

Extracción de Características.

Exploración, selección y limpieza.

La base de datos proporcionada contaba con 171 registros de 106 columnas que almacenaban datos de tipo flotante, entero y objeto respectivamente. Dado el enfoque del análisis se seleccionaron 20 variables cualitativas de las disponibles con la finalidad de extraer características informativas de la muestra recolectada.

Una vez seleccionadas estas columnas se limpiaron los valores nulos mediante la sustitución de estos por la cadena “Unknown”, haciendo referencia al desconocimiento de esos datos. Finalmente, al tratarse de variables cualitativas no fue necesaria la limpieza de valores atípicos.

Hallazgos por variable.

1. `_updated_by`

Si bien 31 personas participaron en el registro de la base de datos, el 51.46% de los registros se tienen gracias a 6 de ellos, siendo Salomon Abdala Perez el principal aportante con un total de 20 registros que representan el 11.69% de los registros totales, como se puede ver en la Tabla 1

	<code>_updated_by</code>	<code>frequency</code>	<code>percentage</code>	<code>cumulative_perc</code>
0	Salomon Abdala Perez	20	0.116959	0.116959
1	Ángel Francisco Altamirano Campos	16	0.093567	0.210526
2	José Zapata	14	0.081871	0.292398
3	Emilio Garza	13	0.076023	0.368421
4	Benjamin Barbosa	13	0.076023	0.444444
5	Pedro Contreras	12	0.070175	0.514620

Tabla 1

2. `232_type_of_store`

Esta columna constaba originalmente de 64 categorías; sin embargo, se seleccionaron solo aquellos registros que aparecían mínimo 5 veces dentro del dataframe. Estos datos se visualizan en la Ilustración 1 con las cinco categorías más mencionadas dependiendo de su frecuencia. Los tipos de tienda más frecuentes en orden fueron Grocery store, fonditas, tiendas que ofrecen en su mayoría productos lácteos, cafeterías y carnicerías.

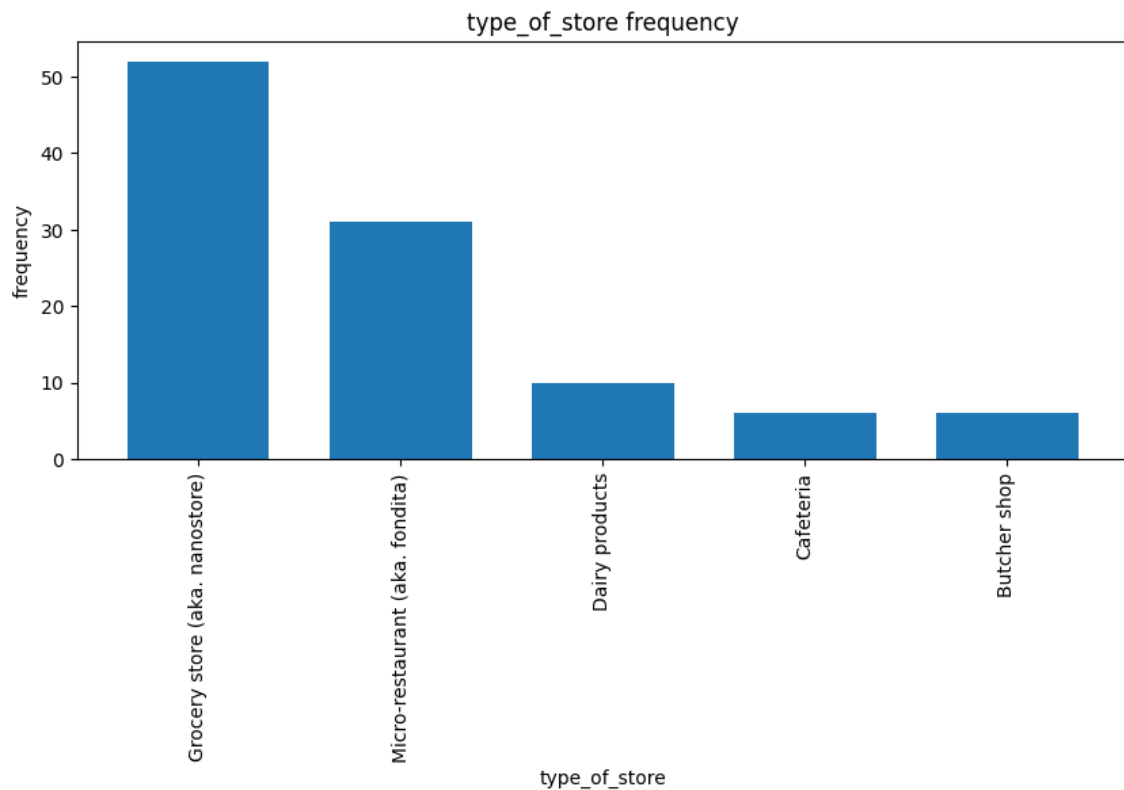


Ilustración 1

3. 108_does_the_micro_retailer_has_a_barred_window_

Esta variable no presentó datos nulos y, a pesar de ser tipo cadena, su interpretación es booleana. Se interpreta del nombre de la variable que esta columna clasifica los registros dependiendo de si el micro_retailer tiene rejas en sus ventanas. La Ilustración 2 presenta las proporciones de la respuesta, donde se observa que solo el 27.5% de los encuestados afirman tener esta medida de seguridad.

does_the_micro_retailer_has_a_barred_window

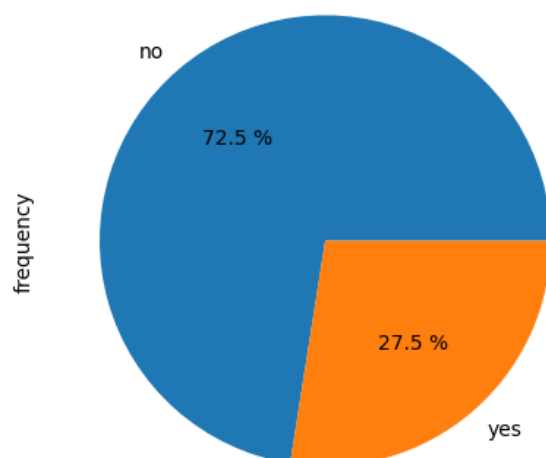


Ilustración 2

4. 99_does_the_micro_retailer_exhibits_products_outside_

Esta variable no presentó datos nulos y, a pesar de ser tipo cadena, su interpretación es booleana. Se interpreta del nombre de la variable que esta columna clasifica los registros dependiendo de si el `micro_retailer` exhibe productos al exterior de su negocio. La Ilustración 3 presenta las proporciones de la respuesta, donde se observa el 70.2% de los encuestados prefiere exhibir sus productos exclusivamente dentro de su establecimiento, posiblemente por los niveles de inseguridad reportados en los últimos años en la zona.

`does_the_micro_retailer_exhibits_products_outside`

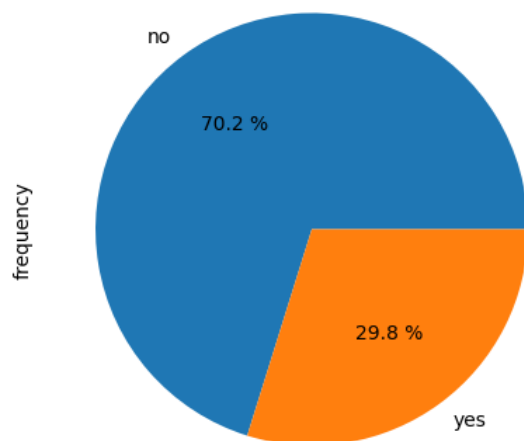


Ilustración 3

5. `102_does_the_micro_retailer_sells_fresh_products_`

Esta variable no presentó datos nulos y, a pesar de ser tipo cadena, su interpretación es booleana. Se interpreta del nombre de la variable que esta columna clasifica los registros dependiendo de si el `micro_retailer` vende productos frescos. La Ilustración 4 describe el comportamiento de las proporciones de las respuestas recopiladas. Un hallazgo relevante resulta al comparar los resultados del análisis de esta variable con los obtenidos de la variable 2 (`232_type_of_store`). 61.4% de los encuestados negaron vender productos frescos cuando en el análisis de la variable 2 se hallaron los tipos de tiendas más frecuentes, donde se esperaría que al menos el top 5 presentado vendiera productos frescos.

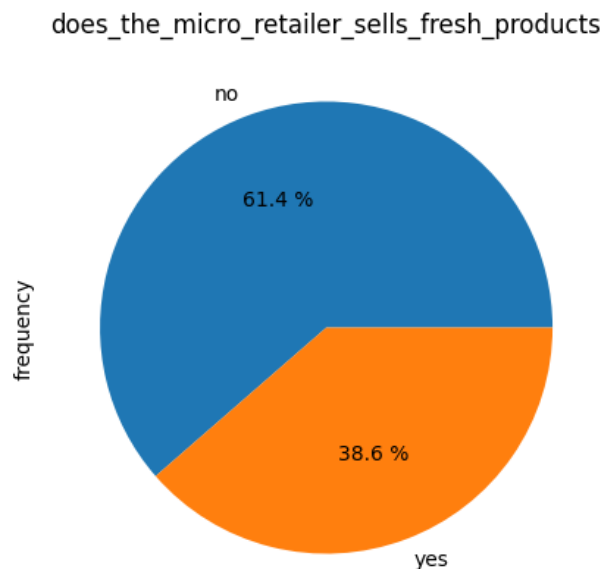


Ilustración 4

6. 28_prefered_payment_method

Esta variable presenta inicialmente un 98.83% de datos nulos, por lo que se descarta su relevancia dentro del análisis dado el alto porcentaje de información faltante. No obstante, la Tabla 2 presenta la frecuencia de sus categorías con fines informativos donde solo 1 micro_retailer prefiere las tarjetas de débito y/o crédito como método de pago y otro el efectivo.

	28_prefered_payment_method	frequency	percentage	cumulative_perc
0	Unknown	169	0.988304	0.988304
1	Tarjeta de débito y/o crédito	1	0.005848	0.994152
2	Efectivo	1	0.005848	1.000000

Tabla 2

7. 30_payment_methods_large_purchases

Al igual que la variable anterior, esta columna presentó un alto porcentaje de valores faltantes, siendo este de 99.41%. Se descarta entonces la relevancia de su aporte al análisis. Sin embargo, la Tabla 3 presenta la frecuencia de sus categorías con fines informativos donde solo 1 micro_retailer prefiere las tarjetas de débito y/o crédito como método de pago de grandes compras.

	30_payment_methods_large_purchases	frequency	percentage	cumulative_perc
0	Unknown	170	0.994152	0.994152
1	Tarjeta de débito y/o crédito	1	0.005848	1.000000

Tabla 3

8. 214_customer_relationship_tools

Esta columna contaba inicialmente con 14 categorías; sin embargo, si seleccionaron únicamente aquellas con una frecuencia de aparición dentro del dataframe mayor a 5 para poder identificar las herramientas más populares para gestionar la relación con el cliente entre los micro_retailers. Si bien se desconoce dicha herramienta en el 44.44% de los casos, las herramientas más populares son Redes sociales, aplicación móvil, ninguna, ambas o cuaderno. Esto se visualiza en la Ilustración 5.

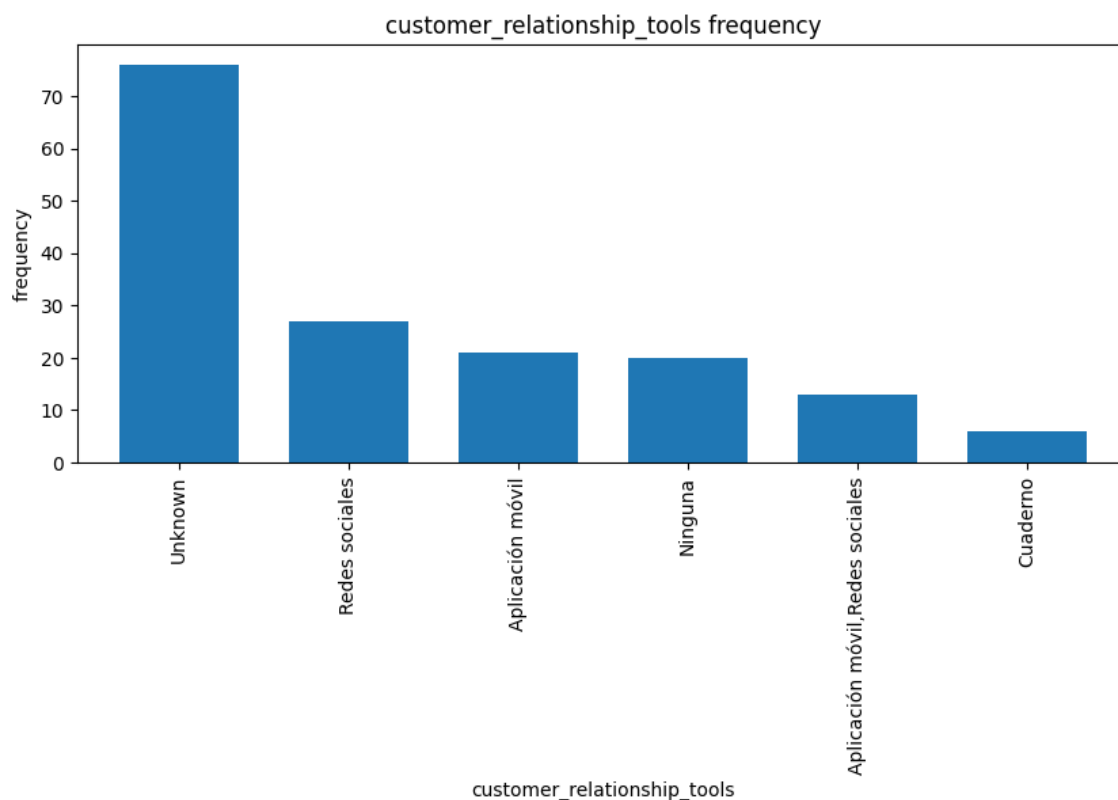


Ilustración 5

9. 33_credit_to_customers

En cuanto al ofrecimiento de crédito a sus clientes, el 46.78% solo admite pagos de contado, se desconoce el criterio del 37.43% y solo el 15.79% cuenta con esta opción de pago para sus clientes. Este hallazgo se visualiza en la Ilustración 6.

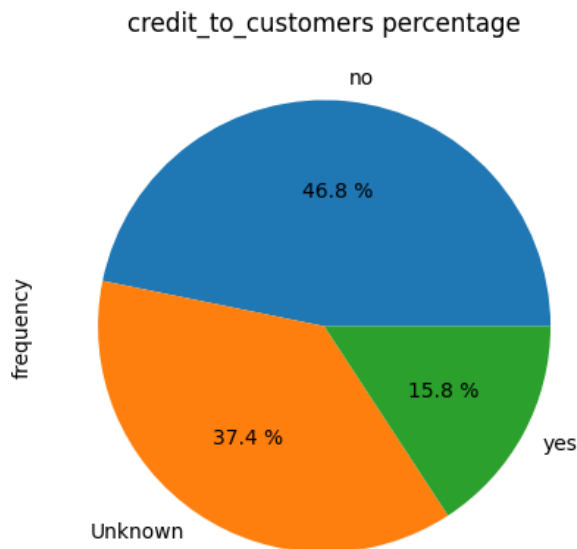


Ilustración 6

10. 193_sales_planning_tools

Esta variable contaba inicialmente con 18 categorías únicas; algunas de ellas constaban de la mezcla de otras, lo que indica que es cualidad fue recopilada como pregunta de selección múltiple. Para fines de este análisis se redujeron las categorías a solo aquellas que tuvieran una frecuencia mayor a 5 registros. Dando como resultado la gráfica presentada en este reporte de Ilustración 7 donde se puede ver que se desconoce la respuesta para la mayoría de los micro_retailers; lo siguiente más popular dentro de las herramientas para planeación de ventas es no utilizar ninguna herramienta para estos fines, seguido de cuaderno, software especializado no especificado y Excel.

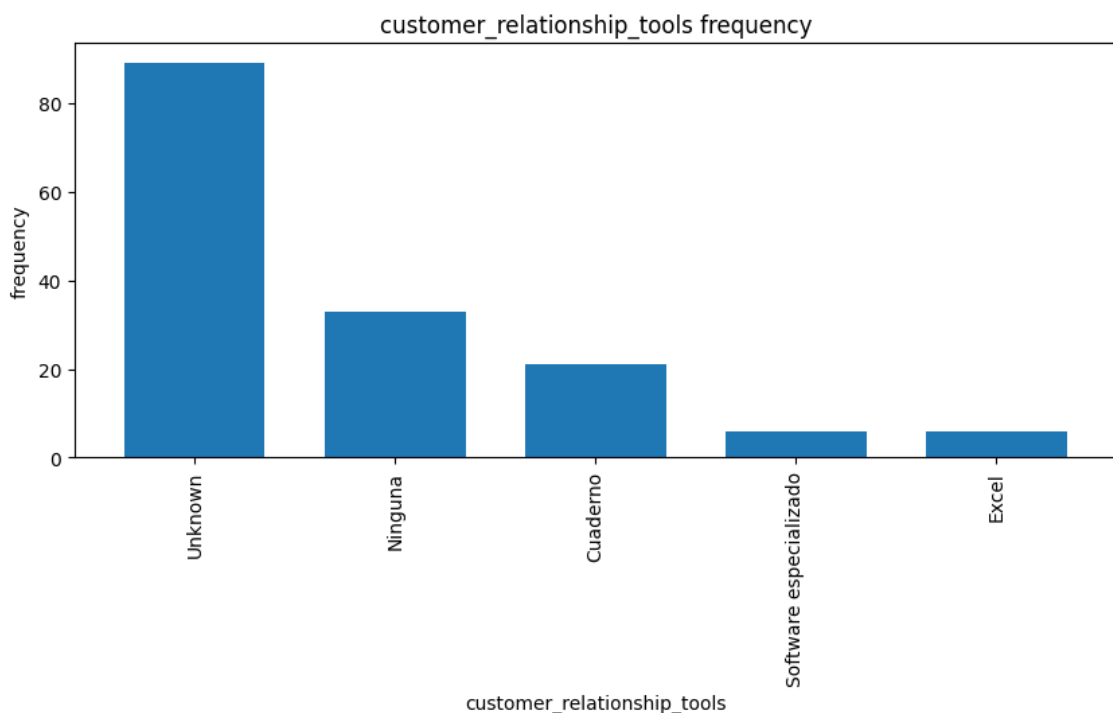


Ilustración 7

11. 64_comission_services

Para el análisis de esta variable cualitativa se seleccionaron aquellos valores únicos con una frecuencia mayor o igual a 5. Después de esta selección, se pudo identificar que se desconoce la tasa de comisión por servicios del 92.5% de los micro_retailers encuestados, el 4.4% dice aplicar una tasa del 1% y el 3.1% una tasa del 5%. Este hallazgo es representado visualmente en la Ilustración 8.

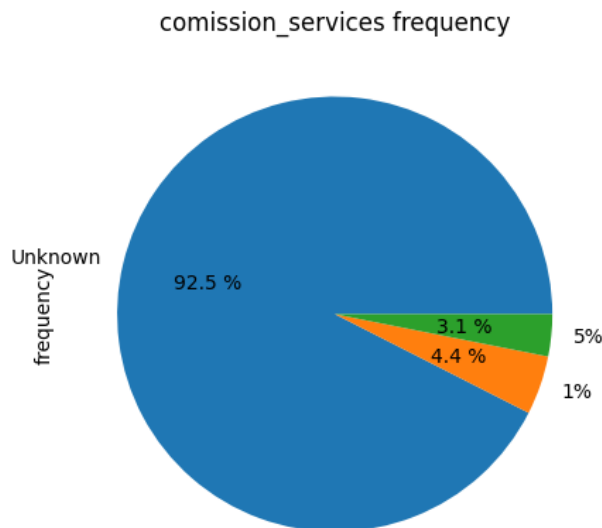


Ilustración 8

12. 313_home_deliveries

Esta variable recopila la respuesta de una pregunta cerrada de opción múltiple donde solo se puede seleccionar una respuesta, además el planteamiento de las posibles opciones aplica una escala de Likert (da Silva, 2023). Se desconoce el criterio del 66.08% de los micro_retailers, pero dentro de los encuestados de los que se conoce la respuesta, la mayoría está totalmente en desacuerdo. Esto se logra observar en la Ilustración 9.

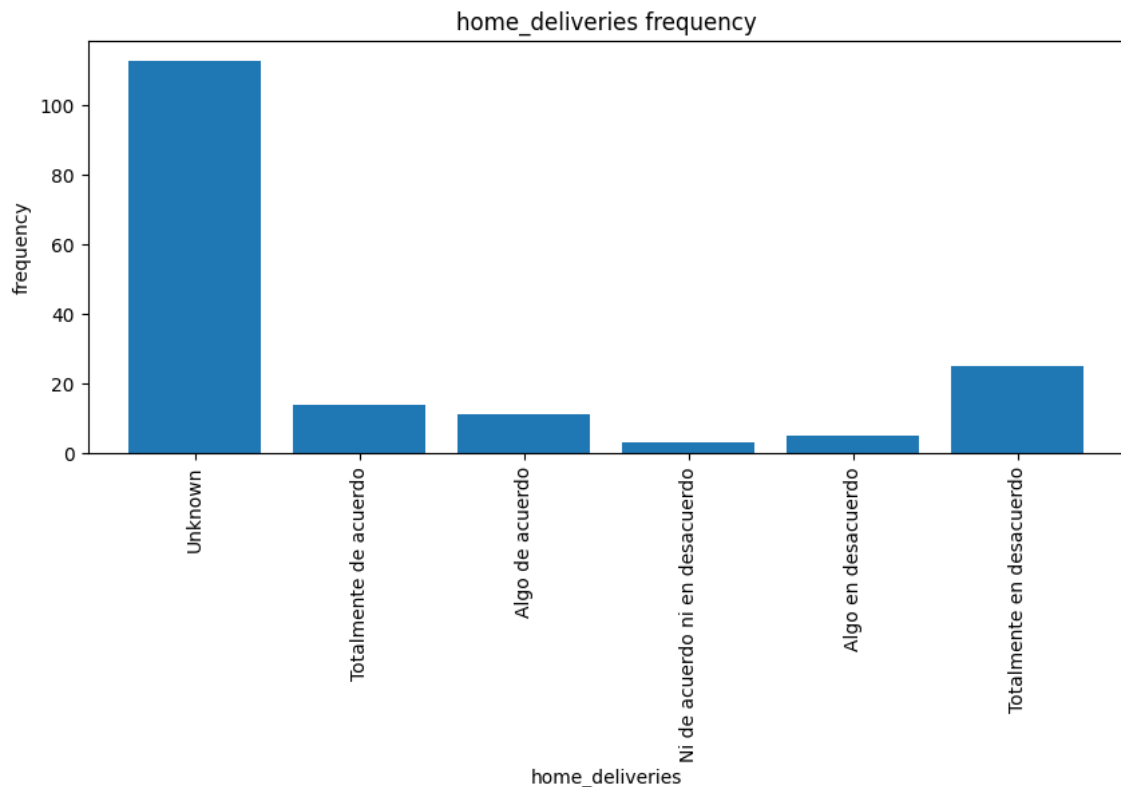


Ilustración 9

13. 185_place_orders_suppliers

Esta columna contaba inicialmente con 100 valores nulos y 20 categorías diferentes. Una vez visualizada su tabla de frecuencia, se infirió que la pregunta se había recopilado como opción múltiple. Se seleccionaron solo las respuestas con una frecuencia mayor a 5 y sin separar por opción, se analizó su frecuencia relativa. Se desconoce el este canal en el 70.9% de los casos. El 14.2% de los micro_retailers realiza este proceso en persona y 5.7% por mensajería instantánea (por ejemplo, por WhatsApp). Esto se aprecia visualmente en la Ilustración 10.

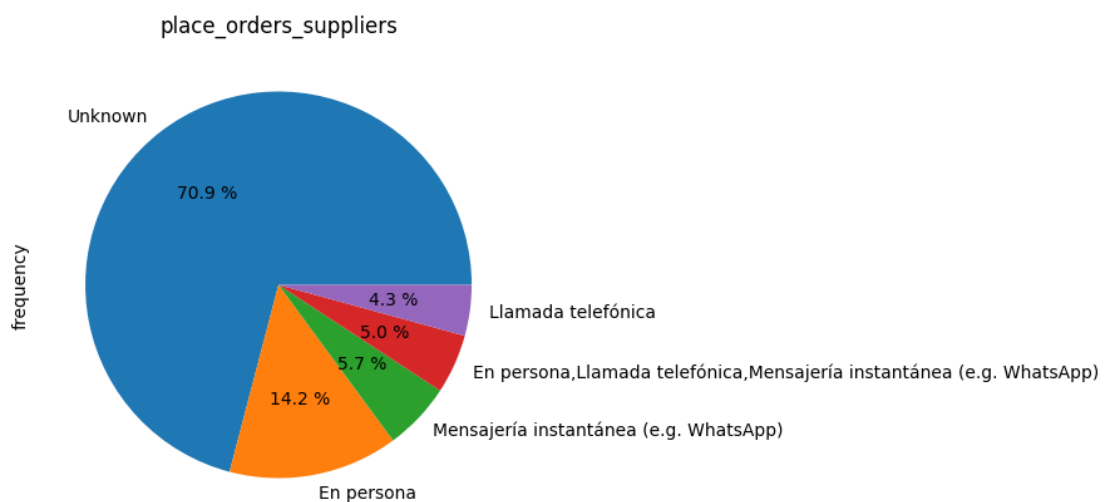


Ilustración 10

14. 53_special_conditions_suppliers

Si bien inicialmente se contaba con 21 categorías únicas para esta columna, se seleccionaron solo aquellas que representaban más de 5% de los casos. Esto redujo las categorías a tres: desconocido, crédito y descuento por compras anticipadas. Desafortunadamente la tasa de datos nulos era alta (72.51%) lo que reduce la relevancia de los resultados obtenidos. En la Ilustración 11 se observa un pico en la frecuencia de la categoría Unknown y una caída para el resto de las categorías. Las dos restantes comparten frecuencia y esto se aprecia en la ausencia de inclinación de una categoría a otra.

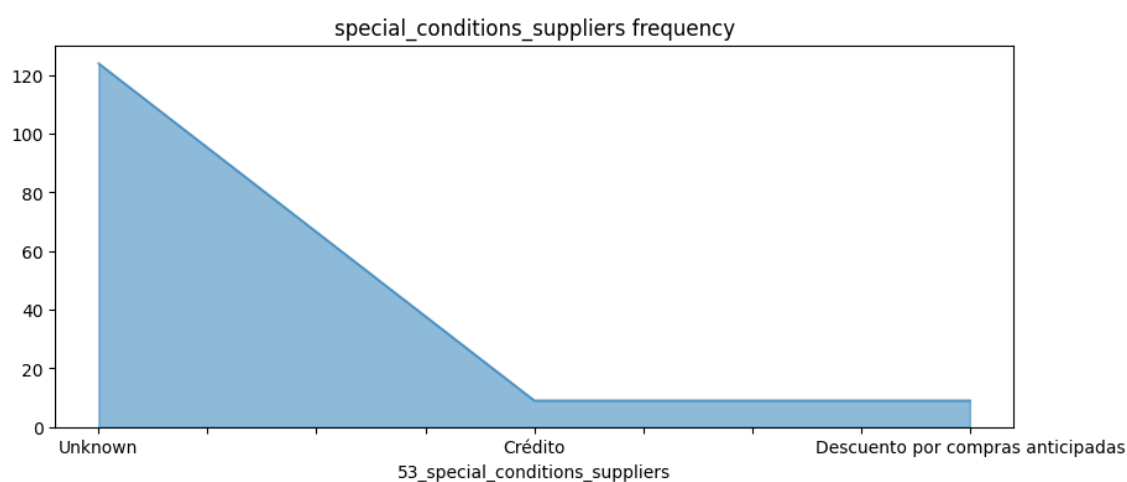


Ilustración 11

15. 192_procurement_planning_tools

El registro de las herramientas empleadas para buscar proveedores y comprar bienes o servicios de esos proveedores incluyendo negociación y licitación, gestión de contratos y gestión de gastos comerciales (Kushner, 2021) contaba inicialmente con 12 categorías; sin embargo, se seleccionaron solo aquellas con una frecuencia mayor a 5. Si bien se desconocen las herramientas utilizadas para estos fines en el 65.2% de los casos, se sabe que un 15.8% utiliza un cuaderno, 15.2% afirma no tener una herramienta específica para estos procesos y un 3.8% utiliza Excel. Esto se puede observar en la Ilustración 12.

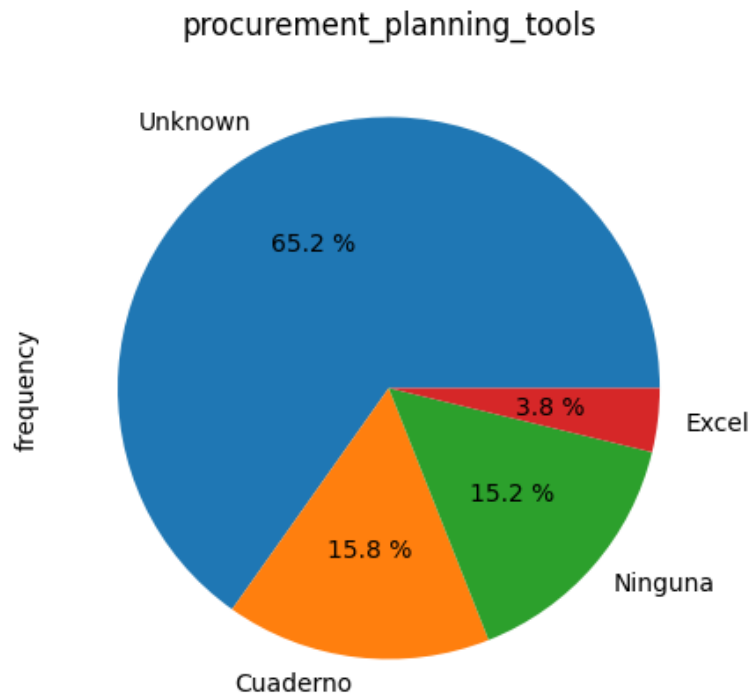


Ilustración 12

16. 277_payment_method_suppliers

Para analizar los métodos de pago a proveedores se identificaron inicialmente 7 categorías únicas dentro de la variable. Una vez observadas se identificó que la pregunta había sido planteada en un formato de selección múltiple por lo que primero se agruparon individualmente las categorías. Si bien esto no es óptimo porque los datos pueden cambiar por archivo, se hizo por beneficio de la visualización. Una vez agrupadas se identificó que la mayoría de los micro_retailers no dieron una respuesta, pero lo más popular es pagar en efectivo, seguido de transferencia bancaria y tarjeta de débito y/o crédito. Este hallazgo es representado por la Ilustración 13.

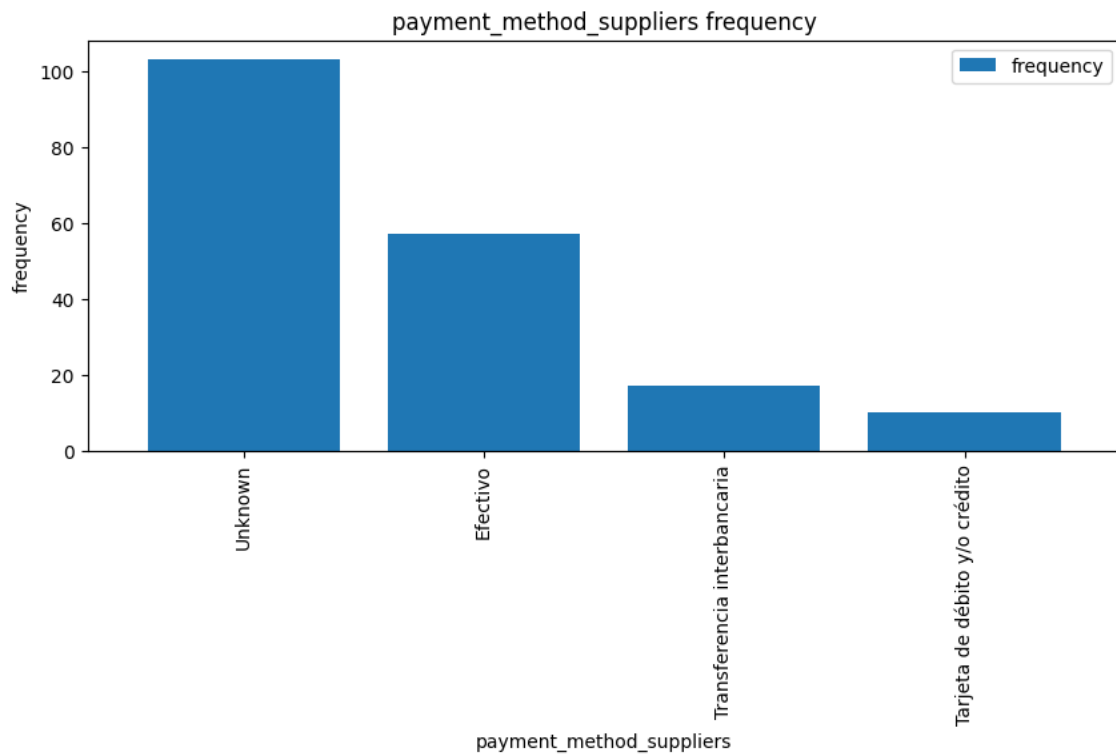


Ilustración 13

17. 56_procurement_changes_pandemic

Una vez comprendido el término procurement, se cuestionó si estos procesos habían tenido algún cambio como consecuencia de la pandemia a lo que el 52.63% omitió su respuesta, 32.16% comentaron que no habían implementado ningún cambio y solo el 15.2% afirmó haber modificado estos procesos en su negocio. Estos hallazgos son visualizados en la Ilustración 14.

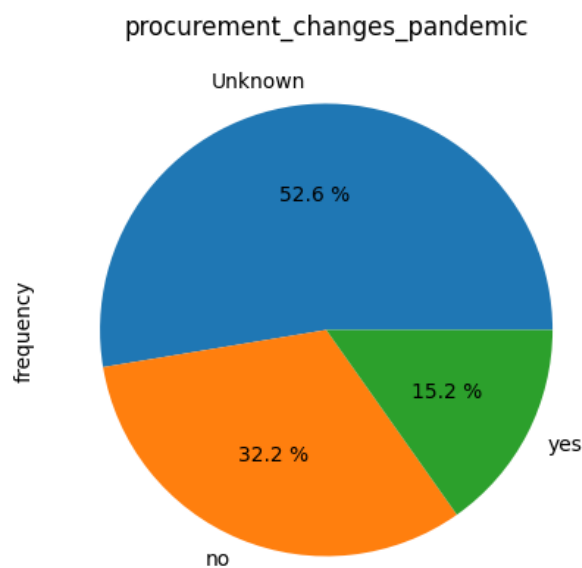


Ilustración 14

18. 157_frequency_organize_shelves

Esta variable contaba con 9 valores únicos incluyendo el reemplazo de nulos por la cadena “Unknown” realizado en un inicio. En este caso el cuestionamiento realizado a los encuestados fue en relación con la frecuencia en la que estos ordenan las repisas presentes en su establecimiento. Ignorando el hecho de que se desconoce esta frecuencia en el 43.86% de los casos, la frecuencia de organización de repisas más popular es semana tras semana, seguida de día tras día y nunca. Esta frecuencia se visualiza en la Ilustración 15.

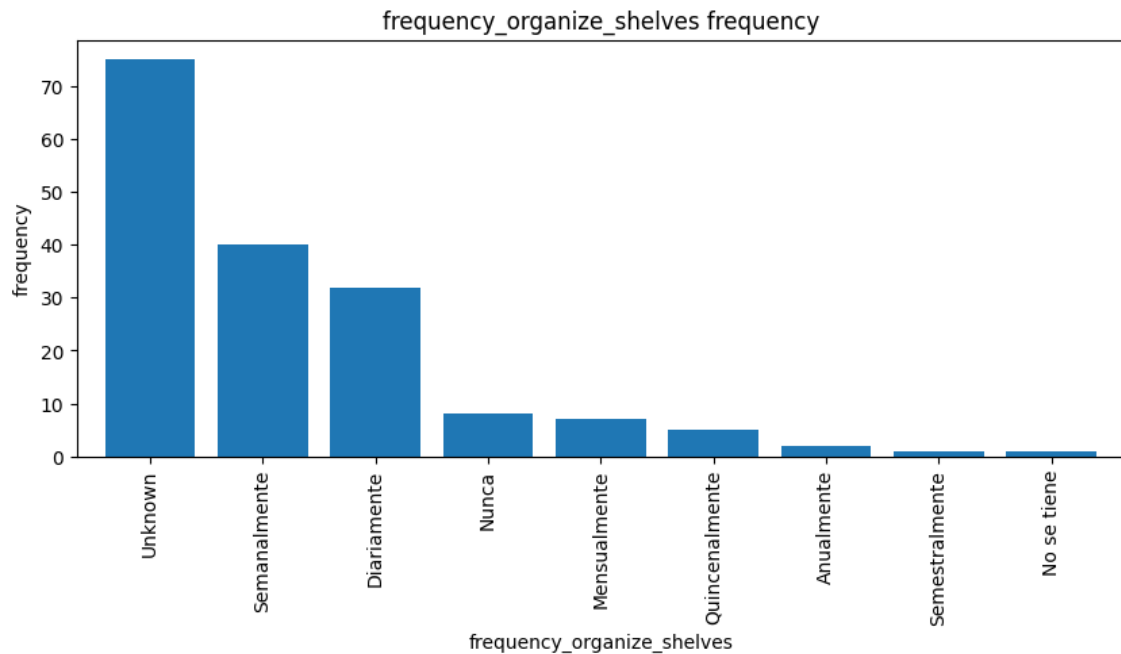


Ilustración 15

19. 158_charact_organize_products_on_shelves

Una vez registrada la frecuencia con la que los micro_retailers organizan sus repisas, se les cuestionó respecto al criterio que emplean para realizar este proceso. Se desconoce el criterio de 46.2% de los casos, pero la mayoría organiza sus repisas por categoría, frecuencia de venta, tamaño, margen de utilidad, precio y refrigeración. Esto se visualiza en la Ilustración 16.

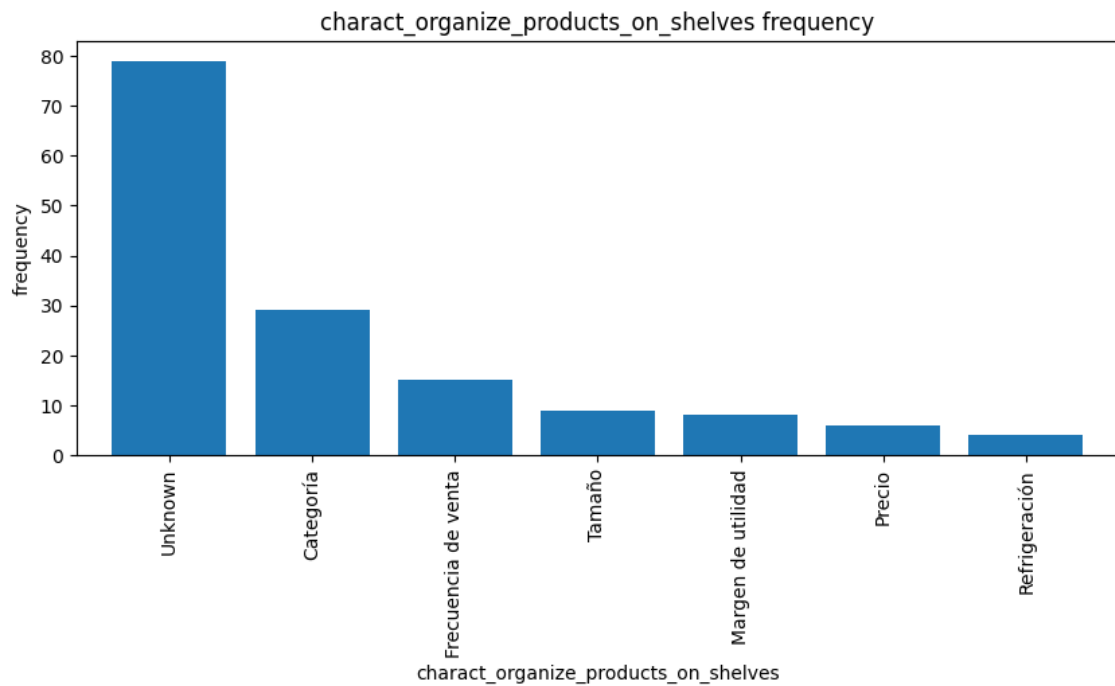


Ilustración 16

20. 234_willingness_participate_workshop

Finalmente, respecto a la voluntad de participar en el taller para el que se realizó la encuesta el 42.7% de los encuestados mantuvieron su respuesta privada, un 28.7% negó estar interesado y otro 28.7% comunicó su interés por participar. Estos hallazgos se visualizan en la Ilustración 17.

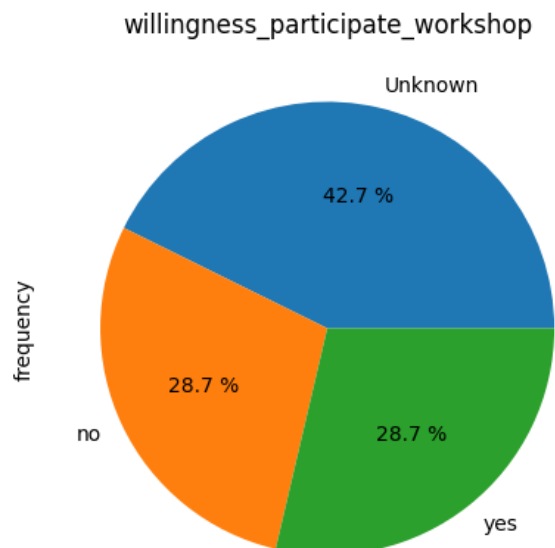


Ilustración 17

Referencias.

Castañón, X. (2019, 25 febrero). *The hidden impact of micro-retailers' survival rate on the logistics cost of consumer-packaged goods companies* - MIT LIFT Lab. MIT LIFT Lab.

<https://liftlab.mit.edu/the-hidden-impact-of-micro-retailers-survival-rate-on-the-logistics-cost-of-consumer-packaged-goods-companies/>

Da Silva, D. (2023, 18 septiembre). *Qué es escala de Likert y cómo aplicarla*. Zendesk MX.

<https://www.zendesk.com.mx/blog/que-es-escala-de-likert/>

Kushner, A. (2023). What is procurement? your guide to the entire process. Approve.com.

[https://www.approve.com/blog/what-is-](https://www.approve.com/blog/what-is-procurement/#:~:text=What%20exactly%20is%20procurement%3F,management%2C%20and%20business%20spend%20management.)

[procurement/#:~:text=What%20exactly%20is%20procurement%3F,management%2C%20and%20business%20spend%20management.](https://www.approve.com/blog/what-is-procurement/#:~:text=What%20exactly%20is%20procurement%3F,management%2C%20and%20business%20spend%20management.)