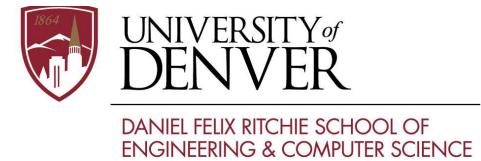Gaper Taxroll Scrubber

Duncan Ferguson
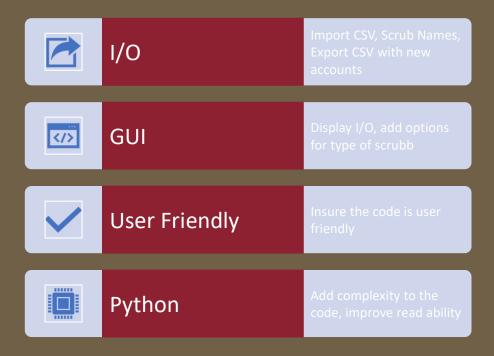
Student ID: 871-641-260

Class: COMP-3006-2

Teacher: Dalton Crutchfield@du.edu

UNIVERSITY *of* DENVER
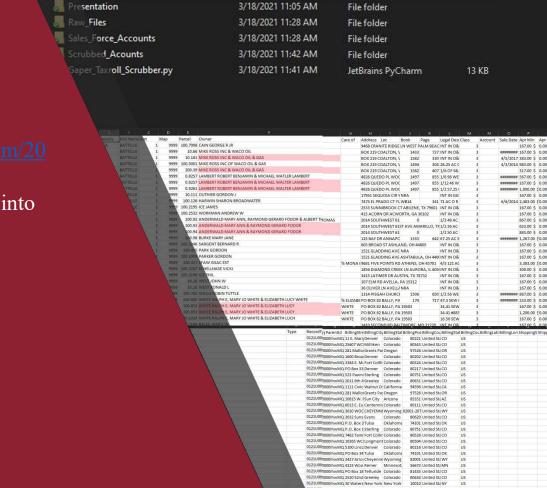
DANIEL FELIX RITCHIE SCHOOL OF
ENGINEERING & COMPUTER SCIENCE

# Scope of Project

UNIVERSITY *of* DENVER

DANIEL FELIX RITCHIE SCHOOL OF
ENGINEERING & COMPUTER SCIENCE

| | |
|---|---|
| **I/O** | Import CSV, Scrub Names, Export CSV with new accounts |
| **GUI** | Display I/O, add options for type of scrubb |
| **User Friendly** | Insure the code is user friendly |
| **Python** | Add complexity to the code, improve read ability |

```python
def lastnameswap(namesplit):
    """This changes the order of the names. Example: Last, First, Middle
    Ending Example: First, Middle, Last. The returned value is in string form"""
    newname = ""  # Creating a new blank string
    z = 1
    while z < len(namesplit):
        newname += " " + namesplit[z]  # adding a space then the name starting from the second position
        z += 1
    newname += " " + namesplit[0]
    newname = newname.lstrip()  # Stripping out the leading space
    return newname


def lastnamesuffixswap(namesplit):
    """This changes the order of the names if it sees a suffix for a swap. Example: Last, First, Middle, Suffix
    Ending Example: First, Middle, Last, Suffix. The returned value is in string form"""
    newname = ""
    z = 1
    while z < len(namesplit) - 1:
        newname += " " + namesplit[z]
        z += 1
    newname += " " + namesplit[0] + " " + namesplit[-1]  # Moves the lastname right before the suffix
    newname = newname.lstrip()
    return newname


def endoption(newnrasum):
    """This function lets you choose whether to save the dictionary or
    print it newnrasum dictionary or saves it to a CSV File"""
    print("\nThe Raw List has been scrubbed\n")
    ending = input("Press 'P' or 'p' to print NRA Sum Dictionary\n"
                   "Press 'S' or 's' to save CSV File\n"
                   "Press 'Any other Key to Exit\n")
    if ending == 's' or ending == 'S':
        makefile(newnrasum)  # This runs makefile()
        print("File is now saved in the Same Directory")
    elif ending == 'p' or ending == 'P':  # Simple Printing of dictionary
        print("Key = New Owner Name : Value = [NRA, Raw Name]")
        print(newnrasum)
    else:
        exit()
```

# The Start

- Creation of the Blog Spot
  - https://duncanfergusonpython.blogspot.com/2021/02/it-hit-me-today.html
  - Blog walks through the in steps that went into creating the Gaper Taxroll Scrubber.
- Starting Data
  - Program File
  - Raw Taxroll File
  - SF Accounts

# Main Function

- Asks user to upload a list
- Ask user to select columns with name
- Ask user to upload a master account file
- Ask user to select column with name
- Ask user to select return Id column
- Ask user where to save new accounts
- Glorified index(match()) with scrubbing abilities and user prompts

```python
def main():
    """Main Function for running the code. Most of the sure prompts are located here."""

    # Asks User which file they would like to use for the Raw Tax Roll File
    rawlist = startoption("Welcome to the Gaper Taxroll.\nTo start we will have to import a
                           "Click the start button to proceed.")

    # Asking User to select the column where the accounts that need to be scrubbed are loca
    rawlist_account_col = HeaderGui("Click the button that contains the \nACCOUNTS/NAME/OWN
                                    list(rawlist.columns.values))

    # Scrubbing names to be in proper form
    names_scrubbed_list = fixnames(rawlist, rawlist_account_col.choice)

    # Asking User to located Salesforce Accounts
    sf_accounts = startoption("We now have to select an update the salesforce accounts file

    # Asking user for column of name to be scrubbed and return Id column
    sf_account_col = HeaderGui("Click the button that contains the \nNAMES for comparing ac
                               list(sf_accounts.columns.values))
    sf_ids = HeaderGui("Click the button with the IDs to be returned", list(sf_accounts.col

    # Looking up the ids on from the selections
    raw_file_w_id = lookup_ids(names_scrubbed_list, sf_ids.choice, sf_accounts, sf_account_

    # Creating new accounts sheet
    new_contacts = create_new_accounts(raw_file_w_id)

    # Prompting user to save new accounts sheet
    endoption(new_contacts)


if __name__ == "__main__":
    main()
```

# GUI Class Creation



- Creation of GUI's to help with user prompts

- Creation of GUI to help with selecting the correct columns

- Class function makes it easy to create new prompts without having to recode tkinter

- Functionality allows text to be passed through into the GUI

# Fixing Names



- Re-wrote code to hand pandas
- Kept lastname and suffix swaps
- Open ability to plug in future code, i.e. machine learning
- Introduction to handling pandas more

# Lookup and Filter



- Panda lookup
  - Simple code
  - Took me forever to figure out and led to rewrites
- Filtering the accounts
  - Taking out the matching IDs
  - Taking out business accounts
  - Dedubbing names

# Save File and Conclusions



```python
def endoption(panda_file):
    """This option asks the user if they would like to save their n
    dataframe and converts it into a csv through the GUI Class to b
    decide = GUI([True, "Where Would you Like to save the new Accou
                        "Note, you will have to check the cities be
                        [True, "Exit"])  # Save option through looping gui
    if decide.button_1_pushed:
        filepath = asksaveasfile(defaultextension=".csv")
        filepath.write(panda_file.to_csv(header=True, index=False,
        exit()
    else:
        exit()
```

- Program is user friendly

- Produces save new account sheet

- Needs more work done to addresses for double names

- Learned a lot about classes, tkinter and pandas

- Can be expanded upon and turned into actual software

- Would love to plug machine learning into it to full go through complexities of names and addresses

- Want to learn how to save it as a python package and make it easy to send to other people



| SF_Account_Names | Address | Street_Address_Raw | City | State |
|---|---|---|---|---|
| ADRIAN L HAUGHT | 5260 DUPONT RD PARKERSBURG, WV 26101 | 5260 DUPONT | PARKERSBURG | WV |
| AGNES M RITZ | 69 HEISKELL AVE WHEELING, WV 26003 | 69 HEISKELL | WHEELING | WV |
| ALEX HENNEN | 4341 MASON DIXON HWY MAIDSVILLE, WV 26541 | 4341 MASON | MAIDSVILLE | WV |
| ALFRED L MANCINI | 1228 PHILIP ST MORGANTOWN, WV 26501 | 1228 PHILIP | MORGANTOWN | WV |
| ALICE HAUGHT | PO BOX 62 WANA, WV 26590 | PO BOX 62 | WANA | WV |
| ALICE BALES GRAIGER | 538 BUTTERNUT LN MATTHEWS, NC 28104 | 538 BUTTERNUT | MATTHEWS | NC |
| ALICE FAYE DEMASTUS | 3 BUCKINHAM PL PINEHURST, NC 28374 | 3 BUCKINHAM | PINEHURST | NC |
| AMANDA M BURNS | 454 MARANDA RUN RD GRAYSVILLE, PA 15337 | 454 MARANDA | GRAYSVILLE | PA |
| ANDREW M CUNNINGHAM | 910 STOUT ST BRIDGEPORT, WV 26330 | 910 STOUT | BRIDGEPORT | WV |
| ANDREW W WORKMAN | 415 ACORN DR ACWORTH, GA 30102 | 415 ACORN | ACWORTH | GA |
| APRIL RICCI | 217 MAIN ST BENTLEYVILLE, PA 15314 | 217 MAIN | BENTLEYVILLE | PA |
| AUBREY L HAUGHT | 5260 DUPONT RD PARKERSBURG, WV 26101 | 5260 DUPONT | PARKERSBURG | WV |
| BARBARA A TAYLOR | PO BOX 93 HUNDRED, WV 26575 | PO BOX 93 | HUNDRED | WV |
| BARBARA ANN BELL TROUP | 1000 GOOD HOPE RD MECHANICSBURG, PA 17050 | 1000 GOOD | MECHANICSBURG | PA |
| BARBARA CORBETT | 626 CABOT CIR JONESVILLE, MI 49250 | 626 CABOT | JONESVILLE | MI |
| BARBARA E SHRIVER | 9050 MASON DIXON HWY WADESTOWN, WV 26590 | 9050 MASON | WADESTOWN | WV |
| BARBARA ELLEN HENDERSON | 1 NETHERLAND BLVD SCHENECTADY, NY 12306 | 1 NETHERLAND | SCHENECTADY | NY |
| BARBARA J HOBDAY | 3197 MAYFLOWER LN SUMTER, SC 29150 | 3197 MAYFLOWER | SUMTER | SC |
| BARBARA P PALMER | 215 BUCKHANNON AVE CLARKSBURG, WV 26301 | 215 BUCKHANNON | CLARKSBURG | WV |
| BARBARA SNYDER | 1336 WILSON CIR MORGANTOWN, WV 26505 | 1336 WILSON | MORGANTOWN | WV |
| BERNARD R SARGENT | 605 BROAD ST ASHLAND, OH 44805 | 605 BROAD | ASHLAND | OH |
| BERTHA ELLEN PRICE | 5260 DUPONT RD PARKERSBURG, WV 26101 | 5260 DUPONT | PARKERSBURG | WV |
| BETTY J CROSS | 524 MILFORD ST MORGANTOWN, WV 26505 | 524 MILFORD | MORGANTOWN | WV |
| BETTY JEAN MCCONNELL | 92 OGDEN FARM RD MOOREFIELD, WV 26836 | 92 OGDEN | MOOREFIELD | WV |
| BETTY LYNCH | 440 DRUMMOND ST MORGANTOWN, WV 26505 | 440 DRUMMOND | MORGANTOWN | WV |
| BETTY M MANNING | 40622 CARLISLE AVE ELYRIA, OH 44035 | 40622 CARLISLE | ELYRIA | OH |
| BILLY (AGT) ATKINS | 3122 N GREYSTONE DR MORGANTOWN, WV 26508 | 3122 N | MORGANTOWN | WV |
| BOB MICHAEL | 96 MIDDLE FORK RD BURTON, WV 26562 | 96 MIDDLE | BURTON | WV |
| BONNIE CAIN | 645 HILLWOOD DR MORGANTOWN, WV 26508 | 645 HILLWOOD | MORGANTOWN | WV |
| BONNY L CAIN | 231 DEER AVE NICEVILLE, FL 32578 | 231 DEER | NICEVILLE | FL |
| BRYAN K SNYDER | 160 ROCKY RIDGE RD WAYNESBURG, PA 15370 | 160 ROCKY | WAYNESBURG | PA |
| C D TENNANT | 5260 DUPONT RD PARKERSBURG, WV 26101 | 5260 DUPONT | PARKERSBURG | WV |
| C RAY MARTIN | 5260 DUPONT RD PARKERSBURG, WV 26101 | 5260 DUPONT | PARKERSBURG | WV |