KrispCall Revenue Attribution Analysis: Project Description

1. Objective

KrispCall Revenue Attribution Analysis is a data analyst test whose purpose is to test the skill to analyze customer data and assign revenue to integration platforms with the help of various attribution models. It aims to deliver specific steps toward a better performance of the platform partnership, marketing strategy, customer success program, and product development of KrispCall, which is a cloud-based communications platform, connected to 25 platforms (e.g., Salesforce, Monday, Zoho).

2. Business Context

KrispCall enables businesses to make calls and send SMS through web, mobile, and third-party integrations. Understanding which platforms drive revenue is critical for:

- Partnership Strategy: Prioritizing integration development.
- Marketing Investment: Allocating spend across platform ecosystems.
- Customer Success: Identifying high-value integration patterns.
- Product Development: Focusing on revenue-driving integrations.

3. Datasets

The analysis uses four datasets, detailed below:

3.1 Business Rules

- Only workspace owners generate revenue (payments/refunds).
- Any workspace user can connect integrations or make calls/SMS.
- Integration platform usage requires an active connection.
- Events must occur after first payment and before refunds.

4. Deliverables

The assignment requires the following outputs:

1. first_touch_attribution.csv: Revenue attributed to the first integration connected postpayment.
2. usage_based_attribution.csv: Revenue distributed by communication volume (call duration+ SMS segments).
3. chosen_attribution.csv: Revenue attributed using a chosen model (e.g., Linear).
4. analysis_summary.md: 500-1000 word report explaining methodology, findings, and recommendations.

5. attribution_analysis.py: Complete Python code for the analysis.

5. Attribution Models

5.1 First-Touch Attribution

- Assigns 100% of net revenue to the first integration connected after the initial payment.
- Attributes to ”No Integration” for customers without connections.

5.2 Usage-Based Attribution

- Distributes revenue based on communication volume (call duration in seconds + SMS seg-ments normalized to 60 seconds per segment).
- Includes web and mobile platforms.

5.3 Chosen Model (Linear)

- Equally distributes revenue across all active integrations per customer.
- Justify selection and compare with other models in the summary.

6. Technical Requirements

- Data Validation: Verify user existence, temporal constraints, and platform availability.
- Attribution Calculations: Handle edge cases (no platform usage, zero revenue), ensure properdata types, round revenue to 2 decimal places.
- Code Quality: Use clear variable names, comprehensive comments, and error handling.
- Visualizations: Generate plots (e.g., histograms, count plots) and a comparison chart.

7. Results Comparison

The attribution models yield distinct insights into platform contributions to revenue, with a total net revenue of $4,775. The comparison table below summarizes revenue for key platforms across the models.

7.1 Comparison of Results

First-Touch model shows most revenue ($2,335) is attributed to ”No Integration” and most of the customers fail to connect integrations on receiving initial payment. Of integrations, monday ($585) and pipedrive ($545) have the highest attribution, which is reflection that they are pre-existing common first connections that lead to acquisition.

Usage-Based, in its turn, changes the center of interest towards real communication, and web ($1,824.29) and mobile ($1,671.71) are in the spotlight because of their everywhere-and-everywhen availability and profoundly high usage values (i.e., 1572 events, mainly calls and SMS). There is a

high level of integration as the use of integration platforms such as zoho ($251.08), pipedrive ($197.33), and salesforce ($184.67) is noted, which implies that significantly integrating specific platforms will promote further interactions.

The linear model offers balanced perspective, and allocates revenue evenly across the available integrations per customer. Still in lead, the "No Integration" ($1,715), but such platforms as monday ($360.83), pipedrive ($308.33), and asana ($273.33) are scored rather equally due to their universal application. The model minimizes the prevalence of non-integration platforms and provides emphasis on a broader array of integrations (e.g., 20 platforms are the ones who are getting revenue).

On the whole, First-Touch is emphasizing acquisition (first integrations), Usage-Based is emphasizing engagement (daily usage), and Linear is emphasizing diversity (multiple integrations). Discrepancies indicate it is not used as early but when used, it is used strongly as in the case of zoho which reports high Usage-Based but zero First-Touch attribution.

7.2 Comparison of Plots

The visualizations provide supporting insights:

- Correlation Heatmap (correlation_heatmap.png): Displays weak correlations between payment/refund amounts and timestamps, indicating refunds are not strongly tied to payment size or timing, which supports the net revenue calculation approach.
- Total Users Histogram (total_users_histogram.png): Shows a right-skewed distribution (mean2.08, median 1), confirming KrispCall's focus on small teams, aligning with high web/mobile usage in the Usage-Based model.
- Platform Count Plot (platform_count_plot.png): Illustrates the top 10 platforms, with weband mobile leading, followed by zoho and pipedrive. Labels are rotated 45 degrees for clarity, highlighting usage patterns that explain the Usage-Based model's results.
- Net Revenue Histogram (net_revenue_histogram.png): Reveals a right-skewed distribution($0-$100 range), with outliers at higher values, explaining why "No Integration" dominates in First-Touch and Linear models (many low-revenue customers without integrations).
- Attribution Comparison Bar Plot (attribution_comparison.png): A horizontal bar plot com-paring revenue across models, with green for First-Touch, blue for Usage-Based, and yellow for Linear. It visually emphasizes web/mobile dominance in Usage-Based and balanced distribution in Linear, aiding quick identification of key platforms.

These plots are generated using Matplotlib and Seaborn, with a Chart.js configuration for interactive rendering of the attribution comparison, allowing zooming and tooltips for detailed exploration.

8. Evaluation Criteria

The evaluation for the revenue attribution can be done using information present in Table 1 with the evaluation criteria in Table 2.

Table 1: Revenue Attribution by Platform and Model (USD)

Platform	First-Touch	Usage-Based	Linear
No Integration	2335.00	0.00	1715.00
monday	585.00	174.67	360.83
pipedrive	545.00	197.33	308.33
asana	0.00	173.33	273.33
activecampaign	290.00	156.00	235.00
freshsales	175.00	178.67	210.00
salesforce	160.00	184.67	195.00
zoho	0.00	251.08	177.50
web	0.00	1824.29	0.00
mobile	0.00	1671.71	0.00
front	150.00	0.00	165.83
engagebay	160.00	110.67	125.00

Table 2: Evaluation Criteria

Category	Description	Weight
Technical Implementation	Correct model implementation, data handling, validation, code clarity	40%
Business Understanding	Accurate interpretation of requirements, edge case handling, revenue accuracy	30%
Analysis Quality	Clear methodology, insightful comparisons, actionable recommendations	20%
Communication	Professional writing, logical structure, effective visualizations	10%