

Trabajo Práctico 1

[75.06/95.58] Organización de Datos
Segundo cuatrimestre de 2020
The Big Four

Mariotti, Franco	102223	fmariotti@fi.uba.ar
Fabbiano, Fernando	102464	ffabbiano@fi.uba.ar
Alasino, Franco Federico	102165	falasino@fi.uba.ar
Colombo, Juan Ignacio	103471	jcolombo@fi.uba.ar

Índice

1. Introducción	3
2. Análisis Exploratorio Inicial	4
2.1. Detección de nulos	5
2.1.1. Tratamiento de nulos de ASP y ASP_(converted)	6
2.1.2. Tratamiento de nulos de Total_Amount	6
2.2. Detección de valores inútiles	6
3. Análisis sobre los vendedores	8
3.1. Top 10 vendedores con más ingresos generados	8
3.2. Top 10 vendedores más eficaces	8
3.3. Top 10 vendedores menos eficaces	9
3.4. Conclusiones	10
4. Análisis sobre la Clientela	11
4.1. Introducción	11
4.2. Más oportunidades creadas, ¿clientes más importantes?	11
4.3. Conclusión	13
5. Restricciones burocráticas que influyen en las ventas	14
5.1. Introducción	14
5.2. Analizando país por país	14
5.3. Conclusiones	16
6. Performance de la empresa a lo largo del tiempo	17
6.1. Introducción	17
6.2. Ver el pasado, para poder predecir el futuro	17
6.3. Meses más calurosos, ¿implican más ventas?	18
7. Análisis sobre Regiones	20
7.1. Introducción	20
7.2. Porcentaje de oportunidades según región	20
7.3. Oportunidades perdidas por región	21
7.4. Relación Oportunidades ganadas y totales	22
7.5. Conclusión	23
8. Análisis sobre Productos	24
8.1. Introducción	24
8.2. Familias de Producto mas Vendidas por Año	24
8.3. Distribución de las Ventas según Región	24
8.4. Productos mas vendidos por Región	26
8.5. Los 10 productos con mayor recaudación	27
8.6. Los 10 productos más y menos vendidos.	27
9. Análisis sobre TRF	29
9.1. Análisis de ventas según TRF	29
10. Análisis sobre rango entrega	30
10.1. Rango de entrega y porcentaje de éxito	30
10.2. Rango de entrega según Regiones	31

11. Análisis sobre requerimiento de aprobación especial	32
11.1. Stages y sus aprobaciones especial	32
11.2. Porcentaje de oportunidades según región	32
11.3. Países con mayor requerimiento de aprobación especial	33
12. Conclusiones generales	35
13. Referencias	36

1. Introducción

El objetivo del siguiente informe es hacer un relevamiento a la empresa "Frio Frio" acerca del análisis del DataSet otorgado por la misma. El mismo consiste en una exploración e interpretación de las distintas oportunidades de venta que la empresa tuvo en el periodo que abarca desde los años 2013 a 2018.

La primer parte del trabajo consistió en el análisis de aquellos registros que contenían valores nulos, o que son indefinidos (léase Nat, None, Other, etc). Esto nos permitió decidir cuáles de ellos descartar y cuáles no, haciendo una fina interpretación de cada una de las variables en cuestión.

Para la segunda parte ya nos encontramos con un DataFrame reducido y con registros válidos en su totalidad. A partir de allí, realizamos un profundo análisis de la relación entre las variables que se encontraban en el data set. Al no poseer un contacto directo con la empresa para consultar dudas, en algunos casos resultó necesario hacer suposiciones para continuar con los análisis propuestos -las mismas se encuentran aclaradas a lo largo del informe para una correcta comprensión del lector-.

Para la tercera y última parte, se procedió a colocar los análisis realizados en visualizaciones que permitiesen interpretar con mayor claridad las relaciones encontradas.

2. Análisis Exploratorio Inicial

Con este análisis inicial, buscamos observar detenidamente los datos disponibles y hacer un pre-procesamiento de los mismos. Aquí, explicaremos los motivos y como se fue desarrollando dicho pre-procesamiento para llegar al Dataframe con el cual trabajamos.

Inicialmente, contabamos con un set de datos que representaba las oportunidades de venta de la empresa. Cada elemento del set representaba un ítem o producto del cual además se incluía la siguiente información:

- ID: id único del registro.
- Región: región de la oportunidad.
- Territory: territorio comercial de la oportunidad.
- Pricing , Delivery _Terms _Quote _Approval: variable que denomina si la oportunidad necesita aprobación especial de su precio total y los términos de la entrega.
- Pricing , Delivery _Terms _Approved: variable que denomina si la oportunidad obtuvo aprobación especial de su precio total y los términos de la entrega.
- Bureaucratic _Code _0 _Approval: variable que denomina si la oportunidad necesita el código burocrático 0.
- Bureaucratic _Code _0 _Approved: variable que denomina si la oportunidad obtuvo el código burocrático 0.
- Submitted _for _Approval: variable que denomina si fue entregada la oportunidad para la aprobación.
- Bureaucratic _Code: códigos burocráticos que obtuvo la oportunidad.
- Account _Created _Date: fecha de creación de la cuenta del cliente.
- Source : fuente de creación de la oportunidad.
- Billing _Country: país donde se emite la factura.
- Account _Name : nombre de la cuenta del cliente.
- Opportunity _Name : nombre de la oportunidad.
- Opportunity _ID : id de la oportunidad.
- Sales _Contract _No : número de contrato.
- Account _Owner : vendedor del equipo comercial responsable de la cuenta cliente.
- Opportunity _Owner : vendedor del equipo comercial responsable de la oportunidad comercial.
- Account _Type : tipo de cuenta cliente.
- Opportunity _Type : tipo de oportunidad.
- Quote _Type : tipo de presupuesto.
- Delivery _Terms : términos de entrega.
- Opportunity _Created _Date : fecha de creación de la oportunidad comercial.
- Brand : marca del producto.
- Product _Type : tipo de producto.

- Size : tamaño del producto.
- Product_Category_B : categoría 'B' del producto.
- Price : precio.
- Currency : moneda.
- Last_Activity : fecha de la última actividad.
- Quote_Expiry_Date : fecha de vencimiento del presupuesto.
- Last_Modified_Date : fecha de ultima modificación en la oportunidad.
- Last_Modified_By : usuario responsable de la última modificación en la oportunidad.
- Product_Family : familia de producto.
- Product_Name : nombre del producto.
- ASP_Currency : moneda del precio promedio.
- ASP : (Average Selling Price) precio promedio a la venta.
- ASP_(converted)_Currency : moneda del precio promedio convertido en la variable.
- ASP_(converted) : precio promedio a la venta convertido a otra moneda.
- Planned_Delivery_Start_Date : límite inferior del rango previsto para la fecha de entrega.
- Planned_Delivery_End_Date : límite superior del rango previsto para la fecha de entrega.
- Month : mes-año de Planned_Delivery_Start_Date.
- Delivery_Quarter : trimestre de Planned_Delivery_Start_Date.
- Delivery_Year : año de Planned_Delivery_Start_Date.
- Actual_Delivery_Date : fecha real de la entrega.
- Total_Amount_Currency : moneda del monto total.
- Total_Amount : monto total.
- Total_Taxable_Amount_Currency : moneda del monto gravado total.
- Total_Taxable_Amount : monto gravado total.
- Stage : variable target. Estado de la oportunidad.
- Prod_Category_A : categoría 'A' del producto.
- TRF : Toneladas de refrigeración. Es una unidad de potencia.

2.1. Detección de nulos

Lo primero que hicimos fue detectar aquellas columnas de tuvieran valores nulos para luego decidir como estos iban a ser tratados en cada caso.

Luego de analizar el set de datos, pudimos ver que teníamos 3 columnas que contenían valores nulos: ASP, ASP_(converted) y Total_Amount.

2.1.1. Tratamiento de nulos de ASP y ASP_(converted)

Vemos que para varios productos hay valores de ASP y ASP_(converted) nulos. Dado que ese valor es importante para poder realizar la conversión a USD porque establece una relación entre USD y una determinada moneda, decidimos reemplazar estos valores nulos por el promedio de los valores de ASP y ASP_(converted) respectivamente, correspondientes a cada moneda.

Ejemplo:

- $ASP_JPY_nulo = ASP_JPY_mean$
- $ASP_converted_JPY_nulo = ASP_converted_JPY_mean$

Ahora que tenemos un valor asignado de ASP y ASP_Converted para todos los items, podemos realizar la conversión para todos ellos. Además, dado que son valores que utilizamos recurrentemente en varios análisis, decidimos agregar dos columnas nuevas con los valores de Total_Amount y Total_Taxable_Amount en USD.

2.1.2. Tratamiento de nulos de Total_Amount

En esta columna también teníamos elementos nulos pero en una cantidad menor, específicamente 59 valores nulos. En este caso, decidimos que lo más indicado era dejar estos valores en el Dataframe y tratarlos de manera particular para cada análisis que quisiéramos, dado que eran pocos casos y que no creíamos correcto eliminar por completo estas filas que nos podían brindar información a pesar de tener un valor de Total_Amount nulo.

2.2. Detección de valores inútiles

Más allá de detectar valores nulos en nuestro set de datos, también identificamos varias columnas que tenían valores válidos pero que a priori no aportaban información útil, como por ejemplo: 'Nat', 'Other' o 'None'. Por lo que decidimos analizar más en detalle dichas columnas para decidir que íbamos a hacer con cada una de ellas.

- Columnas que tienen 100% de datos inútiles, es decir, que son 'None', 'Other' o 'NaT'.
 - Como no aportan ningún tipo de información decidimos descartarlas, dichas columnas son las siguientes:
 1. Last_Activity
 2. Actual_Delivery_Date
 3. Prod_Category_A
- Columnas que tienen más de un 90% de datos inútiles
 - Vemos que la información que brindan estas columnas no es representativa, es decir, cualquier estadística que realicemos sobre estos datos solo representaría una parte muy pequeña del total, dando lugar a análisis poco fundamentados y por lo tanto decidimos descartarlas. Estas columnas son:
 1. Price
 2. Currency (columna asociada a Price)
 3. Brand
 4. Size
 5. Product_Category_B
 6. Product_Type

- Source

- Esta columna tiene un 56 % de datos de son 'None'. En este caso creemos que es un caso similiar a los mencionados anteriormente, tenemos más de la mitad del total de los datos que no nos aportan información. Por esta razón, concluimos que realizar alguna estadística sobre estos seria poco representativa y hasta quizás imprudente ya que podríamos llegar a resultados equívocos.

- Sales_Contract_No

- Esta columna tiene un 56 % de datos de son 'None' o 'Other'. Sin embargo, luego de revisar un poco los datos, nos dimos cuenta que la gran mayoría de estos 'None' o 'Other' aparecen en oportunidades que fueron cerradas sin éxito. Esto tiene sentido, ya que es lógico que si hablamos de un contrato de ventas, este solo exista o esté registrado para aquellas oportunidades que efectivamente se concretaron. Es por esta razón que decidimos considerar la información brindada por esta columna para análisis posteriores.

- TRF

- En esta columna tenemos alrededor de un 70 % de valores que son 0. Creemos que algunos de estos 0 podrían corresponder a la venta de productos relacionados a la instalación de algún equipo de aire acondicionado o a la venta de algún repuesto, por lo que tendría sentido que no tuvieran un valor de TRF asociado.

Por otro lado, teniendo en cuenta que un aire acondicionado promedio de uso cotidiano suele tener entre 0 y 1 TRF, creemos que estos 0 pueden ser producto de un truncamiento en los valores, es decir, resulta lógico que se tuviese una gran cantidad de productos con una cantidad de TRF entre 0 y 1 y al pasarse a un numero entero, se truncaran estos valores y cayeran al cero.

Por estas razones decidimos considerar esta columna para nuestros análisis , sabiendo que el 0 podría llegar a representar alguno de estos dos casos.

3. Análisis sobre los vendedores

En esta sección lo que buscamos es analizar el rendimiento de los vendedores de la empresa viendo cuales son los mejores y peores teniendo en cuenta distintos aspectos.

3.1. Top 10 vendedores con más ingresos generados

En este caso buscamos aquellos vendedores que generaron más ingresos para la empresa.



Figura 1: Top 10 vendedores con mas ingresos generados

Viendo este gráfico uno se vería tentado a decir que Person_Name_8 es sin dudas el mejor vendedor de la empresa ya que podemos ver una amplia diferencia respecto a sus compañeros en la cantidad de ingresos generados. Sin embargo, surgen varias preguntas que ponen en duda esta hipótesis: ¿en cuantas oportunidades generó esos ingresos?, ¿serán esos ingresos producto de una única venta exitosa?, ¿cuántas ventas cerradas sin éxito tiene en su haber?

3.2. Top 10 vendedores más eficaces

Por lo mencionado anteriormente, vemos que para calificar a un vendedor no solo es importante ver cuantos ingresos generó sino que también debemos analizar cuantas oportunidades le asignaron y cuál fue su desempeño respecto a las mismas.

Algo que tenemos que aclarar antes de pasar a ver los resultados obtenidos, es que para evitar tener resultados poco representativos relacionados con el problema planteado en "La ecuación mas peligrosa de la historia", decidimos considerar para este análisis solo aquellos vendedores que tuviesen asignadas como mínimo una cantidad de oportunidades mayor al promedio de oportunidades por vendedor.

Performance de los vendedores con mayor tasa de ventas cerradas exitosamente



Figura 2: Top 10 vendedores mas eficaces

Ahora vemos que la hipótesis mencionada en la sección anterior tiene más fundamentos, porque Person_Name_8 se encuentra dentro del top 10 de vendedores con mayor tasa de éxito, es decir, no fue que tuvo suerte y logró cerrar solo una venta grande sino que esa gran cantidad de ingresos generados fue fruto de una mayoría de ventas exitosas. Por otro lado, vemos que Person_Name_19 tiene una tasa muy alta de ventas cerradas de manera exitosa, algo que lo posiciona muy bien de cara a la disputa por ser el mejor vendedor de la empresa.

A partir de este punto creemos que la elección está en manos de la empresa, que podrá decidir si prevalece la gran cantidad de ingresos generados por Person_Name_8 o la altísima efectividad de Person_Name_19.

3.3. Top 10 vendedores menos eficaces

Así como vimos aquellos vendedores con mejor desempeño, nuestro análisis no estaría completo si no analizáramos a aquellos que tuvieron los peores rendimientos.

Igual que para el caso anterior, solo entraron en consideración aquellos vendedores que tuviesen asignadas una cantidad de oportunidades mayor al promedio de oportunidades por vendedor.

Performance de los vendedores con mayor tasa de ventas cerradas sin éxito



Figura 3: Top 10 vendedores menos eficaces

Estos son los vendedores a los que la empresa debe prestar mayor atención dado que sus performances ,en algunos casos, pueden llegar a ser preocupantes. Creemos que los despenos más inquietantes son aquellos en los que el porcentaje de Closed Lost es mayor al 50 %, dado que indicaría que el vendedor pierde más de la mitad de las oportunidades. La empresa debe tratar estos casos para evitar mayores pérdidas a futuro y aumentar la efectividad.

Algo que nos resultó interesante destacar es que nuevamente vemos al Person_Name_8. Teniendo en cuenta su aparición en este top de vendedores menos efectivos, la idea de considerarlo el mejor vendedor de la empresa vuelve a cuestionarse.

3.4. Conclusiones

Más allá de que la decisiones sean responsabilidad de la empresa, estos análisis le brindan información interesante y necesaria acerca de sus vendedores, para que luego ellos a partir de sus criterios tomen las decisiones adecuadas.

4. Análisis sobre la Clientela

4.1. Introducción

Nuestro objetivo en esta sección fue darle un foco importante a los 15 clientes más importantes en las siguientes categorías:

- Cantidad de oportunidades generadas
- Cantidad de oportunidades cerradas exitosamente
- Cantidad de oportunidades cerradas sin éxito
- Redituabilidad

Elegimos 15 porque consideramos este un número representativo, que permite a la empresa concentrarse en una cantidad reducida de clientes y decidir dónde poner el foco y destinar recursos.

4.2. Más oportunidades creadas, ¿clientes más importantes?

En la figura 4 mostramos el Top 15 de clientes con más oportunidades creadas, independientemente del estado actual de la misma.



Figura 4: Top 15 Clientes con mas Oportunidades

Si bien tener una gran cantidad de oportunidades puede parecer, a priori, que el cliente es de gran interés para la empresa, resulta interesante e importante ver que clientes son los que más oportunidades concretaron, y que clientes no. Una gran cantidad de ventas cerradas y ganadas, nos indica una cuenta que puede traernos ganancias. Por el contrario una con un monto grande de oportunidades perdidas, pone a la empresa en un estado de alerta acerca de si conviene seguir destinando recursos a ese cliente.

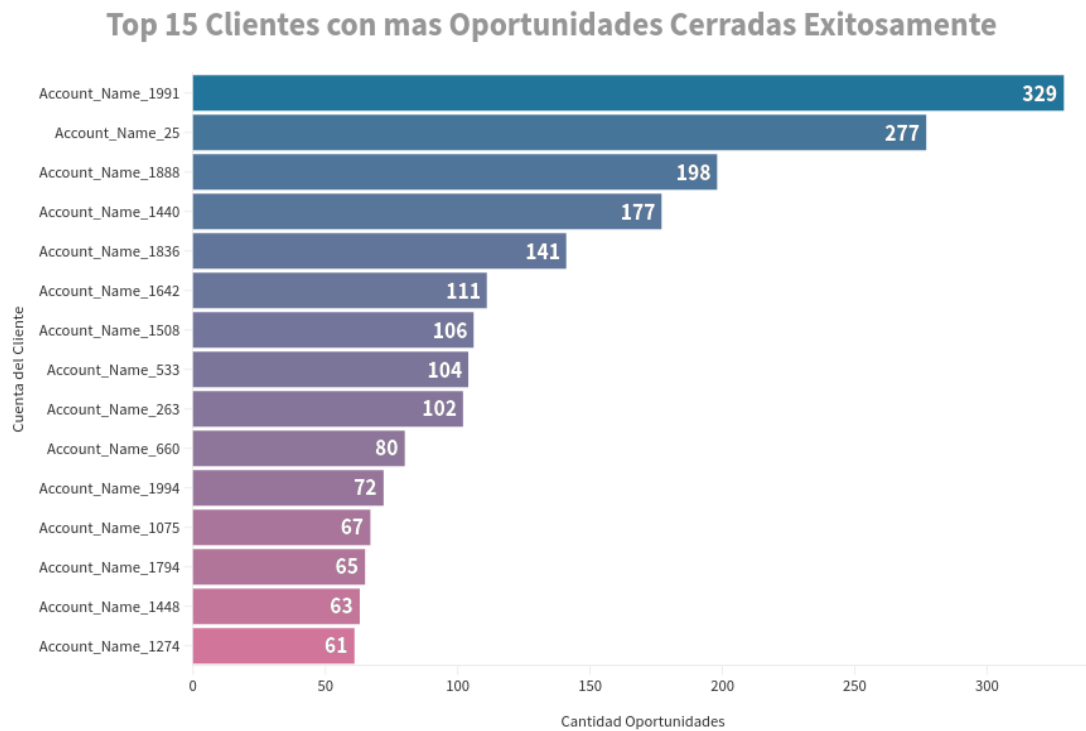


Figura 5: Top 15 Clientes con mas Oportunidades cerradas con Éxito

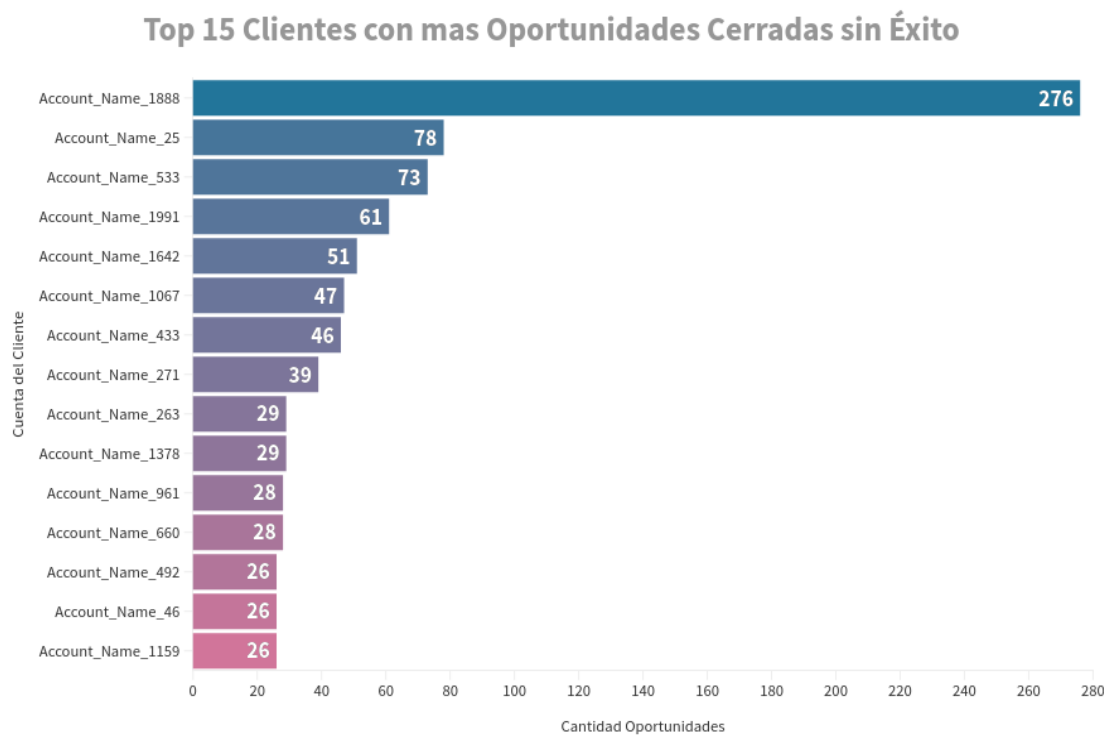


Figura 6: Top 15 Clientes con mas Oportunidades cerradas sin Éxito

De las figuras 5 y 6 podemos ver, por ejemplo, que si bien la cuenta 1888 parece ser a la que la empresa debería dedicarle más recursos ya que es la que mas oportunidades generó, en realidad es a la que debería mirar con más cuidado. Esto se debe a que, a pesar de estar en el Top 15 de los clientes con más oportunidades cerradas con éxito, lidera ampliamente el ranking de las cuentas con más oportunidades cerradas sin éxito. Por ende, podríamos decir que a este cliente hay que observarlo de cerca; invertir muchos recursos podría ser un error, pero dejarlo ir podría significar una gran pérdida en cuanto a ingresos para la empresa.

Para dejar esto último más en claro, la figura 7 nos muestra el Top 15 de los clientes que más ingresos le generaron a la empresa.

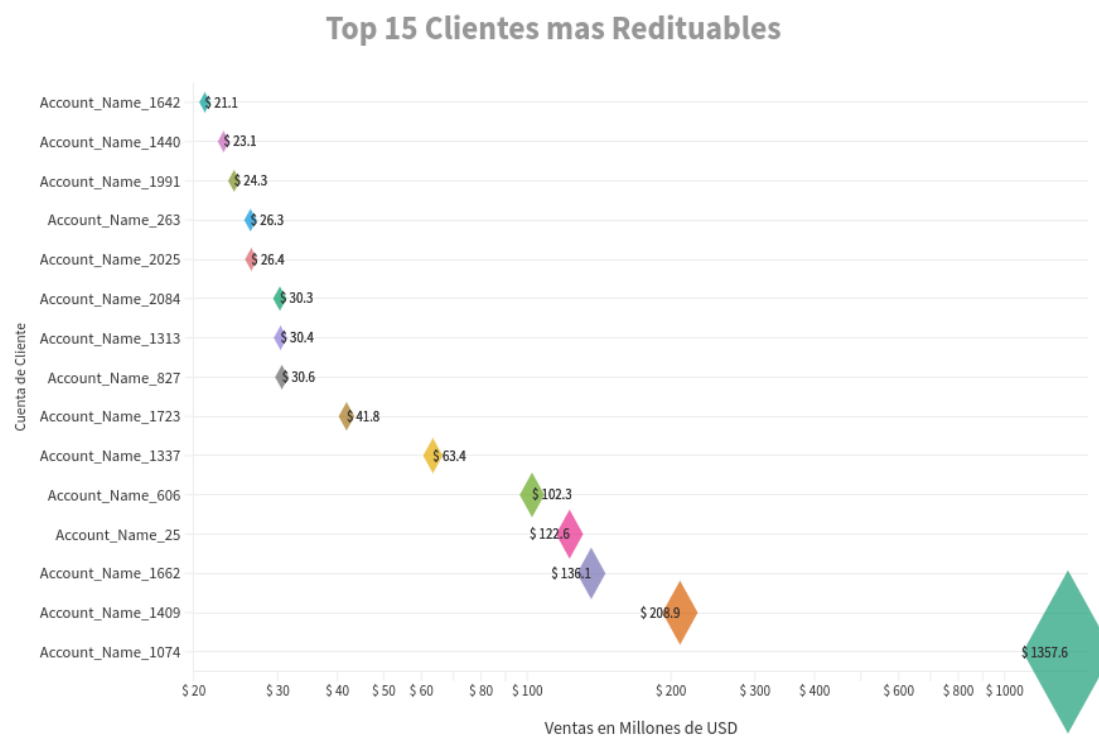


Figura 7: Top 15 Clientes mas redituables

Siguiendo con el análisis de la misma cuenta de antes, la 1888, vemos que esta no figura en el gráfico anterior, lo que confirma aún más lo mencionado anteriormente de que es una cuenta a seguir de cerca para evitar que la empresa invierta esfuerzo y recursos por demás.

4.3. Conclusión

Es importante que para hacer un buen análisis de esta sección, se tomen en cuenta los 4 gráficos en su conjunto. Solo tomar uno como medida, podría llevar a estimaciones de importancia sobre los clientes totalmente erróneas.

5. Restricciones burocráticas que influyen en las ventas

5.1. Introducción

- Requerir del código burocrático 0, ¿afecta una posible venta?
- ¿Qué países son los mas estrictos en términos de burocracia?
- ¿Es conveniente imponer desde la empresa trabas burocráticas, o puede eso afectar las ganancias?

Esas fueron algunas de las preguntas que nos surgieron a la hora de analizar la variable "Necesita aprobación de Código Burocrático 0".

Para llevar a cabo este análisis, decidimos que el "Billing Country" nos daría el país a observar. Nuestra principal teoría, y el motivo que impulsó esta sección, es que los países que imponen o a los que se les imponen trabas burocráticas, ven afectadas la forma en la que culminan sus oportunidades generadas. Esto afecta tanto al cliente, como a la empresa, ya que esta también podría estar sufriendo una disminución en los ingresos pura y exclusivamente por trabas burocráticas, que quizás son innecesarias.

5.2. Analizando país por país

En las figuras 8 y 9, podemos empezar a confirmar lo dicho anteriormente, teniendo en cuenta que los países más oscuros en el mapa, representan un mayor porcentaje de oportunidades perdidas.

Restricciones burocráticas que influyen en las ventas

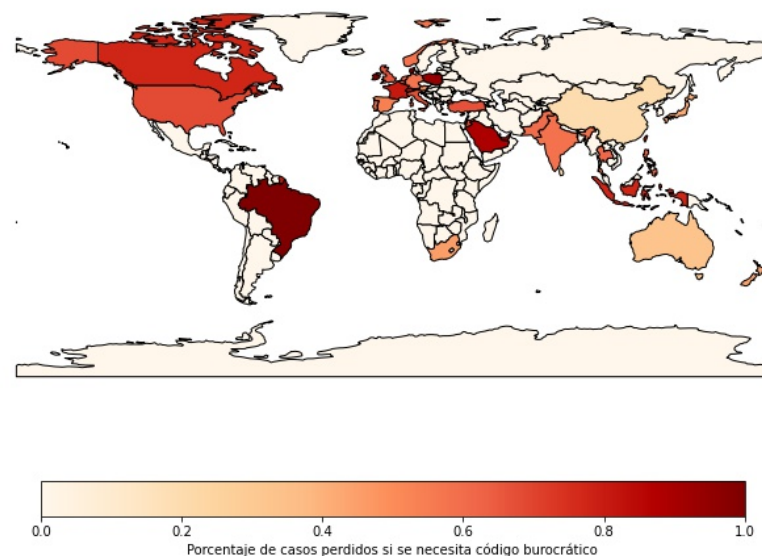


Figura 8: Representación por país de porcentaje de casos perdidos, para aquellas oportunidades que requirieron de aprobación de código burocrático.

Restricciones burocráticas que influyen en las ventas

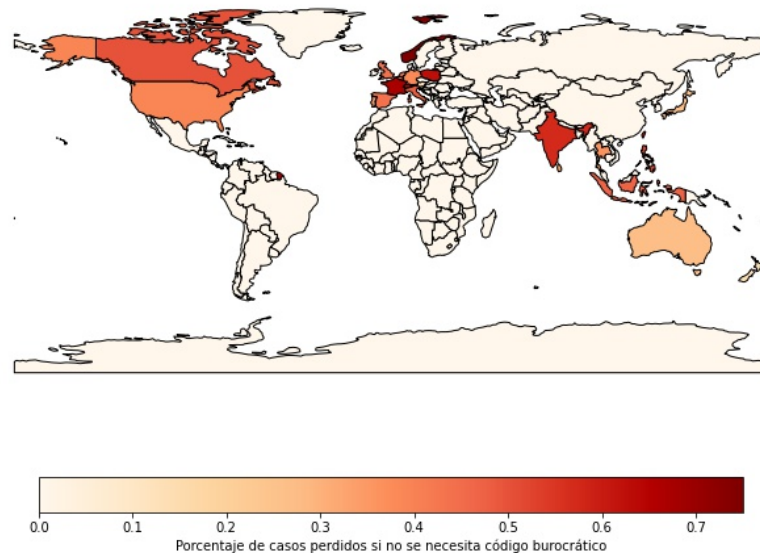


Figura 9: Representación por país de porcentaje de casos perdidos, para aquellas oportunidades que no requirieron de aprobación de código burocrático.

Luego de detenernos a observar detalladamente ambos gráficos, ya podemos hacernos la idea de que lo que propusimos al inicio, es totalmente acertado. El color de la segunda imagen es en casi toda su totalidad más claro que el primero. Para terminar de asentar esta idea, y hacer esta conclusión aún más visible, la figura 10 nos otorga los porcentajes de casos perdidos, según si se requirió o no de la aprobación del código burocrático 0, por país.

Porcentaje de Oportunidades Perdidas según si Tiene o no Código Burocrático

SI=Tiene Código Burocrático - NO=No tiene

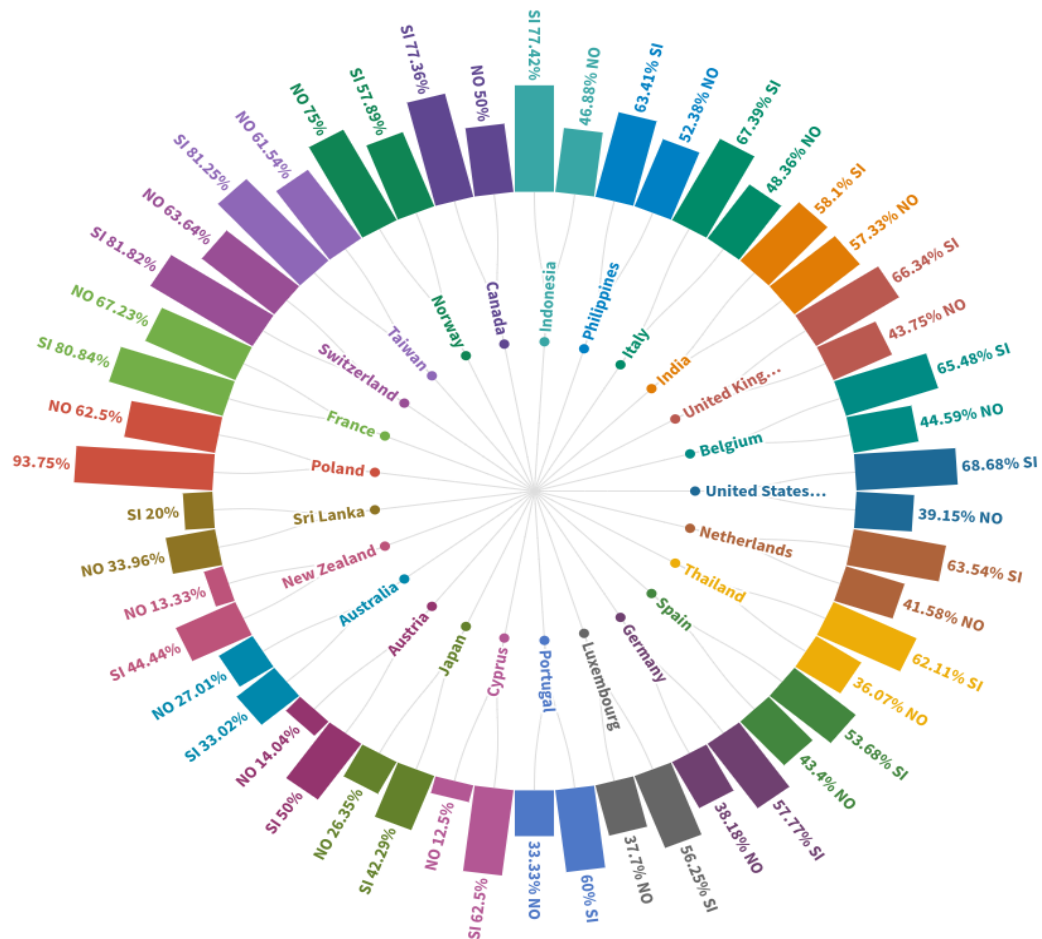


Figura 10: Por país, porcentajes de casos perdidos según si se necesita aprobación de código burocrático 0.

5.3. Conclusiones

De esta sección podemos concluir que debería reverse la necesidad de imponer esta restricción burocrática, ya que es evidente que esta impactando en el rendimiento de la empresa. Quizás este código sea algo sumamente importante, que no es negociable quitar de circulación, pero se podría pensar en reducir la cantidad de oportunidades en las que se aplica.

6. Performance de la empresa a lo largo del tiempo

6.1. Introducción

Hacer un análisis de la performance de la empresa a lo largo de los últimos años, nos puede dar un indicio de cuales políticas repetir y cuales no, de acuerdo al éxito de las mismas.

6.2. Ver el pasado, para poder predecir el futuro

El gráfico 11 busca dar un primer pantallazo de lo que será el análisis de esta sección, mostrándonos la cantidad de oportunidades generadas a lo largo del tiempo. En este punto, aún no nos adentramos en los resultados finales de las mismas.

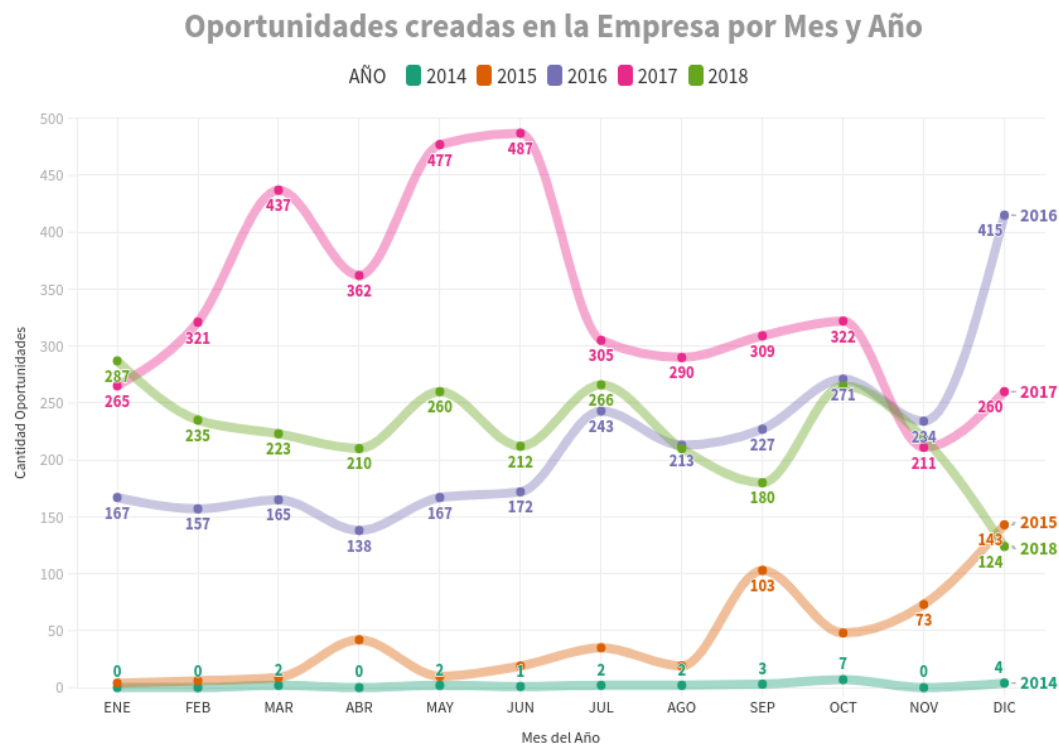


Figura 11: Cantidad de oportunidades generadas por Mes y por Año

Se puede ver que el 2017 fue un gran año y que el 2018 estuvo lejos de poder seguirle el ritmo. Sin embargo, para poder confirmar esto, veamos en la figura 12 que sucedió con esas oportunidades concretadas. Aclaración: Solo tenemos en cuenta aquellas oportunidades que se encuentran en estado Closed Lost o Closed Won, ya que son las más representativas y las que nos importan para este análisis.



Figura 12: Estado de las Oportunidades según el Año

Podemos ver que, efectivamente, no sólo 2017 fue el mejor año en cuanto a oportunidades generadas, sino también en cuanto a oportunidades ganadas.

Vale la pena destacar, que si observamos el gráfico anterior en detenimiento, vemos que la pendiente del segmento que une los años 2016 y 2017, es más pronunciada en la figura que muestra las oportunidades ganadas. ¿Qué nos indica esto? Nos indica que la cantidad de ventas exitosas creció más rápido que la cantidad de ventas cerradas sin éxito, lo cual evidencia una buena performance de la empresa en ese período.

También vemos que entre los años 2015 y 2016 hay una pendiente pronunciada, pero en este caso esto se ve en ambas figuras, lo cual es un indicio de que el ritmo de crecimiento fue similar para aquellas oportunidades cerradas con éxito, y aquellas cerradas sin éxito, lo cual no es ni bueno ni malo.

6.3. Meses más calurosos, ¿implican más ventas?

Dado que la empresa en cuestión esta dedicada a la venta e instalación de equipos de aire acondicionado para grandes superficies, ¿hay algún mes en el que se generen mas oportunidades, debido a que se aproximan épocas con altas temperaturas? Veamos el siguiente gráfico (13).

Cantidad de Oportunidades Promedio Generadas por Mes

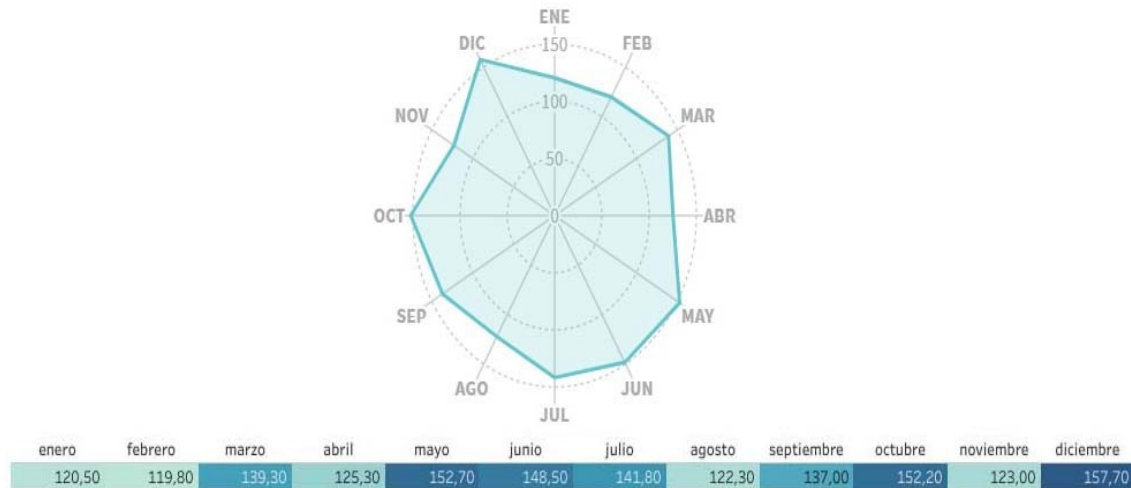


Figura 13: Promedio de Oportunidades Generadas por Mes

Vemos que si bien hay algunos meses en los que se generan más oportunidades que en otros, está muy parejo y no creemos que se pueda concluir algo concreto que valga la pena informar.

7. Análisis sobre Regiones

7.1. Introducción

En esta sección buscamos analizar ciertos aspectos según la Región:

- EMEA: hace referencia a Europa, Medio Oriente y África.
- Japan: hace referencia al país asiático
- Americas: hace referencia a los continente Americanos
- APAC: hace referencia a la parte del mundo dentro o cerca del océano Pacífico occidental.
- Middle East: hace referencia a Oriente Medio

Para empezar con este análisis, el siguiente gráfico (14) muestra la cantidad de dinero (en millones de dólares) que representan las oportunidades generadas para cada una de las regiones, agrupadas por el estado de las mismas.



Figura 14: Estado de las oportunidades según la Región

Vemos que los estadíos de interés son **Closed Lost** y **Closed Won**. Los otros tres, representan una mínima parte de los posible ingresos ya que son estados intermedios por los que van pasando cada una de las distintas oportunidades.

A raíz de este simple gráfico, vemos que **Japan** parece ser una buena región para centrar los recursos de la empresa, debido a la alta recaudación de dinero, y a la poca pérdida del mismo debido a oportunidades cerradas sin éxito.

7.2. Porcentaje de oportunidades según región

Al momento de iniciar una oportunidad, sería interesante saber la probabilidad de que esta sea exitosa según cada región. Esto serviría para poder así elegir la región en la cual crearemos nuestra oportunidad.

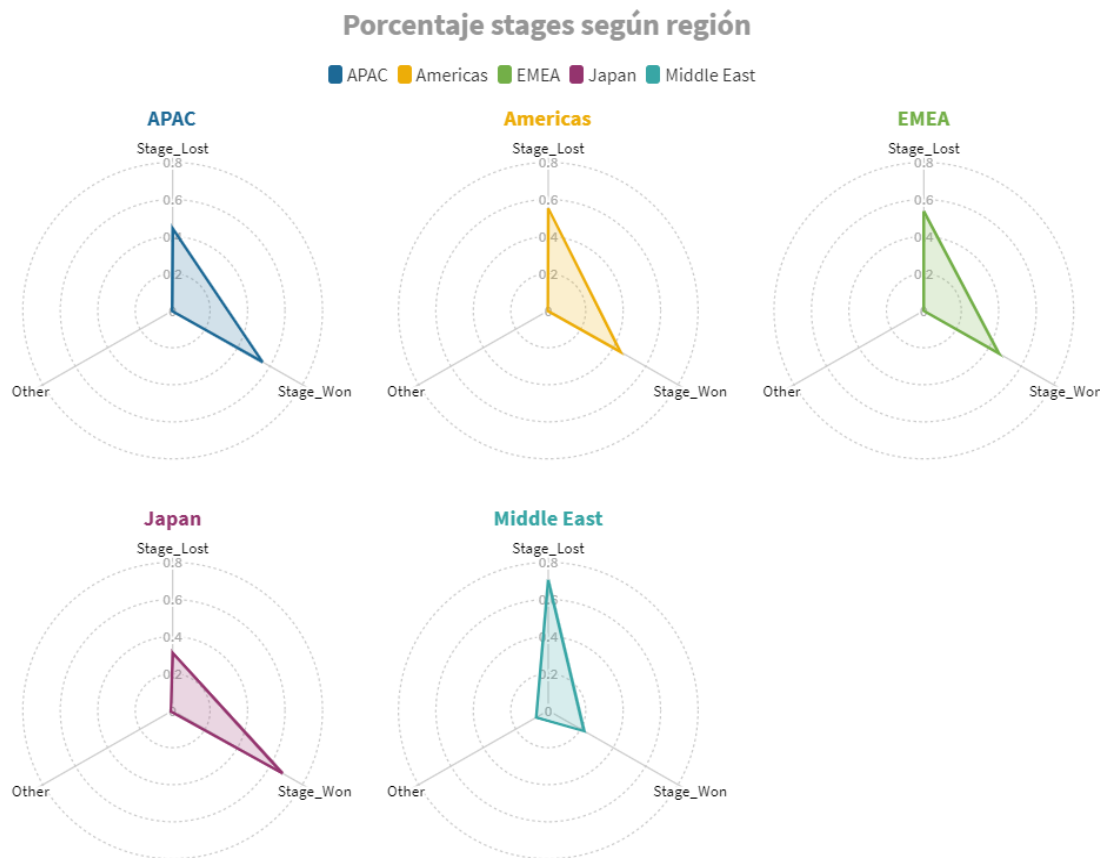


Figura 15: Muestra el porcentaje de cada stage según la región.

Como podemos ver en el gráfico 15 claramente sería a nuestra conveniencia empezar una oportunidad en la región de **Japan** ya que esta tiene una probabilidad de éxito de casi el 80 por ciento. Por otro lado podemos ver que empezar una oportunidad en la región de **Middle East** es una complicación ya que tiene una probabilidad de terminar en Closed Lost de casi el 80 por ciento.

7.3. Oportunidades perdidas por región

En este caso vamos a analizar la cantidad de oportunidades perdidas por región.

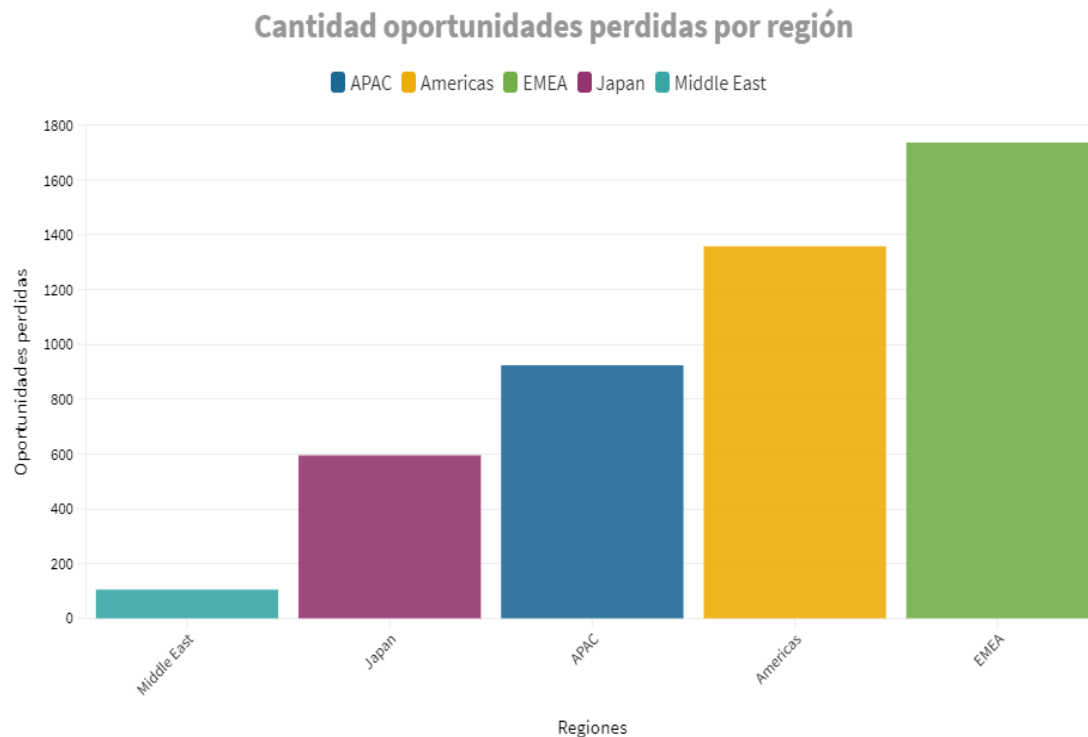


Figura 16: Porcentaje oportunidades perdidas por Región

Como vemos en la figura (16) la región **Middle East** es la que menos oportunidades perdidas tiene, si bien en el análisis anterior vimos que era la región con un porcentaje de pérdidas mayor que los demás. Esto nos muestra como debido a la alta tasa de pérdida de la región, la cantidad de oportunidades es menor al resto, lo cual tiene sentido. Por otro lado vemos que **EMEA** es la región con más oportunidades perdidas, si bien tiene un porcentaje de éxito de mas del 50 por ciento. Esto nos lleva a la siguiente hipótesis. que **EMEA** es la región con más oportunidades.

7.4. Relación Oportunidades ganadas y totales

Para responder a la hipótesis planteada en la sección anterior, debemos poder ver la cantidad de oportunidades totales según cada región.

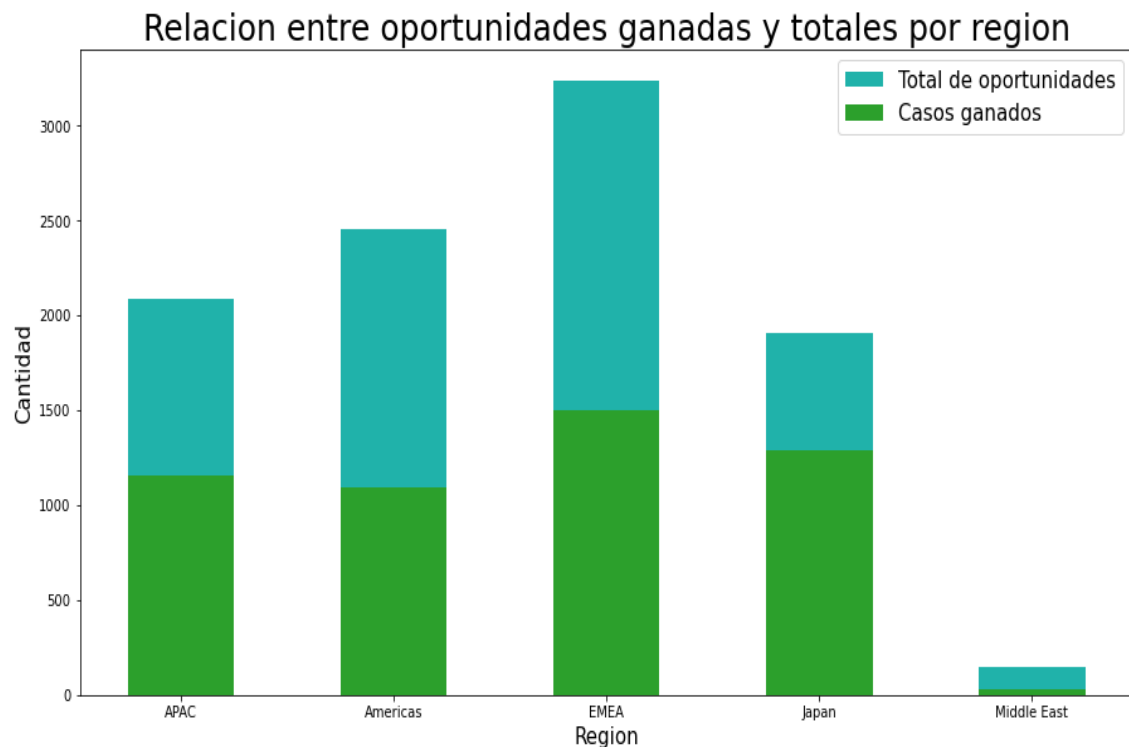


Figura 17: Ventas totales y ganadas por Región

Como vemos en la Figura 17, la hipótesis de que EMEA es la región con mayor cantidad de oportunidades es correcta. Con la ayuda de la misma, podemos ver como la razón de la gran cantidad de oportunidades perdidas, se debe a la gran cantidad de oportunidades.

7.5. Conclusión

Según lo analizado sobre las regiones, podemos decir que **Japan** es la región más exitosa de las 5. Mientras que por otro lado **Middle East** sería la menos exitosa. Es por esto que se le recomienda a un futuro vendedor que al momento de querer iniciar una oportunidad nueva, considere inicialmente hacerlo en **Japan**. Por otro lado, a partir de este análisis surge una duda a ser respondida mas adelante. Hay algún factor que pueda hacer que si bien **Japan** es la región mas exitosa, no sea la que posea mas oportunidades ?.

8. Análisis sobre Productos

8.1. Introducción

El análisis de los productos que una empresa vende es vital para que la misma pueda administrar recursos, tales como la logística para la distribución de los mismos y centralización de plantas de fabricación, de la manera más eficiente.

8.2. Familias de Producto mas Vendidas por Año

En el gráfico 18 se puede ver cuáles fueron las familias de productos más vendidas por año. Algo que se puede ver a simple vista es que en el 2014 la cantidad de ventas es muy baja en comparación con los siguientes. Esto puede deberse a que en ese año la empresa lanzó una línea de productos inicial para evaluar el mercado mientras que en los años siguientes expandió sus líneas de productos incrementando las ventas, o simplemente a que no se cargaron los datos correctamente para ese año.

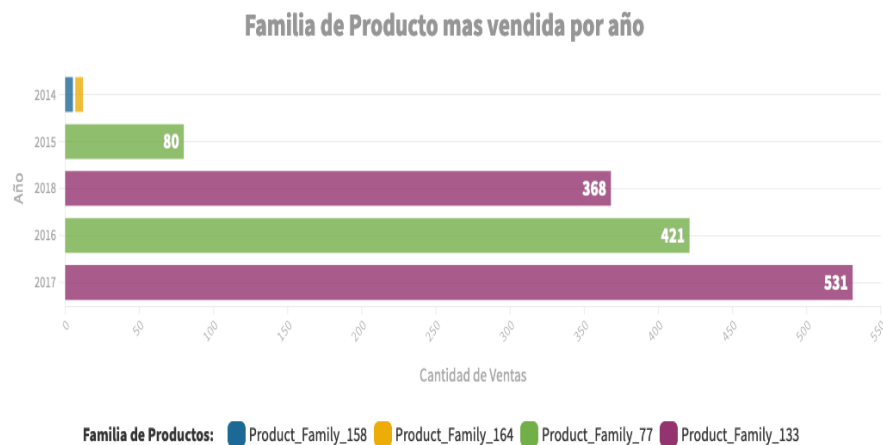


Figura 18: Por año, cantidad de ventas concretadas exitosamente de la familia de producto mas vendida.

La siguiente pregunta que podemos hacernos es: ¿cómo están distribuidas estas ventas en las distintas regiones?, ¿están distribuidas uniformemente o se concentran únicamente en una región determinada?

8.3. Distribución de las Ventas según Región

En el gráfico 19 lo primero que podemos observar es que para el año 2014 y las familias de productos 158 y 164 no hay un gráfico asociado. Esto es porque la cantidad de ventas que tienen son tan pocas que no tiene sentido considerarlas cuando lo que se busca visualizar es una tendencia marcada de las ventas por región.

Ahora lo interesante de ver en este gráfico es que para la región Middle East no hay ventas en la absoluto de estas familias de productos.

Para la familia de producto 77 en los años 2015 y 2016 las ventas se concentraron en las regiones de **Japan** y **EMEA**. Y para la familia de producto 133 las ventas se concentraron en las regiones **EMEA**, **Americas** y **APAC**.

Pero viendo cada gráfico por separado dificulta visualizar claramente si hay una centralización de las ventas o no.

Distribución de Ventas por Región para la Mejor Familia de Productos de cada Año

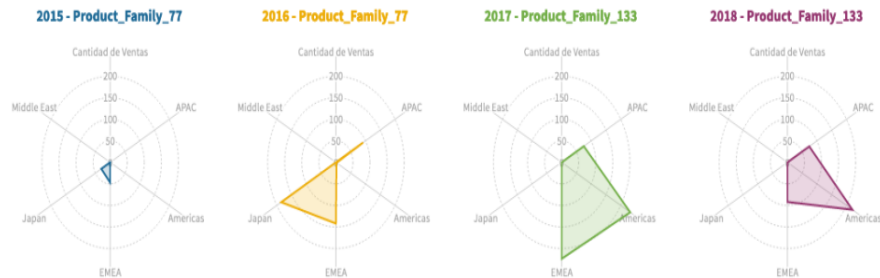


Figura 19: Distribución de ventas por región las Familias de Productos mas vendidas en cada Año

El gráfico de la figura 20 nos da una mejor visualización de la centralización mencionada con anterioridad. Como se puede ver, el solapamiento de las areas se concentra en las áreas correspondientes a las regiones de **Japan**, **EMEA** y **Americas**.

Distribución de Ventas por Región para la Mejor Familia de Productos de cada Año

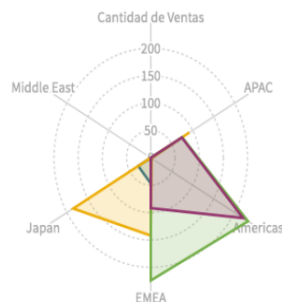


Figura 20: Combinado de la distribución de ventas por región las Familias de Productos mas vendidas en cada Año

Hasta acá podemos concluir que una empresa que está haciendo un análisis de la logística de fabricación y distribución de sus productos para una futura expansión debería considerar que la mayor parte de sus ventas en los últimos años provienen de las regiones ya mencionadas.

8.4. Productos mas vendidos por Región

Por último en la figura 21 se muestra el top 5 de productos mas comprados según la región. A primera vista podemos ver una diferencia muy clara en las ventas correspondiente a la región de Middle East que se representan con la franja violeta delgada y las demás ventas en el resto de las regiones.

En relación al análisis hecho con anterioridad podemos ver este patrón que se repite, donde tenemos una cantidad muy pequeña, prácticamente despreciable, de ventas proveniente de Middle East y una distribución mas marcada de las ventas en las Regiones Japan, EMEA, APAC y Americas.

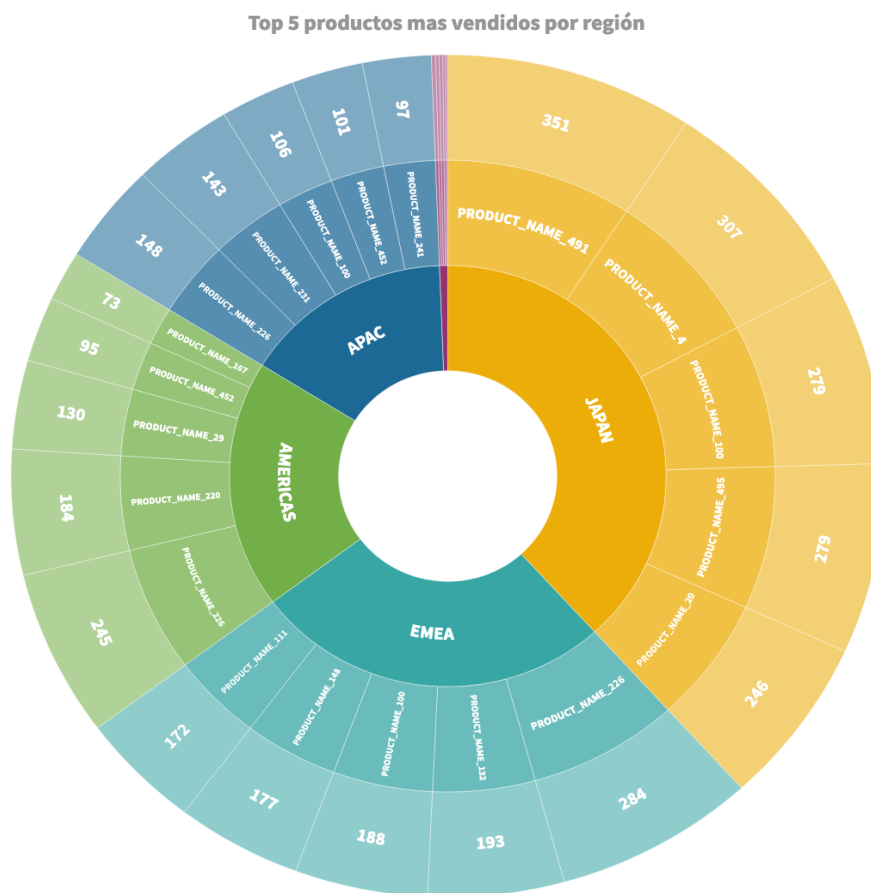


Figura 21: Top 5 de los productos mas comprados por región

Ahora queda en manos de la empresa "Frio Frio" analizar la logística de fabricación y distribución no solo pensando en satisfacer la demanda creciente proveniente de estas regiones sino también pensando en una expansión hacia futuro.

8.5. Los 10 productos con mayor recaudación

Toda empresa busca generar el mayor ingreso posible, es por eso que conocer los productos que más recauden es fundamental. En el gráfico 22 mostramos los 10 productos con mayor recaudación.

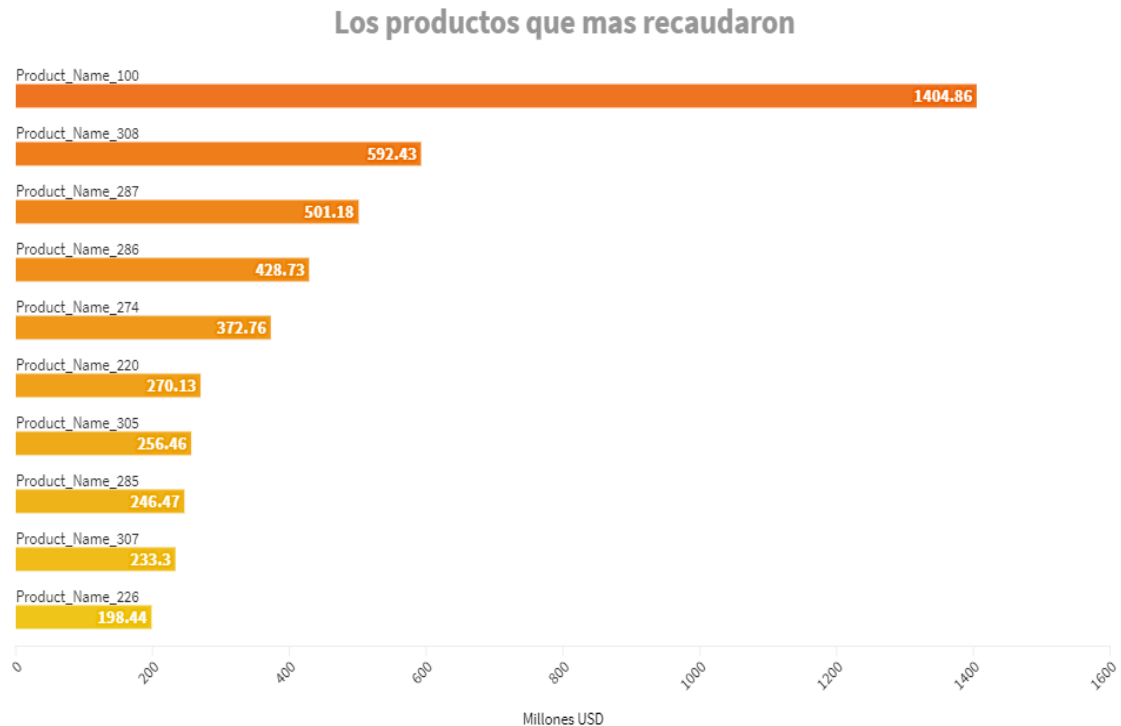


Figura 22: Muestra los 10 productos con mayor recaudación.

Según el gráfico el producto 100 es el que mas recaudo con un total de 1404.86 millones de dólares. Este se encuentra muy por arriba del resto con una diferencia casi del doble con respecto al segundo producto más vendido. Esto hace que el producto 100 sea uno a tener en cuenta, por ejemplo en el caso que se quiera aumentar las ganancias de la empresa.

8.6. Los 10 productos más y menos vendidos.

En esta sección vamos a analizar los 10 productos más vendidos y los 10 menos vendidos. Una pregunta a tener en mente, es si hay alguna relación entre los 10 más vendidos y aquellos que recaudan más.

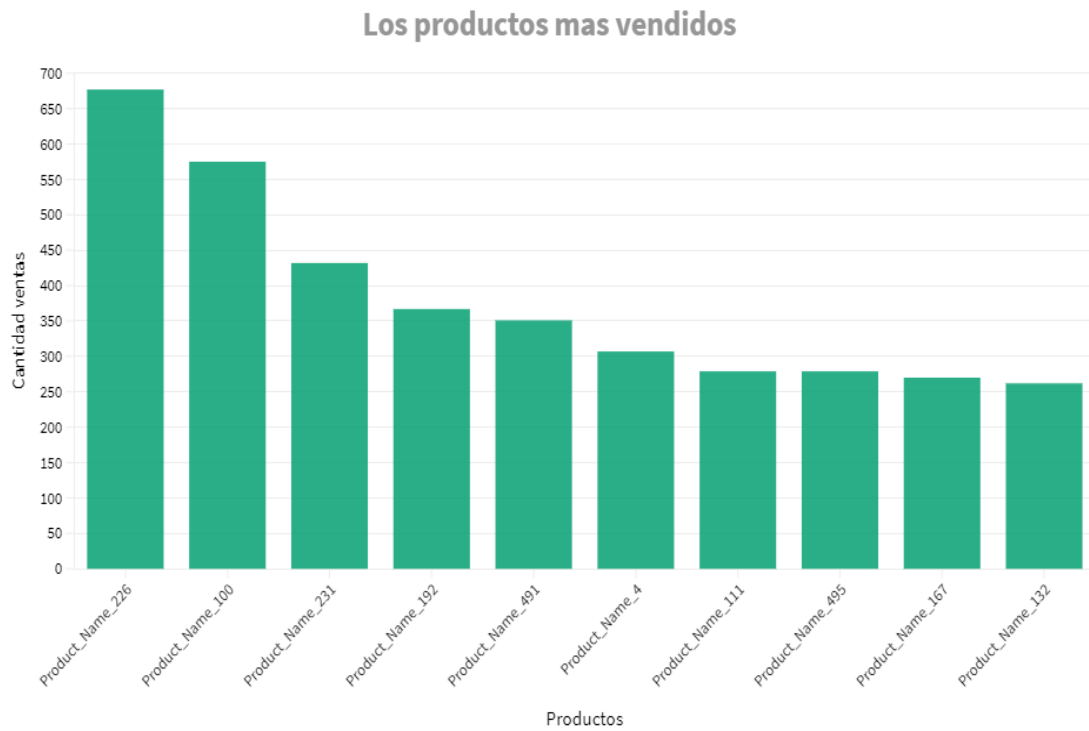


Figura 23: Muestra los 10 productos mas y menos vendidos.

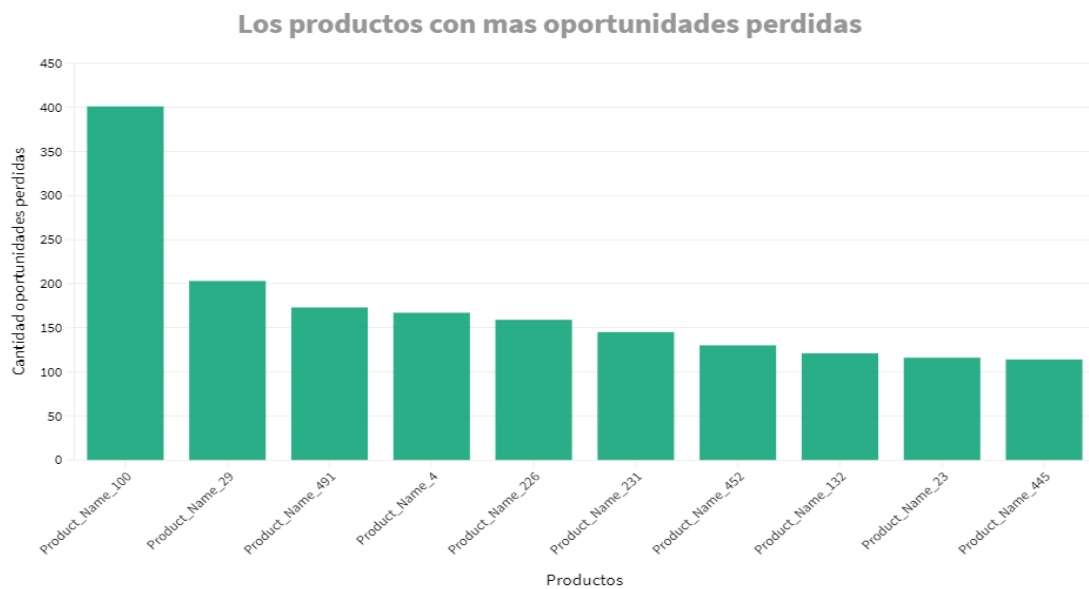


Figura 24: Muestra los 10 productos con mayor cantidad de oportunidades perdidas.

9. Análisis sobre TRF

El foco de los análisis de esta sección estará puesto en las TRF.

9.1. Análisis de ventas según TRF

Para el siguiente análisis decidimos tomar aquellos productos que tuviesen un valor de TRF mayor a 0, por lo explicado en 2.2, ya que no sabemos exactamente a cual de las dos opciones planteadas hace referencia cada valor 0.

En el mismo, buscamos establecer alguna relación entre las TRF y las ventas producidas. ¿Cuáles son los productos que mas se venden? aquellos con mayor o menor TRF?

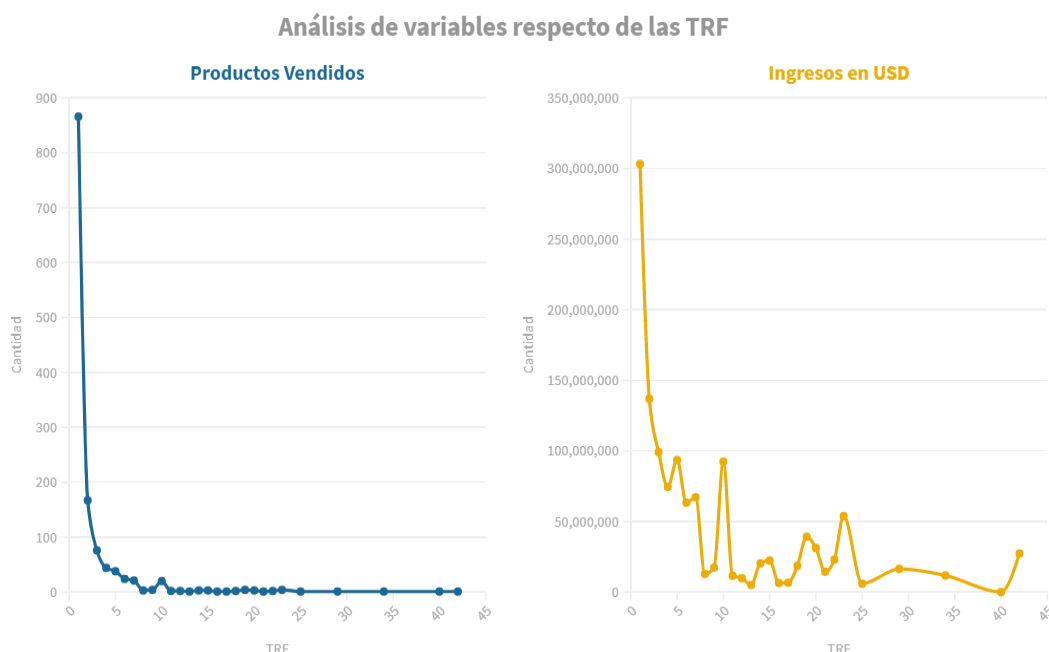


Figura 25: Productos vendidos e ingresos generados según TRF

Con respecto a la cantidad de productos vendidos, vemos que la tendencia de que estas disminuyan conforme aumentan las TRF es bastante razonable. Generalmente, los productos con menor TRF suelen tener un precio menor, lo que los convierte en productos más "populares." accesibles que los de mayor TRF. Además, sabiendo que los aires acondicionados de uso cotidiano suelen tener TRF similares a 1, se explica que estos productos tengan una mayor cantidad de ventas por el simple hecho de tener mas compradores posibles, mientras que los productos de mayor TRF están hechos para ciertos usos específicos, reduciendo así las posibilidades de encontrar compradores.

Con respecto a los ingresos generados, se sigue una lógica similar a la mencionada anteriormente.

Por otro lado podemos ver que hay algunos productos con TRF intermedios o altos que no generan ni muchos ingresos ni muchas ventas. Teniendo en cuenta esto, la empresa podría revisar y replantearse la producción de estos equipos, buscando quizás abaratar los costos para poder hacerlos más accesibles y aumentar las ventas o reducir la oferta o variedad de productos de altas TRF dado que hay algunos de estos que no se venden tanto (ejemplo: en vez de ofrecer un producto con 7 u 8 TRF, directamente producir mas unidades de 10 que pueda cumplan con ambos propósitos).

10. Análisis sobre rango entrega

En esta sección se busca analizar el rango de entrega de las oportunidades. Para obtener el rango se paso a utilizar del set de datos las columnas `Planned_Delivery_Start_Date` y `Planned_Delivery_End_Date`. Con estas columnas se restaba el `end_date` con el `start_date` y en base a esto se obtenía un rango de la duración de la entrega. Con este nuevo dato del rango pasaremos a ver que relaciones presenta, con la probabilidad de éxito de una oportunidad, así como las regiones y los vendedores.

10.1. Rango de entrega y porcentaje de éxito

En esta sección buscamos responder la siguiente pregunta:

- A menor duración de entrega, ¿mayor cantidad de éxito?

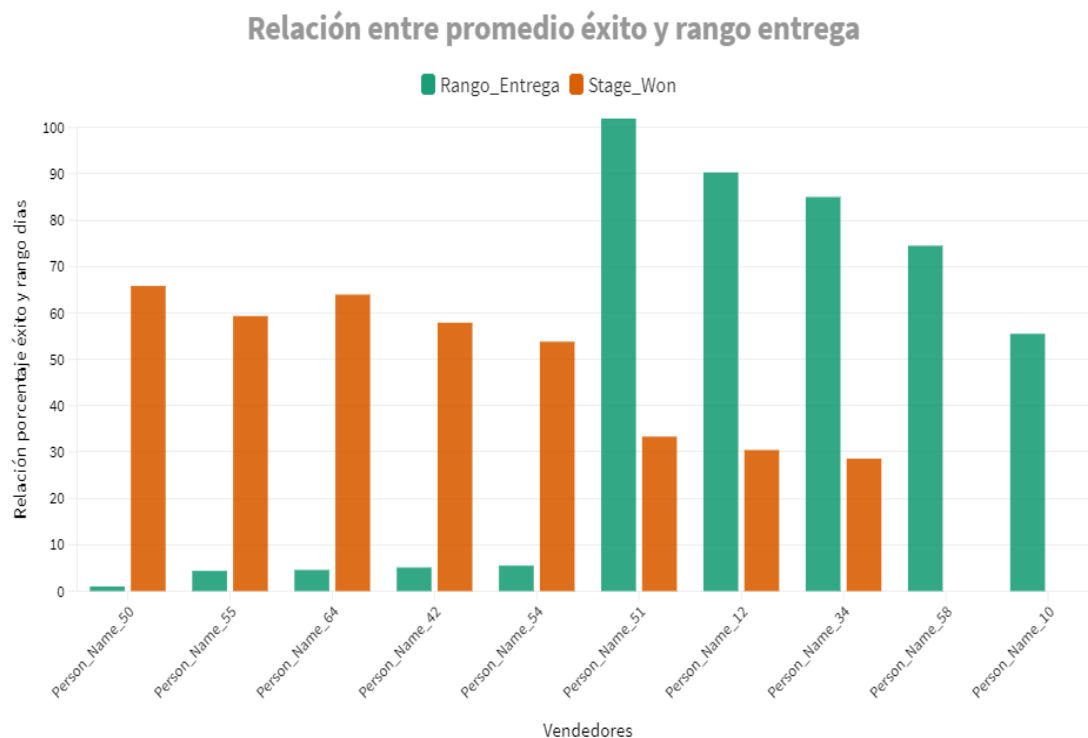


Figura 26: Muestra los 5 mayores y menores rango de entrega de los vendedores y compara con su porcentaje de éxito.

Con el gráfico 26 podemos pasar a responder la pregunta planteada al comienzo de esta sección. Podemos observar claramente que los vendedores que poseen un rango de entrega menor, son los que se encuentran con una mayor cantidad de oportunidades ganadas, mientras que por otro lado los vendedores que poseen un rango de entrega mayor, son aquellos que tienen una cantidad de oportunidades ganadas inferior. Es por esto que la respuesta a nuestra pregunta es afirmativa; mientras más corto sea el período de entrega, mayor probabilidad de éxito tendrá la oportunidad.

10.2. Rango de entrega según Regiones

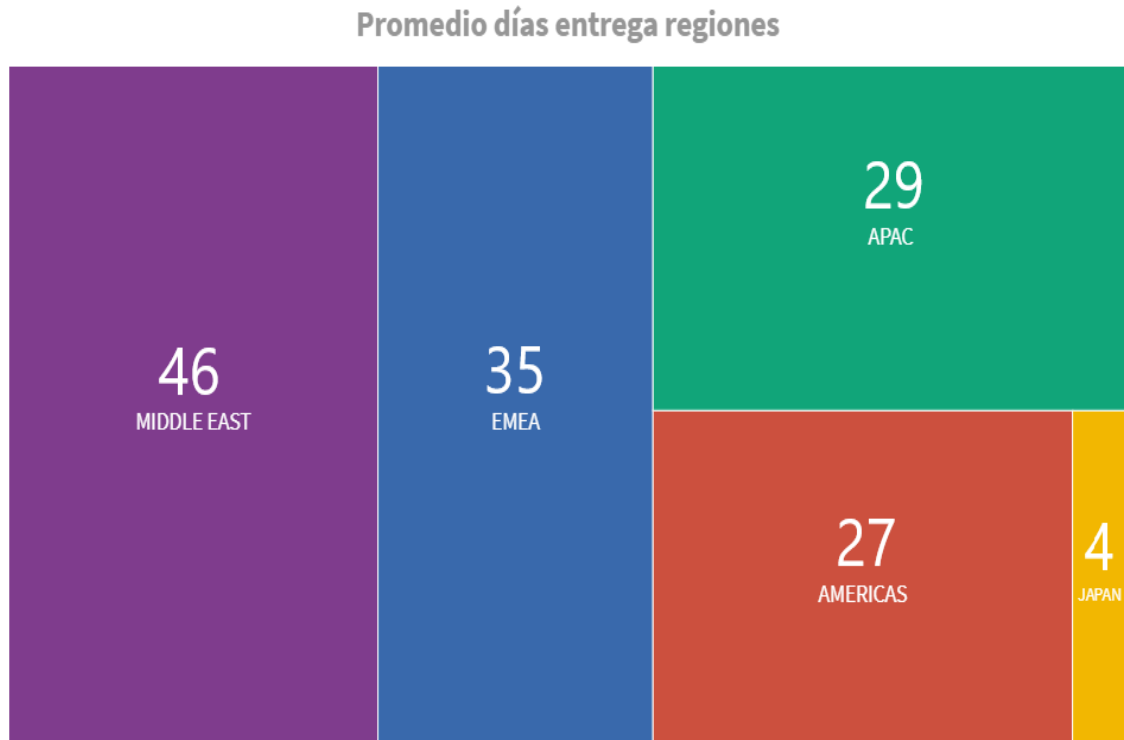


Figura 27: Muestra el promedio de días de una entrega en cada región.

En la Figura de arriba se puede ver el promedio en días de la duración de una entrega por región. Estos datos los podemos relacionar con el porcentaje de éxito de cada región. Como muestra la Figura 27, **Japan** es la región con mayor probabilidad de éxito y a la vez la que tiene una duración de entrega menor, mientras que por otro lado **Middle East** es la que menor porcentaje de éxito tiene y a su vez la que mayor duración de entrega posee. Esto nos muestra una vez mas que a menor duración de entrega mayor probabilidad de éxito tendrá la oportunidad.

11. Análisis sobre requerimiento de aprobación especial

En esta sección vamos a analizar la variable Pricing_Delivery_Terms_Quote_Approval, que indica si la oportunidad requiere una aprobación especial de su precio total y términos de entrega. En este análisis buscamos ver si influye en el éxito de las oportunidades.

11.1. Stages y sus aprobaciones especial

En el gráfico 29 se muestra la cantidad de oportunidades que requieren de aprobación especial según las stages.

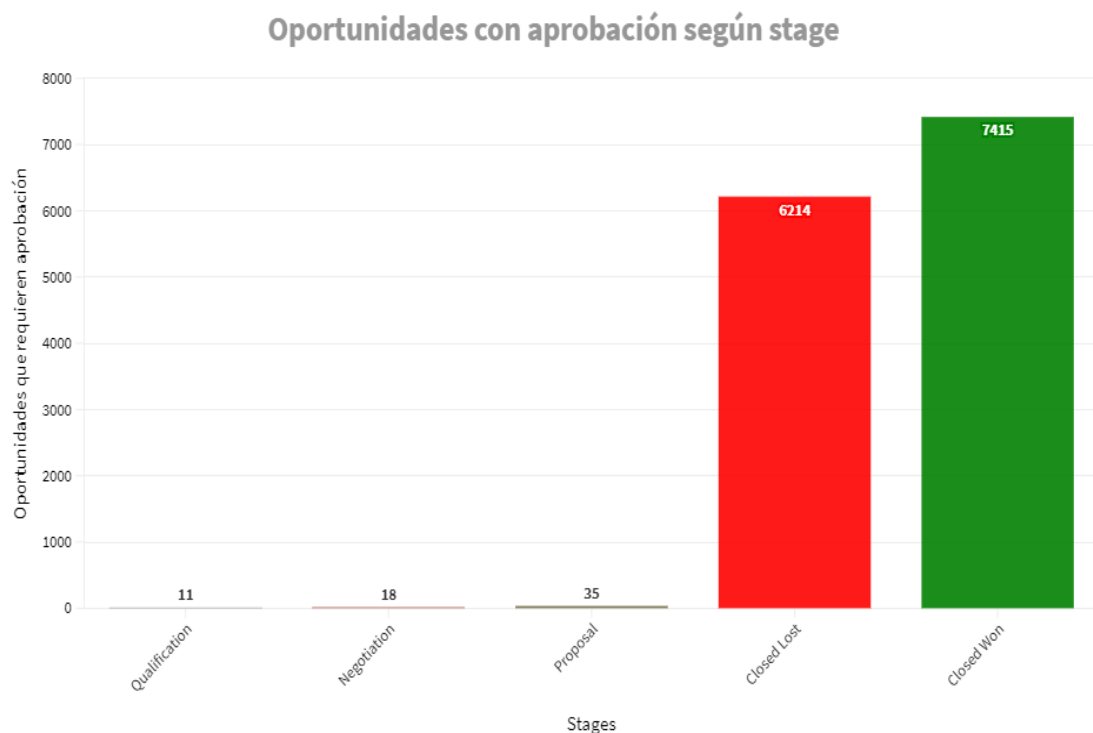


Figura 28: Podemos ver la cantidad de aprobaciones especiales requeridas por stage.

11.2. Porcentaje de oportunidades según región

Una pregunta que nos hicimos es si realmente influía el requerimiento de una aprobación especial sobre el porcentaje de éxito.

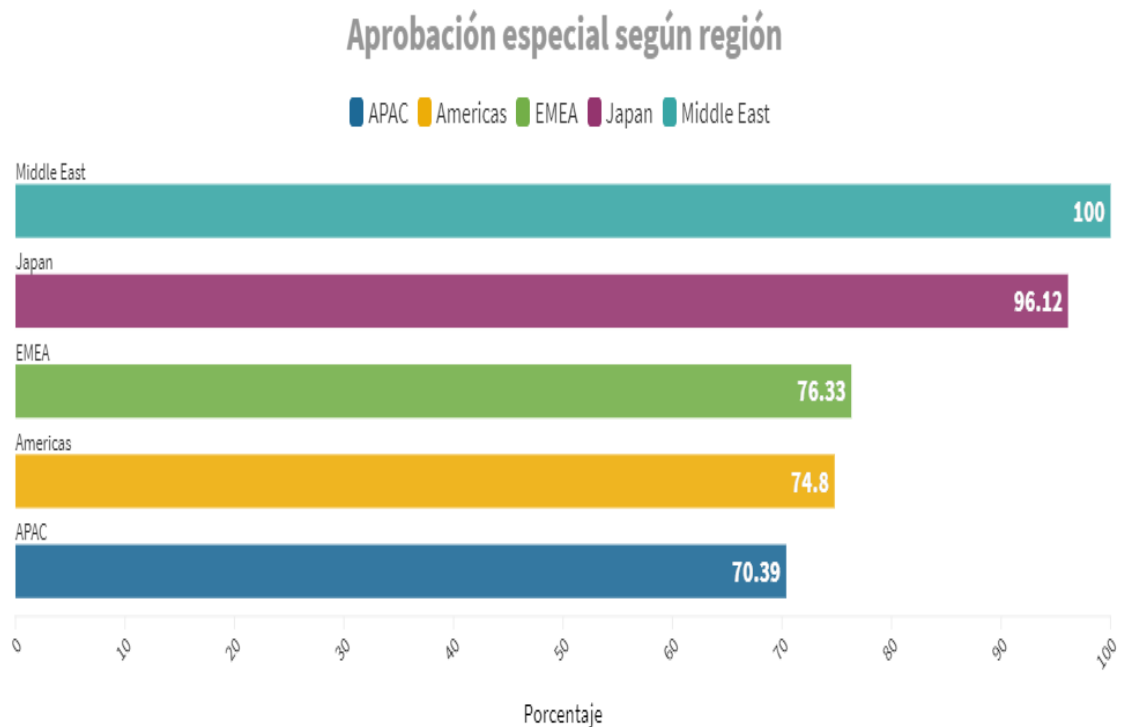


Figura 29: Muestra el porcentaje de requerimiento de aprobación de las oportunidades según región.

En el gráfico 29 podemos ver como **Japan** y **Middle East** son dos regiones con alto porcentaje de requerimiento de aprobación especial. Si estos resultados los comparamos con los del gráfico 17 podemos ver como claramente tanto **Japan** y **Middle East**, son las dos regiones con menor cantidad de oportunidades. Esto nos lleva a pensar si estas dos regiones presentan menor cantidad de oportunidades debido a estas trabas impuestas por el requerimiento de una aprobación.

11.3. Países con mayor requerimiento de aprobación especial

Finalmente en esta sección se busca mostrar los 10 países que presentan una mayor cantidad de oportunidades con aprobación especial requerida.



Figura 30: Muestra los 10 países con la mayor cantidad de aprobaciones especiales.

12. Conclusiones generales

Las conclusiones a las cuales llegamos a partir de los distintos análisis realizados son las siguientes:

- El plazo de entrega influye en el porcentaje de éxito. A mayor plazo , mayor probabilidad de que la venta fracase.
- Las trabas burocráticas influyen negativamente en la creación de oportunidades y el cierre de ventas.
- Los productos con menor TRF son los más populares y los que generan más ingresos.
- El 2017 fue el año donde la empresa tuvo su mejor rendimiento. En el mismo, alcanzó un pico, y fue seguido de un 2018 con bajas notorias en el rendimiento.
- Japón es la región con mayor cantidad de oportunidades y tasa de éxito en ventas, mientras que Middle East registra los menores valores para ambos campos.
- A la hora de pensar en dónde y en quiénes invertir los recursos de la empresa, es importante no dejarse guiar únicamente por un gráfico, sino tener en cuenta todo el escenario.
- Mayor cantidad de oportunidades generadas, no implica mayores ganancias para la empresa, o un cliente importante. Es necesario ver que sucedió con las mismas para poder tomar una decisión de ese calibre, de lo contrario se estarían tomando decisiones en base a análisis incompletos.

13. Referencias

Repositorio de Github,

<https://github.com/FerFabbiano/tp1-OrganizacionDeDatos>

Herramienta de Visualización de Datos Flourish,

<https://flourish.studio/>