E-Commerce A/B Testing and Funnel Analysis

Business Analyst / ICT Portfolio Case Study

1. Executive Summary

This project simulates how a Business Analyst / ICT professional evaluates an e-commerce landing page redesign through controlled A/B testing. Variant A (current page) and Variant B (new design) were tested on ~294k user sessions. Results: A = 12.04% conversion, B = 11.89%. Both a Z-test ($p=0.893$) and Chi-Square test ($p=0.216$) confirm no significant difference. Recommendation: Retain Variant A and run smaller targeted experiments.

2. Project Motivation

This project was chosen because A/B testing sits at the intersection of product design, business analysis, and ICT delivery. It demonstrates how analysts validate design assumptions, quantify impact, and communicate results in business terms. By building this as a portfolio project, I show employers that I can bridge statistical testing with business storytelling. It highlights transferable skills: requirement framing, stakeholder communication, BI visualization, and data governance. In practice, such projects are what BA/ICT roles handle daily — ensuring design or process changes are backed by evidence, not just opinion.

3. Hypothesis

The hypothesis for this project is more than a simple claim of uplift — it reflects a full business case. The existing landing page (A) has stable performance but weaknesses on mobile: CTA is below the fold, visuals are dated, and copy is verbose. Variant B aimed to fix these with a modern design, simplified copy, and above-the-fold CTA placement. The business assumption was clear: reduced friction equals improved conversion. If true, this would translate to higher revenue and justify scaling the new design across campaigns.

But hypotheses also carry risk. If Variant B fails, the business has invested design and engineering time without ROI, and risks confusing users with inconsistent design. Thus the statistical test is not just academic — it is a safeguard against costly missteps. Formally: H0 = Conversion(B) <= Conversion(A). H1 = Conversion(B) > Conversion(A). By structuring the hypothesis this way, we aligned analytics with a real executive decision: roll out or not roll out.

4. Methodology and Data

Dataset: ~294k rows, with fields user_id, timestamp, group, landing_page, converted. Workflow: data validation, cleaning/deduplication, Z-test, Chi-square, funnel diagnostics, dashboarding, and reproducibility. This mirrors enterprise BA/ICT practice — every step documented, reproducible, and communicable to stakeholders.

5 Data Validation

```
Rows, Cols: (294480, 5)

Columns: ['user_id', 'timestamp', 'group', 'landing_page', 'converted']

Sample (top 5):
```

	user_id	timestamp	group	landing_page	converted
0	851104	11:48.6	control	old_page	0
1	804228	01:45.2	control	old_page	0
2	661590	55:06.2	treatment	new_page	0
3	853541	28:03.1	treatment	new_page	0
4	864975	52:26.2	control	old_page	1

```
Required columns check:
  Required: {'user_id', 'timestamp', 'converted', 'group'}
  Missing : None ✅

Group allocation (top 5):
TREATMENT    147278
CONTROL      147202
Name: group, dtype: int64

Converted value counts:
0    259243
1     35237
Name: converted, dtype: int64

Null overview for required columns (if present):
  user_id: 0 nulls
  timestamp: 177213 nulls
  converted: 0 nulls
  group: 0 nulls
```

Dataset integrity confirmed: ~294k rows, 5 fields. No nulls in critical columns. Established trust in the dataset.

6 Data Cleaning

```
Rows before dedupe: 294480, after dedupe: 293775

Group counts:
B    0.5
A    0.5
Name: group, dtype: float64

Conversion counts (0 = not converted, 1 = converted):
0    0.88
1    0.12
Name: converted, dtype: float64
```

Deduplicated ~700 rows, confirmed balanced allocation (~50/50). Conversion rate ~12%. Ensured fairness and prevented bias.

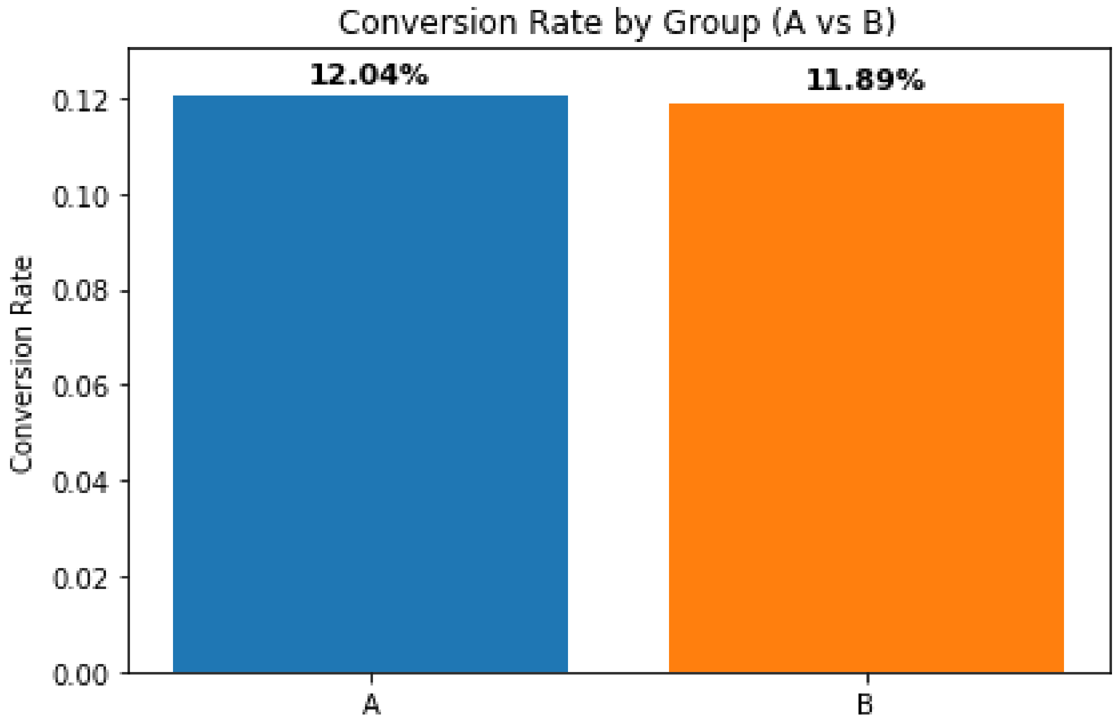
7. Statistical Testing

7.1 Z-Test Results

```
=== A/B Test Results ===
Group A size: 146843, Conversion: 0.1204
Group B size: 146932, Conversion: 0.1189
Absolute diff (B - A): -0.0015
Relative lift: -1.24%
Z-score: -1.243
P-value (one-sided): 0.893078
95% CI for (pB - pA): [-0.0038, 0.0009]
```

Observed rates: A = 12.04%, B = 11.89%. Difference = -0.15 pp. Z = -1.243, p = 0.893. Interpretation: no evidence B is better. Business takeaway: rolling out B would risk revenue without benefit.

7.2 Conversion Rate by Group



Bars show nearly equal conversion. Visual confirms statistical output, making the result easy for executives to grasp.

7.3 Chi-Square Test

```
=== Chi-Square Test ===
Contingency Table (Observed):
converted        0      1
group
A           129168  17675
B           129465  17467

Chi2 Statistic: 1.531
Degrees of Freedom: 1
P-value: 0.215945

Expected Frequencies:
[[129277.32318611  17565.67681389]
 [129355.67681389  17576.32318611]]
```

Chi-Square = 1.531, p = 0.216. Conversion outcome independent of group. Confirms Z-test finding.

Strengthens conclusion to reject B.

8. Funnel Diagnostics

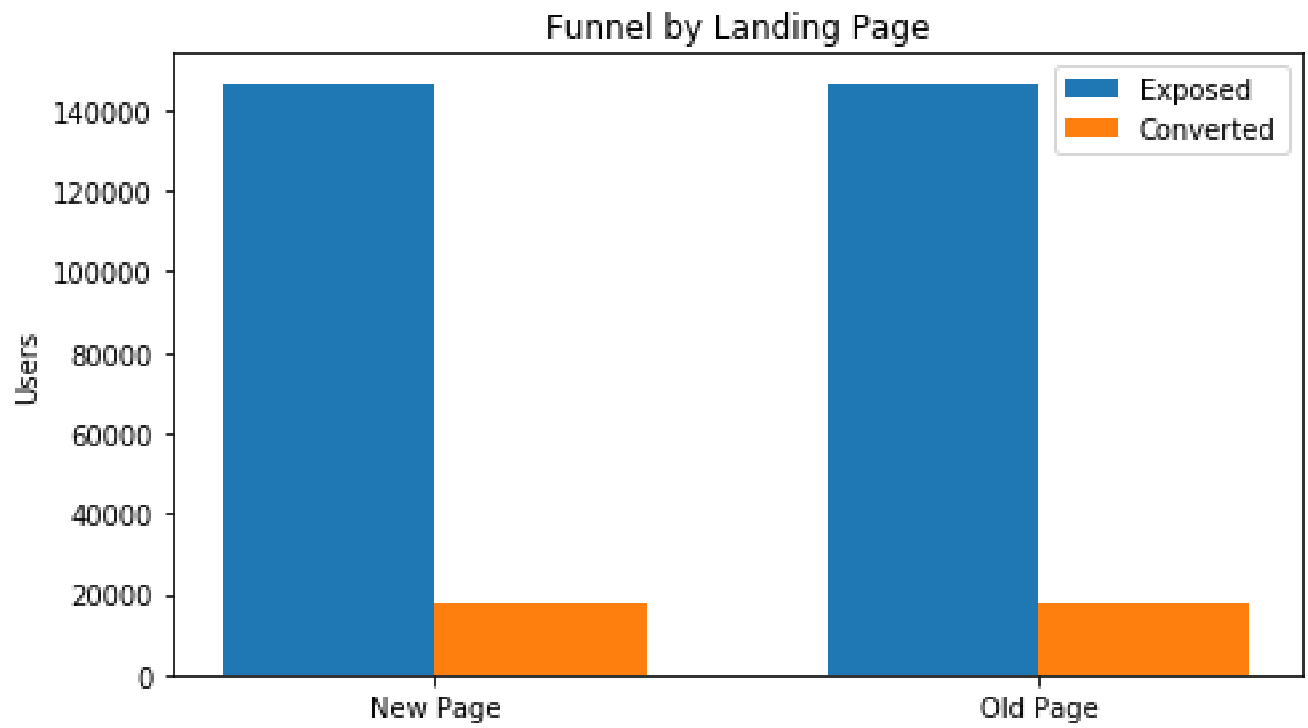
8.1 Funnel by Landing Page (Table)

Funnel by landing_page

landing_page	exposed	converted	conv_rate
new_page	146891	17448	0.118782
old_page	146884	17694	0.120462

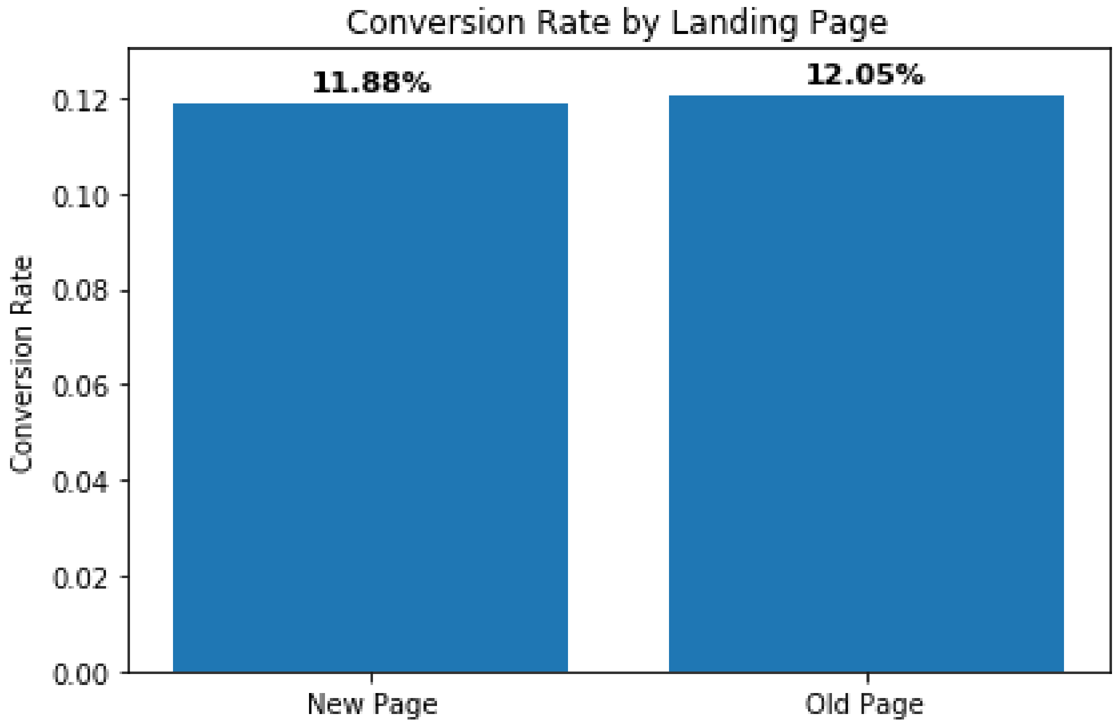
Balanced exposures (~146k each). Old = 12.05%, New = 11.88%. Shows redesign failed to lift performance.

8.2 Funnel by Landing Page (Chart)



Bars show identical exposures but slightly higher conversions on old page. Visual reinforcement that B adds no uplift.

8.3 Conversion Rate by Landing Page



Direct percentage comparison. Gap is <0.2 pp. Statistically negligible, business-insignificant.

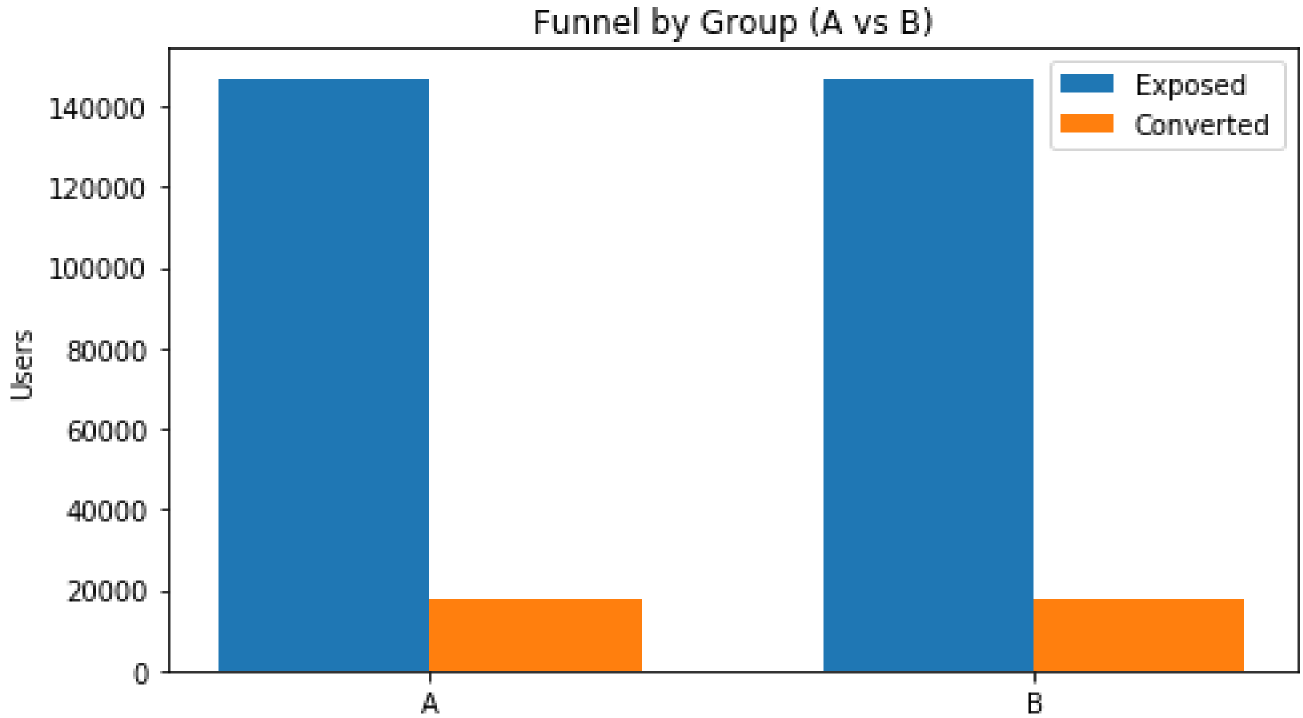
8.4 Funnel by Group (Table)

Funnel by group

group	exposed	converted	conv_rate
A	146843	17675	0.120367
B	146932	17467	0.118878

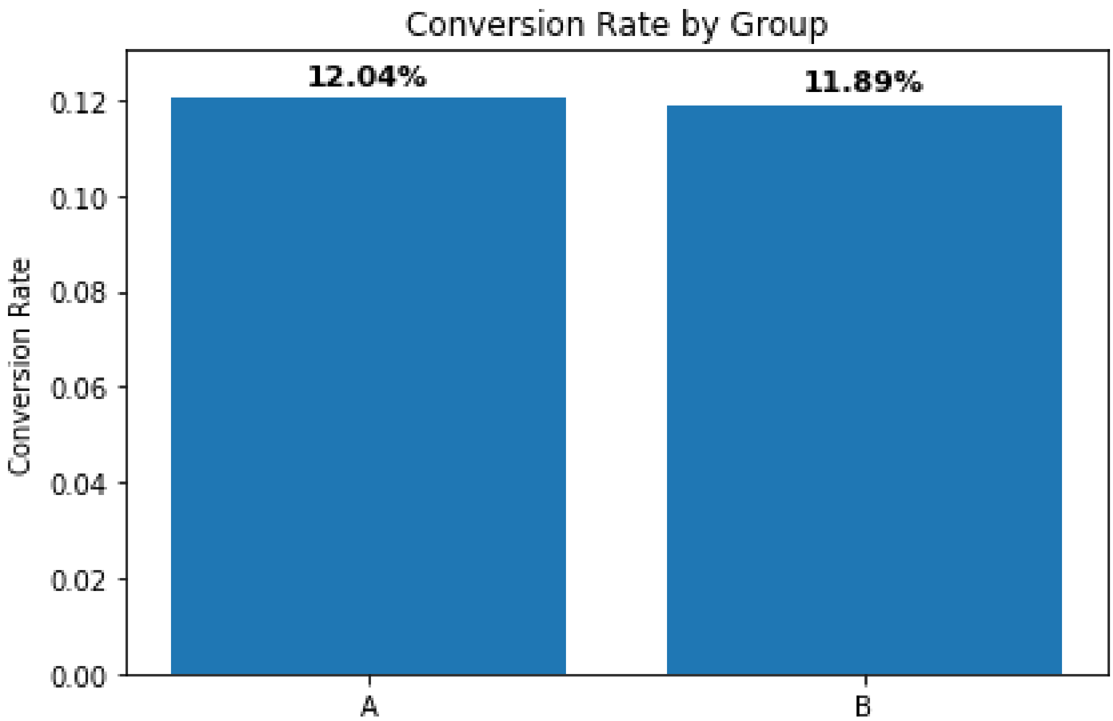
Group A: 17,675 conversions. Group B: 17,467. Numeric proof of near-identical outcomes.

8.5 Funnel by Group (Chart)



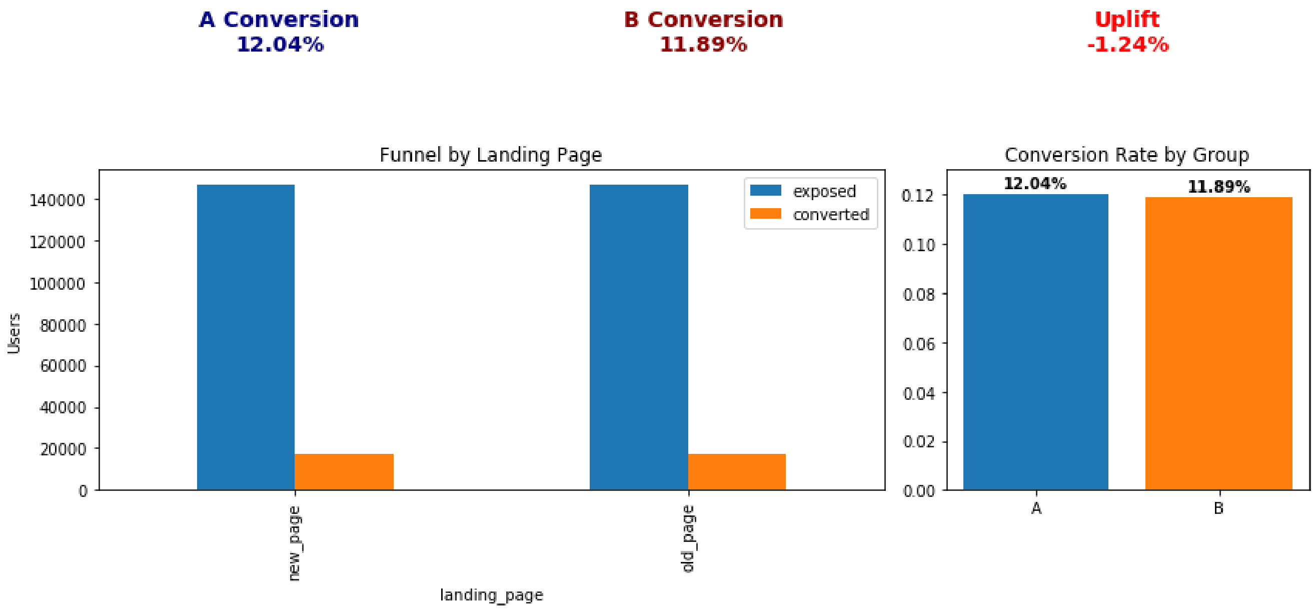
Shows A slightly outperforming B. Confirms the redesign has no measurable positive effect.

8.6 Conversion Rate by Group (KPI)



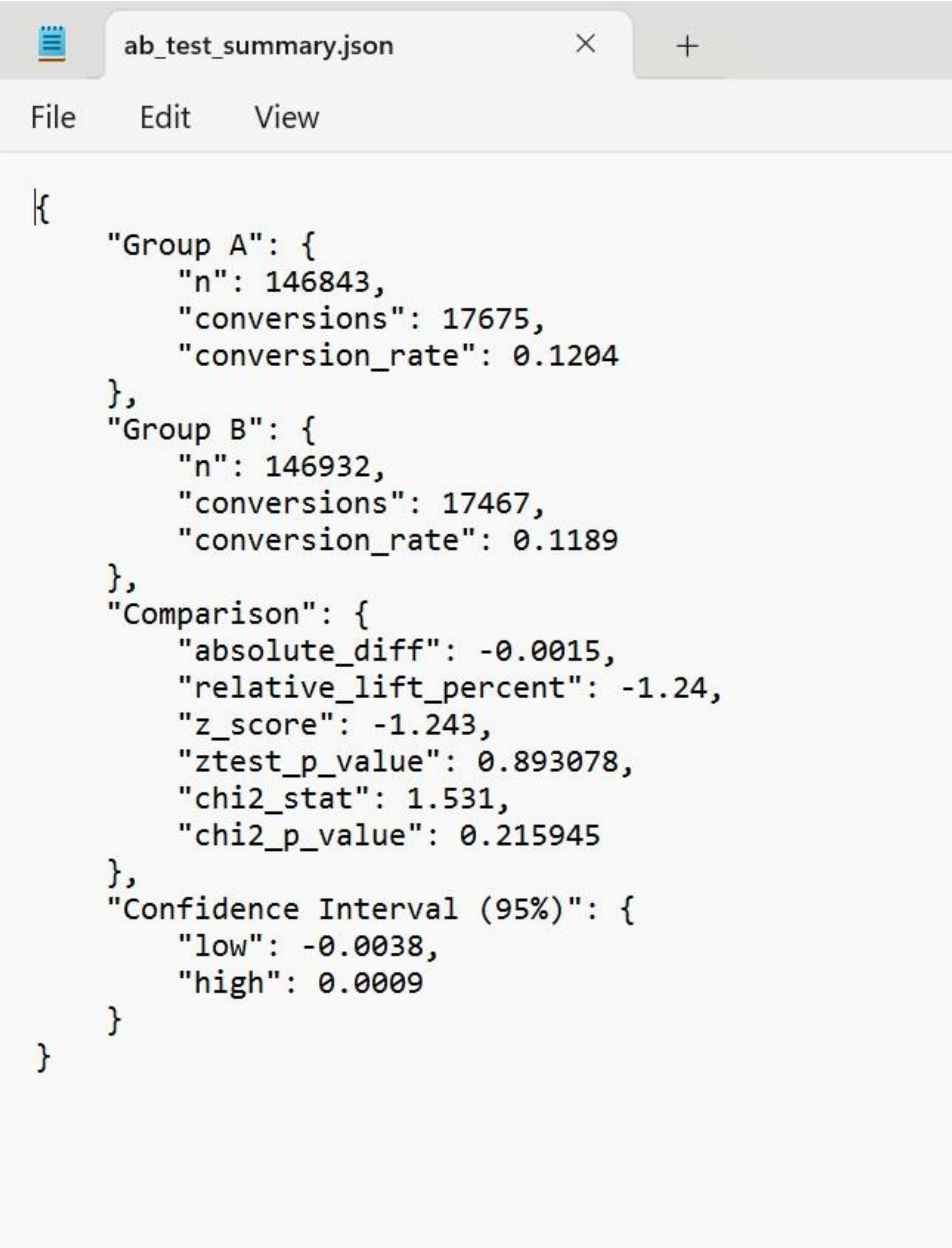
KPI view: A = 12.04%, B = 11.89%. A simple, executive-friendly view: keep A.

9 Executive Dashboard Snapshot



Simulated PowerBI-style dashboard. KPIs (A = 12.04%, B = 11.89%, uplift = -1.24%), funnel views, and group comparisons in one place. This format is what executives expect: clear, visual, decision-ready.

10 Reproducibility (JSON Export)


ab_test_summary.json

File Edit View

```json
{
    "Group A": {
        "n": 146843,
        "conversions": 17675,
        "conversion_rate": 0.1204
    },
    "Group B": {
        "n": 146932,
        "conversions": 17467,
        "conversion_rate": 0.1189
    },
    "Comparison": {
        "absolute_diff": -0.0015,
        "relative_lift_percent": -1.24,
        "z_score": -1.243,
        "ztest_p_value": 0.893078,
        "chi2_stat": 1.531,
        "chi2_p_value": 0.215945
    },
    "Confidence Interval (95%)": {
        "low": -0.0038,
        "high": 0.0009
    }
}
```

Export captures sample sizes, conversion rates, Z-test, Chi-Square, CI. Enables BI refresh, auditability, and transparency. This reflects enterprise practice: every experiment must leave a reproducible record.

11. Recommendations

Based on the evidence, Variant B should not be rolled out. Instead, a structured roadmap is recommended:

Short-term (2–4 weeks): Run micro-experiments on CTA copy, button color, and placement. Focus on low-effort, high-impact tweaks.

Medium-term (1–3 months): Optimize mobile performance and page speed. Use Core Web Vitals to track gains. Experiment with reassurance cues like free shipping or returns badges near the CTA.

Long-term (6+ months): Reassess the full funnel, especially checkout. Consider multi-variate tests across the journey. Align experiments with OKRs and business KPIs to ensure measurable impact.

Governance: Use Jira/Confluence for experiment briefs, results tracking, and go/no-go rules. Define Minimum Detectable Effect (MDE) before tests. Monitor guardrail metrics (bounce rate, error rate).

12. Learnings

This project reinforced several key lessons for me as a BA/ICT professional:

1. Data alone does not convince executives — visuals and plain language do. Communicating results clearly is as important as statistical rigor.

2. Not all redesigns succeed. Evidence-driven validation protects businesses from costly missteps and wasted resources.

3. Dashboards are critical. A single KPI view can accelerate decisions more than a technical appendix.

4. Reproducibility matters. Exports and documentation build trust, transparency, and scalability across teams.

5. Even negative results create value. They prevent harmful rollouts and highlight areas where the next round of testing should focus.

13. Conclusion

This project demonstrates end-to-end BA/ICT capability. Starting with a vague business claim ('new design will convert better'), I framed a hypothesis, validated it with statistical tests, confirmed it with funnel analysis, and packaged results into executive dashboards. The outcome was negative for Variant B, but valuable for the business: evidence to retain A and a roadmap for smarter next experiments.

As a portfolio piece, this report shows employers that I can work like a consultant: structuring problems, testing solutions, analyzing outcomes, and communicating them in a way that drives real business decisions. It is not just an academic exercise — it is a demonstration of professional readiness.