



Modelos estadísticos y probabilísticos

2.^a entrega de proyecto

Satisfacción en compras en línea

Luis Felipe Anaya Parada

Diego Sebastian Rojas Guerrero

INTRODUCCIÓN. -

En el contexto del comercio electrónico en constante evolución, la satisfacción del cliente se ha convertido en un componente crítico para el éxito empresarial. Este proyecto de Modelos Estadísticos y Probabilísticos se centra en explorar y analizar la satisfacción del cliente en las compras en línea. Con el aumento significativo del comercio electrónico en los últimos años, los factores que influyen en la satisfacción del cliente son esenciales para las empresas que desean mantener su competitividad en un mercado cada vez más exigente y dinámico.

Para abordar esta cuestión, hemos llevado a cabo una investigación exhaustiva que combina métodos continuos, discretos y cualitativos. En particular, hemos realizado encuestas detalladas a una muestra representativa de consumidores habituales de compras en línea. Estas encuestas nos han proporcionado una comprensión profunda de las percepciones de los usuarios a la hora de realizar compras en línea.

OBJETIVO. -

El objetivo principal de este proyecto es investigar y comprender la satisfacción de los clientes en sus compras en línea, así como determinar si estas experiencias son óptimas y valen la pena o deben mejorar. Para lograr este propósito, se llevarán a cabo encuestas detalladas para recopilar datos sobre la experiencia de usuario de los consumidores al realizar compras en línea. Se analizarán los factores más que nada de los gustos personales y experiencia de usuario dentro de las cientos de páginas de compras en línea. Tocando temas como plataformas favoritas, categorías de productos preferidos, dinero gastado en un solo mes y experiencias con devoluciones, entre otras.

POBLACIÓN OBJETIVO. –

La población objetivo de este estudio son los jóvenes que realizan compras en línea de manera regular. Se enfocará específicamente en aquellos individuos que tienen experiencia y hábitos establecidos de compra en plataformas de comercio electrónico. Se excluye a aquellos que no realizan compras en línea de forma habitual, ya que el objetivo es analizar la satisfacción de aquellos que están familiarizados con este tipo de transacciones y que representan una parte significativa del mercado digital. Este enfoque permitirá obtener una comprensión más precisa y relevante de la satisfacción del cliente en el contexto específico de los jóvenes consumidores en línea el cual estamos buscando.

DISEÑO DE MUESTREO. –

El diseño de muestreo para este estudio implica la selección de una muestra de 150 personas de la población objetivo de jóvenes que realizan compras en línea. Utilizaremos un enfoque de muestreo aleatorio simple, donde cada individuo en la población tiene la misma probabilidad de ser seleccionado para participar en el estudio. Esta selección se llevará a cabo mediante la distribución de enlaces de encuestas en redes sociales y otros canales en línea. Este diseño nos permitirá obtener una muestra diversa de jóvenes compradores en línea, lo que facilitará un análisis significativo de la satisfacción del cliente en el contexto del comercio electrónico.

METODO DE MEDICION. –

El método de medición empleado en este estudio consistió en la realización de encuestas en línea por vía “Forms” dirigidas a los jóvenes. Este enfoque permitió alcanzar a una muestra amplia y diversa de jóvenes compradores en línea, garantizando la validez y la representatividad de los resultados obtenidos.

INSTRUMENTO DE MEDICIÓN. –

En las preguntas del cuestionario que fue seleccionado como método de medición para esperar las respuestas deseadas minimizamos el rango de error implementando la función que en las preguntas continuas y discretas solo se puedan recibir enteros y en las cualitativas respuestas predeterminadas para evitar la variación exagerada de cada una de las respuestas

al dejarlas libres. En este caso realizamos diez preguntas en esta encuesta, tres cualitativas, tres continuas y tres discretas

PRUEBA PILOTO. –

Los resultados de la prueba piloto inicial, realizada con una muestra de 15 personas seleccionadas aleatoriamente, demostraron una validación satisfactoria de las preguntas y el enfoque. Sin embargo, tras esta prueba piloto, se identificaron áreas de mejora en la claridad y especificidad de algunas preguntas. En consecuencia, se realizaron ajustes en el cuestionario para garantizar una mayor precisión y relevancia de los datos recopilados.

ORGANIZACIÓN DEL TRABAJO DE CAMPO. -

Para este apartado obtuvimos una respuesta muy simple para la organización de los cuestionarios porque afortunadamente contamos con las herramientas como “Forms” que crea la facilidad de contestar los cuestionarios en la instantaneidad de tu teléfono celular, tomando menos de 5 minutos en contestar desde la comodidad de cualquier lugar.

ORGANIZACIÓN DE MANEJO DE DATOS. -

Par manejar cada respuesta con la calidad y orden que deseamos con la herramienta de “Microsoft Forms” como mencionamos anteriormente se utiliza la función de “solo enteros” y “respuestas predeterminadas” para que evitar respuestas erróneas, a la vez las respuestas que parecen ficticias, ya que no contamos con una supervisión total a la hora de que los usuarios respondan las encuestas, si descubrimos un dato ficticio simplemente descartamos la respuesta de cuestionario al completo.

Matriz de la base de datos:

	Que tan frecuente	Plataforma favorita	Experiencia haciendo compras en línea	de productos que mas consumes	Horas al mes buscando	Compras por pedido	Dinero mensual	Compra mas cara	Problemas con paquetes	Articulos devueltos al año
1										
2										
3										

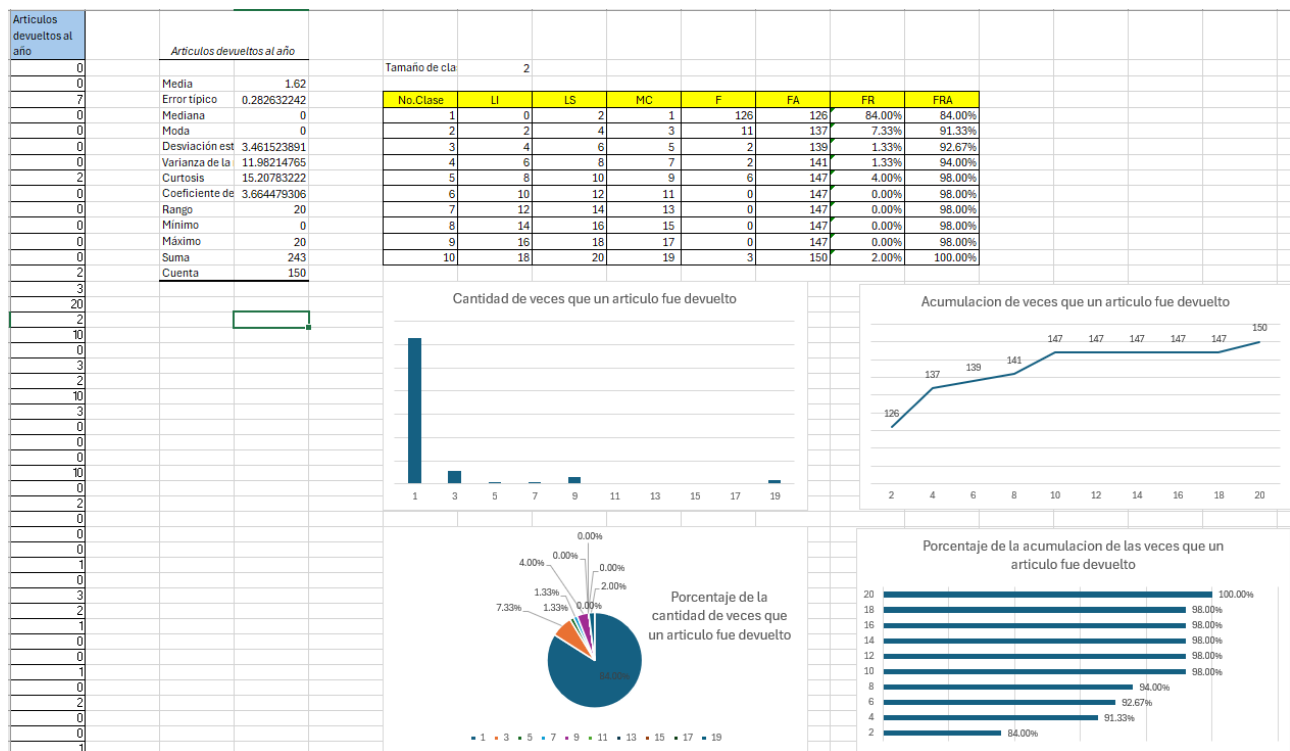
ANÁLISIS DE DATOS. -

Todas las respuestas después de las más de 150 encuestas realizadas fueron de acuerdo con el resultado esperado al tener un margen de error casi nulo en las respuestas de “solo enteros” y un absoluto 0 en los repuestos predeterminados no recibimos ni una respuesta fuera del margen esperado y pudimos hacer el análisis, base de datos y análisis de datos. Trabajamos de acuerdo con lo esperado.

Desarrollo

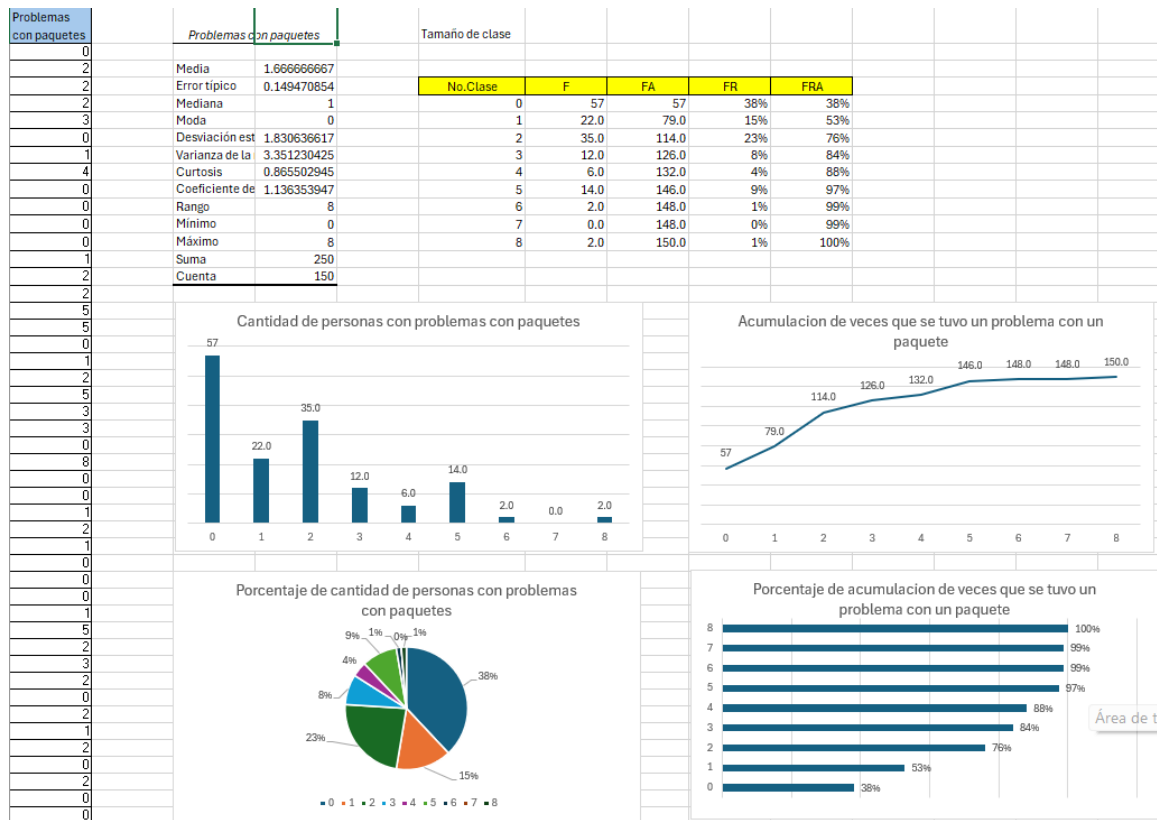
Artículos devueltos en un año. -

En este apartado logramos notar una variabilidad limitada entre las respuestas obteniendo 20 resultados diferentes, pero, sorpresivamente hay un rango que resalto entre todas las respuestas que fue de 0-1 paquetes devueltos a través de todo el año, logrando ver un 84% de las respuestas acorde a este rango. Haciendo notar como los problemas son la devolución no suele ser un problema de escalas mayores.



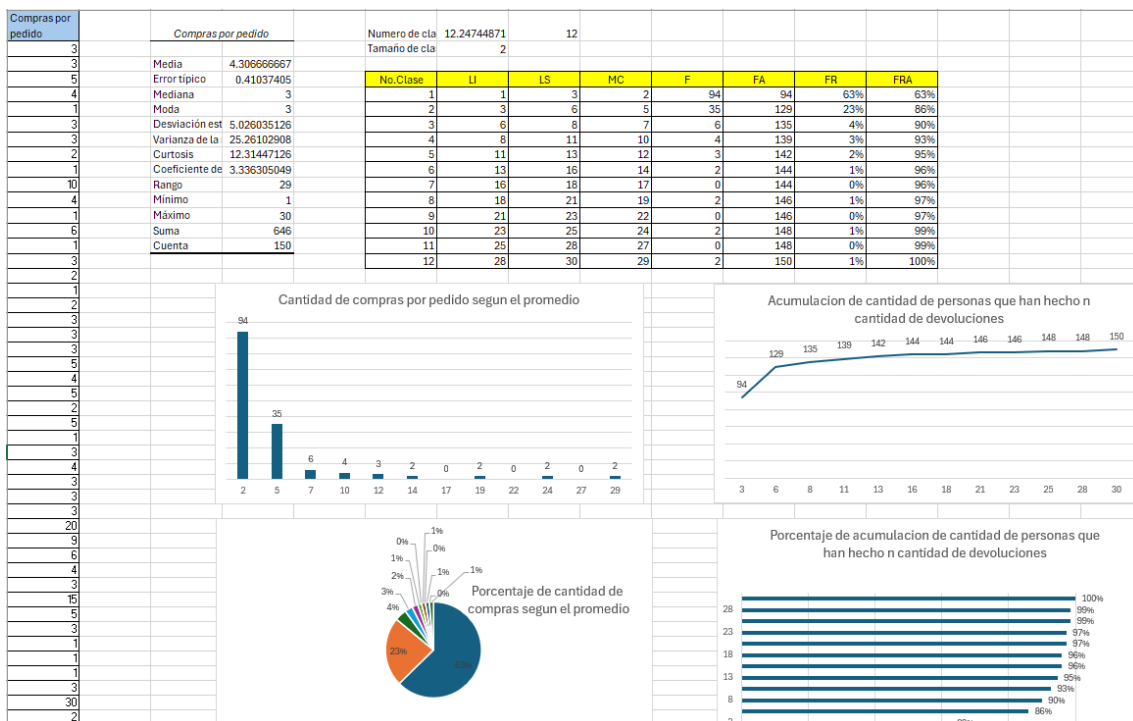
Problemas con entregas de paquetes. -

En este apartado se puede apreciar cómo hay una gran variación entre las respuestas el problema con las entregas es un problema muy resonado con las ultimas páginas de venta online de “mayoreo”, suele haber bastantes problemas con las entregas tardías o no entregadas pudiendo observar cómo un 65% de los usuarios tienen malas experiencia con la entrega de paquetes.



Compras por pedido. -

Podemos notar en este apartado como los usuarios son compradores ocasionales como bien se menciona en la primera pregunta, suelen comprar de 1-2 productos por compra total o “carrito” dejándolo con un 63% de compradores ocasionales y con pedidos pequeños para uso personal pero por otro lado podemos resaltar como hubo 2 personas que hacen 30 compras en un solo pedido podemos llegar a una conclusión un tanto intuitiva de que pueden ser compras para no uso personal si no para vender por otros medio ya sea en otras páginas o presencialmente.



Categoría de productos que más consumes. -

Notablemente en este apartado podemos observar una variabilidad diversa en todas las respuestas predeterminadas pudiendo hacer énfasis en la después de ropa donde podemos adjuntar lo anterior hacer compras en línea de ropa en las nuevas plataformas de “mayoreo” y vender presencial o solamente como uso personal, llevándose un 38% por cierto del total. Algo para resaltar es como el mercado de los alimentos y bebidas aun no es tan fuerte dentro de las compras en línea.

Categoría de productos que mas consumes			
<div> <div>Cuenta de Categoría de productos que mas consumes</div> <div> <div>Etiquetas de fila</div> <div> <div></div> <div>consumes</div> </div> </div> </div>			
Ropa			
Otros	Alimentos y bebidas	5	
Ropa	Electronicos	24	
Ropa	Entretenimiento	22	
Electronicos	Otros	42	
Electronicos	Ropa	57	
Total general		150	
<div> <div>Categorías</div> <div> <div>F</div> <div>FR</div> </div> </div>			
Ropa	Alimentos y bebidas	5	3%
Ropa	Electronicos	24	16%
Ropa	Entretenimiento	22	15%
Ropa	Otros	42	28%
Entretenimiento	Ropa	57	38%
Alimentos y bebidas	Total	150	

Categorías de productos mas consumidos

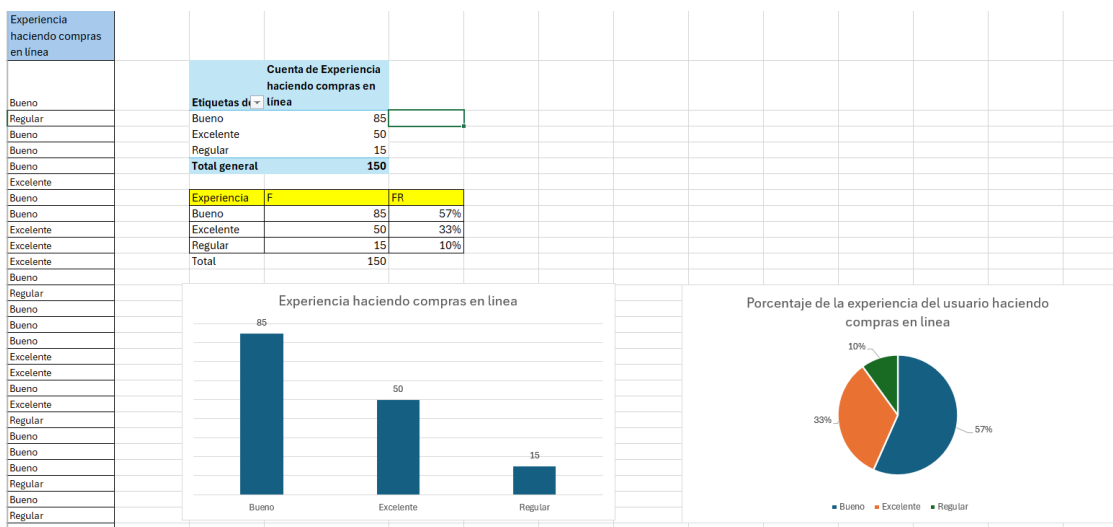
Categoría	Count
Alimentos y bebidas	5
Electronicos	24
Entretenimiento	22
Otros	42
Ropa	57

Porcentaje de categorías mas consumidas

Categoría	Percentage
Alimentos y bebidas	3%
Electronicos	16%
Entretenimiento	15%
Otros	28%
Ropa	38%

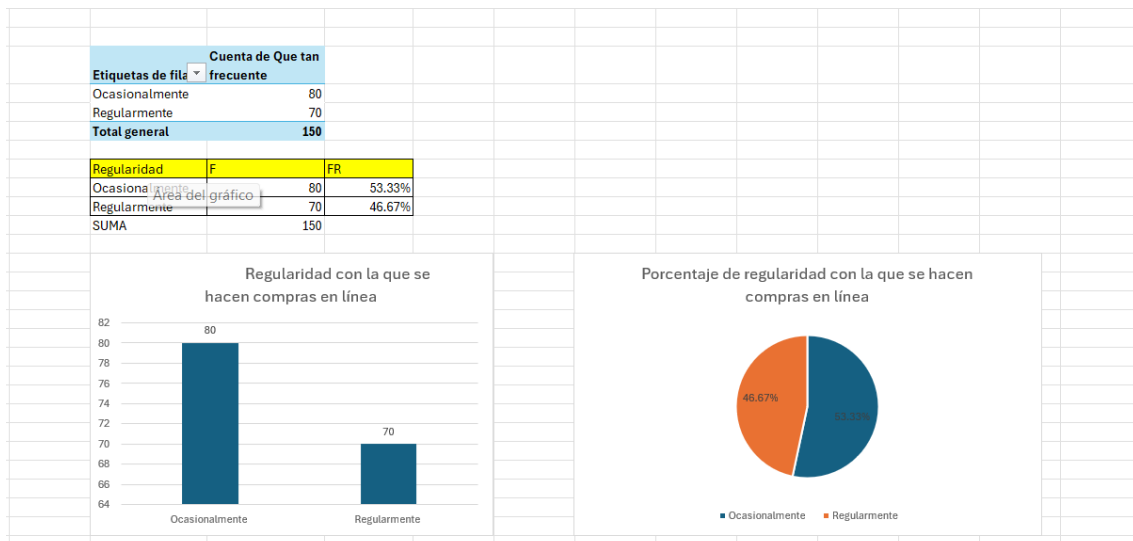
Experiencia de usuario haciendo compras. -

La experiencia de usuario es un punto importante a tomar en cuenta ya que resume la experiencia del usuario en todos los aspectos a la hora de realizar una compra en línea, sorprendentemente la mayoría de personas han tenido una experiencia positiva con ciertos puntos a trabajar pero claramente hay una mayoría que disfruta hacer este tipo de compras más reciente en los miles de sitios que ofrecen la posibilidad de compras en línea llevándose un 90% de opiniones positivas.



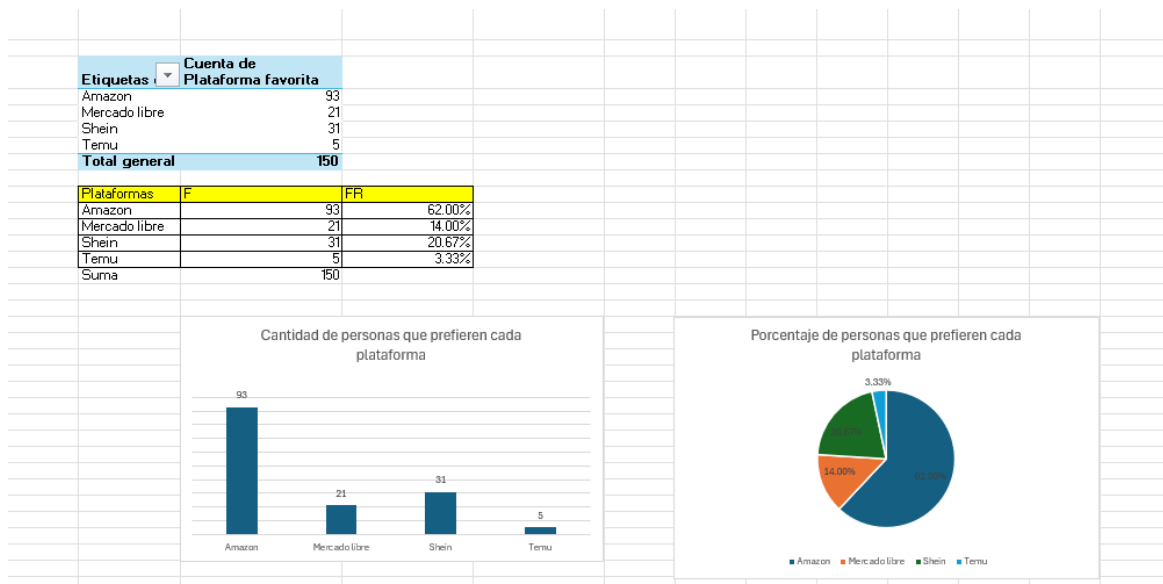
Que tan frecuente los usuarios realizan compras en línea. -

En este apartado analizamos la regularidad con la que los usuarios hacen compras en línea, los resultados arrojaron que el 53.33% de las personas encuestadas lo hacen ocasionalmente, mientras que el 46.67% lo hace con regularidad, haciendo notar que la mayoría de las personas hacen compras en línea de una manera esporádica, no tan seguido.



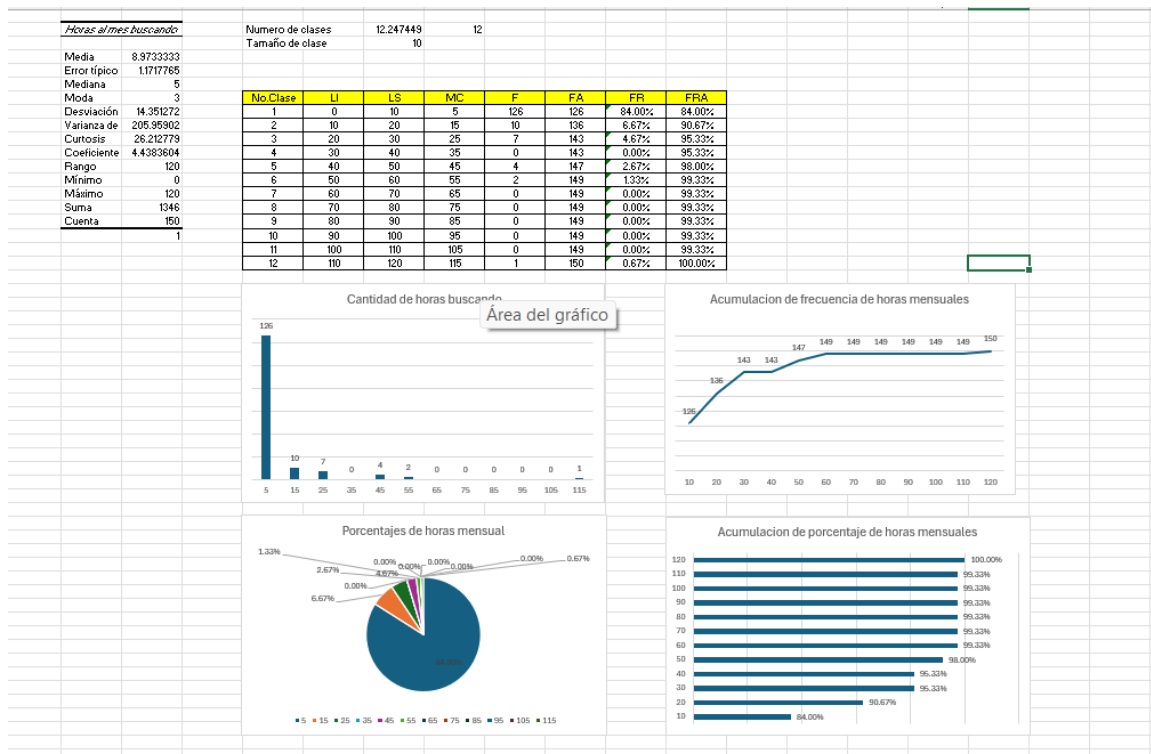
Plataforma donde los usuarios prefieren adquirir sus productos. -

En este apartado podemos notar que no hay mucha variabilidad, ya que, el mayor porcentaje se lo lleva la plataforma de Amazon con un aplastante 62%, mientras que el otro 38% se reparte entre las otras tres plataformas. Amazon desde su salida ha sido una plataforma impecable, siendo la plataforma donde las tiendas prefieren llevar su mercado.



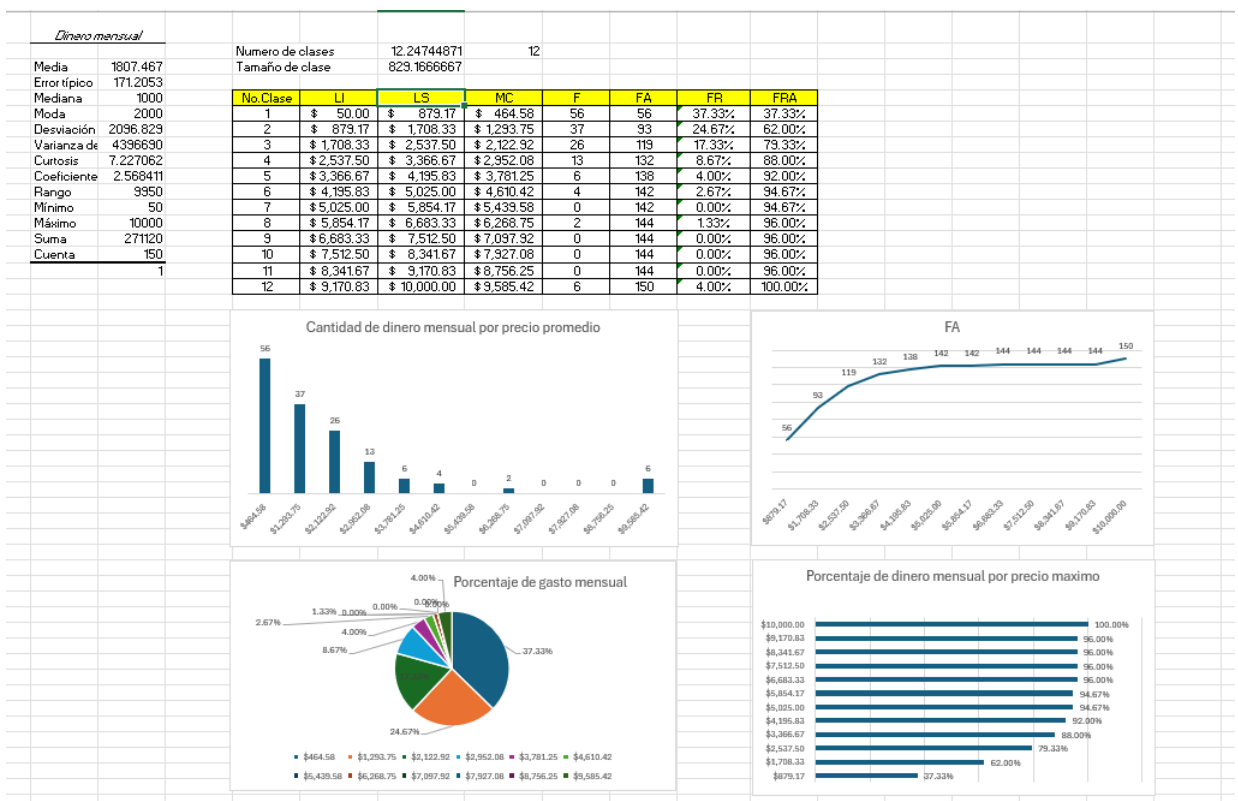
Cantidad de horas al mes que la gente usa buscando productos en línea. -

En este apartado recibimos una variabilidad muy amplia, gente que usa tanto una hora al mes, así como hay gente que usa cien horas al mes, nuestros resultados arrojaron que una gran parte de las personas que formaron parte del muestreo usaron un promedio de 5 horas al mes buscando productos en línea.



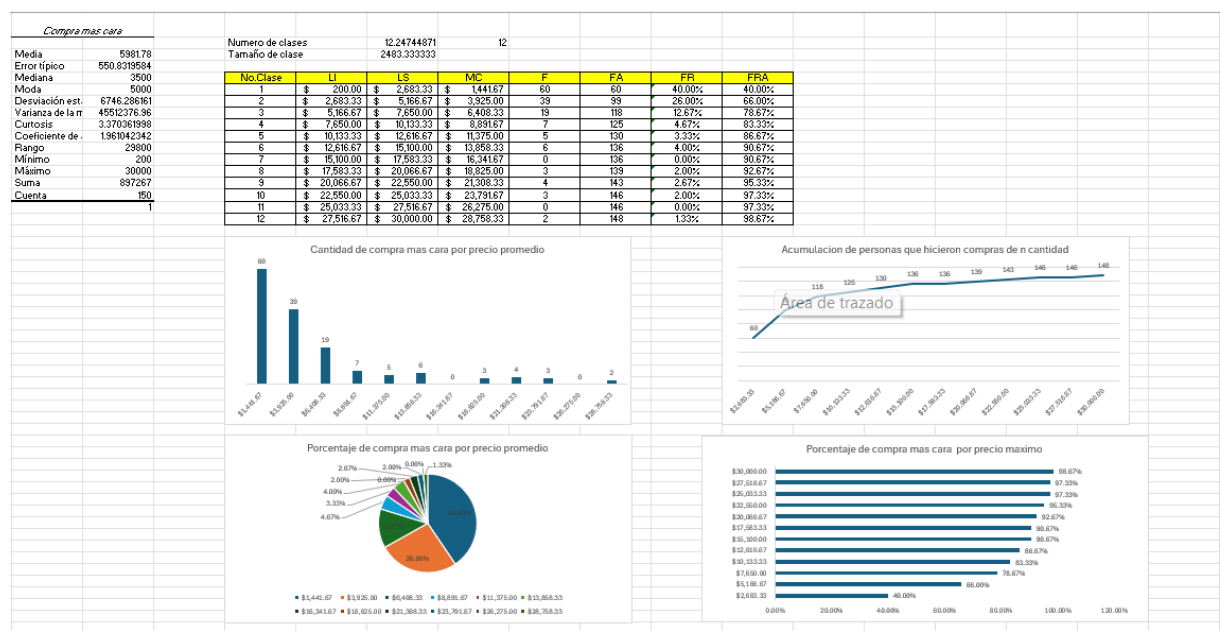
Cantidad de dinero mensual que gastan los usuarios al mes. -

En este apartado, a pesar de que recibimos una variabilidad muy alta, hubo una cantidad que en promedio se repitió más que el resto, los resultados arrojaron que la mayoría de las personas que fueron parte del muestreo, en promedio gastan \$464.56 pesos mensuales en plataformas en línea.



Compras más caras realizadas por el usuario. -

En este apartado, como en su mayoría, recibimos una variabilidad muy alta, la mayoría de las respuestas no era ni remotamente parecida a otra, pero claro que hay una cantidad que destacó y que se repitió más que las demás, los resultados arrojaron que, en promedio, la compra más cara es de \$1,441.67 pesos, esta cantidad se obtuvo con una frecuencia de 60, lo que significa que el 40% de los encuestados, su compra mas cara, en promedio es de \$1,441.67 pesos.



Análisis de las tablas de datos cruzados. –

Cualitativa vs Cualitativa:

Plataformas	Que tan frecuente realizas una compra en línea		Total
	Ocasionalmente	Regularmente	
Amazon	50	43	93
Mercado Libre	9	12	21
Shein	20	11	31
Temu	1	4	5
Total	80	70	150

Los resultados revelaron que Amazon es la plataforma de compras en línea más popular, con un total de 93 personas que la eligieron, seguida de cerca por Shein con 31 y Mercado Libre con 21. Temu, por otro lado, tuvo una menor presencia con solo 5 personas registradas.

Al observar las probabilidades individuales, se encontró que la mayoría de los participantes tienen una probabilidad del 62% de realizar compras en Amazon, seguido de un 21% en Shein, un 14% en Mercado Libre y un 3% en Temu.

Los análisis de probabilidad condicional destacaron la relación entre las plataformas y la frecuencia de compra. Por ejemplo, la probabilidad de comprar en Amazon dado que se compra ocasionalmente es del 62.50%, mientras que la probabilidad de comprar ocasionalmente dado que se compra en Amazon es del 53.76%.

Las probabilidades de intersección mostraron cuántos participantes cumplían simultáneamente dos criterios específicos, como comprar en Amazon y hacerlo ocasionalmente, que representó el 33.33% del total

Cualitativa vs Discreta:

Categorías	Cantidad de problemas con paquetes								Total
	0	1	2	3	4	5	6	8	
Alimentos y bebidas	0	0	3	0	0	0	0	0	5
Electronicos	8	4	2	2	4	2	0	2	24
Entretenimiento	8	2	3	2	0	5	2	0	22
Otros	22	4	10	4	2	0	0	0	42
Ropa	19	10	17	4	0	7	0	0	57
Total general	57	22	35	12	6	14	2	2	150

Se observó que la categoría de Alimentos y Bebidas registró la menor cantidad de inconvenientes, con un total de 5 casos reportados. Por otro lado, la categoría de Electrónicos destacó con la mayor cantidad de problemas, contabilizando un total de 24 casos, lo que sugiere una susceptibilidad significativa en la entrega de estos productos.

Al considerar las probabilidades asociadas, se encuentra que la probabilidad de experimentar un problema con un paquete en la categoría de Ropa es notablemente alta, alcanzando un 38%. Esta cifra indica una propensión sustancial a los desafíos logísticos en la entrega de artículos de vestimenta.

El análisis de probabilidades de intersección proporciona información valiosa sobre la coocurrencia de problemas de envío en categorías específicas y con cantidades determinadas de problemas. Por ejemplo, se observa que la probabilidad de que un paquete de Ropa presente exactamente 2 problemas es del 11.33%, lo que señala una tendencia específica dentro de esta categoría.

Además, las probabilidades incluyentes y excluyentes permiten una comprensión más profunda de la distribución de problemas con paquetes entre varias categorías y combinaciones de estas. Estos análisis revelan la probabilidad de que un paquete tenga al menos un problema en una categoría determinada o en una combinación de categorías, así como la probabilidad de que un paquete tenga problemas en una categoría dada, excluyendo otras categorías.

Cualitativa vs Continua:

Experiencia	Dinero mensual que se gasta																						Total general
	50	60	100	200	300	350	400	500	600	800	900	1000	1200	1500	1700	2000	3000	3800	4000	5000	6000	10000	
Bueno	4	2	2	3	2	0	2	6	4	4	1	8	4	2	1	18	9	2	2	4	2	3	85
Excelente	0	0	0	1	2	2	2	6	7	2	0	11	0	2	0	6	4	0	2	0	0	3	50
Regular	2	0	0	0	0	0	2	0	1	0	0	6	2	0	0	2	0	0	0	0	0	0	15
Total general	6	2	2	4	4	2	6	12	12	6	1	25	6	4	1	26	13	2	4	4	2	6	150

Estos datos representan un análisis de gastos mensuales en relación con la experiencia en un grupo de individuos. Observamos que la mayoría de los gastos mensuales se concentran en el rango de 1000 a 2000 unidades monetarias, con una probabilidad del 16.67% para gastos de 1000 y del 17.33% para gastos de 2000.

Además, la probabilidad de tener una experiencia clasificada como "Bueno" y realizar gastos de 1000 es del 5.33%, mientras que para la misma experiencia y gastos de 2000 es del 12.00%. Esto sugiere que aquellos con una experiencia calificada como "Bueno" tienen una mayor propensión a realizar gastos más altos en comparación con otros grupos de experiencia.

Por otro lado, la probabilidad de que alguien con experiencia clasificada como "Excelente" realice gastos de 1000 es del 7.33%, y para gastos de 2000 es del 4.00%. Esto indica que, aunque aquellos con experiencia "Excelente" también tienden a realizar gastos más altos, la probabilidad es ligeramente menor en comparación con el grupo "Bueno".

Discreta vs Discreta:

Problemas con paquetes	Cantidad de veces al año que productos son devueltos										Total
	0	1	2	3	4	5	7	8	10	20	
0	47	6	0	0	0	0	0	0	4	0	57
1	20	0	0	0	2	0	0	0	0	0	22
2	10	7	11	5	0	0	1	1	0	0	35
3	2	4	0	2	0	2	0	0	2	0	12
4	0	4	2	0	0	0	0	0	0	0	6
5	2	2	5	2	0	0	0	0	0	3	14
6	0	0	2	0	0	0	0	0	0	0	2
8	2	0	0	0	0	0	0	0	0	0	2
Total	83	23	20	9	2	2	1	1	6	3	150

Observamos que las probabilidades individuales varían significativamente, desde un 38.00% para el evento "0" hasta un 1.33% para el evento "8".

Además, al analizar las intersecciones, encontramos que ciertos eventos tienen una probabilidad considerablemente mayor de ocurrir simultáneamente, lo que sugiere una posible relación entre ellos. Por ejemplo, la intersección de los eventos "2" y "2" tiene una probabilidad del 7.33%, mientras que la intersección de los eventos "1" y "5" tiene una probabilidad del 2.67%.

Esto indica que ciertos eventos pueden estar correlacionados en su ocurrencia. Además, al examinar las probabilidades inclusivas y excluyentes, podemos entender mejor la probabilidad de que ocurran eventos dentro de ciertos conjuntos o condiciones. Estos datos son fundamentales para comprender las relaciones probabilísticas entre diferentes eventos y pueden ser útiles en la toma de decisiones y el análisis de riesgos.

Discreta vs Continua:

Horas buscadas	Cantidad de compras por pedido															Total
	1	2	3	4	5	6	7	9	10	12	15	20	25	30		
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	
1	5	2	5	0	2	0	0	0	0	0	2	0	0	0	16	
2	12	0	1	5	2	0	0	0	0	0	0	0	2	0	22	
3	5	8	2	2	0	2	2	0	2	0	0	0	0	0	23	
4	4	0	8	0	0	0	0	0	0	0	0	0	0	0	12	
5	0	2	7	2	6	2	0	0	0	0	0	2	0	0	21	
6	3	0	2	0	0	0	0	2	0	0	0	0	0	0	7	
8	0	0	4	0	0	0	0	0	0	0	0	0	0	0	4	
10	0	2	7	4	4	0	0	0	0	3	0	0	0	0	20	
15	0	2	2	0	2	0	0	0	0	0	0	0	0	0	6	
20	0	0	0	2	0	0	0	0	0	0	0	0	0	2	4	
24	0	3	0	0	2	0	0	0	0	0	0	0	0	0	5	
30	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	
50	2	0	1	1	0	0	0	0	0	0	0	0	0	0	4	
60	0	0	2	0	0	0	0	0	0	0	0	0	0	0	2	
120	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	
Total	34	19	41	16	19	4	2	2	2	3	2	2	2	2	150	

La tabla muestra que la mayoría de las compras se realizan después de buscar entre 1 y 3 horas, con un total de 350 compras.

Las horas de búsqueda adicionales parecen tener un efecto marginal en la cantidad de compras, ya que las cantidades disminuyen a medida que aumenta el tiempo de búsqueda. Por ejemplo, se observa una disminución significativa en la cantidad de compras después de 3 horas de búsqueda, con solo 50 compras registradas después de 5 horas de búsqueda. Sin embargo, la probabilidad de realizar una compra parece estabilizarse después de 3 horas de búsqueda, con un 15% de probabilidad de realizar 2 compras por pedido y un 10% de probabilidad de realizar 1 compra. Por otro lado, las probabilidades de intersección revelan que la probabilidad de realizar 2 compras por pedido después de buscar 3 horas es del 30%, mientras que la probabilidad de realizar 1 compra es del 20%.

Esto sugiere que las personas que buscan durante un período más prolongado tienen una mayor probabilidad de realizar múltiples compras por pedido. En resumen, la tabla muestra que el tiempo de búsqueda influye en la cantidad de compras realizadas por pedido, con una disminución en las compras a medida que aumenta el tiempo de búsqueda, aunque las probabilidades de realizar múltiples compras por pedido se mantienen relativamente estables después de cierto punto.

Continua vs Continua:

Dinero Mensual		Compra mas cara												Total
		\$200.00	\$2,683.33	\$5,166.67	\$7,650.00	\$10,133.33	\$12,616.67	\$15,100.00	\$17,583.33	\$20,066.67	\$22,550.00	\$25,033.33	\$27,516.67	
\$50.00	\$879.17	36	9	4	2	1	2	0	0	2	0	0	0	56
\$879.17	\$1,708.33	13	11	2	2	2	0	0	3	2	0	0	2	37
\$1,708.33	\$2,537.50	5	9	6	2	2	0	0	0	0	2	0	0	26
\$2,537.50	\$3,366.67	4	4	2	0	0	2	0	0	0	1	0	0	13
\$3,366.67	\$4,195.83	2	2	0	0	0	2	0	0	0	0	0	0	6
\$4,195.83	\$5,025.00	0	4	0	0	0	0	0	0	0	0	0	0	4
\$5,025.00	\$5,854.17	0	0	0	0	0	0	0	0	0	0	0	0	0
\$5,854.17	\$6,683.33	0	0	2	0	0	0	0	0	0	0	0	0	2
\$6,683.33	\$7,512.50	0	0	0	0	0	0	0	0	0	0	0	0	0
\$7,512.50	\$8,341.67	0	0	0	0	0	0	0	0	0	0	0	0	0
\$8,341.67	\$9,170.83	0	0	0	0	0	0	0	0	0	0	0	0	0
\$9,170.83	\$10,000.00	0	0	3	1	0	0	0	0	0	0	0	2	6
Total		60	39	19	7	5	6	0	3	4	3	0	4	150

El proyecto analiza la relación entre el dinero mensual que las personas gastan en compras en línea y la probabilidad de que realicen una compra en diferentes rangos de precios. Los datos muestran que la mayoría de las compras se realizan en el rango de \$200 a \$2,683.33, con una probabilidad del 79.31%. A medida que el precio de compra aumenta, la probabilidad de que las personas realicen una compra disminuye significativamente, siendo menos probable que realicen compras en rangos de precios más altos. Esto sugiere que existe una relación inversa entre el dinero disponible para gastar y la disposición a gastar sumas más altas en compras en línea.

Interpretaciones de las probabilidades

1. Devoluciones al año

Distribución Binomial:

Se ha calculado la probabilidad de obtener exactamente 130 devoluciones en un año, que es del 6.29%.

La probabilidad de obtener 125 o menos devoluciones es del 8.499%.

La probabilidad de obtener 135 o más devoluciones es del 2.42%.

La probabilidad de obtener entre 130 y 140 devoluciones es del 15.77%.

Distribución Hipergeométrica:

Se ha calculado la probabilidad de obtener exactamente 66 devoluciones en un año, que es del 7.38%.

La probabilidad de obtener 68 o menos devoluciones es del 99.35%.

La probabilidad de obtener 58 o más devoluciones es del 0.65%.

La probabilidad de obtener entre 57 y 68 devoluciones es del 99.2%.

Distribución Geométrica:

Se ha calculado la probabilidad de que la segunda devolución ocurra en el segundo año, que es del 13.44%.

Distribución de Poisson:

Se ha calculado la probabilidad de que haya exactamente 2 devoluciones en un año, que es del 25.97%.

La probabilidad de que haya 3 o menos devoluciones es del 91.84%.

La probabilidad de que haya 2 o más devoluciones es del 77.82%.

La probabilidad de que haya entre 1 y 3 devoluciones es del 39.99%.

Intervalo de Confianza:

Se ha calculado el intervalo de confianza para la media de devoluciones por año utilizando tanto muestras grandes como pequeñas. Para las muestras grandes, el intervalo de confianza del 90% para la media de devoluciones por año es de [1.155, 2.085]. Para las muestras pequeñas, el intervalo de confianza del 90% para la media de devoluciones por año es de [1.532, 133.374].

Prueba de Hipótesis:

Se han realizado pruebas de hipótesis tanto para muestras grandes como pequeñas sobre la media de devoluciones por año. En ambos casos, se rechaza la hipótesis nula de que la media de devoluciones es mayor a 40 o 2, respectivamente.

2. Problemas con paquetes**Distribución Binomial:**

Se ha calculado la probabilidad de obtener exactamente 30 problemas con paquetes, que es del 5.01%.

La probabilidad de obtener 40 o menos problemas con paquetes es del 4.679%.

La probabilidad de obtener 32 o más problemas con paquetes es del 74.73%.

La probabilidad de obtener entre 30 y 40 problemas con paquetes es del 71.19%.

Distribución Hipergeométrica:

Se ha calculado la probabilidad de obtener exactamente 18 problemas con paquetes, que es del 15.01%.

La probabilidad de obtener 14 o menos problemas con paquetes es del 12.33%.

La probabilidad de obtener 22 o más problemas con paquetes es del 93.91%.

La probabilidad de obtener entre 13 y 22 problemas con paquetes es del 96.4%.

Distribución Geométrica:

Se ha calculado la probabilidad de que el segundo problema con paquetes ocurra en el segundo intento, que es del 17.89%.

Distribución de Poisson:

Se ha calculado la probabilidad de que haya exactamente 2 problemas con paquetes, que es del 26.23%.

La probabilidad de que haya 3 o menos problemas con paquetes es del 91.17%.

La probabilidad de que haya 2 o más problemas con paquetes es del 76.60%.

La probabilidad de que haya entre 1 y 3 problemas con paquetes es del 40.81%.

Intervalo de Confianza:

Se ha calculado el intervalo de confianza para la media de problemas con paquetes utilizando tanto muestras grandes como pequeñas. Para las muestras grandes, el intervalo de confianza del 90% para la media de problemas con paquetes es [1.421, 1.913]. Para las muestras pequeñas, el intervalo de confianza del 90% para la media de problemas con paquetes es [2.951, 49.773].

Prueba de Hipótesis:

Se han realizado pruebas de hipótesis tanto para muestras grandes como pequeñas sobre la media de problemas con paquetes. En ambos casos, se rechaza la hipótesis nula de que la media de problemas con paquetes es mayor a 30 o 4, respectivamente.

3. Compras por pedido

Distribución Binomial:

La probabilidad de realizar exactamente 2 compras por pedido es del 4.25%.

La probabilidad de realizar 8 o menos compras por pedido es del 10.465%.

La probabilidad de realizar 5 o más compras por pedido es del 72.04%.

La probabilidad de realizar entre 4 y 8 compras por pedido es del 70.57%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 3 compras por pedido ocurran es del 31.89%.

La probabilidad de que 1 o menos compras por pedido ocurran es del 10.46%.

La probabilidad de que 8 o más compras por pedido ocurran es del 100.00%.

La probabilidad de que ocurran entre 2 y 7 compras por pedido es del 89.5%.

Distribución Geométrica:

La probabilidad de que la tercera compra por pedido ocurra en el tercer intento es del 3.69%.

Distribución de Poisson:

La probabilidad de que haya exactamente 3 compras por pedido es del 0.76%.

La probabilidad de que haya 4 o menos compras por pedido es del 99.99%.

La probabilidad de que haya 3 o más compras por pedido es del 99.91%.

La probabilidad de que haya entre 2 y 4 compras por pedido es del 0.84%.

Intervalo de Confianza:

Para muestras grandes, el intervalo de confianza del 90% para la media de compras por pedido es [3.632, 4.982].

Para muestras pequeñas, el intervalo de confianza del 90% para la media de compras por pedido es [4.409, 41.255].

Prueba de Hipótesis:

Se rechaza la hipótesis nula de que el promedio de compras por pedido es mayor a 30.

Se rechaza la hipótesis nula de que el promedio de compras por pedido es mayor a 2.

4. Categorías consumidas

Distribución Binomial:

La probabilidad de que exactamente 59 personas consuman principalmente ropa es del 6.29%.

La probabilidad de que 54 o menos personas consuman principalmente ropa es del 5.96%.

La probabilidad de que 60 o más personas consuman principalmente ropa es del 33.51%.

La probabilidad de que entre 52 y 62 personas consuman principalmente ropa es del 64.51%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 consuman principalmente electrónicos es del 13.16%.

La probabilidad de que 25 o menos personas de un total de 150 consuman principalmente electrónicos es del 15.64%.

La probabilidad de que 32 o más personas de un total de 150 consuman principalmente electrónicos es del 84.36%.

La probabilidad de que entre 20 y 36 personas de un total de 150 consuman principalmente electrónicos es del 99.5%.

Distribución Geométrica:

La probabilidad de que la tercera persona en consumir principalmente ropa sea la tercera en ser encuestada es del 14.61%.

5. Experiencia haciendo compras en línea

Distribución Binomial:

La probabilidad de que exactamente 59 personas tengan una experiencia excelente al hacer compras en línea es del 6.90%.

La probabilidad de que 54 o menos personas tengan una experiencia excelente al hacer compras en línea es del 4.84%.

La probabilidad de que 60 o más personas tengan una experiencia excelente al hacer compras en línea es del 17.01%.

La probabilidad de que entre 52 y 62 personas tengan una experiencia excelente al hacer compras en línea es del 61.08%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 tengan una experiencia excelente al hacer compras en línea es del 13.74%.

La probabilidad de que 25 o menos personas de un total de 150 tengan una experiencia excelente al hacer compras en línea es del 5.93%.

La probabilidad de que 32 o más personas de un total de 150 tengan una experiencia

excelente al hacer compras en línea es del 97.19%.

La probabilidad de que entre 20 y 36 personas de un total de 150 tengan una experiencia excelente al hacer compras en línea es del 91.3%.

Distribución Geométrica:

La probabilidad de que la tercera persona que tenga una experiencia excelente al hacer compras en línea sea la tercera en ser encuestada es del 14.81%.

6. Frecuencia de compras

Distribución Binomial:

La probabilidad de que exactamente 59 personas compren regularmente en línea es del 6.52%.

La probabilidad de que 54 o menos personas compren regularmente en línea es del 4.69%.

La probabilidad de que 60 o más personas compren regularmente en línea es del 18.40%.

La probabilidad de que entre 52 y 62 personas compren regularmente en línea es del 58.50%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 compren regularmente en línea es del 12.99%.

La probabilidad de que 25 o menos personas de un total de 150 compren regularmente en línea es del 7.03%.

La probabilidad de que 32 o más personas de un total de 150 compren regularmente en línea es del 96.43%.

La probabilidad de que entre 20 y 36 personas de un total de 150 compren regularmente en línea es del 89.4%.

Distribución Geométrica:

La probabilidad de que la tercera persona que compre regularmente en línea sea la tercera en ser encuestada es del 13.27%.

7. Plataforma favorita

Distribución Binomial:

La probabilidad de que exactamente 59 personas prefieran Mercado Libre como plataforma favorita es del 9.35%.

La probabilidad de que 54 o menos personas prefieran Mercado Libre como plataforma favorita es del 4.98%.

La probabilidad de que 60 o más personas prefieran Mercado Libre como plataforma favorita es del 10.05%.

La probabilidad de que entre 52 y 62 personas prefieran Mercado Libre como plataforma favorita es del 75.58%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 prefieran Mercado Libre como plataforma favorita es del 18.06%.

La probabilidad de que 25 o menos personas de un total de 150 prefieran Mercado Libre como plataforma favorita es del 0.85%.

La probabilidad de que 32 o más personas de un total de 150 prefieran Mercado Libre como plataforma favorita es del 99.15%.

La probabilidad de que entre 20 y 36 personas de un total de 150 prefieran Mercado Libre como plataforma favorita es del 98.3%.

Distribución Geométrica:

La probabilidad de que la tercera persona que prefiera Mercado Libre como plataforma favorita sea la tercera en ser encuestada es del 10.35%.

8. Horas al mes navegando

Binomial:

Se utiliza para calcular la probabilidad de obtener un número específico de éxitos en una secuencia de ensayos, donde cada ensayo tiene dos resultados posibles. Por ejemplo, la probabilidad de obtener exactamente 59 éxitos en 150 ensayos es del 15.26%.

Hipergeométrica:

Describe la probabilidad de obtener un cierto número de éxitos en una muestra sin reemplazo de una población finita. Por ejemplo, la probabilidad de obtener exactamente 28 éxitos en una muestra de 75 cuando se selecciona de una población de 150 con 7 éxitos es del 27.91%.

Geométrica:

Describe la probabilidad de que el primer éxito ocurra en el ensayo k-ésimo en una secuencia de ensayos independientes. Por ejemplo, la probabilidad de que el primer éxito ocurra en el tercer ensayo es del 4.24%.

Normal:

Utilizada para modelar variables continuas, como el tiempo dedicado a buscar en línea. Por ejemplo, la probabilidad de que el tiempo dedicado a buscar en línea sea menor o igual a 9 horas es del 50.07%.

Intervalos de confianza:

Intervalo de confianza para muestras grandes: Se estima el intervalo dentro del cual la media poblacional puede caer con un cierto nivel de confianza. Para este conjunto de datos, el intervalo de confianza del 90% para la media de horas al mes invertidas en compras en línea es [7.05, 10.90].

Intervalo de confianza para muestras pequeñas: Similar al anterior, pero se aplica cuando el tamaño de la muestra es pequeño y se desconoce la desviación estándar de la población. Por ejemplo, el intervalo de confianza del 90% para la media de

horas invertidas al mes en compras en línea es [5.08, 55.16].

Pruebas de hipótesis:

Se realizan pruebas de hipótesis para determinar si las afirmaciones sobre la media de las horas de búsqueda son válidas.

Por ejemplo, una hipótesis nula (H_0) podría ser que el promedio de horas de búsqueda es mayor que 30, mientras que la hipótesis alternativa (H_1) sería que el promedio es menor que 30. Los cálculos conducen a rechazar la hipótesis nula a favor de la hipótesis alternativa.

9. Compra más cara

Distribución Binomial:

La probabilidad de que exactamente 59 personas compren regularmente en línea es del 6.64%.

La probabilidad de que 54 o menos personas compren regularmente en línea es del 4.75%.

La probabilidad de que 60 o más personas compren regularmente en línea es del 17.94%.

La probabilidad de que entre 52 y 62 personas compren regularmente en línea es del 59.31%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 compren regularmente en línea es del 13.23%.

La probabilidad de que 25 o menos personas de un total de 150 compren regularmente en línea es del 30.86%.

La probabilidad de que 32 o más personas de un total de 150 compren regularmente en línea es del 93.33%.

La probabilidad de que entre 20 y 36 personas de un total de 150 compren regularmente en línea es del 65.8%.

Distribución Geométrica:

La probabilidad de que la tercera persona que compre regularmente en línea sea la tercera en ser encuestada es del 14.40%.

Intervalo de Confianza:

Para muestras grandes, el intervalo de confianza para el promedio de horas al mes invertidas en compras en línea está entre \$7,045.93 y \$10,900.73, con una tolerancia de \$17,295.21.

Para muestras pequeñas, el intervalo de confianza para el promedio de horas invertidas al mes en compras en línea está entre \$5,075.15 y \$5,516.70, con una tolerancia de \$4,475.

Pruebas de Hipótesis:

Para la media en una muestra grande, con un nivel de significancia del 10%, se rechaza la hipótesis nula de que el promedio de horas invertidas en compras en línea sea mayor de 30 horas al mes.

Para la media en muestras pequeñas, con un nivel de significancia del 10%, se rechaza la hipótesis nula de que el promedio de horas invertidas en compras en línea sea igual a 8.97.

10. Dinero mensual invertido

Distribución Binomial:

La probabilidad de que exactamente 59 personas gasten menos de 1500 pesos al mes es del 7.54%.

La probabilidad de que 54 o menos personas gasten menos de 1500 pesos al mes es del 4.99%.

La probabilidad de que 60 o más personas gasten menos de 1500 pesos al mes es del 14.90%.

La probabilidad de que entre 52 y 62 personas gasten menos de 1500 pesos al mes es del 65.28%.

Distribución Hipergeométrica:

La probabilidad de que exactamente 28 personas de un total de 150 gasten menos de 1500 pesos al mes es del 14.74%.

La probabilidad de que 25 o menos personas de un total de 150 gasten menos de 1500 pesos al mes es del 22.45%.

La probabilidad de que 32 o más personas de un total de 150 gasten menos de 1500 pesos al mes es del 93.55%.

La probabilidad de que entre 20 y 36 personas de un total de 150 gasten menos de 1500 pesos al mes es del 74.7%.

Distribución Geométrica:

La probabilidad de que la tercera persona que gaste menos de 1500 pesos al mes sea la tercera en ser encuestada es del 14.00%.

Intervalo de Confianza:

Para muestras grandes, el intervalo de confianza para el promedio de dinero gastado al mes está entre \$1,525.86 y \$2,089.07, con una tolerancia de \$508,996.57.

Para muestras pequeñas, el intervalo de confianza para el promedio de dinero gastado al mes está entre \$62.81 y \$162.54, con una tolerancia de \$637.16.

Pruebas de Hipótesis:

Para la media en una muestra grande, con un nivel de significancia del 10%, se acepta la hipótesis nula de que el promedio de dinero gastado al mes es mayor a \$1,500.

Para la media en muestras pequeñas, con un nivel de significancia del 10%, se rechaza la hipótesis nula de que el promedio de dinero gastado al mes sea igual a \$1,807.47.

Interpretación pruebas de independencia para tablas de datos cruzados

Cualitativas vs Cualitativas

Se llevó a cabo un análisis para determinar si la frecuencia con la que los clientes compran está influenciada por la tienda específica en la que realizan sus compras. La prueba buscaba establecer si la elección de una tienda particular afecta la frecuencia de compra de los clientes.

Los resultados del análisis sugieren que no hay suficiente evidencia para rechazar la idea de que la frecuencia de compra depende de la tienda específica. Esto significa que, de acuerdo con los datos y el análisis realizado, la tienda en la que los clientes realizan sus compras tiene una influencia significativa en la frecuencia con la que compran.

Cualitativa vs Discreta:

Se realizó un análisis para evaluar si la cantidad de problemas reportados está influenciada por la categoría del producto. La prueba tenía como objetivo determinar si la categoría del producto afecta el número de problemas que se presentan.

Los resultados del análisis indican que hay suficiente evidencia para concluir que la cantidad de problemas sí depende de la categoría del producto. Esto sugiere que diferentes categorías de productos están asociadas con distintas frecuencias de problemas reportados.

En otras palabras, la categoría a la que pertenece un producto influye significativamente en la cantidad de problemas que se registran. Por lo tanto, es importante considerar la categoría del producto al analizar y gestionar los problemas reportados.

Cualitativa vs Continua:

Se llevó a cabo un análisis para determinar si el dinero mensual gastado por los clientes está influenciado por su experiencia de compra. La prueba tenía como objetivo investigar si la experiencia de compra afecta la cantidad de dinero que los clientes gastan mensualmente.

Los resultados del análisis sugieren que no hay suficiente evidencia para concluir que el dinero mensual gastado dependa de la experiencia de compra. Esto significa que, según los datos analizados, la experiencia de compra no tiene un impacto significativo en la cantidad de dinero que los clientes gastan mensualmente.

En resumen, la conclusión del análisis es que el dinero mensual que los clientes gastan no está relacionado con su experiencia de compra. Otros factores podrían estar influyendo en el gasto mensual, pero la experiencia de compra no parece ser uno de ellos.

Discreta vs Discreta:

Se realizó un análisis para determinar si la cantidad de pedidos devueltos está influenciada por los problemas que se obtienen con los productos. La prueba tenía como objetivo investigar si los problemas reportados afectan la tasa de devoluciones de pedidos.

Los resultados del análisis indican que hay suficiente evidencia para concluir que los pedidos devueltos dependen de los problemas obtenidos. Esto sugiere que la presencia de problemas con los productos está significativamente asociada con una mayor tasa de devoluciones de pedidos.

En otras palabras, los problemas que se presentan con los productos tienen una influencia considerable en la decisión de los clientes de devolver sus pedidos. Por lo tanto, es crucial abordar y resolver los problemas de los productos para reducir la cantidad de devoluciones.

Discreta vs Continua:

Se llevó a cabo un análisis para determinar si la cantidad de compras por pedido está influenciada por el tiempo que los clientes pasan buscando productos. La prueba tenía como objetivo investigar si las horas dedicadas a buscar productos afectan la cantidad de compras realizadas por pedido.

Los resultados del análisis indican que hay suficiente evidencia para concluir que la cantidad de compras por pedido depende de las horas que los clientes pasan buscando productos. Esto sugiere que el tiempo invertido en la búsqueda de productos está significativamente asociado con la cantidad de compras realizadas en cada pedido.

En otras palabras, las horas que los clientes dedican a buscar productos tienen una influencia considerable en la cantidad de compras que realizan por pedido. Por lo tanto, el tiempo de búsqueda es un factor importante para considerar al analizar los hábitos de compra de los clientes.

Continua vs Continua:

Se llevó a cabo un análisis para determinar si el monto de la compra más cara realizada por un cliente está influenciado por el total de dinero que gasta mensualmente. La prueba tenía como objetivo investigar si el gasto mensual total afecta la magnitud de la compra más costosa.

Los resultados del análisis sugieren que no hay suficiente evidencia para concluir que la compra más cara dependa del dinero gastado al mes. Esto significa que, según los datos analizados, el gasto mensual total no tiene un impacto significativo en el monto de la compra más cara realizada por los clientes.

En resumen, la conclusión del análisis es que la cantidad de dinero que un cliente gasta mensualmente no está relacionada con el monto de su compra más cara. Otros factores podrían estar influyendo en el monto de la compra más costosa, pero el gasto mensual total no parece ser uno de ellos.

Interpretación de las pruebas de bondad y ajuste

Artículos devueltos al año:

Los resultados de la prueba chi-cuadrado sugieren que la cantidad de devoluciones podría seguir una distribución de Poisson. Esto se basa en el hecho de que, al comparar los valores observados y esperados de las devoluciones en varias clases, la diferencia no es lo suficientemente grande como para rechazar la hipótesis nula. Es decir, no hay evidencia significativa para afirmar que las devoluciones no se ajustan a una distribución de Poisson. Esto implica que, en términos generales, la frecuencia de devoluciones parece seguir un patrón predecible similar al que se esperaría de un proceso de Poisson.

Problemas con paquetes al año:

Los resultados de la prueba chi-cuadrado sugieren que la cantidad de problemas no sigue una distribución de Poisson. Esto se basa en la discrepancia significativa entre los valores observados y esperados en las distintas clases de problemas. Al calcular la estadística de prueba chi-cuadrado y compararla con el valor crítico para un nivel de significancia del 5%, se encuentra que la estadística de prueba es mucho mayor que el valor crítico. Esto proporciona suficiente evidencia para rechazar la hipótesis nula y concluir que la cantidad de problemas no se ajusta a una distribución de Poisson.

Compras por pedido:

Los resultados de la prueba chi-cuadrado sugieren que la cantidad de pedidos por pedido podría tener un comportamiento aproximadamente de Poisson. Esto se basa en la evaluación de la discrepancia entre los valores observados y esperados en diferentes clases de pedidos por pedido. Al calcular la estadística de prueba chi-cuadrado y compararla con el valor crítico para un nivel de significancia del 5%, se encuentra que la estadística de prueba no supera el valor crítico. Por lo tanto, no hay suficiente evidencia para rechazar la hipótesis nula. En consecuencia, se acepta la hipótesis nula y se concluye que la cantidad de pedidos por pedido podría seguir un comportamiento

aproximadamente de Poisson.

Horas al mes buscando:

Los resultados de la prueba chi-cuadrado indican que el promedio de horas buscando no sigue un comportamiento aproximadamente normal. Esto se deriva de la discrepancia significativa entre los valores observados y esperados en diferentes intervalos de horas buscando. Al calcular la estadística de prueba chi-cuadrado y compararla con el valor crítico para un nivel de significancia del 5%, se encuentra que la estadística de prueba es considerablemente mayor que el valor crítico. Por lo tanto, hay suficiente evidencia para rechazar la hipótesis nula. En consecuencia, se concluye que el promedio de horas buscando no sigue un comportamiento aproximadamente normal.

Dinero mensual gastado:

Los resultados de la prueba chi-cuadrado indican que el dinero mensual gastado no sigue un comportamiento aproximadamente normal. Esto se basa en la discrepancia significativa entre los valores observados y esperados en los diferentes intervalos de dinero gastado. Al calcular la estadística de prueba chi-cuadrado y compararla con el valor crítico para un nivel de significancia del 5%, se encuentra que la estadística de prueba es considerablemente mayor que el valor crítico. Por lo tanto, hay suficiente evidencia para rechazar la hipótesis nula. En consecuencia, se concluye que el dinero mensual gastado no sigue un comportamiento aproximadamente normal.

Compra mas cara:

Los resultados de la prueba chi-cuadrado sugieren que la compra más cara no sigue un comportamiento aproximadamente normal. Esto se deriva de la discrepancia significativa entre los valores observados y esperados en los diferentes rangos de precios de compra. Al calcular la estadística de prueba chi-cuadrado y compararla con el valor crítico para un nivel de significancia del 5%, se encuentra que la estadística de prueba es considerablemente mayor que el valor crítico. Por lo tanto, hay suficiente evidencia para rechazar la hipótesis nula. En conclusión, se rechaza la idea de que la compra más

cara sigue un comportamiento aproximadamente normal.

Regresión Lineal Simple

Los resultados sugieren que, en promedio, por cada paquete enviado, se devuelve aproximadamente 0.16 paquetes. Además, incluso cuando no se envían paquetes, hay un valor base de devoluciones, alrededor de 1.41. Sin embargo, la cantidad de paquetes enviados explica solo alrededor del 9.24% de las devoluciones totales, lo que indica que la relación entre la cantidad de paquetes enviados y la cantidad devuelta es bastante débil.

Regresión Lineal múltiple:

Claro, los resultados indican que las variables analizadas están relacionadas entre sí, pero de manera débil y variable. Esto significa que, aunque las variables como "Compra más cara" y "Artículos devueltos al año" parecen tener algún efecto en el resultado, su influencia no es muy fuerte. Además, parece que los "Problemas con paquetes" no tienen una influencia significativa en el resultado según estos datos. En resumen, el modelo puede explicar solo una pequeña parte de la variabilidad en el resultado, lo que sugiere que otros factores no incluidos en el análisis podrían estar en juego.