

Statistics Tables and Visualizations on Two Spreadsheets

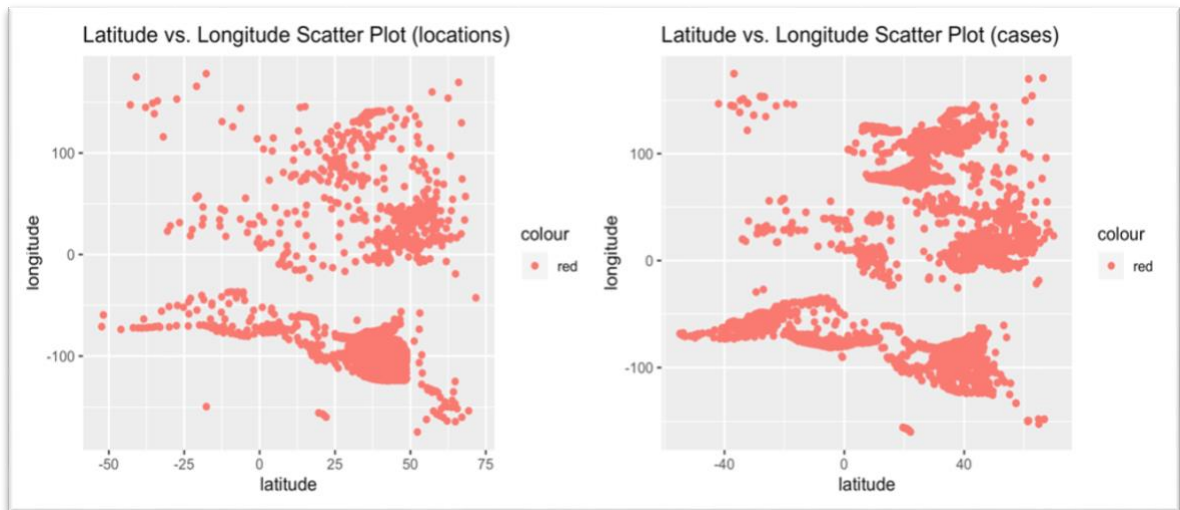
Table 1 – Summary Table of “*processed_individual_cases_Sep20th2020.csv*”

“ <i>processed_individual_cases_Sep20th2020.csv</i> ” (n = 557,364)				
Attributes	Data Type	Number of Missing Values	Number of Unique Values	Visualization
Age	“char”	296,874	362	Word Cloud
Sex	“char”	293,734	2	Pie Chart
Province	“char”	6,568	1,180	Map (plot by latitude and longitude) and Scatter Plot
Country	“char”	24	136	
Latitude	“numeric”	2	8,459	
Longitude	“numeric”	2	8,454	
Date_confirmation	“char”	462	172	Time Series Plot
Additional_information	“char”	522,969	20,689	Word Cloud (on etyma)
Source	“char”	209,191	9,597	
Outcome	“char”	0	4	Pie Chart

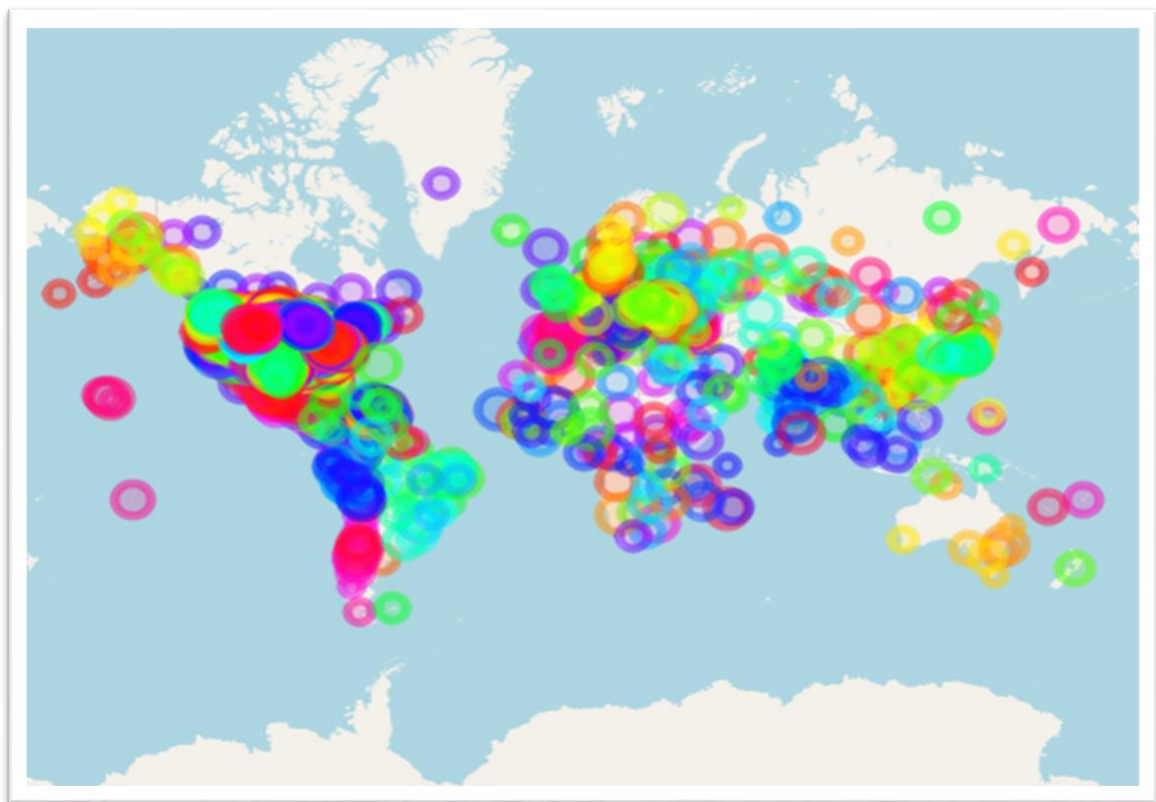
Table 2 – Summary Table of “*processed_location_Sep20th2020.csv*”

“processed_location_Sep20th2020.csv” (n = 3,954)							
Attributes	Data Type	Number of Missing Values	Number of Unique Values	Summary Statistics (For Numerical Data Only)			Visualization
				Min	Median	Max	
Province_State	“char”	168	563				Map
Country_Region	“char”	0	188				
Last_Update	“char”	0	3				
Lat	“numeric”	80	3,874	-52.37	37.94	71.71	Map
Long_	“numeric”	80	3,864	-174.16	-86.88	178.06	
Confirmed	“int”	0	2,116	0	498.5	1,167,496	Histograms
Deaths	“int”	0	583	0	9	37,076	
Recovered	“int”	0	590	0	0	2,577,446	
Active	“numeric”	2	1,939	-2,577,446	421	337,913	
Combined_Key	“char”	0	3,954				
Incidence_Rate	“numeric”	80	3,863	0.0	1,204.4	14,871.2	Histograms
Case.Fatality_Ratio	“numeric”	48	2,625	0.0	1.6334	108.8130	

Graph 1 – Scatter Plot of *latitude* & *longitude* in “location.csv” and “cases.csv”

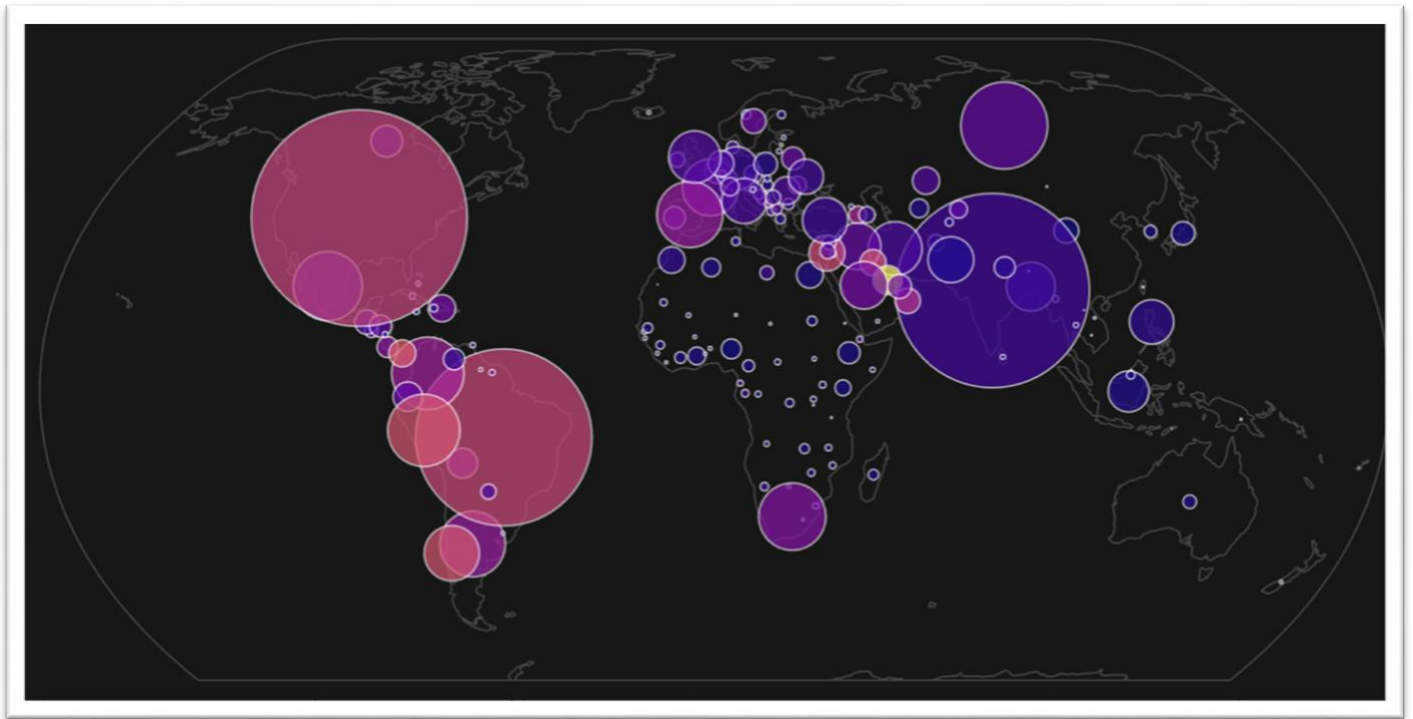


Graph 2 – Locations that have Covid Cases based on “location.csv” – latitude & longitude



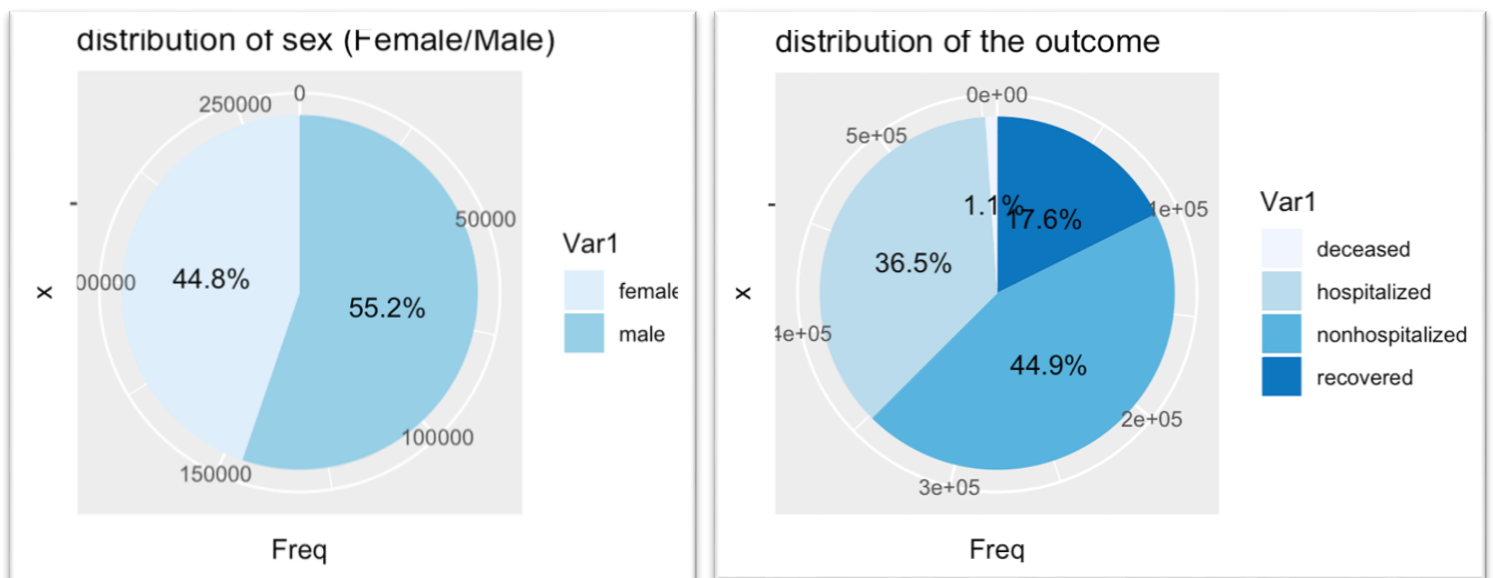
(R Version)

Graph 2.5 –Number of Covid Cases from “location.csv” – latitude & longitude

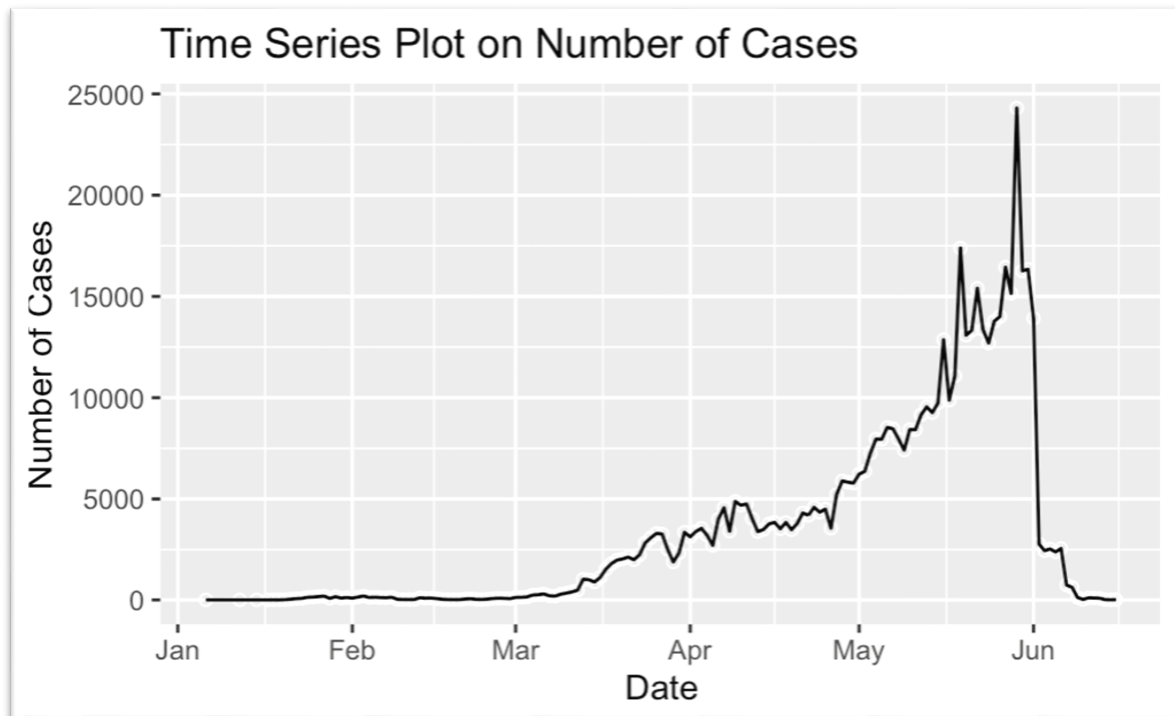


(Python Version)

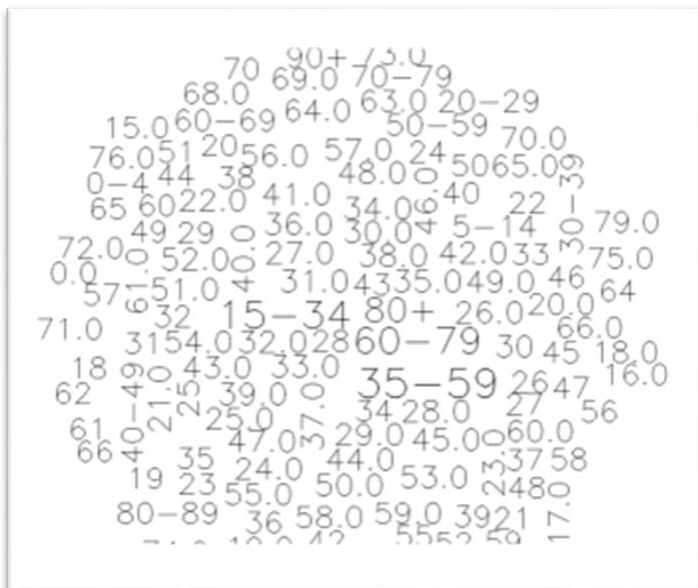
Graph 3 – Pie Charts on the *Sex* and *Outcome* Attributes from “case.csv”



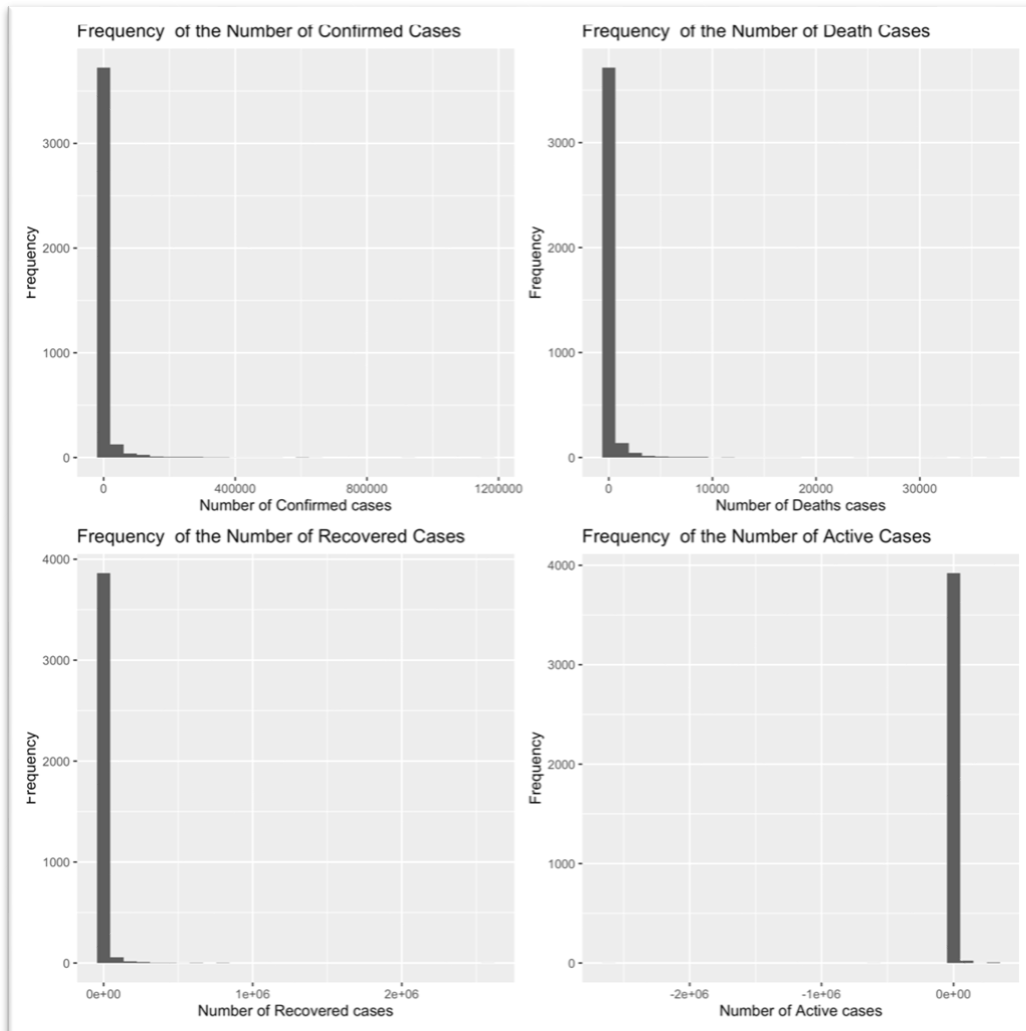
Graph 4 – Time Series Plot on the Aggregated *Date_confirmation* from “case.csv”



Graph 5 – Word Cloud Keywords of *Age* and *Additional_information* from “cases.csv”



Graph 6 – Histograms of *Confirmed, Death, Recovered* and *Active* from “location.csv”



Graph 7 – Histogram of *Incidence* and *Fatality Rates* from “location.csv”

