Workbook

Part 1: Motivation

Problem Statement

- High Costs per click with SEM (Search Engine Marketing)
- Airline industry a Competitive market with Low margins

State the Questions

- Where do we allocate our marketing budget most efficiently?
- How can we reduce Cost/Click, increase revenue and optimize performance?
- Do branded keywords bring in more revenue?
- Are broad or focused keywords more profitable?
- Can assist keywords help increase conversion rate?
- Which search engine delivers the most ROI
- what are the customer segments / search engine –> Specific pattern in buying behavior?

Main Objectives

- Find out Cost, revenue and volume of campaigns
- Find out single-click conversion rate of branded / unbranded keywords?
- Minimize Cost/Click
- Maximize ROA
- Maximize Revenue
- Maximize Single-click conversion
- Maximize Profitability
- Maximize Conversion Rate

What could be a positive outcome?

Part 2: Method

What key resources do we acquire?

Data: - Description

Are all the imported variables important? Useful variables in the dataset (Type: xls)

$campaigns $impressions $click-through $Cost/Click $Revenue $Single-click conversion $Profitability $Conversion Rate

R Libraries

```
# Import Libraries
library(readxl)
library(tidyr)
```

What is our approach to solve the problem?

High level process of steps

Part 3: Mechanics

Inspect & Import data

R tries to import the first sheet of the excel file which resolves in an error. This is why the argument read_excel function has to be used to specify the column.

```
# Inspect sheets of excel-file
excel_sheets('C:/Users/LK/Nextcloud7/Personal/Docs/case-studies/Air France/assets/Air France Case Sprea
```

```
## [1] "DoubleClick" "Copyright"   "Kayak"
```

```
# Import data
kayak <- read_excel("C:/Users/LK/Nextcloud7/Personal/Docs/case-studies/Air France/assets/Air France Cas
                                                    sheet = "Kayak")
```

```
## New names:
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * `` -> ...6
## * ...
```

```
doubleclick <- read_excel("C:/Users/LK/Nextcloud7/Personal/Docs/case-studies/Air France/assets/Air Fran
                          sheet = "DoubleClick")
```

Massaging

```
#Convert to dataframe
doubleclick <- as.data.frame(doubleclick)

# Get a big picture understanding of the data
summary(doubleclick)
```

```
##  Publisher ID       Publisher Name      Keyword ID          Keyword
##  Length:4510        Length:4510        Length:4510        Length:4510
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
```

```
##
##
##
##   Match Type          Campaign         Keyword Group        Category
##  Length:4510        Length:4510        Length:4510        Length:4510
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##  Bid Strategy       Keyword Type          Status          Search Engine Bid
##  Length:4510        Length:4510        Length:4510        Min.   : 0.000
##  Class :character   Class :character   Class :character   1st Qu.: 3.384
##  Mode  :character   Mode  :character   Mode  :character   Median : 6.250
##                                                           Mean   : 5.435
##                                                           3rd Qu.: 6.250
##                                                           Max.   :27.500
##      Clicks        Click Charges      Avg. Cost per Click  Impressions
##  Min.   :    0.0   Min.   :    0.00   Min.   : 0.000      Min.   :      0
##  1st Qu.:    1.0   1st Qu.:    2.31   1st Qu.: 0.825      1st Qu.:     28
##  Median :    4.0   Median :    6.76   Median : 1.650      Median :    176
##  Mean   :  113.7   Mean   :  167.48   Mean   : 1.890      Mean   :   9284
##  3rd Qu.:   19.0   3rd Qu.:   28.49   3rd Qu.: 2.663      3rd Qu.:    844
##  Max.   :34012.0   Max.   :46188.44   Max.   :10.000      Max.   :8342415
##  Engine Click Thru %   Avg. Pos.      Trans. Conv. %     Total Cost/ Trans.
##  Min.   :  0.000     Min.   : 0.000   Min.   :  0.0000   Min.   :   0.00
##  1st Qu.:  1.532     1st Qu.: 1.143   1st Qu.:  0.0000   1st Qu.:   0.00
##  Median :  4.106     Median : 1.594   Median :  0.0000   Median :   0.00
##  Mean   : 11.141     Mean   : 1.930   Mean   :  0.5693   Mean   :  27.61
##  3rd Qu.: 10.917     3rd Qu.: 2.308   3rd Qu.:  0.0000   3rd Qu.:   0.00
##  Max.   :200.000     Max.   :15.000   Max.   :900.0000   Max.   :9597.17
##      Amount          Total Cost       Total Volume of Bookings
##  Min.   :     0   Min.   :    0.00   Min.   :  0.0000
##  1st Qu.:     0   1st Qu.:    2.31   1st Qu.:  0.0000
##  Median :     0   Median :    6.76   Median :  0.0000
##  Mean   :  1034   Mean   :  167.48   Mean   :  0.8734
##  3rd Qu.:     0   3rd Qu.:   28.49   3rd Qu.:  0.0000
##  Max.   :567463   Max.   :46188.44   Max.   :439.0000
```

```
#Look for weird stuff
table(doubleclick$`Match Type`)
```

```
##
## Advanced    Broad    Exact      N/A Standard
##      969     2591       22        48      880
```

```
# The NAs have to be removed.
doubleclick_clean <- na.omit(doubleclick)

# Notice how the number of rows gets reduced
print(nrow(doubleclick_clean))
```

```
## [1] 3286
```

##Descriptive

Plots Statistics Correlation Association

Predictive

Feature Selection Apply ML-Algorithmus

Mechanics

Distribution

Stats

Boxplot

Outlier

Message

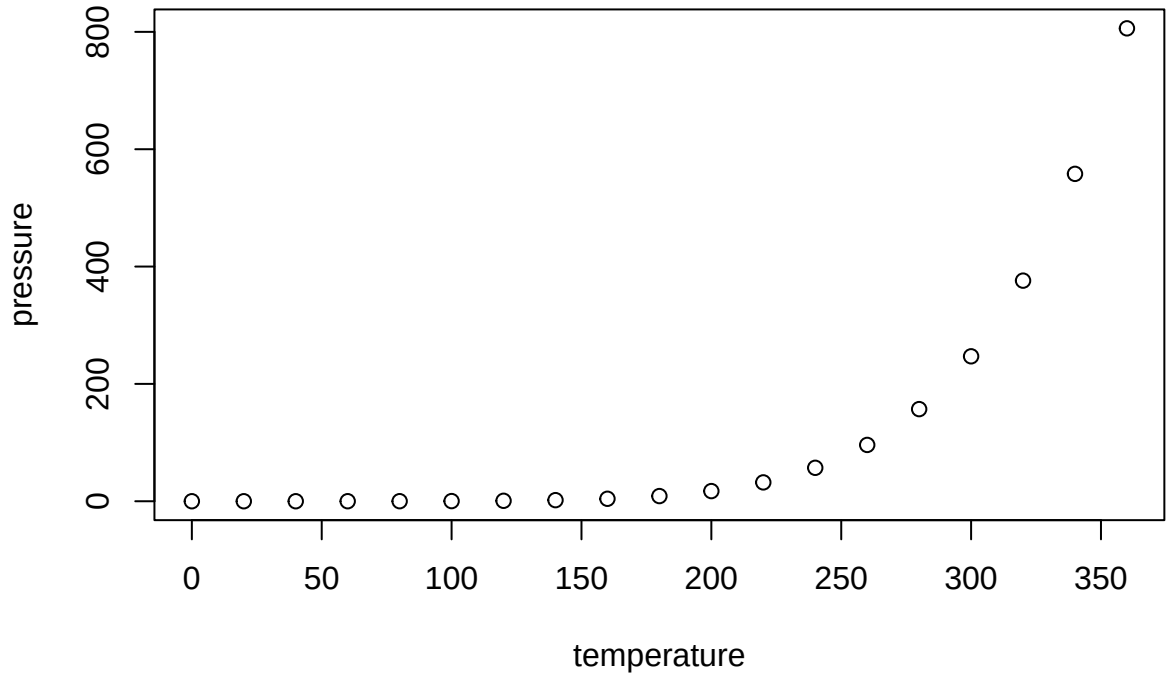
Key Findings

The C-suite of _____ face the following (problem/challenge), which is best solved with _ (solution) having an impact and/or making profits via _____ . The unique advantages/differentiators of the MVP are _____ , when comparing with the following key competitors / alternatives: _____

Next steps

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.