



Proyecto final CoderHouse

Data Engineering

Alumno: Nicolás Rivas



DATA ENGINEERING

ETL

El Pipeline creado tiene como objetivos guardar diariamente el valor de las criptomonedas de la api publica <https://docs.coincap.io/> para el cual se usarán 2 tablas

- La tabla de histórico de valores “cripto_price_history”
 - La misma contará con la siguiente estructura (nombre de campo - tipo de dato)

ID-varchar(255)	name-varchar(255)	priceusd-float8	date_time-timestamp	key_id_date_time-varchar(255)
-----------------	-------------------	-----------------	---------------------	-------------------------------

- Una tabla que hará las veces de temporal para comparar contra la anterior los datos ya existentes “cripto_price_new_data”
 - La misma contará con la siguiente estructura

ID-varchar(255)	name-varchar(255)	priceusd-float8	date_time-timestamp	key_id_date_time-varchar(255)
-----------------	-------------------	-----------------	---------------------	-------------------------------

COMPONENTES

- El trabajo se compone de:
 - Archivo docker-compose.yml en el que están seteadas las configuraciones para correr el airflow webserver, worker, trigger, etc., la base postgresql y en el que se seteo en el worker la carpeta correspondiente llamada tmp files en la que se guardan los archivos temporales de ejecución.
 - Un archivo airflow.cfg con la configuración de airflow básica
 - La configuración del SMTP para el envío de alertas se realizó seteando el email_conn_id en default y creando la conexión en airflow -> admin -> connections
 - El código de python API_Crypto_Extraction_modularizad.py en el cual está desarrollado el DAG y las tareas del mismo
 - Un archivo .env en el mismo nivel del .py para tomar las credenciales desde allí (este no se subirá a github para la entrega).
 - Archivo Crear tablas en DW.sql con el código para la creación de las 2 tablas necesarias para la ejecución del ETL

SQL para creación de las tablas

```
CREATE TABLE ripto_price_history(  
    id            VARCHAR(255) not null,  
    name          VARCHAR(255) not null,  
    priceusd      float8 not null ,  
    date_time     TIMESTAMP not null ,  
    key_id_date_time VARCHAR(255) primary key  
);  
  
CREATE TABLE cripto_price_new_data(  
    id            VARCHAR(255) not null,  
    name          VARCHAR(255) not null,  
    priceusd      float8 not null ,  
    date_time     TIMESTAMP not null ,  
    key_id_date_time VARCHAR(255) primary key  
);
```