D598\_Task 2

Bianca Papapietro

Professor Hanan Swidan

June 2025

WGU

Analytics Programming: D598

```python

import pandas as pd

#help to import the file and give the data a structure

```

```python

df = pd.read\_csv("D598 Data Set.csv")

#upload and read the CSV file

```

```python

df = df.drop\_duplicates()

#remove the duplicates

```

```python

groupbystate\_df = df.groupby("Business State")

#group data by state

```

```python

groupbystate\_df = groupbystate\_df.apply(lambda x: x.reset\_index(drop=True))

#save the newly grouped data by state as a new dataframe

```

```python

df\_new\_filtered = df[df["Debt to Equity"] > 0]

#filter out negative debt to income ration values from the Debt to Equity column

```

```python

df["DTI\_Per\_Business"] = df.groupby("Business ID")["Debt to Equity"].transform("mean")

#calculate the debt to income ration

```

```python

df\_new = df\_new\_filtered.copy()

#save as a new dataframe

```

```python

final\_df = pd.concat([df, df\_new])

#concatenate with the original dataset

```python

final\_df.to\_csv("new\_completely\_processesd\_data.csv", index=False)

#save the new dataset as a new CSV file

```

```python

print("Data processing is complete!")

#print to confirm that the process is complete

```

Data processing is complete!

```python

# Create a new project directory in a different location

import os

project\_dir = os.path.join("C:\\Users\\greet\\Documents", "my\_github\_project")

if not os.path.exists(project\_dir):

os.makedirs(project\_dir)

# Change to the new project directory

os.chdir(project\_dir)

print(f"Current working directory: {os.getcwd()}")

# Initialize a fresh Git repository

!git init

```

Current working directory: C:\Users\greet\Documents\my\_github\_project

Initialized empty Git repository in C:/Users/greet/Documents/my\_github\_project/.git/