

## Parcial 1 Data Product

### Parte 1:

Para la primera parte del parcial, se levantó un contenedor en Docker para Mysql. El objetivo era poder crear en este una base de datos que posteriormente se conectará con RStudio. Para ello fue necesario descargar la imagen de MySQL para posteriormente poder ejecutarlo con Docker y levantar el contenedor.

### Conectar con Mysql

```
CONTAINER ID        IMAGE               COMMAND             CREATED             STATUS              PORTS              NAMES
e8a34775c352      mysql              "docker-entrypoint.s..." 3 minutes ago       Created             3306/tcp           MySQLDPP1
7826bda0563c      rocker/rstudio     "/init"            19 minutes ago      Created             3306/tcp           amazing_gates
a41c3af61689      rocker/rstudio     "/init"            22 minutes ago      Created             3306/tcp           jolly_raman
e758f90fa388      rocker/rstudio:4.0.2 "/init"            22 minutes ago      Up 22 minutes      0.0.0.0:8787->8787/tcp festive_franklin
e2866a8ede5d      jupyter/datascience-notebook "tini -g -- start-no..." 12 days ago         Exited (129) 11 days ago 3306/tcp, 33068/tcp affectionate_elbakyan
a910184c196e      mysql              "docker-entrypoint.s..." 12 days ago         Exited (255) 33 minutes ago 3306/tcp, 33068/tcp unruffled_neitner
1e458bc297ed      jupyter/datascience-notebook "tini -g -- start-no..." 12 days ago         Exited (0) 12 days ago 3306/tcp, 33068/tcp suspicious_wright
8b4de88d7565      busybox            "echo 'hello from da..." 12 days ago         Exited (0) 12 days ago 3306/tcp, 33068/tcp nostalgic_feistel
3f9f9d7c0ea6      busybox            "sh"                12 days ago         Exited (0) 12 days ago 3306/tcp, 33068/tcp optimistic_dirac
237183e42db9      hello-world        "/hello"            12 days ago         Exited (0) 12 days ago 3306/tcp, 33068/tcp vibrant_newton
8f936d528e9b      hello-world        "/hello"            13 days ago         Exited (0) 13 days ago 3306/tcp, 33068/tcp sad_hamilton
bfc4e9951955      hello-world        "/hello"            2 weeks ago         Exited (0) 2 weeks ago 3306/tcp, 33068/tcp charming_engelbart

(base) Rauls-MacBook-Pro-2:~ raucastellanos$ docker container stop [e8a34775c352]
Error response from daemon: No such container: [e8a34775c352]
(base) Rauls-MacBook-Pro-2:~ raucastellanos$ docker container rm [e8a34775c352]
Error: No such container: [e8a34775c352]
(base) Rauls-MacBook-Pro-2:~ raucastellanos$ docker run --name Parcial1DP -p 3302:3306 -e MYSQL_ROOT_PASSWORD=password mysql
2020-08-19 00:25:53+00:00 [Note] [Entrypoint]: Entrypoint script for MySQL Server 8.0.21-1debian10 started.
2020-08-19 00:25:53+00:00 [Note] [Entrypoint]: Switching to dedicated user 'mysql'
```

### Crear la base de datos

```
mysql> create database academatik;
|Query OK, 1 row affected (0.01 sec)

mysql> use academatik;
|Database changed
```

Crear tablas: la creación de tablas y la importación de los datos se realizó en Workbench.

```
• DROP TABLE IF EXISTS Metadata;
• CREATE TABLE Metadata(
  VideoID VARCHAR(100),
  Title VARCHAR(100),
  Descriptions MEDIUMTEXT,
  Details VARCHAR(100)
);

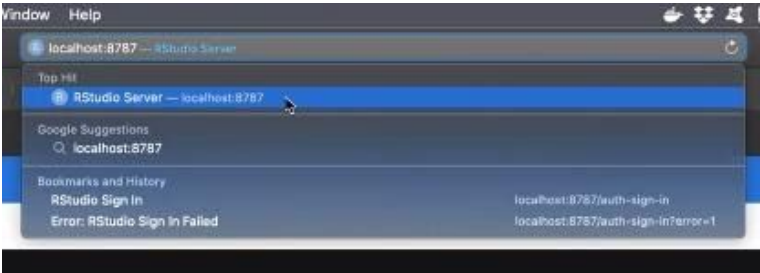
• DROP TABLE IF EXISTS Videos;
• CREATE TABLE Videos(
  Kind VARCHAR(100),
  Etag VARCHAR(100),
  id TINYTEXT,
  Details VARCHAR(100),
  Pubish VARCHAR(100)
);

• DROP TABLE IF EXISTS Stats;
• CREATE TABLE Stats
(
  ID VARCHAR(50),
  Views INT,
  Likes INT,
  Dislikes INT,
  Favorites INT,
  Comments INT
);
```



```
Terminal Shell Edit View Window Help
raulcastellanos ~ - bash -- 15x47
(base) Rauls-MacBook-Pro-2:~ raulcastellanos$ docker ps
CONTAINER ID   IMAGE                                COMMAND                  CREATED        STATUS        PORTS
NAME          obedaeg/rstudio-mysql              "/init"                  About a minute Up About a minute 8787/tcp, 0.0.0.0:8888->8888/tcp
ce7c12296cf4  competent_elbakyan                 "/entrypoint.sh mysql..." 20 hours ago  Up 20 hours (healthy) 3306/tcp, 33060/tcp
766b21cf9432  mysql/mysql-server:latest          "/entrypoint.sh mysql..." 20 hours ago  Up 20 hours (healthy) 3306/tcp, 33060/tcp
BaseDP        a7399bfb5693                       "docker-entrypoint.s..." 20 hours ago  Up 20 hours          3306/tcp, 33060/tcp
vigilant_vaughan  mysql                               "docker-entrypoint.s..." 20 hours ago  Up 20 hours          3306/tcp, 33060/tcp
9dd10917ec26  obedaeg/rstudio-mysql              "/init"                  21 hours ago   Up 21 hours          0.0.0.0:8787->8787/tcp
cranky_lamport  obedaeg/rstudio-mysql              "/init"                  21 hours ago   Up 21 hours          0.0.0.0:8787->8787/tcp
(base) Rauls-MacBook-Pro-2:~ raulcastellanos$
```

Rstudio en localhost:

A screenshot of the 'Sign in to RStudio' web form. It contains two input fields: 'Username:' with the value 'rstudio' and 'Password:' with masked characters. Below the fields is a checkbox labeled 'Stay signed in when browser closes' and a 'Sign In' button. A message at the bottom states: 'You will automatically be signed out after 60 minutes of inactivity.'

Tras acceder a RStudio, se realizó la conexión con la base de datos que contiene la información. La conexión se realizó con la librería RMySQL. Para la parte del análisis se trabajó un Flex Dashboard en el cual se incluyen varios datos generales extraídos de la base de datos, al igual que algunas visualizaciones de los datos.

Como parte inicial se proporciona información sobre el trabajo realizado, y se brinda una pequeña introducción de los datos analizados.

| Análisis de Academatica  |   |  |
|--|---|--|
| Parcial 1 Data Product   | Data Academatica  | Estadísticas del Canal   |
| Datos Generales de los Videos  |   |  |
| <p>Nombres</p> <p>Carlos Morales 20180067</p> <p>Raúl Castellanos 20180052</p> |   | <p>Introducción</p> <p>Estos análisis están hechos con base a los datos proporcionados del canal Academatica. El objetivo del análisis es observar el comportamiento de los usuarios del canal a lo largo del tiempo.</p> <p>Para ello se muestran los datos de las estadísticas generales del canal, al igual que la respuesta de la comunidad hacia los videos a lo largo de los años.</p> |
| <p>Programas Utilizados</p> <p>Docker</p> <p>RStudio</p> <p>MySQLWorkbench</p> | <p>Nombre de los archivos Utilizados</p> <p>academatica_videos.csv</p> <p>academatica_videos_metadata.csv</p> <p>academatica_videos_stats.csv</p> | <p>Comentarios extra</p> <p>Esta página HTML fue hecha por medio de RStudio utilizando la librería de flexdashboards. Los datos utilizados son reales, proporcionados por el usuario propietario del canal Academatica Carlos Zelada.</p>  |

Posteriormente se incluyeron dos páginas para visualización de datos. En la primera se pueden observar todos los datos descriptivos de los videos subidos por el canal Academica, seguido por todas las estadísticas de los videos. Para la segunda se generó una tabla que incorpora todos los datos relevantes de los videos, y separada en cinco pestañas que los agrupan por periodos de tres años desde 2006 hasta la fecha.

Analisis de Academica

Parcial 1 Data Product

Data Academica

Estadísticas del Canal

Datos Generales de los Videos

Show10▼entries

Search:

|   | Title                                 | Descriptions  | Details                      | Fecha      |
|---|---------------------------------------|---|------------------------------|------------|
| 1 | Flexdashboard - Sesion 1 - 2020-08-12 | En esta primera sesion aprenderemos sobre como dise<f1>ar dashboard utilizando la libreria de flexdashboards.   | https://youtu.be/lZliv4xo67c | 2020-08-10 |
| 2 | RnoteBooks parte 2                    | En el taller pasado aprendimos como agregar diferentes elementos de dise<f1>o a nuestro notebook ahora aprenderemos a agregar codigo y correrlo. Crearemos imagenes, usaremos codigo en python, y SQL | https://youtu.be/Gty0tPoESxc | 2020-08-06 |

Showing 1 to 10 of 859 entries

Previous12345...86Next

Show10▼entries

Search:

| ID | Views       | Likes | Dislikes | Favorites | Comments | Fecha      |
|----|-------------|-------|----------|-----------|----------|------------|
| 1  | lZliv4xo67c | 0     | 2        | 0         | 0        | 2020-08-10 |
| 2  | Gty0tPoESxc | 192   | 9        | 0         | 1        | 2020-08-06 |
| 3  | 96niOlwgosg | 72    | 7        | 1         | 0        | 2020-08-10 |
| 4  | kFRAM2f2YS4 | 143   | 6        | 0         | 0        | 2020-08-06 |

Showing 1 to 10 of 859 entries

Previous12Activate Windows...86Next

Analisis de Academática

Parcial 1 Data Product

Data Academática

Estadísticas del Canal

Datos Generales de los Videos

2006 al 2008

2009 al 2011

2012 al 2014

2015 al 2017

2018 al 2020

Show10entries

Search:

| Title                                       | Descriptions  | Details                      | ID          | Views | Likes | Dislikes | Favorites | Comments | Fecha      |
|---|---|------------------------------|-------------|-------|-------|----------|-----------|----------|------------|
| 855 Vitriolo (Hoy)                          | Video de un grupo guatemalteco buenísimo, vitriola,Saludos Canche   | https://youtu.be/pgXkBdcARz4 | pgXkBdcARz4 | 1060  | 9     | 0        | 0         | 1        | 2007-03-25 |
| 856 Demo de una mi clase                    | Ejemplo de series de potencia. Serie de potencia del ln(x) www.chzelada.net   | https://youtu.be/4Z7uRiAhcsM | 4Z7uRiAhcsM | 5970  | 3     | 1        | 0         | 6        | 2006-12-14 |
| 857 TimeLapse desde mi ventana              | ES un time lapse verdadero, ahora si, miren las nubes se miran chileras..   | https://youtu.be/S8CPPg72Lqc | S8CPPg72Lqc | 761   | 7     | 1        | 0         | 1        | 2006-10-09 |
| 858 De reversa de san jose pinula a galileo | El mismo video que el anterior pero visto desde atraz. The same video but seen from behind.   | https://youtu.be/Yuni7gNVgIM | Yuni7gNVgIM | 1655  | 3     | 1        | 0         | 0        | 2006-10-08 |
| 859 De San jose pinula a la Galileo         | Video del trafico que me toca cuando voy de donde vivo hasta la universidad. Video of the traffic I encounter when I drive from home to school. | https://youtu.be/lVRzL12p-SA | lVRzL12p-SA | 6133  | 19    | 4        | 0         | 9        | 2006-10-07 |

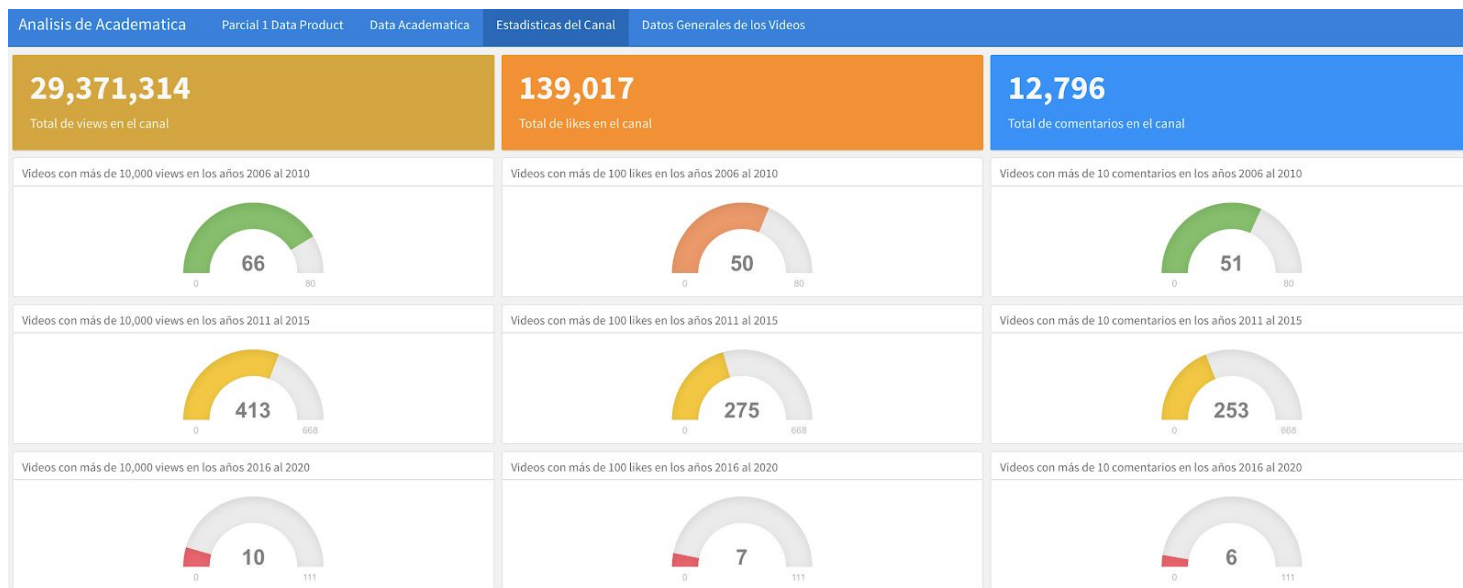
Showing 1 to 5 of 5 entries

Previous

1

Next

Finalmente para la parte del análisis se realizaron dos segmentos. Primero se incluyó una colección de datos históricos del canal, la cual incluye el total de views, el total de likes, y el total de comentarios que tiene históricamente el canal. Seguido a este análisis se realizó una serie de medidores segmentados que reflejan la cantidad de videos por encima de 10,000 views, la cantidad de videos por encima de 100 likes, y la cantidad de videos por encima de 10 comentarios, agrupados por rango de tiempo en periodos de 5 años de 2006 a 2020.



### Conclusiones:

Con base al análisis realizado podemos concluir que el canal de Academática tuvo un crecimiento bastante alto en los años del 2011 al 2015 debido a que en ese lapso se subieron 668 videos a comparación de los otros dos períodos de tiempo. En los últimos 5 años el canal ya no ha tenido las visitas que tenía debido a la falta de constancia de publicar videos y también que pueden encontrarse más videos con el contenido relacionado en YouTube de otros canales que han crecido en los últimos años.

Adicionalmente se observa que las visitas de años antiguos superan las actuales, lo que apunta a que muchos espectadores no se quedan siguiendo el canal. Sin embargo, este puede ser un problema común en canales cuyo contenido es técnico/informativo ya que probablemente muchos solo vean el tema que fue de su interés, y no busquen permanecer en el canal.

### Recomendaciones:

- Intentar retomar el ritmo de creación de contenido, ya que de esta forma se podrá recuperar las estadísticas que mantenía el canal en su momento más alto.
- Aprovechar posibles hot topics para incluirlo en el título del video y enganchar a más espectadores.

- Aprovechar los comentarios que los usuarios realizan para poder hacer más videos con explicaciones de sus preguntas, para que de esta manera se pueda fidelizar a los usuarios y poder obtener más usuarios nuevos