# Documentación del Proyecto TFM-Proyecto

## 1. Introducción

Este documento proporciona una guía detallada sobre la instalación, configuración y ejecución del proyecto 'TFM-Proyecto'. Se explican los pasos para configurar el entorno, instalar dependencias, ejecutar los scripts y subir el proyecto a GitHub.

## 2. Requisitos Previos

Antes de instalar el proyecto, asegúrate de tener los siguientes componentes instalados:

- Python 3.8 o superior

- Git

- Docker y Docker Compose (opcional, para ejecutar Kafka y MySQL en contenedores)

- Apache Kafka

- MySQL

## 3. Clonar el Proyecto

Para obtener una copia local del proyecto, ejecuta el siguiente comando en la terminal:

git clone https://github.com/alexander2295/TFM-Proyecto.git

## 4. Configuración del Entorno Virtual

Se recomienda usar un entorno virtual para manejar las dependencias. Para crearlo, usa:

python -m venv venv

Para activarlo:

- \*\*Windows\*\*: `venv\Scripts\activate`

- \*\*Linux/Mac\*\*: `source venv/bin/activate`

## 5. Instalación de Dependencias

Ejecuta el siguiente comando para instalar las dependencias necesarias del proyecto:

pip install -r requirements.txt

## 6. Configuración de Variables de Entorno

Crea un archivo `.env` en la raíz del proyecto con los siguientes datos:

MYSQL\_HOST=localhost

MYSQL\_PORT=3306

MYSQL\_USER=root

MYSQL\_PASSWORD=tu\_contraseña

MYSQL\_DB=youtube\_data

KAFKA\_SERVER=localhost:29092

## 7. Configuración de Kafka y MySQL

Si usas Docker, ejecuta el siguiente comando para iniciar Kafka y MySQL en contenedores:

docker-compose up -d

## 8. Ejecutar los Reproductores

Para ejecutar los scripts de los reproductores, usa los siguientes comandos:

python scripts1/reproductores/reproductor1.py

python scripts1/reproductores/reproductor2.py

python scripts1/reproductores/reproductor3.py

## 9. Ejecutar el Consumidor

Para consumir los datos de Kafka y almacenarlos en MySQL, ejecuta:

python scripts1/consumidor/consumidor.py

## 10. Ejecutar el Pipeline de Datos

Para ejecutar todo el pipeline de datos, usa el siguiente comando:

python pipeline\_data.py

## 11. Subir el Proyecto a GitHub

Para subir el proyecto a GitHub, sigue estos pasos:

git add .

git commit -m 'Subiendo el proyecto TFM-Proyecto'

git push origin main

## 12. Configurar .gitignore

Crea un archivo `.gitignore` en la raíz del proyecto y agrega lo siguiente para evitar subir archivos innecesarios:

venv/

\*.pyc

\_\_pycache\_\_/

.env

logs/

\*.log

data/

## 13. Conclusión

Este documento proporciona una guía detallada para instalar, configurar y ejecutar el proyecto TFM-Proyecto. Si tienes dudas, consulta la documentación de cada componente.