

Analytic Exercise



GUILHERME GONÇALVES DE LIMA

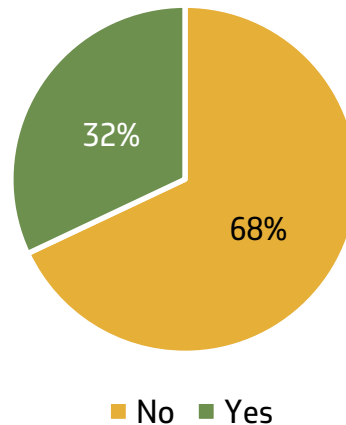
Summary

1 Analysis

2 Next steps

1 ANALYSIS

Percentage of Conversion



Considering **Conversion** as the customer bought at least one product.

32% (16,009) of customers bought a product on the website.

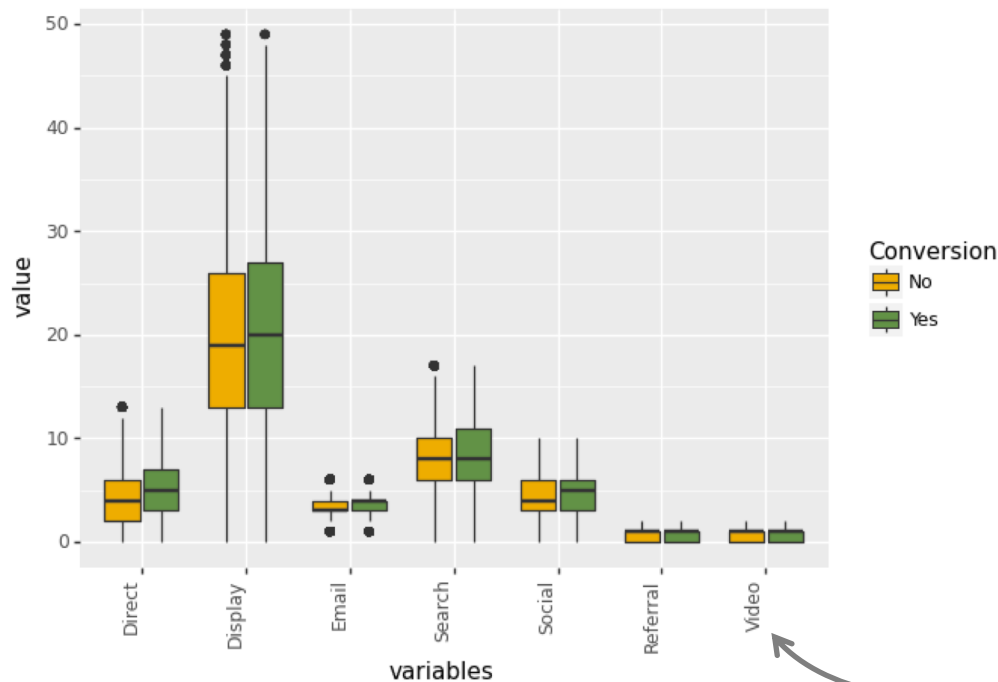
This study will identify the **Touchpoints** that increase the likelihood of *Conversions* and recommend the next steps.

7 Touchpoints:

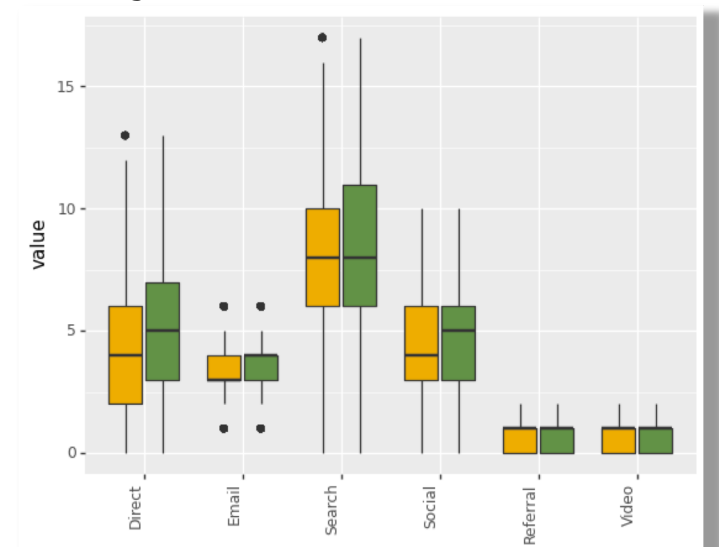
- a) **Direct:** Customer navigated directly to the client's website.
- b) **Display:** Exposure to a display or banner ad.
- c) **Email:** Customer received an email promotion from the client.
- d) **Search:** Customer used a search term tracked by the client.
- e) **Social:** Customer saw an advertisement on a social network platform (e.g., Facebook).
- f) **Referral:** Customer clicked on a link from a partner site that sent them to the client site.
- g) **Video:** Customer watched a client's online video.

Analytic Exercise

Distribution of Touchpoints by Conversion



Better visualization of touchpoints without Display
Range of 0 - 20



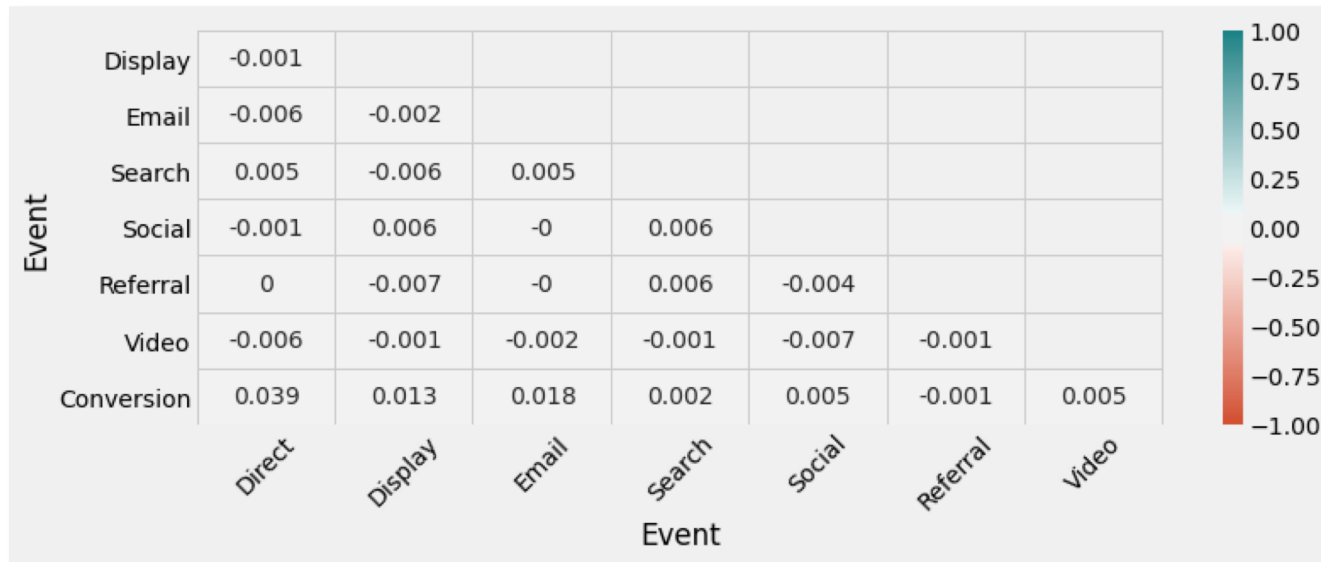
There is a significant difference between *Conversion* and *Non-Conversion* for **Direct, Display, Email, and Social** (Mann-Whitney U test: p-value<5%)

Outliers (395 customers) was removed for this analysis.

Analytic Exercise

Correlation Matrix of Touchpoints

Pearson's Correlation



All Touchpoints are ***weakly correlated*** with each other and with *Conversion*.

Using **Logistic Regression** to model the probability of **Conversion**.

$$\text{Logit}[P(Y = 1)] = \ln \frac{P(Y = 1)}{1 - P(Y = 1)} = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_7 X_7$$

where:

$P(Y = 1)$: is the probability that the customers will buy at least one product (Conversion).

X_1, X_2, \dots, X_7 : are the explanatory variables (Touchpoints).

$\beta_1, \beta_2, \dots, \beta_7$: are the coefficients (effects) of variables.

Analytic Exercise

Logistic Regression results:

Parameter/Variable	Coef.	Std. Err.	z statistic	p-value	CI 95%
β_1 : Direct	-0.0140	0.0033	-4.2978	0.0000	[-0.0204, -0.0076]
β_2 : Display	0.0040	0.0009	4.354	0.0000	[0.0022, 0.0058]
β_3 : Email	0.0711	0.0076	9.3227	0.0000	[0.0561, 0.0860]
β_4 : Search	0.0305	0.0028	10.8004	0.0000	[0.0250, 0.0361]
β_5 : Social	0.0323	0.0045	7.2148	0.0000	[0.0235, 0.0411]
β_6 : Referral	0.0611	0.0174	3.5053	0.0005	[0.0269, 0.0952]
β_7 : Video	0.0460	0.0174	2.6416	0.0083	[0.0119, 0.0802]

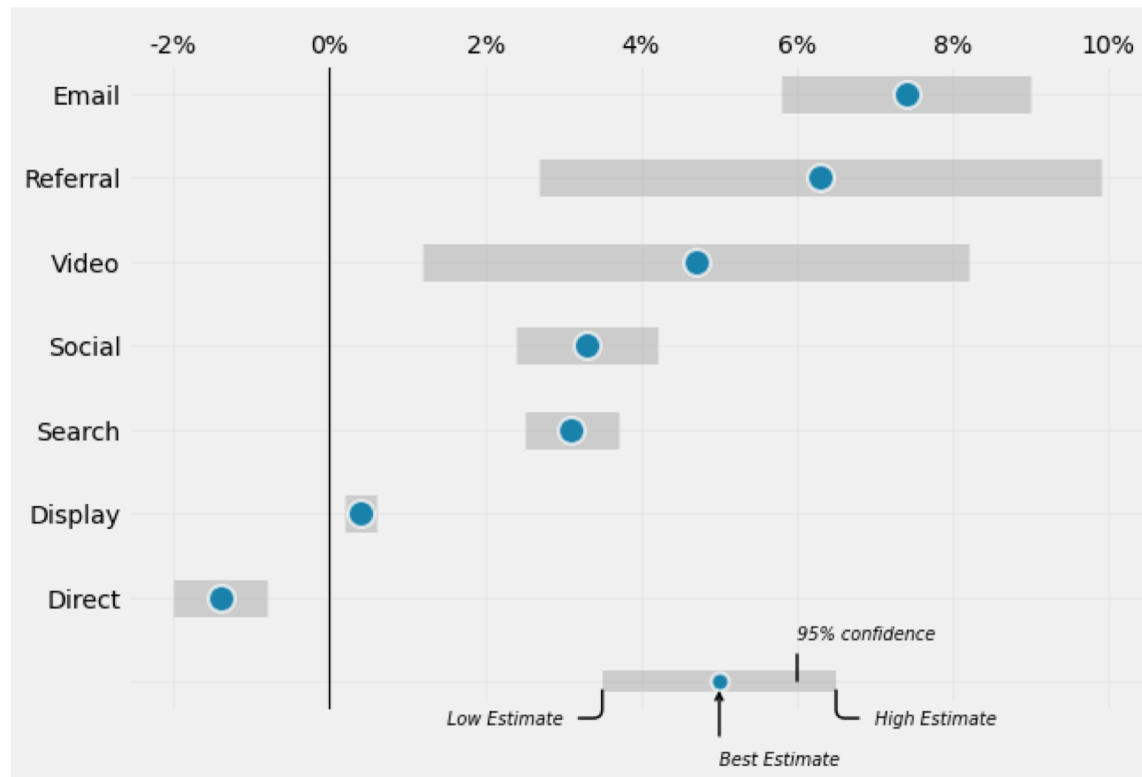
Null deviance:	62,180 on 49,605 degrees of freedom
Residual deviance:	62,499 on 49,598 degrees of freedom
AIC:	62,513

- The regression coefficients gives the **Odds Ratio** that quantifies the strength of the association between the **Conversion** and the **Touchpoints**.
- If the Odds Ratio is **greater than 1**, the *Conversion* probability **increase** for every one unit change in the touchpoint, holding all other variables constant.

Analytic Exercise

The Odds of *Conversion* (versus *Non-Conversion*) increases when a customer receives an email promotion.

The percentage change in the Odds Ratio with 95% confidence intervals for the Touchpoints.



2 NEXT STEPS

Based on this study, we recommend:

- A campaign based on Email.
- Deep study for the Referral and Video touchpoints, for example, which sites the Customers interacted with these events.
- Deep study for Direct touchpoint, for example, the Customers navigate directly on the website searching for promotions or product price reduction.
- Check which touchpoints make the Customer buy more.

Thanks



GUILHERME G LIMA