Student: Andrew Palmer

Date: Mar 21, 2023

Course: CSE 111

Professor: William Clements

Semester Final Project (covering wks 11-13)

Week 11 Submission

Due: Mar 18, 2023

CSE 111 Proposal for a Student Chosen Program

(This is a proposal. All proposals are written with our best

knowledge at the beginning of a project. As we progress from the

beginning to the end of a project, we gain more knowledge, and we

change some of our proposed ideas. This is part of completing any

project and is fine. However, we still must write a proposal so

that our teachers or supervisors know what we intend to do. Please

answer each of the following questions to the best of your

knowledge.)

1. What is the title of your program?

Title: The Merging, Cleaning, Sorting (& Possibly Mapping) of PBC Multi Media’s Printed Phone Directory Dictionaries.

Filenames Used:

directory\_requests.csv

directory\_listings.csv

merged\_directories.py

test\_merged\_directories.py

2. What real-world problem will your program address or help to solve?

My wife, Collette, and I are the second-generation owners of a multimedia company based in Brandon, Manitoba Canada. From 1996 till about 2016, the company (then known as *The Phone Book Company*) was exclusively involved in the business of publishing printed phonebooks, widely recognized as the best of such in the areas in which we operate.

Since that time, a lot of change has happened. Beyond COVID, we’ve expanded the product lines and change the company brand. So now we’re *PBC Multi Media*, and we aim to aid businesses with all sorts of search solutions. This includes being found in print through the phonebooks; it also includes assisting with internet search, online listing management, web development, and more. Indeed, while our products are all very different (ie serving businesses in different ways), they are also very similar (ie aiming to solve the age-old problem of ‘how can I get found by new/existing customers?’).

This program, however, won’t look to address that. Nor will this paper, except for providing a little bit of backstory, as has been done. And now for the matter at hand: *what will this program look to accomplish?*

First a little more backstory:

Recently in our business we have undertaken another first: in the past month we’ve moved from a ‘blanket distribution’ model for the delivery of our printed directories to a ‘eco-friendly, opt-in’ model. In undertaking this, we announced it through a direct mail campaign to 93,500 homes and businesses in the area. The first flyers went out last Tuesday, and to date there’s been a literal non-stop request line of people contacting us, asking to be included on the list. Plus, on average we’re seeing roughly 250 daily via mbphonebooks.ca. All in, we’ve so far past the 2,000 order mark, and we expect to reach about 15,000 – 20,000 when it’s all said and done.

This is where computer programming can help.

Currently we have database of tens of thousands of personal and business landlines, and of thousands of opted-in cell listings. It’s arranged as a dictionary in Python terms. That is, the phone numbers are all unique, while other info (town, name, address) may be the same.

Similarly, we have this now growing secondary csv file. It too can be arranged as a dictionary based on the same criteria. With it, we intend to organize teams of delivery groups to hand deliver the directories right to the doorsteps of those living in towns and cities, and to mail (for a $5 postage fee) to those living rurally. (FYI the mailers that went out spoke to this difference, and some received a code for free delivery).

The ‘real-world’ problems that I will attempt to solve will be to do the following:

1. Clean up the directory\_requests.csv file, or create an altogether new file as much as possible. It’s not that we fear that the data isn’t well-organized as it is, but rather we recognize that when some either went online or mailed in their orders they didn’t provide a physical address. And if that’s the method of the delivery, we want to efficiently check it in the database of listings and append it into the request file if it’s been found there.
2. If the listing doesn’t have a physically deliverable address in the directory database, then the program will need to pass that with a try block function.
3. Following this cleansing, the program will be helpful if it can sort the request directory in the following manner:
   1. By mail/at-home delivery
   2. By town/city/area
   3. By street
4. The real bonus would be if I could get it to actually map the delivery addresses out so that we could give them to a group and they could plan a route accordingly. This may not be possible for me given the time constraints I’m under (and my novice level of expertise in these matters), but if it can be done that would be ideal.

3. What will you learn from developing this program?

I feel like the most important thing I will gain in developing this program is confidence: confidence that I have the ability to write a program from the ground up. This, I believe will be even more ‘useable’ than the functions that are created.

Aside from that, I hope to learn to use dictionaries, lists, and objects in a more meaningful way. And I hope that I can learn how to apply aspects of Python in a manner that I’d not been taught before. What I mean by that is that it’s my hope that I can write programs that would use altogether new elements for me.

I also hope to be able to use this exercise to demonstrate to some of those around me (ie family/friends) some of the things I’ve been learning, and shine a light on the real-world learning benefits of this endeavour (ie of me going back to school).

4. What Python modules will your program use?

(Some examples are: csv, datetime, functools, matplotlib, math,

pandas, pytest, random, requests, and tkinter.)

I expect to use the following Python modules:

1. Csv – for reading and writing to/from csv files
2. Pytest – for testing functions
3. Os – for its clear screen functionality
4. Colorama – for its ability to display in different fonts
5. The Try Block (Try, Except, Else, Finally) – for its ability to find and handle errors smoothly
6. Some form of mapping module, possibly GeoPy, folium or plotly. This could involve the geolocator or nominatim elements from Geopy.

5. Will you separate your Python program into functions that each

perform a single task?

(Remember that the most reusable functions don't get user input and

don't print results but instead have parameters and return a result.

Functions that get user input and print results are important and do

useful work but are not easily reusable.)

(Remember also that it's hard to test functions that get user input

and print results. It's easy to test functions that don't get user

input and don't print results but instead have parameters and return

a result. Therefore, you should write most of your program functions

to have parameters and return a result.)

I absolutely will. Splitting functions into individual tasks is right because it keeps things organized, is testable, is reusable, is more easily readable, and is otherwise the better way to code.

6. Will you write test functions to test at least two of your program

functions?

I absolutely will. And I’ll be running Try blocks to accommodate for functional errors along the way too.