

Air_Travel_Entrega_vSent

Bryan Velazco & Daniel Guardia

2024-11-16

#Sinopsis

El caso se centra en la compañía Ariel Research, que ha desarrollado un modelo para medir el Brand Equity (equidad de marca) utilizando datos de una encuesta masiva realizada a consumidores canadienses. Este concepto intangible, que refleja el valor añadido a una marca por su nombre, fue desglosado en cinco dimensiones principales: familiaridad, unicidad, relevancia, lealtad y popularidad. La empresa busca evaluar si su modelo y base de datos pueden proporcionar métricas significativas y útiles para tomar decisiones estratégicas de marketing.

Problema General y Especifico del caso

Determinar si el modelo desarrollado por Ariel Research es efectivo para medir y analizar la equidad de marca de manera confiable, permitiendo a las empresas tomar decisiones estratégicas basadas en sus resultados.

Identificar la relación entre las dimensiones de la equidad de marca y la lealtad del cliente. Evaluar la pertinencia de utilizar variables binarias para categorizar las dimensiones de la equidad de marca. Analizar si el enfoque de Ariel es aplicable a diferentes industrias y categorías de productos.

Objetivo General y Especifico del caso

Validar y perfeccionar el modelo de Ariel Research para medir la equidad de marca y determinar su utilidad en el diseño de estrategias de marketing basadas en datos.

Realizar análisis estadísticos que confirmen las correlaciones entre las dimensiones del modelo y la lealtad del cliente. Comparar las métricas de equidad de marca entre diferentes categorías (como comida rápida y viajes). Proponer mejoras al modelo basado en los hallazgos para optimizar su aplicabilidad y precisión.

Importando librerías para el análisis del caso

```

library("dplyr")
library("tidyverse")
library("RColorBrewer")
library("gridExtra")
require("knitr")
library(openxlsx)
library("funModeling")
library("nortest")
require(car)
library(ggplot2)
library("gplots")
library(FactoMineR)
library(ade4)
library(ExPosition)
library(factoextra)
library(Rcpp)
library(InPosition)
library(ca)
library(plotly)
library(reshape2)
library(readxl)
library(readxl)

```

Importando DataSet

```

Base1 <- read_excel("AIR TRAVEL.xlsx", sheet = "Sheet1")
glimpse(Base1)

```

```

## Rows: 1,500
## Columns: 17
## $ gender    <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 2, 2, 2, 1, 2, 1, 1, 2, 1...
## $ age       <dbl> 60, 56, 51, 52, 69, 47, 49, 58, 44, 50, 61, 50, 58, 49, 43, 4...
## $ children  <dbl> 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1...
## $ income    <dbl> 4, 3, 4, 3, 2, 2, 1, 2, 4, 4, 2, 3, 4, 1, 4, 2, 3, 2, 3, 4, 4...
## $ famil     <dbl> 1, 10, 10, 7, 8, 9, 8, 10, 9, 7, 9, 9, 8, 3, 9, 8, 10, 8, 9, ...
## $ uniqu     <dbl> 1, 10, 1, 5, 5, 8, 8, 8, 2, 5, 5, 7, 5, 3, 9, 3, 3, 7, 8, 1, ...
## $ relev     <dbl> 5, 10, 10, 4, 8, 8, 9, 10, 5, 5, 7, 7, 5, 3, 9, 6, 7, 6, 9, 1...
## $ loyal     <dbl> 1, 10, 1, 1, 1, 9, 6, 10, 1, 5, 1, 7, 1, 3, 7, 1, 6, 5, 8, 10...
## $ popul     <dbl> 5, 10, 6, 5, 6, 9, 8, 10, 5, 8, 8, 8, 4, 6, 8, 8, 8, 9, 8, 10...
## $ category  <dbl> 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 26, 2...
## $ brand     <dbl> 348, 348, 348, 348, 348, 348, 348, 348, 348, 348, 348, 348, 3...
## $ familbin  <dbl> 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1...
## $ uniqubin  <dbl> 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 1...
## $ relevbin  <dbl> 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 1, 1...
## $ loyalbin  <dbl> 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1...
## $ populbin  <dbl> 0, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1...
## $ region    <dbl> 3, 3, 3, 4, 1, 4, 3, 1, 4, 4, 1, 4, 3, 4, 3, 4, 4, 4, 2, 3, 3...

```

```
tibble(Base1)
```

```
## # A tibble: 1,500 × 17
##   gender age children income famil uniqu relev loyal popul category brand
##   <dbl> <dbl>   <dbl>   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl> <dbl>
## 1     1    60      2     4     1     1     5     1     5     26    348
## 2     1    56      2     3    10    10    10    10    10     26    348
## 3     1    51      2     4    10     1    10     1     6     26    348
## 4     1    52      1     3     7     5     4     1     5     26    348
## 5     1    69      2     2     8     5     8     1     6     26    348
## 6     1    47      2     2     9     8     8     9     9     26    348
## 7     1    49      2     1     8     8     9     6     8     26    348
## 8     1    58      2     2    10     8    10    10    10     26    348
## 9     1    44      2     4     9     2     5     1     5     26    348
## 10    2    50      2     4     7     5     5     5     8     26    348
## # i 1,490 more rows
## # i 6 more variables: familbin <dbl>, uniqubin <dbl>, relevbin <dbl>,
## #   loyalbin <dbl>, populbin <dbl>, region <dbl>
```

```
df_status(Base1)
```

```
##   variable q_zeros p_zeros q_na p_na q_inf p_inf   type unique
## 1   gender      0    0.00  0 0.00    0    0 numeric      2
## 2     age      0    0.00 44 2.93    0    0 numeric     47
## 3 children      0    0.00  0 0.00    0    0 numeric      2
## 4   income      0    0.00  0 0.00    0    0 numeric      4
## 5    famil      0    0.00  9 0.60    0    0 numeric     10
## 6   uniqu      0    0.00  5 0.33    0    0 numeric     10
## 7    relev      0    0.00 38 2.53    0    0 numeric     10
## 8    loyal      0    0.00  0 0.00    0    0 numeric     10
## 9    popul      0    0.00  9 0.60    0    0 numeric     10
## 10 category      0    0.00  0 0.00    0    0 numeric      1
## 11   brand      0    0.00  0 0.00    0    0 numeric      5
## 12 familbin    770  51.33  9 0.60    0    0 numeric      2
## 13 uniqubin   1102  73.47  5 0.33    0    0 numeric      2
## 14 relevbin    905  60.33 38 2.53    0    0 numeric      2
## 15 loyalbin   1265  84.33  0 0.00    0    0 numeric      2
## 16 populbin    710  47.33  9 0.60    0    0 numeric      2
## 17   region      0    0.00  0 0.00    0    0 numeric      4
```

Transformación de la data

```

convert_to_factor <- function(data, column_name, levels) {

  data[[column_name]] <- factor(data[[column_name]], labels = levels)
  return(data)
}

Base1 <- convert_to_factor(Base1, "brand", c("Can Air", "Maple Leaf", "Charter or Low cost", "U
K", "AirUSA"))
Base1 <- convert_to_factor(Base1, "gender", c("Femenino", "Masculino"))
Base1 <- convert_to_factor(Base1, "children", c("Si", "No"))
Base1 <- convert_to_factor(Base1, "income", c("1. <$30k", "2. [$30k-$50k>", "3. [$50k-$75k>",
"4. >= $75k"))
Base1 <- convert_to_factor(Base1, "region", c("1. Maritimes", "2. Quebec", "3. Ontario", "4. Wes
t"))

factor_columns <- c("famil", "uniqu", "relev", "loyal", "popul")
Base1[factor_columns] <- lapply(Base1[factor_columns], as.factor)

# Conversion of binary columns to factors with specified levels
Base1 <- convert_to_factor(Base1, "familbin", c("No familiar", "Familiar"))
Base1 <- convert_to_factor(Base1, "uniqubin", c("No unico", "Unico"))
Base1 <- convert_to_factor(Base1, "relevbin", c("No relevante", "relevante"))
Base1 <- convert_to_factor(Base1, "loyalbin", c("No Leal", "Leal"))
Base1 <- convert_to_factor(Base1, "populbin", c("No Popular", "Popular"))

```

Pruebas/Gráficas Estadísticas

```

dt<-table(Base1$loyal, Base1$brand)
dt

```

```

##
##      Can Air Maple Leaf Charter or Low cost  UK AirUSA
##  1      85      87              122 142      116
##  2      23      15              29 31       28
##  3      10      12              20 21       22
##  4      11       9              15 16       16
##  5      58      56              54 53       59
##  6      27      23              12 12       12
##  7      20      18              10 11       10
##  8      36      41              19 12       11
##  9      20      19              13 5        2
## 10      23      21              6 4         3

```

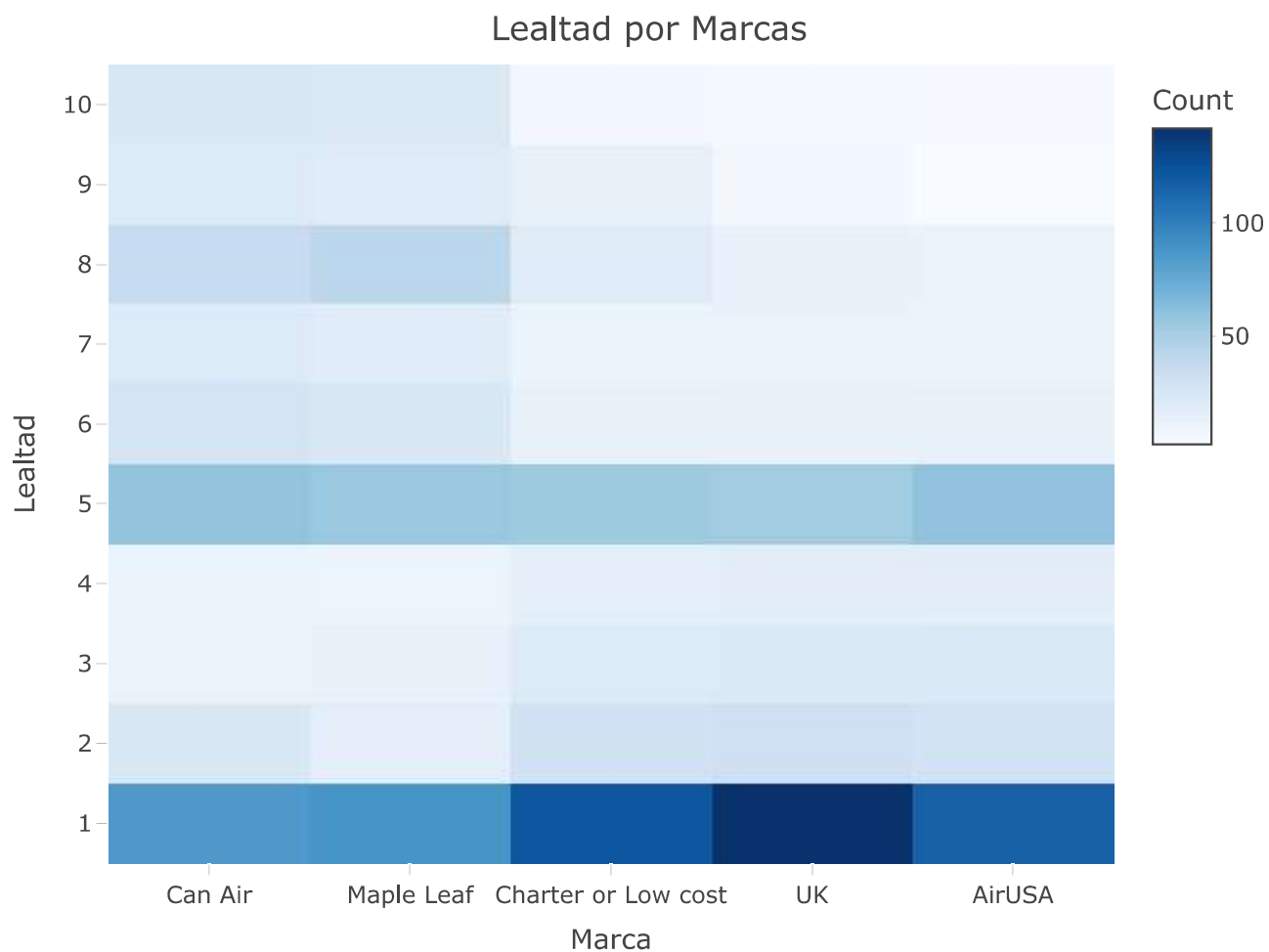
```

#Graficando La relación
# Convert table to a data frame for plotting
dt_df <- as.data.frame(dt)
colnames(dt_df) <- c("Loyalty", "Brand", "Count")

# Generate an interactive heatmap with plotly
fig <- plot_ly(
  data = dt_df,
  x = ~Brand,
  y = ~Loyalty,
  z = ~Count,
  type = "heatmap",
  colors = "Blues"
) %>%
  layout(
    title = "Lealtad por Marcas",
    xaxis = list(title = "Marca"),
    yaxis = list(title = "Lealtad")
  )

# Display the plot
fig

```



```

# Convert Perfil_fila to a data frame for easier plotting
dt.row = dt/rowSums(dt)

mass.row = colMeans(dt.row)

Perfil_fila=rbind(dt.row,mass.row)

Perfil_fila_df <- as.data.frame(Perfil_fila)
Perfil_fila_df$Category <- rownames(Perfil_fila_df)

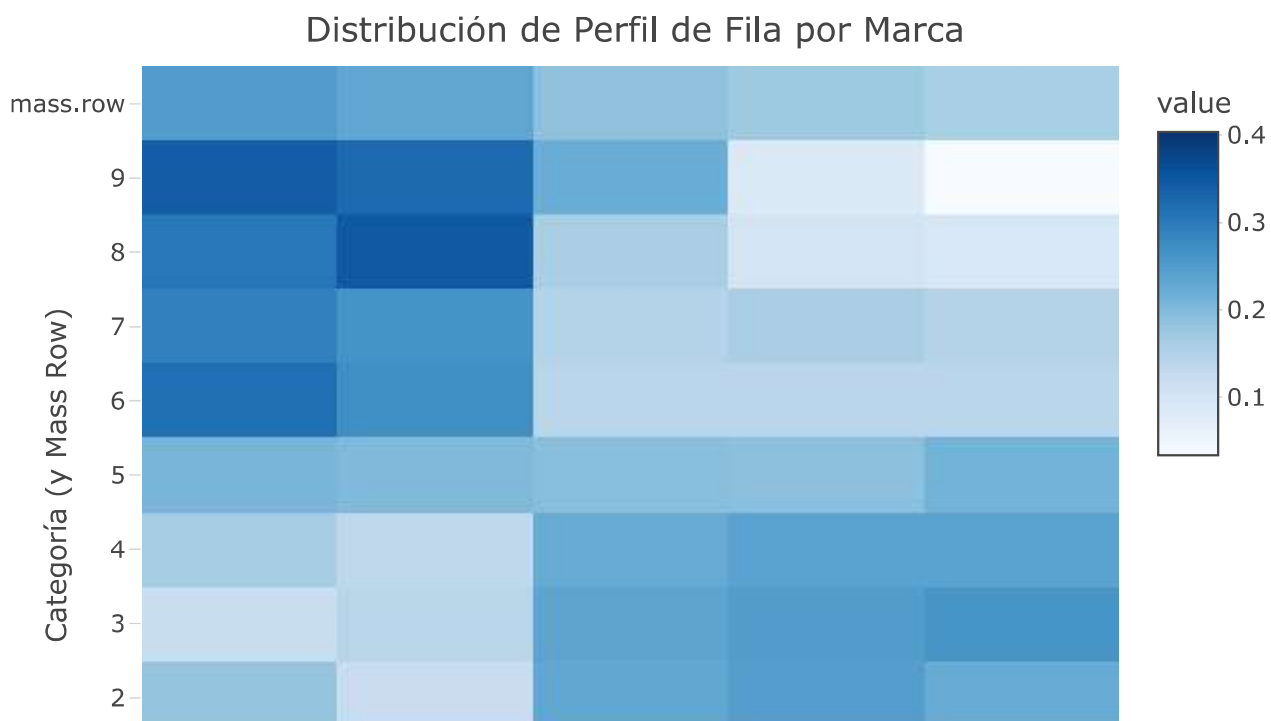
# Melt data for plotting (assuming reshape2 library)

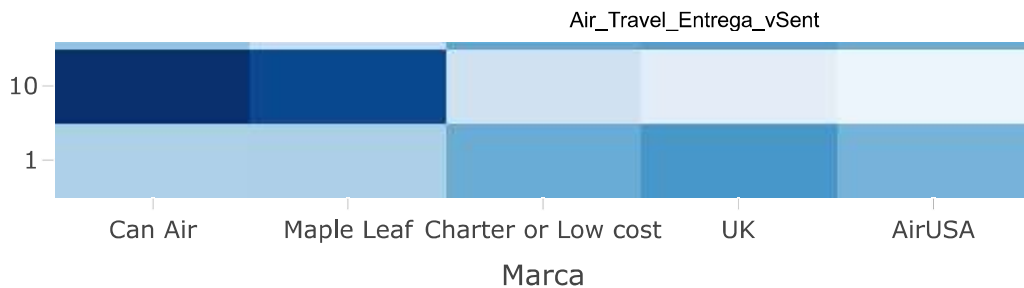
Perfil_fila_melted <- melt(Perfil_fila_df, id.vars = "Category")

# Create an interactive heatmap
fig <- plot_ly(
  data = Perfil_fila_melted,
  x = ~variable,
  y = ~Category,
  z = ~value,
  type = "heatmap",
  colors = "Blues"
) %>%
  layout(
    title = "Distribución de Perfil de Fila por Marca",
    xaxis = list(title = "Marca"),
    yaxis = list(title = "Categoría (y Mass Row)")
  )

# Show the plot
fig

```





```
# Calculate column profile
dt.col <- t(dt) / colSums(dt)
mass.col <- rowMeans(t(dt.col))

# Combine column profile with mass column
Perfil_col <- rbind(dt.col, mass.col)

Perfil_col_df <- as.data.frame(Perfil_col)
Perfil_col_df$Category <- rownames(Perfil_col_df)

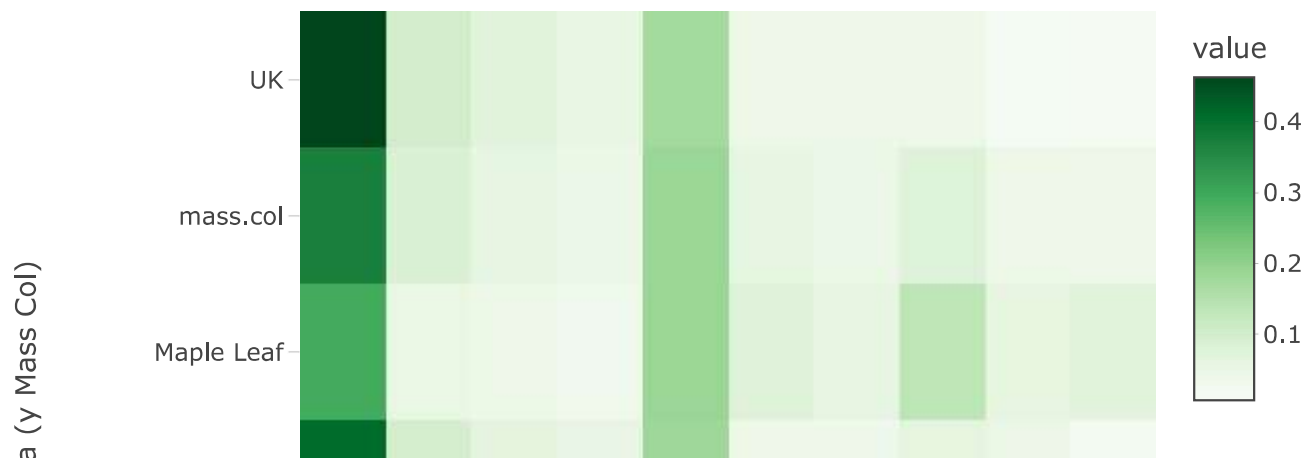
Perfil_col_melted <- melt(Perfil_col_df, id.vars = "Category")

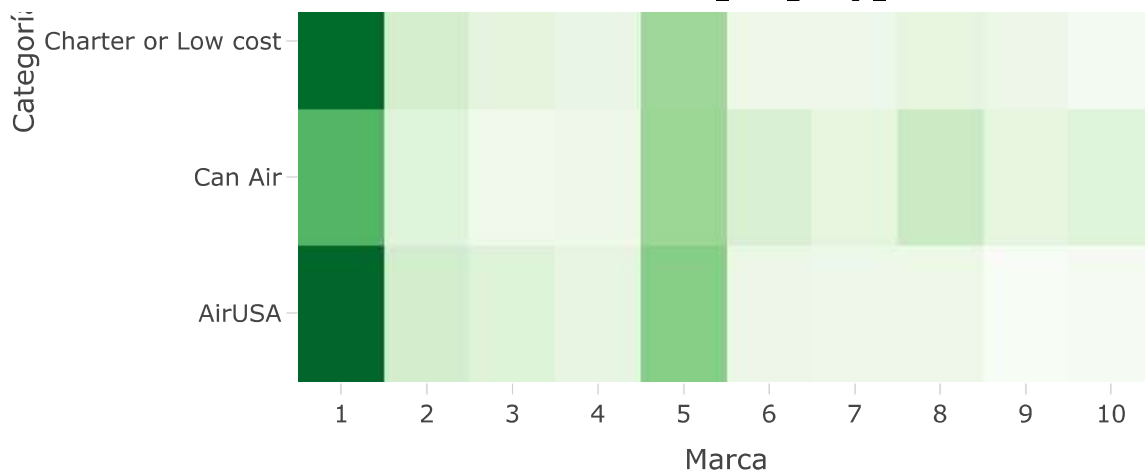
# Plot for Perfil Col (Column Profile)
fig_col <- plot_ly(
  data = Perfil_col_melted,
  x = ~variable,
  y = ~Category,
  z = ~value,
  type = "heatmap",
  colors = "Greens"
) %>%
  layout(
    title = "Distribución de Perfil de Columna por Marca",
    xaxis = list(title = "Marca"),
    yaxis = list(title = "Categoría (y Mass Col)")
  )

# Show the plots

fig_col
```

Distribución de Perfil de Columna por Marca





3. Prueba de Independencia

Para interpretar el análisis de correspondencia, el primer paso es evaluar si existe una dependencia significativa entre las filas y las columnas.

Hipótesis

H0: Asume que no existe asociación entre la marca y la lealtad H1: Asume que existe asociación entre la marca y la lealtad

```
chisq = chisq.test(dt)
chisq
```

```
##
## Pearson's Chi-squared test
##
## data: dt
## X-squared = 143, df = 36, p-value = 1.084e-14
```

```
'#Debido a que el valor de p es menor que el 0.05 ( $\alpha$ ), la decisión es rechazar la hipótesis nula.
#Entonces, con base en la prueba de Chi-Cuadrado, se puede concluir que existe asociación entre
#la marca y la lealtad_bin con el nivel de significancia del 0.05.'
```

```
## [1] "#Debido a que el valor de p es menor que el 0.05 ( $\alpha$ ), la decisión es rechazar la hipótesis nula.
#Entonces, con base en la prueba de Chi-Cuadrado, se puede concluir que existe asociación entre
#la marca y la lealtad_bin con el nivel de significancia del 0.05."
```

4. Análisis de correspondencia


```

# Perform Correspondence Analysis (CA) if not already done
res.ca <- CA(dt, graph = FALSE)

# Extract row and column coordinates and ensure correct naming
row_coords <- as.data.frame(res.ca$row$coord)
colnames(row_coords) <- paste0("Dim.", 1:ncol(row_coords)) # Rename columns to Dim.1, Dim.2, et
c.
row_coords$Category <- rownames(row_coords)

col_coords <- as.data.frame(res.ca$col$coord)
colnames(col_coords) <- paste0("Dim.", 1:ncol(col_coords)) # Rename columns to Dim.1, Dim.2, et
c.
col_coords$Brand <- rownames(col_coords)

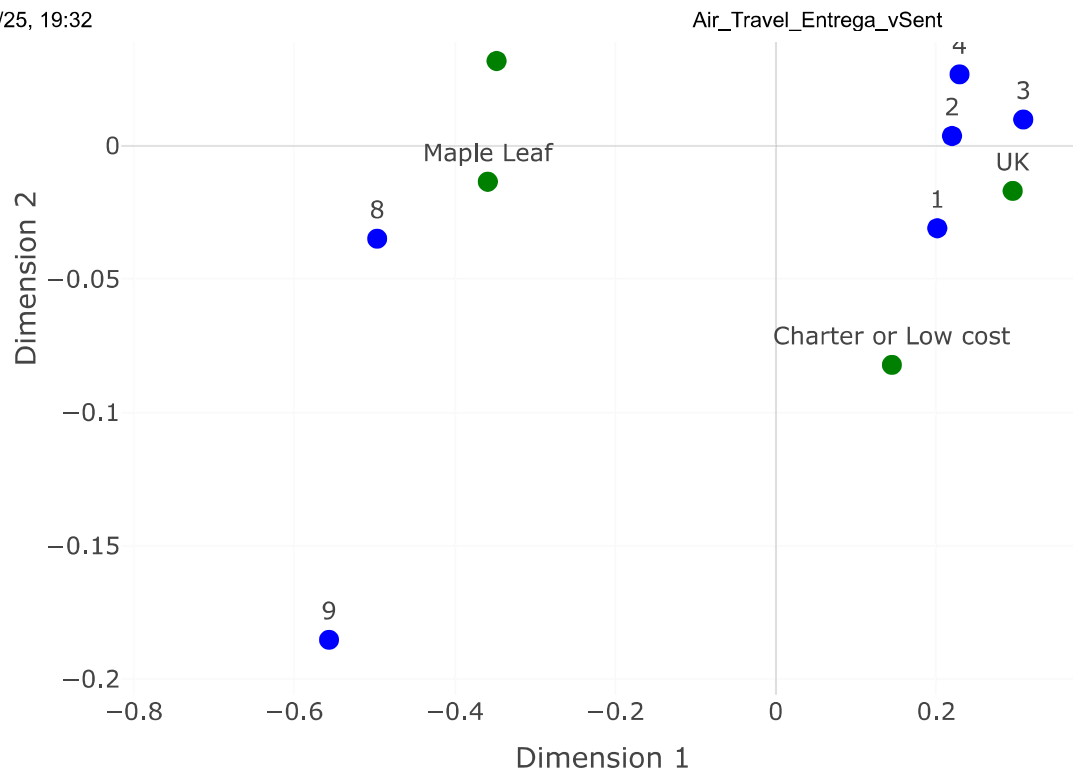
fig_combined <- plot_ly() %>%
  add_trace(
    data = row_coords,
    x = ~Dim.1,
    y = ~Dim.2,
    type = "scatter",
    mode = "markers+text",
    text = ~Category,
    textposition = "top center",
    marker = list(size = 10, color = 'blue'),
    name = "Categories"
  ) %>%
  add_trace(
    data = col_coords,
    x = ~Dim.1,
    y = ~Dim.2,
    type = "scatter",
    mode = "markers+text",
    text = ~Brand,
    textposition = "top center",
    marker = list(size = 10, color = 'green'),
    name = "Brands"
  ) %>%
  layout(
    title = "Correspondence Analysis - Combined Plot",
    xaxis = list(title = "Dimension 1"),
    yaxis = list(title = "Dimension 2")
  )

fig_combined

```

Correspondence Analysis - Combined Plot





4. Análisis de correspondencia

```
# Realizar el Análisis de Correspondencia (CA) si aún no se ha hecho
Base1_filtered <- droplevels(subset(Base1, !(brand %in% c("UK", "AirUSA"))))

dt_fil<-table(Base1_filtered$loyal, Base1_filtered$brand)
dt_fil
```

```
##
##      Can Air Maple Leaf Charter or Low cost
##  1      85      87      122
##  2      23      15      29
##  3      10      12      20
##  4      11       9      15
##  5      58      56      54
##  6      27      23      12
##  7      20      18      10
##  8      36      41      19
##  9      20      19      13
## 10      23      21       6
```

```
#Graficando la relación
# Convert table to a data frame for plotting
dt_df_fil <- as.data.frame(dt_fil)
colnames(dt_df_fil) <- c("Loyalty", "Brand", "Count")
dt_df_fil
```

##	Loyalty	Brand	Count
## 1	1	Can Air	85
## 2	2	Can Air	23
## 3	3	Can Air	10
## 4	4	Can Air	11
## 5	5	Can Air	58
## 6	6	Can Air	27
## 7	7	Can Air	20
## 8	8	Can Air	36
## 9	9	Can Air	20
## 10	10	Can Air	23
## 11	1	Maple Leaf	87
## 12	2	Maple Leaf	15
## 13	3	Maple Leaf	12
## 14	4	Maple Leaf	9
## 15	5	Maple Leaf	56
## 16	6	Maple Leaf	23
## 17	7	Maple Leaf	18
## 18	8	Maple Leaf	41
## 19	9	Maple Leaf	19
## 20	10	Maple Leaf	21
## 21	1 Charter or Low cost		122
## 22	2 Charter or Low cost		29
## 23	3 Charter or Low cost		20
## 24	4 Charter or Low cost		15
## 25	5 Charter or Low cost		54
## 26	6 Charter or Low cost		12
## 27	7 Charter or Low cost		10
## 28	8 Charter or Low cost		19
## 29	9 Charter or Low cost		13
## 30	10 Charter or Low cost		6

```

res.ca1 <- CA(dt_fil, graph = FALSE)

row_coords1 <- as.data.frame(res.ca1$row$coord)
colnames(row_coords1) <- paste0("Dim.", 1:ncol(row_coords1)) # Renombrar columnas a Dim.1, Dim.
2, etc.
row_coords1$Category1 <- rownames(row_coords1)

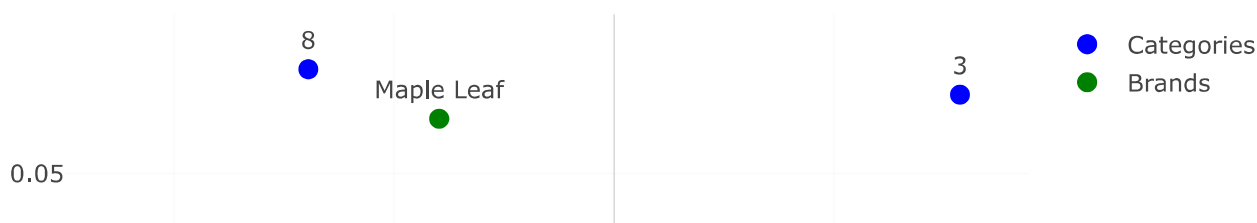
col_coords1 <- as.data.frame(res.ca1$col$coord)
colnames(col_coords1) <- paste0("Dim.", 1:ncol(col_coords1)) # Renombrar columnas a Dim.1, Dim.
2, etc.
col_coords1$Brand1 <- rownames(col_coords1)

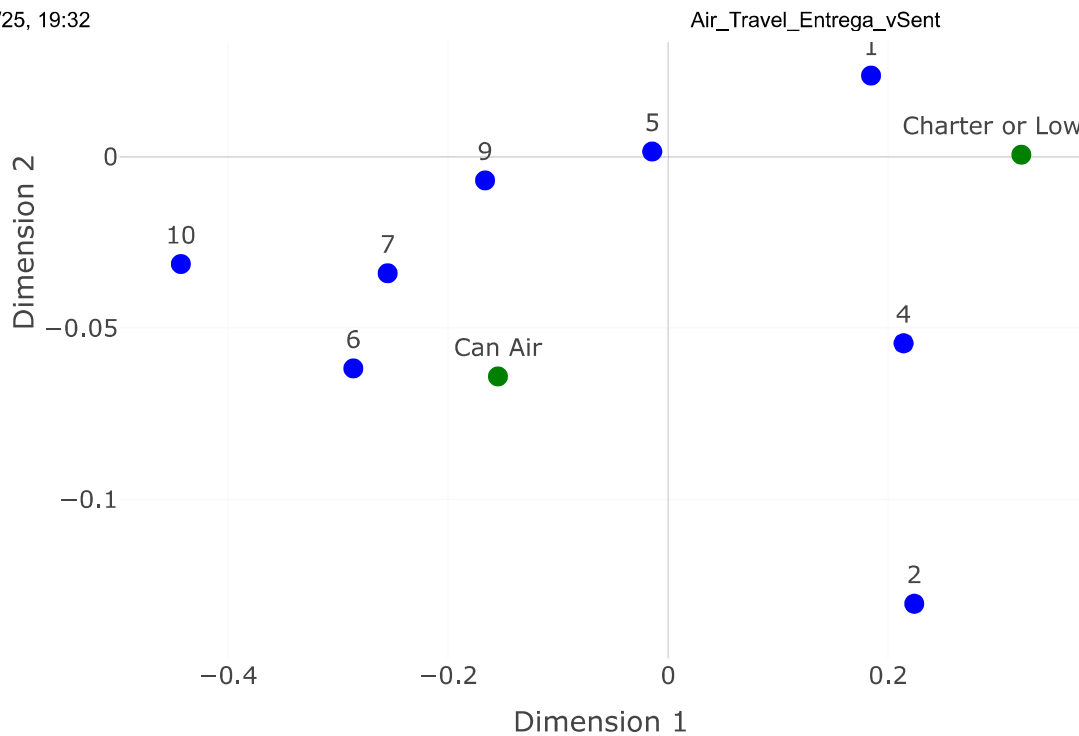
# Gráfico combinado de puntos de fila y columna
fig_combined1 <- plot_ly() %>%
  add_trace(
    data = row_coords1,
    x = ~Dim.1,
    y = ~Dim.2,
    type = "scatter",
    mode = "markers+text",
    text = ~Category1,
    textposition = "top center",
    marker = list(size = 10, color = 'blue'),
    name = "Categories"
  ) %>%
  add_trace(
    data = col_coords1,
    x = ~Dim.1,
    y = ~Dim.2,
    type = "scatter",
    mode = "markers+text",
    text = ~Brand1,
    textposition = "top center",
    marker = list(size = 10, color = 'green'),
    name = "Brands"
  ) %>%
  layout(
    title = "Correspondence Analysis - Combined Plot",
    xaxis = list(title = "Dimension 1"),
    yaxis = list(title = "Dimension 2")
  )

# Mostrar el gráfico interactivo combinado
fig_combined1

```

Correspondence Analysis - Combined Plot





Concluisiones

En promedio, la marca Can Air tiene mayor lealtad con un 24.74%, le siguen Maple Leaf, Charter or Low Cost, Uk mientras que AirUSA es la marca con menor lealtad con un 16.04%.

1.

El análisis de correspondencia revela patrones claros en la lealtad del cliente hacia diferentes marcas. UK se asocia principalmente con categorías de lealtad baja (1 a 4), lo que sugiere que esta marca atrae a clientes con menor fidelidad. Can Air se vincula con categorías de lealtad intermedia (6 y 7), mientras que AirUSA se encuentra en una posición aislada, asociándose con niveles de lealtad relativamente bajos, con valores cercanos a 3, 4 y 5. Maple Leaf muestra una conexión con la categoría de lealtad media-alta (8), mientras que Charter or Low cost se vincula con la categoría de lealtad más baja, reforzando la idea de una percepción de menor compromiso de sus clientes. Estos resultados brindan información valiosa para orientar estrategias de fidelización según el perfil de cada marca.

2. El análisis de correspondencia revela patrones claros en la lealtad del cliente hacia diferentes marcas. Can Air se sigue vinculando con categorías de lealtad intermedia (6 y 7). Maple Leaf sigue mostrando una conexión con la categoría de lealtad media-alta (8), Mientras que Charter or Low cost se sigue vinculando con la categoría de lealtad más baja, reforzando la idea de una percepción de menor compromiso de sus clientes.

Observación: Al eliminar las categorías UK y AirUSA, las relaciones entre las demás aerolíneas y niveles de lealtad se mantienen similares, lo que indica que la estructura general de la matriz de datos sigue capturando la varianza principal y que estas categorías no tenían un impacto significativo en la disposición global de las asociaciones. ”

Recomendaciones

Recomendamos el enfoque de compromiso con el cliente ya que encontramos que esta variable podría significar un mayor grado de lealtad.