

SONY AIESEC TASK DEBRIEF

Luiz C. E. Amoras Jr.
Junior Data Science Analyst Application

Summary



- 1 SQL Task Description
 - 1.1 SQL Task Solution Problem #01
 - 1.3 SQL Task Solution Problem #02
 - 1.3 SQL Task Solution Problem #03
- 2 Python Task Description
 - 2.1 Python Task Solution Problem



SQL TASK

1 - SQL Task - Description

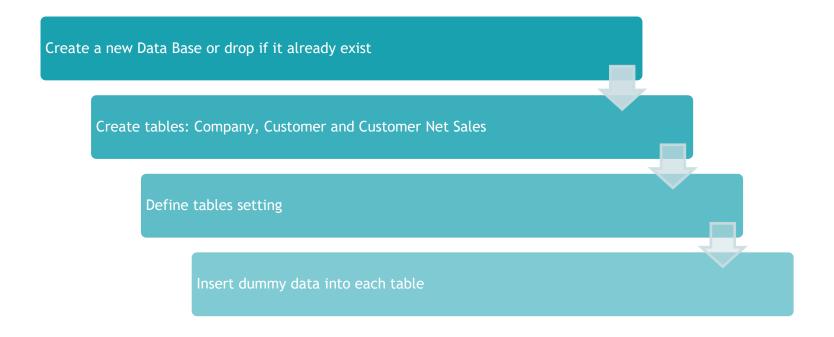


- 1. Prepare tables and populate with Dummy data:
 - Company table fields (CompanyID, CompanyName)
 - Customer belongs to Company fields (CustomerID, CustomerName, CompanyID)
 - Customer has Net Sales fields (Customer ID, Quarter [data like 2019Q1], Amount)
- 2. Make Selection that will show the Net Seales for Company ID 1.
- 3. Create pivot that will show Customers as rows, quarters as columns and Net Sales Amounts as Values as table valued function, parameter is Company ID.

1.1 - SQL Task - Solution Problem #01

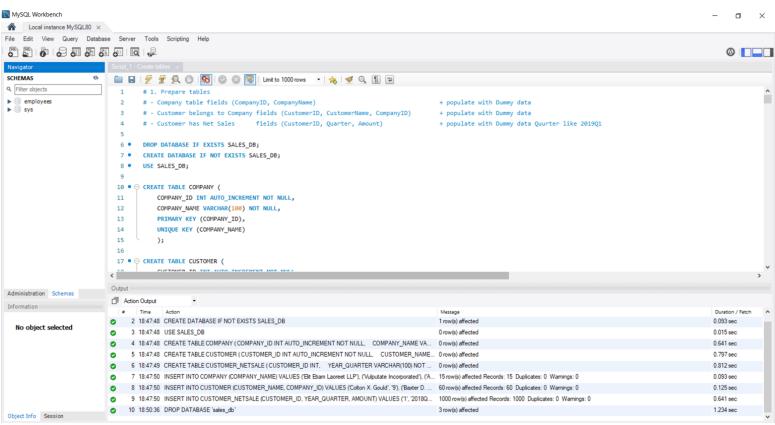


1. Prepare tables and populate with Dummy data:



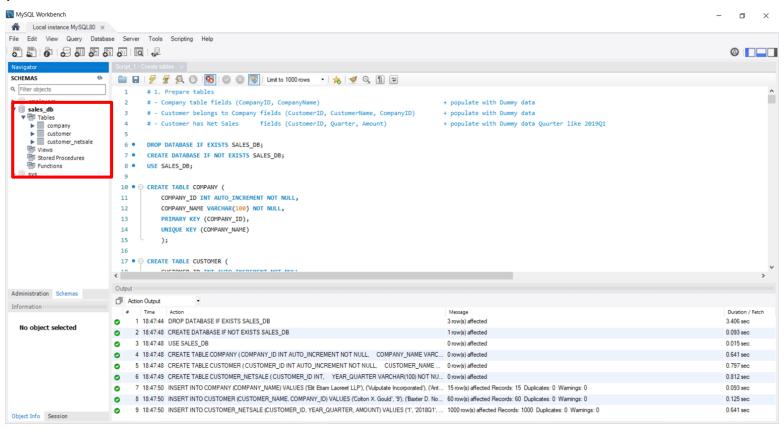
1.1 - SQL Task - Solution Problem #01





1.1 - SQL Task - Solution Problem #01





1.2 - SQL Task - Solution Problem #02

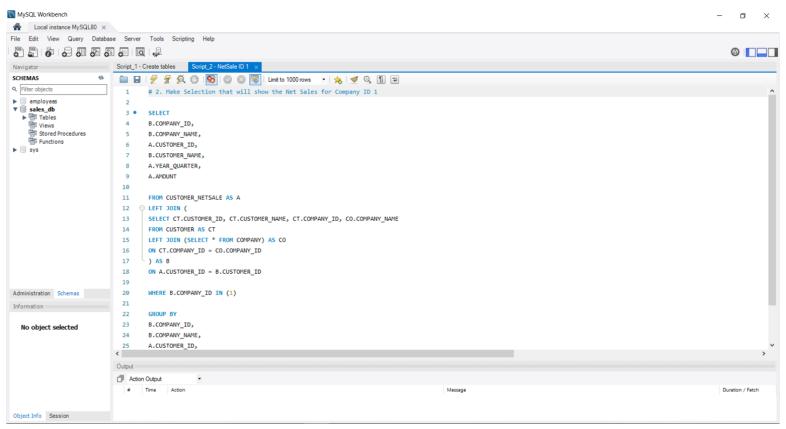


2. Make Selection that will show the Net Seales for Company ID 1.



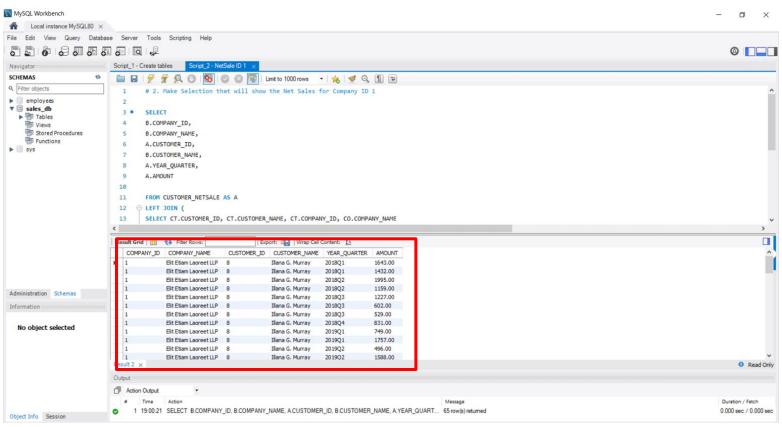
1.2 - SQL Task - Solution Problem #02





1.2 - SQL Task - Solution Problem #02





1.3 - SQL Task - Algorithm Problem #03

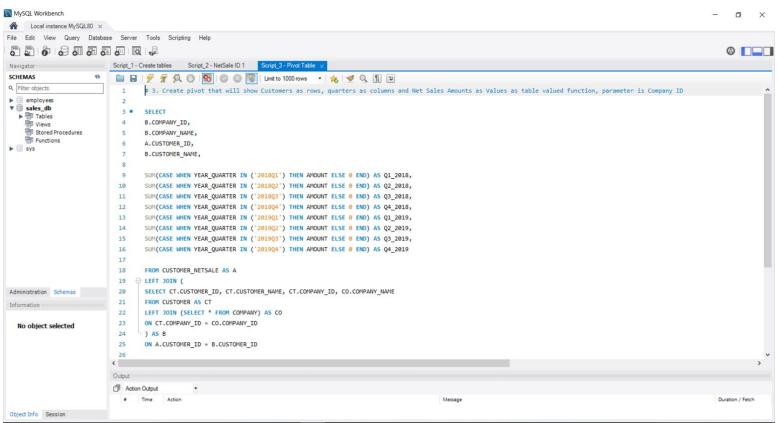


3. Create pivot that will show Customers as rows, quarters as columns and Net Sales Amounts as Values as table valued function, parameter is Company ID.



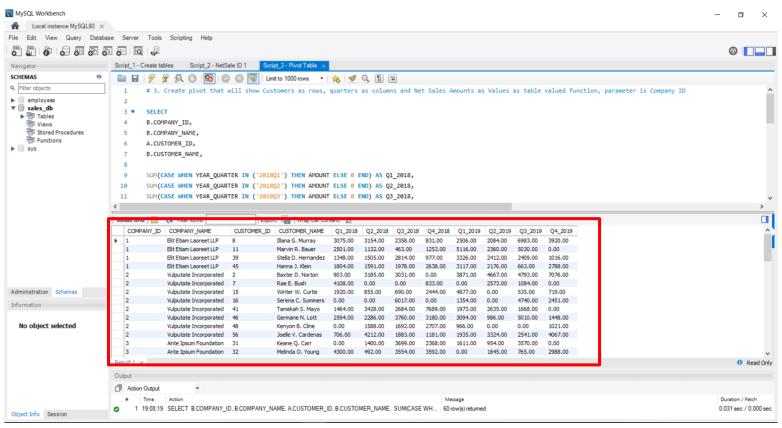
1.3 - SQL Task - Solution Problem #03





1.3 - SQL Task - Solution Problem #03







PYTHON TASK

- -

2 - Python Task



Task description: Design a VBA/Python automation which would create a report for the business according to the below details.

Source Files(There is an automatic daily process creating refreshed source files each day with updated data):

- **Source 1**: BW_CPFR_OUTLET_HU10.XLS (Sell Out Data)
- **Source 2**: BW_SDS_HU10.XLS (Net Sales Data)

Final Report File:

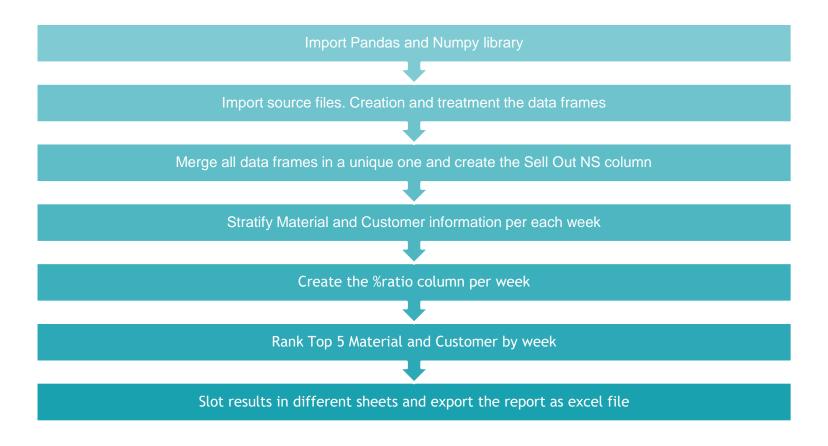
• Top5NS_Report.xlsm (to be created daily)

Show data as values only (no formula reference to the datasheets) presenting the following:

- 1. List the top 5 SELL OUT NS generating materials of each week available in the source file. Also show the %Ratio they represent of the total NS of that week. Please use 0 (zero) Average NS wherever the data is not available for the material. Key between Source 1 and Source 2 file is MATERIAL.
- 2. Which are the top 5 NS generating customers of each week in the last 3 closed weeks? (excluding the last partial week).

The automation should run daily in English environment.

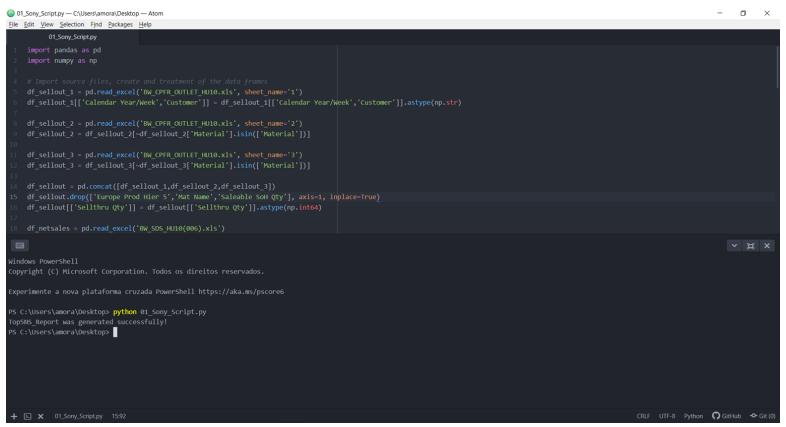




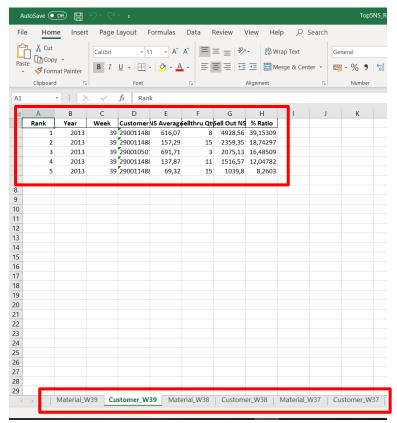


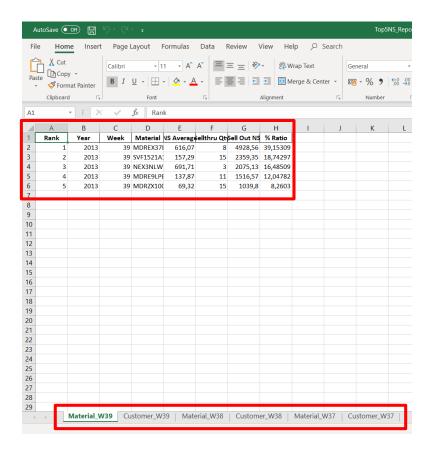
```
@ 01_Sony_Script.py — C:\Users\amora\Desktop — Atom
File Edit View Selection Find Packages Help
          01_Sony_Script.py
    import pandas as pd
    import numpy as np
    df sellout 1 = pd.read excel('BW CPFR OUTLET HU10.xls', sheet name='1')
    df sellout 1[['Calendar Year/Week', 'Customer']] = df sellout 1[['Calendar Year/Week', 'Customer']].astype(np.str)
    df_sellout_2 = pd.read_excel('BW_CPFR_OUTLET_HU10.xls', sheet_name='2')
    df sellout 2 = df sellout 2[~df sellout 2['Material'].isin(['Material'])]
    df sellout 3 = pd.read excel('BW CPFR OUTLET HU10.xls', sheet name='3')
    df sellout 3 = df sellout 3[~df sellout 3['Material'].isin(['Material'])]
    df sellout.drop(['Europe Prod Hier 5', 'Mat Name', 'Saleable SoH Qty'], axis=1, inplace=True)
    df_sellout[['Sellthru Qty']] = df_sellout[['Sellthru Qty']].astype(np.int64)
    df netsales = pd.read excel('BW SDS HU10(006).xls')
    df netsales.columns = df netsales.iloc[0]
    df netsales.drop(['Calendar Year/Month', 'Europe Prod Hier 3', 'Sales Quantity', 'P5 Net Sales LC', 'Material Name'], axis=1, inplace=True)
    df netsales.rename(columns={'P5 Net Sales EUR': 'NetSales'}, inplace=True)
    df netsales[['NetSales']] = df netsales[['NetSales']].astype(np.float64)
    df source = (pd.merge(df netsales, df sellout, on=['Calendar Year/Week', 'Material', 'Customer'], how='outer'))
    df source = df source.replace(np.nan, 0, int==True)
    df source = (df source.groupby(['Calendar Year/Week', 'Customer', 'Material'], as index=False)
                 .agg({'Sellthru Qty':'sum', 'NetSales':'mean'})
                 .rename(columns={'NetSales':'NS Average'}))
    n column = df source['Calendar Year/Week'].str.split('.', n=1, expand = True)
    df source['Week'] = n column[0]
    df source['Year'] = n column[1]
     off source = df source reindex(columns=['Year' 'Week' 'Customer' 'Material' 'NS Average' 'Sellthru Otv'])
                                                                                                                                                                     CRLF UTF-8 Python GitHub - Git (0)
```











SONY

make.believe