My Final Project

Moataz Basyouni

February 2020

1 Abstract

Talk about the program and the data and what was found out about the data after running the program on it.

Contents

List of Figures

2 Introductiont

Talk about the purpose of the program and what it does. Talk about who can use it and its importance.

3 Methods

Talk about the code used and how I made it. Talk about how its run by Python through the terminal.

```
#!/usr/bin/env python

def openfile(fname):
    import numpy
    import pandas as pd
    tmp_data = pd.read_csv(fname)
    data = tmp_data.to_numpy()
openfile('patientdata.csv')
#defines a function openfile so that whenever the function
#is called with the name of a file, it will open that file

def maxage(fname):
    import numpy
    import pandas as pd
    tmp_data = pd.read_csv(fname)
    data = tmp_data.to_numpy()
    age = data[:,4]
    age_max = numpy.max(age)
    print("The oldest patient is ", age_max, " years old.")
maxage('patientdata.csv')
#this function looks through the fourth column in the file
#and finds the max value. Sifting through the fourth column
#which includes the ages, and finding the max value will help
#find the oldest person in the list

print("\n")

def display(source, numlines = 3):
    print("Here is a preview of the file:")
    with open(source) as f:
        for i, line in enumerate(f):
            print(line.strip())
```

```
            if i == numlines:
                break
display('patientdata.csv')
#This function gives a preview of the file with the number of lines
#specified. If number of lines isn't specified, it will show a max
#of 3 lines.

print("\n")

def displaydict(source, numlines = 3):
    print("Here is a preview of the file in a dictionary")
    import csv
    with open(source) as f:
        reader = csv.DictReader(f)
        for i, row in enumerate(reader):
            print(dict(row))
            if i == numlines:
                break
displaydict('patientdata.csv')
#This gives a perview of the file in a dictionary.

print("\n")

def importgenderdata(source, output, gender):
    print("Here is a preview of the female patients only")
    import csv
    assert gender != "Female" or "Male", "Please ensure first letter of gender is capitalize
    with open (source) as fr:
        reader = csv.DictReader(fr)
        header = reader.fieldnames
        with open (output, "w") as fw:
            writer = csv.DictWriter(fw, fieldnames = header, delimiter = ",")
            for row in reader:
                if row["sex"] == gender:
                    writer.writerow(row)
importgenderdata('patientdata.csv', 'femaledata.csv', 'Female')
with open('patientdata.csv') as f:
    for i, line in enumerate(f):
        print(line.strip())
        if i == 3:
            break
#This function imports the female data only into a new file,
#it then previews some of it as done before
print("\n")

def hist(source):
```

```
    import numpy
    import pandas as pd
    tmp_data = pd.read_csv(source)
    data = tmp_data.to_numpy()
    import matplotlib.pyplot as plt
    ages = data[:,4]
    x = [ages]
    plt.style.use('ggplot')
    plt.hist(x, bins = 10)
hist('patientdata.csv')
#This functions uses data from column 4 (ages), and builds
#a histogram with it. It can be used as a visual representation
#of all the ages of the patients

print("\n")

def julydates():
    import csv
    import re
    with open("patientdata.csv") as fr:
        dictread = csv.DictReader(fr, delimiter = ',')
        header = dictread.fieldnames
        dates = []
        for row in dictread:
            dates.append(row['treatmentDate'])
    print('Here are the treatment dates which occured in July')
    date_regex = re.compile(r'7/\d*/\d{4}')
    uniqdates = set(dates)
    for date in uniqdates:
        if re.match(date_regex, date):
            print(date)

julydates()
#This function displays only the treatment dates which
#occured in July of 2015
```

4 Results

Talk about the outputs of my program after inputting my dataset.

4.1 Figures

Insert the histograms and discuss them

5 Conclusion

summarize every section

6 References

Reference stackoverflow and the other sources used