

ASSIGNMENT NO: 1

Group Members:

Name	GR.No & Roll No.
Vikram Shinde	21920048 (322065)
Pranav Wagh	21810340 (322061)
Samruddhi Desai	21920035 (322063)
Avdhut Sagare	21920143 (322073)

Problem Statement: Design Decision support system for any suitable application

Objectives:

- Understanding the dataset
- Plotting graph
- Finding out Statistical calculations such as mean, median, mode etc.

Theory:

A decision support system (DSS) is a computerized program used to support determinations, judgments, and courses of action in an organization or a business. A DSS sifts through and analyzes massive amounts of data, compiling comprehensive information that can be used to solve problems and in decision-making.

1. Functions:

- A. Min ()- Minimum Values of particular data
- B. Max ()- Maximum Values of data set
- C. S.D ()- Standard Deviation of data
- D. Avg () – Average of particular data

- E. Median ()- Median of particular data
- F. Count()- Counting no.of data

2. Data Visualisation:

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data

3. Data Warehouse:

Data warehousing is the process of constructing and using a data warehouse. A data warehouse is constructed by integrating data from multiple heterogeneous sources that support analytical reporting, structured and/or ad hoc queries, and decision making.

Output:

Dataset used: Covid-19 dataset

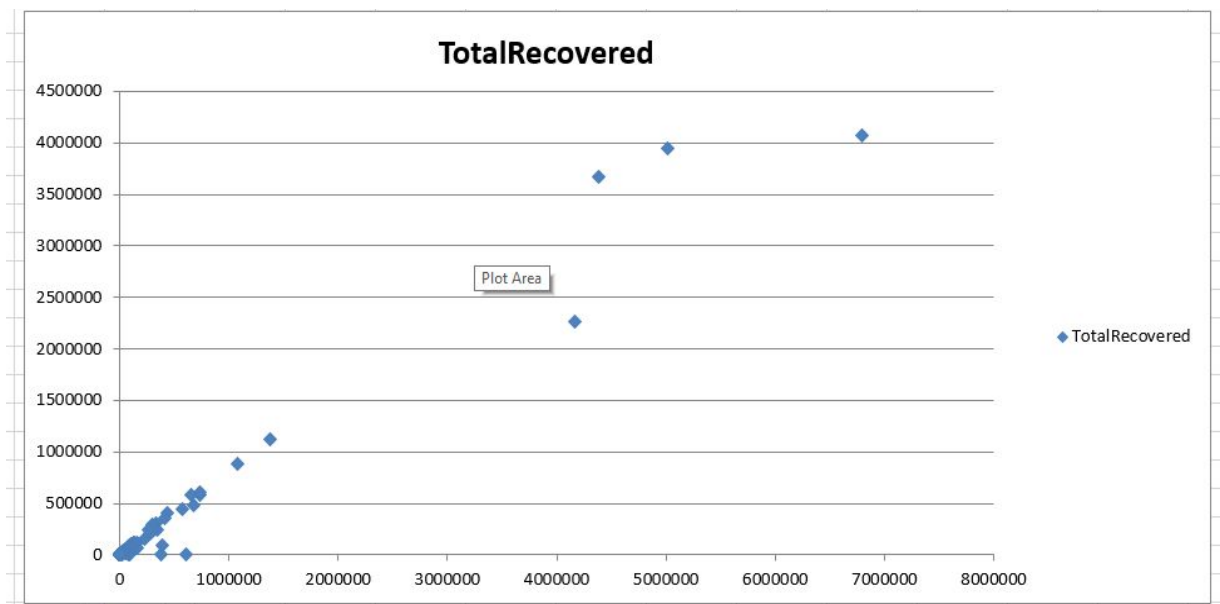
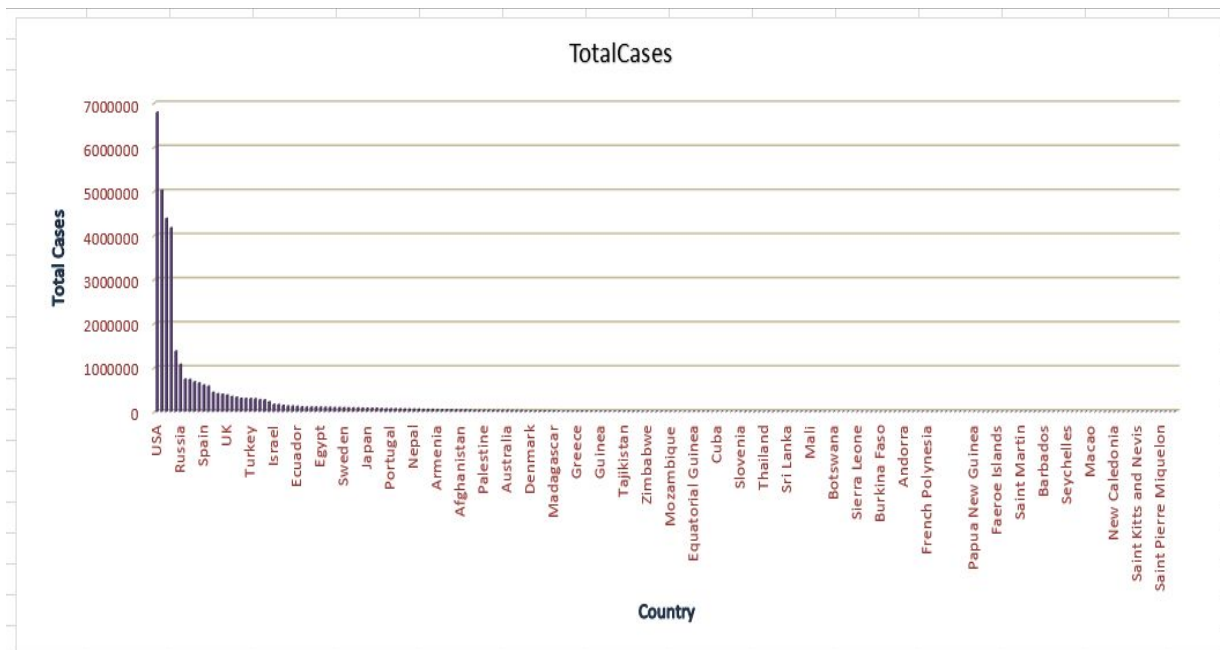
Dataset contains covid 19 worldwide spread till 6th sep 2020, it contains 220 rows and 12 attributes like The total tests conducted, total positive cases and total deaths across all the countries in the world

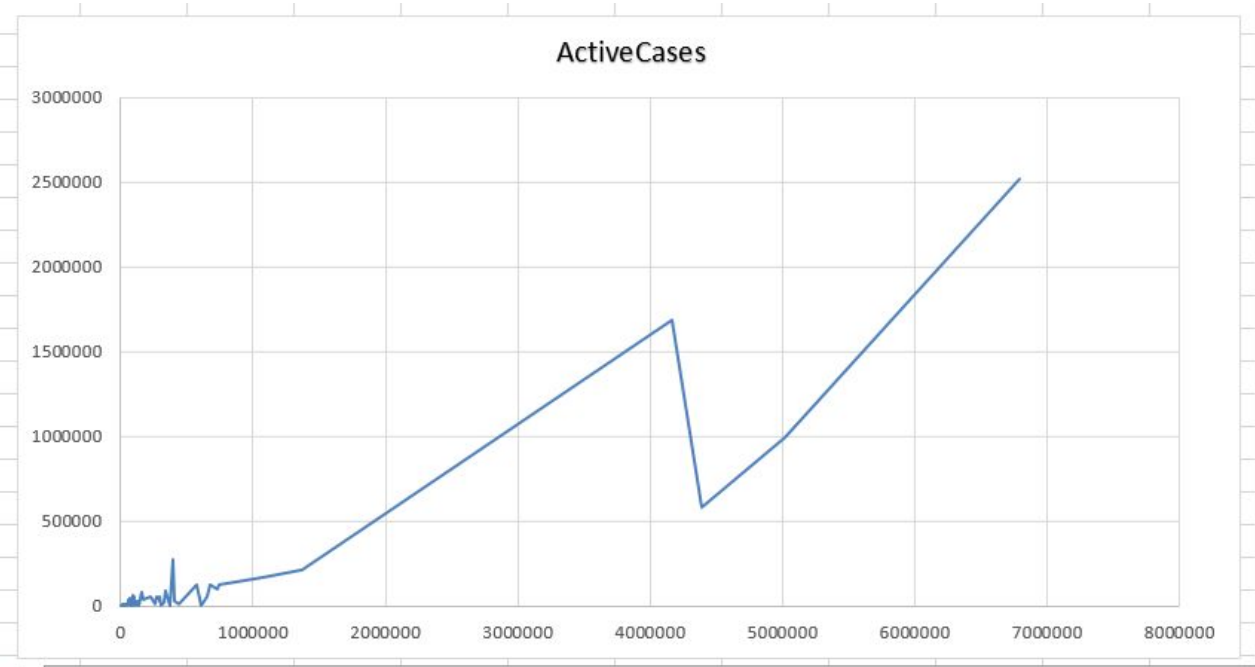
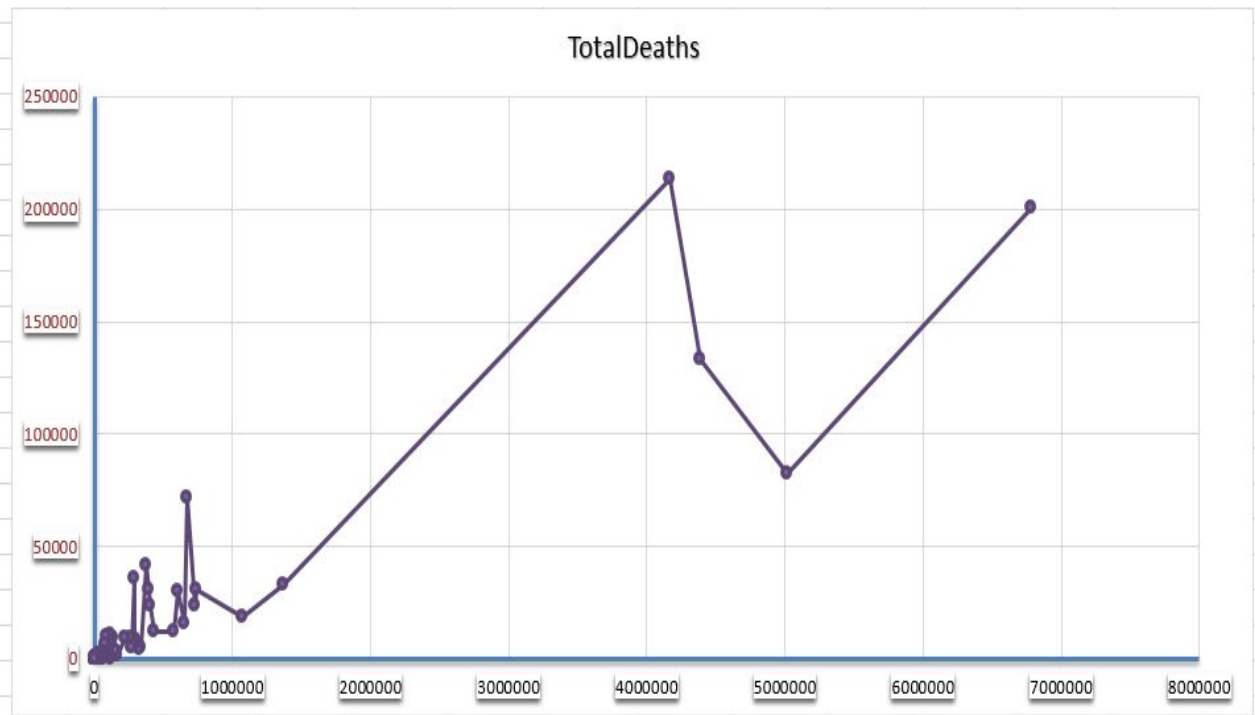
1	Country	TotalCases	TotalDeaths	TotalRecovered	ActiveCases	Serious,Critical	Tot Cases/1M pop	Deaths/1M pop	TotalTests	Tests/1M pop	Population
2	USA	6788147	200197	4068086	2519864	14165	20483	604	93631958	282526	331409864
3	India	5020359	82091	3942360	995908	8944	3631	59	59429115	42977	1382826609
4	Brazil	4384299	133207	3671128	579964	8318	20596	626	14505652	68141	212875572
5	Europe	4169717	213547	2270067	1686103	6678	null	null	null	null	null
6	Africa	1372731	33126	1121263	218342	1428	null	null	null	null	null
7	Russia	1073849	18785	884305	170759	2300	7358	129	41122307	281760	145947739
8	Peru	738020	30927	580753	126340	1460	22319	935	3567927	107901	33066529
9	Colombia	728590	23288	607978	97324	863	14287	457	3213298	63011	50995500
10	Mexico	676487	71678	481068	123741	2747	5235	555	1520463	11767	129214113
11	South Africa	651521	15641	583126	52754	539	10957	263	3940217	66263	59463283
12	Spain	603167	30004			1157	12900	642	10756835	230050	46758621
13	Argentina	577338	11852	438883	126603	3049	12750	262	1602403	35387	45282165
14	Chile	437983	12040	409944	15999	905	22871	629	2887903	150801	19150427
15	Iran	407353	23453	349984	33916	3811	4837	278	3613891	42913	84215113
16	France	395104	30999	89891	274214	759	6050	475	10000000	153130	65304045
17	UK	374228	41664			106	5507	613	20292025	298584	67960923
18	Bangladesh	341056	4802	245594	90660	null	2067	29	1756746	10645	165030582
19	Saudi Arabia	326930	4338	305022	17570	1286	9361	124	5817955	166583	34925213
20	Pakistan	303089	6393	290760	5936	571	1367	29	3024987	13640	221765383
21	Iraq	298702	8166	233346	57190	570	7393	202	1948531	48225	40404908
22	Turkey	294620	7186	261260	26174	1327	3485	85	8742535	103428	84527459
23	Italy	289990	35633	214645	39712	201	4798	590	9943944	164518	60442889
24	Philippines	260467	4653	207353	53083	1040	2453	43	2161465	20700	109982134

Statistic Calculation:

	A	B	C	D	E	F	G	H
1	Column names	mean	mode	median	std	max	min	
2	Total Cases	147003.945	26	5387	645599.5	6431160	3	
3	Total Deaths	5773.55897	1	175	23679.92	209853	1	
4	Total recover	103749.776	651	3627	437151.2	3707000	3	
5	Total test	2725766.91	900	168153	13673223	1.6E+08	120	
6								
7								
8								
9								
10								
11	Paramters	Correlation						
12	case vs death	0.91860923						
13	case vs recover	0.98646442						
14	test vs recover	0.47608499						
15	test vs death	0.46946496						
16	pop vs cases	0.47982255						
17								

Graphical representation:





Inference:

- Most Corona cases are in USA,India,Brazil,Europe
- Maximum number of deaths have occurred in USA
- There is strong correlation between population of country and the total tests done.
- Total cases of Russia are more than total cases of UK still UK has more deaths than Russia thus maybe Health system of UK lacked somewhere.
- Ratio of Total cases by total deaths is highest for India.5. Total recovery rate is high than death rate in most of the countries.