Namgung_Min_midterm

Min Namgung

3/24/2020

1. Introduction

1.1. Variables: Explain the variables you are presenting in your report.
To start with, I will explain the original datasets. The original datasets have two char datatype variables, and the rest of the variables are all num datatypes. This project runs based on three datasets, which are Covid_Confirmed, Covid_Deaths, and Covid_Recovered. Each original datasets have the same amount of observatinos(482), variables (64), and the same columns. Among the columns, the time series is accumulated by the date.

Original Datasets :

- Covid_Confirmed : this data frame contains the original time_series_19-Covid_Confirmed dataset
- Covid_Deaths : this data frame contains the original time_series_19-Covid_Deaths dataset
- Covid_Recovered : this data frame contains the original time_series_19-Covid_Recovered dataset
    - Province/State : char, this variable contains the state and province name of each country
    - Country/Region : char, this variable contains the country names
    - Lat : num, this variable contains each country's location's latitude
    - Long : num, this variable contains each country's location's longtitude
    - 1/22/20 : num, this variable contains total number of cases (confirmed, deaths, or recovered) by the date and country
    - 1/23/20 : num, this variable contains total number of cases (confirmed, deaths, or recovered) by the date and country
    - …..
    - 3/21/20 : num, this variable contains total number of cases (confirmed, deaths, or recovered) by the date and country

2.1 Trend of cases worldwide : Tidy data

- Covid_Confirmed_worldwide:
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases by each date
- Covid_Deaths_worldwide :
    - Date : All of the date series, from original data, has been moved into a date column
    - Deaths : Total number of worldwide covid deaths cases by each date
- Covid_Recovered_worldwide :
    - Date : All of the date series, from original data, has been moved into a date column
    - Recovered : Total number of worldwide covid recovered cases by each date
- all_worldwide_covid_data :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases by each date, and it is from Covid_Confirmed_worldwide dataset.
    - Deaths : Total number of worldwide covid deaths cases by each date, and it is from Covid_Deaths_worldwide

    - Recovered : Total number of worldwide covid recovered cases by each date, and it is from Covid_Recovered_worldwide

2.1 Trend of cases worldwide : Draw ggplot graph

- all_worldwide_covid_data_for_graph :
    - Date : All of the date series, from original data, has been moved into a date column, formatted to date datatype
    - Case : Verify which cases, among confirmed, deaths or recovered, by date.
    - Case_value : Show the number of worldwide cases by date and case

2.2 Case worldwide by date

- all_worldwide_covid_data_for_table :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases by each date
    - Deaths : Total number of worldwide covid deaths cases by each date
    - Recovered : Total number of worldwide covid recovered cases by each date
    - Active : Total worldwide active cases (confirmed - deaths - recovered) by date
    - Recovery % : Total worldwide recovery % (recovered/confirmed) by date
    - Morailty % : Total worldwide morailty % (deaths/confirmed) by date
- all_worldwide_covid_data_for_table_format
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases by each date and format it by adding “,”
    - Deaths : Total number of worldwide covid deaths cases by each date and format it by adding “,”
    - Recovered : Total number of worldwide covid recovered cases by each date and format it by adding “,”
    - Active : Total worldwide active cases (confirmed - deaths - recovered) by date and format it adding by “,”
    - Recovery % : Total worldwide recovery % (recovered/confirmed) by date
    - Morailty % : Total worldwide morailty % (deaths/confirmed) by date

2.3 cases by country table

- Covid_Confirmed_by_country
    - Country : All countries listed from the original data frame
    - Confirmed : All total covid confirmed cases by country
- Covid_Deaths_by_country
    - Country : All countries listed from the original data frame
    - Deaths : All total covid deaths cases by country
- Covid_Recovered_by_country
    - Country : All countries listed from the original data frame
    - Recovered : All total covid Recovered cases by country
- all_covid_cases_by_country
    - Country : All countries listed from the original data frame
    - Confirmed : Total number of covid confirmed cases by each country
    - Deaths : Total number of covid deaths cases by each country
    - Recovered : Total number of covid recovered cases by each country
- all_covid_cases_by_country_for_table
    - Country : All countries listed from the original data frame
    - Confirmed : Total number of covid confirmed cases by each country in descending order
    - Deaths : Total number of covid deaths cases by each country

    - Recovered : Total number of covid recovered cases by each country
    - Active : Total active cases (confirmed - deaths - recovered) by country
    - Recovery % : Total recovery % (recovered/confirmed) by country
    - Morailty % : Total morailty % (deaths/confirmed) by country
- all_covid_cases_by_country_for_table_format
    - Country : All countries listed from the original data frame
    - Confirmed : Total number of covid confirmed cases by each country in descending order, and format it by adding “,”
    - Deaths : Total number of covid deaths cases by each country, and format it by adding “,”
    - Recovered : Total number of covid recovered cases by each country, and format it by adding “,”
    - Active : Total active cases (confirmed - deaths - recovered) by country, and format it by adding “,”
    - Recovery % : Total recovery % (recovered/confirmed) by country
    - Morailty % : Total morailty % (deaths/confirmed) by country

2.4 Trends of cases excluding China : Table

- Covid_Confirmed_exclude_china :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases excluding China by each date
- Covid_Deaths_exclude_china :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid deaths cases excluding China by each date
- Covid_Recovered_exclude_china :
    - Date : All of the date series, from original data, has been moved into a date column
    - Recovered : Total number of worldwide covid recovered cases excluding China by each date
- all_covid_cases_exclude_China :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : Total number of worldwide covid confirmed cases excluding China by each date
    - Deaths : Total number of worldwide covid deaths cases excluding China by each date
    - Recovered : Total number of worldwide covid recovered cases excluding China by each date

2.4 Trends of cases excluding China : Draw ggplot graph

- all_covid_cases_exclude_China_for_graph :
    - Date : All of the date series, from original data, has been moved into a date column
    - Case : Verify which cases, among confirmed, deaths or recovered, excluding China by date.
    - Case_value : Show the number of worldwide cases excluding China by date and case

2.5 Case worldwide by date excluding China

- all_data_exclude_China_for_table :
    - Country : All countries excluding China from original data
    - Confirmed : Total number of covid confirmed cases excluding China by each country, and format it by adding “,”
    - Deaths : Total number of covid deaths cases excluding China by each country, and format it by adding “,”
    - Recovered : Total number of covid recovered cases excluding China by each country, and format it by adding “,”
    - Active : Total active cases (confirmed - deaths - recovered) excluding China by country, and format it by adding “,”
    - Recovery % : Total recovery % (recovered/confirmed) excluding China by country
    - Morailty % : Total morailty % (deaths/confirmed) excluding China by country

3.1 Trends in Us : Tidy data

- Covid_Confirmed_US_by_date :
  - Date : All of the date series, from original data, has been moved into a date column
  - Confirmed : Total number of US covid confirmed cases by each date
- Covid_Deaths_US_by_date :
  - Date : All of the date series, from original data, has been moved into a date column
  - Deaths : Total number of US covid Deaths cases by each date
- Covid_Recovered_US_by_date :
  - Date : All of the date series, from original data, has been moved into a date column
  - Recovered : Total number of US covid Recovered cases by each date
- all_covid_case_data_US :
  - Date : All of the date series, from original data, has been moved into a date column
  - Confirmed : Total number of US covid confirmed cases by each date
  - Deaths : Total number of US covid Deaths cases by each date
  - Recovered : Total number of US covid Recovered cases by each date

3.1 Trends in Us : Draw ggplot graph

- all_covid_case_US_for_graph :
  - Date : All Date frame from original data is in date column and formatting date datatype
  - Case : Verify which cases, among confirmed, deaths or recovered, in US by date.
  - Case_value : Show the number of US cases by date and case

3.2 Tables of cases in US by date

- all_covid_case_data_US_for_table
  - Date : All Date frame from original data is in date column and formatting date datatype
  - Confirmed : Total number of covid confirmed cases in US by each date, and format it by adding “,”
  - Deaths : Total number of covid deaths cases in US by each date, and format it by adding “,”
  - Recovered : Total number of covid recovered cases in US by each date, and format it by adding “,”
  - Active : Total active cases (confirmed - deaths - recovered) in US by each date, and format it byadding “,”
  - Recovery % : Total recovery % (recovered/confirmed) in US by each date
  - Morailty % : Total morailty % (deaths/confirmed) in US by each date

3.3 Tables of cases in US by state

- Covid_Confirmed_US_by_state :
  - State : All state names within US
  - Confirmed : Total number of covid confirmed cases in each state
- Covid_Deaths_US_by_state :
  - State : All state names within US
  - Deaths : Total number of covid deaths cases in each state
- Covid_Recovered_US_by_state :
  - State : All state names within US
  - Recovered : Total number of covid recovered cases in each state
- all_US_covid_data_by_state :
  - State : All state names within US
  - Confirmed : Total number of covid confirmed cases in each state
  - Deaths : Total number of covid deaths cases in each state
  - Recovered : Total number of covid recovered cases in each state

- all_US_covid_data_by_state_for_table :
    - State : All state names within US
    - Confirmed : Total number of covid confirmed cases in each state, and format it by adding “,”
    - Deaths : Total number of covid deaths cases in each state, and format it by adding “,”
    - Recovered : Total number of covid recovered cases in each state, and format it by adding “,”
    - Active : Total active cases (confirmed - deaths - recovered) in each US date, and format it by adding “,”
    - Recovery % : Total recovery % (recovered/confirmed) in each US date
    - Morailty % : Total morailty % (deaths/confirmed) in each US date

4.1 Covid predictions cases by country

- all_data_by_country_rate :
    - Country : all country names from the original data set
    - active % : active % by total worldwide active cases (active / sum of all worldwide active cases) * 100
    - expected recovery : calculation (active/recovery %) and round to three decimal points
    - expected mortailty : calculation (active/ mortailty %) / 100 and round to three decimal points
- Total_active: this is a variable which contains all sum of active elements from each country
- all_data_by_country_rate_format :
    - Country : all country names from the original data set
    - active % : active % by total worldwide active cases (active / sum of all worldwide active cases) * 100 in descending order, and format it by adding “,”
    - expected recovery : calculation (active/recovery %) and round to three decimal points, and format it by adding “,”
    - expected mortailty : calculation (active/ mortailty %) / 100 and round to three decimal points, and format it by adding “,”

4.2 China and Italy Covid Confirmed cases by date

- Covid_Confirmed_china :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : All covid confirmed cases filtered by China by date
- Covid_Confirmed_Italy :
    - Date : All of the date series, from original data, has been moved into a date column
    - Confirmed : All covid confirmed cases filtered by Italy by date
- China_Italy_Covid_Confirmed :
    - Date : All of the date series, from original data, has been moved into a date column
    - China_Confirmed : All covid confirmed cases filtered by China by date
    - Italy_Confirmed : All covid confirmed cases filtered by Italy by date
- China_Italy_Covid_Confirmed_for_graph:
    - Date : All of the date series, from original data, has been moved into a date column
    - Cases : Verify either china_confirmed or Italy_confirmed
    - Confirmed : All covid confirmed cases filtered by either China or Italy by date

1.1 Import library

This project imports 4 different libraries, tidyverse, readr, dplyr, and ggplot2.

1.2 Import data

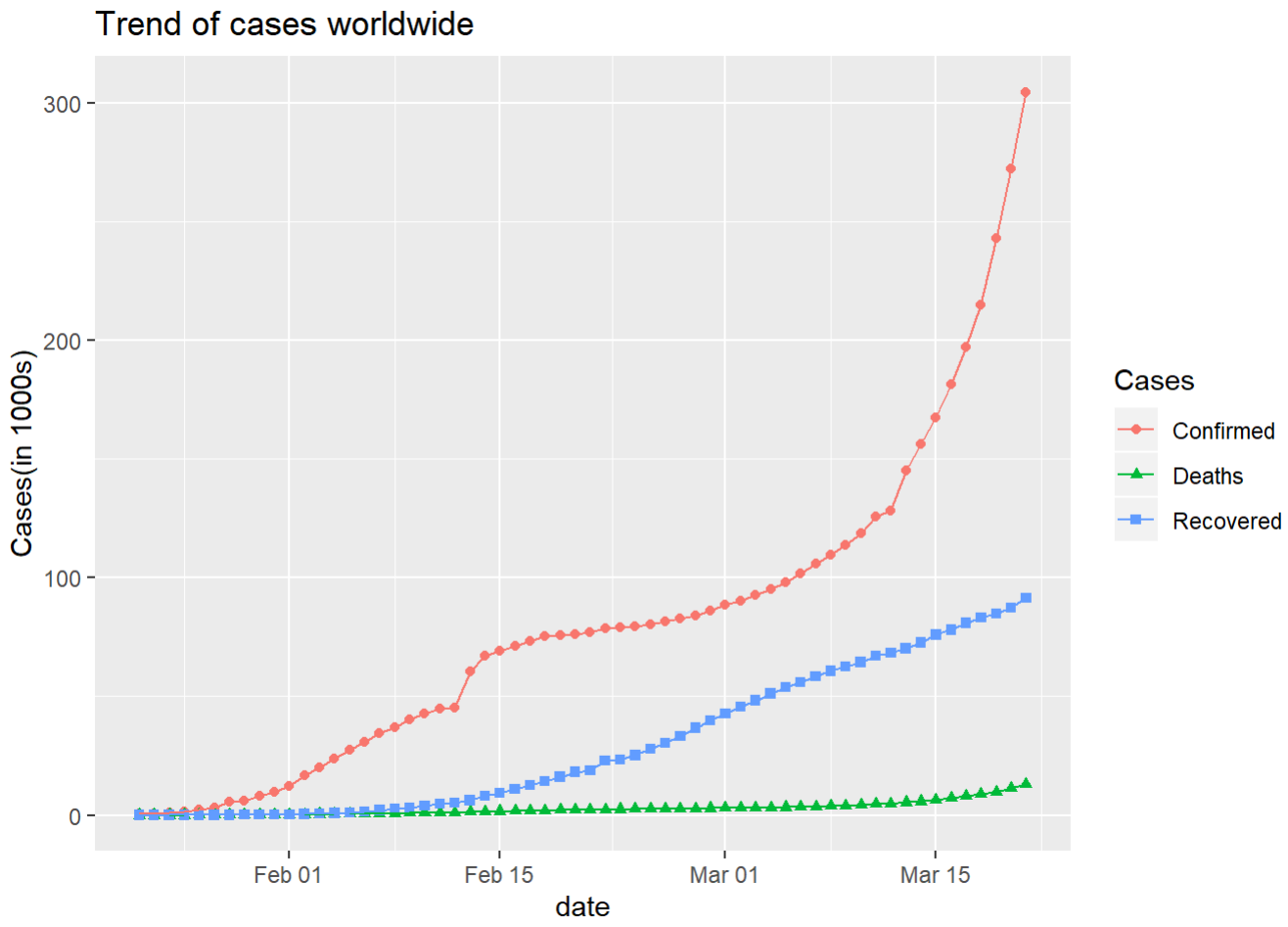
The three datasets are imported by the read_csv() function: covid_confirmed variable have covid-confirmed.csv, Covid_Deaths variable have covid-deaths.csv and covid_recovered variable have covid-recovered.csv dataset.

1.3 Tidy data

To tidy each datasets, I drop the lat and long columns, since it won't be used in the project. Then, rename Country/Region column to Country, and Province/State column to State. This is because it makes it easier to join or filter data from these two columns. Also, I intentionally leave the State column with NA rows, because even though the State column will be used, the NA rows in the State column will not be used in this project. So, the NA rows will not affect the results.

2.1 Trend of cases worldwide

Before I draw the graph, tidying the three datasets is necessary. I tidy each case (each data frame) by each date, so all the date series are now under the date column. Hence, **Covid_Confirmed_worldwide, Covid_Deaths_worldwide, Covid_Recovered_worldwide** are combined into total cases by each date. Then, add these three tidied datasets into one data frame, **all_worldwide_covid_data**. so the **all_worldwide_covid_data** has date, confirmed, deaths and recovered columns.



In **all_worldwide_covid_data_for_graph** data frame, covid cases are under the cases column. In the cases column, it verifies cases among confirmed, deaths, or recovered with the number of worldwide cases. The graph **Trend of cases worldwide** shows that confirmed cases is much higher than deaths and recovered cases.

However, worldwide deaths is the second most popular graph compared to recovered cases. By this graph, I found that the worldwide covid confirmed cases are greatly increasing.

2.2 Cases worldwide by date

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-01-22	555	17	28	510	5.045	3.063
2020-01-23	653	18	30	605	4.594	2.757
2020-01-24	941	26	36	879	3.826	2.763
2020-01-25	1,434	42	39	1,353	2.720	2.929
2020-01-26	2,118	56	52	2,010	2.455	2.644
2020-01-27	2,927	82	61	2,784	2.084	2.802
2020-01-28	5,578	131	107	5,340	1.918	2.349
2020-01-29	6,166	133	126	5,907	2.043	2.157
2020-01-30	8,234	171	143	7,920	1.737	2.077
2020-01-31	9,927	213	222	9,492	2.236	2.146
2020-02-01	12,038	259	284	11,495	2.359	2.152
2020-02-02	16,787	362	472	15,953	2.812	2.156
2020-02-03	19,881	426	623	18,832	3.134	2.143
2020-02-04	23,892	492	852	22,548	3.566	2.059
2020-02-05	27,635	564	1,124	25,947	4.067	2.041
2020-02-06	30,817	634	1,487	28,696	4.825	2.057
2020-02-07	34,391	719	2,011	31,661	5.847	2.091
2020-02-08	37,120	806	2,616	33,698	7.047	2.171
2020-02-09	40,150	906	3,244	36,000	8.080	2.257
2020-02-10	42,762	1,013	3,946	37,803	9.228	2.369
2020-02-11	44,802	1,113	4,683	39,006	10.453	2.484
2020-02-12	45,221	1,118	5,150	38,953	11.389	2.472
2020-02-13	60,368	1,371	6,295	52,702	10.428	2.271
2020-02-14	66,885	1,523	8,058	57,304	12.048	2.277
2020-02-15	69,030	1,666	9,395	57,969	13.610	2.413
2020-02-16	71,224	1,770	10,865	58,589	15.255	2.485
2020-02-17	73,258	1,868	12,583	58,807	17.176	2.550

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-02-18	75,136	2,007	14,352	58,777	19.101	2.671
2020-02-19	75,639	2,122	16,121	57,396	21.313	2.805
2020-02-20	76,197	2,247	18,177	55,773	23.855	2.949
2020-02-21	76,823	2,251	18,890	55,682	24.589	2.930
2020-02-22	78,579	2,458	22,886	53,235	29.125	3.128
2020-02-23	78,965	2,469	23,394	53,102	29.626	3.127
2020-02-24	79,568	2,629	25,227	51,712	31.705	3.304
2020-02-25	80,413	2,708	27,905	49,800	34.702	3.368
2020-02-26	81,395	2,770	30,384	48,241	37.329	3.403
2020-02-27	82,754	2,814	33,277	46,663	40.212	3.400
2020-02-28	84,120	2,872	36,711	44,537	43.641	3.414
2020-02-29	86,011	2,941	39,782	43,288	46.252	3.419
2020-03-01	88,369	2,996	42,716	42,657	48.338	3.390
2020-03-02	90,306	3,085	45,602	41,619	50.497	3.416
2020-03-03	92,840	3,160	48,228	41,452	51.947	3.404
2020-03-04	95,120	3,254	51,170	40,696	53.795	3.421
2020-03-05	97,882	3,348	53,796	40,738	54.960	3.420
2020-03-06	101,784	3,460	55,865	42,459	54.886	3.399
2020-03-07	105,821	3,558	58,358	43,905	55.148	3.362
2020-03-08	109,795	3,802	60,694	45,299	55.279	3.463
2020-03-09	113,561	3,988	62,494	47,079	55.031	3.512
2020-03-10	118,592	4,262	64,404	49,926	54.307	3.594
2020-03-11	125,865	4,615	67,003	54,247	53.234	3.667
2020-03-12	128,343	4,720	68,324	55,299	53.235	3.678
2020-03-13	145,193	5,404	70,251	69,538	48.385	3.722
2020-03-14	156,094	5,819	72,624	77,651	46.526	3.728
2020-03-15	167,446	6,440	76,034	84,972	45.408	3.846
2020-03-16	181,527	7,126	78,088	96,313	43.017	3.926
2020-03-17	197,142	7,905	80,840	108,397	41.006	4.010
2020-03-18	214,910	8,733	83,207	122,970	38.717	4.064

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-03-19	242,708	9,867	84,854	147,987	34.961	4.065
2020-03-20	272,166	11,299	87,256	173,611	32.060	4.152
2020-03-21	304,524	12,973	91,499	200,052	30.047	4.260

The **all_worldwide_covid_data_for_table** shows worldwide covid cases by date, along with covid active, recovery % and morailty %. The table above is shown from **all_worldwide_covid_data_for_table_format** after formatting **all_worldwide_covid_data_for_table**. By this table, I found that Covid confirmed cases are greatly growing and the death rates remain relatively steady.

2.3 Cases by country table

Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
China	81,305	3,259	71,857	6,189	88.380	4.008
Italy	53,578	4,825	6,072	42,681	11.333	9.006
US	25,489	307	0	25,182	0.000	1.204
Spain	25,374	1,375	2,125	21,874	8.375	5.419
Germany	22,213	84	233	21,896	1.049	0.378
Iran	20,610	1,556	7,635	11,419	37.045	7.550
France	14,431	562	12	13,857	0.083	3.894
Korea, South	8,799	102	1,540	7,157	17.502	1.159
Switzerland	6,575	75	15	6,485	0.228	1.141
United Kingdom	5,067	234	67	4,766	1.322	4.618
Netherlands	3,640	137	2	3,501	0.055	3.764
Belgium	2,815	67	263	2,485	9.343	2.380
Austria	2,814	8	9	2,797	0.320	0.284
Norway	2,118	7	1	2,110	0.047	0.331
Sweden	1,763	20	16	1,727	0.908	1.134
Denmark	1,420	13	1	1,406	0.070	0.915
Portugal	1,280	12	5	1,263	0.391	0.938
Canada	1,278	19	10	1,249	0.782	1.487
Malaysia	1,183	4	114	1,065	9.637	0.338
Australia	1,071	7	26	1,038	2.428	0.654
Brazil	1,021	15	2	1,004	0.196	1.469

Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Japan	1,007	35	232	740	23.039	3.476
Czechia	995	0	6	989	0.603	0.000
Israel	883	1	36	846	4.077	0.113
Ireland	785	3	5	777	0.637	0.382
Pakistan	730	3	13	714	1.781	0.411
Cruise Ship	712	8	325	379	45.646	1.124
Luxembourg	670	8	0	662	0.000	1.194
Turkey	670	9	0	661	0.000	1.343
Chile	537	0	6	531	1.117	0.000
Poland	536	5	1	530	0.187	0.933
Greece	530	13	19	498	3.585	2.453
Finland	523	1	10	512	1.912	0.191
Ecuador	506	7	3	496	0.593	1.383
Qatar	481	0	27	454	5.613	0.000
Iceland	473	1	22	450	4.651	0.211
Indonesia	450	38	15	397	3.333	8.444
Singapore	432	2	140	290	32.407	0.463
Thailand	411	1	42	368	10.219	0.243
Saudi Arabia	392	0	16	376	4.082	0.000
Slovenia	383	1	0	382	0.000	0.261
Romania	367	0	52	315	14.169	0.000
India	330	4	23	303	6.970	1.212
Peru	318	5	1	312	0.314	1.572
Philippines	307	19	13	275	4.235	6.189
Estonia	306	0	1	305	0.327	0.000
Russia	306	1	12	293	3.922	0.327
Bahrain	305	1	125	179	40.984	0.328
Egypt	294	10	41	243	13.946	3.401
South Africa	240	0	0	240	0.000	0.000
Iraq	214	17	51	146	23.832	7.944

Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Croatia	206	1	5	200	2.427	0.485
Mexico	203	2	4	197	1.970	0.985
Panama	200	1	0	199	0.000	0.500
Colombia	196	0	1	195	0.510	0.000
Lebanon	187	4	4	179	2.139	2.139
Slovakia	178	1	0	177	0.000	0.562
Kuwait	176	0	27	149	15.341	0.000
Serbia	171	1	1	169	0.585	0.585
Bulgaria	163	3	3	157	1.840	1.840
Armenia	160	0	1	159	0.625	0.000
Argentina	158	4	3	151	1.899	2.532
Taiwan*	153	2	28	123	18.301	1.307
United Arab Emirates	153	2	38	113	24.837	1.307
San Marino	144	20	4	120	2.778	13.889
Algeria	139	15	32	92	23.022	10.791
Latvia	124	0	1	123	0.806	0.000
Costa Rica	117	2	2	113	1.709	1.709
Dominican Republic	112	2	0	110	0.000	1.786
Uruguay	110	0	0	110	0.000	0.000
Hungary	103	4	7	92	6.796	3.883
Morocco	96	3	3	90	3.125	3.125
Vietnam	94	0	17	77	18.085	0.000
Bosnia and Herzegovina	93	1	2	90	2.151	1.075
Andorra	88	0	1	87	1.136	0.000
Jordan	85	0	1	84	1.176	0.000
North Macedonia	85	0	1	84	1.176	0.000
Cyprus	84	0	0	84	0.000	0.000
Brunei	83	0	2	81	2.410	0.000
Lithuania	83	1	1	81	1.205	1.205
Moldova	80	1	1	78	1.250	1.250

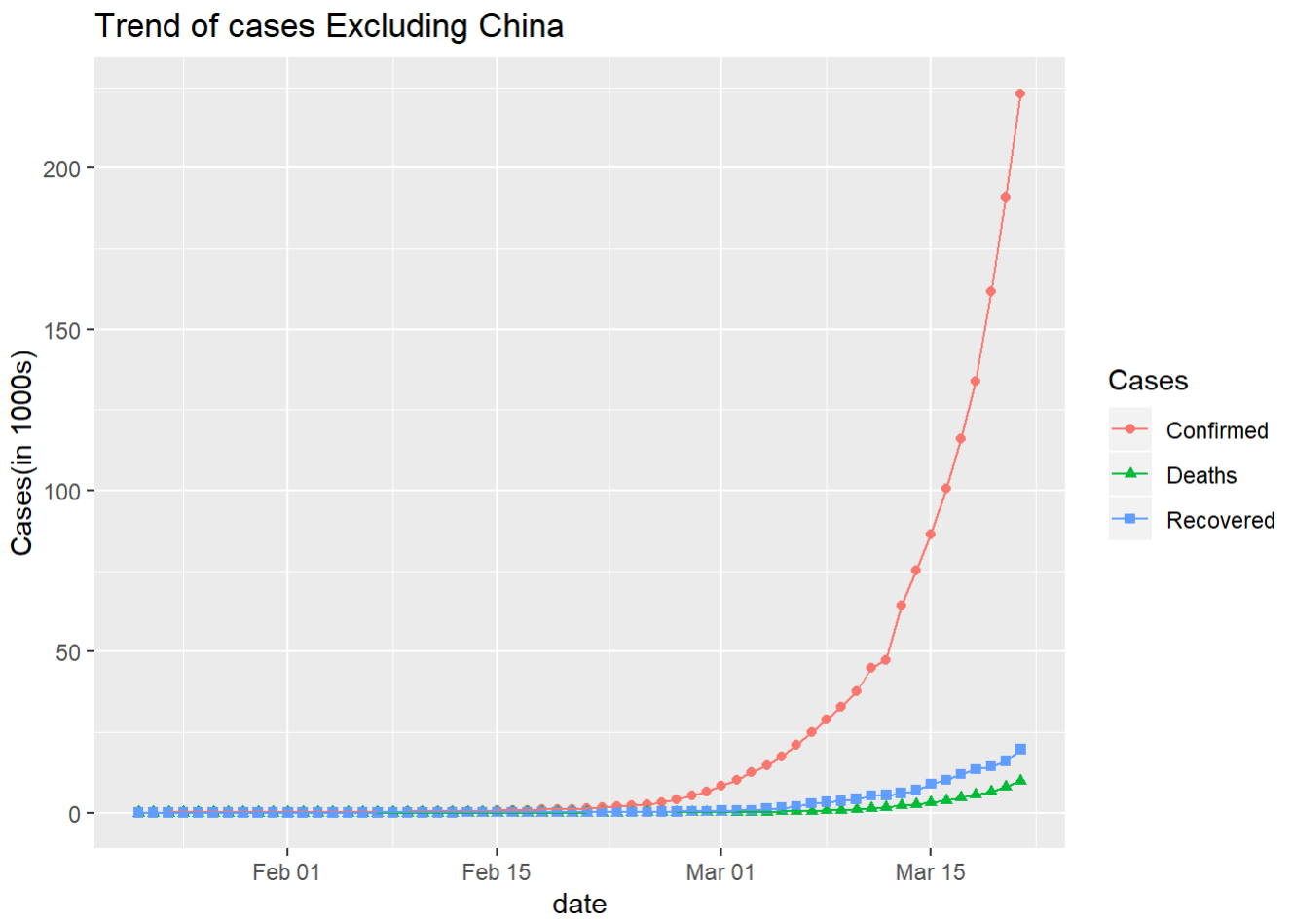
Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Sri Lanka	77	0	1	76	1.299	0.000
Albania	76	2	2	72	2.632	2.632
Belarus	76	0	15	61	19.737	0.000
Malta	73	0	2	71	2.740	0.000
Venezuela	70	0	0	70	0.000	0.000
Burkina Faso	64	2	5	57	7.812	3.125
Tunisia	60	1	0	59	0.000	1.667
Azerbaijan	53	1	11	41	20.755	1.887
Cambodia	53	0	1	52	1.887	0.000
Kazakhstan	53	0	0	53	0.000	0.000
New Zealand	52	0	0	52	0.000	0.000
Oman	52	0	12	40	23.077	0.000
Georgia	49	0	1	48	2.041	0.000
Trinidad and Tobago	49	0	1	48	2.041	0.000
Senegal	47	0	5	42	10.638	0.000
Ukraine	47	3	1	43	2.128	6.383
Uzbekistan	43	0	0	43	0.000	0.000
Liechtenstein	37	0	0	37	0.000	0.000
Martinique	32	1	0	31	0.000	3.125
Cameroon	27	0	0	27	0.000	0.000
Bangladesh	25	2	3	20	12.000	8.000
Afghanistan	24	0	1	23	4.167	0.000
Honduras	24	0	0	24	0.000	0.000
Congo (Kinshasa)	23	1	0	22	0.000	4.348
Nigeria	22	0	1	21	4.545	0.000
Cuba	21	1	0	20	0.000	4.762
Bolivia	19	0	0	19	0.000	0.000
Ghana	19	1	0	18	0.000	5.263
Paraguay	18	1	0	17	0.000	5.556
Guatemala	17	1	0	16	0.000	5.882

Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Rwanda	17	0	0	17	0.000	0.000
Jamaica	16	1	2	13	12.500	6.250
Togo	16	0	1	15	6.250	0.000
Cote d’Ivoire	14	0	1	13	7.143	0.000
Kyrgyzstan	14	0	0	14	0.000	0.000
Mauritius	14	1	0	13	0.000	7.143
Montenegro	14	0	0	14	0.000	0.000
Maldives	13	0	0	13	0.000	0.000
Monaco	11	0	0	11	0.000	0.000
Mongolia	10	0	0	10	0.000	0.000
Ethiopia	9	0	0	9	0.000	0.000
Guyana	7	1	0	6	0.000	14.286
Kenya	7	0	0	7	0.000	0.000
Seychelles	7	0	0	7	0.000	0.000
Barbados	6	0	0	6	0.000	0.000
Equatorial Guinea	6	0	0	6	0.000	0.000
Tanzania	6	0	0	6	0.000	0.000
Bahamas, The	4	0	0	4	0.000	0.000
Gabon	4	1	0	3	0.000	25.000
Suriname	4	0	0	4	0.000	0.000
Cabo Verde	3	0	0	3	0.000	0.000
Central African Republic	3	0	0	3	0.000	0.000
Congo (Brazzaville)	3	0	0	3	0.000	0.000
El Salvador	3	0	0	3	0.000	0.000
Liberia	3	0	0	3	0.000	0.000
Madagascar	3	0	0	3	0.000	0.000
Namibia	3	0	0	3	0.000	0.000
Zimbabwe	3	0	0	3	0.000	0.000
Angola	2	0	0	2	0.000	0.000
Benin	2	0	0	2	0.000	0.000

Country	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Bhutan	2	0	0	2	0.000	0.000
Guinea	2	0	0	2	0.000	0.000
Haiti	2	0	0	2	0.000	0.000
Kosovo	2	0	0	2	0.000	0.000
Mauritania	2	0	0	2	0.000	0.000
Nicaragua	2	0	0	2	0.000	0.000
Saint Lucia	2	0	0	2	0.000	0.000
Sudan	2	1	0	1	0.000	50.000
Zambia	2	0	0	2	0.000	0.000
Antigua and Barbuda	1	0	0	1	0.000	0.000
Cape Verde	1	0	0	1	0.000	0.000
Chad	1	0	0	1	0.000	0.000
Djibouti	1	0	0	1	0.000	0.000
East Timor	1	0	0	1	0.000	0.000
Eritrea	1	0	0	1	0.000	0.000
Eswatini	1	0	0	1	0.000	0.000
Fiji	1	0	0	1	0.000	0.000
Gambia, The	1	0	0	1	0.000	0.000
Holy See	1	0	0	1	0.000	0.000
Nepal	1	0	1	0	100.000	0.000
Niger	1	0	0	1	0.000	0.000
Papua New Guinea	1	0	0	1	0.000	0.000
Saint Vincent and the Grenadines	1	0	0	1	0.000	0.000
Somalia	1	0	0	1	0.000	0.000
Uganda	1	0	0	1	0.000	0.000

The **all_covid_cases_by_country** shows total covid cases by country, along with covid active, recovery % and morailty %. The table above is shown from **all_covid_cases_by_country_desc** after formatting of descending order in confrimed cases and formatting of confirmed, deaths, recovered, and active columns in **all_worldwide_covid_data_for_table**. By this table, I found that China contains the most confirmed cases and the highest recovery rate.

2.4 Trend of Cases Excluding China



To draw the above graph, the data sets are tidied in a table of **all_covid_cases_exclude_China**. Tidying all datasets filtered by excluding China, and adding Confirmed, deaths, and recovered cases in one data frame. In **all_covid_cases_exclude_China** data frame, covid cases are under the cases column excluding China. In the cases column, it verify cases among confirmed, deaths, or recovered with the number of worldwide cases, excluding China, by date. The graph **Trend of cases Excluding China** shows that confirmed cases is much higher than deaths and recovered cases. Deaths and recovered cases do not have very many differences.

2.5 Cases worldwide by date, excluding China

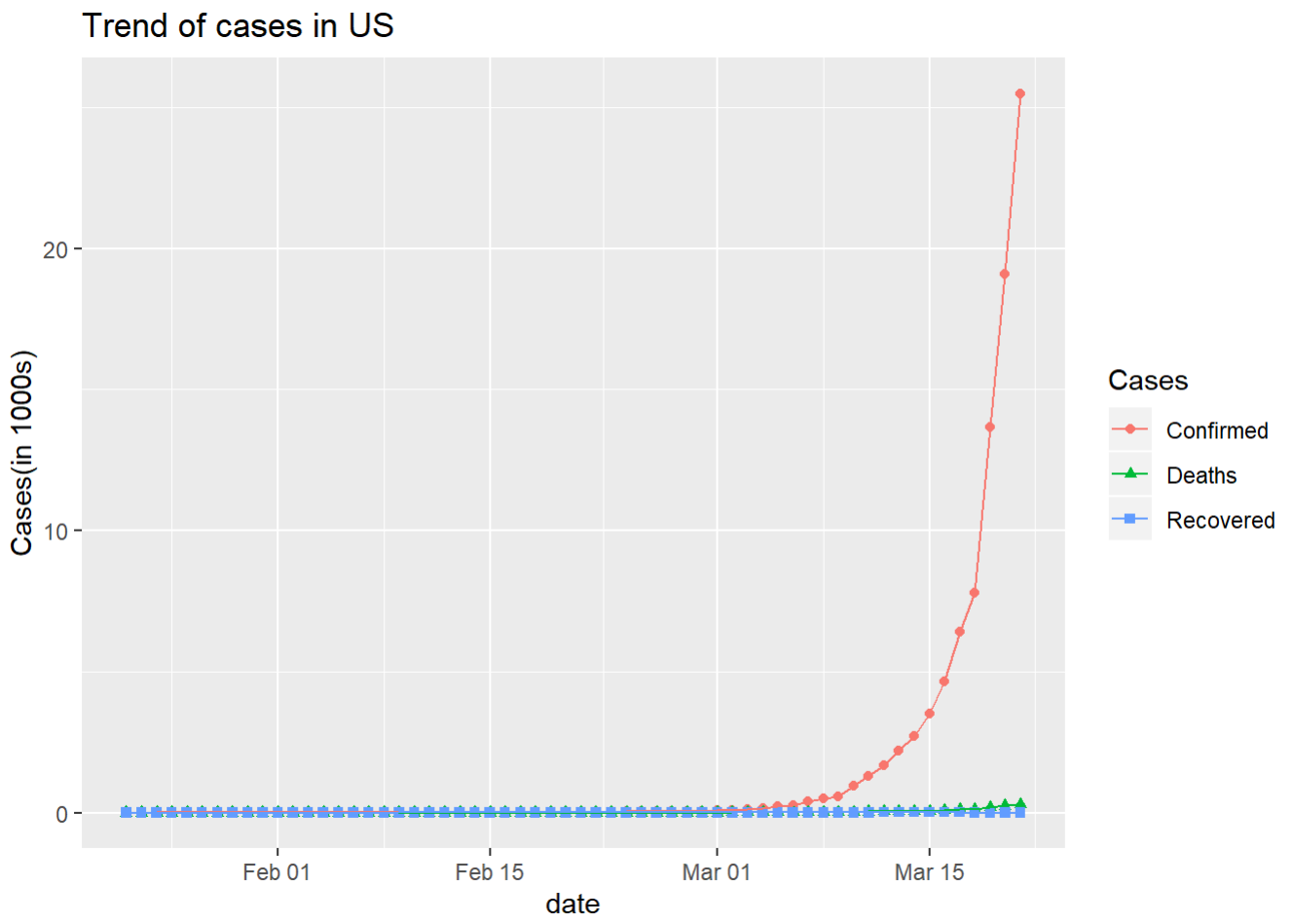
Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-01-22	7	0	0	7	0.000	0.000
2020-01-23	10	0	0	10	0.000	0.000
2020-01-24	21	0	0	21	0.000	0.000
2020-01-25	28	0	0	28	0.000	0.000
2020-01-26	43	0	3	40	6.977	0.000
2020-01-27	50	0	3	47	6.000	0.000
2020-01-28	69	0	6	63	8.696	0.000
2020-01-29	79	0	6	73	7.595	0.000

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-01-30	93	0	8	85	8.602	0.000
2020-01-31	125	0	8	117	6.400	0.000
2020-02-01	147	0	9	138	6.122	0.000
2020-02-02	157	1	9	147	5.732	0.637
2020-02-03	165	1	9	155	5.455	0.606
2020-02-04	185	1	9	175	4.865	0.541
2020-02-05	195	1	9	185	4.615	0.513
2020-02-06	230	1	10	219	4.348	0.435
2020-02-07	281	1	12	268	4.270	0.356
2020-02-08	306	1	20	285	6.536	0.327
2020-02-09	321	1	25	295	7.788	0.312
2020-02-10	408	1	28	379	6.863	0.245
2020-02-11	416	1	47	368	11.298	0.240
2020-02-12	462	1	68	393	14.719	0.216
2020-02-13	473	2	78	393	16.490	0.423
2020-02-14	527	2	81	444	15.370	0.380
2020-02-15	617	3	97	517	15.721	0.486
2020-02-16	711	4	110	597	15.471	0.563
2020-02-17	824	4	121	699	14.684	0.485
2020-02-18	925	4	146	775	15.784	0.432
2020-02-19	1,020	6	159	855	15.588	0.588
2020-02-20	1,120	9	163	948	14.554	0.804
2020-02-21	1,273	13	186	1,074	14.611	1.021
2020-02-22	1,578	15	187	1,376	11.850	0.951
2020-02-23	1,943	24	207	1,712	10.654	1.235
2020-02-24	2,327	34	212	2,081	9.110	1.461
2020-02-25	2,659	43	229	2,387	8.612	1.617
2020-02-26	3,229	53	300	2,876	9.291	1.641
2020-02-27	4,154	68	347	3,739	8.353	1.637
2020-02-28	5,192	82	382	4,728	7.357	1.579

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-02-29	6,655	104	462	6,089	6.942	1.563
2020-03-01	8,437	124	554	7,759	6.566	1.470
2020-03-02	10,170	171	748	9,251	7.355	1.681
2020-03-03	12,579	213	778	11,588	6.185	1.693
2020-03-04	14,734	271	1,169	13,294	7.934	1.839
2020-03-05	17,345	333	1,504	15,508	8.671	1.920
2020-03-06	21,094	416	1,921	18,757	9.107	1.972
2020-03-07	25,051	486	2,819	21,746	11.253	1.940
2020-03-08	28,972	702	3,306	24,964	11.411	2.423
2020-03-09	32,701	865	3,690	28,146	11.284	2.645
2020-03-10	37,705	1,123	4,223	32,359	11.200	2.978
2020-03-11	44,944	1,454	5,359	38,131	11.924	3.235
2020-03-12	47,411	1,548	5,423	40,440	11.438	3.265
2020-03-13	64,248	2,224	6,055	55,969	9.424	3.462
2020-03-14	75,117	2,626	6,964	65,527	9.271	3.496
2020-03-15	86,443	3,237	9,017	74,189	10.431	3.745
2020-03-16	100,494	3,909	10,178	86,407	10.128	3.890
2020-03-17	116,084	4,675	12,042	99,367	10.374	4.027
2020-03-18	133,808	5,492	13,452	114,864	10.053	4.104
2020-03-19	161,552	6,618	14,319	140,615	8.863	4.097
2020-03-20	190,916	8,046	15,990	166,880	8.375	4.214
2020-03-21	223,219	9,714	19,642	193,863	8.799	4.352

The **all_data_exclude_China_for_table** shows the total covid cases by country, along with covid active, recovery % and morailty %. The table above is shown from **all_data_exclude_China_for_table** after formatting confrimed cases and formatting confirmed, deaths, recovered, and active columns in **all_worldwide_covid_data_for_table**. By this table, I found that the initial statement of covid confirmed cases excluding China were not as severe as they are currently. Also, confirmed cases are much higher than recovery case, so I assume covid is difficult to completely recover from.

3.1. Trend in the U.S.



To draw the above graph, the three data sets are combined into one dataset, **all_covid_case_data_US** data frame. In the combined datasets, US covid cases are under the cases column. In the cases column, it verify cases among confirmed, deaths, or recovered, with the number of US covid cases by date. In **all_covid_case_US_for_graph** data frame, covid cases are under the cases column. In the cases column, it verifies cases among confirmed, deaths, or recovered with the number of US cases. The graph **Trend of cases in US** shows that confirmed cases is much higher than deaths and recovered cases. And US deaths and recovered cases contain very little change by date.

3.2 Table of cases in the U.S. by date

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-01-22	1	0	0	1	0.000	0.000
2020-01-23	1	0	0	1	0.000	0.000
2020-01-24	2	0	0	2	0.000	0.000
2020-01-25	2	0	0	2	0.000	0.000
2020-01-26	5	0	0	5	0.000	0.000
2020-01-27	5	0	0	5	0.000	0.000
2020-01-28	5	0	0	5	0.000	0.000
2020-01-29	5	0	0	5	0.000	0.000

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-01-30	5	0	0	5	0.000	0.000
2020-01-31	7	0	0	7	0.000	0.000
2020-02-01	8	0	0	8	0.000	0.000
2020-02-02	8	0	0	8	0.000	0.000
2020-02-03	11	0	0	11	0.000	0.000
2020-02-04	11	0	0	11	0.000	0.000
2020-02-05	11	0	0	11	0.000	0.000
2020-02-06	11	0	0	11	0.000	0.000
2020-02-07	11	0	0	11	0.000	0.000
2020-02-08	11	0	0	11	0.000	0.000
2020-02-09	11	0	3	8	27.273	0.000
2020-02-10	11	0	3	8	27.273	0.000
2020-02-11	12	0	3	9	25.000	0.000
2020-02-12	12	0	3	9	25.000	0.000
2020-02-13	13	0	3	10	23.077	0.000
2020-02-14	13	0	3	10	23.077	0.000
2020-02-15	13	0	3	10	23.077	0.000
2020-02-16	13	0	3	10	23.077	0.000
2020-02-17	13	0	3	10	23.077	0.000
2020-02-18	13	0	3	10	23.077	0.000
2020-02-19	13	0	3	10	23.077	0.000
2020-02-20	13	0	3	10	23.077	0.000
2020-02-21	15	0	5	10	33.333	0.000
2020-02-22	15	0	5	10	33.333	0.000
2020-02-23	15	0	5	10	33.333	0.000
2020-02-24	51	0	5	46	9.804	0.000
2020-02-25	51	0	6	45	11.765	0.000
2020-02-26	57	0	6	51	10.526	0.000
2020-02-27	58	0	6	52	10.345	0.000
2020-02-28	60	0	7	53	11.667	0.000

Date	Confirmed	Deaths	Recovered	active	recovery..	morailty..
2020-02-29	68	1	7	60	10.294	1.471
2020-03-01	74	1	7	66	9.459	1.351
2020-03-02	98	6	7	85	7.143	6.122
2020-03-03	118	7	7	104	5.932	5.932
2020-03-04	149	11	7	131	4.698	7.383
2020-03-05	217	12	7	198	3.226	5.530
2020-03-06	262	14	7	241	2.672	5.344
2020-03-07	402	17	7	378	1.741	4.229
2020-03-08	518	21	7	490	1.351	4.054
2020-03-09	583	22	7	554	1.201	3.774
2020-03-10	959	28	8	923	0.834	2.920
2020-03-11	1,281	36	8	1,237	0.625	2.810
2020-03-12	1,663	40	12	1,611	0.722	2.405
2020-03-13	2,179	47	12	2,120	0.551	2.157
2020-03-14	2,727	54	12	2,661	0.440	1.980
2020-03-15	3,499	63	12	3,424	0.343	1.801
2020-03-16	4,632	85	17	4,530	0.367	1.835
2020-03-17	6,421	108	17	6,296	0.265	1.682
2020-03-18	7,783	118	0	7,665	0.000	1.516
2020-03-19	13,677	200	0	13,477	0.000	1.462
2020-03-20	19,100	244	0	18,856	0.000	1.277
2020-03-21	25,489	307	0	25,182	0.000	1.204

The **all_covid_case_data_US_for_table** shows the total covid cases in US by date, along with covid active, recovery % and morailty %. The table above is shown from **all_covid_case_data_US_for_table** after the formatting of confirmed, deaths, recovered, and active columns in **all_covid_case_data_US_for_table**. By this table, I found that recovery cases, starting March 18 are no longer recorded.

3.3 Table of cases in the U.S. by state

State	Confirmed	Deaths	Recovered	active	recovery..	morailty..
New York	11,710	60	0	11,650	0	0.512
Washington	1,793	94	0	1,699	0	5.243

State	Confirmed	Deaths	Recovered	active	recovery..	morailty..
California	1,364	24	0	1,340	0	1.760
New Jersey	1,327	16	0	1,311	0	1.206
Michigan	788	5	0	783	0	0.635
Illinois	753	6	0	747	0	0.797
Florida	659	13	0	646	0	1.973
Louisiana	585	16	0	569	0	2.735
Texas	581	5	0	576	0	0.861
Massachusetts	525	1	0	524	0	0.190
Georgia	507	14	0	493	0	2.761
Pennsylvania	396	2	0	394	0	0.505
Colorado	390	4	0	386	0	1.026
Tennessee	371	1	0	370	0	0.270
Wisconsin	282	4	0	278	0	1.418
North Carolina	253	0	0	253	0	0.000
Ohio	248	3	0	245	0	1.210
Connecticut	194	4	0	190	0	2.062
Maryland	193	2	0	191	0	1.036
South Carolina	171	3	0	168	0	1.754
Nevada	161	2	0	159	0	1.242
Virginia	156	2	0	154	0	1.282
Mississippi	140	1	0	139	0	0.714
Minnesota	138	1	0	137	0	0.725
Utah	136	0	0	136	0	0.000
Alabama	131	0	0	131	0	0.000
Indiana	128	4	0	124	0	3.125
Arkansas	122	0	0	122	0	0.000
Arizona	118	1	0	117	0	0.847
Oregon	114	5	0	109	0	4.386
Kentucky	87	3	0	84	0	3.448
District of Columbia	77	1	0	76	0	1.299

State	Confirmed	Deaths	Recovered	active	recovery..	morailty..
Missouri	74	3	0	71	0	4.054
Maine	70	0	0	70	0	0.000
Iowa	68	0	0	68	0	0.000
Rhode Island	66	0	0	66	0	0.000
Kansas	57	2	0	55	0	3.509
New Hampshire	55	0	0	55	0	0.000
Oklahoma	53	1	0	52	0	1.887
Diamond Princess	49	0	0	49	0	0.000
Delaware	45	0	0	45	0	0.000
New Mexico	43	0	0	43	0	0.000
Nebraska	38	0	0	38	0	0.000
Hawaii	37	0	0	37	0	0.000
Idaho	36	0	0	36	0	0.000
Vermont	29	2	0	27	0	6.897
North Dakota	28	0	0	28	0	0.000
Grand Princess	23	0	0	23	0	0.000
Wyoming	23	0	0	23	0	0.000
Montana	21	0	0	21	0	0.000
Puerto Rico	21	1	0	20	0	4.762
Alaska	15	0	0	15	0	0.000
Guam	15	0	0	15	0	0.000
South Dakota	14	1	0	13	0	7.143
West Virginia	8	0	0	8	0	0.000
Virgin Islands	3	0	0	3	0	0.000

To tidy data by US states, I drop the rows which has cities with area code, which is the same as the sample 2.3 question. This is because the rows end up to be 0, in the very last date column, which accumulates all the cases by area. The **all_US_covid_data_by_state_for_table** shows the total covid cases in each US state, along with covid active, recovery % and morailty %. The table above is shown from **all_US_covid_data_by_state_for_table** after the formatting of confirmed, deaths, recovered, and active columns in **all_US_covid_data_by_state_for_table**.

4. Perform an analysis of your own

Table title: **Worldwide Covid active %, expected recovery and expected mortailty by descending order in active %**

country	active..	expected.recovery	expected.mortailty
Italy	21.335	37.661	47.392
US	12.588	0.000	209.153
Germany	10.945	208.732	579.259
Spain	10.934	26.118	40.365
France	6.927	1,669.518	35.586
Iran	5.708	3.082	15.125
Korea, South	3.578	4.089	61.752
Switzerland	3.242	284.430	56.836
China	3.094	0.700	15.442
United Kingdom	2.382	36.051	10.320
Netherlands	1.750	636.545	9.301
Austria	1.398	87.406	98.486
Belgium	1.242	2.660	10.441
Norway	1.055	448.936	63.746
Sweden	0.863	19.020	15.229
Denmark	0.703	200.857	15.366
Portugal	0.631	32.302	13.465
Canada	0.624	15.972	8.399
Malaysia	0.532	1.105	31.509
Australia	0.519	4.275	15.872
Brazil	0.502	51.224	6.835
Czechia	0.494	16.401	0.000
Israel	0.423	2.075	74.867
Ireland	0.388	12.198	20.340
Japan	0.370	0.321	2.129
Pakistan	0.357	4.009	17.372
Luxembourg	0.331	0.000	5.544
Turkey	0.330	0.000	4.922
Chile	0.265	4.754	0.000

country	active..	expected.recovery	expected.mortailty
Poland	0.265	28.342	5.681
Finland	0.256	2.678	26.806
Greece	0.249	1.389	2.030
Ecuador	0.248	8.364	3.586
Qatar	0.227	0.809	0.000
Iceland	0.225	0.968	21.327
Indonesia	0.198	1.191	0.470
Slovenia	0.191	0.000	14.636
Cruise Ship	0.189	0.083	3.372
Saudi Arabia	0.188	0.921	0.000
Thailand	0.184	0.360	15.144
Romania	0.157	0.222	0.000
Peru	0.156	9.936	1.985
Estonia	0.152	9.327	0.000
India	0.151	0.435	2.500
Russia	0.146	0.747	8.960
Singapore	0.145	0.089	6.263
Philippines	0.137	0.649	0.444
Egypt	0.121	0.174	0.714
South Africa	0.120	0.000	0.000
Croatia	0.100	0.824	4.124
Panama	0.099	0.000	3.980
Mexico	0.098	1.000	2.000
Colombia	0.097	3.824	0.000
Bahrain	0.089	0.044	5.457
Lebanon	0.089	0.837	0.837
Slovakia	0.088	0.000	3.149
Serbia	0.084	2.889	2.889
Armenia	0.079	2.544	0.000
Bulgaria	0.078	0.853	0.853

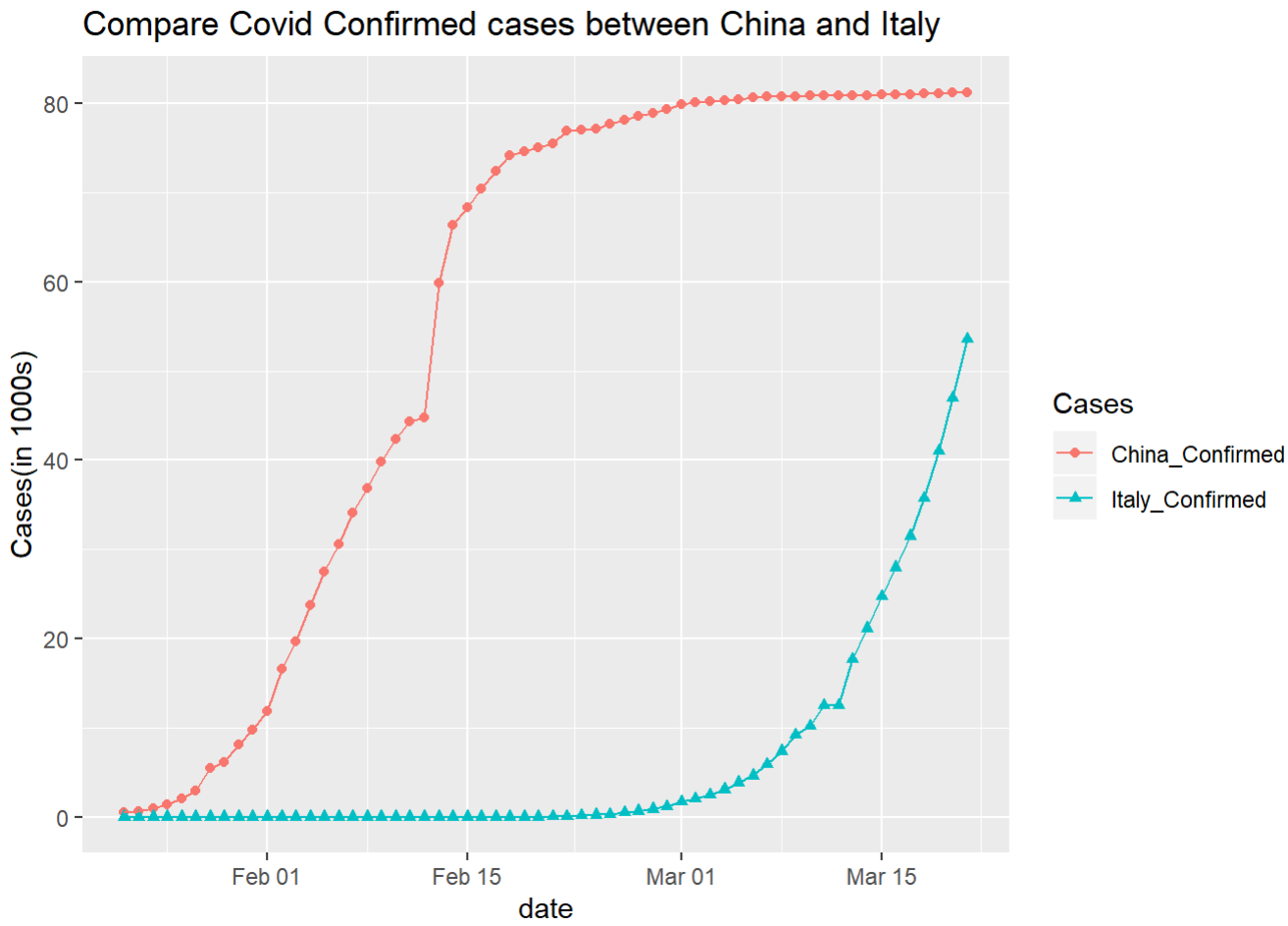
country	active..	expected.recovery	expected.mortailty
Argentina	0.075	0.795	0.596
Kuwait	0.074	0.097	0.000
Iraq	0.073	0.061	0.184
Taiwan*	0.061	0.067	0.941
Latvia	0.061	1.526	0.000
San Marino	0.060	0.432	0.086
United Arab Emirates	0.056	0.045	0.865
Costa Rica	0.056	0.661	0.661
Dominican Republic	0.055	0.000	0.616
Uruguay	0.055	0.000	0.000
Algeria	0.046	0.040	0.085
Hungary	0.046	0.135	0.237
Morocco	0.045	0.288	0.288
Bosnia and Herzegovina	0.045	0.418	0.837
Andorra	0.043	0.766	0.000
Jordan	0.042	0.714	0.000
North Macedonia	0.042	0.714	0.000
Cyprus	0.042	0.000	0.000
Brunei	0.040	0.336	0.000
Lithuania	0.040	0.672	0.672
Moldova	0.039	0.624	0.624
Vietnam	0.038	0.043	0.000
Sri Lanka	0.038	0.585	0.000
Albania	0.036	0.274	0.274
Malta	0.035	0.259	0.000
Venezuela	0.035	0.000	0.000
Belarus	0.030	0.031	0.000
Tunisia	0.029	0.000	0.354
Burkina Faso	0.028	0.073	0.182
Cambodia	0.026	0.276	0.000

country	active..	expected.recovery	expected.mortailty
Kazakhstan	0.026	0.000	0.000
New Zealand	0.026	0.000	0.000
Georgia	0.024	0.235	0.000
Trinidad and Tobago	0.024	0.235	0.000
Senegal	0.021	0.039	0.000
Ukraine	0.021	0.202	0.067
Uzbekistan	0.021	0.000	0.000
Azerbaijan	0.020	0.020	0.217
Oman	0.020	0.017	0.000
Liechtenstein	0.018	0.000	0.000
Martinique	0.015	0.000	0.099
Cameroon	0.013	0.000	0.000
Honduras	0.012	0.000	0.000
Afghanistan	0.011	0.055	0.000
Congo (Kinshasa)	0.011	0.000	0.051
Bangladesh	0.010	0.017	0.025
Nigeria	0.010	0.046	0.000
Cuba	0.010	0.000	0.042
Bolivia	0.009	0.000	0.000
Ghana	0.009	0.000	0.034
Paraguay	0.008	0.000	0.031
Guatemala	0.008	0.000	0.027
Rwanda	0.008	0.000	0.000
Togo	0.007	0.024	0.000
Kyrgyzstan	0.007	0.000	0.000
Montenegro	0.007	0.000	0.000
Jamaica	0.006	0.010	0.021
Cote d’Ivoire	0.006	0.018	0.000
Mauritius	0.006	0.000	0.018
Maldives	0.006	0.000	0.000

country	active..	expected.recovery	expected.mortailty
Monaco	0.005	0.000	0.000
Mongolia	0.005	0.000	0.000
Ethiopia	0.004	0.000	0.000
Guyana	0.003	0.000	0.004
Kenya	0.003	0.000	0.000
Seychelles	0.003	0.000	0.000
Barbados	0.003	0.000	0.000
Equatorial Guinea	0.003	0.000	0.000
Tanzania	0.003	0.000	0.000
Bahamas, The	0.002	0.000	0.000
Suriname	0.002	0.000	0.000
Gabon	0.001	0.000	0.001
Cabo Verde	0.001	0.000	0.000
Central African Republic	0.001	0.000	0.000
Congo (Brazzaville)	0.001	0.000	0.000
El Salvador	0.001	0.000	0.000
Liberia	0.001	0.000	0.000
Madagascar	0.001	0.000	0.000
Namibia	0.001	0.000	0.000
Zimbabwe	0.001	0.000	0.000
Angola	0.001	0.000	0.000
Benin	0.001	0.000	0.000
Bhutan	0.001	0.000	0.000
Guinea	0.001	0.000	0.000
Haiti	0.001	0.000	0.000
Kosovo	0.001	0.000	0.000
Mauritania	0.001	0.000	0.000
Nicaragua	0.001	0.000	0.000
Saint Lucia	0.001	0.000	0.000
Zambia	0.001	0.000	0.000

country	active..	expected.recovery	expected.mortailty
Sudan	0.000	0.000	0.000
Antigua and Barbuda	0.000	0.000	0.000
Cape Verde	0.000	0.000	0.000
Chad	0.000	0.000	0.000
Djibouti	0.000	0.000	0.000
East Timor	0.000	0.000	0.000
Eritrea	0.000	0.000	0.000
Eswatini	0.000	0.000	0.000
Fiji	0.000	0.000	0.000
Gambia, The	0.000	0.000	0.000
Holy See	0.000	0.000	0.000
Nepal	0.000	0.000	NaN
Niger	0.000	0.000	0.000
Papua New Guinea	0.000	0.000	0.000
Saint Vincent and the Grenadines	0.000	0.000	0.000
Somalia	0.000	0.000	0.000
Uganda	0.000	0.000	0.000

The **all_data_by_country_rate** shows the active %, expected recovery and expected mortailty by each country. Active % is calculated by “active / **Total_Active**”, expected recovery amount is calculated by active / recovery %" and expected mortailty amount is calculated by “active / mortailty %”. This table has two prediction models, which are expected recovery and expected mortailty columns. By this table, I found that Italy currently has the most covid active %, and US has the second most active %.



The **Compare Covid Confirmed cases between China and Italy** graph shows the comparison between China's covid confirmed case and Italy's covid confirmed case by date. China and Italy have the most covid confirmed cases compared to other countries and this is the reason I choose these two countries. China's confirmed cases sharply increases around Feburary 15, and Italy's confirmed cases suddenly increases around March 13-15. By this graph, I found that even though Italy's covid confirmed cases appeared later than China, Italy's covid confirmed cases increase quickly. Whereas China's covid confirmed cases seem to be slowing down since March.

5. Conclusion

5.1. What did you learn about COVID-19?

- Covid Recovery rate is pretty low based on the analysis. Also, I can find that the number of confirmed cases per day is not increasing as much as it did during the dates, Feburary 1st through Feburary 15th, in China. Other counties excluding China seem to have lots of confirmed cases one month later than China.

5.2. Do your findings match what you hear in the media?

- Yes. This is because the covid 19 first appeared in China before Feburary, which did not greatly affect other countries. A month or month and a half later, lots of covid confirmed cases appear in other countries. As I analyzed the data in this project, other country's covid infection increased around a month later than they appear in China.

5.3. If you were to explain and advise university officials on the spread of this virus, how would you describe it?

- Worldwide, the number of covid confirmed cases greatly increased after around two months. The death rate begin to increas increased around half a month after the steady increases of confirmed cases. Comparing

the data collected for the death rate with the confirmed cases rate, after two months the confirmed cases exponentially grow, while the death rate continue to steadily increasse. Since the spread of virus is quick, people should stay at home.