**Parcial 1**

**Diego Ralón y Miguel Suárez**

Introducción

Durante el desarrollo de este trabajo podrán encontrar como conectamos un conteiner de Mysql donde creamos tablas de 3 archivos csv dentro de un Data Base con un sesión de Rstudio con Docker.

Con esta conexión creada cargamos los datos de el conteiner de Mysql a la sesión de R para poder trabajar con ellos.

Con los data frames creados trabajamos un Flexdashboard para nuestro análisis de datos tuviera una mejor visualización.

En este trabajo presentaremos todo el proceso más el análisis de la base de datos proporcionada.

***`1. Conexión de Conteiner Mysql con Rstudio***

Para realizar esta conexión utilizamos Docker, la cual tuvo la función de crear la conexión, crear el conteiner de Mysql y realizar la conexión con el conteiner de la sesión de Rstudio.

* 1. *Creación de la conexión*

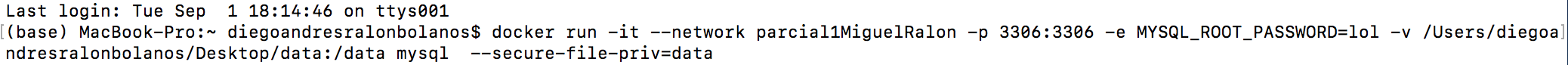
Para la creación de la conexión usamos en nuestra terminal el comando

$ docker network create -d bridge my-bridge-network

Le asignamos el nombre MiguelRalonParcial1.

Esta conexión nos funcionara para conectar los dos conteiner que debemos conectar para poder trabajar con los datos en una sesión de R

* 1. Creación de la base de datos de Mysql



En una nueva terminal corrimos este comando. Al principio del Docker Run asignamos a que conexión corriera nuestro Mysql. También dentro de la creación del Mysql colocamos el path de nuestro data para que estuviera dentro del conteiner. Sin embargo, para que se pueda trabajar la data debemos de crearle tablas donde podamos cargar la data para después exportarlas como un data frame en la sesión de Rstudio.

Captura de pantalla de un celular

Descripción generada automáticamente

Captura de pantalla de un celular con texto

Descripción generada automáticamente

Para la creación de las tablas teníamos dos maneras de realizarlas: Dentro del conteiner en la terminal o con Workbench.

En nuestro caso utilizamos Workbench porque nos facilitaba la manera de trabajar.

* 1. Creación de la conexión con Rstudio

Para la creación de la sesión de Rstudio usamos este código en una nueva terminal:

Con este conteiner corrimos el código: Docker run junto al network

**obedaeg/rstudio-mysql**

***2. Rstudio y flexdashboard:***

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Para trabajar en nuestro Flexdashboards utilizamos estas librerias, que nos ayudan a presentar un trabajo más ordenado y con más produndidad:

* Dplyr: Se utiliza para una mejor manipulación de datos
* DT: Se utiliza para crear tablas interactivas
* Ggplot2: Se utiliza para la creación de gráficas

Nota: También utilizamos la librería de flexdashboards

Captura de pantalla de un celular

Descripción generada automáticamente

Utilizamos un Query para cargar los datos dentro de un data frame usando dbGetQuery. Creamos tres data frames diferentes:

* As: Estadísticas de los videos de academática
* Am: Metadata de los videos de academática
* Av: Información general sobre los videos de academática

Adicional a estos tres data frames nos dimos la libertad de mezclar dos data sets con la función merge.

Para lograr esto renombramos la columna de el data frame As Id a video\_id para hacerla coincidir con la columna del mismo nombre perteneciente a el data set Am.

Con la función merge logramos crear un nuevo data frame llamado asam.

2.1 Análisis del flexdashboard

Al tener los data frames creados no dedicamos a entender la información presentada para poder analizar de manera correcta y poder brindar conclusiónes que brinden un valor a academática.

2.1.1 Primer tab del flexdashboard

Captura de pantalla de un celular

Descripción generada automáticamente

El primer tab del flexdashboard lo utilizamos como un repaso general de lo que estaremos análizando en el resto del trabajo.

Primero creamos un Data Table de as donde podemos entender de donde van a salir los numeros que vamos a utilizar para el resto del análisis.

Nuestro análsis empieza en la parte de abajo. Allí podran enonctrar un resúmen general de los número de academática:

* Visitas totales: Cuantas visitas tiene Academática sumando todos sus videos
* Likes totales: Cuantos likes en total tiene Academática
* Dislikes totales: Cuantos Dislikes totales tiene Academática
* Comentarios totales: Cuantos comentarios

Consideramos que era importante demostrar con la cantidad de datos que tenemos para darle una validez a nuestro análisis, asi como presentar los datos a la persona que le interese conocer los datos de academática.

2.1.2 Segundo tab del flex dashboard

Captura de pantalla de un celular

Descripción generada automáticamente

NOTA: Para tener mejor visualización de las gráficas recomendamos tener el HTML abierto

En este tab realizamos dos gráficas de barras: la cual nos indican datos sobre los mejores videos y los peores videos de Academatica.

La primera gráfica de los mejores videos tomamos como límite que tengan un mínimo de un millón quinientos mil visitas ya que a este punto se le puede considerar a un video viral.

En los cuales encontramos que los siguientes videos tiene esta cantidad o más de vistas:

* Videos de ley de orden
* Ejercicios resueltos por integrales por partes
* Limites de funciones de dos variables
* Introducción a circuitos eléctricos
* Integrales de línea
* Funciones de varias variables
* Ejercicios resueltos derivadas parciales
* Ejemplos integrales por sustitución trigonométrica
* Ejemplos de integrales por fracciones parciales
* Demostración de límites por definición delta eplison
* Curvas de nivel
* Calculo de la resistencia para un LED.

La segunda gráfica sobre los peores videos tomamos como límite que no tengan más de quinientas vistas porque consideramos que por el alcance que tiene YouTube, quinientos es un número muy bajo.

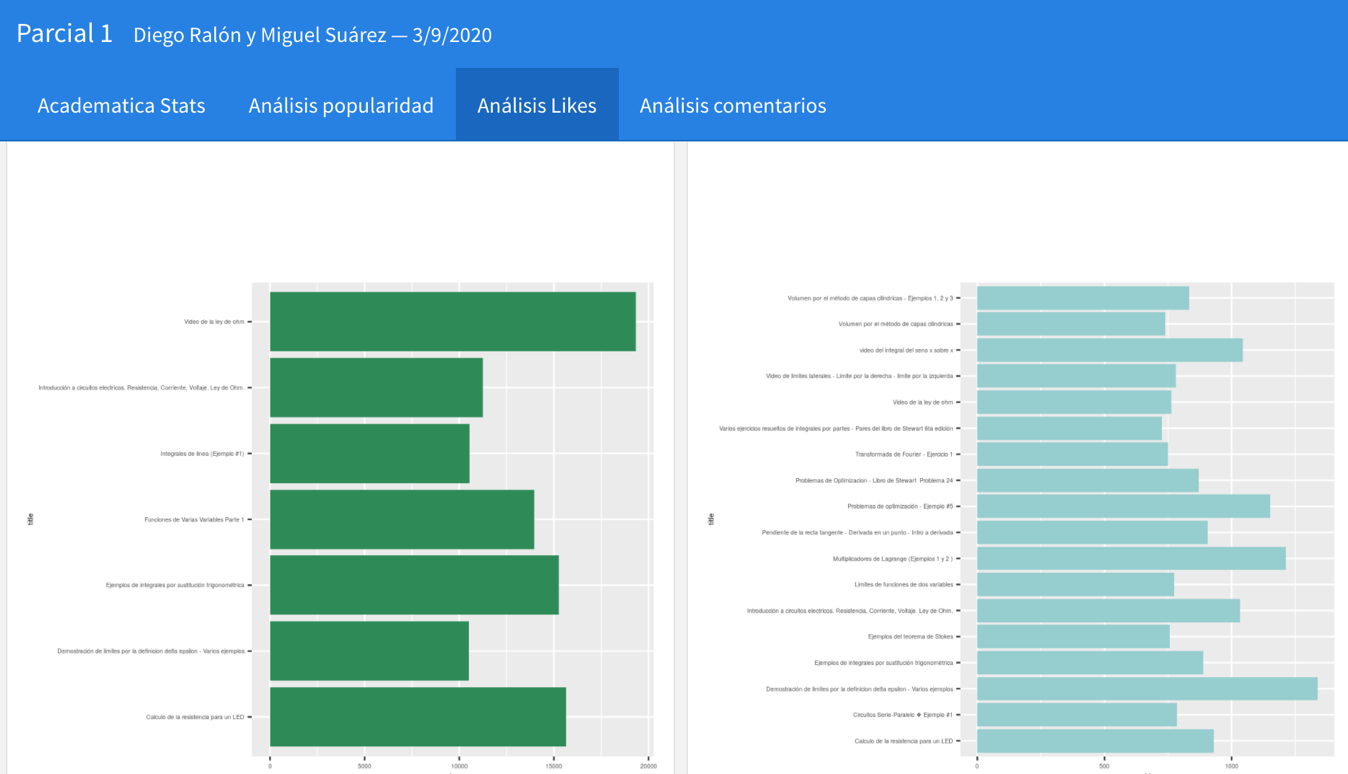
En los cuales encontramos que los siguientes videos cuentan con menos de quinientas visitas:

* Tutorial 2 python
* Strongvpn
* Recetalibre
* Que portero
* Proyecto final UFM
* Probabilidad condicional
* Introducción al curso de fiabilidad
* Flex dash board
* Como instalar Python
* Que es y para que sirve la fiabilidad

Con este análisis podemos concluir que los videos con una orientación magistral enfocados en la matemática son los más exitosos en visitas para Academática. Mientras que los menos visitados por los títulos suponemos que son videos de muchos años atrás y sin ningún tipo de orientación específica.

Por lo cual llegamos a nuestra primera conclusión que es realizar más videos con orientación matemática y magistral ya que estos son los que más se viralizaron en la plataforma.

2.1.3 Tercer tab de flex dashboard



En ese tab analizamos por medio de gráficas de barras los likes y dislikes que a tenido academática en sus videos.

La primera gráfica nos presenta los videos más likeados de academática con un límite de cien mil likes.

Dentro de los cuales podemos encontrar:

* Video de ley de ohm
* Introducción a circuitos eléctricos
* Integrales de línea
* Funciones de varias variables
* Ejemplos de integrales por sustitución trigonométrica
* Demostración de límites por la definición delta épsilon
* Cálculo de resistencia de un LED

En la segunda gráfica nos presenta los videos con más dislikes de academática con un límite de 700 dislikes.

Los cuales los que más destacan son:

* Video del integral del seno x sobre x
* Problemas de optimización
* Multiplicador de Langrange
* Demostración por límites por la definición de delta elipson

Este caso nos pareció bastante interesante ya que dentro del rango de likes y dislikes todos los videos ahora si tiene un enfoque matemático y magistral.

Sin embargo, muchos de los casos con más dislikes no quiere decir que el video no sea aceptado por la comunidad, ya que en el caso de Delta Elipson este también es de los más likeados y de los más vistos del canal.

Muchas veces cuando los videos tienen un alance muy grande es mucho más probable que tenga más dislikes, por lo cual este no es una indicador muy confiable de la calidad del video.

Nuestra recomendación en este caso es que siempre tenemos que tener en cuenta el alcance del video antes de tomar una decisión sobre el contenido a base de los dislikes de la gente.

2.1.4 Análisis comentarios

Una captura de pantalla de una red social

Descripción generada automáticamente

En este tab análisamos la cantidad de comentarios que tienen los videos de academática.

En la primera gráfica podemos encontrar los videos más comentados de academática con un límite de más de 700 comentarios

Los cuales son:

* Ley de ohm
* Ejercicios resueltos de integrales por partes
* Teorema de Stokes
* Problemas de optimización
* Limites de funciones de dos variables
* Introducción a circuitos eléctricos
* Integrales de línea
* Funciones de varias variables
* Ejemplos de integrales por sustitución trigonométrica
* Demostración de limites por definición de delta elipson
* Calculo de resistencia para un led

En la segunda gráfica podemos encontrar los videos con menos comentarios con un límite de menos de 8 comentarios.

En este caso tenemos una cantidad muy extensa de comentarios con menos de 8 comentarios.

Podemos concluir de este tab que los videos con más comentarios son de la mismo enfoque matemático y magistral lo cual puede ser una razón para que la gente interactúe más con este tipo de videos.

Tomando en cuenta que para comentar se necesita tener una cuenta de YouTube y probablemente muchos de ellos no lo tengan, creemos que es un buen Insight que los videos con enfoque en las matemáticas hacen que las personas interactúen más con los videos.

3. Conclusiones y recomendaciones:

Podemos concluir de una manera muy fuerte que los videos de matemáticas y magistrales de Academática dominan las estadísticas de los videos. Por lo cual recomendamos lo siguiente:

* Realizar más videos de este estilo porque son los más populares
* El alcance que tienen este tipo de videos nos ayuda a entender que quieren los usuarios con los likes y dislikes
* Este tipo de videos hacen que la gente interactúe más con el video y deje comentarios con opiniones o recomendaciones.
* Los dislikes no definen la calidad del video