

Appingways Internship Report

This document gives a brief description about the files contained in the folder

This folder consists of the solutions of problems given in test.docx

Libraries included in the solutions :

- 1.pandas
- 2.numpy
- 3.matplotlib
- 4.pylab
- 5.geopandas
- 6.datetime

If libraries are not present in your system install them by using
(sudo pip install library-name)

Programs of python are written in Python 2

Description of the files:

1. pyth1.py :This file consists solution of 1st Question in python.

Input: 2 lists

Output: List containing common elements in 2 lists

Elements of list 1 which are not elements of list 2.

In this program I used inbuilt functions intersection and not in to find common elements and elements not in list 1 and not in list 2.

2. Pyth21.py : This file consists of solution of 2nd Question in python.
No input , **Output:** Number of Thursdays between 1990 to 2000.

In this program I used datetime library to find the week day of the first day of year. Later I checked number of Thursdays in each year and added them.

3. Jvs.js : This file consists of solution for Java Script problem .
Input : List

Output: List which has elements of input list except zero.

In this program I first took a list as input and then created a null list and then added the elements of list 1 which are not equal to zero.

4. Usecase2.py: This file consists of solution of 1st problem in use case 2.

Input : india-census data file

Output : Geographic map indicating literacy rates of each state.

In this first extract the state name, population and literacy and then group by state and divide population by literacy to get literacy rate. Now extract the geographic boundaries from github repo and then plot India Map and then literacy rates of each state.

5. usecase22.py: This file consists solution of problem 2 in use case 2

Input: india-census dataset

Output : Data frame consisting data of similar districts.

In this program first I extracted data of Tamil Nadu and Bihar and then normalized all the columns (except District name , state name , District code). Then I took the dot product and then set a threshold value of 1 and checked if the value is less than threshold value then the districts are similar and appended the value to the dataframe

6.usecase23.py: This file consists solution of problem 3 in use case 3

Input: india-census dataset

Output: Scattered map showing the state and its mobile penetration rate with respect to agricultural workers.

In this program I extracted the State Name , District Name

Agricultural_workers(AGRW) and

Households_with_Telephone_Mobile_Phone_Mobile_only(TMO) and store them in a dataframe

Then grouped the data with respect to the states and then divided TMO with AGRW.

Then plot data of states and Mobile penetration rate using pylab and matplotlib.

7 figure_1.png consists of output of 3rd problem in use case 2

8 figure_1-1.png consists of output to 1st problem in use case 2