



# Non Graded Challenge : DDL, DML, and Basic DQL

FTDS Phase 0 - Week 2

# Contents



Objective	_____	<b>03</b>
Dataset	_____	<b>04</b>
Task	_____	<b>05</b>
Submission	_____	<b>07</b>

## *Non Graded Challenge*

# Objective

This Non Graded Challenge was created to evaluate the concept of Pandas with the aim of:

1. Able to load data with Pandas
2. Able to handle missing values in the dataset
3. Able to perform column manipulation
4. Able to make queries on the dataset
5. Able to perform grouping on the dataset

*Non Graded Challenge*

## **Dataset**

You can access the data here:

[https://github.com/FTDS-learning-materials/phase-0/raw/main/src/data\\_ngc.csv](https://github.com/FTDS-learning-materials/phase-0/raw/main/src/data_ngc.csv)

## *Non Graded Challenge*

# Task

This Non-Graded Challenge is done in notebook format with several mandatory criteria as follows:

1. The library used is Pandas.
2. The project is considered complete and ready for assessment if, when "Run All" is executed in the notebook, all cells successfully run to the end.
3. The notebook's content should follow the outline below:
  - a. Introduction  
The introduction section must contain the identity, background of the problem, and the problem statement to be discussed.
  - b. Import necessary libraries  
The first cell in the notebook must contain and only contain all the libraries used in the project.
  - c. Data Loading  
This section contains the queries that have been created and the data loading process, followed by simple data exploration.

## *Non Graded Challenge*

### **Task**

d. Data Cleaning

This section contains the data preparation process, such as data cleaning, before further data exploration. The cleaning process may include assigning new names to each column, filling in missing values, deleting unused columns, and so on.

e. Data Exploration

This section contains data exploration on the dataset above using queries, grouping, and so on.

f. Conclusion

In this final section, there must be conclusions that reflect the results obtained by comparing them with the problem statement written in the introduction section.

*Non Graded Challenge*

## **Submission**

1. Save this assignment with filename : *h8dsft\_ngc\_pandas.ipynb*.
2. Push your answer into your own GitHub repository.



# Thank You