A/B Test Analysis for Foodtech Company

Yu-Jen Wang

Table of contents

Abstract

This study examines whether larger food images on restaurant menus improve conversion rates through A/B testing. We evaluated data preprocessing methods (removing missing values, imputation, and stratified sampling) to ensure reliability.

Key Findings:

- Larger images → Higher conversion
- Stratified sampling → Stability & reduced bias
- Event type & platform → UX impact

Business Impact:

- Better images → Higher engagement & conversions
- Segmentation (iOS vs. Android) → Improved UX
- A/B testing expansion → Long-term validation

Experimental Background

- **Objective** → Test if larger menu images boost conversions.
- **Hypothesis** → Bigger images increase engagement & orders.
- **Key Metrics** → Conversion rate (successful orders).
- **Challenges** → Missing data, bias, sampling strategies.
- **Analysis** → Preprocessing, Statistical Testing, UX impact.

Experimental Procedure

1. Data Preprocessing

- Handle missing values (final_order_status, shop_id)
- Convert datetime format, remove duplicates

2. Data Processing Methods

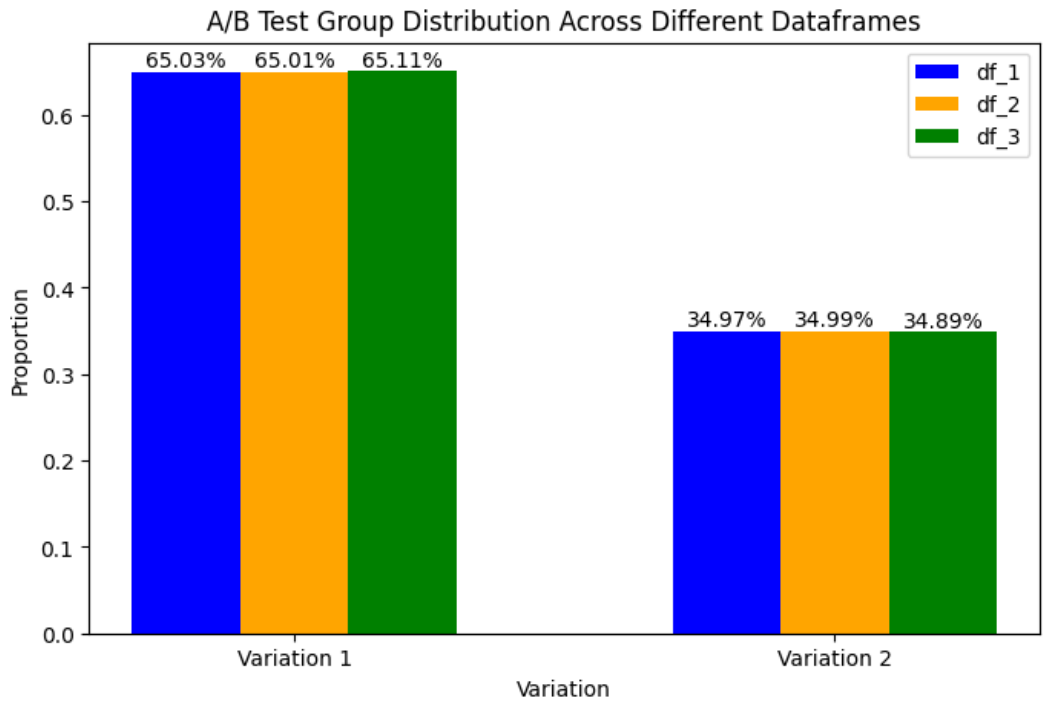
- df_1 → Original data
- df_2 → Drop shop_id missing values
- df_3 → Drop final_order_status missing values
- df_3 = df_4 → Only keep df_3

3. Variance Analysis

- Check event_id, session_id, user_id impact
- Ensure variation balance (A/B groups)
- Analyze platform & event type effects

Impact of Key Variables on A/B Testing

Dataset	Duplicate event_id	Unique session_id	Unique user_id	Total records
df_1 (Original Data)	603	179294	100000	326921
df_2 (Drop shop_id NaN)	561	168215	96293	315842
df_3 (Drop final_order_status NaN)	169	52418	42338	17063



Experimental Procedure

4. Statistical Testing

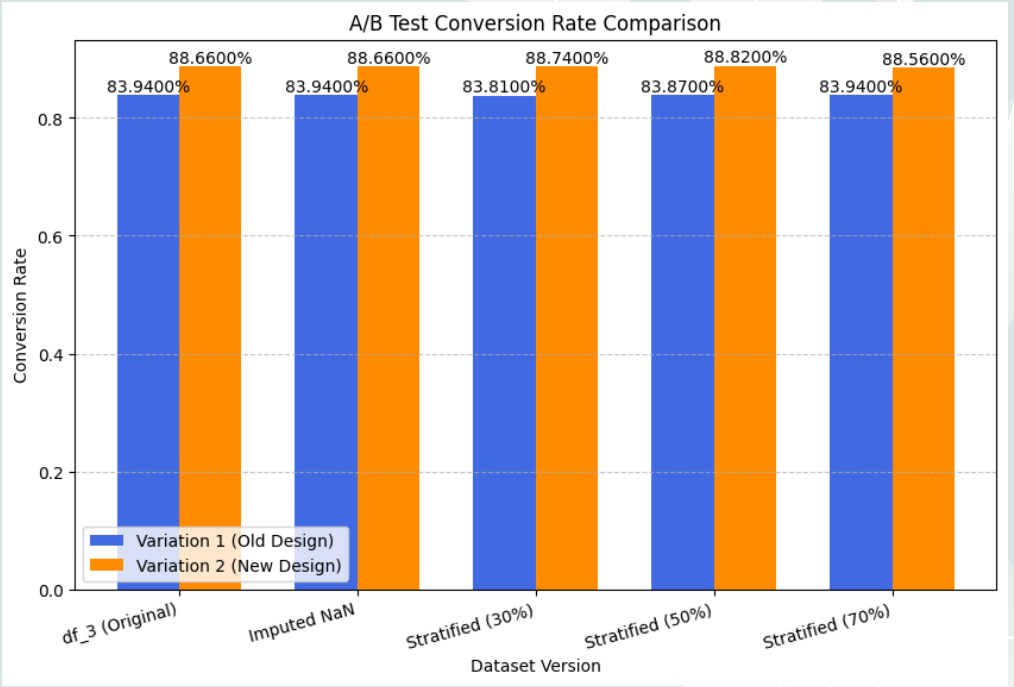
- Conversion rate comparison
- T-test & Chi-square test to measure significance

5. Robustness Test: Stratified Sampling (frac=0.5)

- Ensured stability & reduced bias

Key Findings & Statistical Analysis

- Larger images significantly increased conversion rates
- Stratified sampling confirmed result stability
- NaN Imputation & Stratified Sampling produced consistent results
- T-Test and Chi-Square Test confirmed statistical significance ($p < 0.05$)



Dataset	T-Statistic	P-Value
df_3 (Original)	27.68	3.80e-168 (**Significant**)
Stratified Sampling (30%)	15.83	3.01e-56 (**Significant**)
Stratified Sampling (50%)	20.59	6.35e-94 (**Significant**)
Stratified Sampling (70%)	22.64	3.14e-113 (**Significant**)

Dataset	Chi-Square Statistic	P-Value
df_3 (Original)	698.27	7.11e-154 (**Significant**)
Stratified Sampling (30%)	226.89	2.84e-51 (**Significant**)
Stratified Sampling (50%)	384.67	1.19e-85 (**Significant**)
Stratified Sampling (70%)	468.35	7.32e-104 (**Significant**)

Final Conclusion

- Larger food images significantly increased conversion rates across all dataset versions.
- Statistical analysis confirmed a significant improvement ($p < 0.05$).
- Stratified sampling ensured stability & reduced bias, confirming the robustness of results.
- Platform differences (iOS vs. Android) impact user behavior, suggesting potential UX optimization.

Business Recommendation

Short-Term Actions (Immediate Implementation)

- Deploy larger images across all restaurant menus to maximize conversion rates.
- Monitor key UX events (entry_to_shop, order_paid) to further optimize the user journey.
- Refine the design for mobile platforms, addressing differences in iOS vs. Android engagement.

Long-Term Strategies (Future Optimization)

- Expand A/B testing to other UI elements, such as image positioning and restaurant page layout.
- Segment analysis for different user groups (e.g., frequent vs. occasional buyers).
- Explore personalization strategies, such as AI-driven recommendations based on user preferences.

Future Considerations: Limitations & Recommendations

Test design limitations

1. **Test Duration** – The 6-day test period (11/25 - 11/30) may be too short to capture long-term user behavior and repurchase trends.

2. **Randomization** – No detailed verification of user randomization; potential imbalances in user behavior could skew results.

3. **External Factors** – Black Friday (11/29) promotions and time-based activity peaks (lunch/dinner) may have influenced conversion rates.

Improvements: Extend test duration, validate randomization, and control for external influences to ensure reliable results.

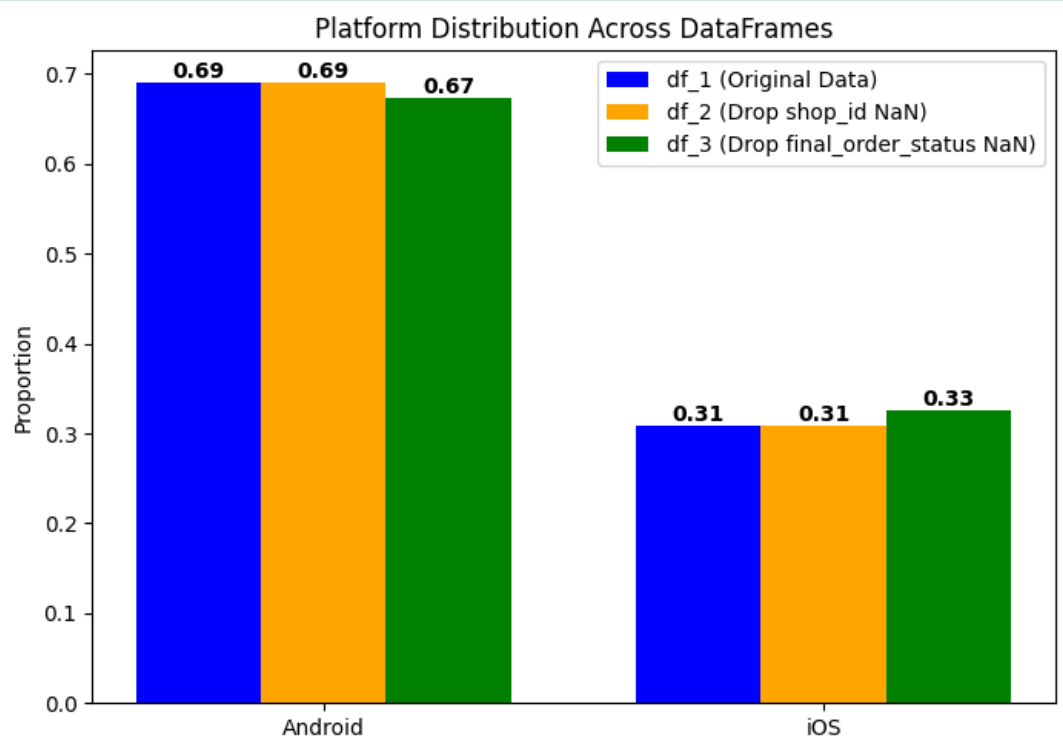
Thank you for listening

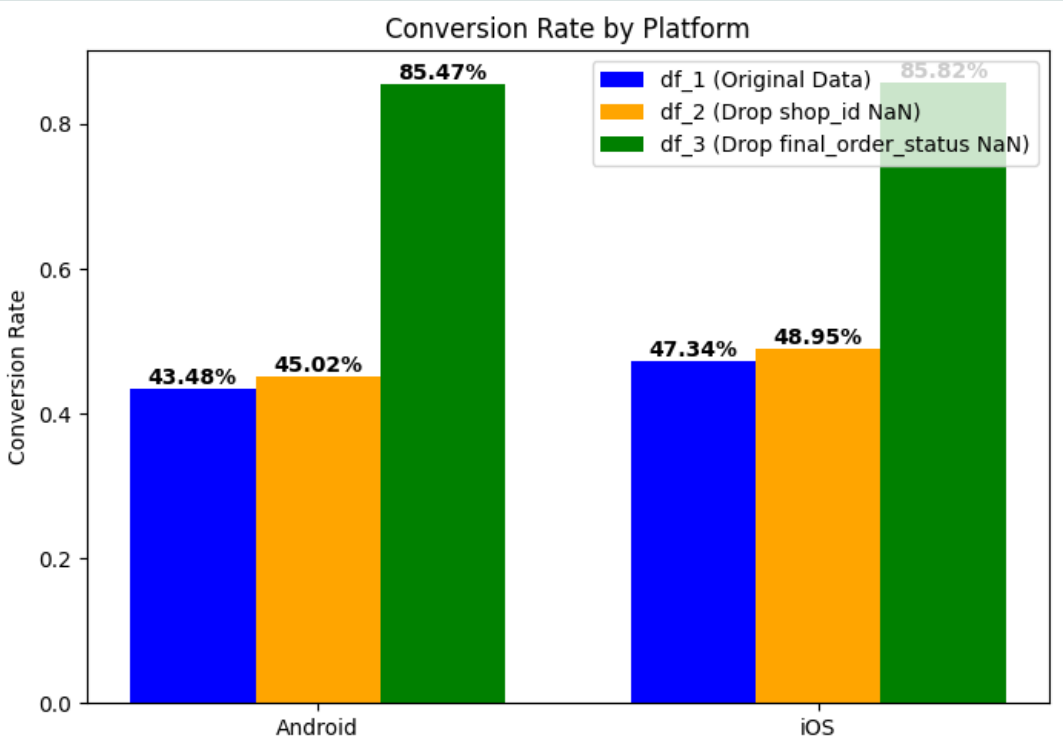
Yu-Jen Wang

ztex060033153@gmail.com

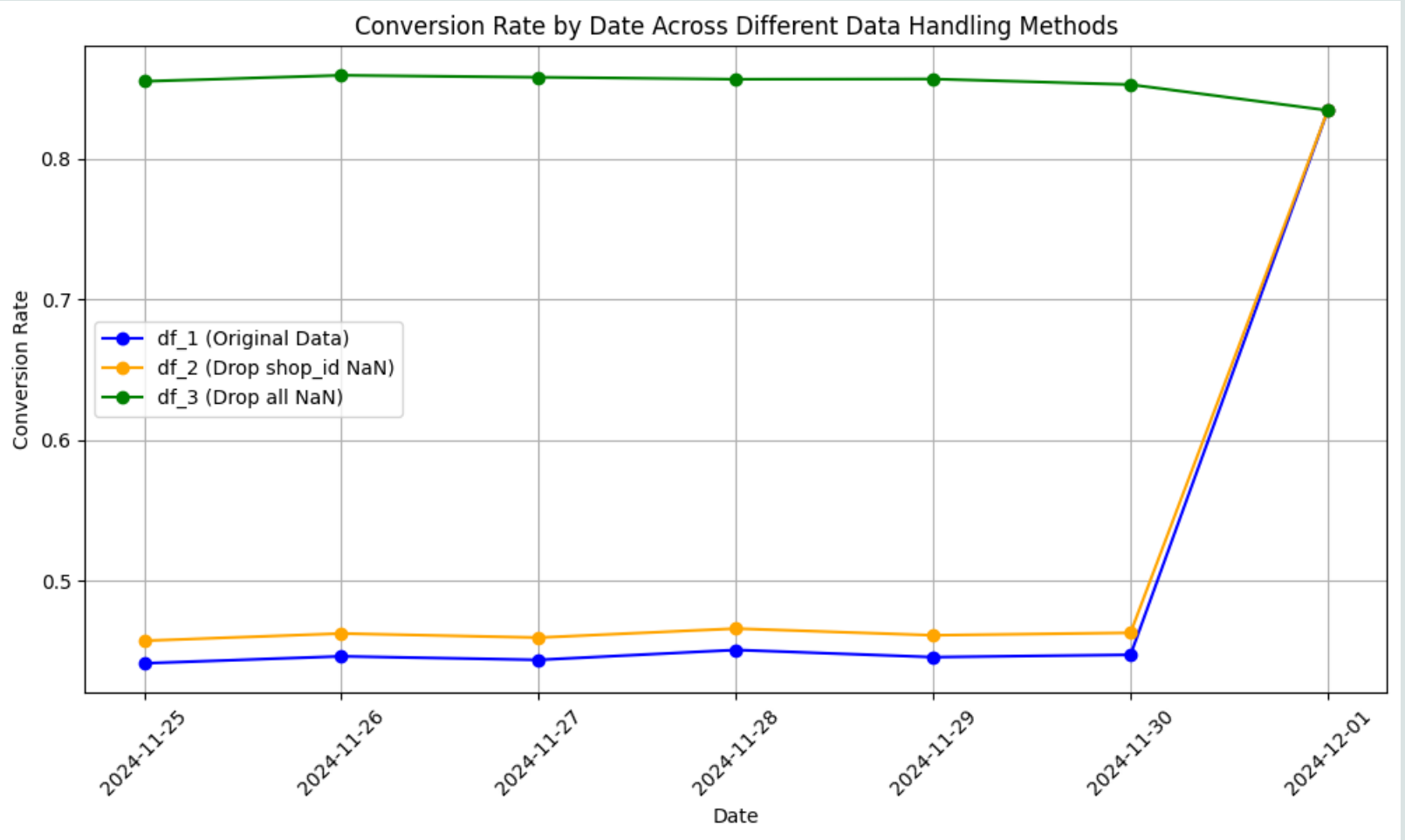
Additional Supplement plots

Platform (iOS vs. Android) & Conversion Rate

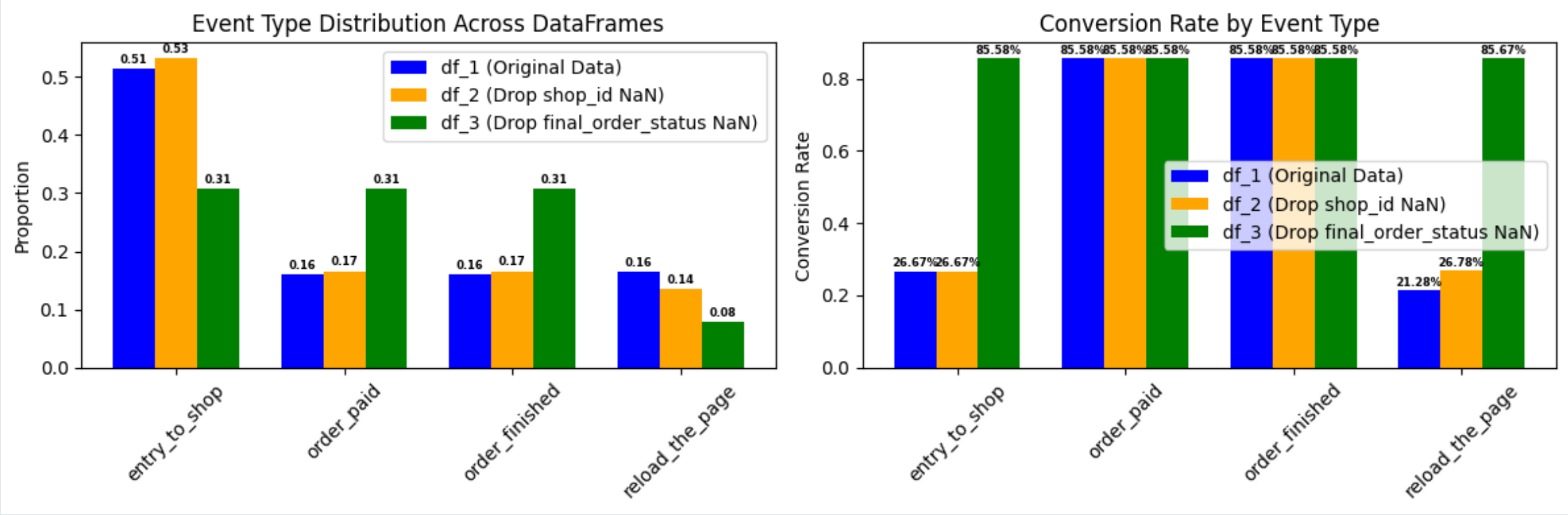




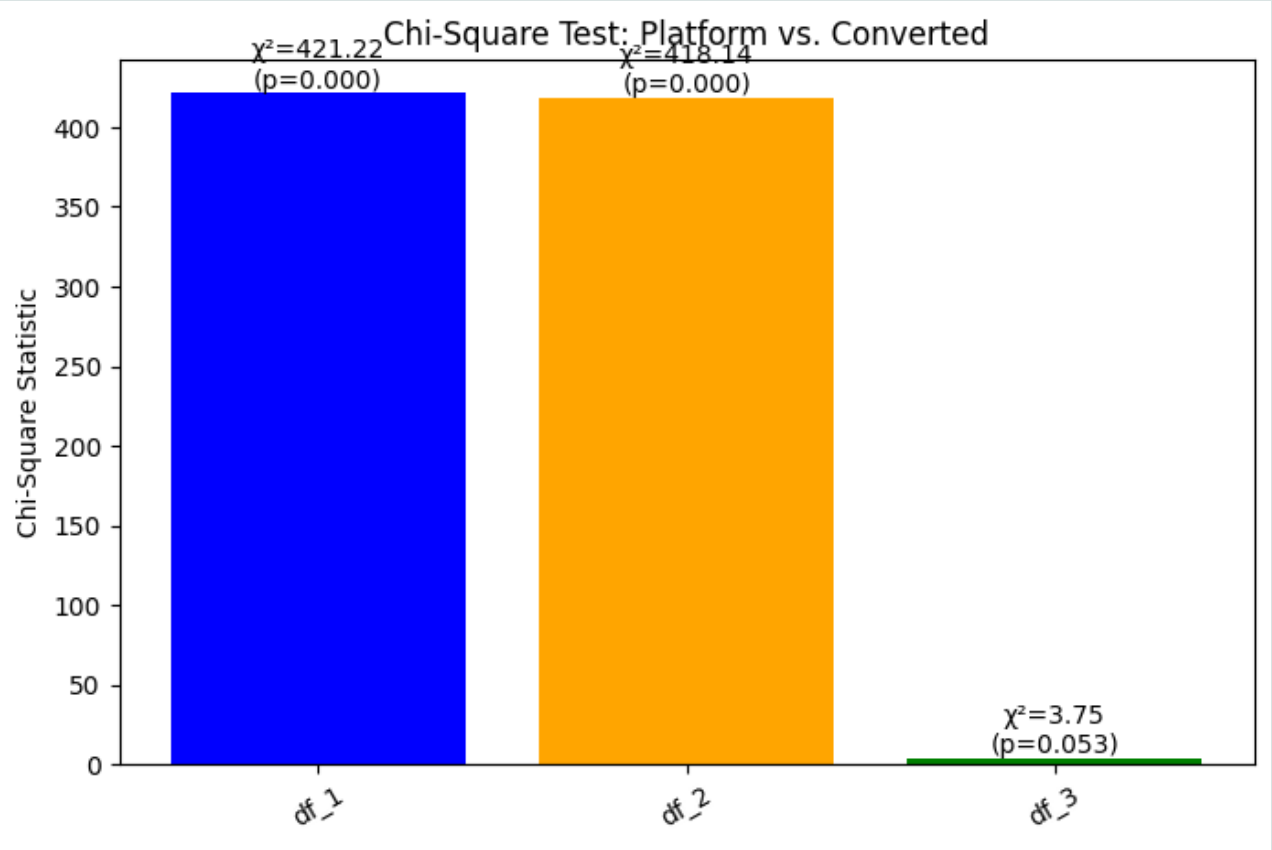
datetime_event (time) & conversion rate



Event_type & conversion rate



Chi-Square Test: Platform vs. Converted



Chi-Square Test: Event Type vs. Converted