

PAKISTAN ESA KIUN HA?

Analysis of algorithm



JULY 2, 2020

SUBB BY: ZIA UL REHMAN SUBB TO: Dr. Junaid Akhtar

ABSTRACT

To check the state of specific country or continent economically, socially and regarding all perspectives can be determined and analyzed only it is known that on which position that country is lying on the globe of world. In this project named as "PAKISTAN aesa q ha" we have to use different indicators which indicates the condition of Pakistan regarding good and bad factors. By the help of indicators and ranking of Pakistan among different countries we can know that "Pakistan aesa q Ha".

Contents

INTRODUCTION	2
OBJECTIVE	2
DATA ANALYSIS	2
STATEMENT OF PROPLEM	3
AIMS	3
BASIC TECHNIQUES	3
DEFINITIVE NEEDS	4
OPERATIONAL REQUIREMENTS	4
UP AND RUNNING	4
CODING	5
Description Of Given Functions	5
Ranking Code	5
Ranking Country	5
Pakistan Aesa Q Hy	6
SUMMARY	6
ANALYTICAL SURVEY	7
CONCLUSION	7
ACKNOWI FDGFMFNT	9

INTRODUCTION

Ranking of country is best option to check its state. Good indicators helps to increase the condition of country and helps to be a successful and developed. It means country is not depending on other countries regarding any help and support. Like big GDP rate of country is better for its progress. Bad indicators should be less in number if a country is willing to be developed. Bad indicators are negative factors and cause of damage in economy. Indicators indicate the perspectives of anything. For this purpose first of all ranking of Pakistan is helpful and other than this comparison indicates the state of country. Python programming language is emerging day by day. Different data structures are also need of it. In this project set, list and different dictionaries are to be used. Duplications of country is not good fact so to overcome this fault, set is being used which is good data structure. For changeable and ordered sequence "list" has to be used. And for optimization and retrieval of values dictionaries are helpful data structure. With the help of these data structures and file reading ranking is major task to find out the rank of any country of world.

OBJECTIVE

The major objective and purpose is to handle big amount of data which contains many different values and to align them in a sequence and order. We should well aware of programming languages and basic data structures. Python is updated language are in this project requirement is to work in python in a specified way.

DATA ANALYSIS

A CSV contained a huge data about almost two hundred twenty seven countries in a gap minder frame. With good and bad indicators. A large unfiltered dataset and a basic code is given. Different countries have different indicators detail throughout a large period of time. Data file is covering huge data of centuries like 1980 to 2010 for example. Different variables are meaningly related to each other and help to be compared regarding comparisons of countries. Data of countries is given different row and column wise. Row and column wise relation is helpful to compare regarding time too. By relation and plotting in a simple manner can to understand in a simple way. Some factors are permanently and few are temporary which play an important role

together. If Average of bad indicators is more than good and effective ones so it will definitely draw back and will negative point and threat for a developing country.

STATEMENT OF PROPLEM

The major problem was to align countries in an order. Resolving faults and difficulties regarding data. First given file is to be read by file handling than observe which data structures are helpful to be used. We have to find out in which point Pakistan is lying among given two hundred twenty seven countries. Some indicators are good with comparison of other country but eventually for another that country is going down with comparison of indicators of next country. So graph is difficult but interesting to plot. So ranking is helpful for a country. Max of indicators supports to tackle it. After ranking we have to compare some countries and then analyses why and which one is going up and which is going down.

AIMS

Tackling the problem of real world is main objective of today's technical life of human. So the basic needs and aims were to tackle the problem in the given data of project.

- 1. Data was huge and can mixed up together, so arranging in sequence was basic aim so that algorithm and techniques should impose on every bigger size of data file.
- **2.** Grip on programming language and installation of header files and libraries should be strong enough.
- **3.** Algorithm should take less space and less time to implement. Memory consuming issue should not be a fault of project.
- **4.** The code should be working and fulfill the entire given requirement and should not halt at any points (corner points of algorithm should be taken care of).

BASIC TECHNIQUES

If size of any related data is big enough so first iteration is to arrange well.

Dictionaries, which are being used in project and basic needs to store time and work efficiently. The time complexity of dictionary is B (O (1)). The data was given in a commaseparated file. So there was a big chance to repeat same data for different countries and

that's why list and set were important data structures to lessen this problem in a while. Because if data is repeated which is not require so it would be negative point of implementation. Logic should be strong and grip on language "like python which is being used "should probably strong.

DEFINITIVE NEEDS

An algorithm should completely definitive and well grounded. In this project ranking of country is major work. So if we are doing for one specific country Pakistan, it should full fill requirements and needs of a user. In this data file if more than two hundred countries are involved so this is not big level of analysis. An efficient algorithm is not dependable on volume. Its complexity should low and working speed should fast. And if data of countries or states or continents is given in millions of number along hundreds and millions of indicators so algorithm should implement on it with no error or logic problem. It means logic should be so reliable and complexity should minimum.

OPERATIONAL REQUIREMENTS

In every ranking algorithm, it should be taken into consideration that the user can check the ranking system with every aspect that is possible in the given data. For example Pakistan can be compared with every other country and all of its indicators can be checked with every other country and the max values that were given in order to clarify that the algorithm is ranking correctly so that it satisfy the user. If a user want to check a specific, country and a specific rank it can be done easily and efficiently. To do this task fast and easy we used dictionaries for that, with values of country name and indicator name.

All the basic operations are added in this project, which can be used to drive some more advance operations.

UP AND RUNNING

The functions which are included in program should work efficiently and their logic should not complex and should be easy to understand. Variables names should be genuine and powerful so these factors increase the effect of work. Functions in the program are general and can be modified easily for further use.

CODING

First of all some of the basic code was provided to us by our professor which help us to understand the basics of python and to get the view of data provided and how to manipulate it.

The basic provided functions are described below.

- 1. DataTypeConvertion
- 2. FetchIndices
- 3. FetchColumnData
- 4. FetchData

Description Of Given Functions

DataTypeConvertion Convert the data into any types according to the requirements. Like a list of string to integer data type for years (1952 -2012). Same indicators values from string to float values.

FetchIndices fetch the index for a specific value like for Pakistan its fetch each row number where Pakistan exits

FetchColumnData fetch all the data from a given column like in program we used **set** for countries name to get countries name without repetition by giving the index of countries to fetchcoloumndata.

FetchData get the data from 2d_list. We pass column index and row index and its fetch the data from it.

At the last a dictionary was provided countries Dict whose key values were Country Name and Indicator Name and it returns the values of that country and indicator.

Ranking Code

Now here starts the ranking system. By using the countries Dict provided by the professor I compute the average sum of all countries and there indicators and store them in a file mycsv.

This file was created to extract the required data from Gapminder file and to get a visual idea that my program is working fine.

After that the remaining code written in project.py file. Now I read the data from the file and make a list from it and make a dictionary countries Dict, which gives the average sum of country indicator values. While creating the file I also stored max and min values of indicator for future use. In order to make the code clean I separated the functions in functions.py file. All the code in commented to understand it easily

Ranking Country

Now by using two more dictionaries rank_by_indicator and total_rank_dict. I rank each country with respect to its indicator and give a rank value to it.

By using them, I detect Top 10 Countries and Pakistan Rank and user can find the rank of desired country.

Pakistan Aesa Q Hy

In order to give answer to this question I have divided indicators in 4 different parts Best, Good, Bad and Worst.

Bad indicators are in worst condition if they have high value and Good indicators are in worst condition if they have low value with respect to max values.

User can check which indicators are in worst condition and affecting the rank of value and which are in best and good condition and compare them with the indicators of Top 10 and other countries to get the idea 'Pakistan Aesa Q Hy'

Indicators Which Are In Worst/Bad Conditions In Pakistan

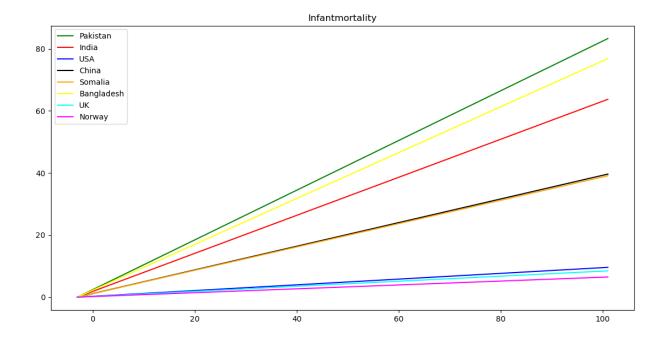
('Pakistan', 'ChildrenPerWoman') Worst

('Pakistan', 'Infantmortality') Worst

('Pakistan', 'Tradebalance') Bad

Graphical Representation

For user to get the idea fast I also provided the graphical representation in it.



SUMMARY

Project contained some basic operational requirements like if a country is going up or either going down so which factors and which indicators are effecting it in a good or bad way. Ranking of

countries, than comparison with other countries, find out there total max, conversion of data types and file reading was basic need.

ANALYTICAL SURVEY

I have used different ways to get the best result possible and not to miss even a single decimal value. For that, I use float data types now user can get the precise results form it. Values contain 7-8 decimal places. So even a single point can affect the ranking system and give accurate ranking. Float values also help to prevent the collisions between rankings.

I have divided indicators into 4 categories Best, Good, Bad and Worst and design a scale for each category. Now user can get idea by theoretical result without numerical result because it's easy to understand and give the reason which indicators are affecting the most to go up and down.

Number of Indicators For Pakistan

Number of **Best** Indicators: 22

Number of Good Indicators: 24

Number of Bad Indicators: 1

Number of Worst Indicators: 2

Number of Indicators For Norway Number of Indicators For China

Number of **Best** Indicators: 26

Number of Good Indicators: 20

Number of Bad Indicators: 0

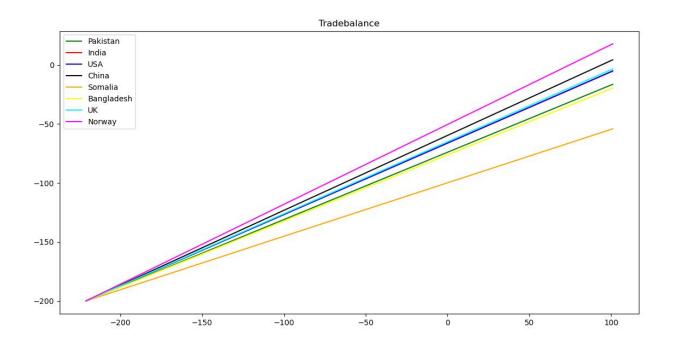
Number of Worst Indicators: 3

Number of **Best** Indicators: 20

Number of Good Indicators: 23

Number of Bad Indicators: 4

Number of Worst Indicators: 2



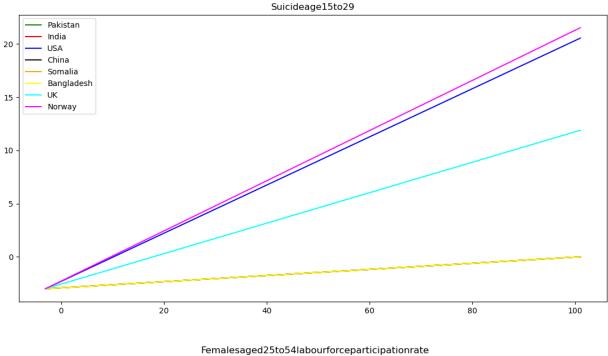
CONCLUSION

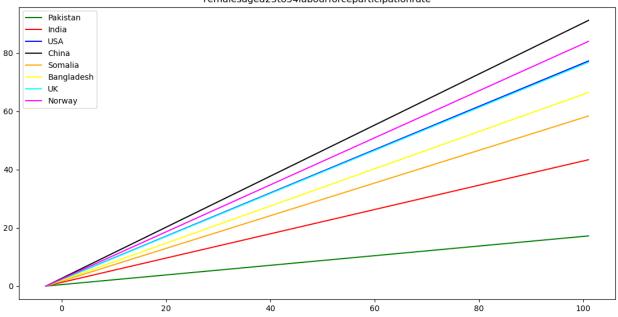
With every passing day world is moving towards technology and advancement and data science is a broad field. Data Science is used to read the given data and design an algorithm for it so that can

easily understand it. This project help us to understand how data it manipulated to get the required result and the field of Data Science is working. How to Design an algorithm and which difficulties can occur in it. So overall its help us a lot in our studies and future understanding. These are the result after my algorithm on given data. I have seen that some of the bad Indicators are high in Top Countries. Given below the graphical representation can give you the idea about it.

Top 10 Countries

Rank Number	Country Name	Rank Value	
1	United States of America	2097	
2	Norway	1843	
3	Luxembourg	1756	
4	Denmark	1739	
5	Sweden	1733	
6	Canada	1727	
7	Belgium	1719	
8	Switzerland	1714	
9	Finland	1696	
10	Netherlands	1692	
129	Pakistan	848	





ACKNOWLEDGEMENT

To achieve success in any field of life is not only the success of a worker oneself but also it is the fruit of efforts of all people belonging to one. I am highly thankful to DR. Junaid who gave me a glowing opportunity to perform this project and by his help, I became able to do it well. Also thankful to my parents and all my supporters.