🧪 A/B Analysis Report — Vanguard Digital Redesign

---

1. Introduction and Project Context

Vanguard, a leading investment firm, has redesigned its digital experience with the goal of improving the conversion process of new financial products through its web platform. To validate the impact of this redesign, an A/B experiment was implemented: one group of users (control) continued to see the original design, while another group (test) accessed the new design.
The company's expectation was clear: to increase the completion rate by at least 5% as the minimum threshold to consider the redesign successful.

2. Data Description

For the analysis, four databases were used:

- **df_final_demo**: demographic information of the users.
- **df_exp_cli**: assignment of each client to the control or test group.
- **df_web_1** and **df_web_2**: web navigation logs that allow reconstruction of user steps in the digital funnel.
  The data were integrated and cleaned to transform navigation into processable sequences, a key step to evaluate friction, times, and errors.

3. Analysis Objective

Determine whether the digital redesign has produced statistically significant improvements in:

- Funnel completion rate (full conversion).
- Error rate during the process.
- Temporal efficiency of the digital journey.
- Impact on different user segments (age, tenure, balance).

4. Exploratory Analysis

User demographic profile

- **Age**: Two usage peaks around 35 and 57 years old, indicating a user base familiar with digital tools.

- **Client tenure**: The central 50% of users have between 8 and 16 years as clients, suggesting an experienced and loyal base.

- **Balance**: Although there are clients with million-dollar balances (outliers), most have between 40,000 and 140,000 USD, reflecting an average saving profile.

5. Error Definition and Key Metrics

The navigation funnel was reconstructed for each user and technical and friction errors were defined as:

- **Step repetition**: users stuck on the same stage (e.g., multiple step_2).

- **Backtracking**: inverse navigation (e.g., from step_3 to step_1).

- **Zero duration between steps**: indicates anomalous records or system errors.

- **Abandonment**: users who do not reach the final confirm stage.

Metrics analyzed

- **Completion rate**: percentage of users who reach the end of the funnel (confirm).

- **Error rate**: percentage of users with at least one anomalous event.

- **Process duration**: average time to complete the funnel.

6. Hypothesis Testing

Statistical tests were conducted to evaluate differences between groups:

- Z-test for proportions for completion and error rates.

- Comparisons of mean times with t-tests or non-parametric analysis according to distribution.

- Segmented evaluations by age, balance, and tenure to detect significant variations.

7. Results and Conclusions

Main results

- **Completion rate:**
    - Control: 74.8%
    - Test: 77.6%
    - Absolute improvement: +2.8% (statistically significant), although below the target threshold of 5%.
- **Error rate:**
    - Lower in the test group for all defined types, especially backtracking and abandonment.
- **Process duration:**
    - Test group: average of 4.1 min
    - Control group: 4.3 min
    - The difference was not always significant, but suggests an efficiency improvement.

8. Final Conclusion and Recommendation

The digital redesign (Test group) has shown statistically significant improvements in key process aspects, but does not meet all of Vanguard's defined operational effectiveness criteria. Below is the final balance between the two versions:

✅ **Clear advantages of the Test group:**

- Higher completion rate: 69.3% vs. 65.6% (significant difference, $p < 0.001$).
- Lower overall technical error rate: 0.076% in Test vs. 0.193% in Control ($p < 0.001$).
- Substantial improvement in critical errors at the "confirm" stage:
    - Error reduction from 0.577% to 0.066% (+0.51 pp improvement, with 95% CI largely above the minimum threshold).

⚠️ **Limitations of the Test group:**

- Worse first-attempt performance: 43.7% vs. 47.4%.
    - Indicates more friction and less initial clarity in the new design.
- Slightly higher total average time ($p < 0.001$), though medians and percentiles are similar.
- Worse times in key steps like step_1 (+5 s) and confirm (+23 s), which may reflect bottlenecks.

🧠 **Business Considerations:**

- Several improvements are statistically significant, but do not reach the company's minimum required threshold of 5%.
    - Example: the +3.7 pp improvement in completeness, although significant, does not exceed the expected threshold.
    - However, the improvement in critical errors does exceed the operational impact threshold.

🧭 **Final Recommendation**

Adopt the Test redesign as the new standard, but with specific adjustments, especially in the step_1 and confirm steps where:

- Although there is greater stability (fewer errors),
- There is more friction (more time, lower first-attempt success rate).