ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE INGENIERÍA DE SISTEMAS

INGENIERÍA EN SISTEMAS INFORMÁTICOS Y DE COMPUTACIÓN

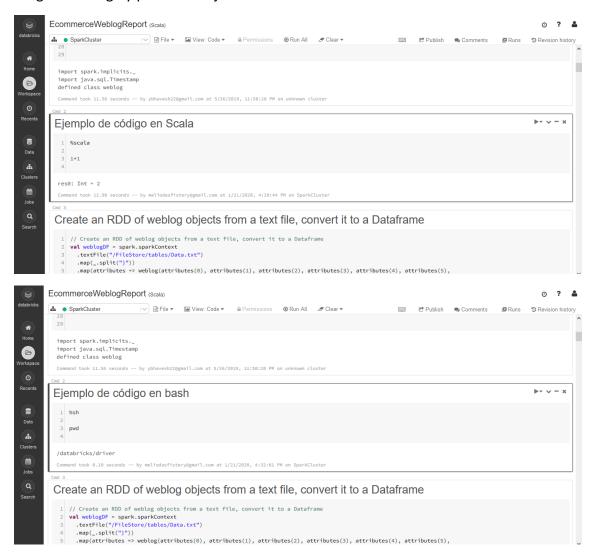
PERÍODO ACADÉMICO: 2019B

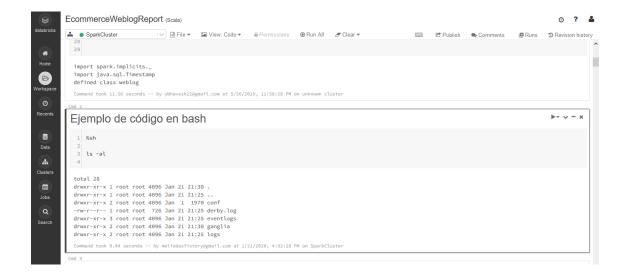
PROFESORA: Cindy López

FECHA DE ENTREGA:21/01/2020

Preparación del entorno

Carga del código y prueba de ejecución en el notebook



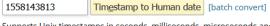


Exploración de los datos



The current Unix epoch time is 1579642915

Convert epoch to human-readable date and vice versa



Supports Unix timestamps in seconds, milliseconds, microseconds and nanoseconds.

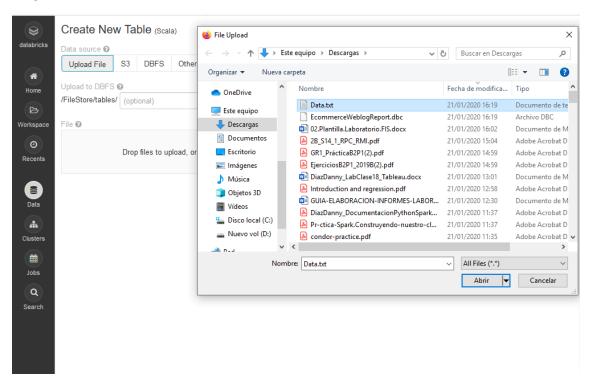
Assuming that this timestamp is in seconds:

GMT : Saturday, May 18, 2019 1:43:33 AM Your time zone: Friday, May 17, 2019 8:43:33 PM GMT-05:00

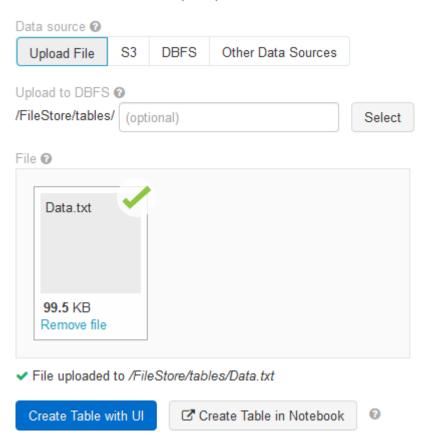
Relative : 8 months ago



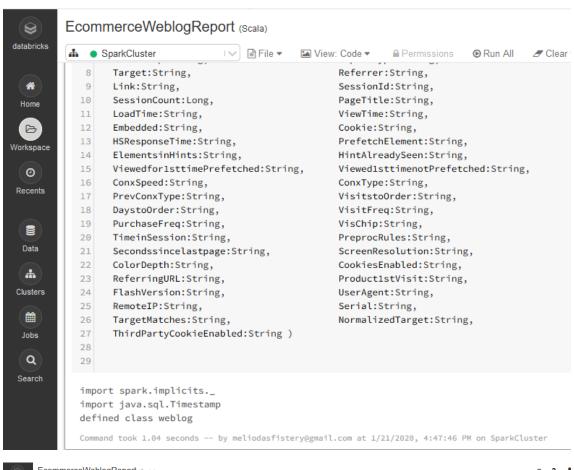
Carga de datos

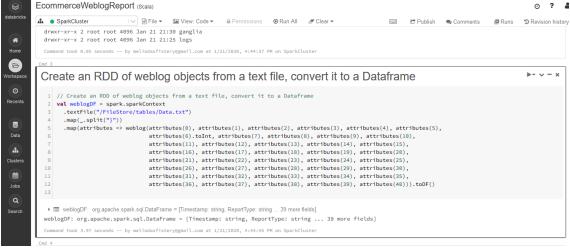


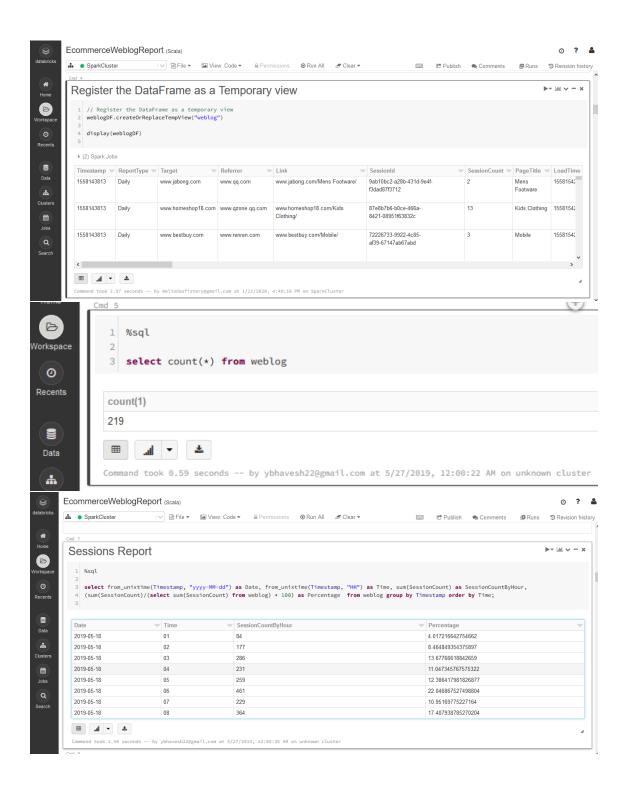
Create New Table (Scala)

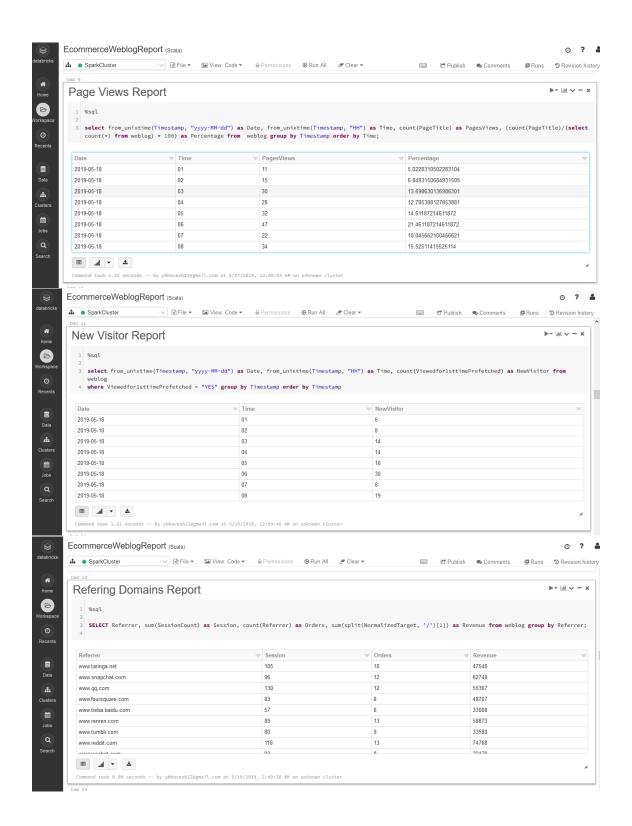


Ejecución del código









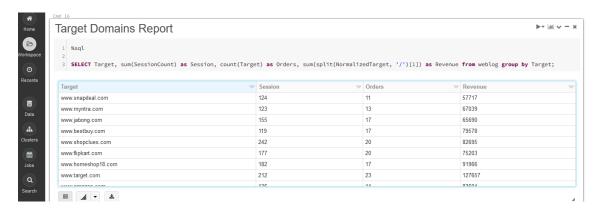
Taller

Compilar, analizar y explicar los siguientes datos de análisis web:

•Informe de dominio de destino

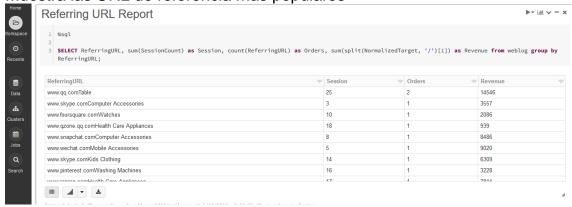
Los dominios de destino son sitios web en los que los usuarios finales compran sus productos.

Extrae datos del dinero generado por la compra en cada sitio web y el número de órdenes por sesión.



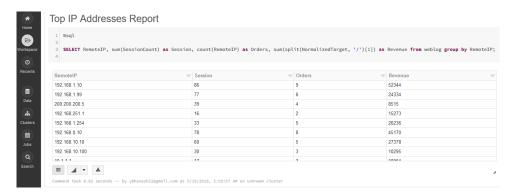
•Reporte de URL de referencia

Muestra las URL de referencia más populares



•Reporte del top de direcciones IP

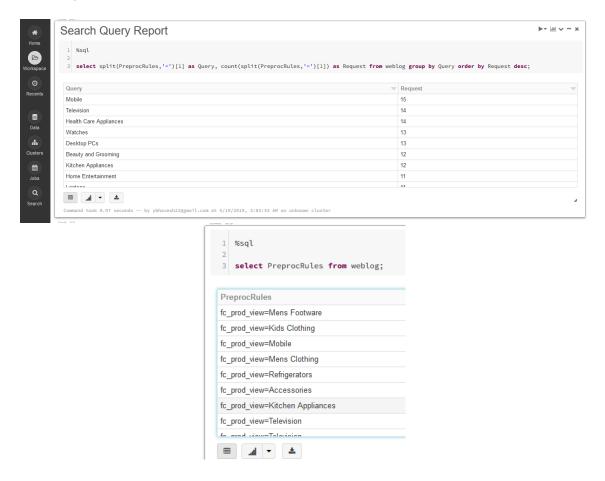
Este informe clasifica las direcciones IP de los visitantes que acceden a su sitio web en términos de cantidad de sesiones. La IP registrada es la ip pública y las IP remotas representan una orden por aparición.



•Reporte de cadenas de búsqueda

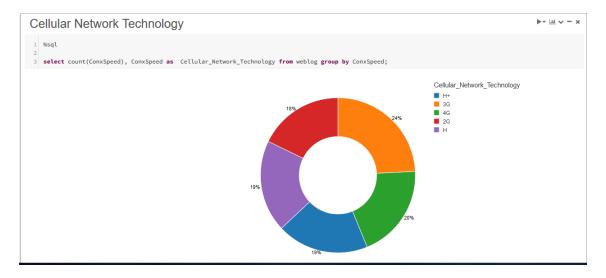
Las consultas de búsqueda son las palabras clave ingresadas en los motores de búsqueda de Internet que proporcionan resultados que dirigen a los usuarios finales a su sitio web.

Este informe muestra las principales consultas de búsqueda que llevaron a los usuarios a su sitio y le permite comparar el número de visitas a la página recibidas por cada consulta de búsqueda.



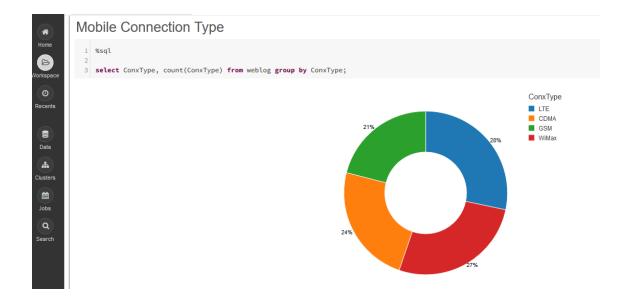
•Tecnología de red celular

Muestra las tecnologías de internet de los celulares que ingresan al sitio web. Todos con porcentajes de uso cercanos.



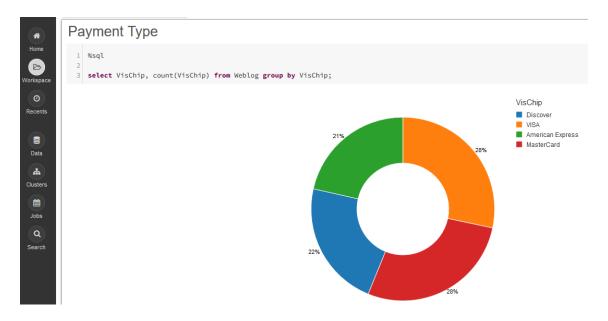
•Tipo de conexión móvil

Indica el tipo de conexión por celular en donde LTE es el más popular de todos.



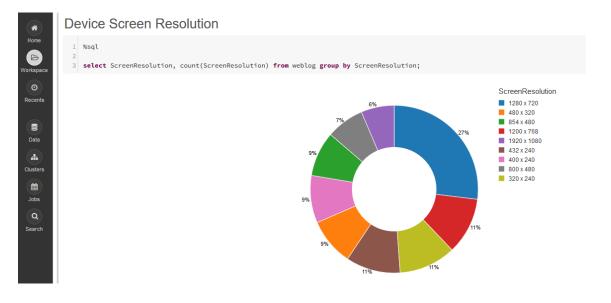
•Tipo de pago

Analiza el tipo de tarjeta usada para realizar el pago, siendo el menos popular Américan Express y el más popular VISA.



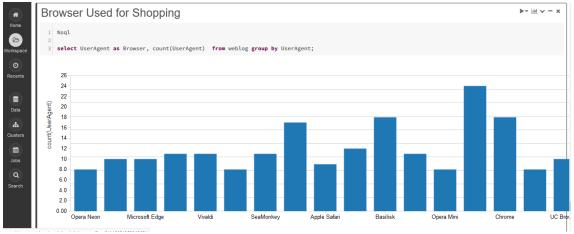
•Resolución de pantalla

Clasifica las resoluciones de pantalla de todos los dispositivos que acceden al sitio web. El más grande es 1920x1080 que puede ser un monitor muy grande hasta una Smart tv. Y el más pequeño 320x240.



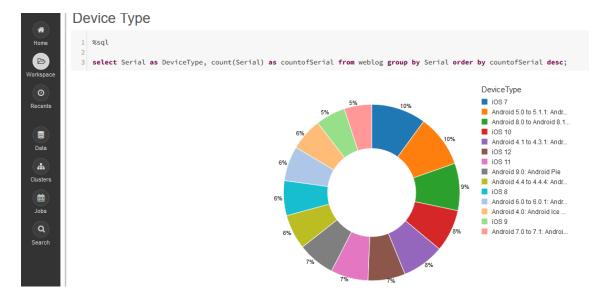
•Browser utilizado para comprar

Obtiene el navegador a partir del metadata de agente de usuario para ver cuál se repite más. Siendo PaleMoon el navegador más usado.



Tipo de dispositivo

Identifica el sistema operativo y la versión de un celular. Habiendo un empate entre IOS 7 y Android 5.0-5.1



Publicar Notebook

El notebook ha sido publicado con una disponibilidad de 6 meses de están visible para todos, debido a la versión gratuita de databriks.

https://databricks-prod-

 $\frac{cloud front. cloud. databricks. com/public/4027ec902e239c93eaaa8714f173bcfc/6414905125001365/259197671035644/3786480379013860/latest. html$