# Air France Business Case

Data Analysis and Insights

Team 1 – MSBA2

17 December 2020

# Overview

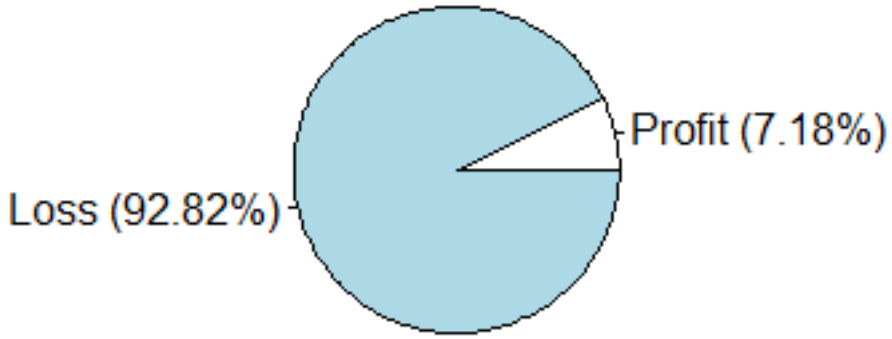
A total of 4510 SEM campaigns

- 2079 unique keywords
- 348 unique keywords IDs
- 7 unique publisher names
- 7 unique bid strategies (1 NA)
- 24 unique types of campaigns - 3 branded

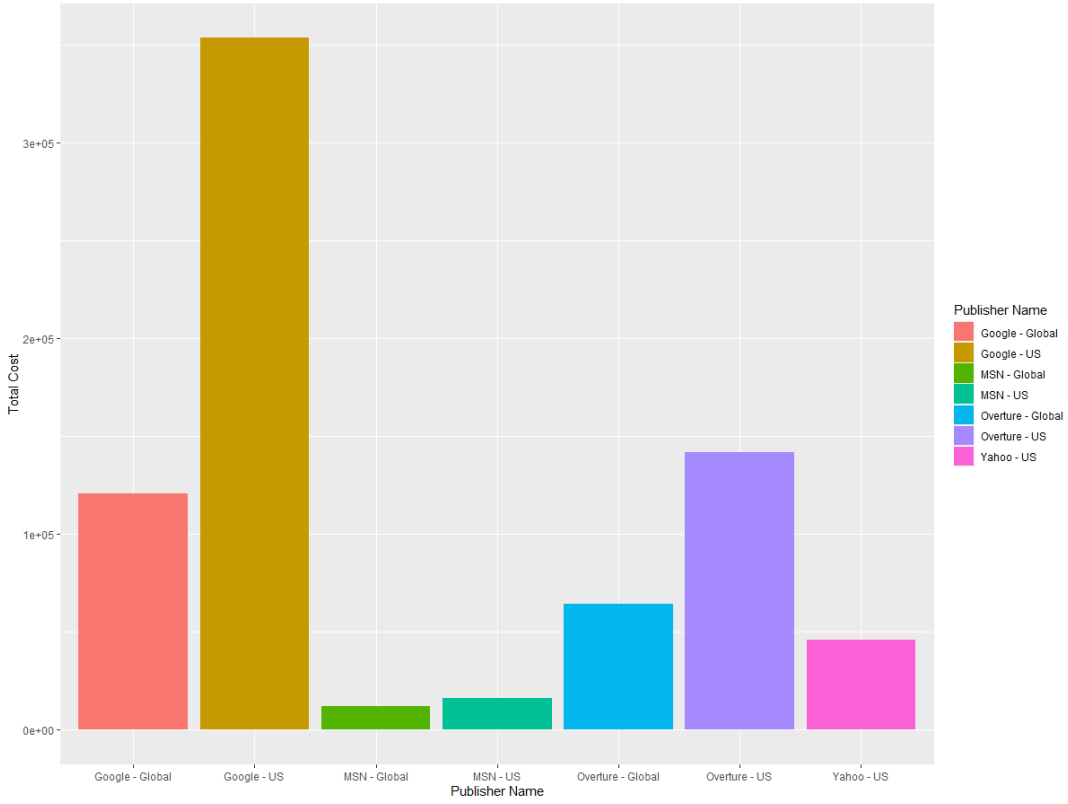
- 4000 total bookings
- Total revenue generated: 4.7 million
- Total cost of campaigns: 0.8 million

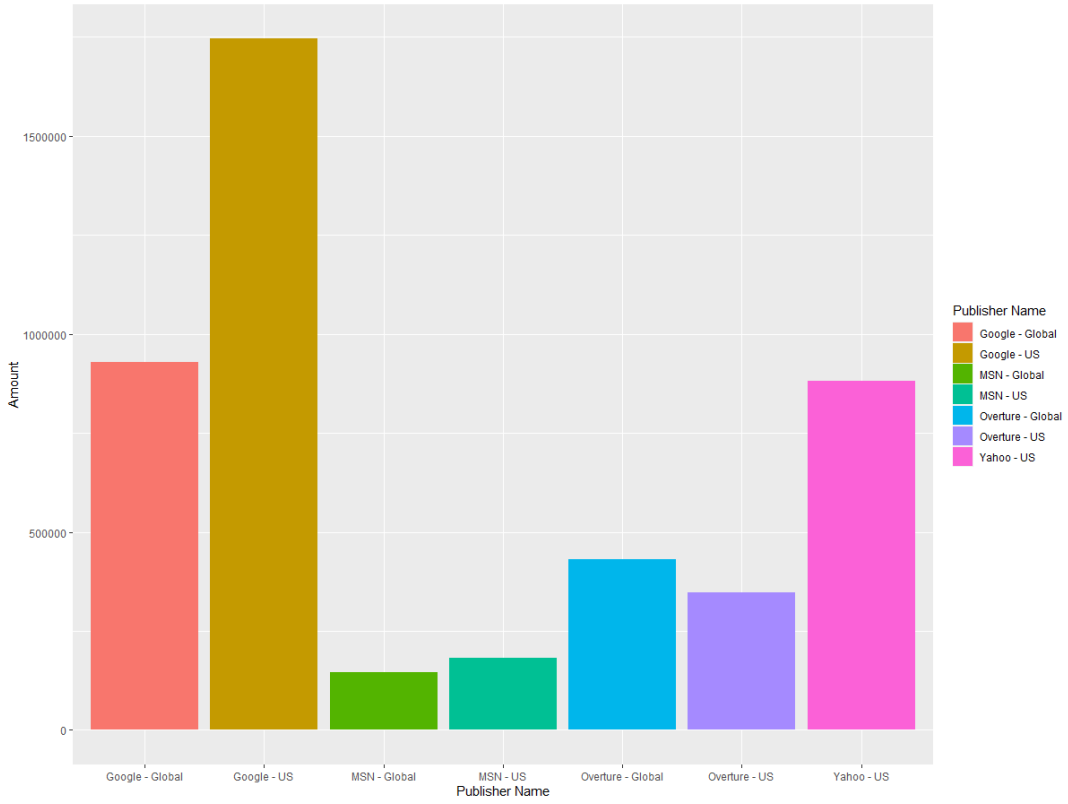
# Problem Statement

Only 7% of total SEM campaigns (4510) produced profit for Air France.

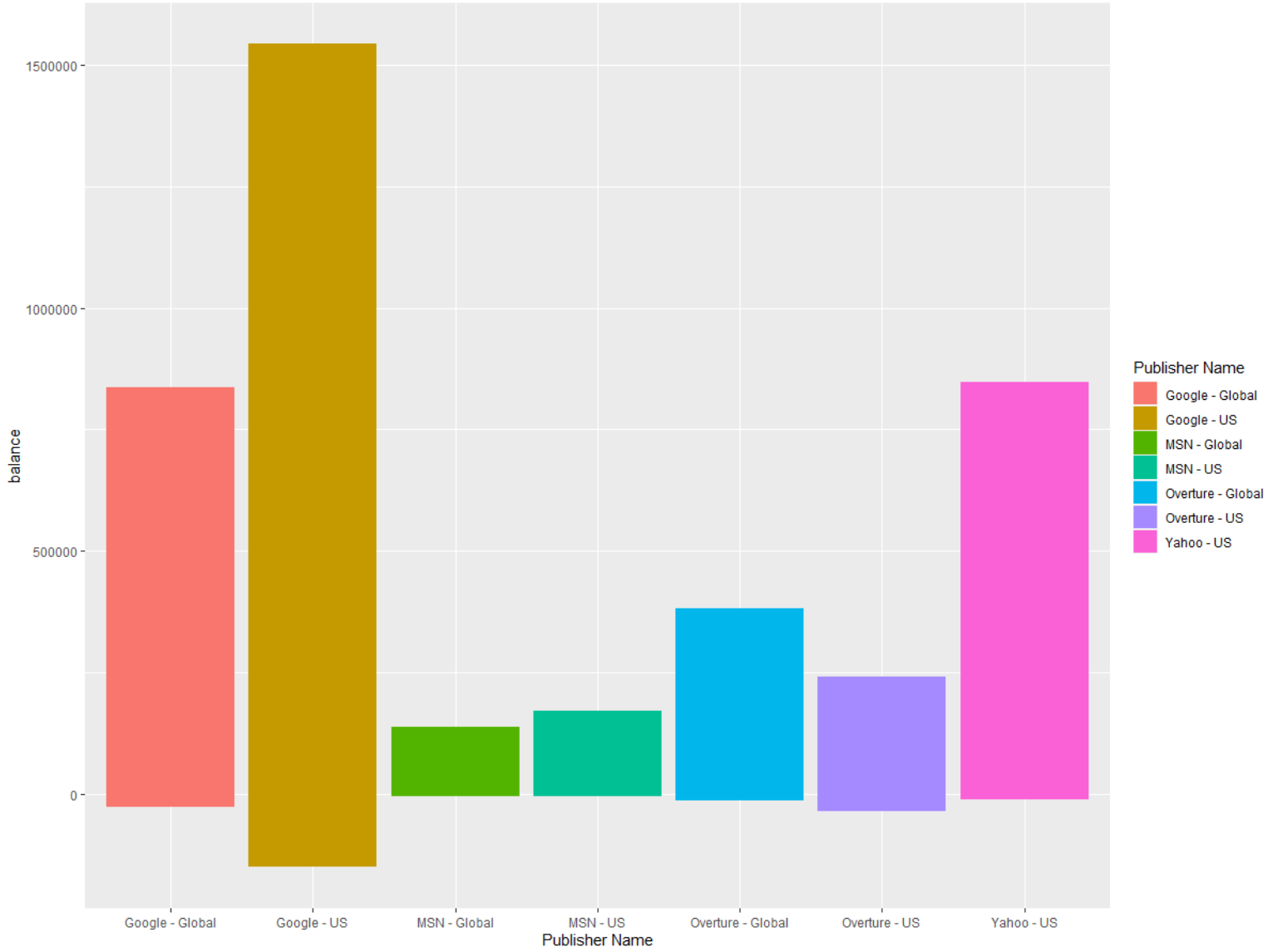


# Publisher

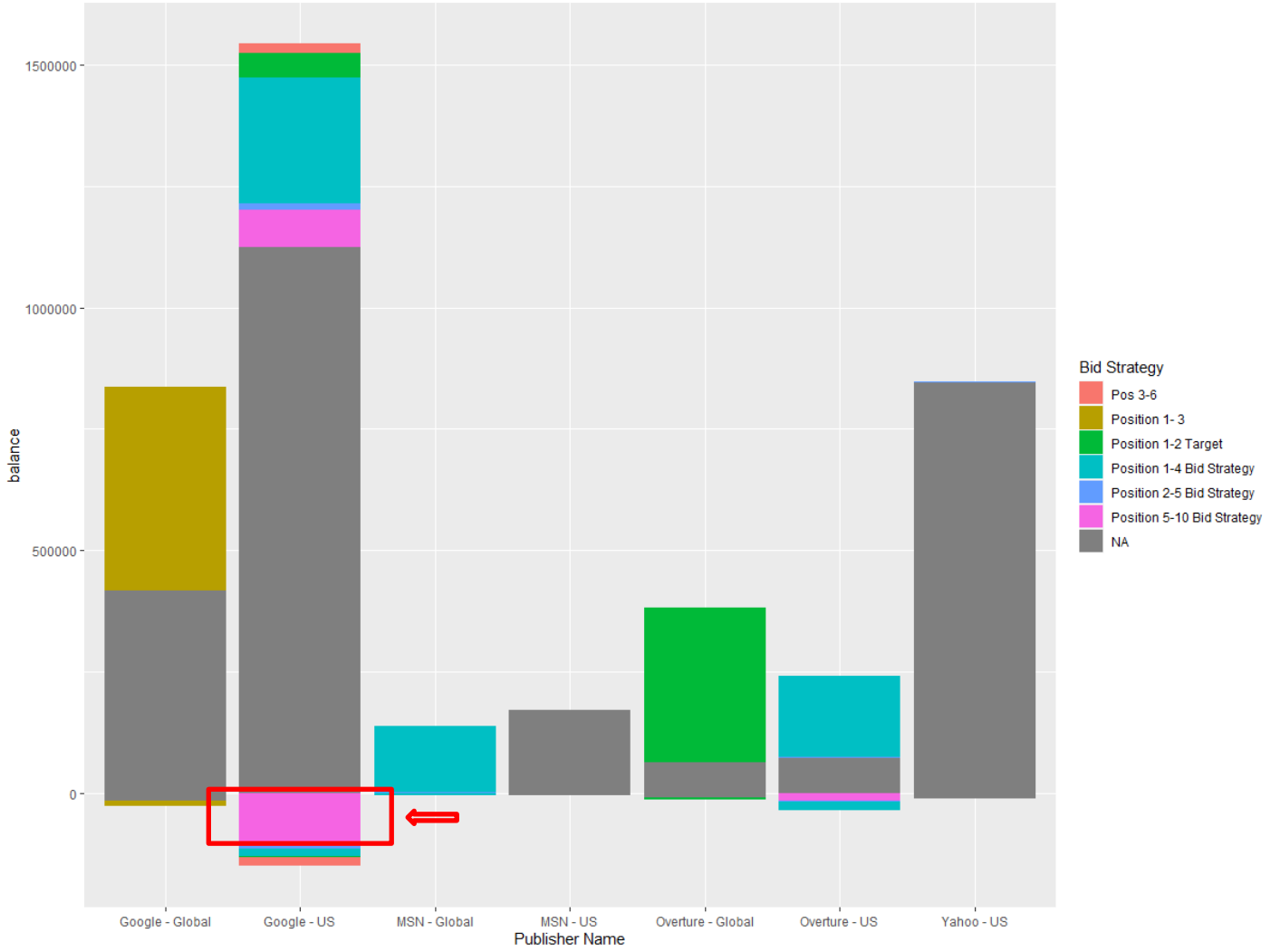

Total Cost
3e+05
2e+05
1e+05
0e+00
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US
Publisher Name
Publisher Name
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US


Amount
1500000
1000000
500000
0
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US
Publisher Name
Publisher Name
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US

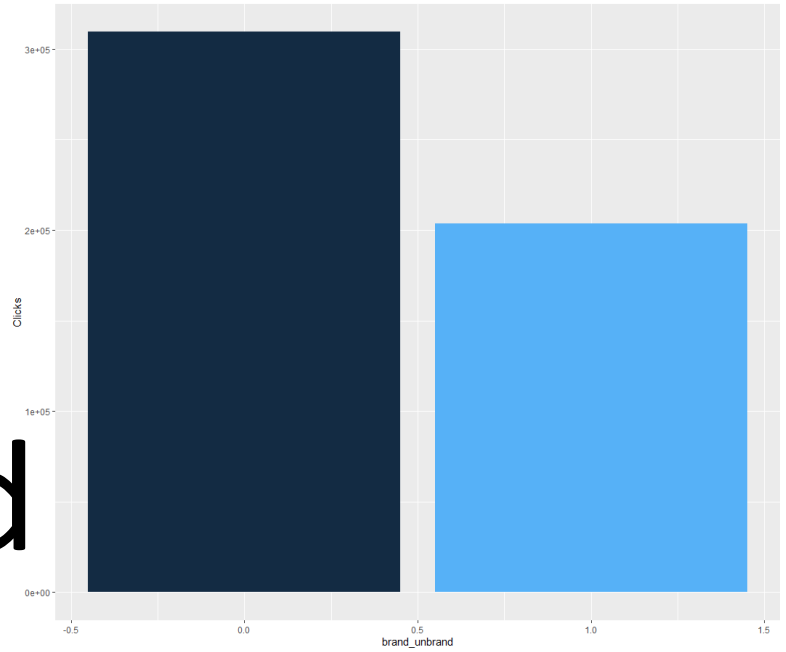
# Profit & Loss

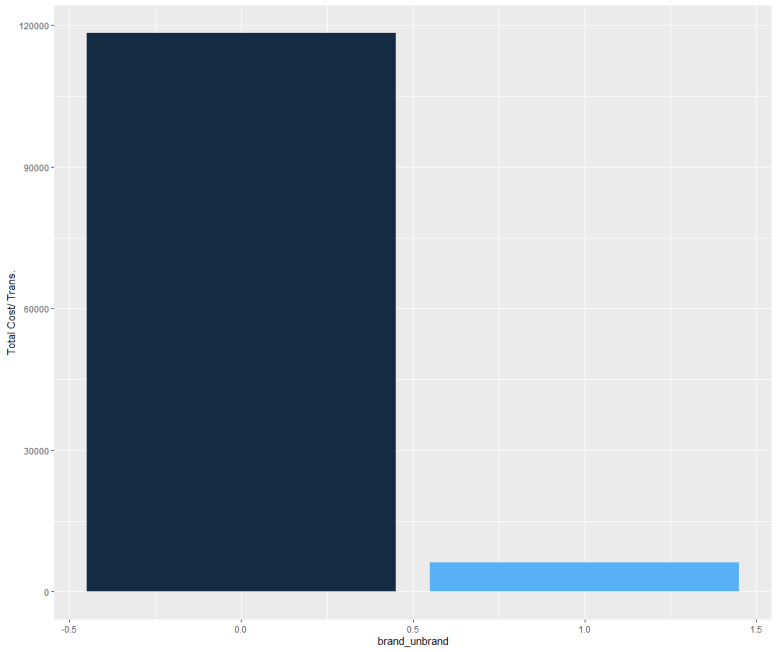

1500000
1000000
500000
0
balance
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US
Publisher Name
Publisher Name
Google - Global
Google - US
MSN - Global
MSN - US
Overture - Global
Overture - US
Yahoo - US

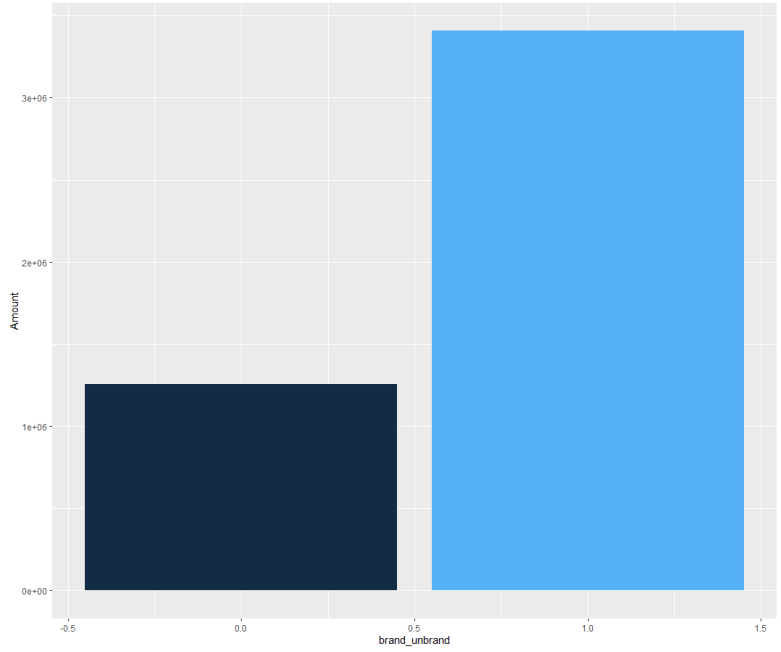
# Bid Strategy – Google US

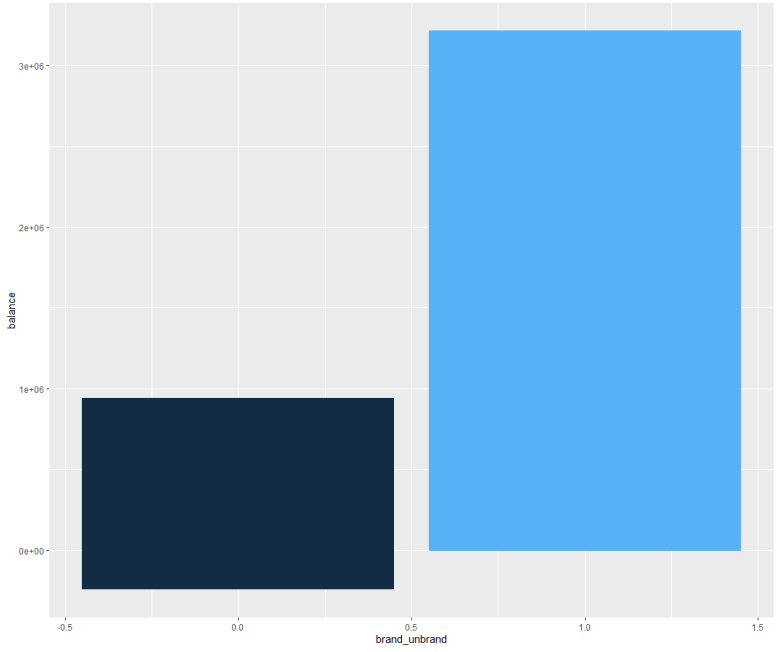


# Unbranded Vs. Branded



# Testing Metrics

## Logistic Regression Model:

```
Call:
glm(formula = balance_binary ~ `Search Engine Bid` + Clicks +
    `Trans. Conv. %`, family = "binomial", data = my_df)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-8.4904  -0.0913  -0.0782  -0.0781   3.4815

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)        -4.4630920  0.2703954 -16.506  < 2e-16 ***
`Search Engine Bid` -0.2126791  0.0524584  -4.054 5.03e-05 ***
Clicks              0.0010798  0.0002004   5.388 7.14e-08 ***
`Trans. Conv. %`   11.2613542  0.7596009  14.825  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2330.53  on 4509  degrees of freedom
Residual deviance:  315.02  on 4506  degrees of freedom
AIC: 323.02

Number of Fisher Scoring iterations: 13
```

## Confusion Matrix:

```
Confusion Matrix and Statistics

          Reference
Prediction    0    1
         0 2927   16
         1   12  202

               Accuracy : 0.9911
                 95% CI : (0.9872, 0.9941)
    No Information Rate : 0.9309
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.9304

 Mcnemar's Test P-Value : 0.5708

            Sensitivity : 0.9959
            Specificity : 0.9266
         Pos Pred Value : 0.9946
         Neg Pred Value : 0.9439
             Prevalence : 0.9309
         Detection Rate : 0.9271
   Detection Prevalence : 0.9322
      Balanced Accuracy : 0.9613
```

**Result: Search Engine Bid – Clicks – Transaction Conversion Rate (%).**

# Business Insights & Key Takeaways:

- Invest more on SEM through Google US and Global, with Yahoo.
- Branded over Unbranded.
- 1-4 Bid Strategy over 5-10 Bid strategy.
- Be ware of missing data.

Prepared by Carlos Andres Restrepo Ayala, Daniel Puello, Tendai Dube, Lavini Raj and Romi Singh