PDF to Excel Converter Script Documentation

Introduction:

This script is designed to process one or more PDF files, extract specific data from them using regular expressions, and export the extracted data to an Excel file. The script is built using Python and the Tkinter library for the user interface.

Prerequisites:

- Python: You need to have Python installed on your system.
- Required Libraries: Make sure you have the following Python libraries installed:
1)tkinter
2)PyPDF2
3)re
4)pandas

Usage:

- **Select PDF Files:**
  Click the "Process PDF Files" button to select one or more PDF files that you want to process.

- **Extraction Process:**
  The script will extract data from the selected PDF files using regular expressions. You can define regular expressions for specific data fields in the code.

- **Data Compilation:**
  The extracted data from all selected PDF files is compiled into a single DataFrame, combining data from all PDFs.

- **Download Excel File:**
  1)The "Download Excel" button is initially disabled.
  2)After processing PDF files, click the "Download Excel" button to save the compiled data as an Excel file.
  3) You can specify the file location and name for the Excel file.

Example Usage:

- Click "Process PDF Files" and select one or more PDF files containing data.
- After processing, click "Download Excel" to save the extracted data as an Excel file.
- Specify the file location and name for the Excel file.

Troubleshooting:

- If you encounter any issues or errors, ensure that you have the required Python libraries installed.