

```
In [1]: import pandas as pd
import numpy as np
import os
from sqlalchemy import create_engine
```

```
In [2]: #Reading Excel File using pandas
olist_database = pd.ExcelFile('olist business database.xlsx')
olist_database
```

```
Out[2]: <pandas.io.excel._base.ExcelFile at 0x21973f83eb0>
```

```
In [3]: #Listing out all sheets in the Excel file
olist_database.sheet_names
```

```
Out[3]: ['olist_Customer_table',
'olist_geolocation_table',
'olist_order_items_table',
'olist_order_payment_table',
'olist_order_reviews_table',
'olist_orders_table',
'olist_products_table',
'olist_sellers_table',
'olist_product_category']
```

```
In [4]: #Creating extract function for ETL process
def extract_from_excel_sheet(excelfilename):
    query = olist_database.parse(excelfilename)
    return query
```

```
In [5]: #Writing load function for ETL process
def load_to_sql_server(df, tablename, db_engine):
    return df.to_sql(tablename, db_engine, if_exists = 'replace', index = False)
```

```
In [6]: #Putting ETL all together, this code extracts tables/sheets from the Excel file and parses the tables to SQL
#where the database relationships will be established.
def extract_and_load(excelfilename, tablename):
    df = extract_from_excel_sheet(excelfilename)

    load_to_sql_server(df, tablename, db_engine)
```

```
In [7]: #Getting connection details to SQL server using environemtal variables
SERVER = os.environ.get('MS SQL SERVER NAME')
DRIVER = os.environ.get('MS SQL SERVER DRIVER')
database_name = 'olist_db'
```

```
In [8]: #Estalishing connection with database
connection = f'mssql://{SERVER}/{database_name}?driver={DRIVER}'
```

```
In [10]: db_engine = create_engine(connection)
```

```
In [11]: #Finally lets load all tables into SQL SERVER
extract_and_load('olist_Customer_table', 'staging_olist_Customer_table')
extract_and_load('olist_geolocation_table', 'staging_olist_geolocation_table')
extract_and_load('olist_order_items_table', 'staging_olist_order_items_table')
extract_and_load('olist_order_payment_table', 'staging_olist_order_payment_table')
extract_and_load('olist_order_reviews_table', 'staging_olist_order_reviews_table')
extract_and_load('olist_orders_table', 'staging_olist_orders_table')
extract_and_load('olist_products_table', 'staging_olist_products_table')
extract_and_load('olist_sellers_table', 'staging_olist_sellers_table')
extract_and_load('olist_product_category', 'staging_olist_product_category')
```

```
In [12]: extract_and_load('olist_products_table', 'staging_olist_products_table')
```

```
In [13]: connection.close()
```

```
In [14]:
```

```
In [15]:
```

```
In [16]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js