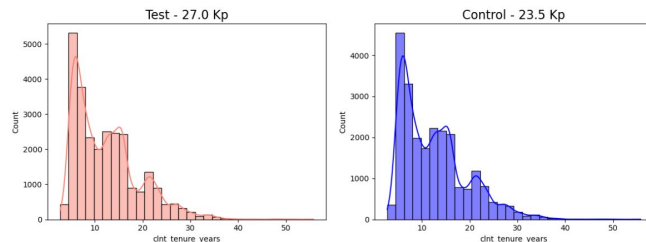
$$H_0 : \mu_{\text{Test}} = \mu_{\text{Control}} = \mu_{\text{All}}$$

$$H_1 : \mu_{\text{Test}} \neq \mu_{\text{Control}} \neq \mu_{\text{All}}$$

Anova test result : F-statistic 1.495 / p-value 0.224

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

There is **no statistically significant difference** in the mean client tenure years between the Test, Control, and All groups.





Thank you !

Vanguard CX team
Benoit-Emmanuel