

Actividad 4 (Extracción de Características)

Como uno de los principales y más importantes puntos previos a realizar los análisis de las variables fue la limpieza de la base de datos, se realizó la limpieza de valores nulos y los valores atípicos de cada una de las variables, una vez limpia la base de datos, se procedió al análisis:

1. Variable `_updated_by`

Para esta variable, se utilizó un filtro con el cual se obtienen los valores más relevantes de la variable, en este caso la frecuencia, posteriormente se obtuvo una gráfica de barras en donde uno de los principales hallazgos de esta variable fue que Salomón Abdala Perez fue quien más veces se presentó en esta columna, seguido por Ángel Francisco y José Zapata.

2. Variable `33_credit_to_customers`

De igual manera, se utilizó un filtro con el cual se obtienen los valores más relevantes de la variable, posteriormente se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 78% de las personas contestaron que sí al rubro de 'credit to customers', mientras el otro 21% respondió negativamente.

3. Variable `232_type_of_store`

Aplicando el filtro con el cual se obtienen los valores más relevantes de la variable, se obtuvo una gráfica de barras en donde 'Grocery store' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por Micro-restaurant y Dairy products.

4. Variable `108_does_the_micro_retailer_has_a_barred_window_`

De igual manera, se utilizó un filtro con el cual se obtienen los valores más relevantes de la variable, posteriormente se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 72% de las personas contestaron que sí al rubro de 'credit to customers', mientras el otro 27% respondió negativamente.

5. Variable `99_does_the_micro_retailer_exhibits_products_outside_`

Ahora bien, para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 70% de las personas contestaron que sí al rubro específico, mientras el otro 29% respondió negativamente.

6. Variable `102_does_the_micro_retailer_sells_fresh_products_`

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 61% de las personas contestaron que sí al rubro específico, mientras el otro 38% respondió negativamente.

7. Variable `184_store_devices`

Para esta variable se obtuvo una gráfica de barras en donde 'Smartphone' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por Analog cash y POS system.

8. Variable `hola_somos_estudiantes_del_tec_de_monterrey_estamos_realiz`

Ahora bien, para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 66% de las personas contestaron que sí al rubro específico, mientras el otro 33% respondió negativamente.

9. Variable `le_comento_que_toda_esta_informacin_es_confidencial`

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 95% de las personas contestaron que sí al rubro específico, mientras el otro 4% respondió negativamente.

10. Variable `5_change_store_space_last_year`

Para esta variable se obtuvo una gráfica de barras en donde 'No cambió' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por Aumentó 25% y Disminuyó 15%.

11. Variable `6_change_employees_average_salary_last_year`

Para esta variable se obtuvo una gráfica de barras en donde 'No cambió' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por Aumentó 15% y Aumentó 50% o más.

12. Variable `20_reviews_finances_monthly`

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 74% de las personas contestaron que sí al rubro específico, mientras el otro 25% respondió negativamente.

13. Variable `49_inventory_records`

Para esta variable se obtuvo una gráfica de barras en donde 'En papel' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por 'En excel' y 'Software especializado'.

14. Variable `18_sales_records`

Similar a la variable anterior, para esta se obtuvo una gráfica de barras en donde 'En papel' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por 'Software especializado' y 'No lo hago'.

15. Variable `19_tax_id`

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 66% de las personas contestaron que sí al rubro específico, mientras el otro 33% respondió negativamente.

16. Variable `310_burnout`

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 57% de las personas contestaron que sí al rubro específico, mientras el otro 42% respondió negativamente.

17. Variable 186_internet_connection

Para esta variable se obtuvo una gráfica de pastel en donde nos arrojó que alrededor del 71% de las personas contestaron que sí al rubro específico, mientras el otro 28% respondió negativamente.

18. Variable 210_sales_channels

Para esta variable se obtuvo una gráfica de barras en donde 'Redes Sociales (Fb, Insta, Whats)' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por 'WhatsApp' y 'Redes Sociales (Fb, Insta)'.

19. Variable 189_payment_methods

Para esta variable se obtuvo una gráfica de barras en donde 'Efectivo' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por 'Efectivo, tarjeta de crédito, tarjeta de débito, Transferencia interbancaria' y 'Efectivo, tarjeta de crédito y/o tarjeta de débito'.

20. Variable 214_customer_relationship_tools

Para esta variable se obtuvo una gráfica de barras en donde 'Redes Sociales' fue el dato dentro de esta variable que más frecuencia obtuvo, seguido por 'Ninguna' y 'Aplicación Móvil'.