

1 Estudio de las Ventas de una Ferretería

Autor: Diana Chacón Ocariz

1.1 Contexto:

Se trata de estudiar las ventas de una pequeña ferretería a partir del 01/01/2020.

Esta empresa maneja un poco más de 3.000 productos distintos. Poseen un software de gestión genérico que les provee una gran cantidad de reportes, básicamente tablas con números, difíciles de analizar (un reporte puede constar de varias decenas de páginas).

Por otro lado, los montos manejados están en Bs (Bolívares, moneda nacional venezolana). Durante el período de estudio, las grandes tasas de inflación, además de un cambio de moneda (eliminación de 6 ceros), dificultan el análisis. Por eso ha sido necesario transformar los montos a dólares según la tasa del día.

Desde hace unos 4 años trabajo con esta empresa, conozco cómo funciona y el tipo de productos que maneja. Eso me motivó a emprender este proyecto.

1.2 Objetivos del negocio:

Tener más visibilidad sobre las ventas para poder mejorar el proceso de compras y la toma de decisiones en general:

- Analizar objetivamente las ventas.
- Identificar los productos más vendidos y que no pueden faltar (productos indispensables).
- Determinar los productos que podrían entrar en rotura de stock al final de un período.
- Predecir las ventas futuras.

1.3 Objetivos académicos:

- Estudiar un caso real, con datos reales y cuyo resultado pueda ayudar a alguien a resolver un problema.
- Demostrar que la ciencia de datos puede ayudar a cualquier empresa, incluso PYMES.
- Conocer y practicar el uso de herramientas de ciencia de datos.

1.4 Fuentes de datos:

Los datos provienen de reportes sacados del software de gestión de la empresa. Se trata de archivos .xls que contienen sólo los datos de reportes sobre facturas (2020), ventas por producto (2021 y 2022) e inventario a la fecha.

Por razones de seguridad, para este estudio eliminamos toda la información concerniente a los clientes. Cualquier uso de los datos fuera de este estudio, deberá ser notificado a la empresa.

También se utilizan los datos del [histórico de la tasa de cambio del dólar \(https://monitordolarvenezuela.com/historial#2a\)](https://monitordolarvenezuela.com/historial#2a) que se obtiene a través de un [scraper \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB5_scraper_tasa_dolar.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB5_scraper_tasa_dolar.ipynb).

Más detalles de los archivos utilizados, en el [NB1: Carga y limpieza de los Datos \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB1_analisis_ventas_load.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB1_analisis_ventas_load.ipynb).

1.5 Notebooks del Estudio:

1. [NB1: Carga y limpieza de los Datos \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB1_analisis_ventas_load.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB1_analisis_ventas_load.ipynb): Carga, limpieza y transformación de los datos. Almacenamiento de los datos en archivos .parquet y .xlsx para uso en los otros notebooks y Google Data Studio.
2. [NB2: EDA y Visualización de los Datos \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB2_analisis_ventas_eda.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB2_analisis_ventas_eda.ipynb): Exploración de los datos y análisis para responder preguntas del negocio.
3. [NB3: Reporte Análisis de Ventas \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB3_analisis_ventas_reporte.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB3_analisis_ventas_reporte.ipynb): Reporte de Negocio, dirigido a los directivos de la empresa con conclusiones y recomendaciones.
4. [NB4: Predicción de Ventas - Estudio de una Serie Temporal \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB4_prediccion_ventas.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB4_prediccion_ventas.ipynb): Estudio de series temporales y predicciones de ventas.
5. [NB5: Scraper Tasa de Cambio Dólar \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB5_scraper_tasa_dolar.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB5_scraper_tasa_dolar.ipynb): Scraper para obtener el histórico de la tasa de cambio del USD.

1.6 Reporte de Análisis de Ventas:

Reporte interactivo destinado a los gerentes de la empresa:

1. [Reporte en Google Data Studio \(https://datastudio.google.com/reporting/a4e9910d-447d-417a-80a9-a1e5ae4a4b22\)](https://datastudio.google.com/reporting/a4e9910d-447d-417a-80a9-a1e5ae4a4b22).

1.7 Hallazgos y resultados:

1. Hemos identificado algunos errores a nivel de los datos debidos a malas prácticas al momento de introducirlos: Reutilización de códigos de productos, productos mal clasificados dentro de las líneas de productos. Acciones correctivas serán aplicadas por el cliente para mejorar la calidad de los datos.
2. El cliente tiene hoy una visión más clara sobre cómo evolucionan las ventas, los momentos de mayor flujo de clientes, las líneas de productos y los productos que más se venden...
3. La definición e identificación de lo que son los productos indispensables (aquellos con ventas regulares y que no pueden faltar) es hasta ahora lo más útil del estudio, pues permite al cliente mejorar su proceso de compras.

1.8 Limitaciones y dificultades:

EL proyecto aún no ha terminado. Falta:

1. Buscar datos sobre las últimas entradas en stock de los productos para identificar aquellos que no se venden.
2. Terminar los reportes dirigidos a los gerentes de la empresa (NB3 y Reporte Google Data Studio).
3. Terminar el estudio de la serie temporal y afinar los modelos, tanto para la predicción de ventas en USD como la predicción de las ventas de los productos indispensables.
4. Automatizar el proceso una vez validado el estudio.

Lo que más se me ha dificultado ha sido todo lo concerniente al estudio de las series temporales y las predicciones: Significado de los gráficos, elección de los modelos y ajuste de los parámetros de los mismos para conseguir los mejores resultados.

1.9 Instrucciones para la ejecución del proyecto:

1. Debe ejecutarse primero el NB1 que se encarga de cargar los datos brutos (raw). Los archivos con los datos se encuentran en el subdirectorio **datos/in**
2. En el mismo subdirectorio, también se encuentra el archivo **tasa_dolar.csv** que contiene el histórico de las tasas de cambio del día hasta el 01/03/2022. Si este archivo no está presente, puede ejecutarse el NB5
3. Una vez ejecutado el NB1, obtendremos en el subdirectorio **datos/out** todos los archivos con los datos necesarios para la ejecución de los notebooks NB2, NB3 y NB4. La ejecución de cada notebook es independiente

2 Reporte

2.1 Resumen

Estudio de las ventas del 01/01/2020 al 20/02/2022:

1. 3 métricas nos interesan:
 - A. Ventas en USD
 - B. Volumen de Ventas
 - C. Afluencia de Clientes (dada por la cantidad de facturas)
2. De manera general las ventas han ido aumentando, en particular a partir del 2do semestre del 2021.
3. Este año 2022 empezó con fuerza y hemos tenido las mayores ventas del período
4. Los días de más afluencia de clientes son los martes y miércoles, seguido de los jueves y lunes. Principalmente en las mañanas de 10am a 12pm.
5. Dentro de las líneas de productos más vendidas (ventas en USD y/o volumen de ventas), se encuentran aquellas relacionadas con los sistemas de riego, lo que confirma la especialización de la empresa en este tipo de productos.

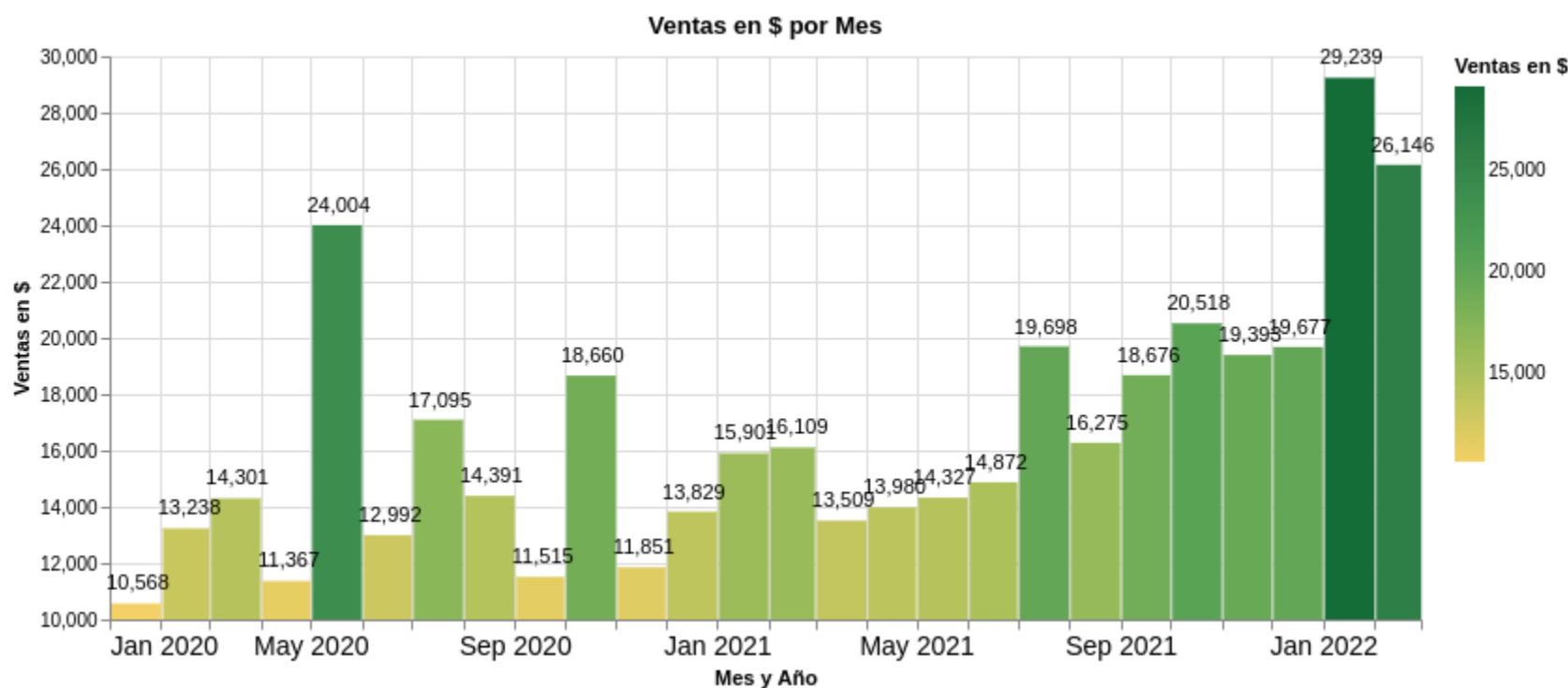
6. De los 3.000 productos en existencia, 236 son considerados **indispensables**, pues registran ventas en al menos el 75% de los meses del período estudiado.
7. De los 236 productos indispensables, 37 podrían entrar en rotura de stock durante marzo del 2022.

2.2 Detalles del Análisis

Resumen General de las Ventas en \$

El período de estudio va de: 2020-01-06 00:00:00 a 2022-02-26 00:00:00
El total de ventas en \$ del período es: 432128.73
El N° de Facturas emitidas durante el período es: 25880
El promedio mensual de ventas en \$ es: 16620.34
El promedio diario es: 706.09

Out[6]:

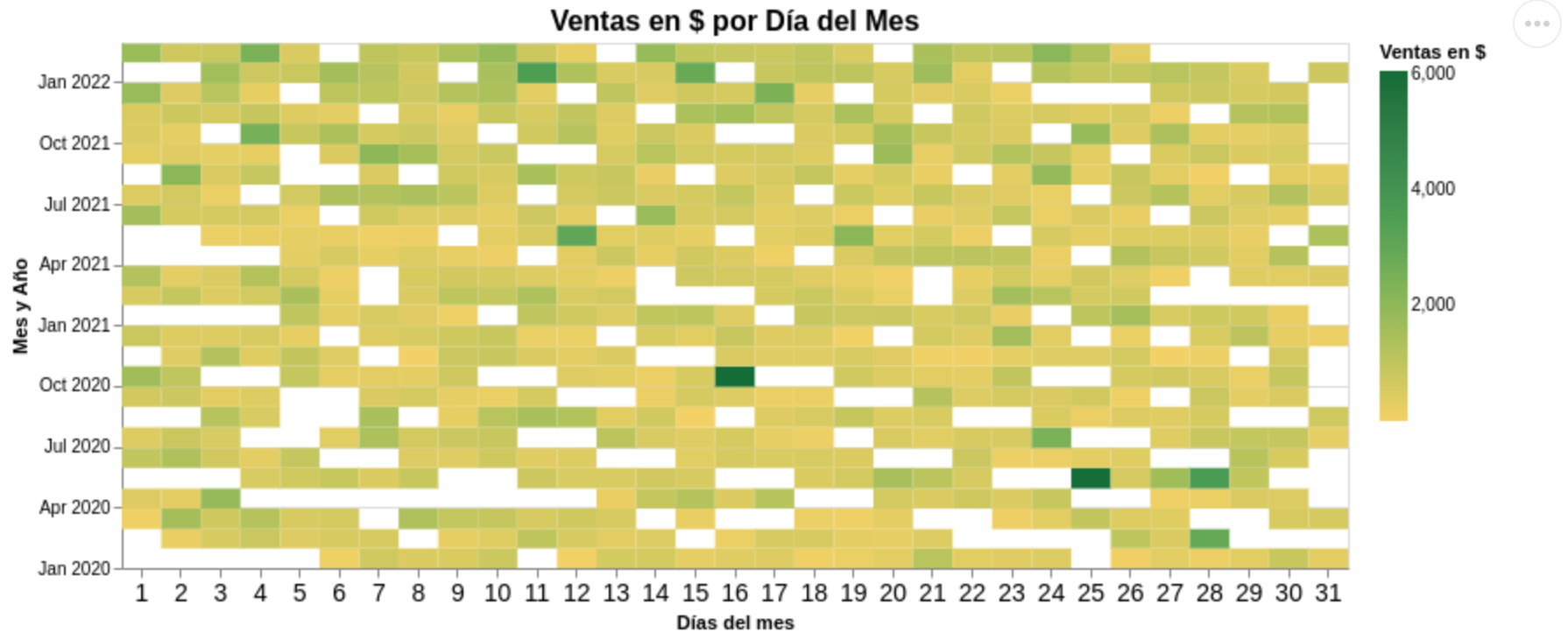


La evolución de las ventas en USD por mes muestra un aumento significativo a partir del 2do semestre del 2021.

Los picos observados en mayo y octubre del 2020 corresponden a ventas excepcionales que se registraron en esos meses.

Es de resaltar que los meses con mayores ventas del período, corresponden a los meses que van del 2022.

Out[20]:



Cada cuadrito corresponde a un día de venta en USD. Mientras más oscuro/verde, mayor las ventas. Los días en blanco corresponden a días sin ventas (fines de semana, feriados, etc.).

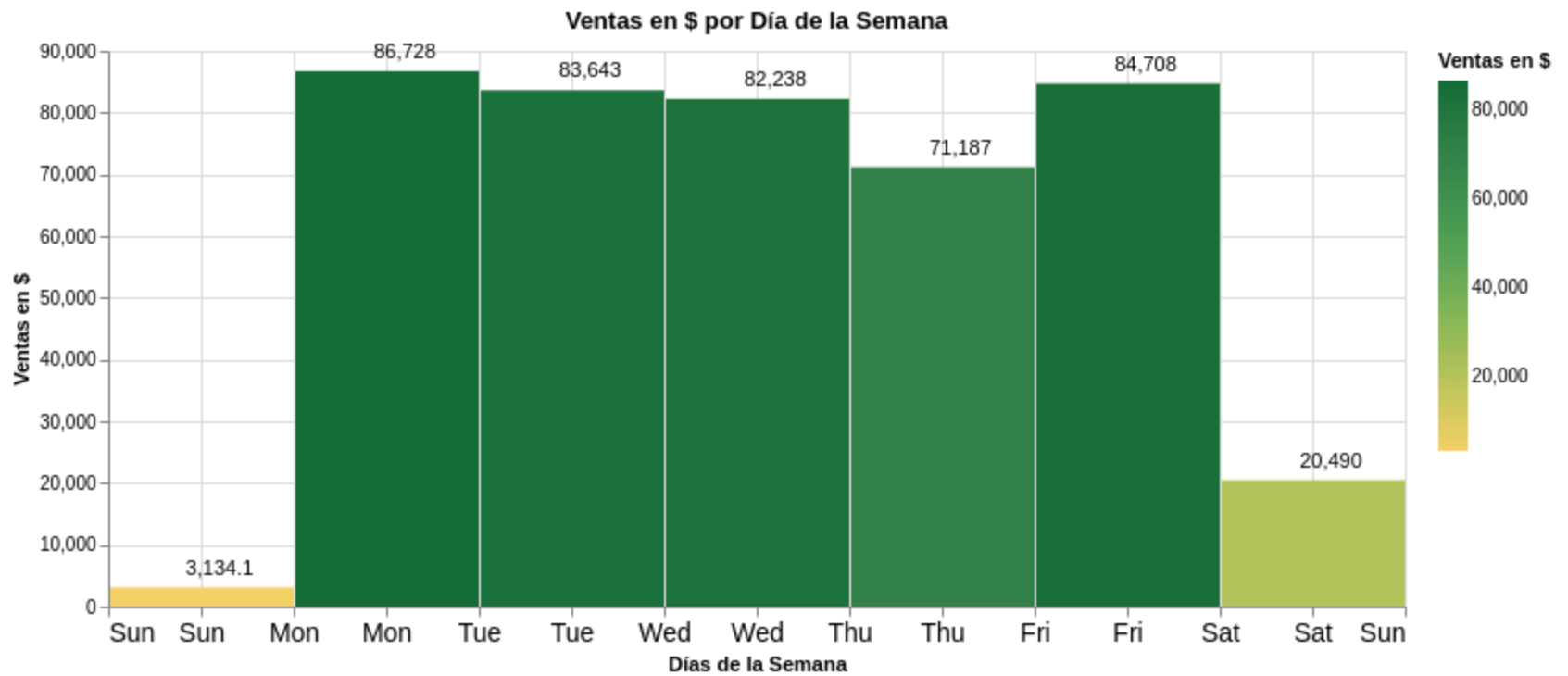
Las bajas ventas los primeros meses del 2020 se explican por los numerosos días sin ventas, debidos a la pandemia y feriados como el carnaval y semana santa.

Los cuadros más oscuro corresponden al 25/05/2020 y el 16/10/2020 donde se registraron ventas excepcionales.

A partir de julio del 2021 comenzamos a ver un poco más de cuadros oscuros, es decir, un aumento en las ventas.

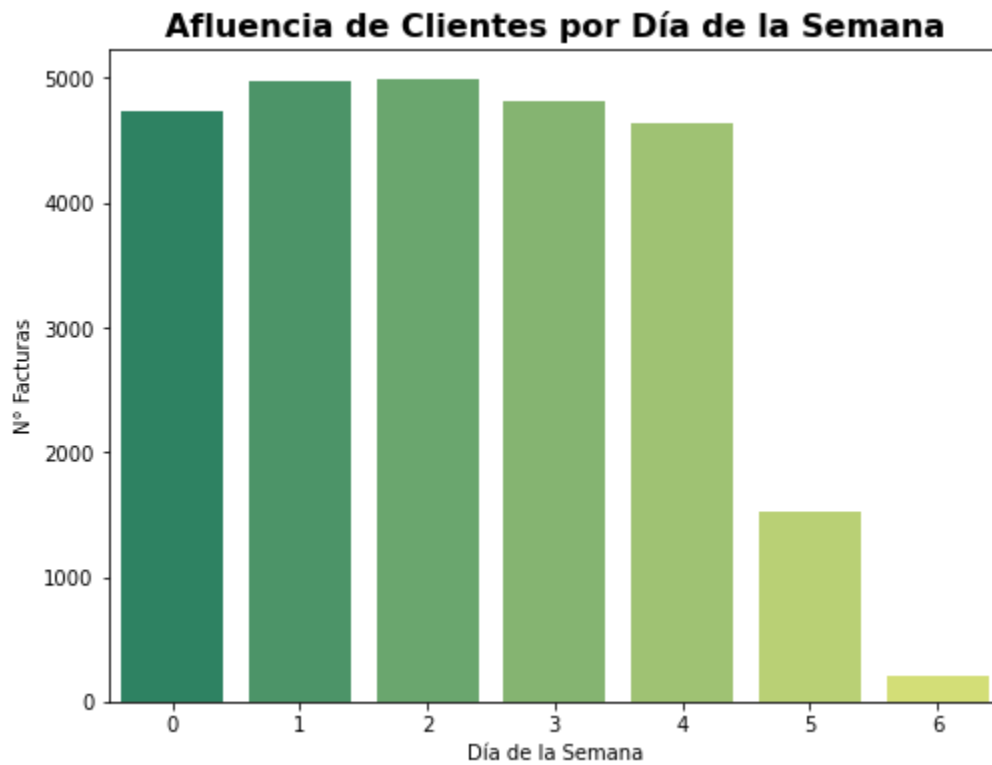
En 2020 podemos observar mayores ventas en la 2da quincena de los meses. Pero a partir de julio del 2021, las ventas se vuelven más uniformes a lo largo de todo el mes.

Out[8]:



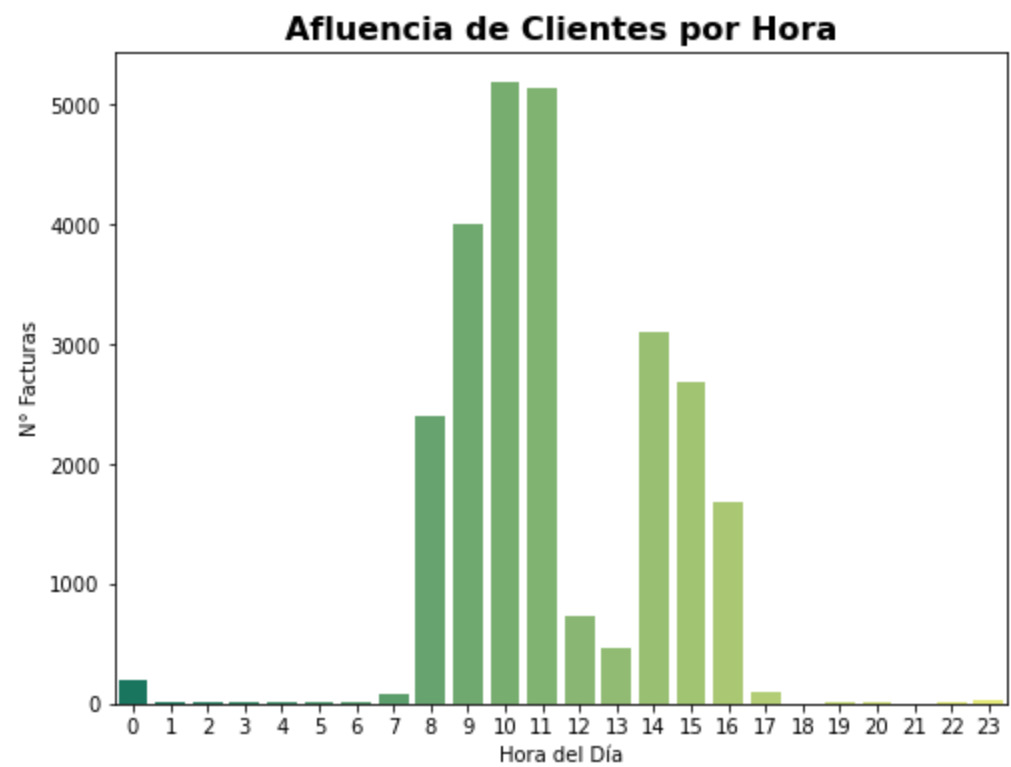
Los lunes y viernes son los días donde observamos mayores ventas. Los sábados es normal tener menos ventas porque sólo se trabaja medio día. Curiosamente el jueves es un día con pocas ventas.

Out[9]: Text(0.5, 0, 'Día de la Semana')

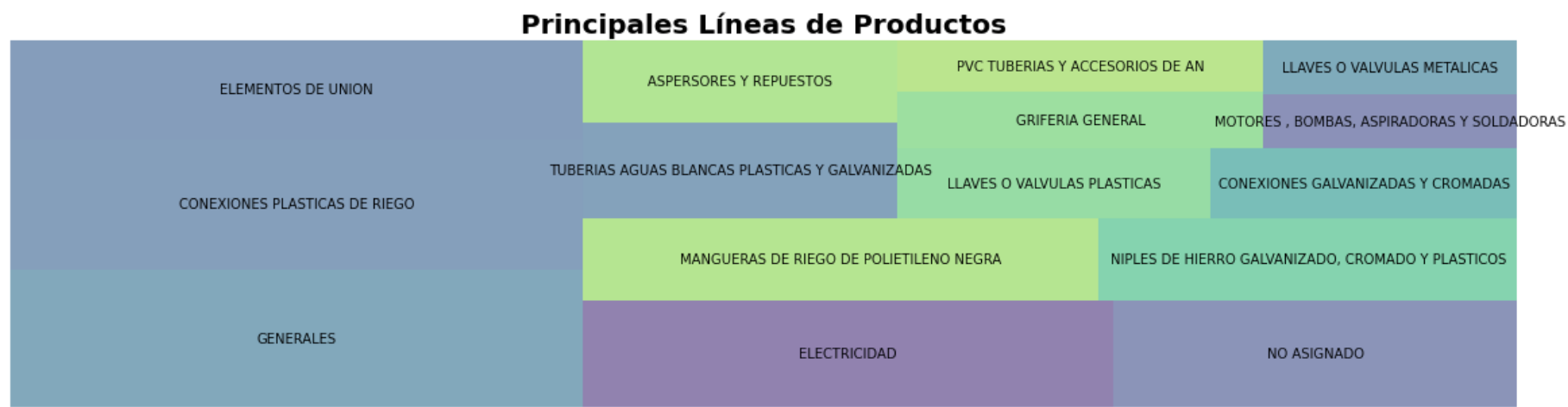


Martes y Miércoles son los días con mayor afluencia de clientes, seguidos del jueves y el lunes. Hacia el fin de semana disminuye la afluencia.

Out[10]: Text(0.5, 0, 'Hora del Día')



En cuanto a las horas, definitivamente las mañanas son más concurridas, en particular de 10am a 12pm.



Estas son las 15 principales líneas de producto (de un total de 34). Son las más importantes porque generan la mayor cantidad de ventas en USD y/o en volumen.

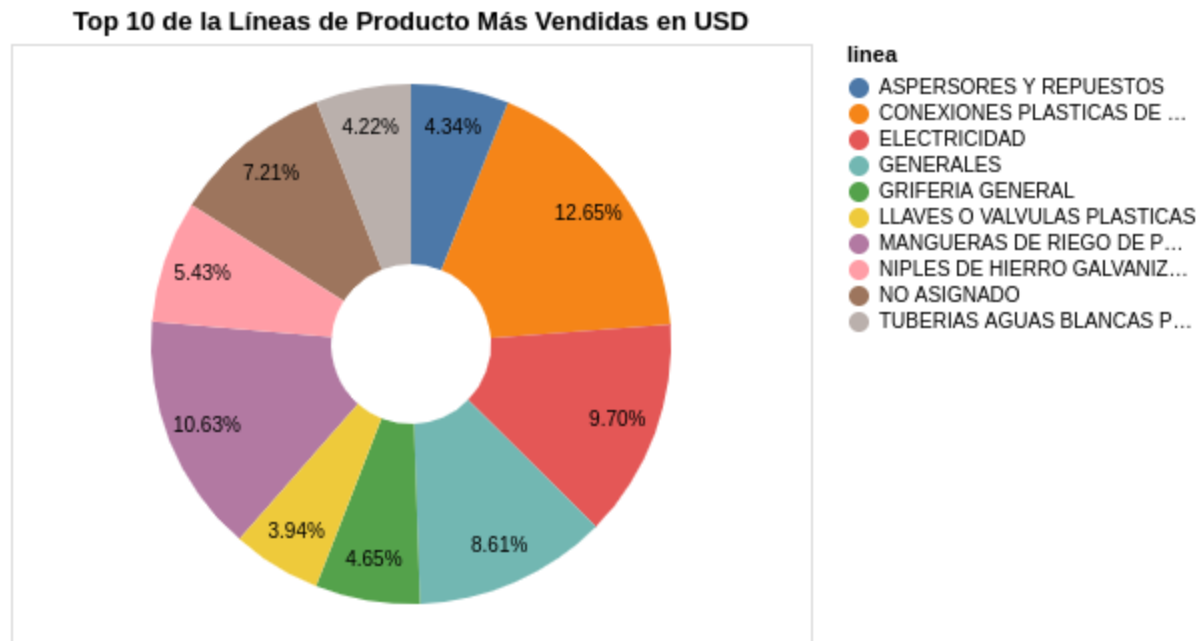
Las líneas:

- Conexiones plásticas de riego
- Mangueras de riego de polietileno
- Aspersores y respuestos
- Motores, bombas, aspiradoras y soldaduras

Incluyen los productos utilizados en los acueductos y sistemas de riego.

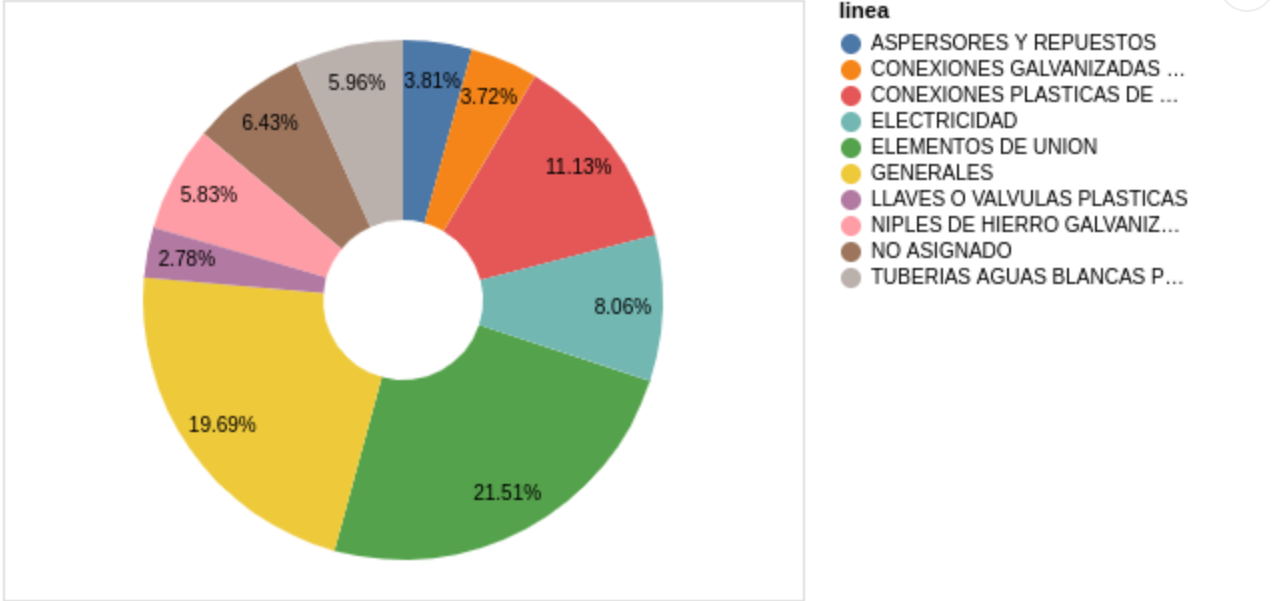
Esto confirma que la empresa es vista por los clientes como especialista en esta materia.

Out[12]:



Out[13]:

Top 10 de la Líneas de Producto Más Vendidas en Volumen



Definimos un **producto indispensable** aquel que se vende regularmente y no puede faltar.

Para identificarlos, calculamos las ventas mensuales de cada producto. Aquellos que tuvieron ventas en al menos 75% de los meses del período (11 meses de 14), los consideramos indispensables.

Son 236 y esta es la lista así como la cantidad por línea de producto:

Lista de Productos Indispensables:
=====

Out[15]:

	cod	producto	linea	stock	fecha_stock	promedio
0	00001	PEGA DE CONTACTO/PEGA ZAPATERA (90 ML) ENVASADO	GENERALES	15.50	2022-03-03	4.0
1	00005	NIPLE PLASTICO 1" * 13CMS	NIPLES DE HIERRO GALVANIZADO, CROMADO Y PLASTICOS	2.00	2022-03-03	39.0
2	00008	ALAMBRE C-17.5 LISO HG KILOGRAMO VICSON	ALAMBRES Y MALLAS	8.75	2022-03-03	15.0
3	00010	LLAVE 1/2" ESFERICA GRINACA COLOMBIA	LLAVES O VALVULAS METALICAS	20.00	2022-03-03	5.0
4	00012	BAILARINA PUNTO NEGRO AGRO APN	ASPERSORES Y REPUESTOS	95.00	2022-03-03	19.0
...
231	06114	CEMENTO BLANCO (POR KG) ARGOS COLOMBIANO	CONSTRUCCIÓN	105.00	2022-03-03	12.0
232	06196	ABRAZADERA METALICA 3/8" 6MM-13MM	ELEMENTOS DE UNION	91.00	2022-03-03	12.0
233	06300	CURVA 3/4" NEGRA POLIETILENO ELECTRICIDAD ECON...	GENERALES	239.00	2022-03-03	10.0
234	06306	REGADOR 1/2" MINI WLOBBER MINI BAILARINA DIAME...	ASPERSORES Y REPUESTOS	76.00	2022-03-03	36.0
235	06315	TUBO PLASTICO 1/2"*6MTS AZUL SHEDULE INPLAVEN	TUBERIAS AGUAS BLANCAS PLASTICAS Y GALVANIZADAS	67.50	2022-03-03	8.0

236 rows × 6 columns

Lista de Líneas de Productos y la cantidad de Productos Indispensables que poseen:
=====

```

Out[17]: linea
CONEXIONES PLASTICAS DE RIEGO          49
GENERALES                               24
ELEMENTOS DE UNION                       24
NIPLES DE HIERRO GALVANIZADO, CROMADO Y PLASTICOS 19
ELECTRICIDAD                            17
CONEXIONES GALVANIZADAS Y CROMADAS      17
PVC TUBERIAS Y ACCESORIOS DE AN         15
GRIFERIA GENERAL                         11
TUBERIAS AGUAS BLANCAS PLASTICAS Y GALVANIZADAS 11
NO ASIGNADO                              8
DISCOS                                   5
ASPERSORES Y REPUESTOS                   5
LLAVES O VALVULAS METALICAS              5
LLAVES O VALVULAS PLASTICAS              5
CHUCHERIAS                              4
MECHAS, BROCAS, SIERRA COPA              3
SEGURIDAD & HIGIENE                      2
HERRAMIENTAS MANUALES                   2
MANGUERAS GENERALES NO DE POLIETILENO    2
MANGUERAS DE RIEGO DE POLIETILENO NEGRA  2
CONSTRUCCIÓN                             2
CABULLAS                                 2
TUBERIA DE COBRE Y CONEXIONES DE BRONCE  1
ALAMBRES Y MALLAS                        1
Name: cod, dtype: int64

```

Lista de Productos Indispensables que podrían entrar en Rotura de Stock:

=====

Esta es la lista de los productos indispensables que podrían entrar en **rotura de stock** durante el mes de marzo 2022, es decir, la cantidad en stock, es menor al promedio mensual de ventas:

```
/home/diana/.local/lib/python3.8/site-packages/pandas/core/frame.py:3607: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy (https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
self._set_item(key, value)
```

Out[19]:

	cod	producto	promedio	stock	faltan
163	02584	LLAVE PLASTICA 1/2" ROSCADA	62.0	0.00	62.00
154	02282	TORNILLO DRYWALL N°8*2"	83.0	39.00	44.00
1	00005	NIPLE PLASTICO 1" * 13CMS	39.0	2.00	37.00
39	00312	RIEGO ABRAZADERA PLASTICA 40MM * 3/4" AGRO/ F...	66.0	36.00	30.00
100	00901	NIPLE PLASTICO 3/4" * 70 CTMS AZUL/GRIS CLASE A	40.0	19.00	21.00
60	00468	CABILLA 3/8" 10MM * 6 MTRS ESTRIADA (POR UNIDAD)	32.0	12.00	20.00
140	01697	NIPLE PLASTICO 1/2" * 7.5CTMS/3"	13.0	0.00	13.00
160	02494	ABRAZADERA METALICA 17-18MM TIPO TORNILLO SIN FIN	8.0	-4.00	12.00
120	01232	BUSINHG PLASTICO REDUCIDO 3/4" * 1/2" AZUL/BL...	17.0	6.00	11.00
200	04129	CONFITERIA CARAMELO CHAO/ LOKIÑO/ MIST	55.0	44.00	11.00
175	03335	PLASTICO NEGRO CONTINUO 6MTS ANCHO 150 MICRAS(...	10.0	0.00	10.00
134	01593	CAJETIN METALICO RECTANGULAR 2*4" 1/2" REFORZADO	11.0	2.00	9.00
144	01826	NIPLE PLASTICO 1 1/2"*15CM SCHEDULE 80 URAPLAST	26.0	18.00	8.00
32	00247	DISCO 4 1/2" CORTE METAL EXTRA FINO DEWALT DW8...	7.0	0.00	7.00
59	00463	RIEGO ADAPTADOR MACHO 20 MM * 1/2" AGRO/MOM	23.0	16.00	7.00
2	00008	ALAMBRE C-17.5 LISO HG KILOGRAMO VICSON	15.0	8.75	6.25
113	01153	TOMA DOBLE 270 2*15A POLARIZADA CON TAPA BLANCO	12.0	6.00	6.00
188	03809	R. MANILLA METALICA AMERICANA 8" SAM	5.0	-1.00	6.00
93	00802	TEIPE ELECTRICO 3/4"*10MTS NEGRO COBRA ORIGINAL	6.0	0.00	6.00
43	00348	TEE HG 1/2"	19.0	14.00	5.00
178	03454	LLAVE CHORRO 1/2" SAM ROSCADA/LISA	7.0	2.00	5.00
195	04040	TAPA HUECO PARA FREGADERO CROMADO CON TORNILLO...	4.0	0.00	4.00
227	05942	MECHA 3/16" HSS BOSCH	3.0	0.00	3.00

	cod	producto	promedio	stock	faltan
224	05135	GRIFERIA LAVAMANOS LLAVE SAM MANILLA AMERICANA...	3.0	0.00	3.00
218	04792	BROCHA 1" AZUL CEBRA	3.0	0.00	3.00
210	04404	NIPLE PLASTICO 1/2" * 10CTMS/4" PVC AZUL	9.0	6.00	3.00
185	03653	CONECTOR PARA CABLE COAXIAL ROJO 1*100 RG6	5.0	2.00	3.00
26	00198	CLAVO 2 1/2" x 11 (POR KILO)	3.0	0.00	3.00
121	01257	NIPLE PLASTICO 2" * 15 CTMS URAPLAST	14.0	12.00	2.00
189	03810	LLAVE CHORRO 1/2" BRONCE GRINACA CON CONEXION ...	10.0	8.00	2.00
61	00487	CLAVO 3" x 10 (POR KILO)	2.0	0.00	2.00
145	01838	RAMPLUG PLASTICO GRIS 1/2" * 2-1/4"	21.0	19.00	2.00
226	05941	ABRAZADERA METALICA 8MM-16MM	7.0	5.00	2.00
228	05964	A.C TUBO 1/2" * 3MTS DE PEGAR TUBRICA	5.0	3.60	1.40
75	00615	REGULADOR GAS TIPO POLL ENSAMBLADO MARIPOSA CA...	3.0	2.00	1.00
96	00858	HOJA PARA SEGUETA 18 DIENTES METAL GRUESO LENO...	10.0	9.00	1.00
37	00291	ANILLO HG 1"	11.0	10.00	1.00

2.3 Conclusiones y recomendaciones:

1. Las ventas en USD tienden a aumentar con el tiempo. Esto puede ser una buena noticia, sin embargo hay que ser prudentes con los imprevistos y eventos que podrían cambiar la tendencia.
2. Para mejorar la calidad de los datos, recomendamos verificar las líneas de producto **GENERALES** y **NO ASIGNADOS** para clasificar mejor los productos.
3. Es recomendable garantizar la presencia y disponibilidad de todo el personal los martes y miércoles y en particular todos los días de 10am a 12pm, ya que son los momentos de mayor afluencia de clientes.
4. Aunque los jueves sea un día de buena afluencia de clientes, las ventas en USD son bajas. Podrían crearse incentivos como promociones especiales ese día.
5. Faltan datos para determinar los productos que no se venden, específicamente la fecha de la última entrada en stock. Una vez identificados estos productos, debemos crear estrategias para venderlos.
6. La identificación de **productos indispensables** y en particular saber si pueden entrar en rotura de stock, nos permitirá priorizar las compras y aumentar la satisfacción de los clientes al tener estos productos siempre disponibles.
7. Más adelante, las predicciones sobre las ventas nos podrán ayudar a mejor planificar el presupuesto y las compras.

Para más detalles del análisis consultar: [NB2: EDA y Visualización de los Datos \(https://github.com/dchaconoca/proyecto-ventas/blob/master/NB2_analisis_ventas_eda.ipynb\)](https://github.com/dchaconoca/proyecto-ventas/blob/master/NB2_analisis_ventas_eda.ipynb).

O también el [Reporte interactivo en Google Data Studio \(https://datastudio.google.com/reporting/a4e9910d-447d-417a-80a9-a1e5ae4a4b22\)](https://datastudio.google.com/reporting/a4e9910d-447d-417a-80a9-a1e5ae4a4b22).

