



Presentación

Nombres: Wilver M. Yena García, Elizabeth Encarnación Otaño, Raúl Taveras

Matricula: 2024-0903, 2022-1962, 2024-1436

Materia: Analítica Descriptiva I

Maestro: Nerbin Noel Jiménez German

Contenido

Introducción	1
Análisis de los Visitantes Retornantes.	7
Identificando el mes mas activo en caso de Análisis a futuro.	8
ANÁLISIS CON LA CORRELACIÓN DE PEARSON.	9
Conclusión	10



Introducción

El presente informe tiene como finalidad analizar el conjunto de datos **“Online Shoppers Purchasing Intention”**, utilizando herramientas de analítica descriptiva con Python. Los datos se recogen información sobre el comportamiento de los usuarios en sitios de comercio electrónico, permitiendo explorar patrones de navegación, tiempo de permanencia, tasas de rebote, y la intención de compra. Este análisis se llevó a cabo siguiendo las instrucciones académicas del laboratorio de Analítica Descriptiva.



Descripción del Conjunto de Datos

El conjunto de datos "Online Shoppers Purchasing Intention" es una colección de datos relacionados con los patrones de compra y el comportamiento del consumidor en el contexto de las compras en línea. Fue creado mediante encuestas a compradores en línea y la recopilación de datos a partir de sus respuestas.

Estos datos contienen variables como:

- **Administrative:** Páginas visitadas con fines administrativos.
- **Administrative_Duration:** Tiempo total en páginas administrativas.
- **Informational:** Páginas informativas visitadas.
- **Informational_Duration:** Tiempo total en páginas informativas.
- **ProductRelated:** Páginas visitadas relacionadas con productos.
- **ProductRelated_Duration:** Tiempo en páginas de productos.
- **BounceRates** y **ExitRates:** Tasas de rebote y salida.
- **PageValues:** Valor de las páginas vistas antes de la conversión.
- **SpecialDay:** Cercanía de la visita a una fecha especial.

Este conjunto de datos es útil para el análisis del comportamiento del cliente.



Análisis de Duraciones.

Podemos observar que la mayoría de los usuarios pasan menos tiempo dentro de las páginas administrativas, pero en casos extremos pueden elevar la media.

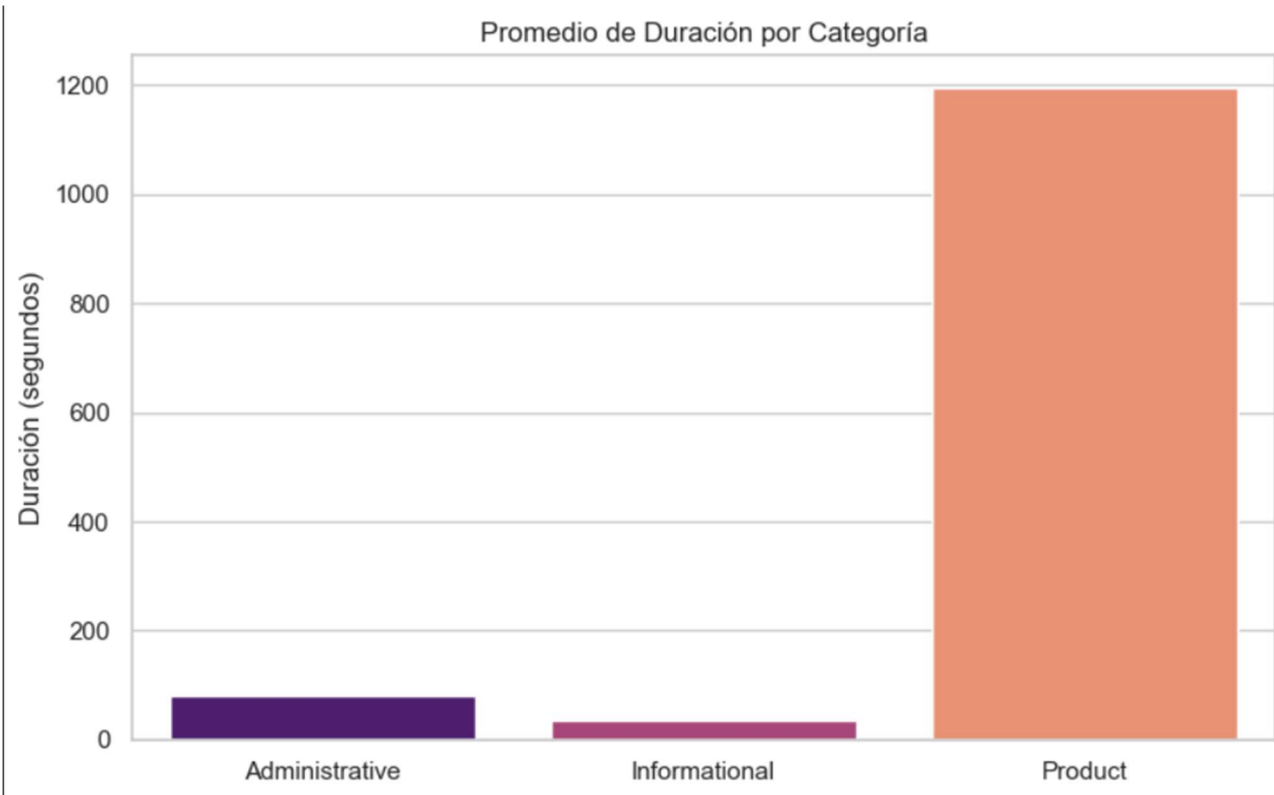
En páginas informativas, los tiempos son más bajos, pero igual ocurre con valores extremos.

Las páginas relacionadas con productos tienen los tiempos más largos en promedio, lo cual tiene sentido considerando que el usuario dedica tiempo a examinar las opciones de compra.

	Administrative_Duration	Informational_Duration	ProductRelated_Duration
count	12330.000000	12330.000000	12330.000000
mean	80.818611	34.472398	1194.746220
std	176.779107	140.749294	1913.669288
min	0.000000	0.000000	0.000000
25%	0.000000	0.000000	184.137500
50%	7.500000	0.000000	598.936905
75%	93.256250	0.000000	1464.157214
max	3398.750000	2549.375000	63973.522230



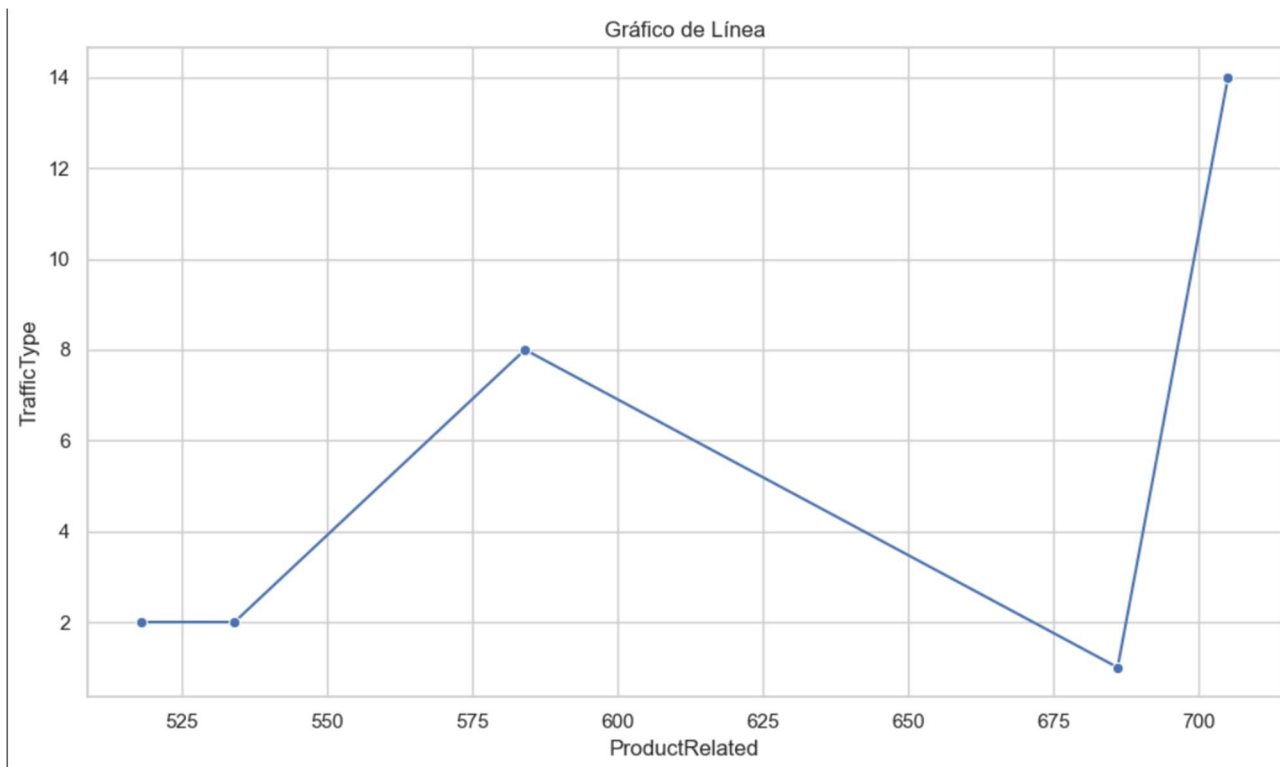
GRÁFICO DEL PROMEDIO DE LAS DURACIONES POR CATEGORÍAS.





Análisis de los 5 productos más visitados.

Para determinar las sesiones con mayor número de páginas visitadas, por la cantidad números relacionadas a nuestro producto.

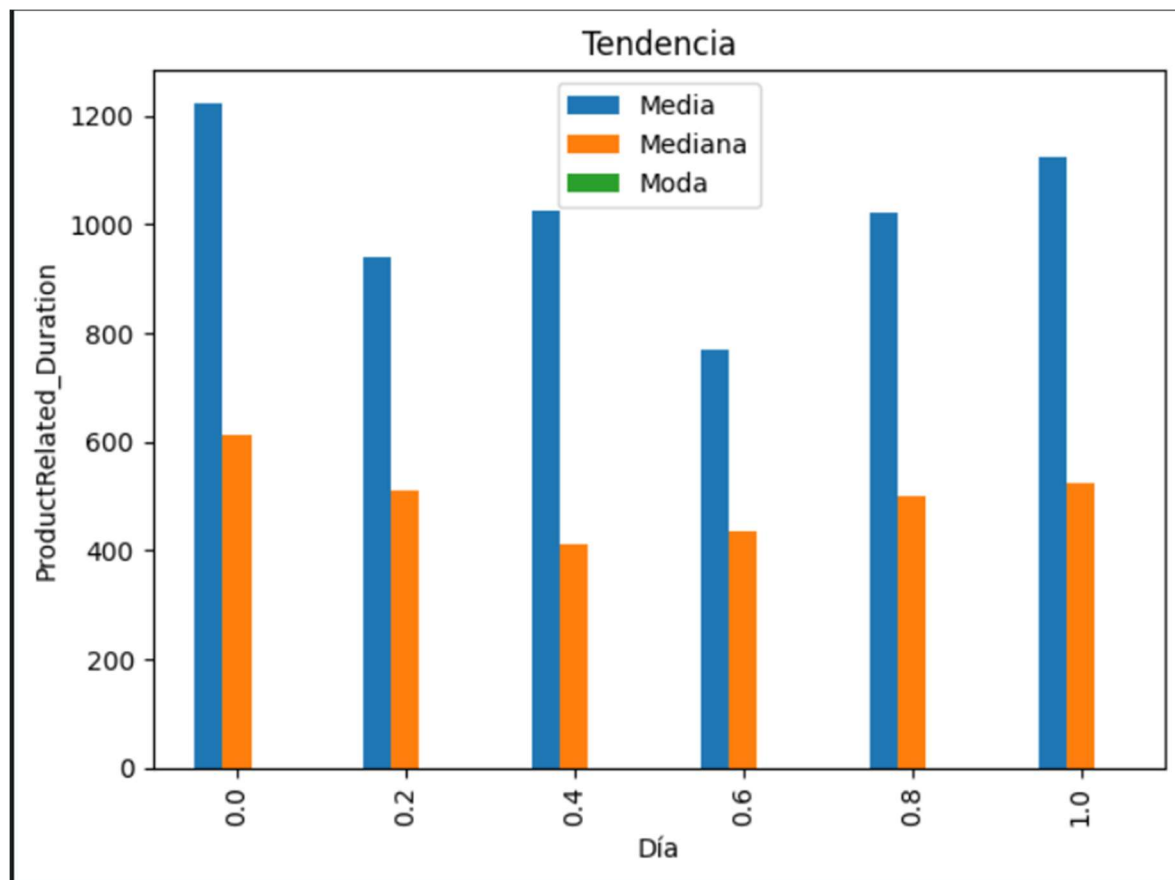


Se observa una relación entre la cantidad de productos visitados y los tipos de tráfico. Por esta razón, se seleccionamos estas dos variables — **ProductRelated** y **TrafficType**— para representar en la gráfica los productos más visitados según el tipo de tráfico. Esto permite identificar patrones de comportamiento de los usuarios en función del tipo de tráfico que los llevó a visitar los productos.



Interpretación cuando se acerca los días especiales.

Cuando se aproxima un día especial, la duración de navegación en páginas de productos aumenta, lo cual podría estar vinculado a compras anticipadas de regalos o búsquedas más extensas.



En el área del código explicamos, porque no tenemos valor en la moda.

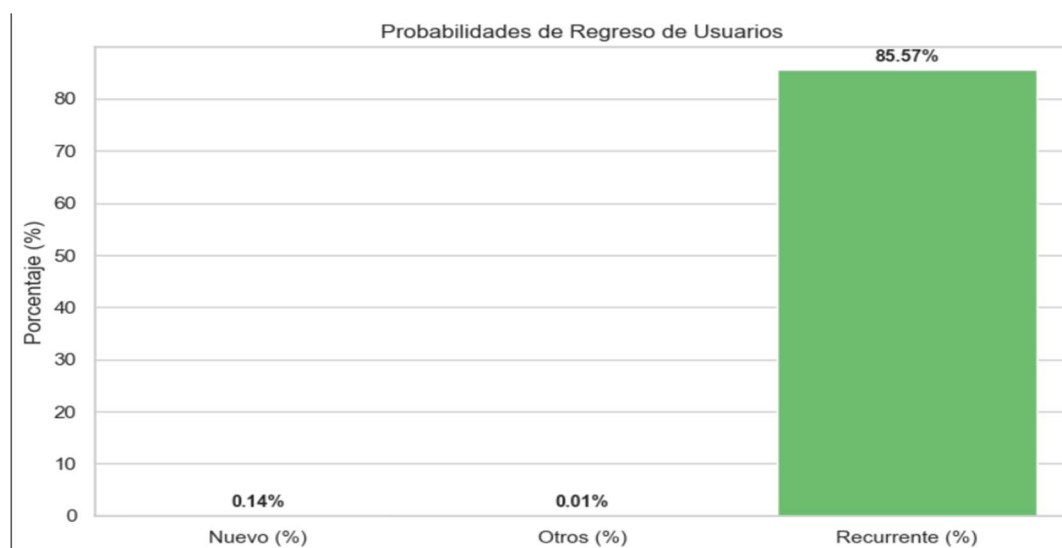


Análisis de los Visitantes Retornantes.

Como podemos observar en la tabla de contingencia, el total de visitantes recurrentes supera significativamente al de los nuevos visitantes, lo que indica que aproximadamente el 85% de los usuarios que interactúan con el sitio web ya son clientes frecuentes. Este dato es sumamente relevante, ya que refleja un nivel alto de fidelización y una base de usuarios consolidada.

TrafficType	1	2	3	4	5	6	7	8	9	10	...	12	13	14	15	16	17	18	19	20	All
VisitorType																					
New_Visitor	37	936	140	99	149	42	5	175	9	22	...	0	6	1	2	1	0	1	1	17	1694
Other	7	11	7	0	1	1	0	1	0	3	...	0	1	0	0	0	0	0	0	52	85
Returning_Visitor	2407	2966	1905	970	110	401	35	167	33	425	...	1	731	12	36	2	1	9	16	129	10551
All	2451	3913	2052	1069	260	444	40	343	42	450	...	1	738	13	38	3	1	10	17	198	12330

GRÁFICO EN PORCENTAJE DEL ANÁLISIS.

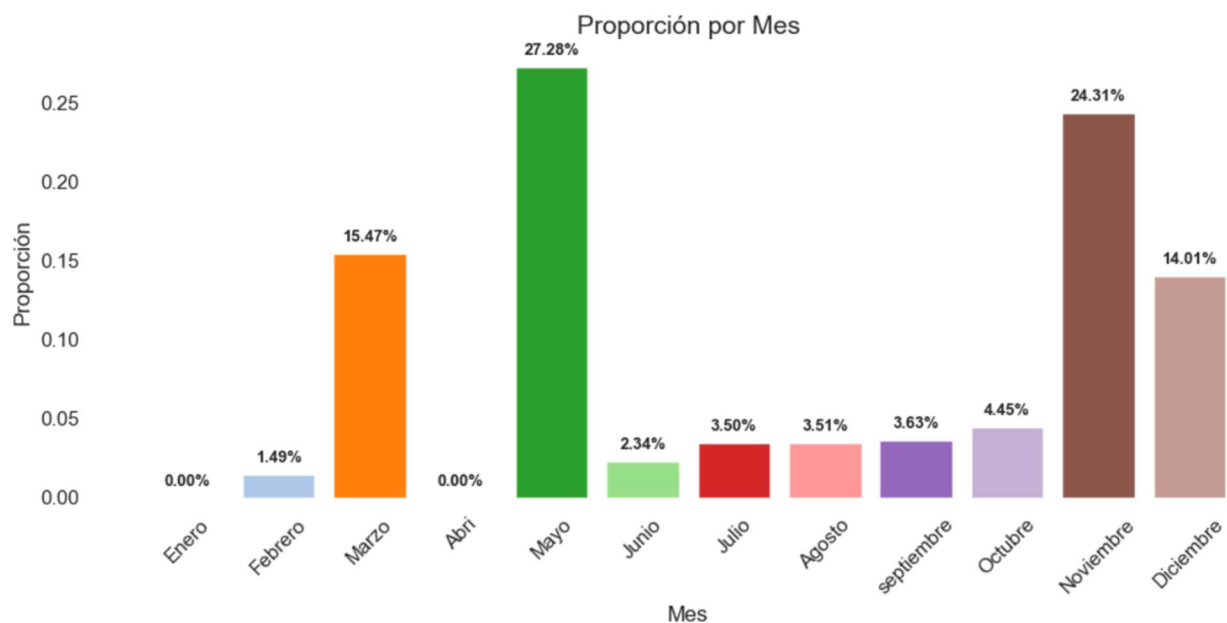




Identificando el mes mas activo en caso de Análisis a futuro.

Aquí tenemos la observación donde nos damos a entender que el mes más activo es **Mayo** teniendo un **27%** y le sigue **Noviembre** con un **24%**, en caso de que queramos implementar algunas nuevas ideas seria considerable para estos meses para facilitar si es buena o mala la decisión tomada.

GRAFICA PARA IDENTIFICAR EL MES MÁS ACTIVO.

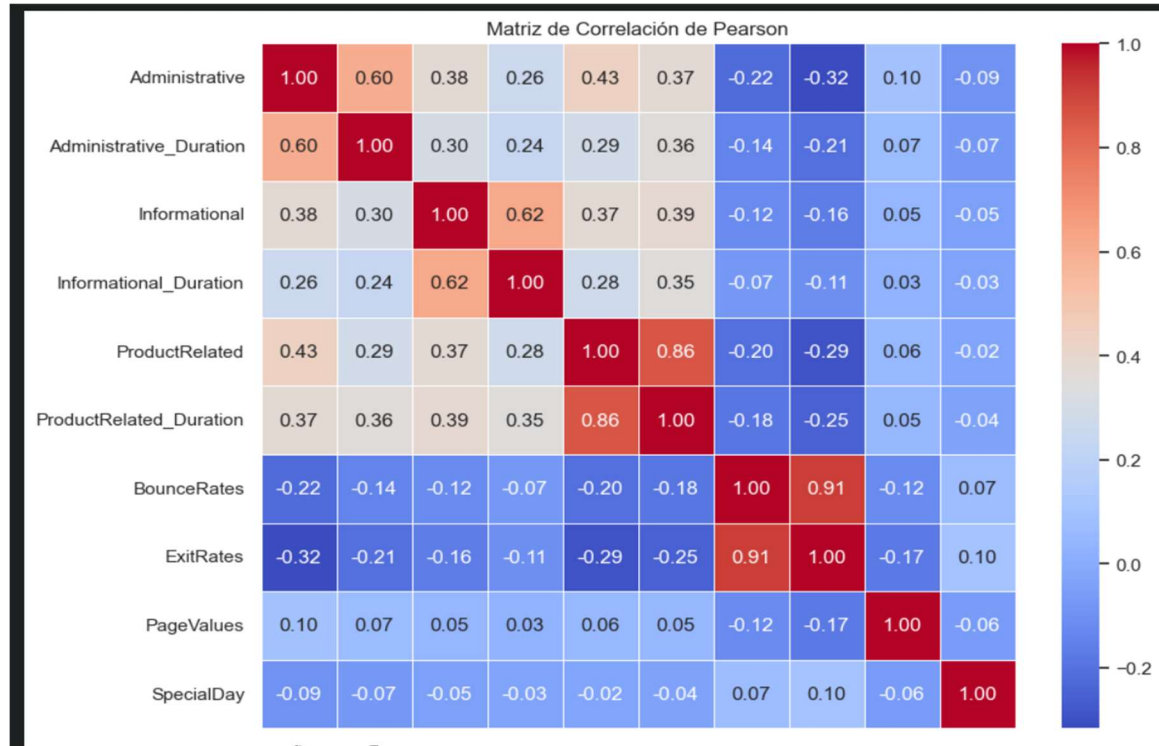




ANÁLISIS CON LA CORRELACIÓN DE PEARSON.

Aplicando la correlación de Pearson entre las variables numéricas que había seleccionado de nuestros datos, las relaciones que vemos en la gráfica que se identificaron fueron clave. Un ejemplo simple es la relación fuerte que existe entre la variable **Product_Related** y **Product_Related_Duration**, lo que nos hace entender que, a mayor cantidad de páginas de producto, mayor es el tiempo de navegación en dichas páginas.

GRÁFICA DE MATRIZ DE CORRELACIÓN.





Conclusión

A través de este análisis, nuestro equipo logró identificar elementos clave del comportamiento de los usuarios en plataformas del comercio electrónico, lo que nos permitió comprender mejor los factores que pueden influir directamente en su intención de compra. Al examinar con detenimiento variables como la duración de las visitas en páginas de producto, la cercanía a fechas especiales (como festividades o eventos promocionales) y la frecuencia con la que un usuario regresa al sitio, obtuvimos una perspectiva más clara sobre cómo se configuran los patrones de navegación con miras a una posible conversión.

La aplicación de técnicas de análisis descriptivo, apoyadas en visualizaciones gráficas cuidadosamente seleccionadas, nos permitió no solo resumir tendencias generales, sino también detectar comportamientos específicos que podrían pasar desapercibidos con un enfoque superficial. Esto fue clave para lograr una interpretación integral de los datos y para generar hipótesis valiosas de cara a futuros estudios o aplicaciones prácticas en estrategias de marketing digital.