

DSC 200 – Data Wrangling
Lab 4: Working with Excel Files

Goals:

- Read data from Excel files
- Use different Python Libraries to open and manipulate Excel files

Instructions:

The linked excel file contains data collected from countries regarding child abuse in those respective countries. Several categories of child abuse occurred. For each category of abuse, the dataset includes total number of children who were affected. You are required to write a program that reads this Excel file and creates a CSV file containing data for the following headings:

CountryName, CategoryName, CategoryTotal

where *CountryName* is the name of the country or area, *CategoryName* is the name of child abuse events, and *CategoryTotal* is the total number of children in that category. All these data can be found in the Excel sheet referenced above and linked in the Canvas assignment.

Note the following:

- For cells that contain “-” the count is zero. If the count for any category is zero, that row should not be included in the final dataset.
- The script should use English names for the countries and category names.
- If a country has no entries for a category, do not include that country in the output CSV file.
- Name the file using the format ***your_nku_username.csv***.
- You should use any of the Python Libraries we discussed in class for this assignment. This include ***openpyxl*** and ***xlrd***.

The program should print the number of rows, excluding the header, in the CSV output file to the console after creating it.

What to submit: Submit only the Python script for creating the CSV file. Name the file using the following format: ***your_groupnumber_Lab4.py***