How to convert multiple excel files (same directory) into a single CSV file ?

Import all necessary modules

```
In [ ]: import pandas as pd
        import os
        import glob
        import shutil
```

Function for converting an .xlsx file to .csv file

```
In [ ]: global count_of_records
        count_of_records = 0

        def xlsx_to_csv(excel_filename):

            global count_of_records

            # Load the excel file in dataframe
            df = pd.read_excel(excel_filename, header=0, usecols=[0,1,2,3,4,5,6,7,8,11,12,13,21,22]) # Necessary columns
            print("Shape : " + str(df.shape))
            count_of_records = count_of_records + df.shape[0]

            # Write a CSV file from the dataframe
            csv_filename = excel_filename.replace('.xlsx', '.csv')
            df.to_csv(csv_filename, encoding='utf-8', index=False)
```

Getting all the files in a directory and converting into .csv by calling the above function

```
In [ ]: # Need os & glob modules

        os.chdir("C:\\Users\\Md Moniruzzaman\\Desktop\\Test")
        # This will change the current directory to this location
        # Can avoid it by giving the full path to "glob"

        for f in glob.glob("*.xlsx"):
            xlsx_to_csv(f)
            print(f + " Is converted")
```

Merging all the CSV files into a single CSV file

```
In [ ]: # import all csv files from folder

        csv_files =  glob.glob("C:\\Users\\Md Moniruzzaman\\Desktop\\Test\\*.csv")

        with open('Demo.csv', 'wb') as outfile:
            for i, fname in enumerate(csv_files):
                with open(fname, 'rb') as infile:
                    if i != 0:
                        infile.readline()  # Throw away header on all but first file

                    # Block copy rest of file from input to output without parsing
                    shutil.copyfileobj(infile, outfile)
                    print(fname + " has been merged.")
```

```
In [ ]:
        print(count_of_records)
        df = pd.read_csv("Demo.csv")
        df.shape
        df.columns
        df.columns = df.columns.str.replace(" ","_")
        df.columns
```