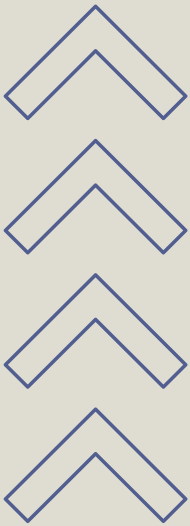




PROYECTO GRUPAL - HENRY

GRUPO 6 - SEMANA 3

E-COMMERCE - DATASET BY OLIST




Julio 2022





ÍNDICE

• Equipo	1
• Introducción	2
• Objetivo	2
• Presentación de KPI's	3
• Presentación de KPI's	4
• Presentación de KPI's	5
• Presentación de KPI's	6
• Presentación de KPI's	7
• Organización	8
• Conclusiones	9





EQUIPO DE TRABAJO



LEONARDO ARGAÑARAS

Periodista, estudiante de Data Science
Product Manager | Data Analyst



JOAN DEMIAN GODOY

Licenciado y estudiante de Data Science.
Ingeniero de datos del proyecto.



EDGAR MORALES

Físico y Estudiante de Data Science en Henry.
Científico de Datos en el Proyecto.



KAREN BANEAS

Ingeniera en Sistemas. Estudiante de Data Science.
Ingeniera de datos en el proyecto.



FACUNDO OPPIDO

Estudiante de Ing. en Sistemas, alumno de SoyHenry en la carrera de Data Science.
Ingeniero de datos del proyecto.



INTRODUCCIÓN

A continuación, se presentará el trabajo realizado en la semana 3 del proyecto Olist. Donde como equipo de trabajo, nos organizamos para poder realizar los reportes de los KPI's propuestos en su momento y de la implementación de un modelo de Machine Learning.

OBJETIVO

El objetivo de esta semana fue lograr llevar a cabo la conexión de la base de datos creada en la semana 2 del proyecto con la herramienta de visualización PowerBI, para así poder lograr manipular los datos de las tablas y poder llevar a cabo la ejecución de los KPI's.

Respecto al modelo de Machine Learning:

¿Cuál es la puntuación probable que obtenemos de los clientes?

Buscamos estimar, basándose en los datos sobre el producto y el pedido, cuál será la puntuación de los clientes.

Hipótesis

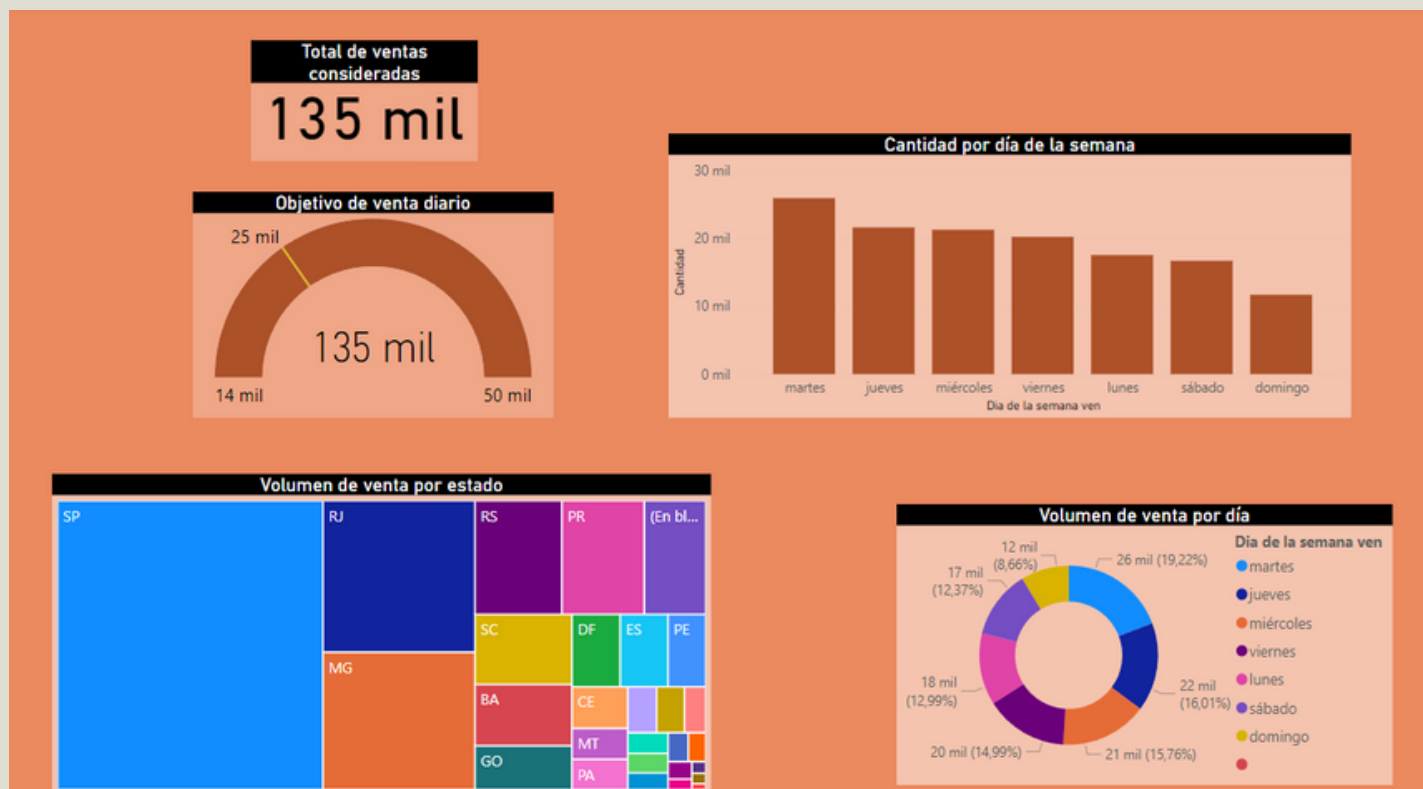
Nuestra hipótesis principal es que el producto y la forma en que se ha realizado el pedido y la entrega pueden influir en la puntuación de las reseñas de los clientes.

Procedimiento

Para responder a la pregunta debemos cargar los datos de cada pedido hasta la fase de entrega. Con eso, debemos implementar un modelo que estime cuál será la puntuación dada por el cliente en la fase de revisión.

PRESENTACION DE KPI'S

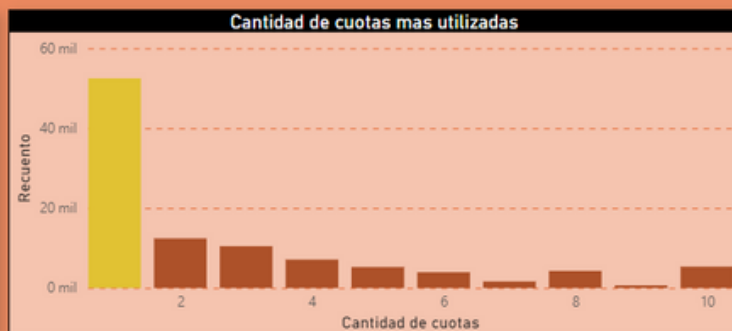
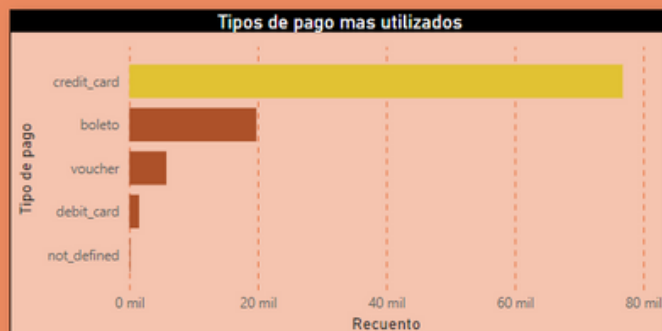
VOLUMEN DE VENTA POR DÍA DE LA SEMANA



Esta métrica permite conocer cuáles son los días con mayores ventas, por lo que se puede aplicar un **KPI para generar un aumento de las ventas**, aplicando acciones de marketing, en las jornadas en donde disminuyen las compras.

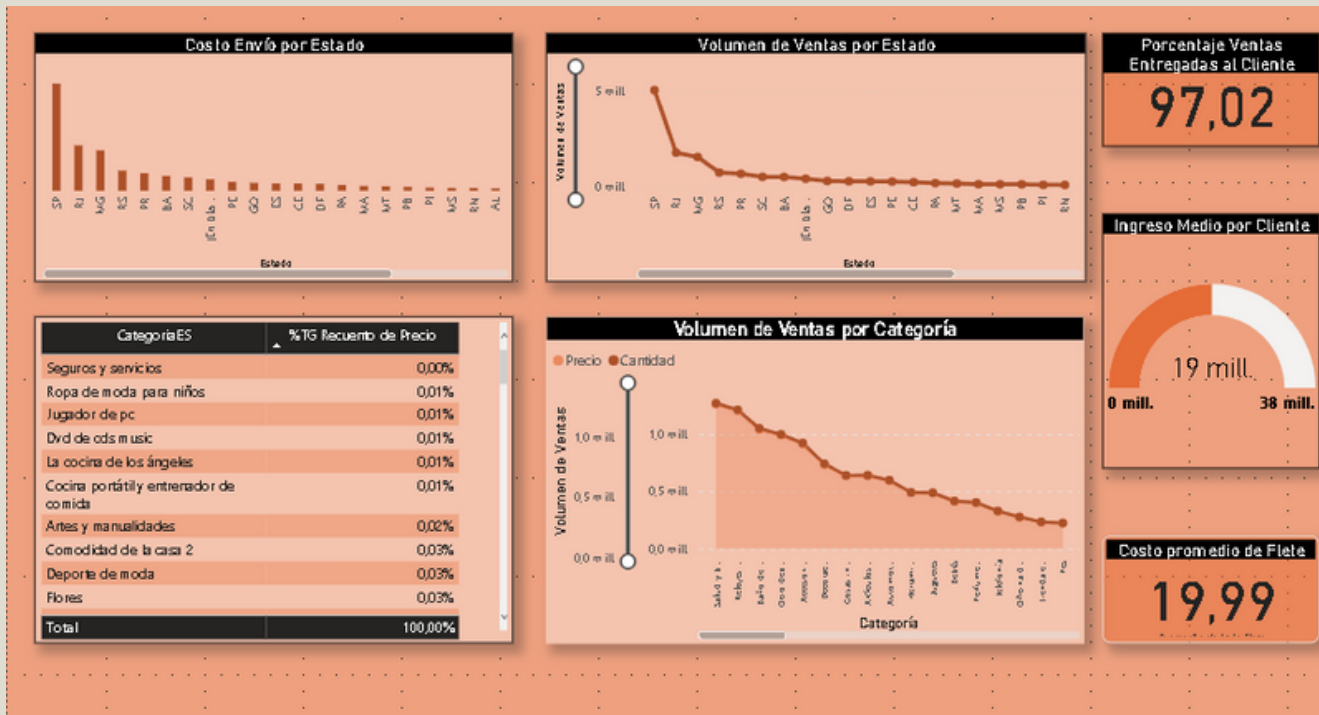
Además, se puede observar cuáles son los Estados con mayores compras también con la intención de implementar acciones para mejorar las ventas en otras regiones geográficas con el objetivo de **aumentar el volumen de ventas durante los próximos meses**.

ANÁLISIS DE LOS PAGOS DE LOS CLIENTES



En este Dashboard se puede analizar de forma dinámica cuales son los medios de pago más utilizados en Brasil. En color amarillo se resalta la tarjeta de crédito y el pago en una cuota, que son el medio que el mercado brasileño espera tener disponible. Adicionalmente, y con el mismo medio de pago, son frecuentemente utilizadas las cantidades de cuotas que figuran a la derecha. Se agregan en la parte de abajo los diez estados y las diez categorías con mayor volumen de venta, para focalizar alguna estrategia comercial, de ser necesario.

ANÁLISIS DE LOS ENVÍOS



Con el objetivo de ayudar al negocio a tener un panorama claro de su eficiencia en el área de ventas tanto en volumen de ventas como eficiencia de entrega de sus productos.

El ingreso medio por cliente es muy importante refleja lo que nuestros clientes están gastando tanto por estado como por categoría así:

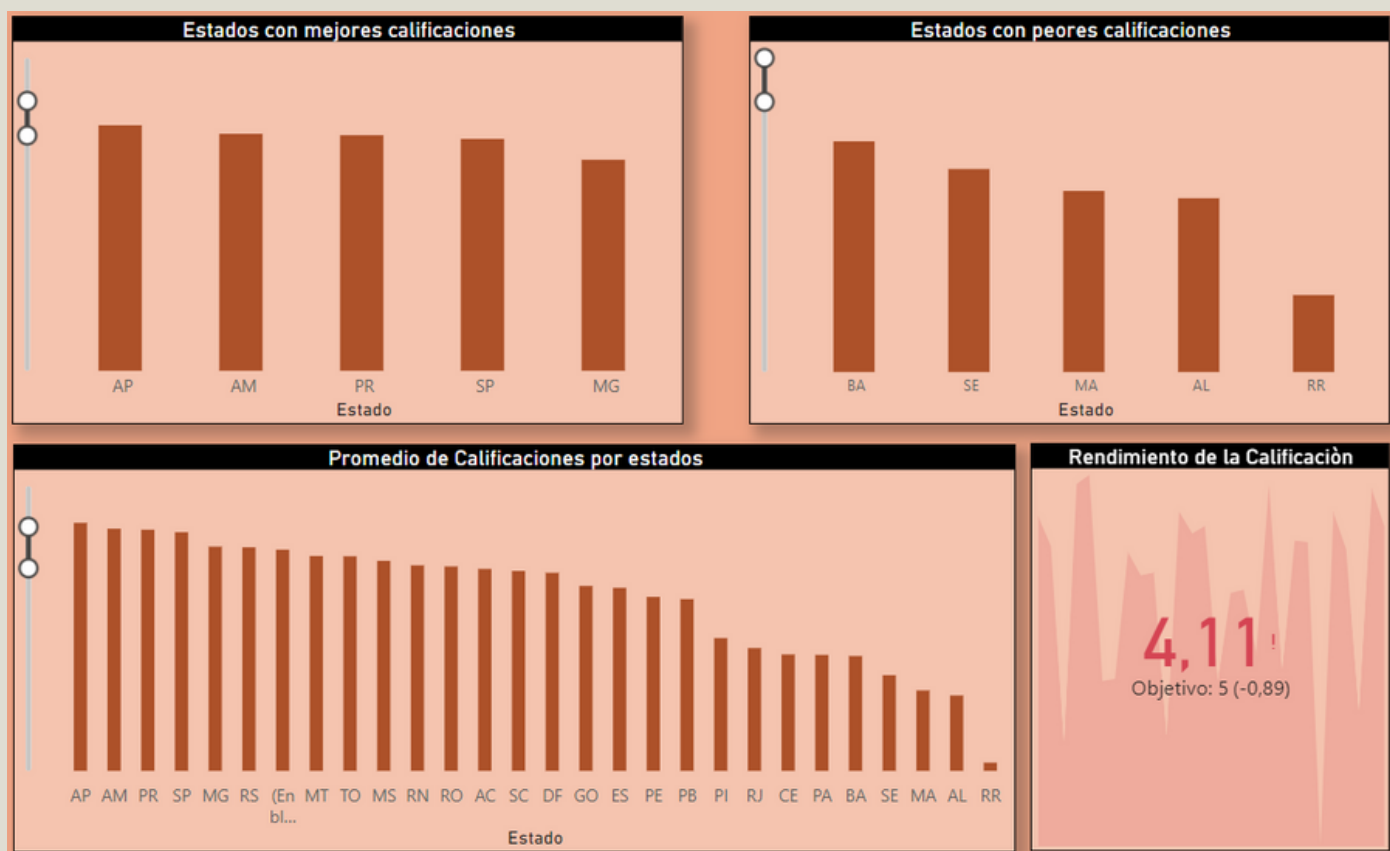
- Ingreso medio por cliente
- El costo promedio de flete
- Eficiencia de ventas de acuerdo al porcentaje de ventas ya entregadas al cliente
-

Todos estos datos pueden filtrarse por estado y categoría

Tenemos una visual mediante graficos :

- Costo de envío y volumen de ventas por estado
- Porcentaje de movimientos de ventas y volumen de las mismas por categoría

ANÁLISIS DE LAS RESEÑAS



Los gráficos presentados nos muestran el promedio de las reseñas que tuvieron los clientes por estado sobre los diferentes tipos de ordenes realizadas. En base a esta información realizamos un KPI que tiene como objetivo la mediana de las calificaciones, la cual es 5, y tiene en cuenta el promedio de las calificaciones por estado.

De esta forma logramos ver que tan bien o mal estamos teniendo en cuenta las calificaciones dadas por los clientes en relación a los estados.

MAPAS DE BRASIL

A continuación se encuentran cuatro mapas de Brasil con la siguiente información:

- Heatmap de los clientes
- Heatmap de los vendedores
- Ubicación de los clientes
- Ubicación de los vendedores

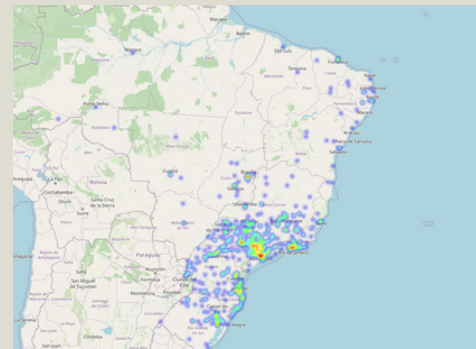
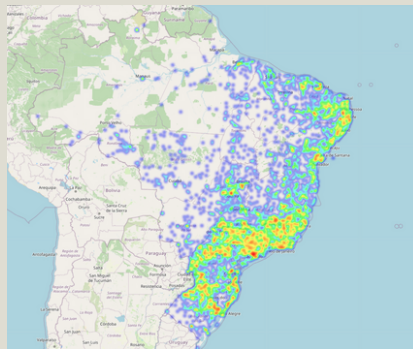
Los heatmaps y los mapas de ubicaciones, muestran la misma información respectivamente para vendedores y clientes, para diferentes tipos de análisis convendrá volcarse por uno u otro.

Si bien, esta información deberá ser analizada en profundidad por personal del área logística de la empresa, de forma preliminar se puede apreciar el desbalance que existe en las concentraciones de clientes y vendedores, quedando despoblada la parte noreste del país de estos últimos, pero notando una distribución más uniforme de los primeros.

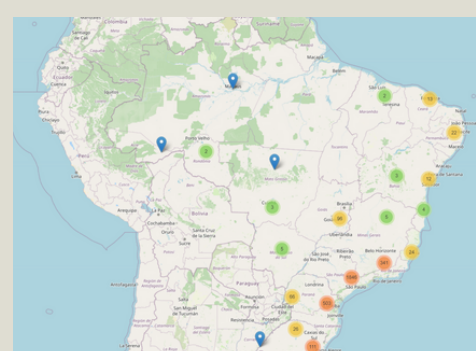
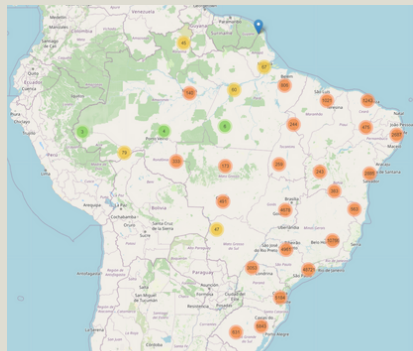
Clientes

Vendedores

Heatmaps

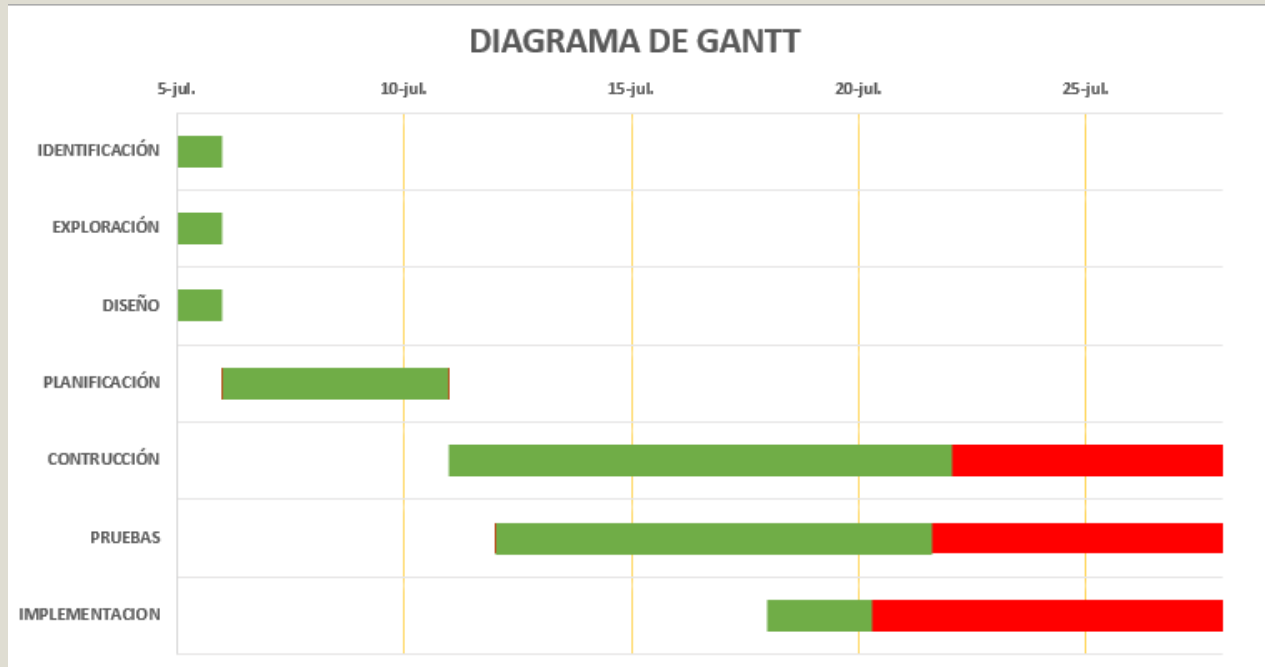


Ubicaciones



ORGANIZACIÓN

DIAGRAMA DE GANTT



TRELLO





>>>> CONCLUSIONES

De la Hipótesis principal, la entrega del producto es el factor que más influye en la puntuación de las reseñas de los clientes. Pero para hacer más notorio este hecho generamos nuevas características en nuestro conjunto de datos, en particular, se otuvo que la puntuación asignada por el cliente es inversamente proporcional al retraso de la entrega.

Por otro lado, Random Forest fue el modelo con mayor precisión, obteniendo un error cuadrático medio de 0.41, que para nuestro objetivo es suficientemente preciso para predecir puntuaciones a partir de las características de las ventas.