



A decorative background featuring a grid pattern with various colored brushstrokes in shades of red, orange, yellow, green, and blue.

Coderhouse

Curso Data Analytics

Comisión 42295

Profesora: Marina Pross
Tutora Adjunta: Rocío López
Tutora: Nicole Vollmar

Proyecto Travel Company

Data Analysis

Autora: Valentina Avariano

31/08/2023

Nombre y Apellido: Valentina Avariano
Comisión del curso: 42295
Fecha: 31/08/2023

ENTREGA FINAL: Proyecto Travel Company

Data Analytics

1. TABLA DE VERSIONES

| | |
|--------------------|------------|
| Versión 1.0 | 01/06/2023 |
| Versión 2.0 | 22/06/2023 |
| Versión 3.0 | 10/08/2023 |
| Versión 4.0 | 31/08/2023 |

2. DATASET

La base de datos utilizada para este trabajo se encuentra disponible a través del siguiente link:

[Dataset.TravelCompany](#)

Por favor haga clic y siga las instrucciones para dirigirse al enlace.

3. INTRODUCCIÓN Y TEMÁTICA DE DATOS

Con el fin de realizar el presente trabajo, se tomó información proveniente de una base de datos de la empresa "Trips & Travel.Com", la cual quiere habilitar y establecer un modelo de negocio viable para ampliar la base de clientes. La compañía ahora planea lanzar un nuevo producto, es decir, Paquete de Turismo de Bienestar. El turismo de bienestar se define como un viaje que permite al viajero mantener, mejorar o impulsar un estilo de vida saludable y apoyar o aumentar la sensación de bienestar.

Sin embargo, esta vez la empresa quiere aprovechar los datos disponibles de los clientes existentes y potenciales para hacer que los gastos de marketing sean más eficientes. El período a examinar comienza desde el 20/02/21 20 de Febrero 2021 hasta el 09 de Mayo de 2023.

4. HIPÓTESIS

El propósito de esta investigación es analizar los datos disponibles de los clientes existentes y potenciales para poder reducir los gastos de marketing para poder lanzar un nuevo paquete vacacional, es decir, Paquete de Turismo de Bienestar.

Adicionalmente, se explorará si la ocupación de los clientes, el género, su lugar de residencia, el estado civil y sus sueldos promedios, pueden determinar un perfil de cliente potencial para el lanzamiento del nuevo producto de la compañía de viajes.

5. ALCANCE

Este proyecto se ha desarrollado con la finalidad de alcanzar un nivel táctico de análisis, ya que podrá ser utilizado por ejecutivos y directores de la empresa para apuntar control de gastos, reducciones y foco de inversiones.

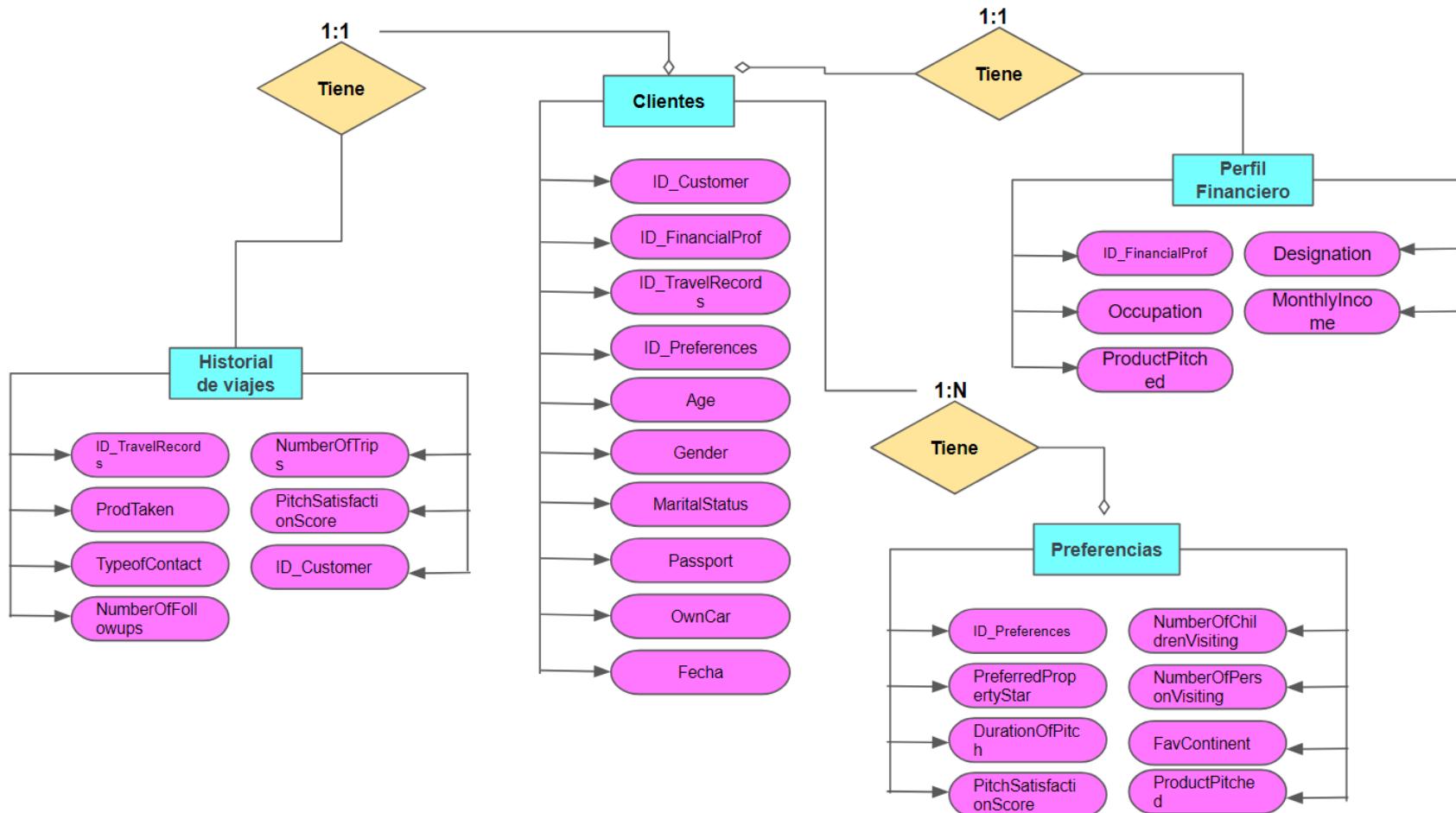
El aprovechamiento de los resultados podrá proporcionar información y diversas perspectivas que permitan mejorar el proceso de tomar decisiones vinculadas a lanzamiento de productos turísticos.

6. HERRAMIENTAS TECNOLÓGICAS UTILIZADAS

Se hizo uso de los siguientes programas para dar forma a este proyecto:

- **Google Sheets y Excel:** para el dataset.
- **Presentaciones de Google:** para el diagrama DER y fondos utilizados en el Dashboard
- **Power BI Desktop:** para la realización de los dashboards.

7. DIAGRAMA ENTIDAD-RELACIÓN



8. LISTADO DE TABLAS

A continuación se presentará un listado de las tablas que se llevarán a cabo para analizar el conjunto de datos, así como una breve descripción y la definición de las claves primarias y foráneas según corresponda:

| CLIENTES | |
|---|---------------|
| Descripción: Posee los datos principales de los potenciales clientes que dejaron su consulta sobre paquetes de viaje a través de plataformas digitales de la compañía. | |
| Campo | Tipo de Clave |
| ID_Customer | int |
| ID_FinancialProf | int |
| ID_TravelRecords | int |
| ID_Preferences | int |
| Age | int |
| Gender | nvarchar(20) |
| MaritalStatus | nvarchar(20) |
| Passport | int |
| OwnCar | int |
| Fecha | datetime |

| PERFIL FINANCIERO | |
|--|---------------|
| Descripción: Corresponde a la información financiera de los potenciales clientes junto a su perfilamiento como viajero. | |
| Campo | Tipo de Clave |
| ID_FinancialProf | int |
| Occupation | nvarchar(20) |
| ProductPitched | nvarchar(20) |
| Designation | nvarchar(20) |
| MonthlyIncome | int |

| HISTORIAL DE VIAJE | |
|---|---------------|
| Descripción: Posee la información correspondiente al historial de viajes que han tenido o no, aquellos clientes que dejaron su consulta. | |
| Campo | Tipo de Clave |
| ID_TravelRecords | int |
| ProdTaken | nvarchar(20) |
| TypeofContact | nvarchar(20) |
| NumberOfFollowups | int |
| NumberOfTrips | int |

| | |
|------------------------|-----|
| PitchSatisfactionScore | int |
| ID_Customer | int |

| PREFERENCIAS | |
|--|---------------|
| Descripción: Corresponde a la información sobre las preferencias en cuanto a tiempos, estadías y lugares que derivaron los potenciales clientes en sus consultas. | |
| Campo | Tipo de Clave |
| ID_Preferences | int |
| PreferredPropertyStar | int |
| DurationOfPitch | int |
| PitchSatisfactionScore | int |
| NumberOfChildrenVisiting | int |
| NumberOfPersonVisiting | int |
| FavContinent | nvarchar(20) |
| ProductPitched | nvarchar(20) |

9. LISTADO DE COLUMNAS POR TABLAS

En esta sección se compartirán las diversas listas de las tablas que conformarán la base de datos de este proyecto, junto al detalle de las columnas definiendo: campos, tipos de campos y claves, para profundizar en la comprensión de las mismas.

| CLIENTEs | | |
|------------------|---------------|---------------|
| Campo | Tipo de Campo | Tipo de Clave |
| ID_Customer | int | PK |
| ID_FinancialProf | int | FK |
| ID_TravelRecords | int | FK |
| ID_Preferences | int | FK |
| Age | int | - |
| Gender | nvarchar(20) | - |
| MaritalStatus | nvarchar(20) | - |
| Passport | int | - |
| OwnCar | int | - |
| Fecha | datetime | - |

| PERFIL FINANCIERO | | |
|-------------------|---------------|---------------|
| Campo | Tipo de Campo | Tipo de Clave |
| ID_FinancialProf | int | PK |
| Occupation | nvarchar(20) | - |
| ProductPitched | nvarchar(20) | - |
| Designation | nvarchar(20) | - |
| MonthlyIncome | int | - |

| HISTORIAL DE VIAJES | | |
|------------------------|---------------|---------------|
| Campo | Tipo de Campo | Tipo de Clave |
| ID_TravelRecords | int | PK |
| ProdTaken | nvarchar(20) | - |
| TypeofContact | nvarchar(20) | - |
| NumberOfFollowups | int | - |
| NumberOfTrips | int | - |
| PitchSatisfactionScore | int | - |
| ID_Customer | int | FK |

| PREFERENCIAS | | |
|--------------------------|---------------|---------------|
| Campo | Tipo de Campo | Tipo de Clave |
| ID_Preferences | int | PK |
| PreferredPropertyStar | int | - |
| DurationOfPitch | int | - |
| PitchSatisfactionScore | int | - |
| NumberOfChildrenVisiting | int | - |
| NumberOfPersonVisiting | int | - |
| FavContinent | nvarchar(20) | - |
| ProductPitched | nvarchar(20) | - |

10. TRANSFORMACIONES DATASET

Las transformaciones que se hicieron en el dataset fueron las siguientes:

- a) Se depuró la base de información que no sería utilizada, como la fila de "CityTier".

| | |
|----------|---|
| F | |
| CityTier | |
| 3 | |
| c | 1 |
| c | 1 |
| c | 1 |

b) Todo se encontraba en una misma tabla como se puede observar:

| A1 | Date | CustomerID | ProdTaken | Age | TypeofContact | CityTier | DurationOfPitch | Occupation | Gender | NumberOfPersons | NumberOfFollowups | ProductPitched | PreferredProper | MaritalStatus | NumberOfTrips | Passport | PitchSatisfactionScore | OwnCar | Num |
|----|----------|------------|-----------|-----|-----------------|----------|-----------------|----------------|--------|-----------------|-------------------|----------------|-----------------|---------------|---------------|----------|------------------------|--------|-----|
| 1 | 20/02/21 | 200001 | 1 | 41 | Self Enquiry | 3 | 6 | Salaried | Female | 3 | 3 | 3 Deluxe | 3 Single | 1 | 1 | 2 | 1 | 1 | |
| 2 | 02/09/21 | 200001 | 0 | 49 | Company Invited | 1 | 14 | Male | 3 | 4 | 4 Deluxe | 4 Divorced | 2 | 0 | 3 | 3 | 1 | | |
| 3 | 20/08/21 | 200002 | 1 | 37 | Self Enquiry | 1 | 8 | Free Lancer | Male | 3 | 3 | 4 Basic | 3 Single | 7 | 1 | 3 | 0 | | |
| 4 | 24/03/21 | 200003 | 0 | 33 | Company Invited | 1 | 9 | Salaried | Female | 2 | 3 | 3 Basic | 3 Divorced | 2 | 1 | 5 | 1 | | |
| 5 | | 200004 | 0 | 32 | Self Enquiry | 1 | 8 | Small Business | Male | 2 | 3 | 3 Basic | 4 Divorced | 1 | 0 | 5 | 1 | | |
| 6 | 22/03/21 | | | | | | | | | | | | | | | | | | |

Se agruparon los datos en 4 grandes tablas: Clientes, Historial de Viajes, Perfil Financiero y Preferencias.

c) Clientes:

OBJETO TRAVEL - Editor de Power Query

El Power Query Editor muestra la tabla 'Clientes' con las siguientes columnas: Date, CustomerID, Age, Gender, MaritalStatus, Passport, OwnCar y ID_PerfilFinan. Los valores en la columna 'OwnCar' son 1, 0 y 1 respectivamente para las tres filas.

Donde se crearon los ID correspondientes al resto de las tablas ya que se contempla Clientes como la tabla principal:

| 1 ² ₃ ID_PerfilFinan | 1 ² ₃ ID_TravelRec | 1 ² ₃ ID_Preferencias |
|--|--|---|
| 1 | 1000 | 1001 |
| 1 | 1002 | 1004 |
| 0 | 1004 | 1007 |
| | | 1003 |
| | | 1007 |
| | | 1011 |

Adicional a ello, también se creó la tabla de Calendario para poder segmentar la información acorde a temporalidades:

Calendario

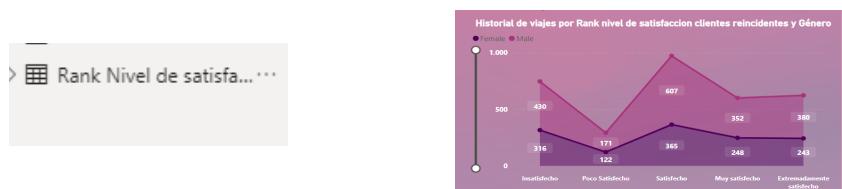
- Año
- Date
- Dia
- Fecha SK
- Mes
- Semestre

d) Historial de Viajes:

Asimismo se agruparon los datos para Historial de Viajes de la siguiente manera

| 1 ² ₃ ProdTaken | 1 ² ₃ TypeofContact | 1.2 NumberOfFollowups | 1 ² ₃ ID_TravelRec | 1 ² ₃ NumberOfTrips | 1 ² ₃ PitchSatisfactionScore | 1 ² ₃ CustomerID | ABC 123 NivelDeSatisfaccion |
|---------------------------------------|---|-----------------------|--|---|--|--|-----------------------------|
| 1 | 1 Self Enquiry | 3 | 1001 | 1 | 2 | 200000 | Poco Satisfecho |
| 2 | 0 Company Invited | 4 | 1004 | 2 | 3 | 200001 | Satisfecho |
| 3 | 1 Self Enquiry | 4 | 1007 | 7 | 3 | 200002 | Satisfecho |

Adicionalmente se elaboró una tabla con para identificar categóricamente el nivel de satisfacción de los clientes en sus experiencias de compras anteriores, yendo en una escala de "Insatisfecho" hasta "Extremadamente satisfecho" lo cual resultó en lo siguiente:



e) Perfil Financiero:

Se agruparon los datos para Perfil Financiero de la siguiente manera:

| ID_PerfilFinan | Designation | ProductPitched | Occupation | MonthlyIncome |
|----------------|-------------|----------------|-------------|---------------|
| 1 | Manager | Deluxe | Salaried | 20993 |
| 2 | Manager | Deluxe | Salaried | 20130 |
| 3 | Executive | Basic | Free Lancer | 17090 |
| 4 | Executive | Basic | Salaried | 17909 |

f) Preferencias

Se agruparon los datos para Preferencias de viaje de la siguiente manera:

| ID_Preferencias | FavContinent | NumberOfChildrenVisiting | PreferredPropertyStar | DurationOfPitch | ProductPitched | NumberOfPersonVisiting |
|-----------------|--------------------|--------------------------|-----------------------|-----------------|----------------|------------------------|
| 1 | 1003: Europa | 0 | 3 | 6: Deluxe | | 3 |
| 2 | 1007: Europa | 2 | 4 | 14: Deluxe | | 3 |
| 3 | 1011: Asia | 0 | 3 | 8: Basic | | 3 |
| 4 | 1015: SouthAmerica | 1 | 3 | 9: Basic | | 2 |
| 5 | 1019: NorthAmerica | 0 | 4 | 8: Basic | | 2 |

11. MODELO RELACIONAL POWER BI

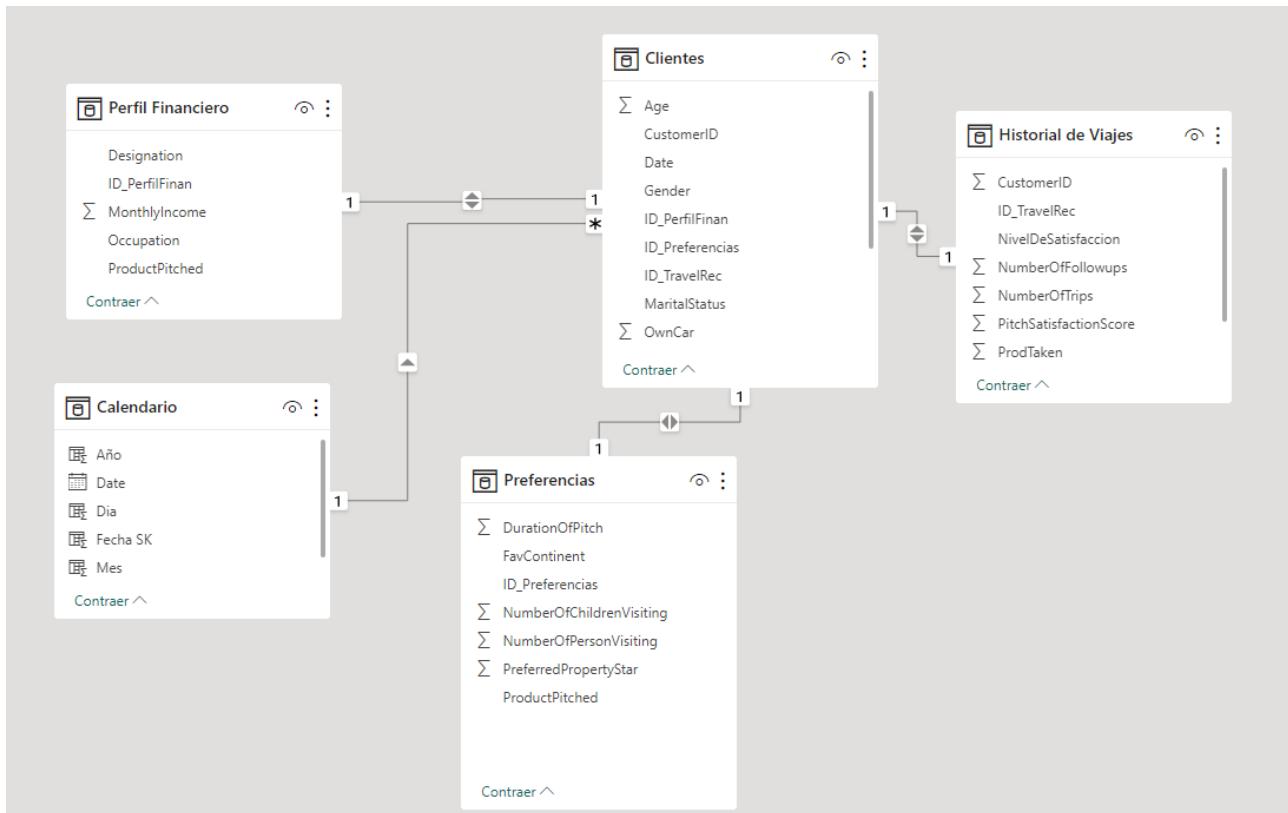
- Tabla Clientes:

- Se agrega la columna “Rango_edad” para agrupar por grupos etarios a los clientes que realizaron consultas en el sitio web.

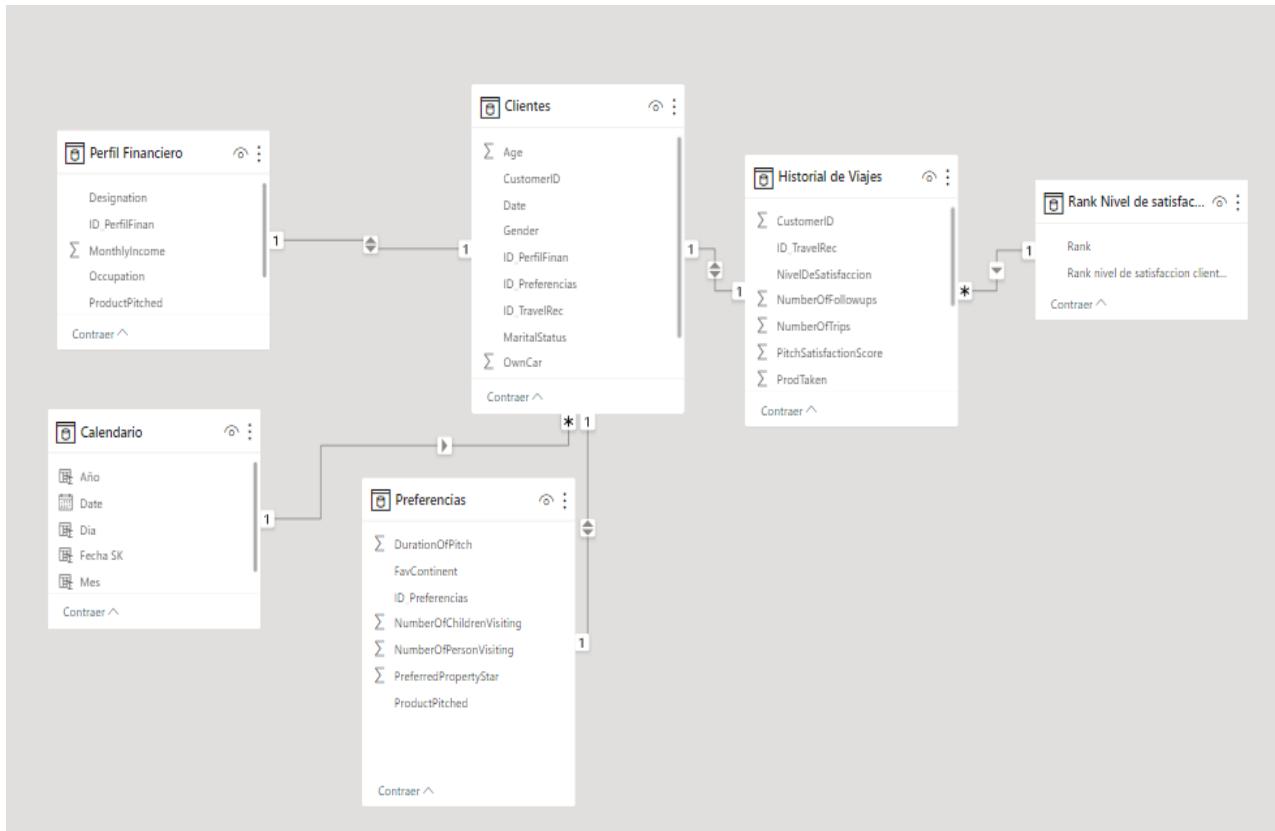
- Tabla Historial de Viajes:

- Se agrega la columna “Viajes” para segmentar por cantidad de viajes anteriores con la compañía de aquellos clientes que consultaron en el sitio web pero que ya han disfrutado de algunos servicios de viajes con la empresa.

Acorde a ello, el diagrama entidad-relación quedó graficado en Power BI de la siguiente forma:



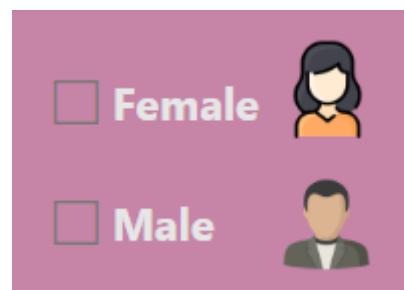
En la última versión se agrego una tabla puente de Rank Historial de viajes, resultando en:



Versión 4.0

12. SEGMENTACIONES DE DASHBOARD

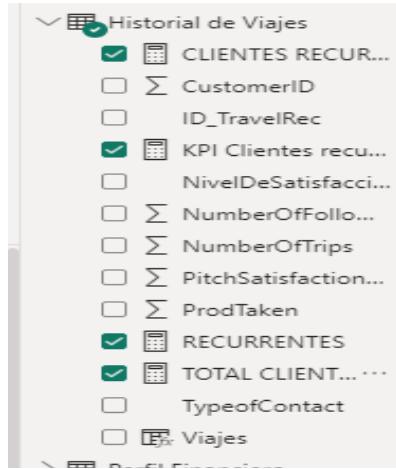
11.1. Gender: Segmentación a través de opciones junto a íconos. Se creó a través de una lista vertical con selección única o múltiple:



11.2. Año: utilizado para filtrar fechas y datos respecto a las consultas y datos brindados por los clientes potenciales. Se creó a través de un control deslizante:



13. MEDIDAS CALCULADAS



Las medidas calculadas realizadas se encuentran en la tabla de Historial de Viajes y son las siguientes:

- CLIENTES RECURRENTES = `COUNT('Historial de Viajes'[ProdTaken])`
- RECURRENTES = `CALCULATE([CLIENTES RECURRENTES],'Historial de Viajes'[ProdTaken]=0)`
- TOTAL CLIENTES = `COUNT(Clientes[CustomerID])`
- KPI Clientes recurrentes = `DIVIDE([RECURRENTES],[TOTAL CLIENTES])`

Se crearon con la finalidad de determinar el KPI de tasa de retención de clientes, en el cual se pueda visualizar el porcentaje de clientes que han vuelto a solicitar servicios de la agencia de viajes en más de una ocasión, por lo cual se mide el nivel de engagement de la base modelo.



Por otro lado, se agregó otra medida para categorizar los rangos de viajes que tuvieron los clientes con la agencia:

```
Viajes = IF('Historial de Viajes'[NumberOfTrips] = 0, "0", IF(AND('Historial de Viajes'[NumberOfTrips] >= 1, 'Historial de Viajes'[NumberOfTrips] <= 3), "De 1 a 3", IF(AND('Historial de Viajes'[NumberOfTrips] >= 4, 'Historial de Viajes'[NumberOfTrips] <= 8), "De 4 a 9", IF('Historial de Viajes'[NumberOfTrips] > 9, "Mas de 10", "Otros"))))
```

De esta forma se pudo clasificar y visualizar en 4 opciones de viajes anteriores:



14. VISUALIZACIÓN DE DATOS

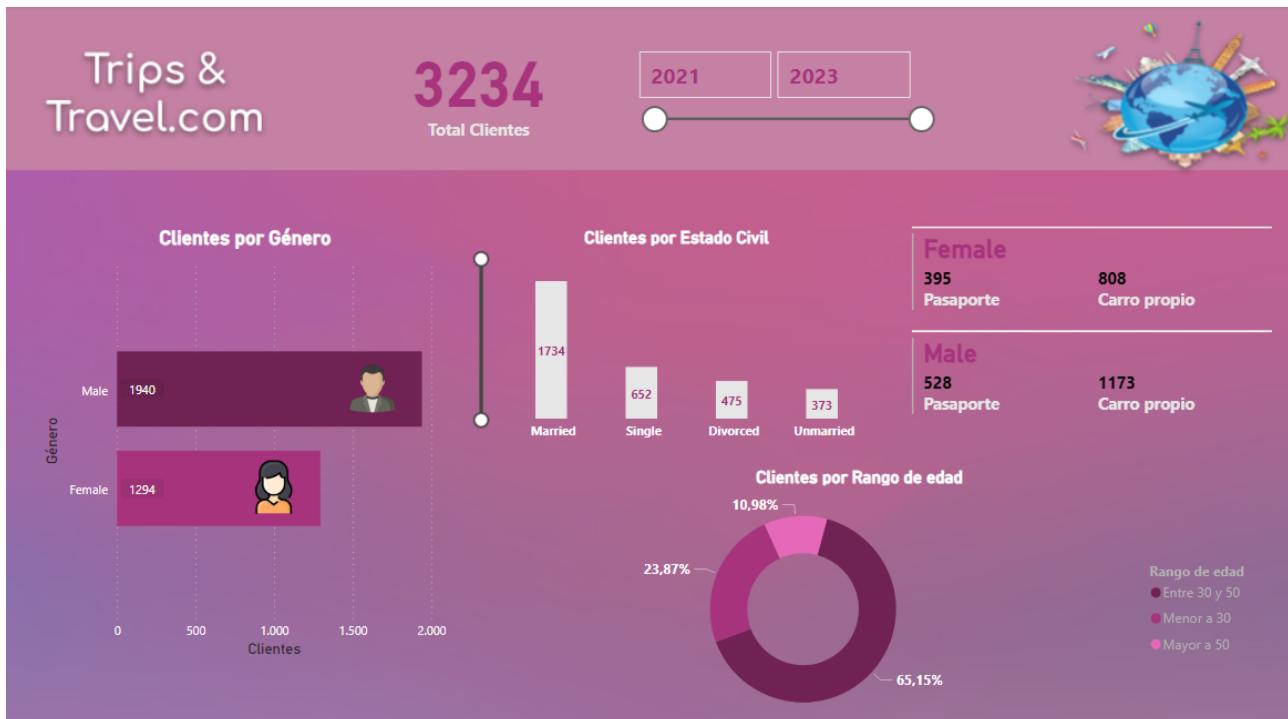
Portada

Se crea con la finalidad de obtener de forma interactiva todos los tableros en una misma página. En primer lugar se encuentra el de Clientes, luego Historial de Viajes, siguiente Perfil Financiero y por último Preferencias. Al hacer clic en cualquiera de ellas, los redirige al dashboard elegido.

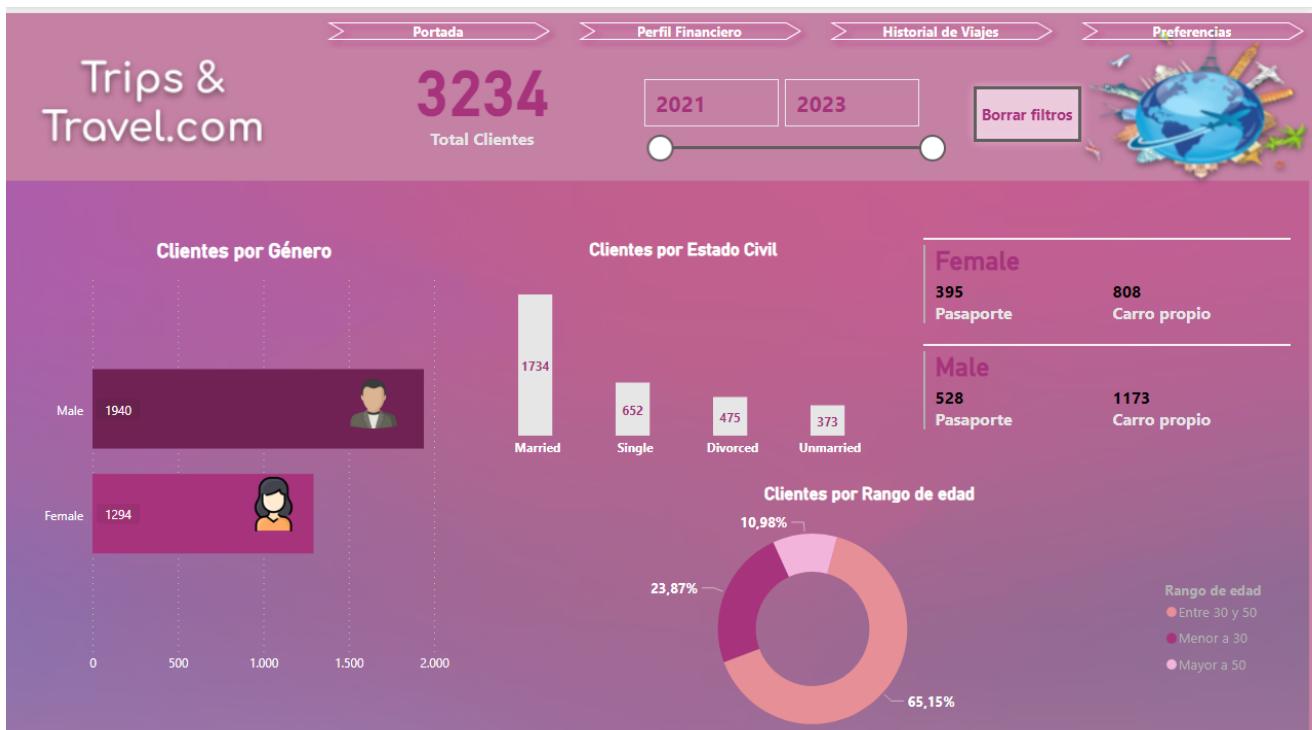


Clientes

Este tablero contiene la información correspondiente a los potenciales clientes que han realizado consultas y se encuentran en la base de datos de la empresa. Los datos mas relevantes que se pueden constatar son: Gender, Estado Civil, si el cliente cuenta con pasaporte y/o auto propio, así como la agrupación por edades.



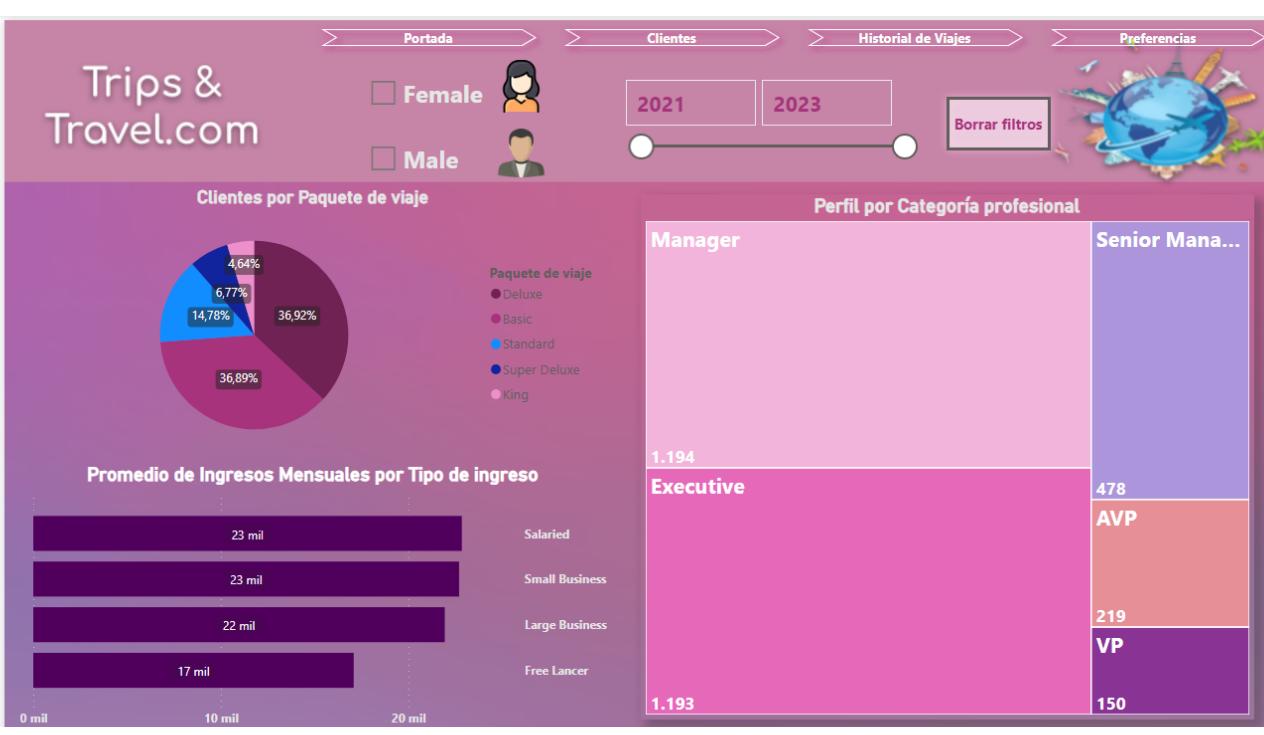
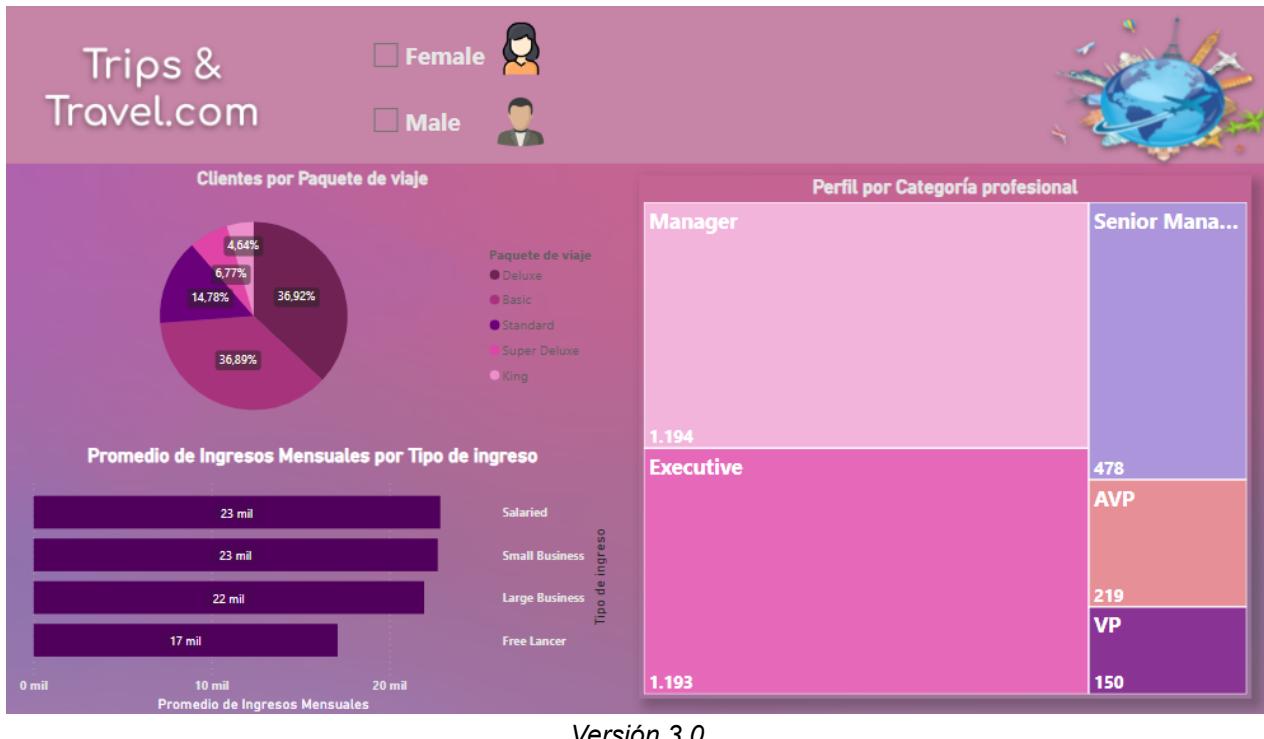
Versión 3.0



Versión 4.0

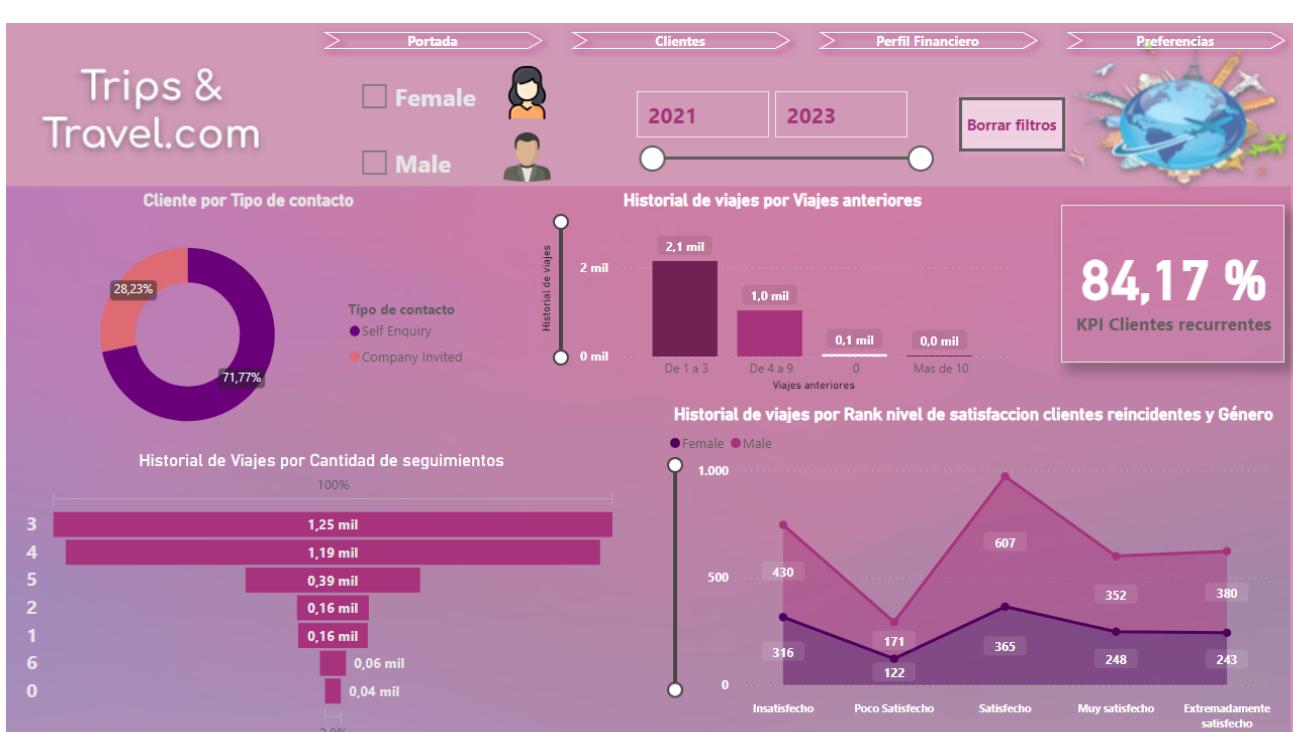
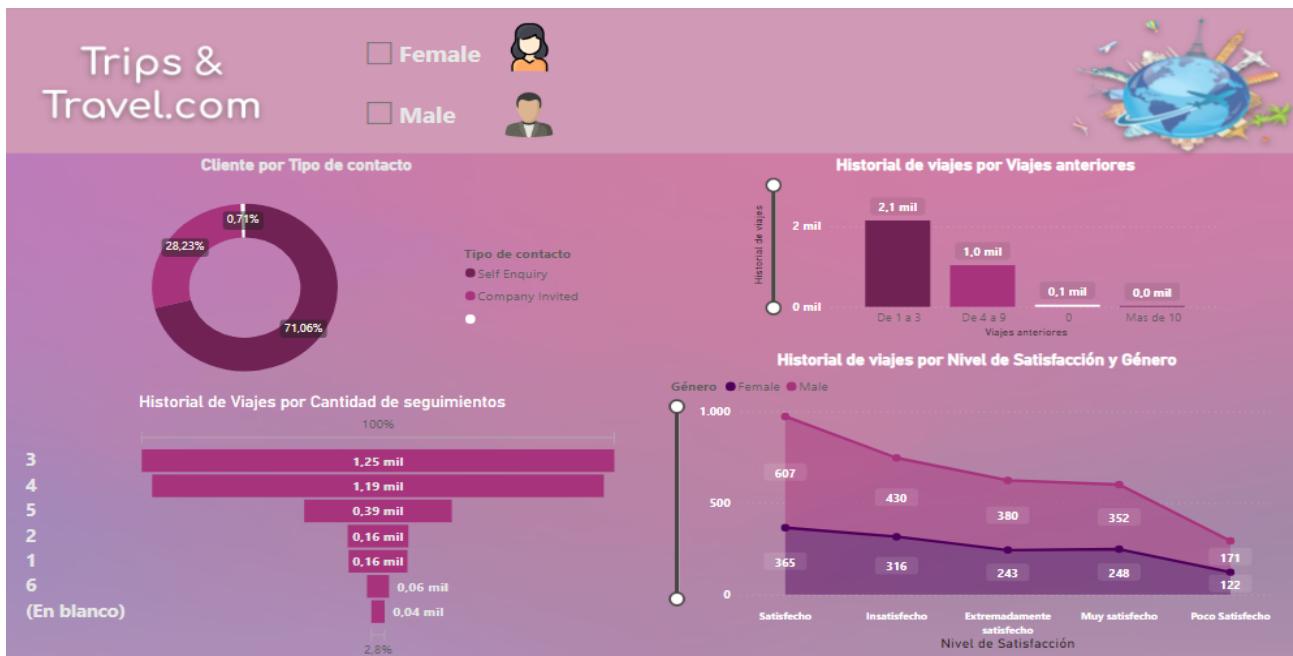
Perfil Financiero

Este tablero tiene como principal objetivo sintetizar la información de los clientes que dan a entender sus diferentes perfiles financieros para poder obtener mayor detalle sobre las capacidades económicas de los mismos. Dentro de las variables más relevantes se encuentran las elecciones del tipo de paquete de viaje que prefieren, así como la información promedio de sus sueldos mensuales y la categorización por perfil profesional según sus ocupaciones.



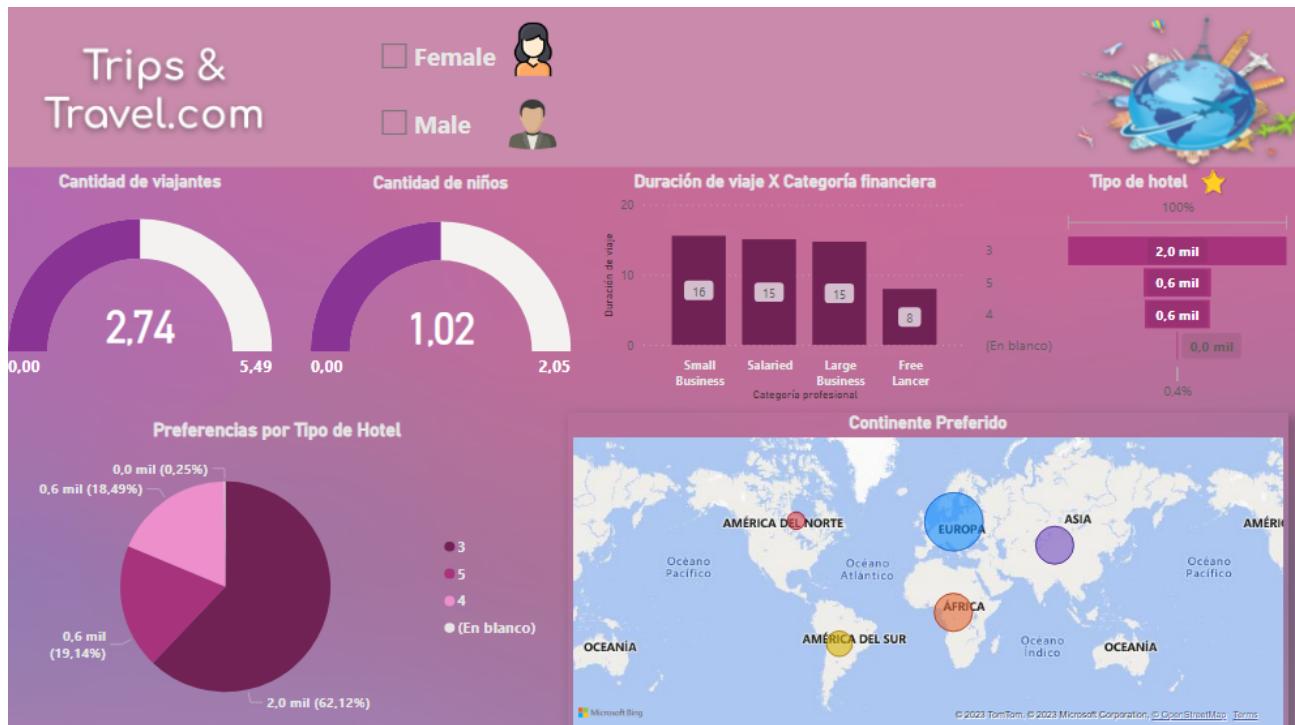
Historial de viajes

Este tablero contiene la información correspondiente al historial de viajes anteriores que los clientes han contratado con la empresa Trips & Travel.com, lo cual arroja información sobre potenciales reingresos de clientes, dando cuenta de nivel de engagement que han tenido según sus experiencias. Por lo que dentro de las variables a considerar se encuentran el nivel de satisfacción de experiencias pasadas, la cantidad de seguimientos por parte del área comercial así como el tipo de contacto realizado.



Preferencias

Este tablero contiene la información correspondiente a las condiciones que prefieren los clientes al momento de realizar una reserva de viaje. Dentro de las principales, se encuentran las cantidades de personas y niños potenciales que asistieron al viaje, así como el tipo de hotel y los continentes favoritos para su elección turística.



15. CONCLUSIONES

- La base de datos de la empresa posee 3.234 clientes desde el 2021 hasta 2023.
- La mayoría de los clientes que realizaron consultas directas o indirectas sobre algún paquete turístico de la agencia fueron personas casadas de entre 30 a 50 años que se encuentran en relación laboral de dependencia o con negocios propios.
- Alrededor del 30% cuenta con pasaporte disponible para viajar al extranjero.
- >50% de los clientes posee auto propio.
- El tipo de paquete de viaje que prefieren los clientes corresponde al Deluxe y Básico. Así como la estadía en un hotel 3 estrellas.
- El ingreso mensual promedio del sueldo de los clientes supera los 20mil \$
- Más de 2mil clientes ya han adquirido de 1 a 3 productos/servicios en la empresa y han estado satisfechos con los mismos.
- En promedio los clientes consultan por un paquete para 3 personas con al menos 1 niño incluido en el viaje. Siendo Europa el continente predilecto para elegir viajar.

16. RECOMENDACIONES

Para disminuir los gastos de marketing de la empresa, se puede reducir la población target a partir de las siguientes condiciones:

- ➔ Grupo etario de 30 a 50 años.
- ➔ Debido a que solo un 30% cuenta con pasaporte para poder viajar al extranjero, deberán apuntar a paquetes nacionales + paquetes a Europa para aquellos que sí tengan la posibilidad.
- ➔ Incluir acuerdos donde en el paquete incluye el alquiler de autos.
- ➔ Paquetes llenos de confort pero económicos ya que pueden apuntar a hoteles 3 estrellas.

17. REFERENCIAS

<https://www.kaggle.com/datasets/susant4learning/holiday-package-purchase-prediction>

Fecha de consulta: 30/05/2023