|  |  |  |
| --- | --- | --- |
| **Nombre**: Kathya, Gerardo, Guillermo, Emiliano | | **Matrícula**: |
| **Nombre del curso:** Ingeniería de datos masivos | **Nombre del profesor**:  **Miguel de Jesús Martínez Felipe** | |
| **Fecha:** | Evidencia 1 | |

Avance 1 Evidencia 1

<https://www.canva.com/design/DAFzo8c7n2o/s-QtnUDwa8JJAYkXqi958Q/edit?utm_content=DAFzo8c7n2o&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton>

Instrucciones Avance 2

7. Una vez realizado lo anterior, accede nuevamente a los datos de Instacart obtenidos en el avance anterior.

8. Utiliza un Notebook de Jupyter para desarrollar lo siguiente:

Llevar a cabo un análisis exploratorio de los datos con la finalidad de entender y localizar hallazgos realizando al menos 3 diferentes visualizaciones, las cuales tiene como punto de partida, hacerse preguntas como:

· ¿Qué productos se venden más?

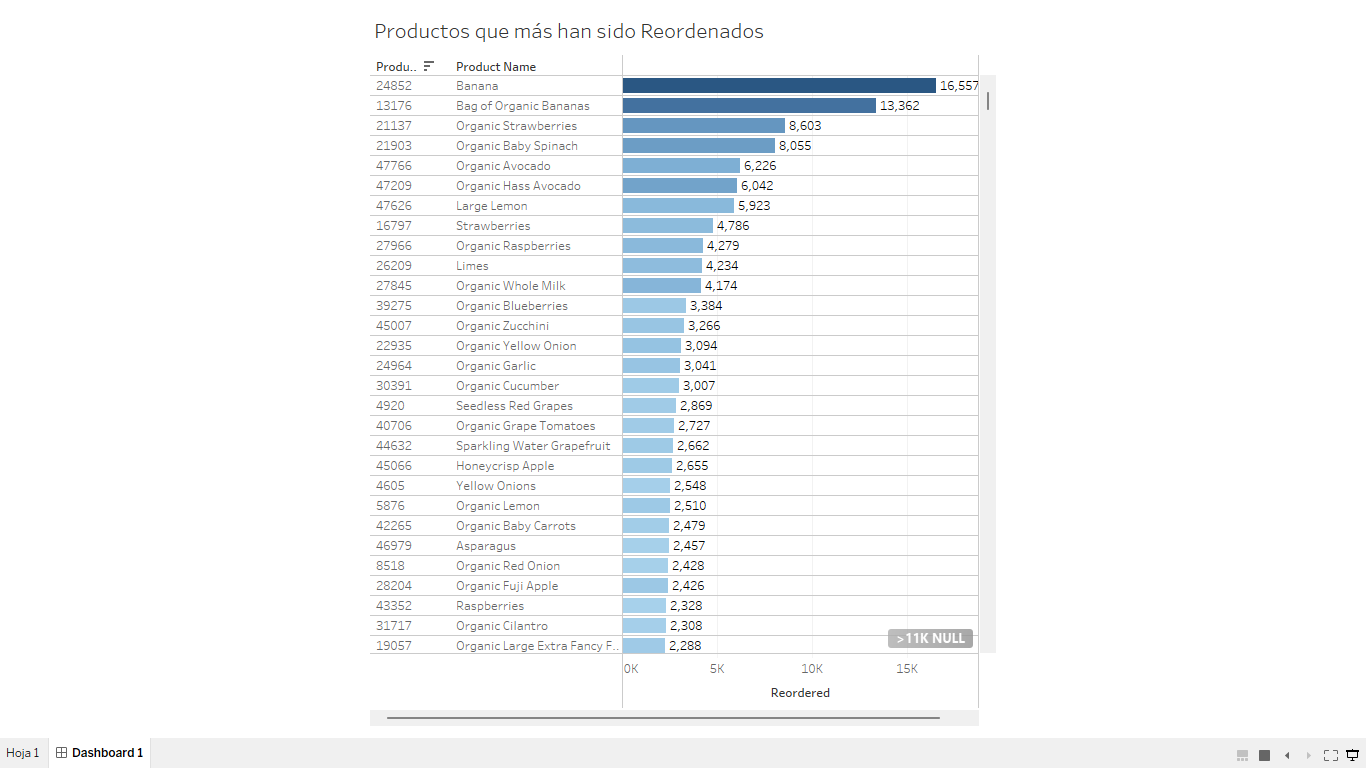
· ¿Qué productos se venden por día y por hora?

9. Incluye en tu archivo los comentarios con la explicación de tu código y hallazgos.

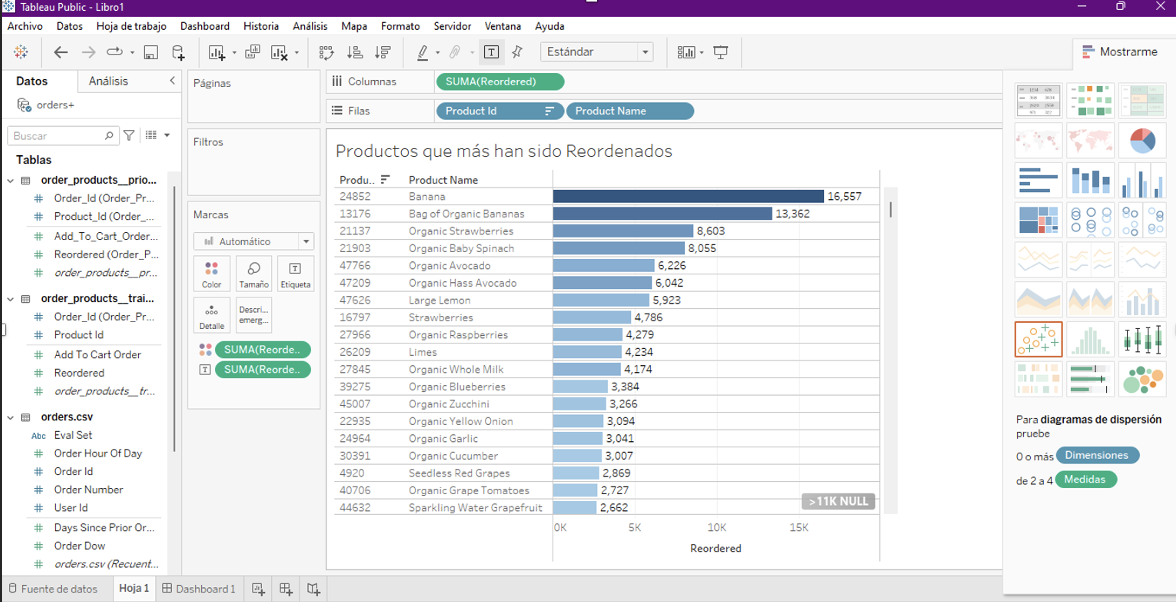
10. Haz entrega de los archivos electrónicos con las actividades realizadas de los temas 3 y 4.

***Visualización 1***

Imagen visualización



Captura de pantalla del código para generarlo.



Explicación del código e interpretación de gráfica

Para la creación de esta gráfica se utilizó como filas los id de los productos y el nombre de cada uno de los productos para poder identificarlos con mayor facilidad al momento de ver la gráfica, y para las columnas se hizo la suma de cuantas veces se repitió el valor de 1 en Reordered (donde los valores son 0 y 1), de esta forma poder observar mediante la gráfica cuales son los productos que más se venden, lo que podría permitir a Instacart tomar decisiones basadas en los productos con mayor número de ventas, como es ofrecer mas promociones a esos productos, o darles mayor importancia en la forma en la que se acomodan los productos a la vista de los clientes, de igual forma permite a la empresa conocer que productos tiene mayor demanda y así poder llevar un control más vigoroso del inventario en esos productos.

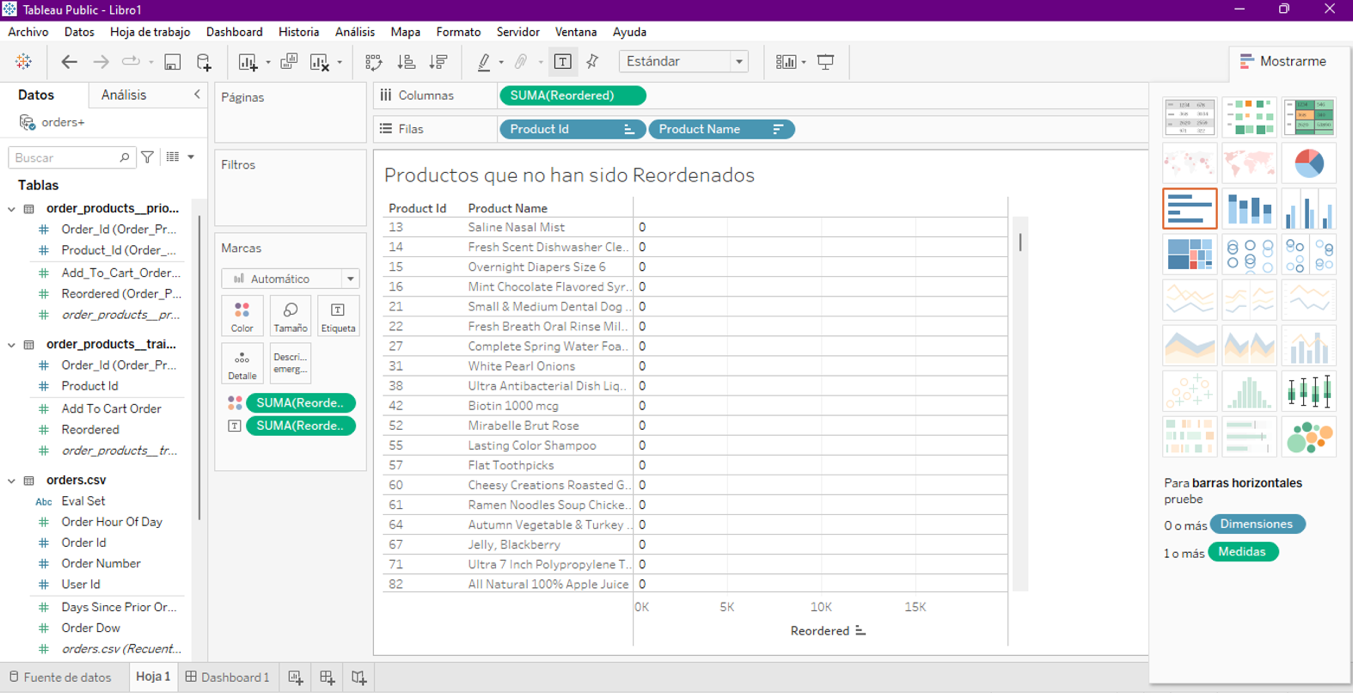
***Visualización 2***

Imagen visualización

Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente

Captura de pantalla del código para generarlo.



Explicación del código e interpretación de gráfica

Para la creación de esta gráfica se utilizó como filas los id de los productos y el nombre de cada uno de los productos para poder identificarlos con mayor facilidad al momento de ver la gráfica, y para las columnas se hizo la suma de cuantas veces se repitió el valor de 1 en Reordered (donde los valores son 0 y 1) pero se invirtieron los valores en su acomodo, de forma que ahora se pueden observar que productos se venden poco o no se venden, esto permite a Instacart tomar decisiones basadas en los productos que no se venden o que tienen poca demanda, de esta forma la empresa podría elegir dejar de vender esos productos o hacer el inventario de esos productos con menor frecuencia y así poder ahorrar dinero, o buscar dar promociones con productos con altas ventas y poder equilibrar las ventas.

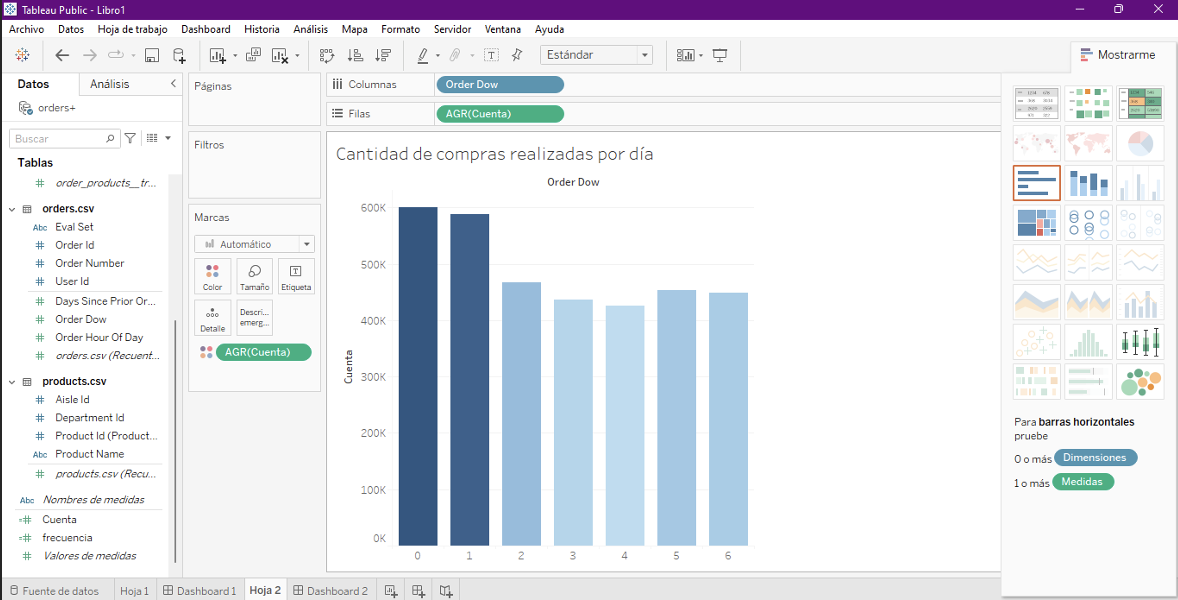
***Visualización 3***

Imagen visualización

Gráfico, Gráfico de barras

Descripción generada automáticamente

Captura de pantalla del código para generarlo.



Explicación del código e interpretación de gráfica

Para el valor de las columnas se usó el valor de ORDER DOW el cual señala en que día de la semana en que se realizó una orden. Y para las columnas se creó una función para realizar la cuenta de cuantas veces se repiten los valores.

Conocer los días en que más se venden productos y en los que menos se venden podría permitir a la empresa realizar estrategias de venta, como podría ser ofrecer promociones los días que más venden y aumentar más las ventas, o dar las promociones en los días con menos ventas y poder así aumentar las ventas en esos días.

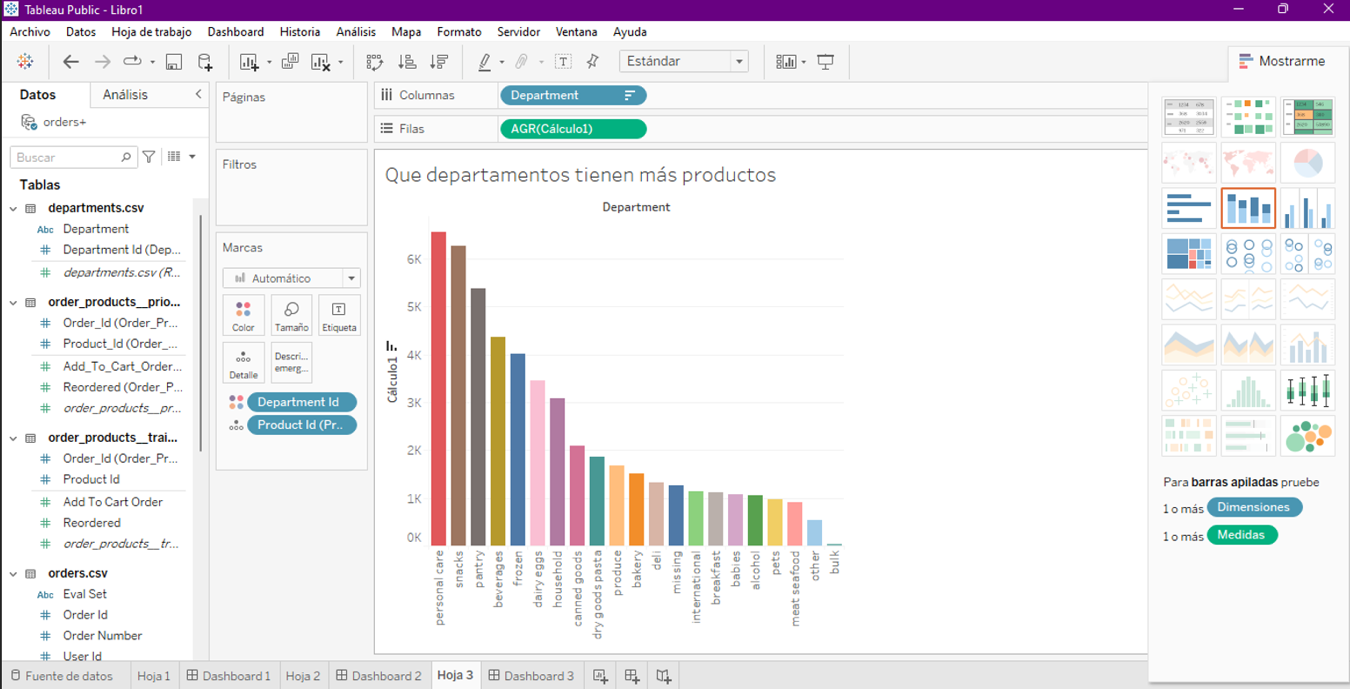
***Visualización 4***

Imagen visualización

Gráfico, Gráfico de barras

Descripción generada automáticamente

Captura de pantalla del código para generarlo.



Explicación del código e interpretación de gráfica

Para crear esta tabla se usó en las columnas los nombres de los departamentos y en las filas se creó una funciona que contara cuantas veces se repite el id de los departamentos en las ordenes registradas.

De esta forma se obtiene una tabla que muestra que departamentos tienen los productos que más se venden dentro de la tienda.

Esta tabla permite conocer que departamentos venden más productos, de forma que la empresa pueda tomar decisiones orientadas a los departamentos y sus ventas, como aumentar los productos de los departamentos, o reducir la cantidad de productos en los departamentos que menos venta

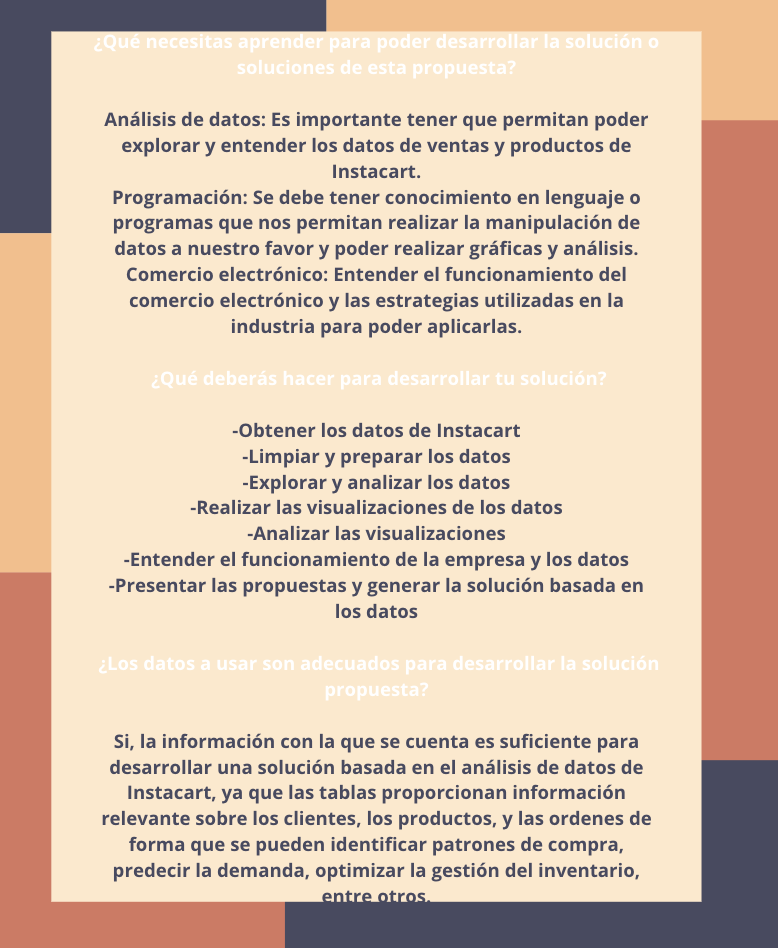
**Entrega final evidencia 1**

1. Utiliza el Notebook de Jupyter con un análisis exploratorio de los datos que creaste en el avance anterior.
2. Prepara una presentación para contar la historia de los hallazgos que encontraste.

<https://www.canva.com/design/DAFzsdcPLpc/s-EsmC_Be7a41GSpYQN5GQ/edit?utm_content=DAFzsdcPLpc&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton>

1. Elaborar una reflexión del proyecto que incluya la respuesta a las siguientes preguntas:

Proyecto



Datos

