

ENTREGA FINAL DEL PROYECTO MAVEN ANALYTICS

Alumnos:

AGUIRRE, Matías Nicolás. DÍAZ, Micaela. DZYSIUK, Matías Lucas.

Fecha: 19/12/2021



Tabla de contenido

TABLA DE VERSIONADO	2
INTRODUCCIÓN	3
OBJETIVOS	3
HERRAMIENTAS TECNOLÓGICAS UTILIZADAS	3
TEMÁTICA SELECCIONADA	4
DIAGRAMA ENTIDAD-RELACIÓN	4
LISTADO DE TABLAS.	5
MOCKUPS	7
TRANSFORMACIÓN DE DATOS	8
COLUMNAS Y MEDIDAS CALCULADAS	9
ANÁLISIS FUNCIONAL DEL TABLERO	14
CLUSTERIZACIÓN DE CONSUMIDORES	18
CONCLUSIONES	19
LINEAMIENTOS FUTUROS	19
ANEXO 1 – DIAGRAMA ENTIDAD-RELACIÓN	20
ANEXO 2 – DIAGRAMA ENTIDAD-RELACIÓN FINAL	21
ANEXO 3 – MODELO RELACIONAL	22

TABLA DE VERSIONADO

Nombre de la versión	Fecha	Desafío	Descripción
1	01/11/2021	13	Creación del modelo relacional
2	03/11/2021	15	Dashboard con las primeras medidas calculadas
3	08/11/2021	16	Dashboard con las primeras segmentaciones
4	16/11/2021	17	Dashboard con medidas calculadas y solapas definitivas
5	24/11/2021	22	Dashboard con modificaciones para la presentación
6	19/12/2021	-	Dashboard con recomendaciones de la presentación final

19/12/2021

INTRODUCCIÓN

En el presente documento se detallan distintos elementos y procedimientos para el abordaje de la temática seleccionada como proyecto final a analizar, se indican la temática analizada, el modelo E-R las tablas confeccionadas con su tipo de dato, el modelo relacional realizado con la herramienta Power Bl[®], las transformaciones realizadas junto a las medidas y columnas calculadas establecidas.

OBJETIVOS

Generales

- Aprender conceptos generales de las bases de datos para luego trabajar sobre bases de datos relacionales.
- Diseñar y desarrollar proyecto de data analytics con la herramienta Power BI utilizando DAX, M y Power Query.

Específicos

- Conocer modelo de negocios de la empresa Maven Marketing e identificar información requerida por el usuario para facilitar el proceso de toma de decisiones.
- Generar un modelo relacional y transformar los datos para que puedan ser usados por la herramienta de visualización.
- Diseñar KPIs que permitan evaluar de manera rápida e intuitiva los resultados de las campañas de marketing.
- Implementar técnicas de storytelling para la presentación de los datos trabajados y validar la información con el usuario final del tablero.

HERRAMIENTAS TECNOLÓGICAS UTILIZADAS

Las herramientas tecnológicas utilizadas durante el desarrollo del trabaja práctico integrador son:

Google Sheet. Utilizado como fuente de datos en la web, la cuenta del equipo contiene los datos del cual se sirve el dashboard.

Power BI. Servicio de análisis de datos de Microsoft, utilizado para la transformación y limpieza de los datos como así también el modelado de los mismos.

Word. Editor de texto de Microsoft utilizado para la confección del informe con la documentación requerida.

19/12/2021



Miro. Plataforma de trabajo colaborativo utilizada para la confección del modelo E-R.

Meet. Servicio de videollamadas de Google, utilizado para las reuniones grupales.

TEMÁTICA SELECCIONADA

Campaña de marketing - Maven Analytics

El objetivo será analizar los resultados de una campaña de marketing digital de 2240 clientes de Maven Marketing realizada durante el período de 2012 a 2014. Los resultados incluyen perfiles de clientes, preferencias de productos, éxitos/fracasos de campaña y rendimientos del canal. Se pretende analizar qué factores están relacionados con el perfil del cliente medio, identificar cuál es la campaña más exitosa, cuáles son los canales de menor rendimiento, cuales factores se relacionan con el número de compras web y cuales productos tienen mayor preferencia.

DIAGRAMA ENTIDAD-RELACIÓN

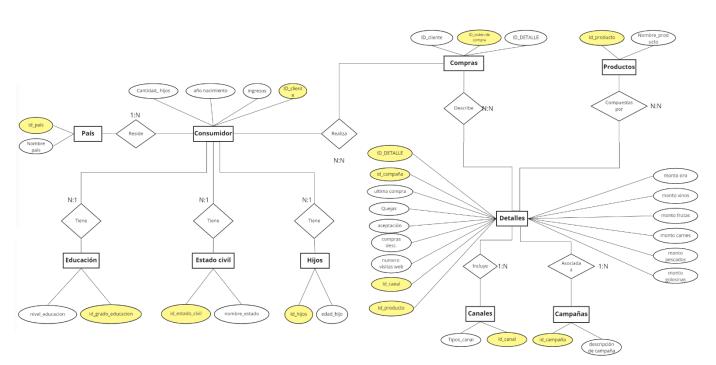


Figura 1.Diagrama EER de los resultados de la campaña de marketing. Fuente: elaboración propia. VER ANEXO 1

El resultado final del modelo E-R puede consultarse en los anexos.

LISTADO DE TABLAS

A continuación, se listan las tablas creadas a partir del diagrama entidad relación para los resultados de la campaña de marketing.

Tabla 1. Tabla de consumidores. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	ID_cliente	DECIMAL
	ingresos	DECIMAL
	año nacimiento	INT
	cantidad_hijos	INT

Tabla 2. Tabla de detalles Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_detalle	DECIMAL
Foreign key	id_campaña	DECIMAL
	ultima compra	BIT
	quejas	DECIMAL
	aceptación	INT
	compras desc.	BIT
	numero visitas web	INT
Foreign key	id_canal	DECIMAL
Foreign key	id_producto	DECIMAL
	monto oro	DECIMAL
	monto vino	DECIMAL
	monto frutas	DECIMAL
	monto carnes	DECIMAL
	monto pescados	DECIMAL
	monto golosinas	DECIMAL

Tabla 3. Tabla de compras. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	ID_orden de compra	DECIMAL
Foreign key	ID_DETALLE	DECIMAL
Foreign key	ID_cliente	DECIMAL

Tabla 4. Tabla de países de los consumidores. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_pais	DECIMAL
	Nombre pais	TEXT

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_grado_educacion	DECIMAL
	nivel_educacion	VARCHAR(100)

Tabla 6. Tabla del estado civil de los consumidores. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_estado_civil	DECIMAL
	nombre estado	TEXT

Tabla 7. Tabla que indica la cantidad de hijos en casa de los consumidores. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_hijos	DECIMAL
	edad_hijo	INT

Tabla 8. Tabla de los productos consumidos. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_producto	DECIMAL
	nombre producto	TEXT

Tabla 9. Tabla de los canales utilizados. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_canal	DECIMAL
	tipos_canal	TEXT

Tabla 10. Tabla de las campañas realizadas. Fuente: elaboración propia.

Tipo de clave	Nombre de campo	Tipo de dato
Primary key	id_campaña	DECIMAL
	descripción de campaña	VARCHAR(100)

MOCKUPS

La maqueta del dashboard es un prototipo del tablero final a crear, a continuación, se presentan las solapas diseñadas (sin carátula).

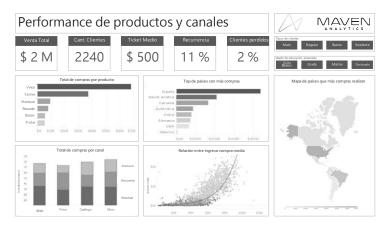


Figura 2. Solapa 1 con datos de las compras. Fuente: Elaboración propia.



Figura 3. Solapa 2 con datos de los clientes. Fuente: Elaboración propia.

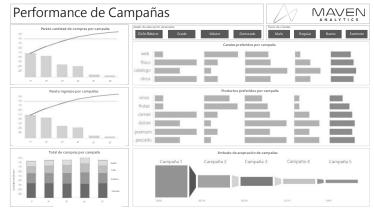


Figura 4. Solapa 3 con datos de las campañas. Fuente: Elaboración propia.

19/12/2021

TRANSFORMACIÓN DE DATOS

En esta sección, se detallan los cambios efectuados al construir el modelo relacional de la base de datos.

Tabla producto

Se cargaron las tablas consumidor, compras, productos, campañas y diccionario.

La tabla que contaba con todos los datos en conjunto fue eliminada.

Las primeras dos filas de cada tabla fueron eliminadas, excepto en diccionario en la cual eliminamos 3. Las primeras filas fueron utilizadas como encabezado.

Se modificó el nombre de la tabla *campañas* a *Aceptación_Campaña* y anulamos la dinamización de las diferentes campañas, quedando en una sola columna llamada *id_campania*. Se renombraron las campañas, numerándolas del 1 al 6.

Se duplicó la tabla antes mencionada, llamando al duplicado *Campañas_info, también se* eliminaron todas las columnas salvo *id_campania*. Se quitaron los datos duplicados y se agregó una columna a partir de ejemplos llamada Nombre_campania en la cual se detalla el nombre de cada una de ellas.

La misma mecánica se replicó para crear la tabla *Aceptación_info*, *Productos_info* y *Canal_info*, *Pais_info*, *Educación_info* y *Marital_info*.

También se creó la tabla Calendario para poder, a futuro, en caso que se requiera aplicar funciones de inteligencia de tiempo.

Por último, realizamos las conexiones de las tablas, donde todas las relaciones de las mismas se dan de 1 a muchos.

Anexamos una captura del modelo. VER ANEXO 2

19/12/2021

COLUMNAS Y MEDIDAS CALCULADAS

A continuación, se detallan los procedimientos efectuados a la base de datos de los resultados de la campaña de marketing para la obtención de medidas y columnas calculadas con la herramienta Power BI[®]

Medidas calculadas de Solapa 1 – Productos y Canales

Se crearon medidas calculadas que sirven para mostrar tarjetas informativas que se adapten a los filtros que serán aplicados a la presentación. Las medidas creadas fueron las siguientes:

1) Cantidad total de clientes

Esta medida tiene como objetivo informar de manera rápida la cantidad de clientes alcanzados durante las 6 campañas de marketing. Se calcula como el recuento total de id de consumidor.

```
cant_total_clientes = COUNT(Consumidor[ID])
```

2) Cantidad compras total

Esta medida se utiliza para informar la cantidad de compras totales realizada, es importante destacar que su utilización alcanza todo su potencial cuando se la acompaña de filtros como por canal de compras o educación del consumidor. Se calcula como la suma total de compras de todos los canales.

```
cant_compras_total = SUM(Compras[Compras_por_canal])
```

3) Sueldo promedio del consumidor

Con el objetivo de conocer la potencialidad de compra de los consumidores e informarlo de manera rápida, se realiza el cálculo promedio de los sueldos de los consumidores.

```
sueldos_promedio = AVERAGE(Consumidor[Income])
```

4) Ticket promedio

Esta medida tiene el objetivo de informar cuantas unidades monetarias se gastan en promedio por cada compra, para poder dar cuenta del tamaño del ticket de compra.

```
ticket_promedio =
VAR productos_vendidos =
    SUM ( Productos[Cant_producto] )
VAR compras_totales = [cant_compras_total]
RETURN
    DIVIDE ( productos vendidos, compras totales )
```

Es un indicador clave ya que siempre será más fácil incrementar el importe medio de cada pedido antes que captar nuevos clientes, además, las acciones que puedes realizar para incrementar tu ticket medio suelen ser mucho más baratas que acciones de captación como Cross-selling y Upselling.

5) Porcentaje de conversión

Esta medida tiene como objetivo informar sobre el porcentaje de visitantes de la página web realizan al menos una compra. Es un indicador clave ya que se relaciona con la experiencia del cliente en el proceso de compra. Representa una métrica clave ya que de su valor se desprende la decisión de lanzar o no campañas de captación, el valor mínimo recomendado es del 3 al 5%.

```
%conversion =
VAR ComprasMes =
    CALCULATE ( COUNT ( Compras[ID] ), Compras[Recency] < 30 )
VAR NumeroVisitasMes =
    SUM ( Compras[NumWebVisitsMonth] )
VAR NuevasCompras = ComprasAdd[Valor ComprasAdd]
RETURN
    ( ComprasMes + ComprasAdd[Valor ComprasAdd] ) / NumeroVisitasMes )</pre>
```

A partir de la adición del parámetro **NuevasCompras** se puede visualizar de forma rápida y precisa cuantas nuevas compras se necesitan para aumentar este indicador y así desencadenar otro tipo de decisiones a futuro.

6) Acumulado total de productos

Esta medida tiene como objetivo formar la línea del eje secundario del Pareto que indica el monto acumulado, lo que permite mostrar cuales 20% de productos representan el 80% del monto recaudado.

```
)
RETURN
DIVIDE ( acumulado, total )
```

7) Variación anual de subscriptores

Esta medida tiene como objetivo informar sobre la variación de subscriptores respecto del año pasado

Medidas calculadas de Solapa 2 - Clientes

Se crearon medidas calculadas con el objetivo de mostrar información relevante para la caracterización de los tipos de clientes.

1) Edad promedio de los consumidores

Esta medida tiene como objetivo mostrar la edad de los clientes, dependiendo del tipo de venta que cierre o el tipo de cliente.

```
edad_promedio = AVERAGE(Consumidor[Edad])
```

2) Nuevos clientes mensuales

Esta medida informa cuantos clientes nuevos se incorporaron a la compañía en los últimos 30 días

```
Nuevos Clientes =
CALCULATE(COUNT(Consumidor[Dt_Customer]), Consumidor[Antiguedad] < 30 )</pre>
```

3) Nuevos clientes mensuales

Esta medida informa cuantos clientes nuevos se incorporaron a la compañía en los últimos 30 días.



Medidas calculadas de Solapa 3 - Campañas

Se crearon medidas calculadas con el objetivo de mostrar información en las gráficas de Pareto

1) % total acumulado (cantidad)

Esta medida tiene como objetivo mostrar cuales 20% de los tipos de clientes abarcan el 80 % de la cantidad de compras. Para calcular esta medida, previamente se debe calcular el ranking de los tipos de clientes con mayor cantidad de ventas ordenado de mayor a menor a través de la siguiente medida:

```
Ranking clusteres =
RANKX (
    ALLSELECTED ( Consumidor[Tipos de consumidores] ),
    [cant_compras_total],
    DESC
)
```

Luego, se procede a calcular la medida que acumula los valores de ese ranking para obtener la línea acumulada del gráfico secundario.

```
% Total acumulado (cantidad) =
VAR acumulado =
    CALCULATE (
        [cant_compras_total],
        TOPN (
            [Ranking clusteres],
            ALLSELECTED ( Consumidor[Tipos de consumidores] ),
            [cant_compras_total], DESC
        )
    )
VAR total =
    CALCULATE (
        [cant_compras_total],
        ALLSELECTED ( Consumidor[Tipos de consumidores] )
    )
RETURN
    DIVIDE ( acumulado, total )
```

De la misma forma se realizaron los cálculos para el 80-20 de montos respectos de tipos de clientes.

2) Variación anual de subscripciones

Muestra la variación respecto del mismo período de año pasado, sirve para hacer una comparación más objetiva y evaluar de forma rápida si aumentan los nuevos clieentes.

```
% Var. subs =
VAR SubsAnuales = COUNT(Consumidor[Dt_Customer])
```

```
>>
```

3) % de aceptación de campañas

Esta medida permite al gráfico tipo funnel presentar el porcentaje de aceptación e las campañas de marketing, calcula el porcentaje de clientes que aceptaron determinada campaña y se calcula como el cociente entre aceptados sobre el total de clientes de cada campaña.

```
% Aceptación =
VAR acep_campania =
    SUM ( 'Aceptacion_Campaña'[Acep_campania] )
VAR total_aceptacion =
    CALCULATE (
        SUM ( 'Aceptacion_Campaña'[Acep_campania] ),
        ALLSELECTED ( 'Aceptacion_Campaña' )
    )
RETURN
    DIVIDE ( acep_campania, total_aceptacion )
```

Columnas calculadas

Se crearon columnas calculadas para evaluar expresiones a todos los registros por igual con el objetivo de enriquecer el análisis. Las columnas creadas las dos siguientes:

1) Grupo Etario

Tiene como objetivo identificar el grupo etario al cual pertenece el cliente y convertir las variables numéricas a categóricas. A continuación, se presenta el código necesario para calcular la columna.

2)Buena venta

Esta columna tiene como objetivo identificar si una compra es buena (cuando el valor de compra es mayor o igual al ticket promedio) o regular (cuando el valor de compra es menor al ticket promedio.

Buena_venta = IF(Productos[Cant_producto]>=[ticket_promedio], "Buena", "Regular")

ANÁLISIS FUNCIONAL DEL TABLERO

Público objetivo – Nivel de aplicación

El nivel de aplicación del tablero diseñado, se corresponde con el nivel táctico, ya que permite a los directivos de rango medio, evaluar planes de acción concretos del departamento de marketing en un horizonte temporal de corto plazo.

Generales

El día 8/11/2021 se efectuaron los siguientes cambios en diferentes solapas de nuestro proyecto a través del programa Power BI:

Se creó la pestaña portada, donde se colocó el título, subtítulo y el logo. También se crearon otras tres pestañas: "Productos y canales", "Clientes" y "Campañas".

Pestaña de productos y canales

Esta solapa permite identificar de manera general, la situación actual de las campañas y el relacionamiento de las ventas de los diferentes productos de acuerdo con el país, canal de venta, ingresos y educación.

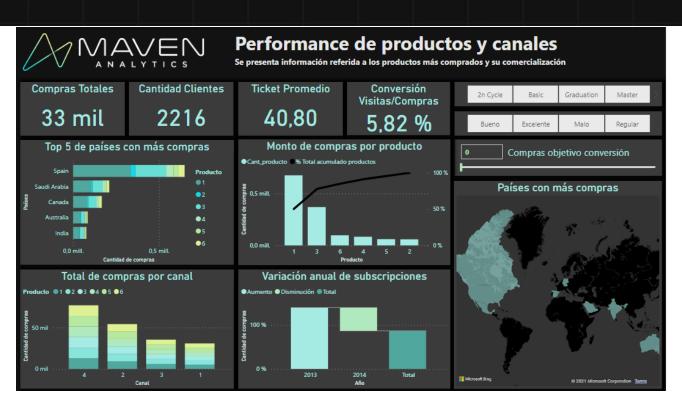


Figura 5. Solapa de productos y canales. Fuente: elaboración propia.

En la parte superior se ubican los indicadores claves para el entendimiento del desempeño de las campañas, de izquierda a derecha puede nombrarse, la cantidad de compras realizadas, la cantidad de clientes alcanzados, el ticket promedio de compra de cada cliente y el porcentaje de clientes que visita la página y termina comprando un producto.

Abajo se crea una gráfica de barras apiladas horizontales en base a los tipos de productos y su cantidad de venta, la cual puede asociarse a la gráfica de productos por país donde nos permite conocer en qué países se consume mayormente cada producto.

Siguiendo el patrón Z, en la zona superior del medio se observa un gráfico de Pareto 80-20 dónde puede observarse el 20% productos que representan el 80% del monto recibido.

El mapa coroplético ayuda a visualizar de una forma más sencilla y dinámica lo planteado anteriormente sobre los países. También, en la parte inferior izquierda podemos observar gracias a las gráficas apiladas verticales, los canales mayormente elegidos por los clientes al contar con una gráfica de compras por canal.

Siguiendo el patrón Z en la parte media inferior puede observarse la variación anual de suscriptores a los, evidenciando un decremento en 2014 respecto del 2013.

Por último, los filtros permiten segmentar la información descrita en base al grado de educación de los clientes y en base a su característica de compra. También el parámetro permite calcular cuando aumenta la conversión con la incorporación de nuevos clientes, esto es de suma importancia ya que el paso previo para lanzar una campaña de captación de clientes es asegurarse de poseer un alto porcentaje de conversión (5% - 8%).

Pestaña de clientes

Esta solapa permite conocer de manera más especifica que la solapa anterior a los consumidores ya que en las diferentes gráficas y filtros observamos todas las variables que los identifican.

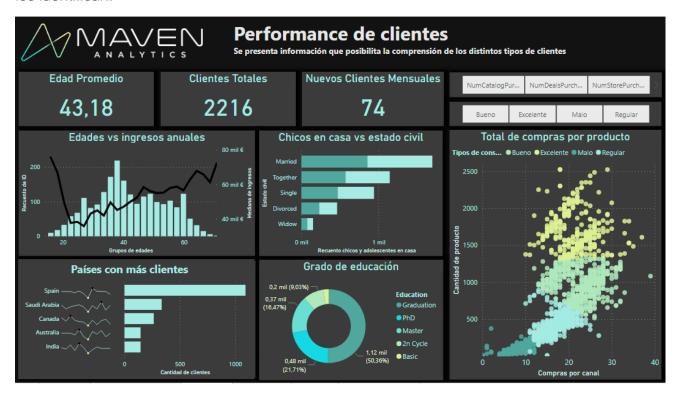


Figura 6. Solapa de clientes. Fuente: elaboración propia.

En la parte superior se muestran tarjetas con información sobre los clientes totales alcanzados, la edad promedio de los mismos y los nuevos clientes mensuales.

En primer lugar, puede visualizarse un histograma de edades vs la mediana de los ingresos, esto es muy importante ya que permite evidenciar a potenciales clientes con un poder de compra específico sin que sea afectado por valores extremos ya que se trabaja con la mediana.

En la zona superior central la gráfica de recuento de niños y adolescentes en casa de acuerdo con su estado civil permite observar las tendencias de acuerdo con cómo se forma el hogar de cada cliente.

A la derecha, se presenta una gráfica de dispersión que permite conocer el comportamiento de los clientes buenos, malos, regulares y excelentes (clasificación que se explica en el apartado "Clusterización de clientes")

Siguiendo el patrón Z, en la parte inferir izquierda puede observarse un top 5 de los países del cual provienen la mayoría de clientes, esto junto al historial de subscripciones que tiene como finalidad enriquecer la información del top.



En la parte inferior central la gráfica de anillo permite conocer la proporción de los clientes con determinada formación educativa. Esta información puede segmentarse por tipo de venta realizada y por cada topología de cliente, permitiendo caracterizarlos de manera integral.

Pestaña de campañas

En esta última solapa se visualiza el desempeño de cada una de las campañas de acuerdo con distintas variables.

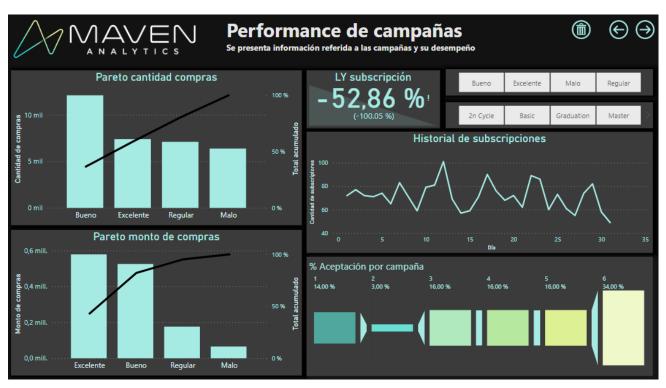


Figura 7. Solapa de campañas Fuente: elaboración propia.

Se puede observar en la parte izquierda un par de gráficas de Pareto que complementan su información sobre las ventas y montos en base al tipo de cliente. Luego un indicador del aumento de las suscripciones respecto del mismo periodo del año anterior.

También se muestra un historial de subscripciones a lo largo del tiempo que permiten evaluar cuantos nuevos clientes se suman en distintas unidades de tiempo (día, mes, trimestre y año)

En la parte inferior izquierda puede visualizarse el porcentaje de aceptación de las campañas a través de la gráfica funnel descargada del Marketplace que muestra de manera intuitiva cuales fueros las campañas más exitosas y cuales no.

A su vez en la pestaña podemos filtrar tanto por grado de educación del cliente como por tipo de cliente. De esta forma ver en que influyen estas variables con los resultados de las mismas.

CLUSTERIZACIÓN DE CONSUMIDORES

El día 24/11/2021 se realizó un clúster a partir de datos de cantidad de productos comprados, monto de los productos, ingresos de los clientes y edad. Esto permite direccionar campañas y productos al agrupar a los clientes en clientes malos, regulares, buenos y excelentes.

Malo: Cliente que compra poco y por un bajo monto.

Regular: Cliente que compra una cantidad moderada y monto moderado.

Bueno: Cliente que compra una mayor cantidad de productos, pero su diferencia principal es el monto de compra.

Excelente: Cliente que compra una amplia cantidad de productos con los montos más altos.

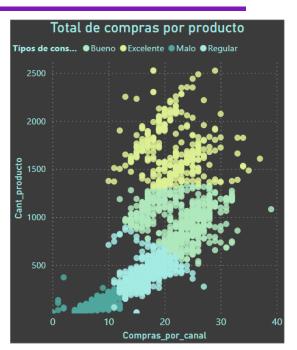


Figura 8. Gráfico de dispersión de los tipos de clientes. Fuente: elaboración propia.

CONCLUSIONES

Se logró abordar un proyecto de data analytics de forma integral, permitiendo conocer el modelo de negocios de la empresa Maven Marketing a través de un tablero que muestra información de lo más general a lo más específico.

De la solapa "Productos y Canales" se observa que el 80% del monto facturado se debe al producto 1 (vino) y al producto 3 (carnes). El mercado que más consume estos productos es el español y lo realiza a través de tiendas físicas.

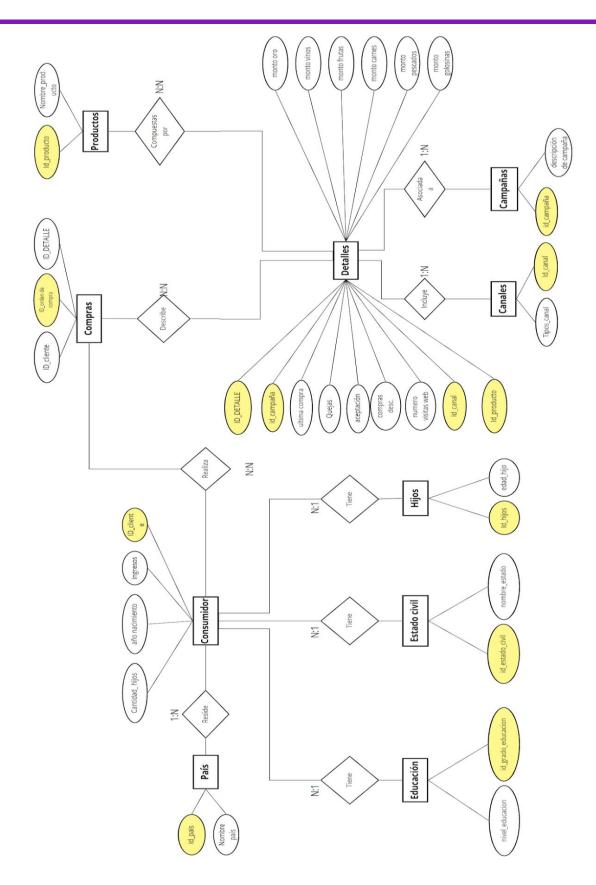
De la solapa "Clientes" y "Campañas" se observa que los clientes buenos y excelentes representa el 82,06 % del monto facturado, estos clientes compran los entre 10 y 30 productos con los montos más elevados, son graduados y la mayor cantidad de ellos tienen entre 35 y 50 años.

Se observa que los clientes regulares y malos compran muchos productos pero no son de valor representativo, esto debe tenerse en cuenta para reducir costos, ya que estas compras son las que contribuyen en mayor medida en los costos de comercialización.

LINEAMIENTOS FUTUROS

Se recomienda avanzar en análisis predictivo de cantidad de subscriptores y tipo de los mismos para poder planificar con antelación estrategias que permitan a la compañía satisfacer las necesidades de los clientes.

ANEXO 1 - DIAGRAMA ENTIDAD-RELACIÓN



ANEXO 3 - MODELO RELACIONAL

