

Vanguard ABTESTING



MARÍA MARTÍNEZ
SANTIAGO LARREA

INDEX

- 1 - Introduction
- 2 - Data Cleaning
- 3 - EDA
- 4 - Hypothesis Testing
- 5 - Client Behaviour Visualizations
- 6 - KPI Analysis



INTRODUCTION

Context

Vanguard tested a redesigned interface with in-context prompts against its traditional online process to enhance client experience.

Objective

The goal is to determine whether the new design increases process completion rates compared to the existing interface.

Tools used

The tools used were Python (`pandas`, `slicy.stats` or `math`) and Tableau.

DATA CLEANING



1

Clean the four initial datasets, remove spaces, format data types, and remove duplicates. Fill null values in demo as 'No participation'

2

Concatenate the final web CSVs and merge the demo datasets into client ID to find out which variant they belong to (Test, control, or No participation).

3

Divide the demo dataset into Test, Control, and No participation to perform EDA based on the variant.

4

Concatenate the control and test files in a final DF to be able to draw general conclusions about only the participants in the experiment and download the 4 CSV

EDA



Data description

- Scope: all variants combined; sources = demographics, experiment assignment, web events.
- Sample: 70,609 clients, 96,773 visits, 744,641 events.
- Journey: start → step_1 → step_2 → step_3 → confirm.

Univariate

- Client completion rate: 67.70% (47,800 / 70,609).
- Visit completion rate: 54.48%.
- Time-to-complete: median 4.12 min (mean 6.39; max 300.25).
- User mix: Gender — U 34.2%, M 33.6%, F 32.2%.
- Age — 50–59 22.0%, <30 18.8%, 30–39 18.6%, 40–49 17.6%, 60–69 16.9%.

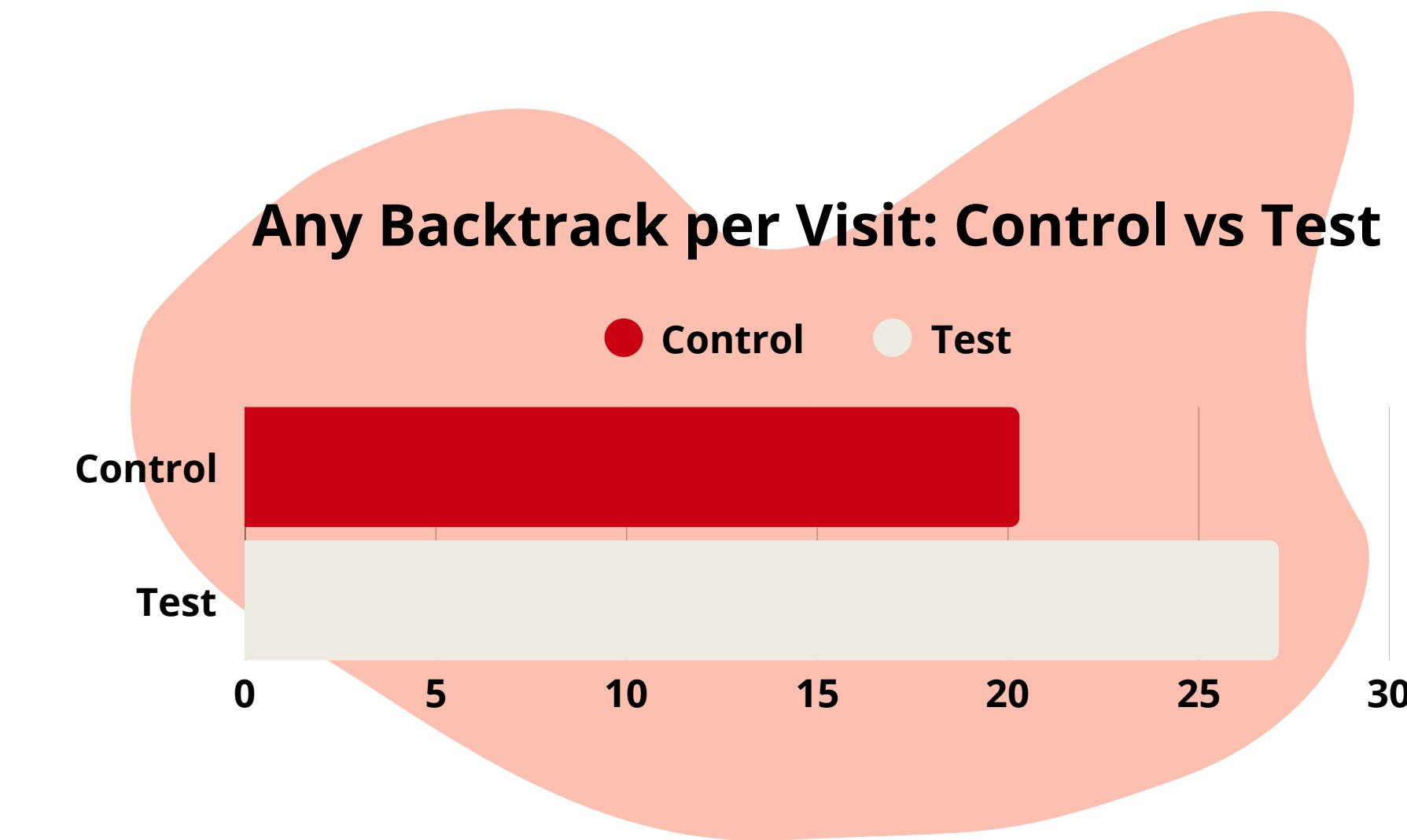
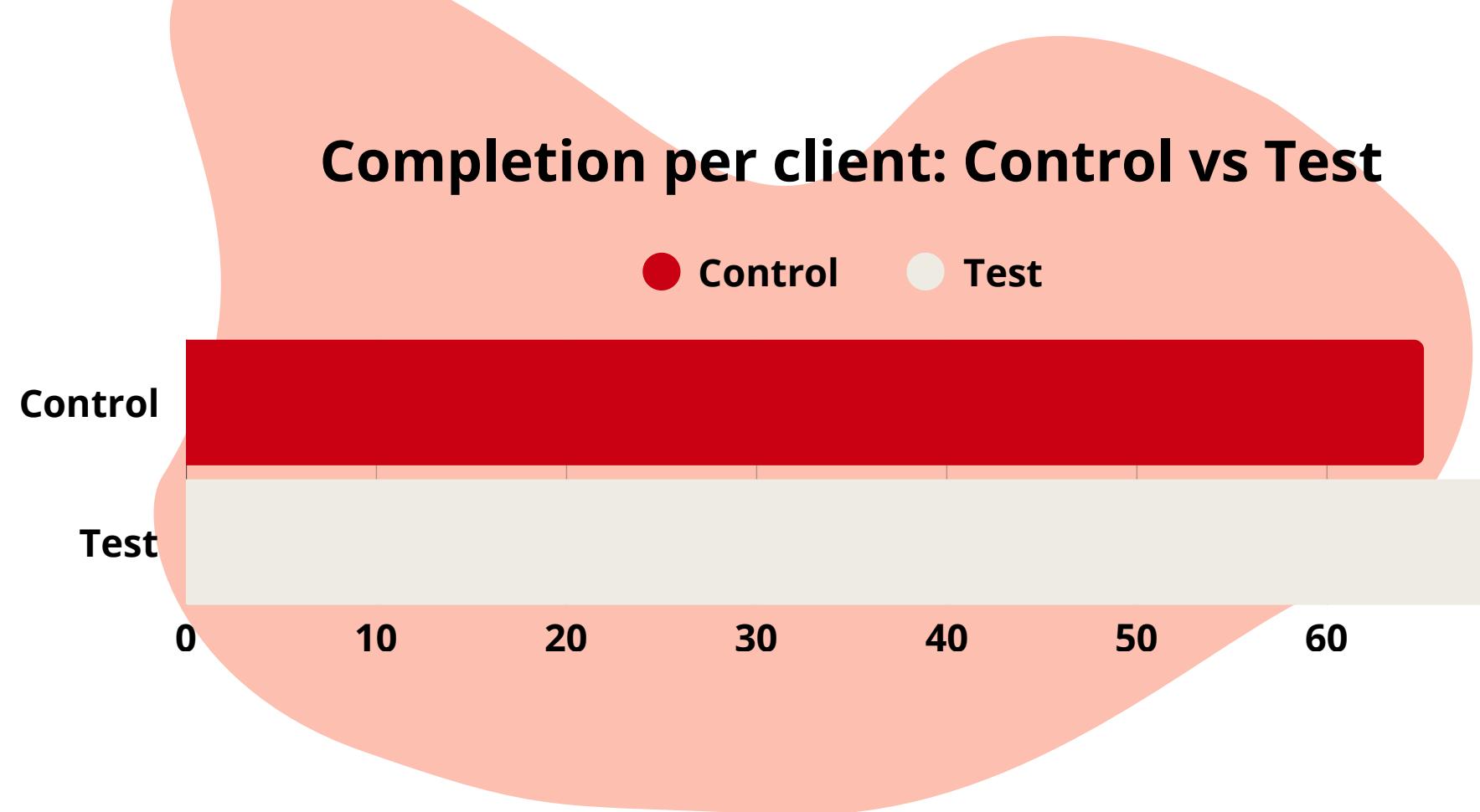
Bivariate

- Funnel drop-offs (vs previous step): start→1 11.23%, 1→2 7.94%, 2→3 6.30%, 3→confirm 10.84%.
- Conversion by age band: <30 71.41%, 30–39 68.93%, 40–49 69.23%, 50–59 67.61%, 60–69 65.34%, 70+ 54.70%.
- Conversion by tenure: <2 yrs 70.37%, 2–5 yrs 68.67%, 5+ yrs 67.49%.
- First-visit dynamics: first-visit_conv 54.93% vs later visits 36.43%; 79.97% of all completions happen on the first visit.

Key insights

- Strong overall completion (67.7%), with biggest leaks at entry (start→1) and final confirmation (3→confirm).
- Speed matters: median TTC ≈ 4.1 min; slowest transitions at 2→3 and 3→confirm.
- Younger and newer-tenure clients convert more; conversion declines with age and tenure.
- First visit is critical (≈80% of completions) → prioritize first-session guidance and friction removal.

HYPOTHESIS TESTING

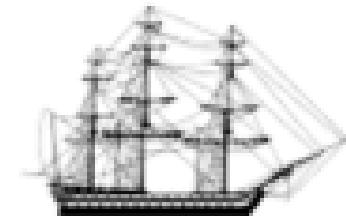


CLIENT BEHAVIOUR VISUALS

Variation

- (Todo)
- Control
- Test

General analysis



Avg age per group

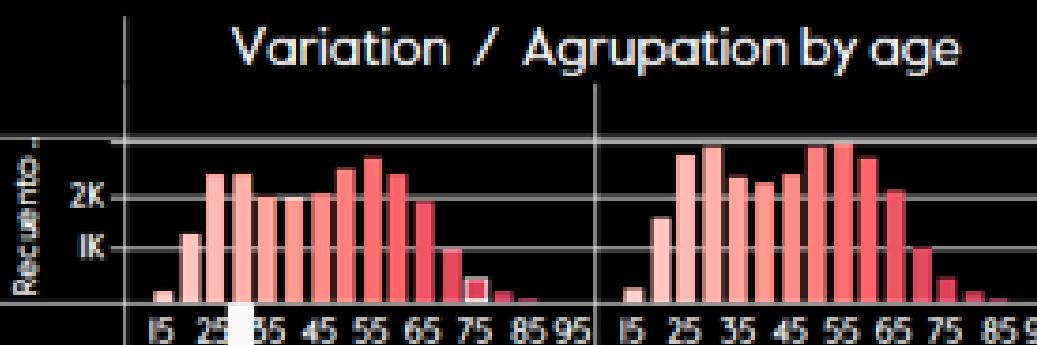
Variation	Mean ..	Media ..	Max ..	Min ..	P25 AG ..	P75 AG ..
Control	47.26	48.00	96.00	17.00	33.00	60.00
Test	46.89	47.00	96.00	17.00	33.00	59.00

Comparison by 4 main metrics

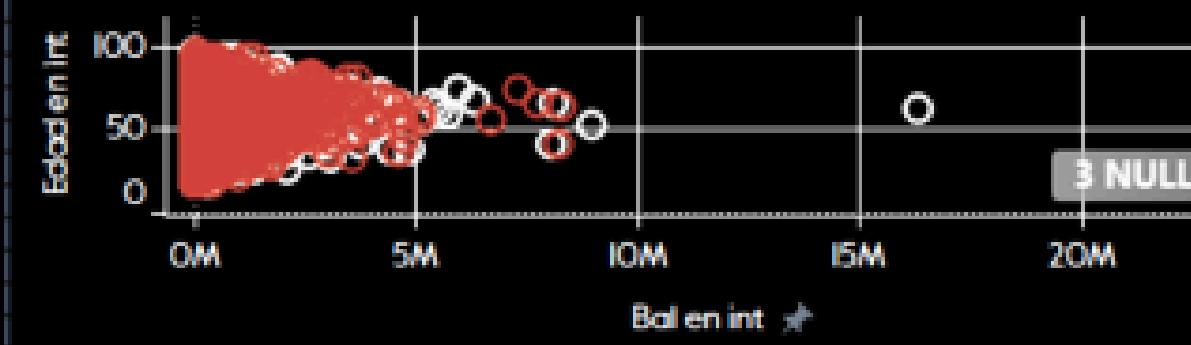
Variation	Mean age	Prom. Logs ..	Prom. Calls ..	Prom. Bal e ..
Control	47.26	6	3	150.147
Test	46.89	6	3	148.962

Comparison age by Control/Test

Variation / Agrupation by age

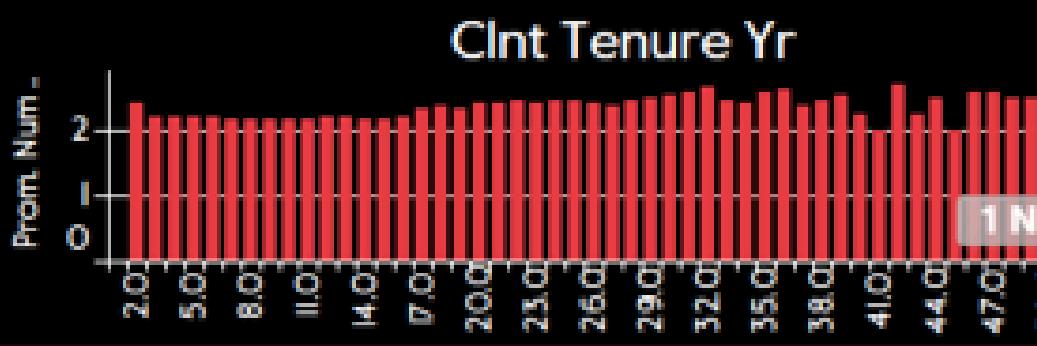


Relation between age and balance

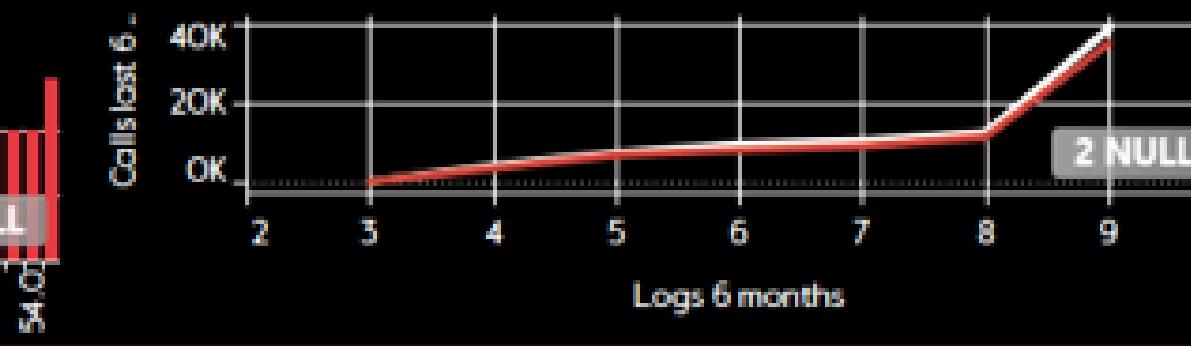


Clt Tenure Year by Num Acts

Clt Tenure Yr

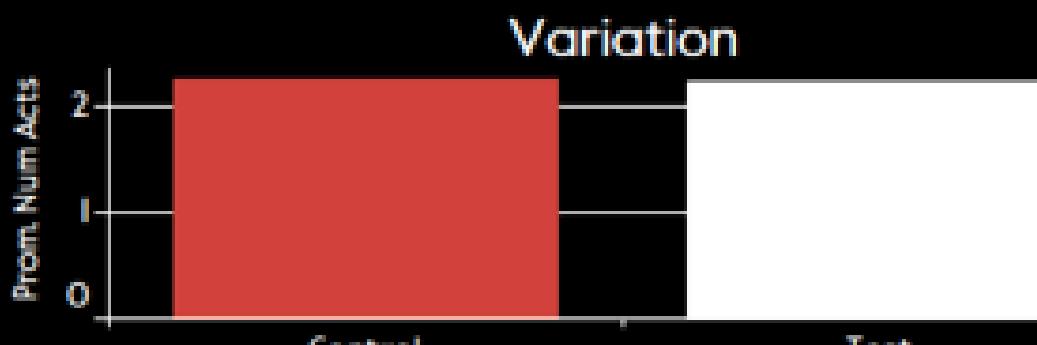


Relation between logs and calls



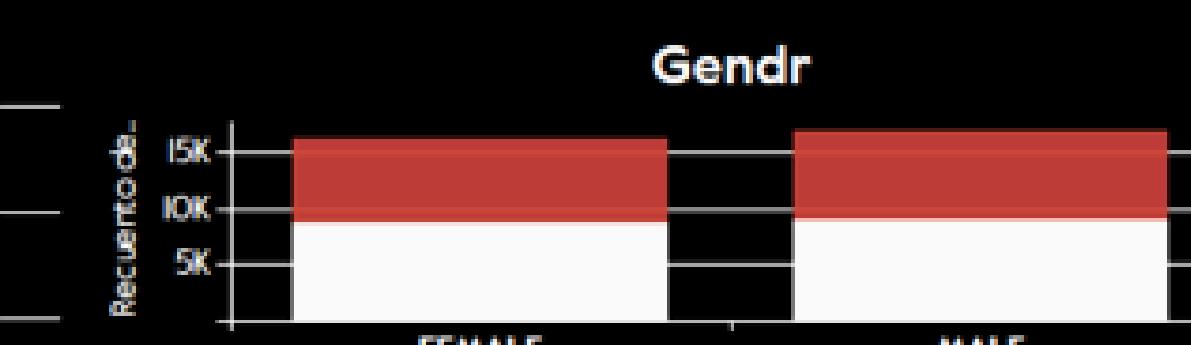
Avg accounts per group

Variation



Comparison by Gender

Gendr



CHALLENGES & LEARNINGS



CHALLENGES

- **Data integration:** combining multiple sources (visits, events, errors) and keeping joins consistent (step, group, client_id).
- **Complex KPIs:** some metrics (backtracks, completion, reach) required custom calculations and intermediate tables.
- **Visualization alignment:** translating outputs from notebook → Tableau while preserving result accuracy.
- **Stakeholder communication:** adapting technical metrics (p-values, statistical tests) into intuitive visuals.
- **Tool limitations:** handling differences between Python, Tableau, etc. (e.g., percentage formats, dataset joins).

PROBLEMS

- **Preprocessing is key:** early data cleaning and normalization avoids downstream errors.
- **Validation matters:** always cross-check numbers between the notebook and Tableau before publishing dashboards.
- **Storytelling first:** selecting only the key visuals (Completion, Backtracks, Error rates) makes insights clear.
- **Visualization design:** consistent brand colors and typography enhance communication.
- **Iterative approach:** build step-by-step (KPI1 → KPI4), validate, then integrate into the final dashboard.

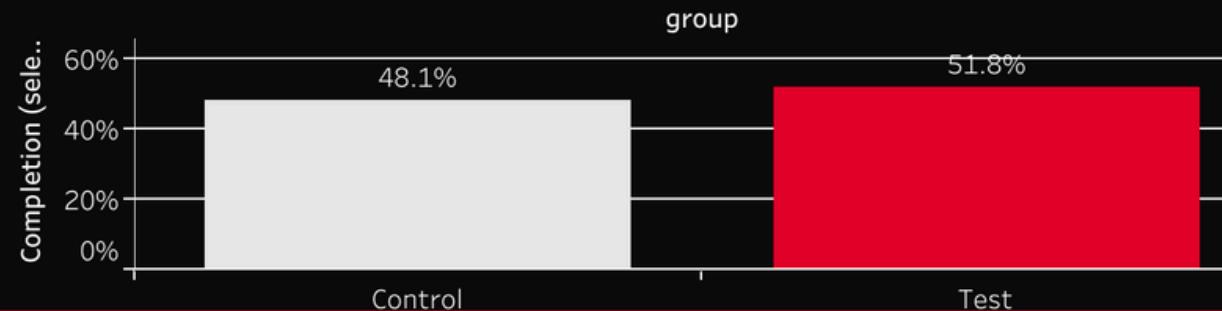
KPI ANALYSIS

KPI analysis

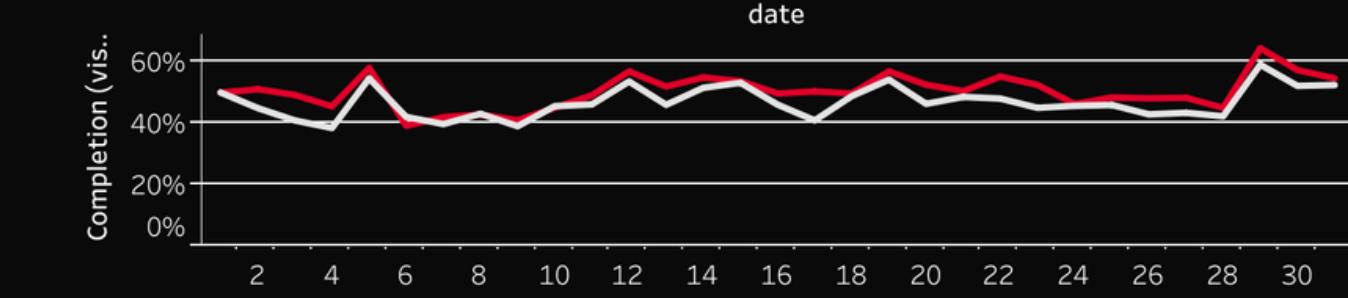


Vanguard®

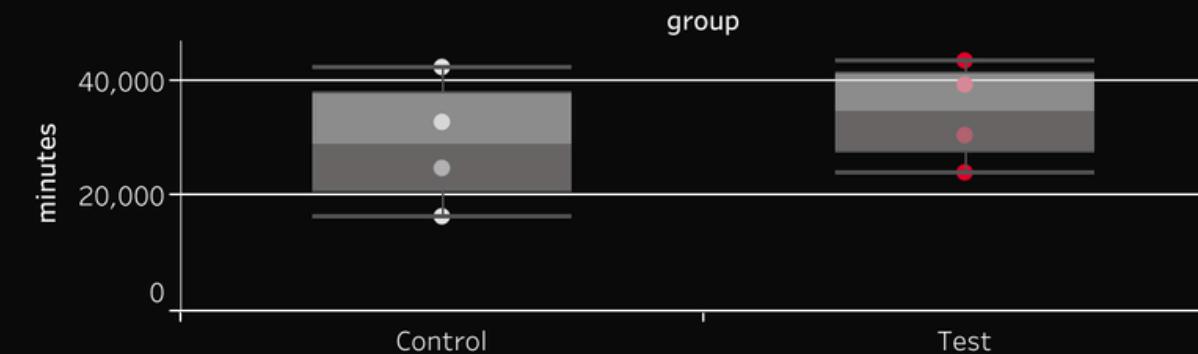
Completion Rate: Test vs Control



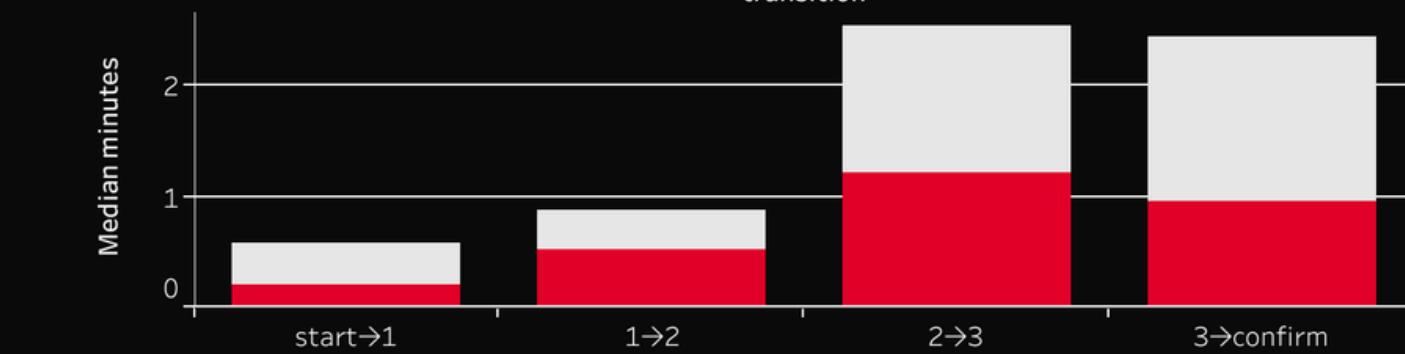
Daily Evolution of Completion Rate by Group



Distribution of Visit Duration (minutes) by Group



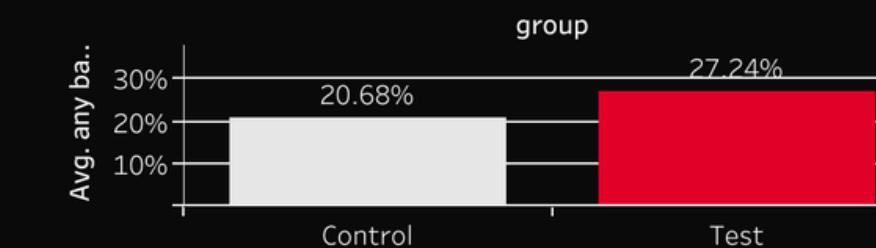
Median Time per Step Transition (minutes)



Error Rate Heatmap by Step (Control vs Test)



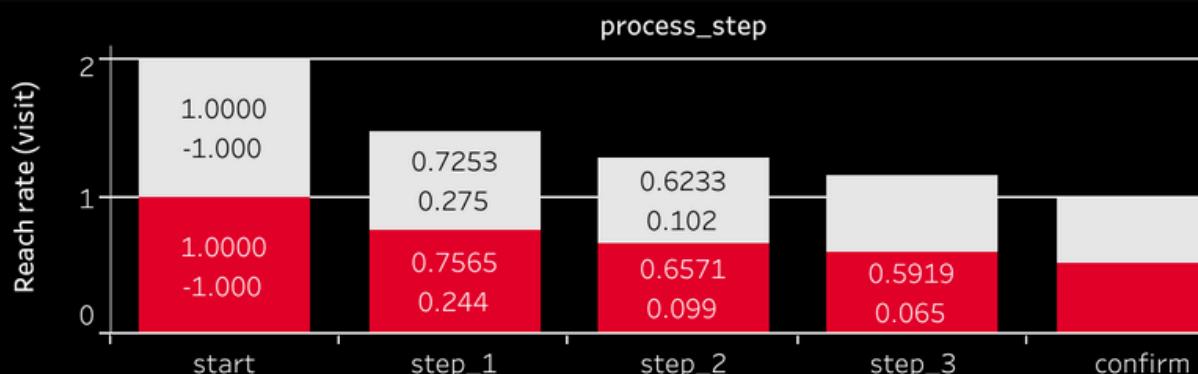
Share of Visits with Any Backtrack (Control vs Test)



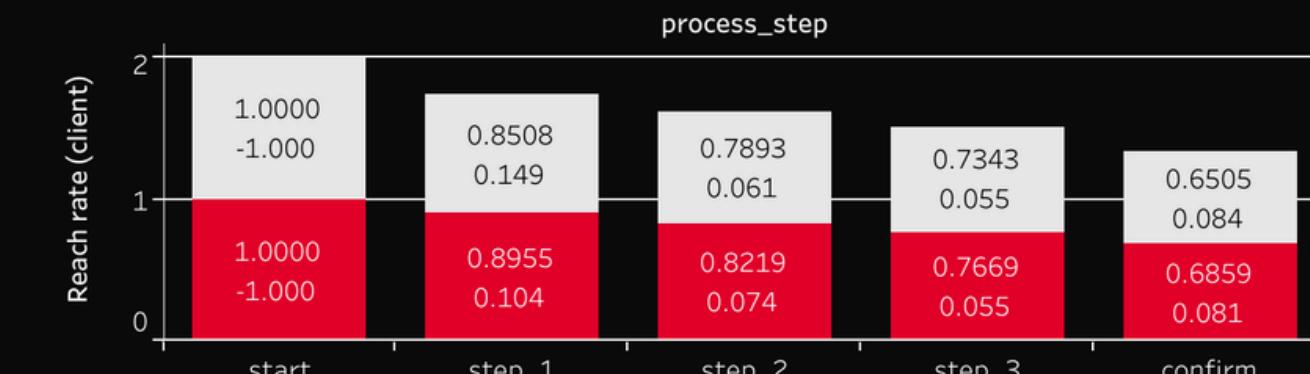
Step Error Rate Comparison: Control vs Test



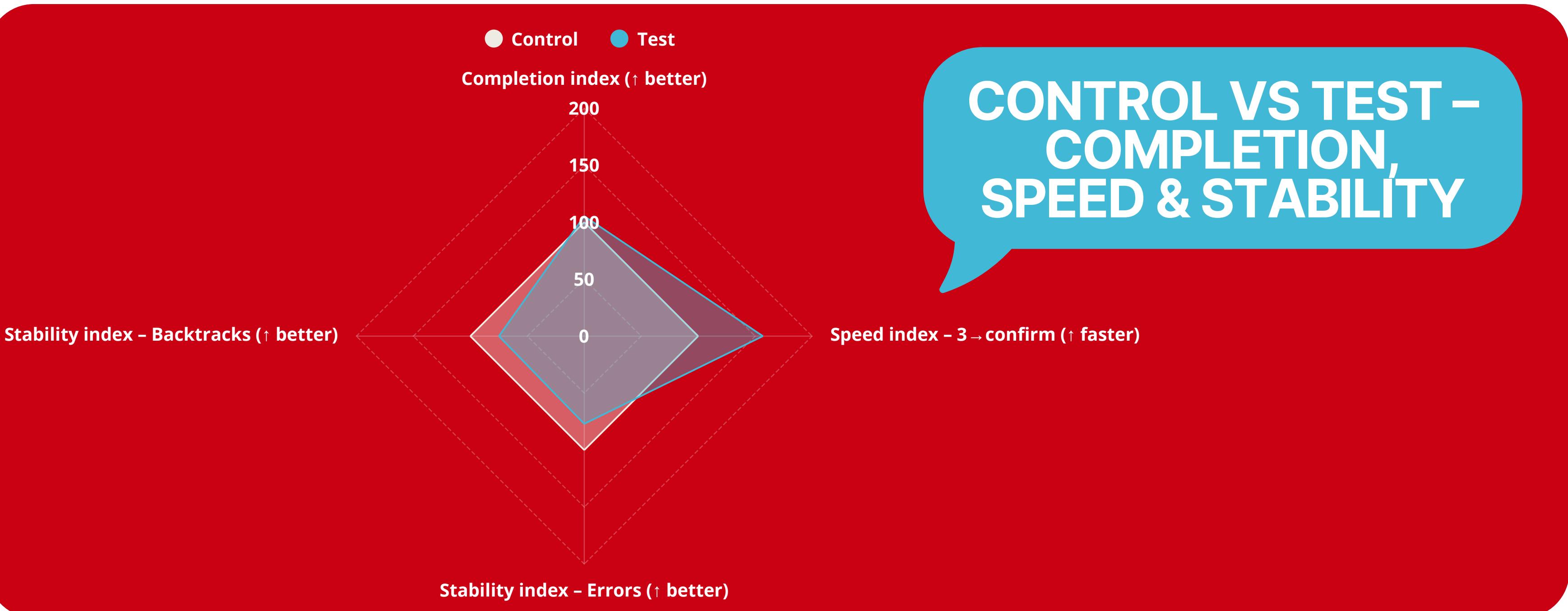
Step Reach Rate (Visits) and Drop-offs



Step Reach Rate (Clients) and Drop-offs

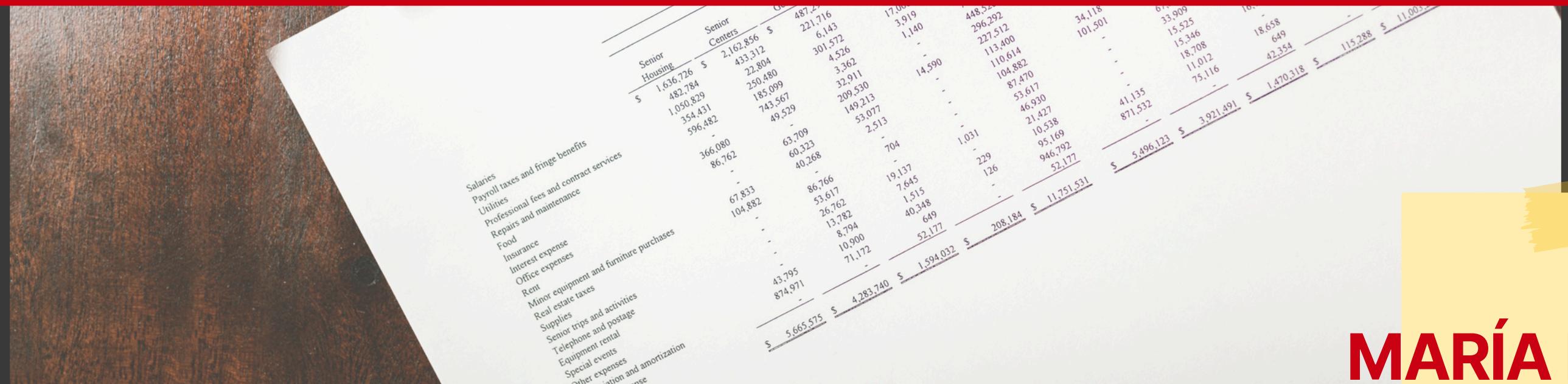


FINAL CONCLUSIONS





THANKYOU!



	Senior Housing	Senior Centers	General	Other
Salaries	\$ 1,636,726	\$ 2,162,856	\$ 221,716	\$ 487,270
Payroll taxes and fringe benefits	482,784	433,312	6,143	3,919
Utilities	1,050,829	22,804	301,572	1,140
Professional fees and contract services	354,431	250,480	4,526	448,541
Repairs and maintenance	596,482	185,099	3,362	296,292
Food	743,567	49,529	209,530	227,512
Insurance	366,080	63,709	53,077	113,400
Interest expense	86,762	60,123	2,513	110,614
Office expenses	67,833	40,268	149,213	104,852
Rent	104,882	86,766	14,590	87,470
Minor equipment and furniture purchases	-	53,617	14,590	53,617
Real estate taxes	-	26,762	-	46,930
Supplies	-	13,782	-	21,427
Senior trips and activities	-	8,794	-	10,538
Telephone and postage	-	10,900	-	95,169
Equipment rental	-	52,177	-	946,792
Special events	-	649	-	52,177
Other expenses	-	-	-	41,135
Depreciation and amortization	-	-	-	871,532
Total	\$ 5,665,575	\$ 4,283,740	\$ 1,594,032	\$ 1,470,318
				\$ 115,288
				\$ 11,003

**MARÍA MARTÍNEZ
SANTIAGO LARREA**