

**REPORT ON DATA PROCESSING OF COVID-19**

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD**



Faculty:

**Dr UMA SHESHADRI**

**DATA STRUCTURES (CS102)**

Project done by:

**SATWIK-19BCS083**

**RUTWIK -19BCS084**

**DHATRI KIRAN-19BCS093**

**MANOJ SAHIT-19BCS110**

**VEERA GOWTHAM-19BEC010**

**WITTY ELIJAH-19BEC011**

**J.K.NAVEEN-19BEC018**

# Purpose of research topic study:

* The sole and main purpose of doing this project is to know different types of data structures and to learn implementation and using them according to situation.
* When we have trillions of bytes of data it is useless unless we know how to handle it. Handling data includes storing, accessing, and editing it.
* For handling data there are various data structures some of them are:

🡪Arrays 🡪Queue 🡪Dictionary

🡪Linked list 🡪Tree

🡪Stack 🡪Graph

* These are some of the data structures that are commonly used. They are studied and analysed in depth further in this article.

# Observations:

**Data structures**: It is a way to organize and store data so that it can be used efficiently.

**Types of data structures:**

* Primitive data structure
* Non-primitive data structure

**Primitive data structure:**

The data structures which are supported in machine level and have pre-defined behaviour and specifications

EX: int, float etc

**Non-primitive data structure:**

The data structures that are derived from primitive data structures.

These are further divided into two types:

**Linear data structure**:

The data structures that store the data in a linear way fall under this category. Arrays, linked list, stack, queue are some linear data structures.

**Non-linear data structure:** The data structures trees and graphs come under this category as the data isstored in them in random way and retrieving is also same.

* Handling of data is tough than we think. Application of data structures in real life is inhigh demand. The term data science is often used but we can say that data structure is a fundamental topic for approaching data science.
* Each data structure haveits own and unique characteristic that other data structure do not possess for example linked list size is dynamic whereas size of array (though we can change it dynamically it is a long process) is fixed though both are linear data structures.
* So, rather than selecting a data structure blindly we have to know the requirement and should have clear picture on how to use it otherwise our data may become corrupted or we may lose the data.
* For implementation of datastructures we can use any programming language but we used python because using it is quite easy.

# Name of data files used:

* time\_series\_covid\_19\_confirmed
* time\_series\_covid\_19\_deaths
* time\_series\_covid\_19\_recovered
* AgeGroupDetails

# Websites used for reference:

* <https://www.worldometers.info/coronavirus/>
* <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>
* <https://www.statista.com/topics/5994/the-coronavirus-disease-covid-19-outbreak/>

# Data types used in solving problems:

* Lists
* Dictionaries
* Xlsx, text file’s
* Strings

# Result:

* As you start running the program two bar-graphs will be displayed which are drawn using matplotlib library. First graph show’s death rate in high risk countries and second graph show’s death rate in 183 countries.
* Later you will be asked a country as input to see the graphs plotted between number of deaths, recovered, confirmed vs dates for that country.
* Further you will have a choice to see the list of dangerous countries either in a text file or as a list in terminal.
* Further you will be asked for input for a country name to see the graphs drawn with average weekly deaths, confirmed cases, recovered.
* By typing the name of country you can find whether it is a danger country or a safe country for next two years.
* At last a pie chart will be displayed containing different age groups and percentage of deaths in that age group.

# Draw backs of our code:

* As the program takes the data input directly from excel files the data should be updated in excel files regularly.
* But for available data our code gives an accurate and precise visual representation.

THE END