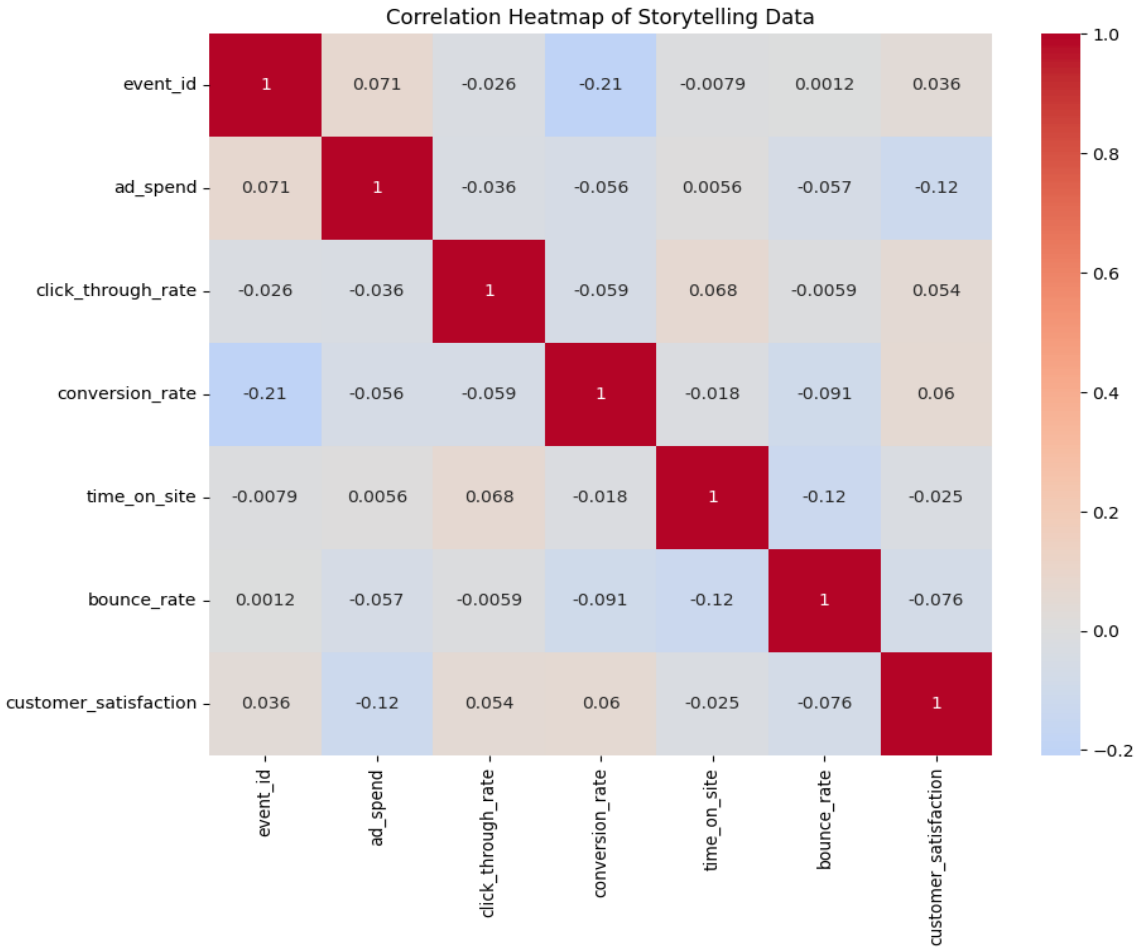
Data Storytelling

Sample Dataset (First 5 rows):

event_id	ad_spend	click_through_rate	conversion_rate	region	time_on_site	bounce_rate	customer_satisfaction
1	5745.07	0.18	0.1	South	203.58	0.31	66
2	4792.6	0.16	0.04	North	176.12	0.69	84
3	5971.53	0.15	0.04	South	223.95	0.67	62
4	7284.54	0.16	0.09	South	216.68	0.22	90
5	4648.77	0.1	0.03	East	251.88	0.55	88

Data Correlation Heatmap:



Task 1: Analyze Conversion Rate Drivers

Description:
Investigate factors influencing conversion rates in different regions.

Steps:
• Create box plots of conversion rate by region using seaborn
• Analyze mean and variability of conversion rates across regions
• Identify potential regional outliers
• Consider additional factors such as ad spend

Task 2: Click-Through Rate Analysis

Description:
Explore relationships between ad spend and click-through rate (CTR).

Steps:
• Use scatter plot with regression line to analyze correlation
• Calculate Pearson correlation between ad spend and CTR
• Identify thresholds where CTR increases or plateaus
• Consider using color to show conversion rate on scatter plot

Task 3: Customer Satisfaction Analysis

Description:
Analyze customer satisfaction and its correlation with bounce rate.

Steps:
• Use heatmap to display correlation matrix
• Plot scatter plot with bounce rate on x-axis and satisfaction on y-axis
• Look for thresholds where bounce rate impacts satisfaction
• Highlight any inverse relationships

Task 4: Time-on-Site Impact

Description:
Investigate how time spent on site influences conversion rate.

Steps:
• Create violin plot of conversion rates for different time-on-site categories
• Bin time-on-site values into short, medium, and long visits
• Analyze mean conversion rates within each category
• Perform ANOVA to check significance

Task 5: Conversion Patterns in High Ad Spend

Description:
Analyze conversion rate for high ad spend campaigns.

Steps:
• Create histogram of conversion rates for ad spend > $7000
• Calculate conversion rate statistics for high spend
• Compare with low ad spend conversions
• Use violin plot for deeper visualization

Task 6: Bounce Rate Clustering

Description:
Use clustering to categorize bounce rate patterns.

Steps:
• Standardize variables related to site interaction
• Apply K-means clustering to group bounce rates
• Visualize clusters in scatter plot
• Analyze cluster centers and interpret behavior

Task 7: Conversion Rate by Time on Site

Description:
Analyze conversion efficiency by time spent on site.

Steps:
• Create scatter plot with time on site vs. conversion rate
• Fit regression line to show relationship
• Use heatmap to show high and low conversion regions
• Summarize key insights for storytelling

Task 8: Bounce Rate vs. Satisfaction

Description:
Explore impact of high bounce rates on satisfaction.

Steps:
• Use violin plot to show satisfaction for high vs. low bounce rates
• Calculate mean satisfaction for each bounce rate category
• Identify trends and discuss potential improvements
• Provide summary statistics and visual insights

Task 9: Click-Through Efficiency

Description:
Analyze ad spend efficiency based on CTR.

Steps:
• Calculate CTR/ad_spend ratio
• Create distribution plot of efficiency
• Identify high-efficiency ad spends
• Generate summary statistics
• Profile high-efficiency outliers

Task 10: Comprehensive Performance Dashboard

Description:
Compile top visualizations to build a data story dashboard.

Steps:
• Identify most impactful visualizations for each metric
• Create combined plots to summarize findings
• Build a narrative with visuals to guide through insights
• Discuss how each metric influences overall performance