



```
In [1]: import pandas as pd
import os
from sqlalchemy import create_engine
import logging
import time

logging.basicConfig( filename="logs/ingestion_db.log",
                     level = logging.DEBUG,
                     format = "%(asctime)s - %(levelname)s - %(message)s",
                     filemode = "a"
                     )

engine = create_engine('sqlite:///inventory.db')
def ingest_large_csv(file, table_name, engine, read_chunksize=100000, insert_chunksize=10000):
    for chunk in pd.read_csv(file, chunksize=read_chunksize):
        chunk.to_sql(table_name, con=engine, if_exists='append', index=False, chunksize=insert_chunksize)
    print(f"Inserted {len(chunk)} rows into {table_name}...")

def ingest_db(df, table_name, engine):
    """ this function will ingest the dataframe into database table"""
    df.to_sql(table_name, con = engine, if_exists = 'replace', index = False)

def load_raw_data():
    """ this function will load the CSVs as dataframe and ingest into db... """
    start = time.time()
    for file in os.listdir("data"):
        if file.endswith(".csv"):
            logging.info(f"Ingesting {file} in db ")
            ingest_large_csv("data/"+file, file[:-4], engine)
    end = time.time()
    total_time = (end - start)/60
    logging.info('-----Ingestion Complete-----')
    logging.info(f'\nTotal Time Taken: {total_time} minutes')
if __name__ == '__main__':
    load_raw_data()
```


In []: