



COVID-19 Mortality Prediction

- Group Wonder Penta (WashU)
- Ziwei Shen
- Shiyan Yang
- Stacey Shi
- Vivian Zhong

12/16/2022



Agenda

1	Background + Outcome
2	Data Description
3	Model Construction
4	Recommendation
5	Q&A
6	Appendix

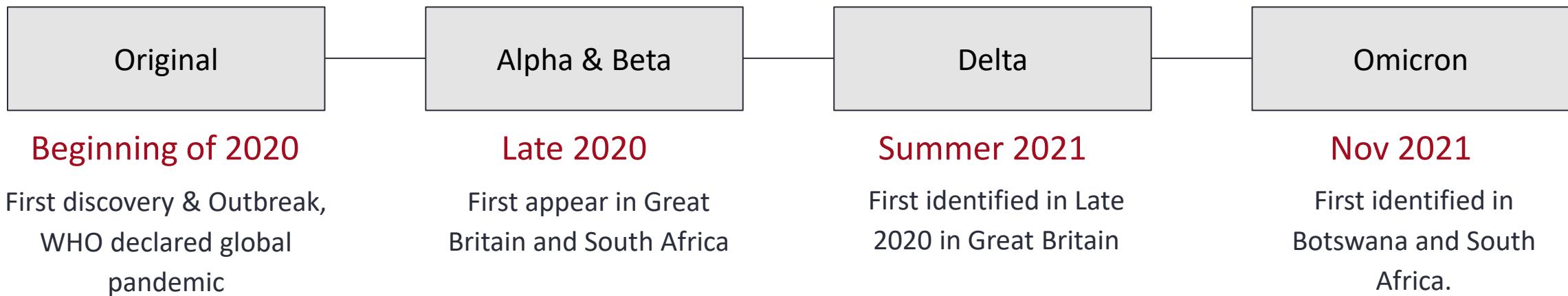
Background +
Outcome

01

Background + Outcome

Problem Background: COVID-19

- SARS-CoV-2
- Respiratory droplets, spread when breaths, talk, or sneeze
- Several variants



Background + Outcome

Problem Background: Company

- Global life and health reinsurance company
- Business:
 - Reinsurance
 - Living benefits reinsurance
 - Group reinsurance
 - Health reinsurance
 - Financial solutions
 - Facultative underwriting
- Location
 - US and Latin America
 - Canada
 - Asia
 - Europe, Middle East and Africa

Background + Outcome

Problem Background: Problem Definition

- ▶ How to predict mortality rate due to COVID-19
 - Approach death by bass model and SIRD
 - Estimate the relationship between death and factors such as vaccine, gender, race, seasonal change, etc
- ▶ \$11 million per 10,000 general population deaths

Background + Outcome

Outcome: Model Comparison



Bass Models

For U.S. & For States
Use Death Only
Daily/Weekly/Monthly



SIRD Model

For U.S. population
Include cases



GBM

For U.S. population
Include Vaccine
Daily/Weekly/Monthly

Outcome – Validation Result

Outcome: Model Comparison

Predictions	Next Day Forecast 11/30	7-Days Forecast 11/21-11/27	Next Month Forecast November 2022
U.S. Bass Model	356	2728 (11/23-11/29)	11057
State Bass Model	184	Daily Model: 1312 Weekly Model: 1798	Daily: 5755 Weekly: 7525 Monthly: 8926
SIRD	384	2572	12191
Actual	577	2074	9985

Outcome – Future Prediction

Outcome: Model Prediction

	Next Month Forecast December 2022	Next Year Forecast 2023
U.S. Bass Model	Daily: 10373 Weekly: 9955 Monthly: 9799	Daily: 59074 Weekly: 64740 Monthly: 56526
State Bass Model	Daily: 5484 Weekly: 6500 Monthly: 7877	Daily: 44377 Weekly: 40318 Monthly: 44527
SIRD	11189	63827
Range	5484 ~ 11189	40318~64740

Data Description | 02

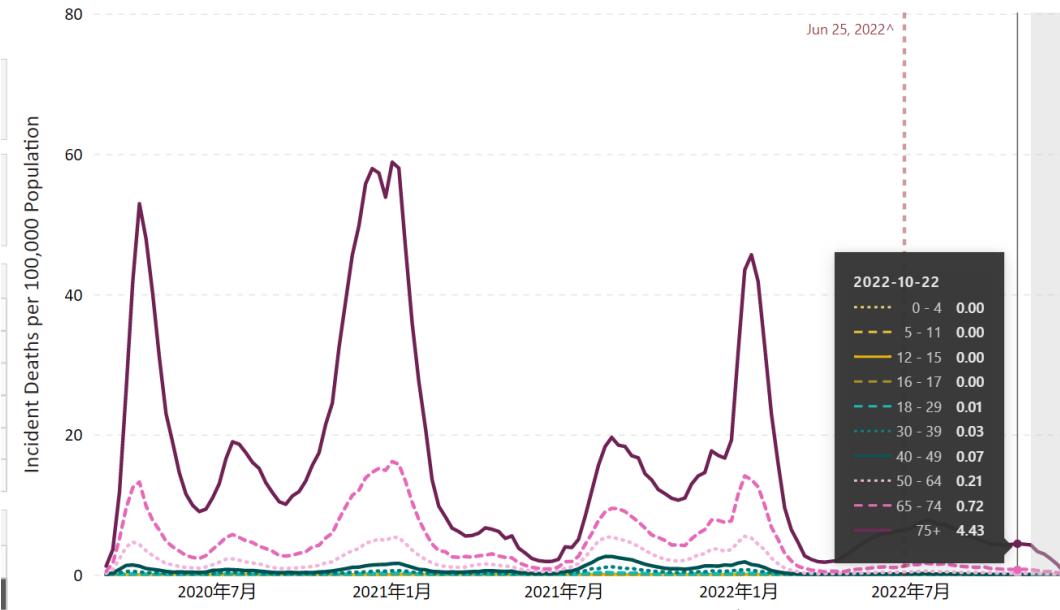
Data Description—Development timeline



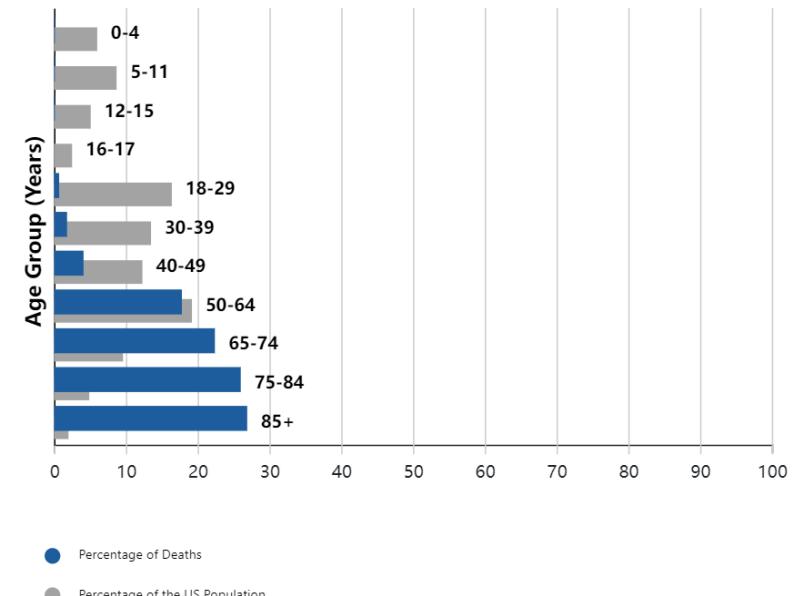
- Happened around Jan 2020
- Twice peak time—Jan 2021 and Jan 2022
- Number of daily deaths decreased to below 1k
- Reached the best situation ever

Data Description—Death by AGE

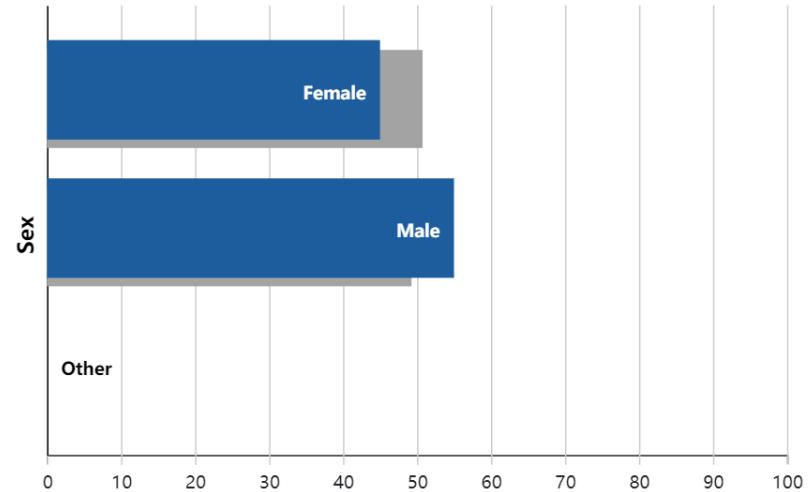
Same development pattern for all ages



The older people are, the higher the death rate

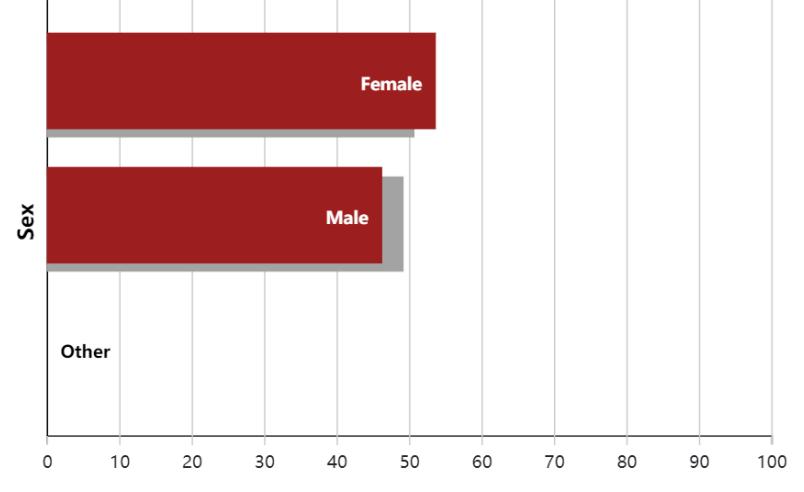


Data Description—Death by GENDER



Deaths by Sex

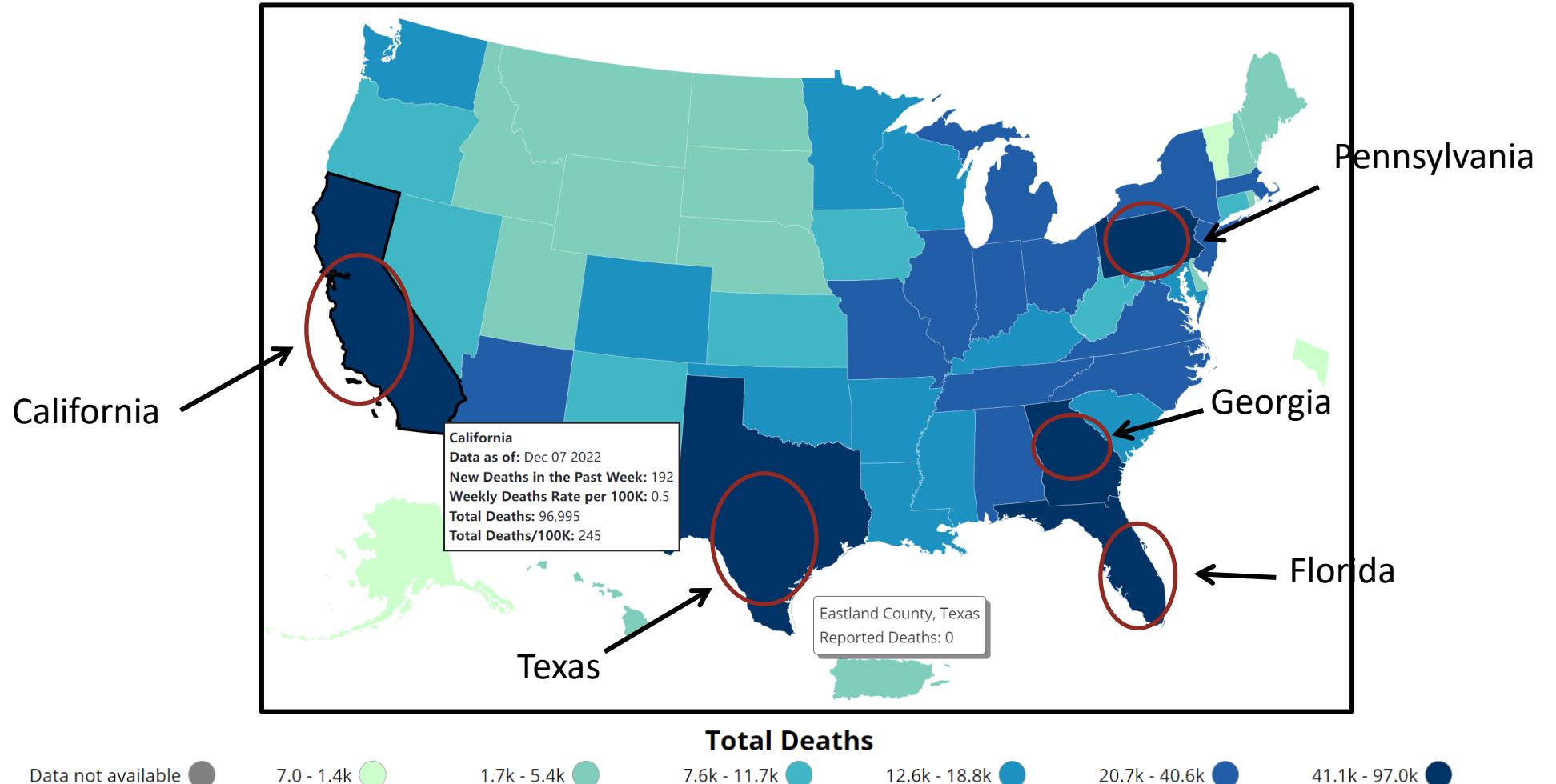
- Percentage of Deaths, All Age Groups
- Percentage of the US Population, All Age Groups



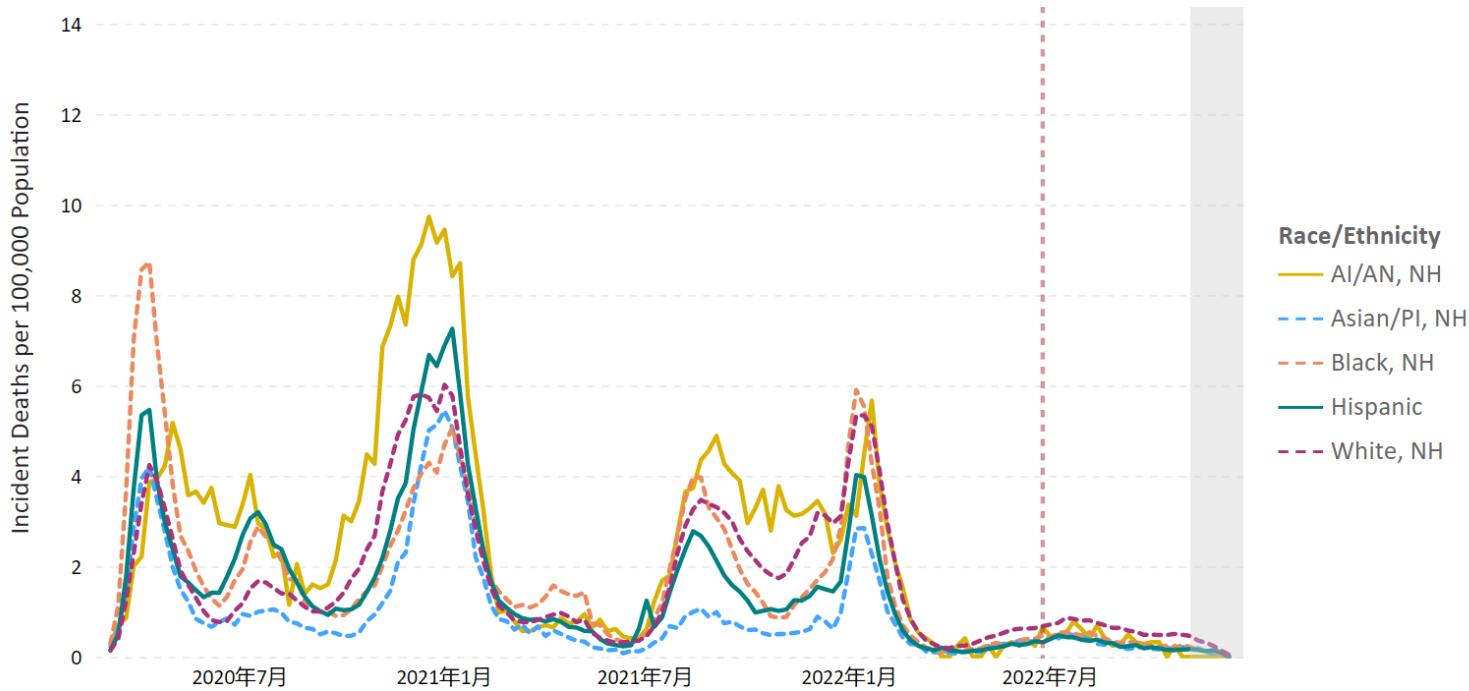
Cases by Sex

- Percentage of Cases, All Age Groups
- Percentage of the US Population, All Age Groups

Data Description—Death by REGION



Data Description—Death by RACE



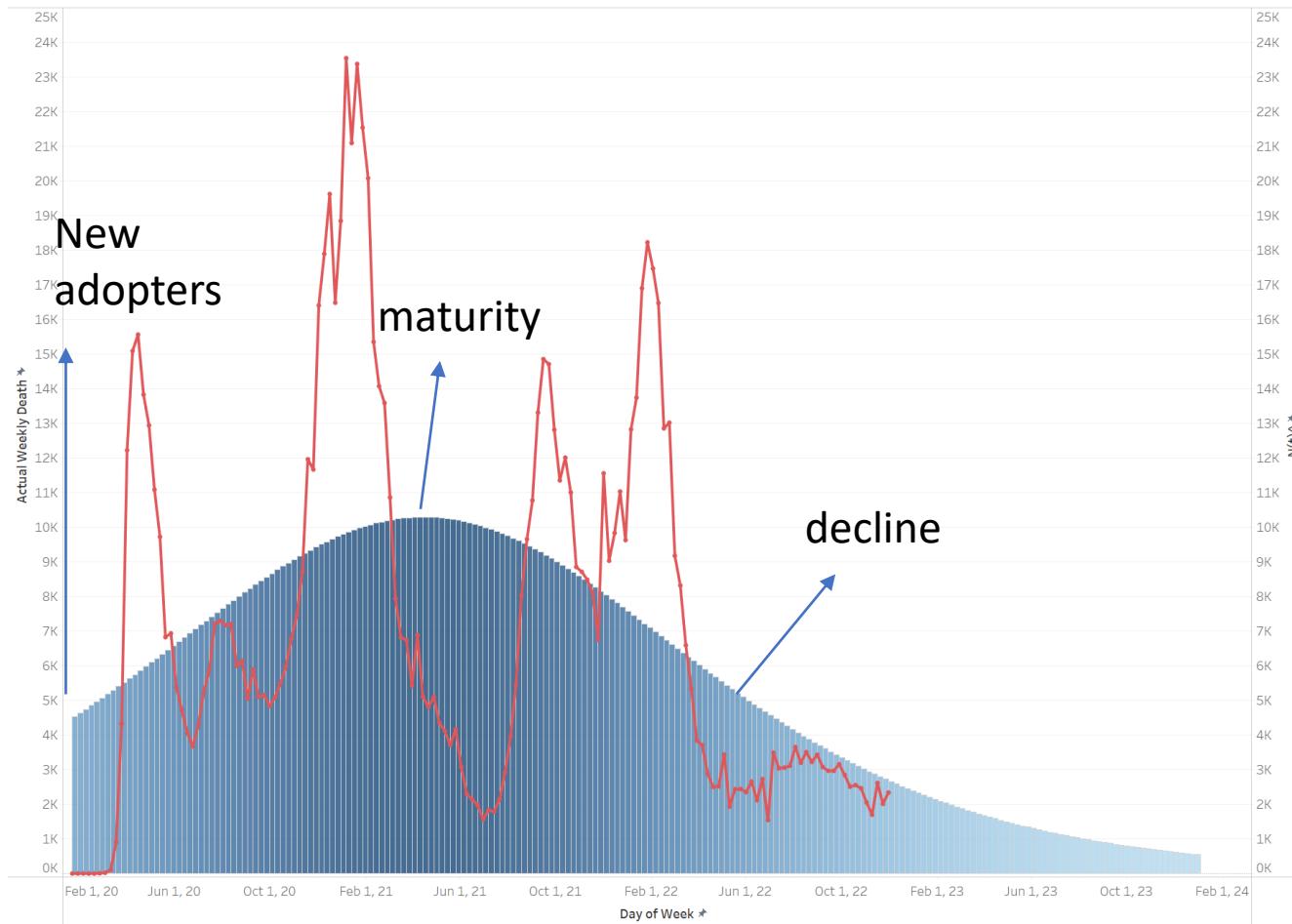
- No clear difference
- There was one time that American Indians and Alaska Natives had terrible number of death
- Asian race have relative better situation
- Death of black race from COVID-19 was most in the beginning.

03

Models Construction

Bass Model Overview

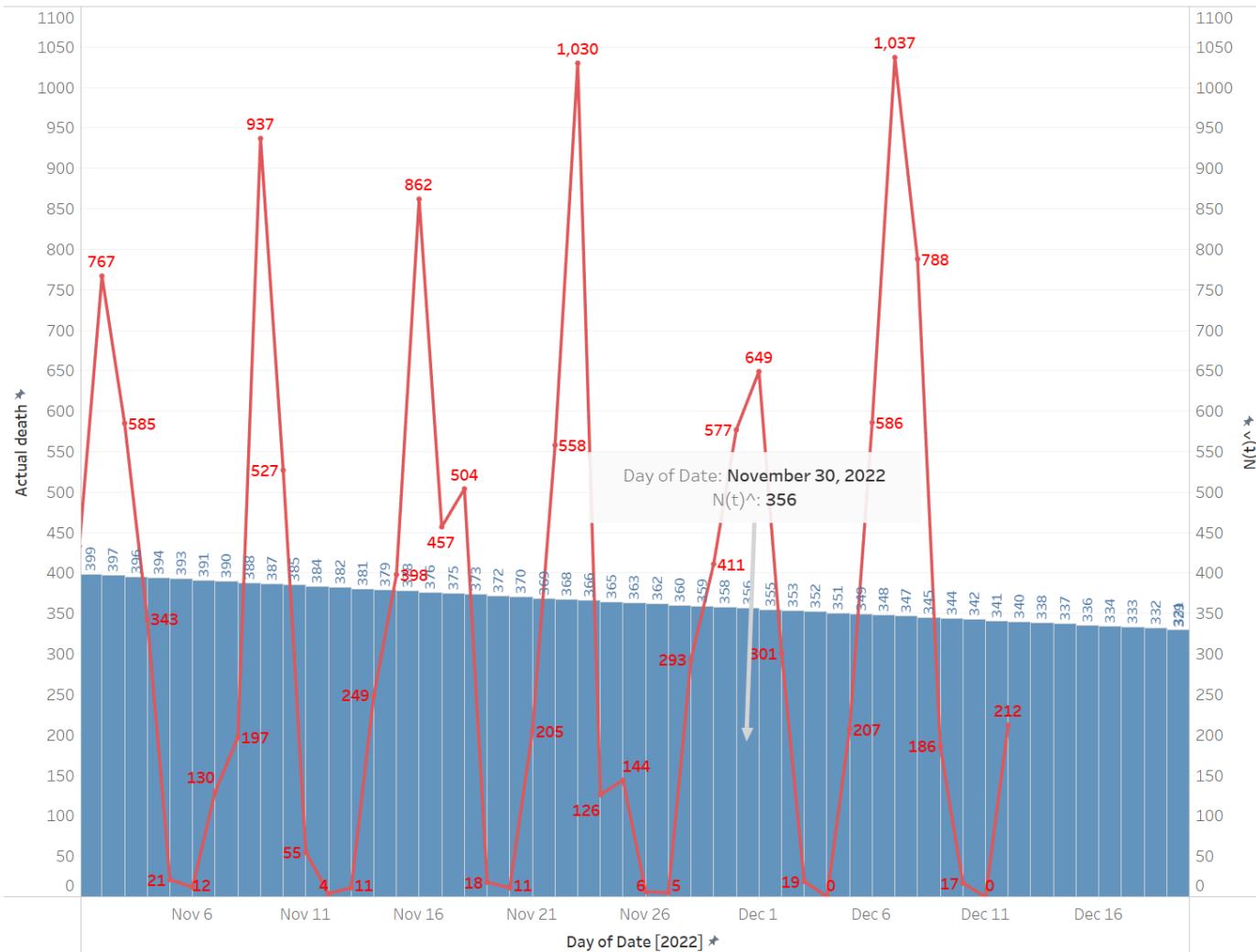
Like a new product adoption model in marketing



- The Bass Model approach and fit the model for COVID in a **bell shape**.
- Approximate the p , q , m with the smallest SSE between prediction and actual death in Solver:
 - p : coefficient of intrinsic susceptibility
 - q : coefficient of infectable susceptibility
 - m : susceptible population, the total number of people who will die from COVID
- Like a new product adoption with **new adopters, maturity and decline stage**
- Bass model believes we are in a **decline stage of death**

Bass Model - Daily

Predict death daily from Bass Model is not very accurate.



Red Line: Actual Death

The reporting of daily death is fluctuating. Daily prediction is not very accurate. Hospitals may not update records every day.

Blue Bar: Predicted Death on Nov. 30, 2022

356

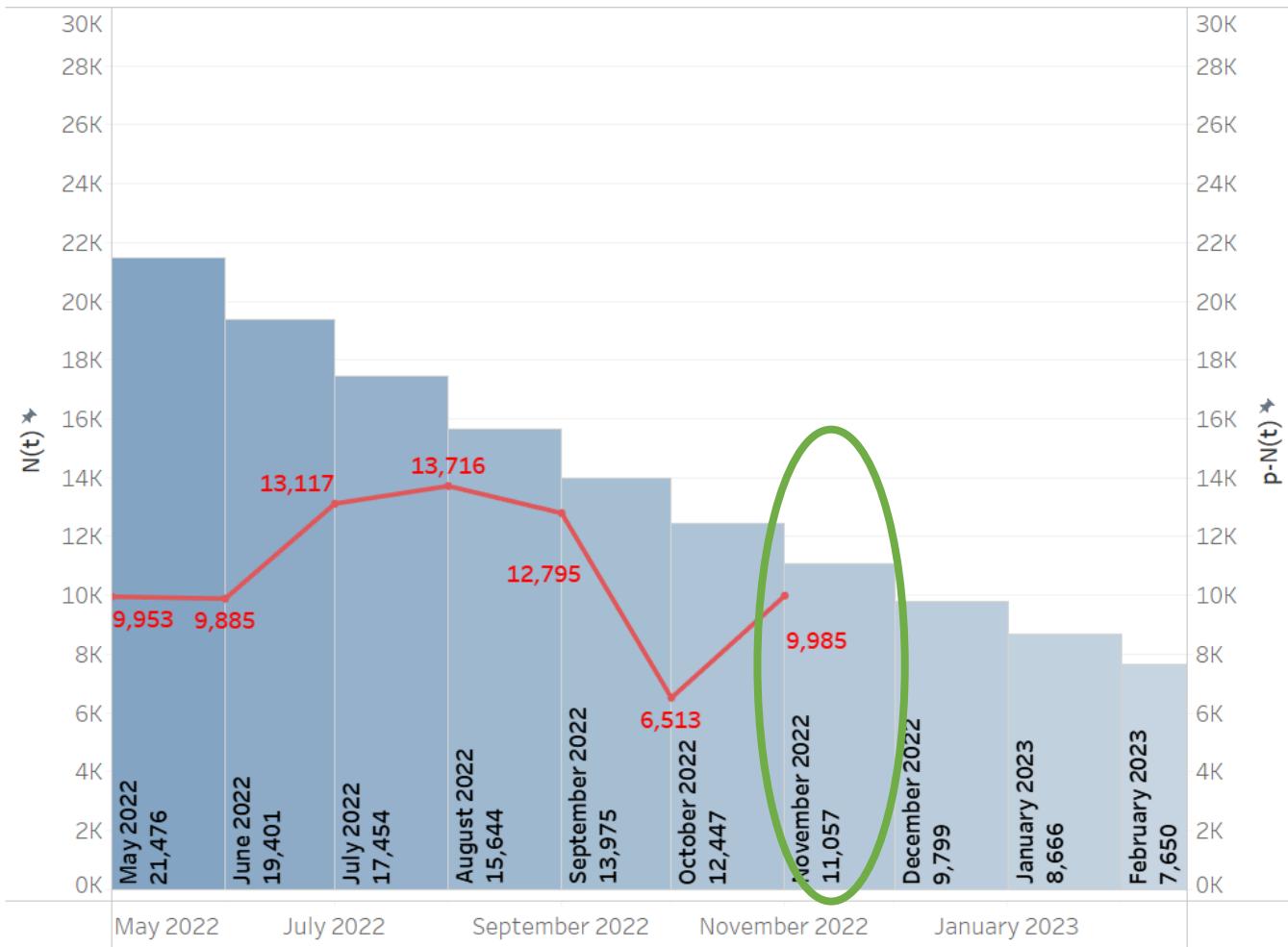
Average Difference between test and validation

$$577 - 356 = 221$$

The prediction is more of average of recent days instead of approaching the actual number of that day.

Bass Model - Monthly

The difference between actual and prediction is smaller than daily model



Actual Death in November 2022

9985

Predicted Death in November 2022

11,057

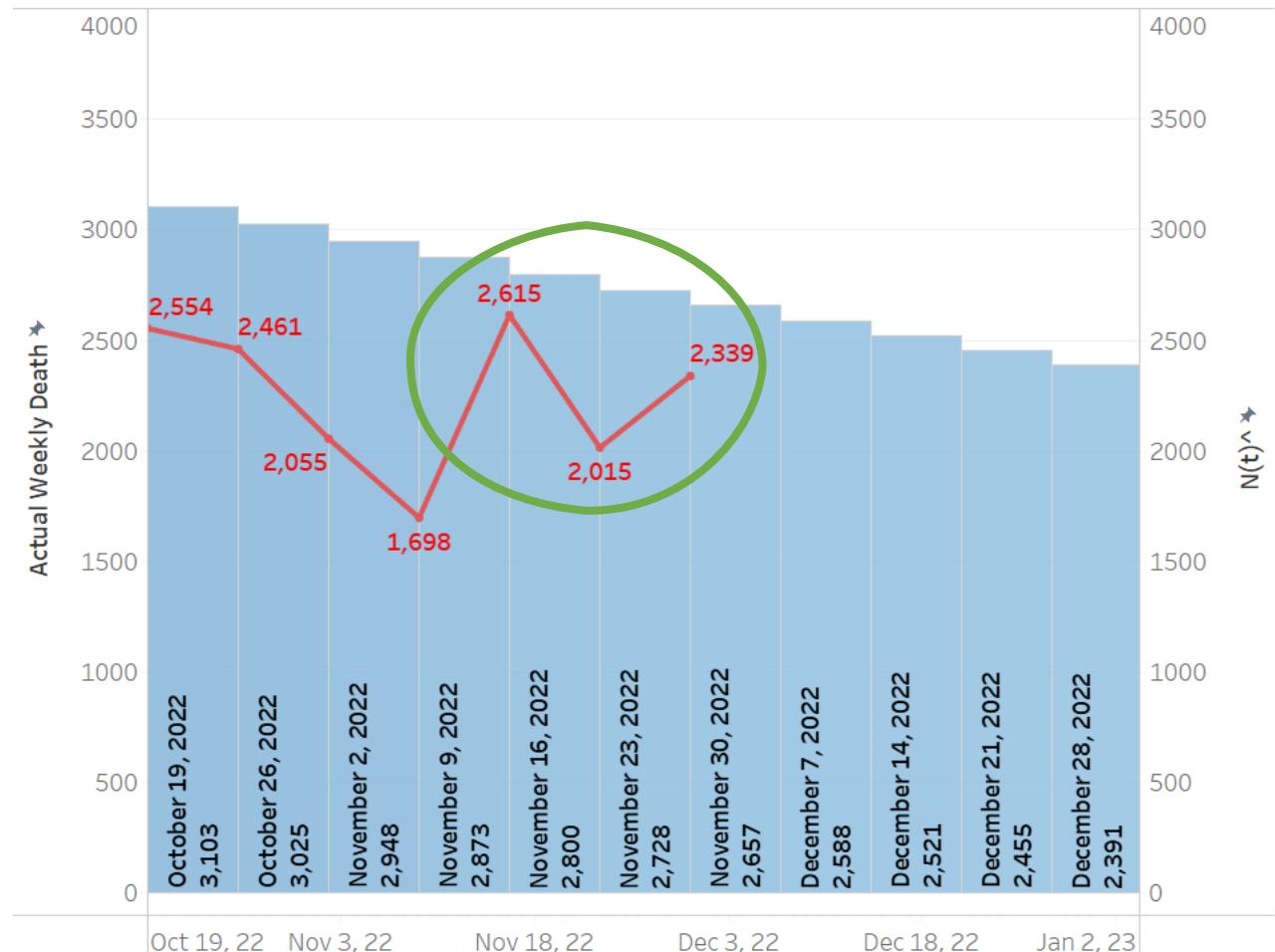
Difference between test and validation

1072

Prediction for next month is the most accurate

Bass Model - Weekly

The most accurate among US bass models



Red Line: Actual Death

2055; 2615; 2015; 2339

Blue Bar: Predicted Death in those three weeks

2948; 2800; 2728; 2657

Difference between test and validation

645

Having Weekly level in bass model approaching total death in the U.S. is the most accurate

Bass Model Prediction

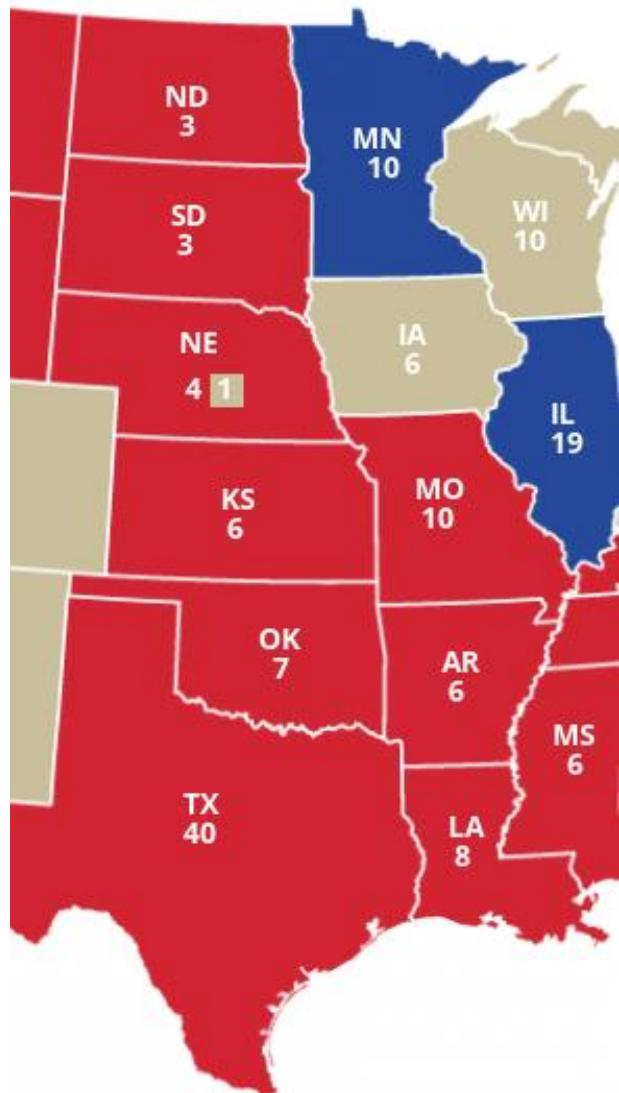
All bass models predict that in 2025 COVID-19 death would fade out



~60,000 deaths -> \$60m-70m extra claims



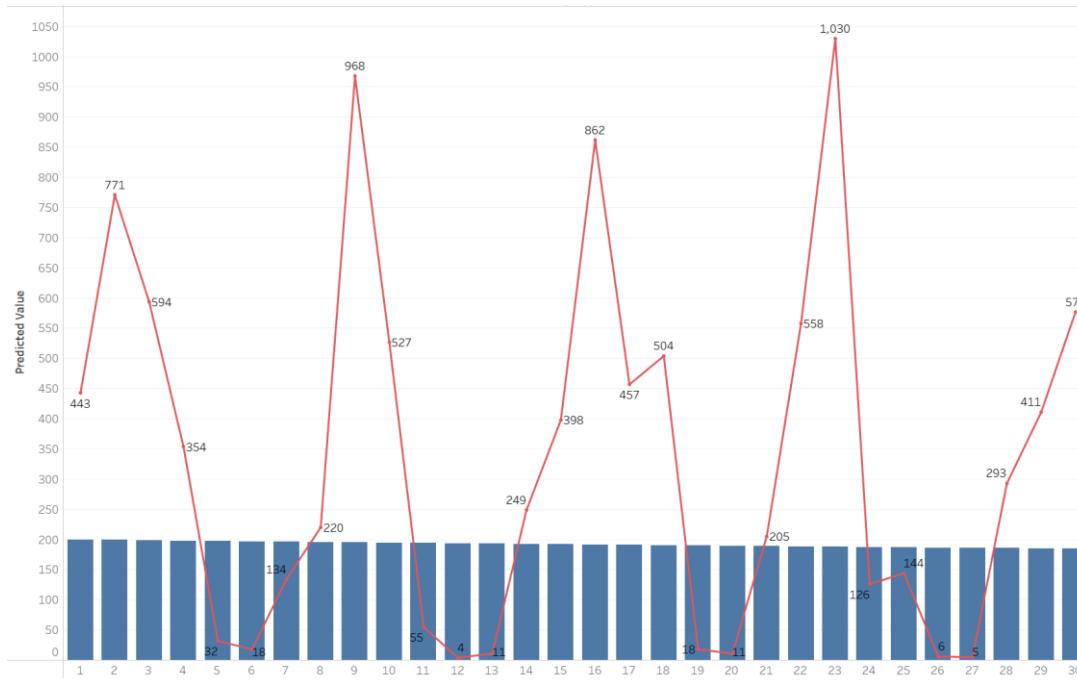
	12/17/2022	12/17/2022-12/23/2022	December 2022	2023	2024	2025
Daily Bass Model Prediction	333	2305	10373	59074	12129	2374
Weekly Bass Model Prediction		2455	9955	64740	14137	2962
Monthly Bass Model Prediction			9799	56526	11402	2202



Predict Death in State Degree

- Run Bass model in every state
- Make predictions through **daily, weekly, monthly** levels
- Utilize python to find the optimal solution of parameters(p, q, m) for each state
- **Sum up** all predicted data from **all states**
- **Compare** the predicted outcomes and the real value

Bass Model on State degree-Daily



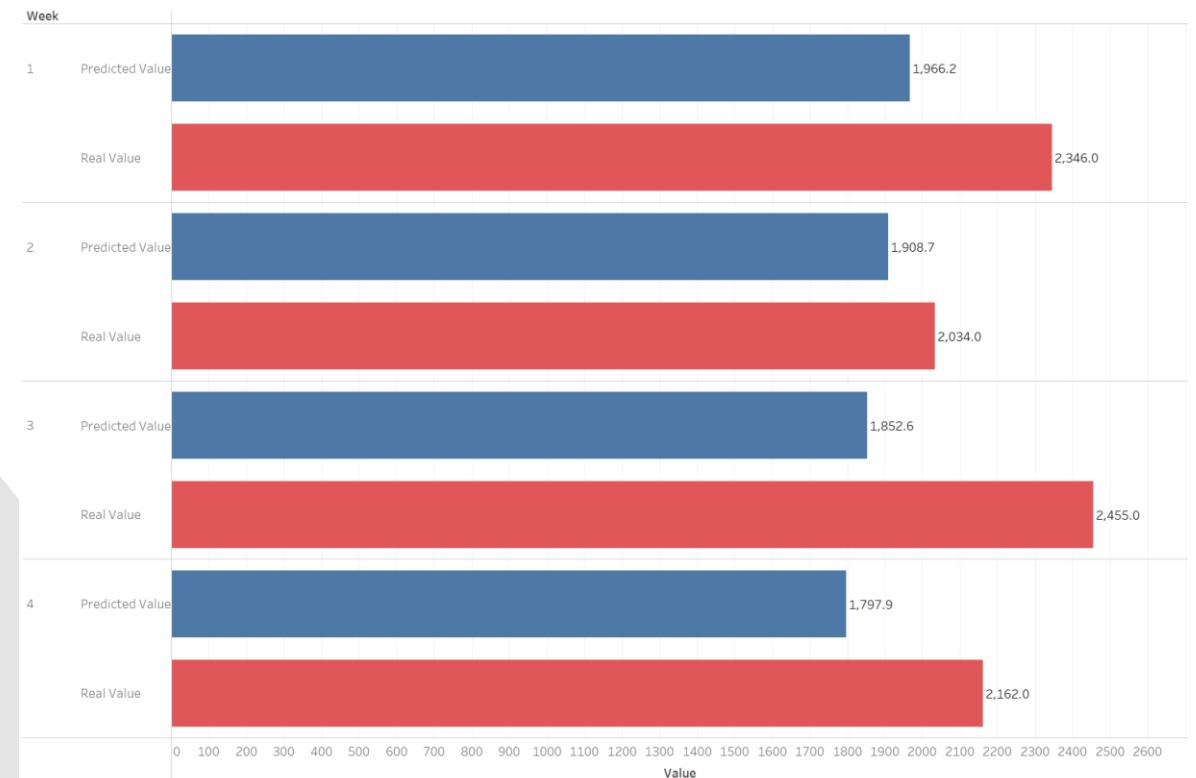
- Difference between predicted value and real value is big
- The predicted values have a smooth line without big fluctuations
- Perform poorly compared to the U.S. degree

- In the validation month of November, differences between predicted and actual deaths are big (>50)

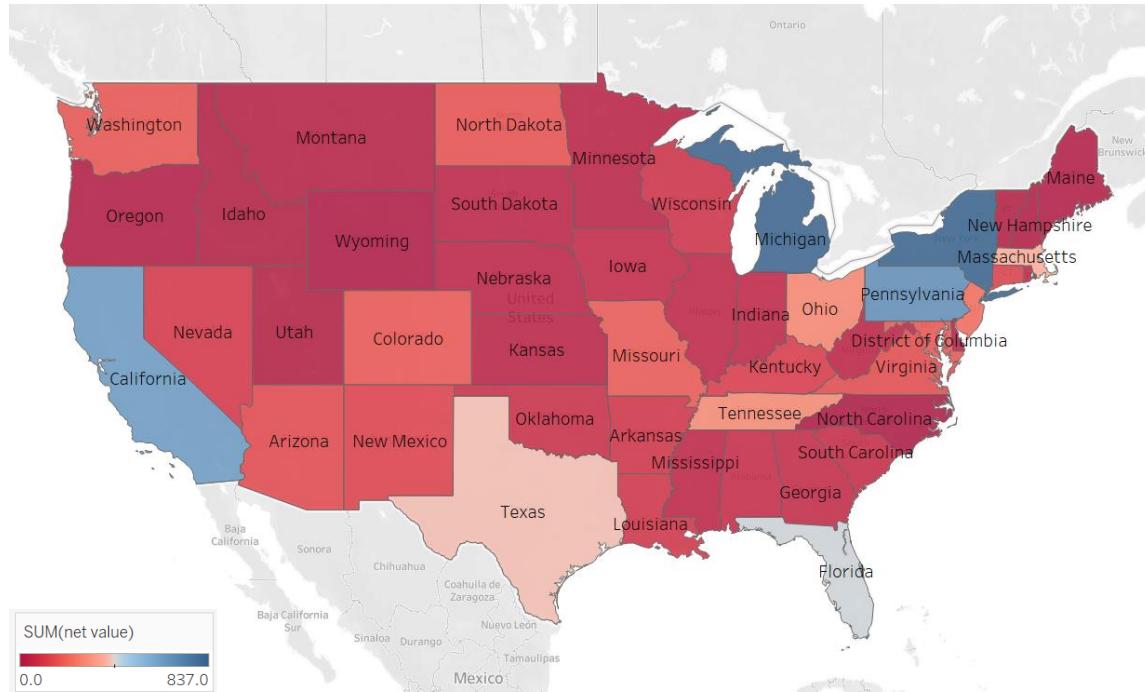
Week o...	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Week 45			X -243	X -572	X -396	X -156	X 165
Week 46	X 179	X 62	✓ -24	X -773	X -332	X 139	X 190
Week 47	X 182	X -56	X -206	X -670	X -266	X -314	X 172
Week 48	X 178	✓ -16	X -370	X -842	X 61	✓ 43	X 180
Week 49	X 181	X -108	X -226	X -393			

Bass Model on State degree- Weekly

- Predicted death in November 2022--
1966, 1908, 1852, 1797
- Quite similar between predicted values and
real values
- Performs better than the daily model
- A decreasing tendency for predicted model
while real value is fluctuating



Bass Model on State degree-Monthly



Red areas have better performance in approaching bass model

Predicted Death in November 2022

8916

Difference between predicted death and real death:

1068

No fit area: not suitable for Bass Model

California, Michigan, New York

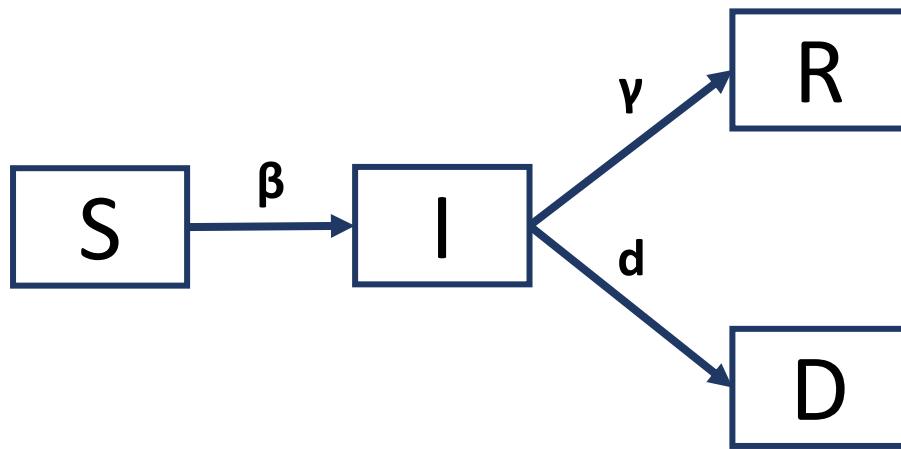
- These states have a huge difference between predicted and actual death when running a bass model

Compare All State Degree Bass Models

	2022/11/30	Real Death	2022/11/21-2022/11/27	Real death	2022/11	Real death
Daily Bass model	184		1312		5755	
Weekly Bass model		577	1798	2074	7525	9985
Monthly Bass model					8926	

SIRD Model Intro

The Susceptible-Infectious-Recovered-Deceased model



S = The number of susceptible individuals
(S_0 : number of initial susceptible individuals)

I = The number of infectious individuals
(I_0 : the number of individuals who get infected)

Beta = Contact Rate

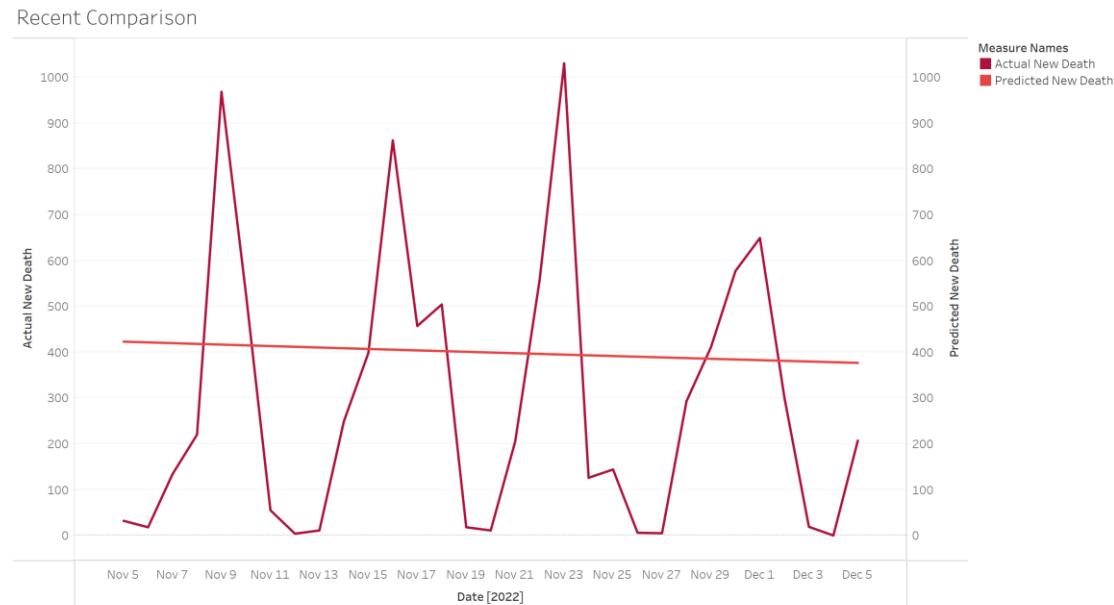
Gamma = Recovery Rate

d = Death Rate

SIRD Model

Comparisons and Predictions

	2022/11/30	Real Death	2022/11/21-2022/11/27	Real death	2022/11	Real death
SIRD	384	577	2572	2074	12191	9985



Daily Predictions for Dec 16, 2022

360

Monthly Prediction for Dec 2022

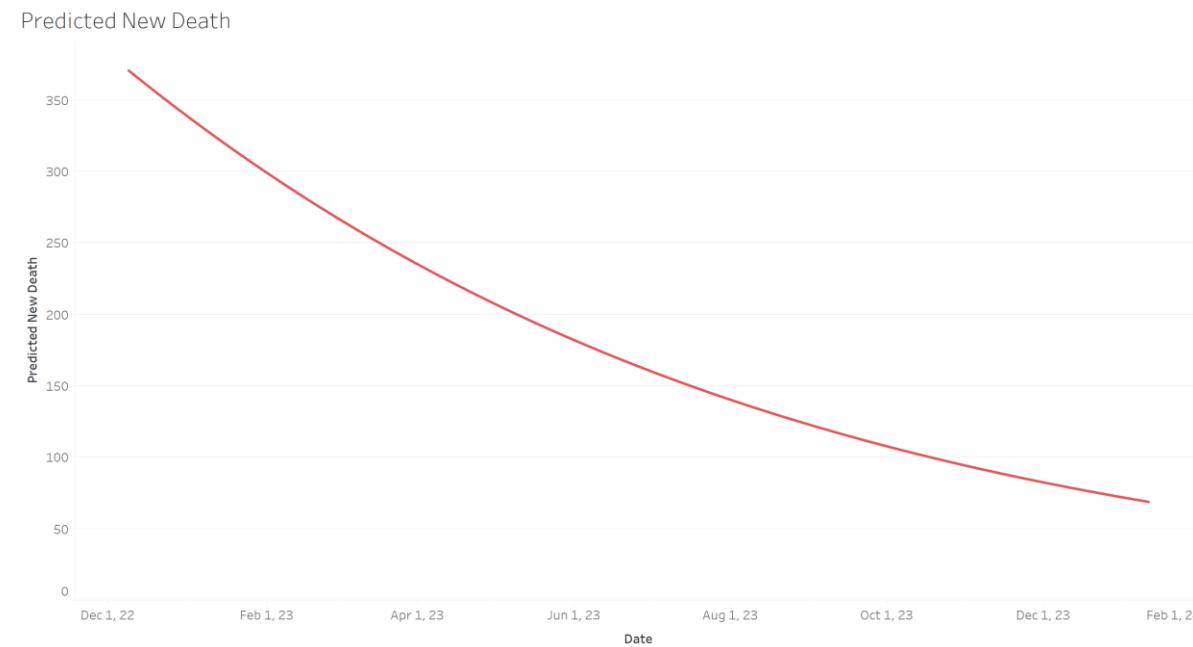
11189

Yearly Prediction for 2023

63827

SIRD Model

Predicted Future Trends



Model Trends for Future Deaths

Decreasing Trend

Parameters for S0 + I0

18.3 M:
Total number of people who will be influenced by COVID-19

Generalized Bass Model (GBM)

Allow seeing the influence of Vaccines

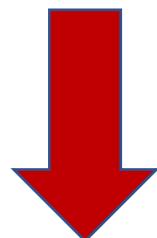
$$A(t) = M \frac{1 - \exp(-(p+q)t^*)}{1 + \frac{q}{p} \exp(-(p+q)t^*)}$$



p	0.000652209
q	0.007680922
M	669025.1021
b	-34.63850985

$$t^* = t + b * \ln [\text{Vaccine}(t) / \text{Vaccine}(1)]$$

$$- \frac{*}{-} +$$



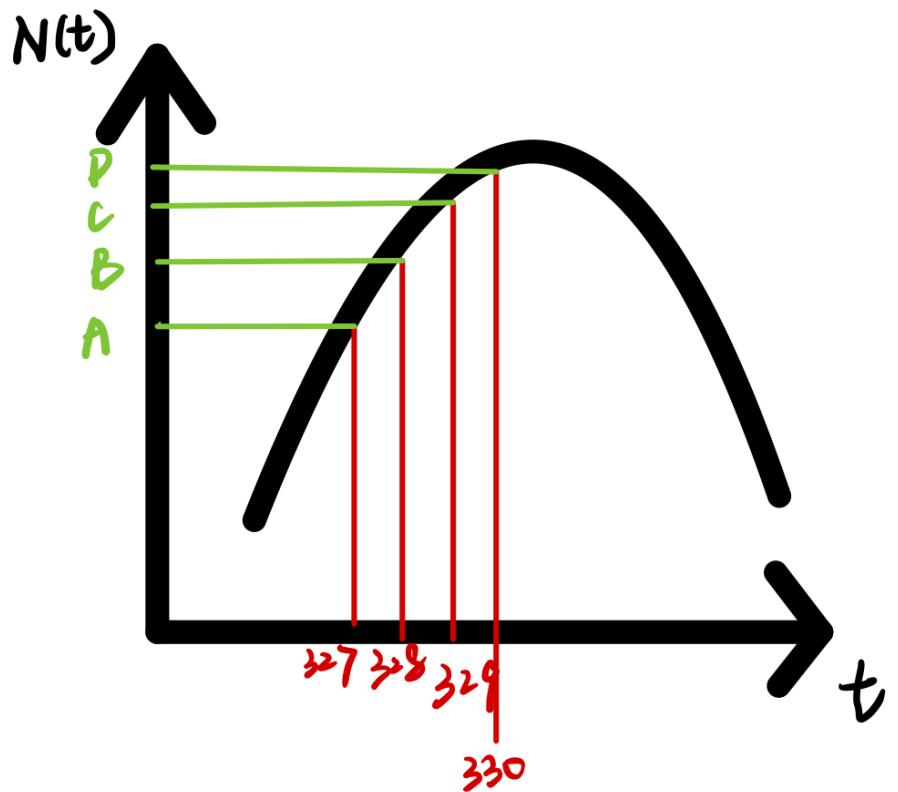
b is negative!

$$t^* < t$$

Generalized Bass Model (GBM)

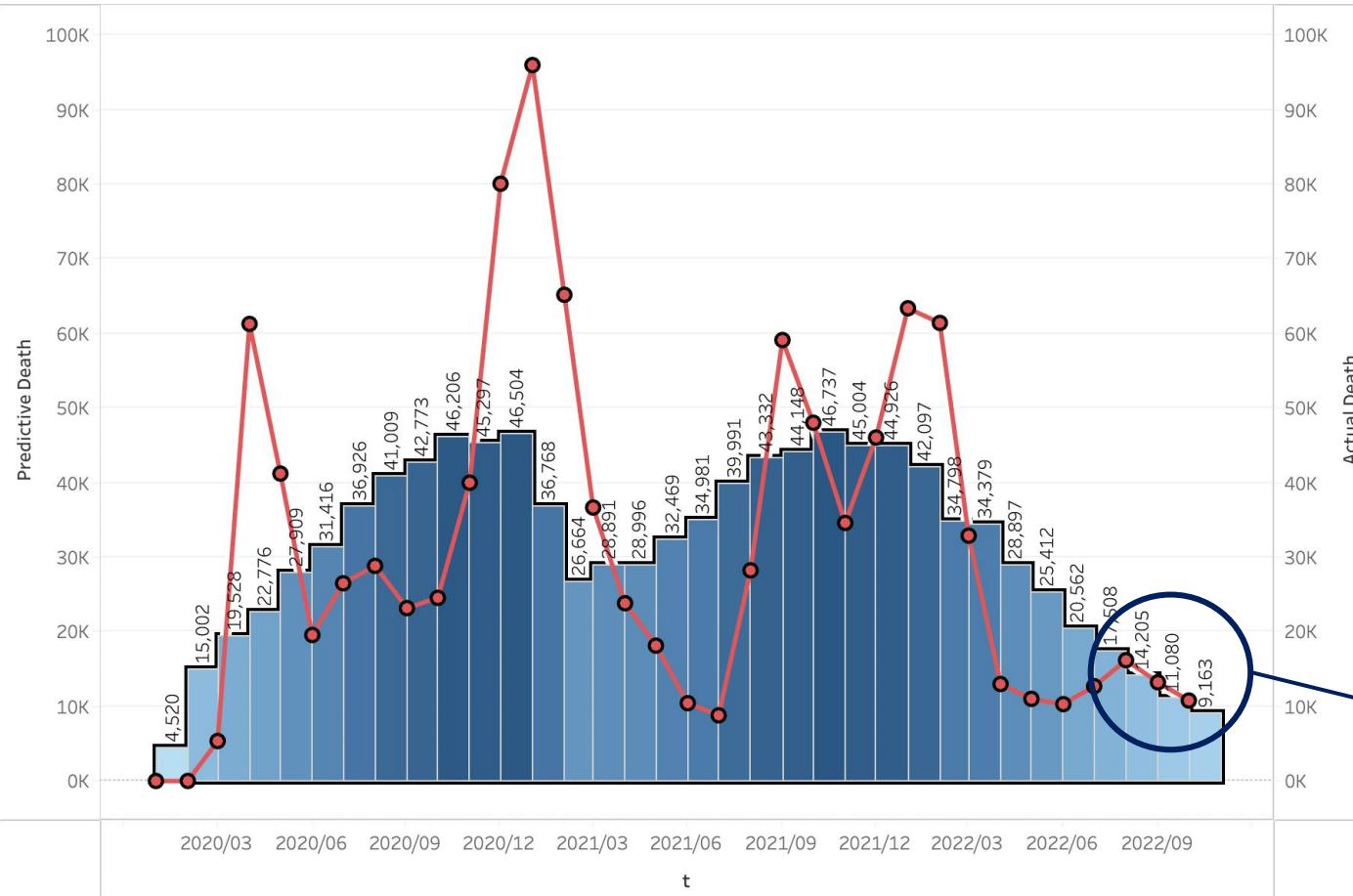
Vaccines effectively slower the speed of new death.

	t	t^*	
B	328	327.40	A
C	329	> 327.21	A
D	330	326.35	A



Generalized Bass Model (GBM)

Prediction accuracy evaluation

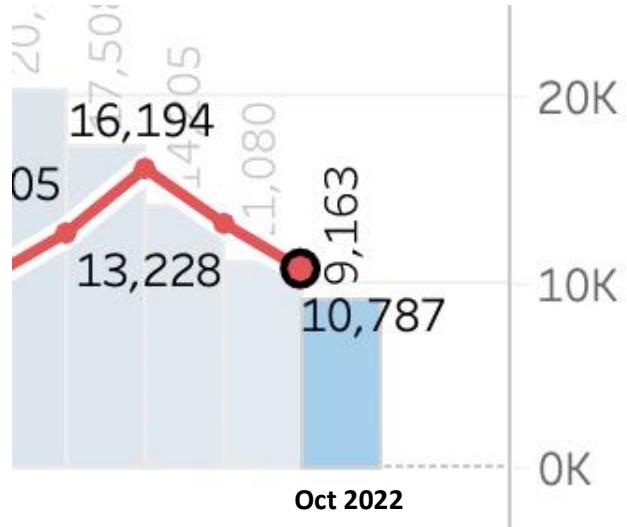


Predictive total death: October 1st–October 31st

9163 Blue Bar chart

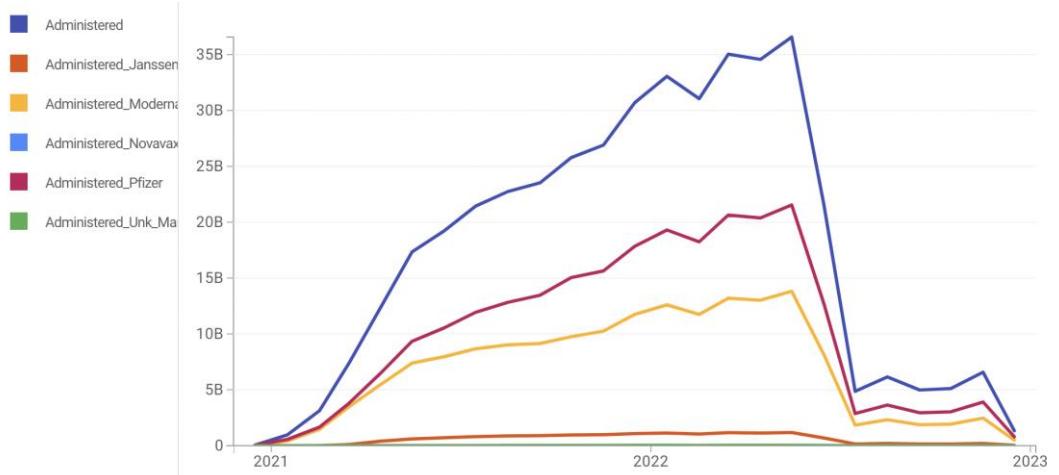
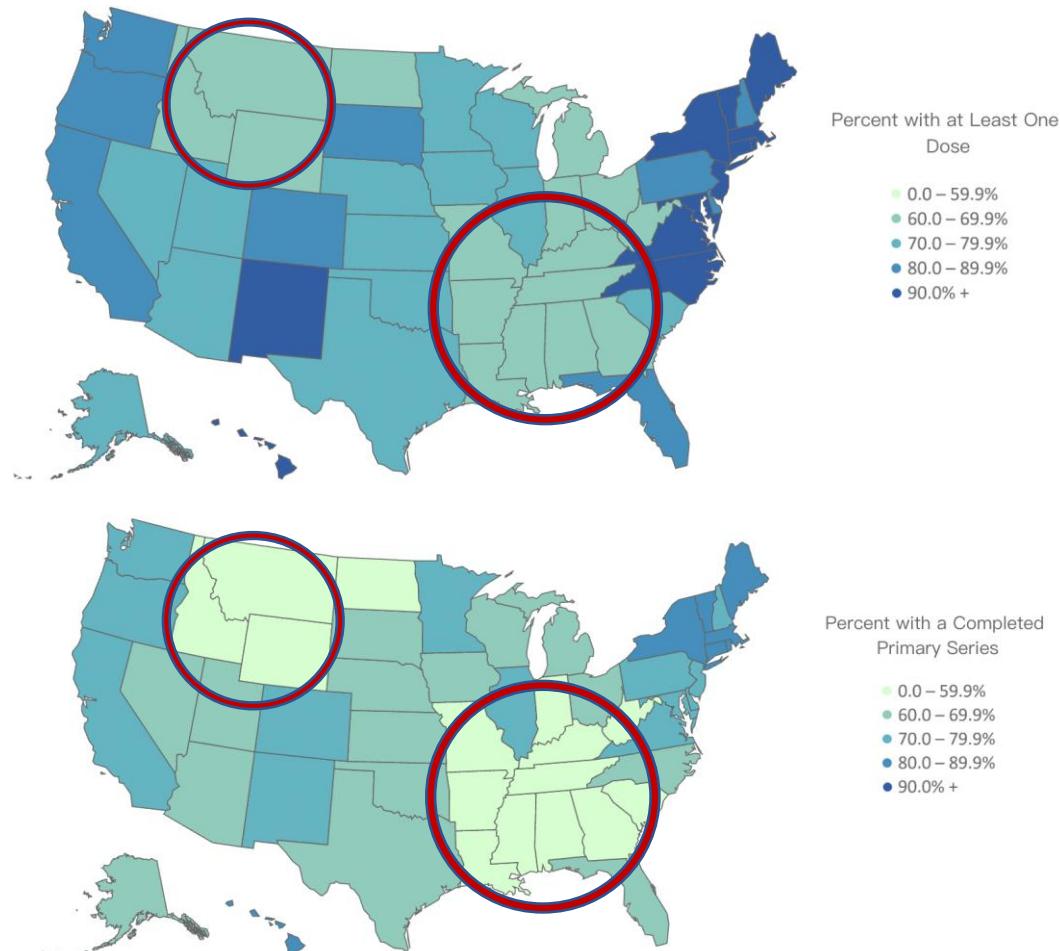
Actual total death:

10787 Red Line chart



Generalized Bass Model (GBM)

Overview of Vaccines

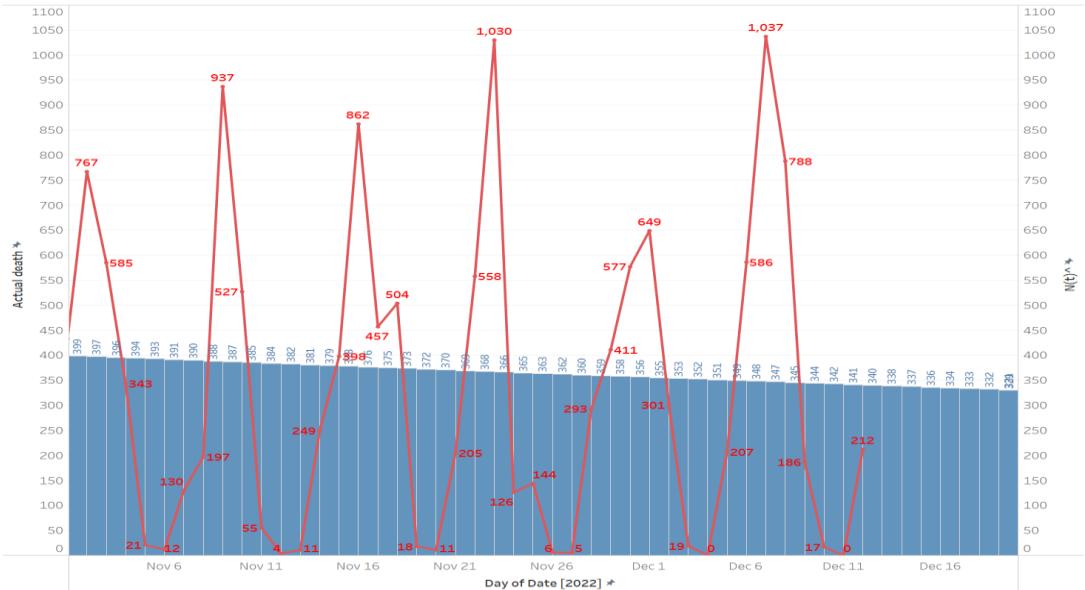


Focus more on both the north and southeast areas since these areas have lower percentage of vaccine administered.

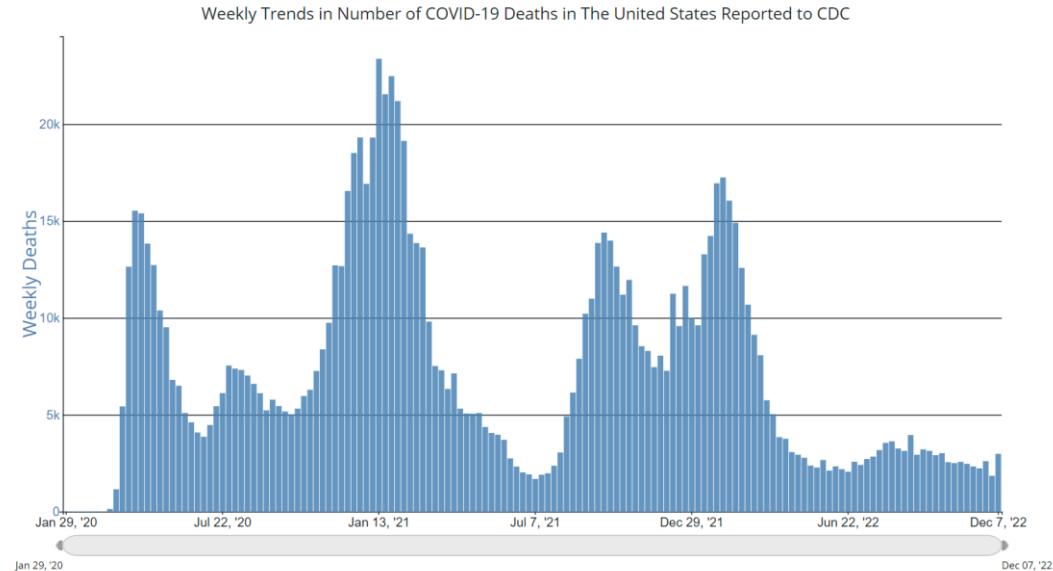
04

Recommendation

Focus on monthly/weekly forecast



The data of daily deaths comes from JHU dashboard. The report of COVID deaths is more weekly based and fluctuating.



CDC only reports weekly deaths/cases on the official website now.

https://covid.cdc.gov/covid-data-tracker/#trends_weeklydeaths_select_00

Use SIRD/state level bass model as a prediction model

Models	Next day (11.30)	7-days (11.21-11.27)	next month 11
US Level-Bass model	356	2728	11057
State Level-Bass model	184	1312 daily 1798 weekly	5755 daily 7525 weekly 8926 monthly
SIRD	384	2572	12191
Actual	577	2074	9985

- Next day — SIRD
- 7-days — State Level Bass model (Weekly basis)
- Next month — State Level Bass model (Monthly basis)

Focus on separating into smaller areas with bass prediction will yield a more accurate prediction (US level—State level—City level)

Models	Next day (11.30)	7-days (11.21- 11.27)	next month 11
US Level-Bass model	356	2728	11057
State Level-Bass model	184	1312 daily 1798 weekly	5755 daily 7525 weekly 8926 monthly
SIRD	384	2572	12191
Actual	577	2074	9985

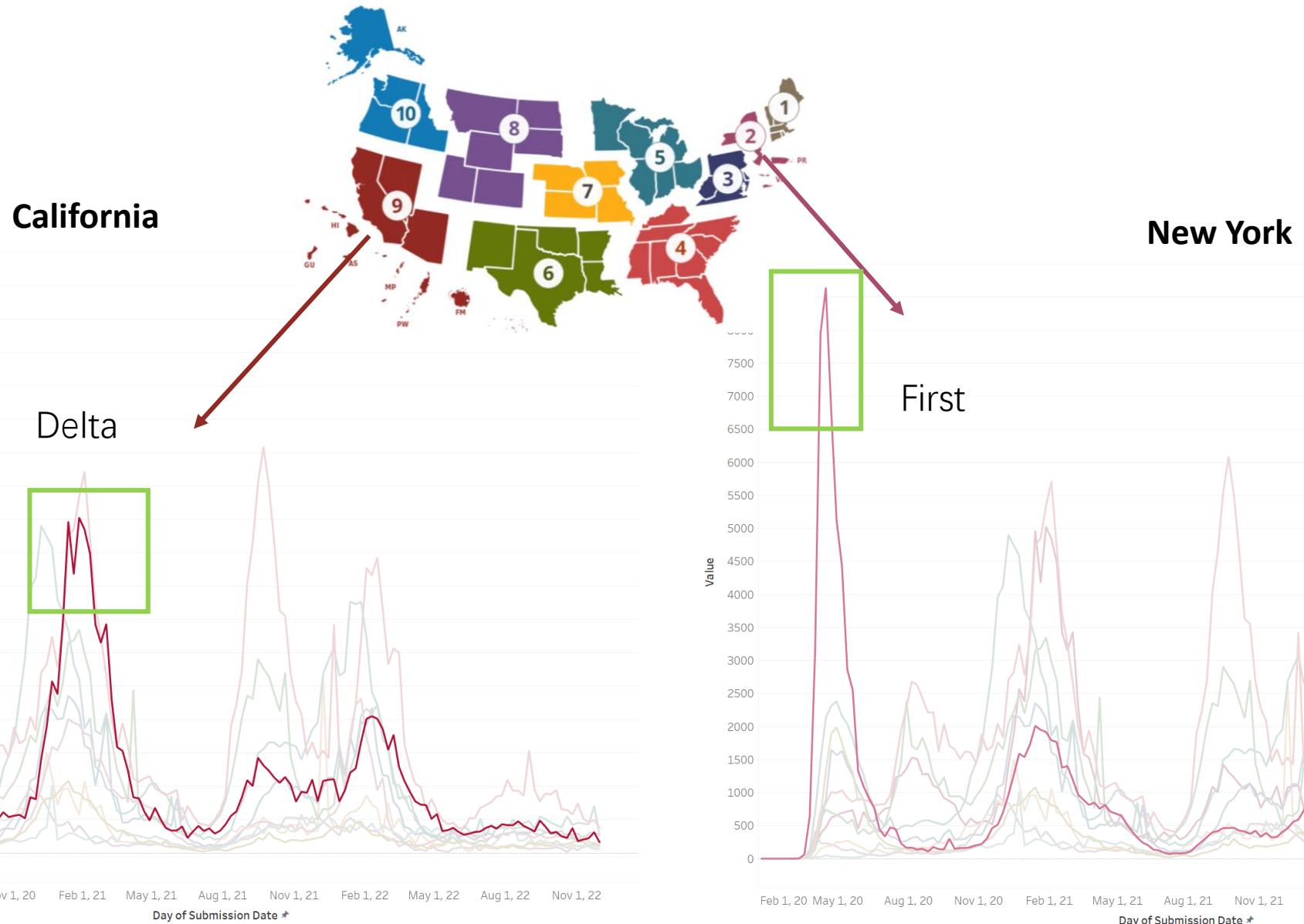


Overall, bass model in a state level yield a more accurate prediction. We could focus on city level in the future

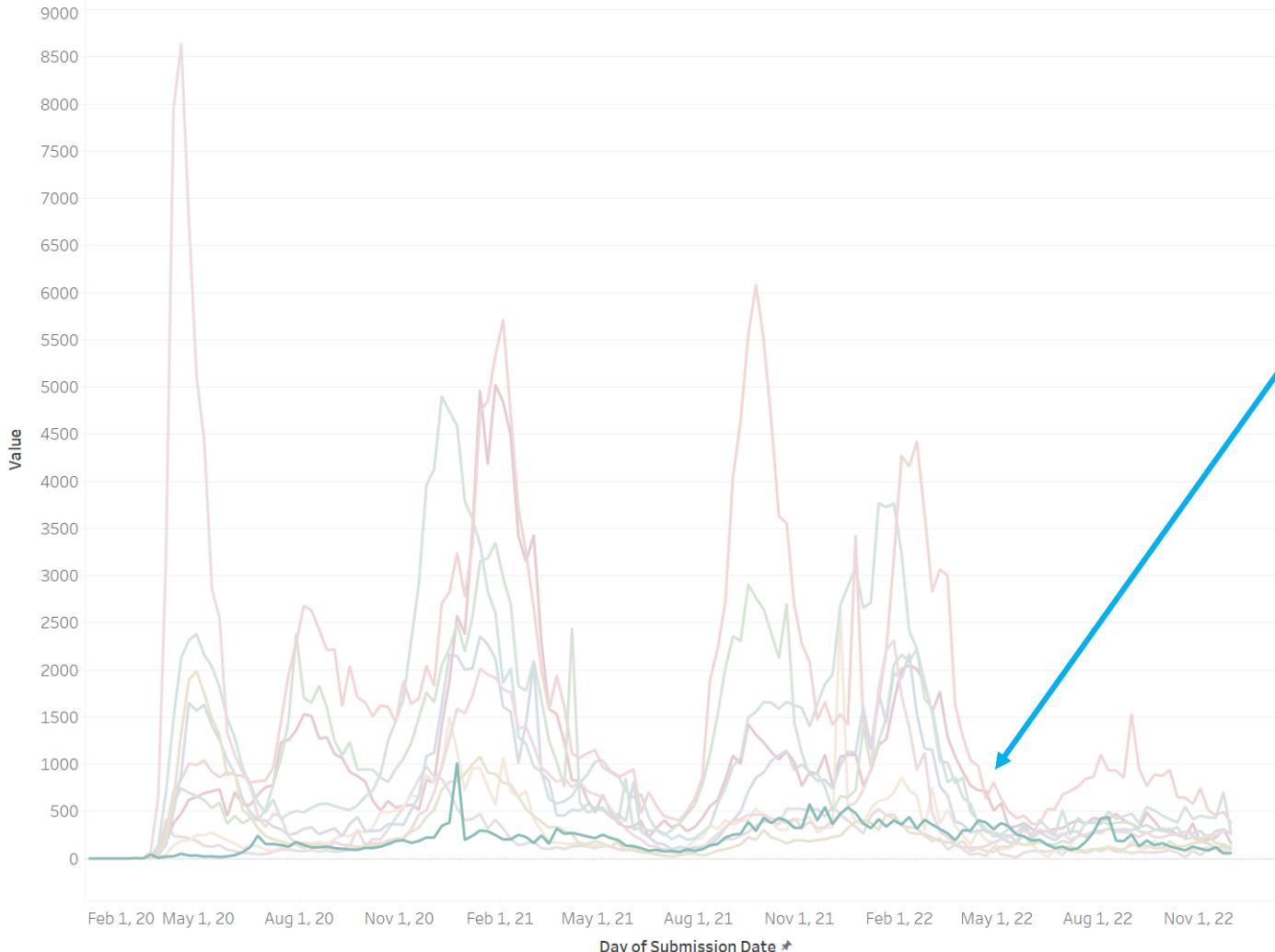


Running model in a regional level and then aggregate data together in a US level would be better.

Promote reinsurance in certain areas like Seattle and Florida



Promote reinsurance in certain areas like Seattle and Florida

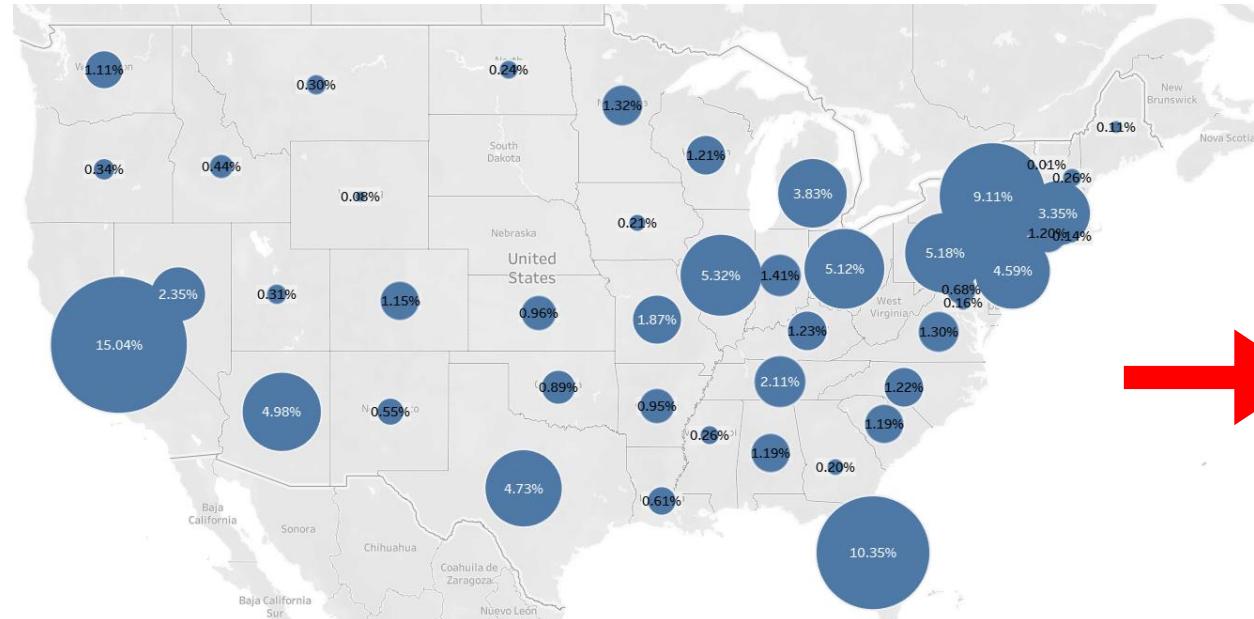


Regions that have suffered deaths from previous wave of COVID variant would have fewer deaths in the next wave



Seattle area may encounter a great loss of deaths if a new strong variant is introduced because it has never had a huge wave of COVID breakout.

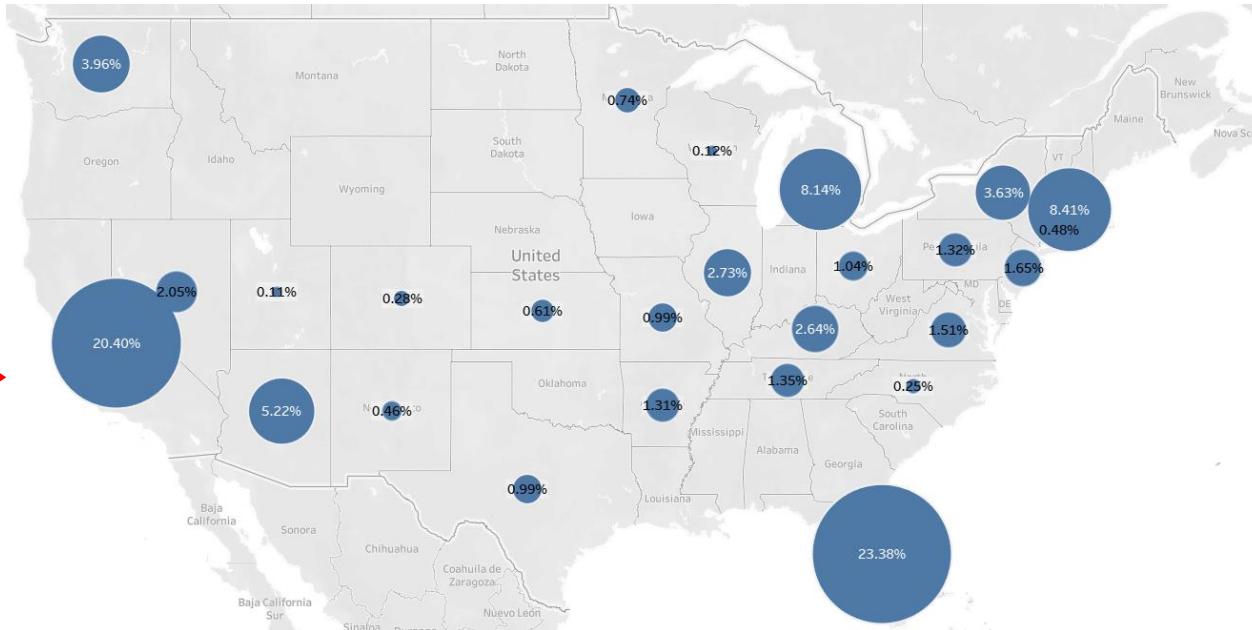
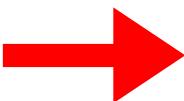
Promote reinsurance in certain areas like Seattle and Florida



Before June 2022



New York area and California have a high percentage of death contributed before.

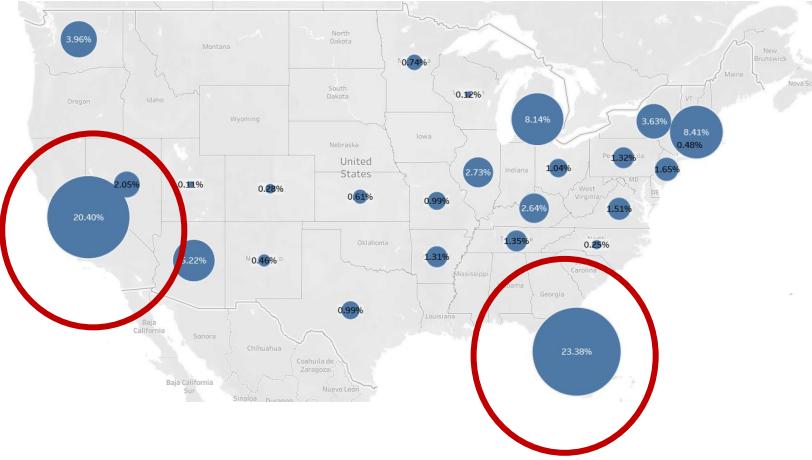
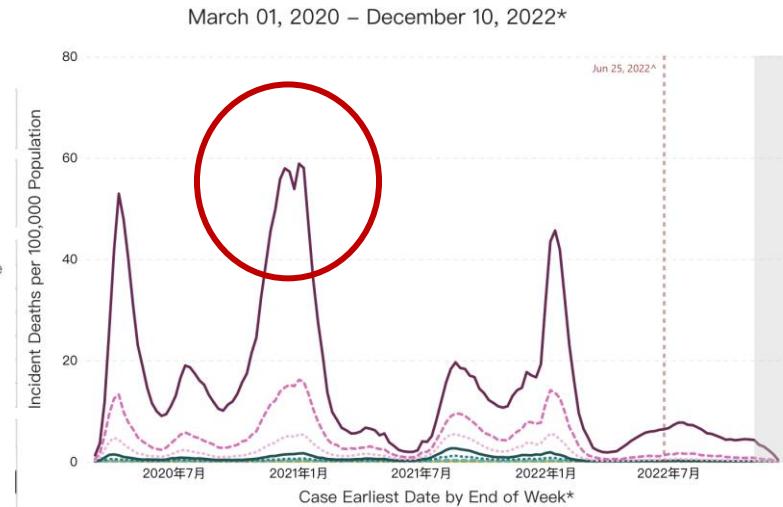
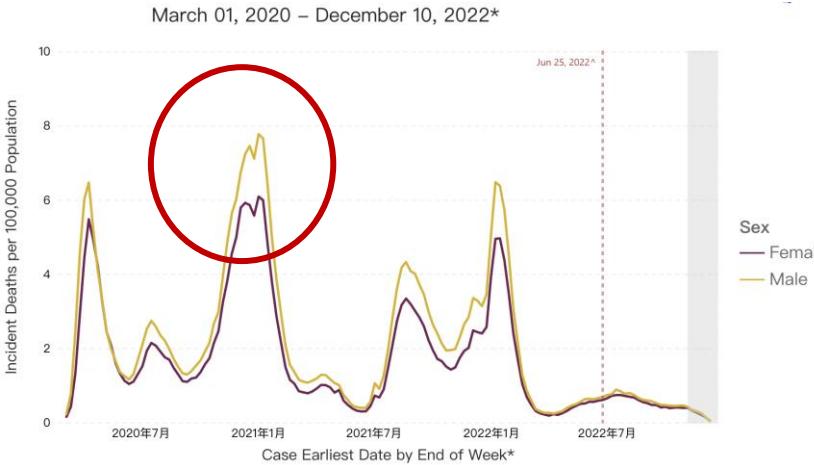


Last 6 Month



Florida has increasing percent of death in the last 6 month which may be due to a higher percentage of people over 65





It is obvious that compared to females, males have a higher mortality rate due to Covid-19.

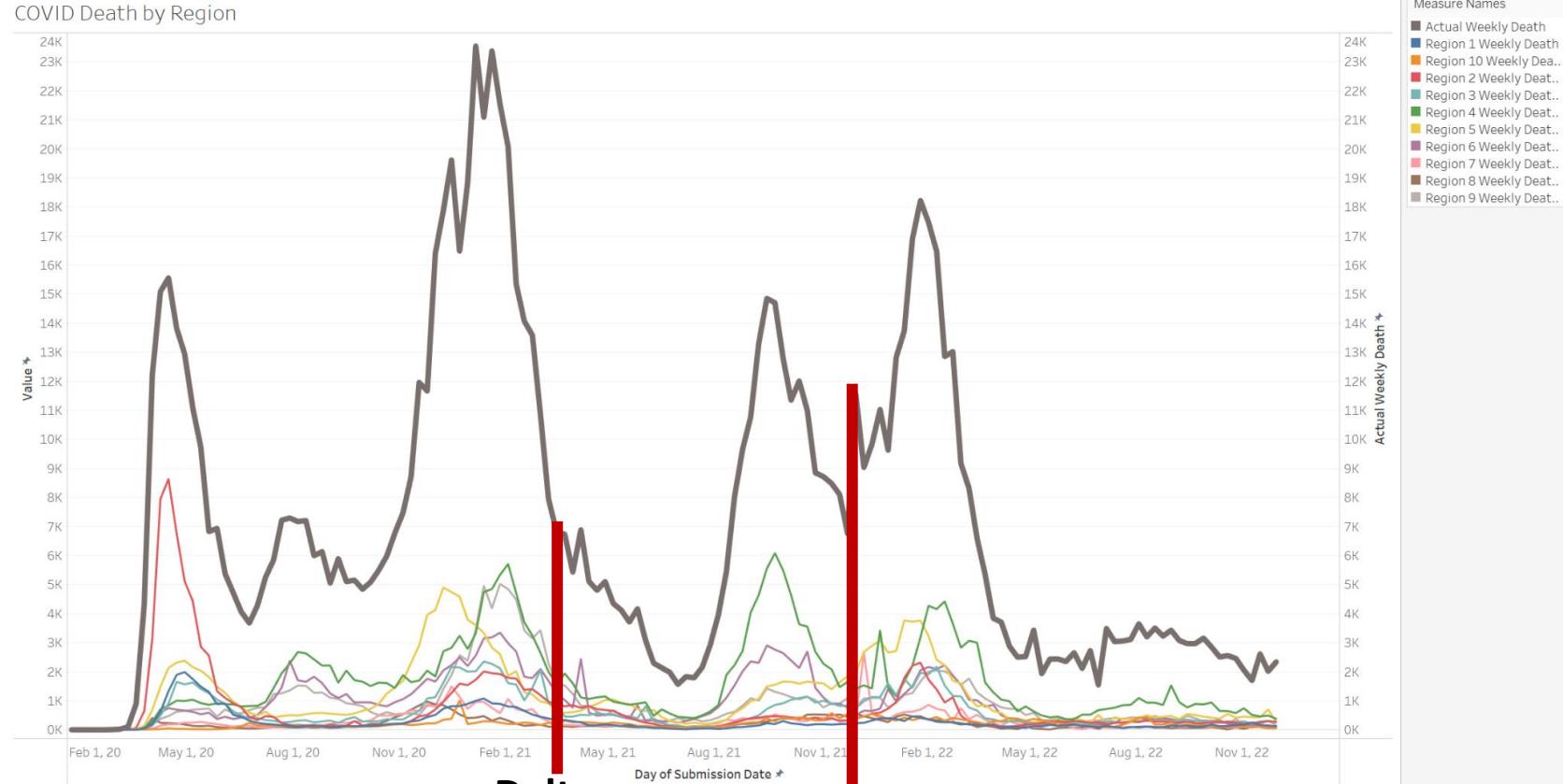


**Among different age groups,
people aged 75+ are the most
vulnerable group.**



Both FL and CA are experiencing a high percentage of the death rate.

Pay great attention to the new variant of covid virus entrance

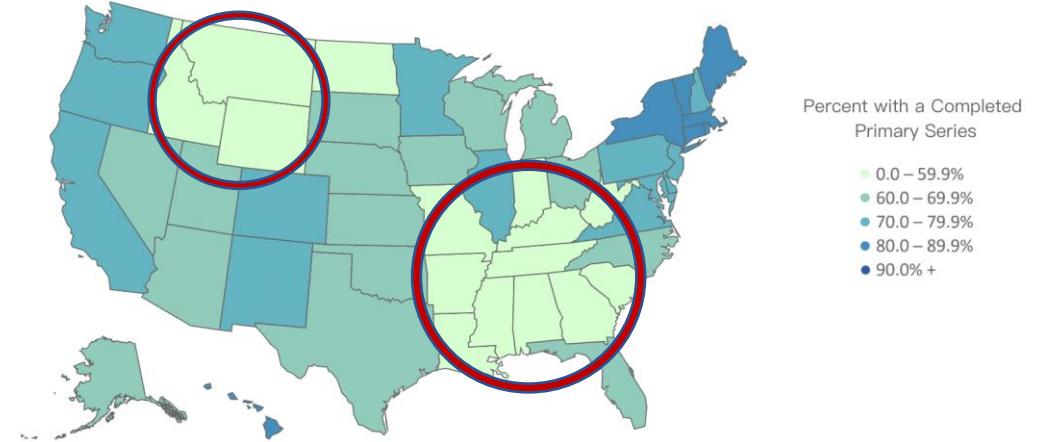
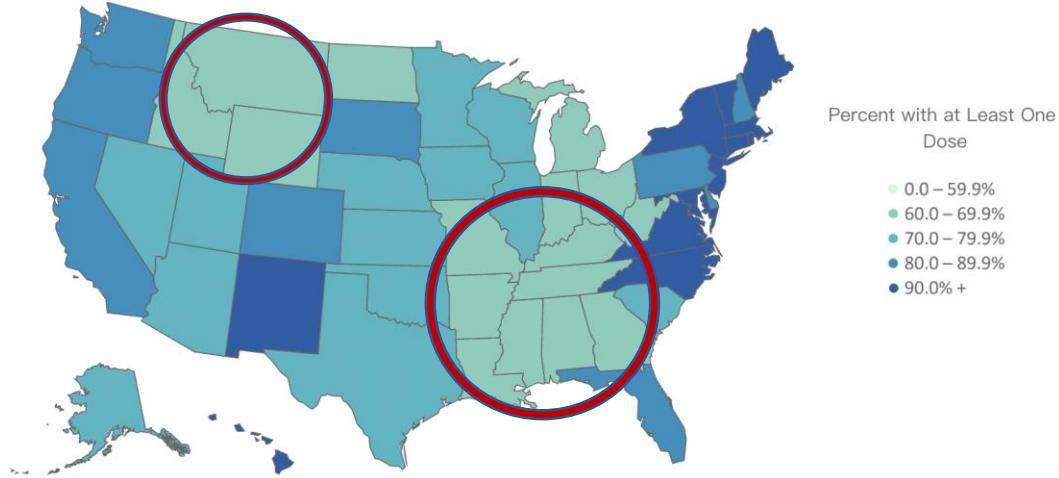


As soon as a **new variant** of
COVID-19 virus enters U.S.,
an **outbreak** follows.

The first detected case in the U.S. was in **March** 2021

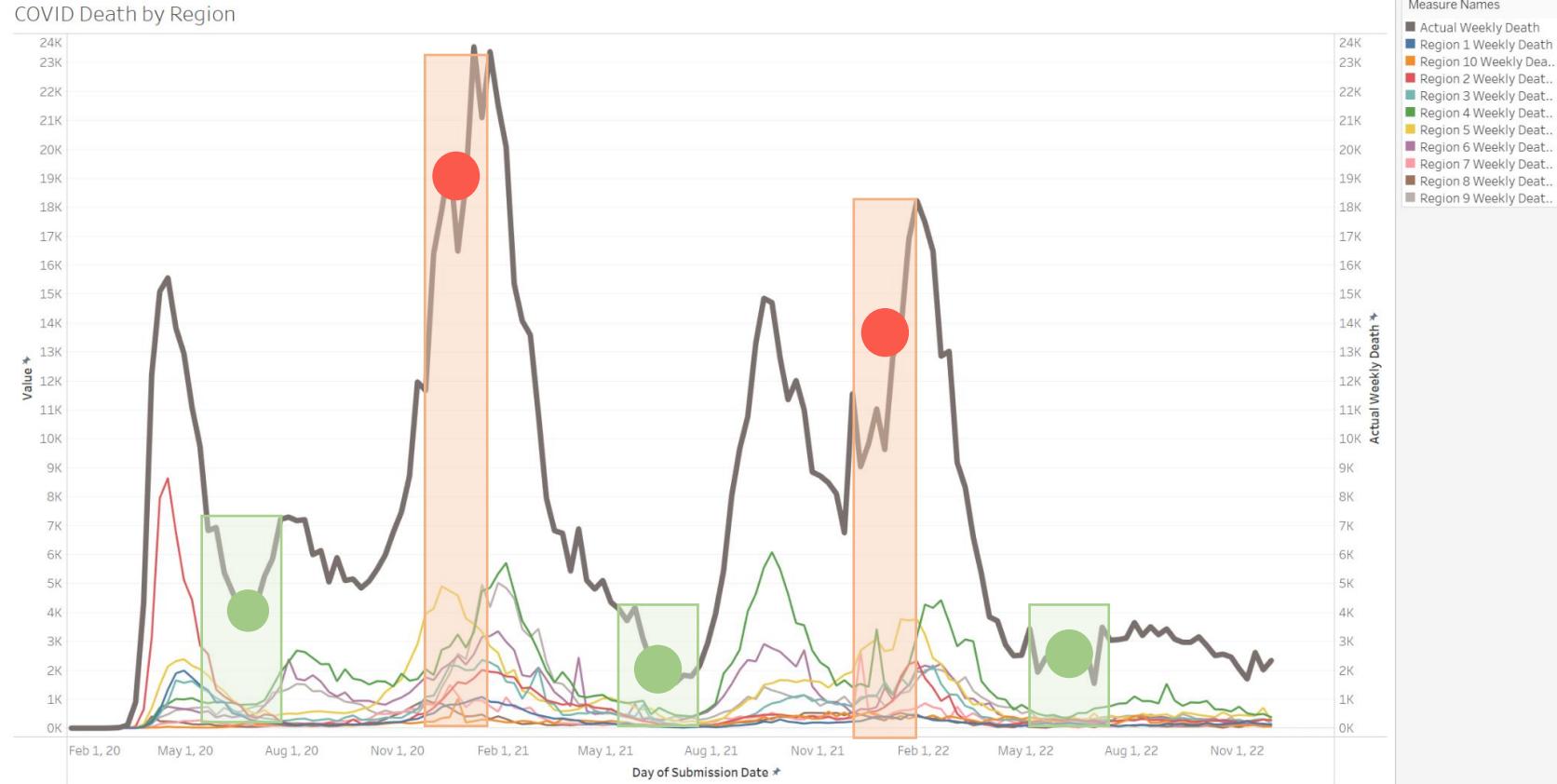
The first detected case in the U.S. was in **November** 2021

Targeting the north and southeast areas for business. (vaccine)



Since the number of deaths **decreases** with the **vaccine** effects, the company should target the north and southeast areas that have a lower percentage in both "completed primary series" and "with at least one dose".

Focus on seasonal patterns (winter).



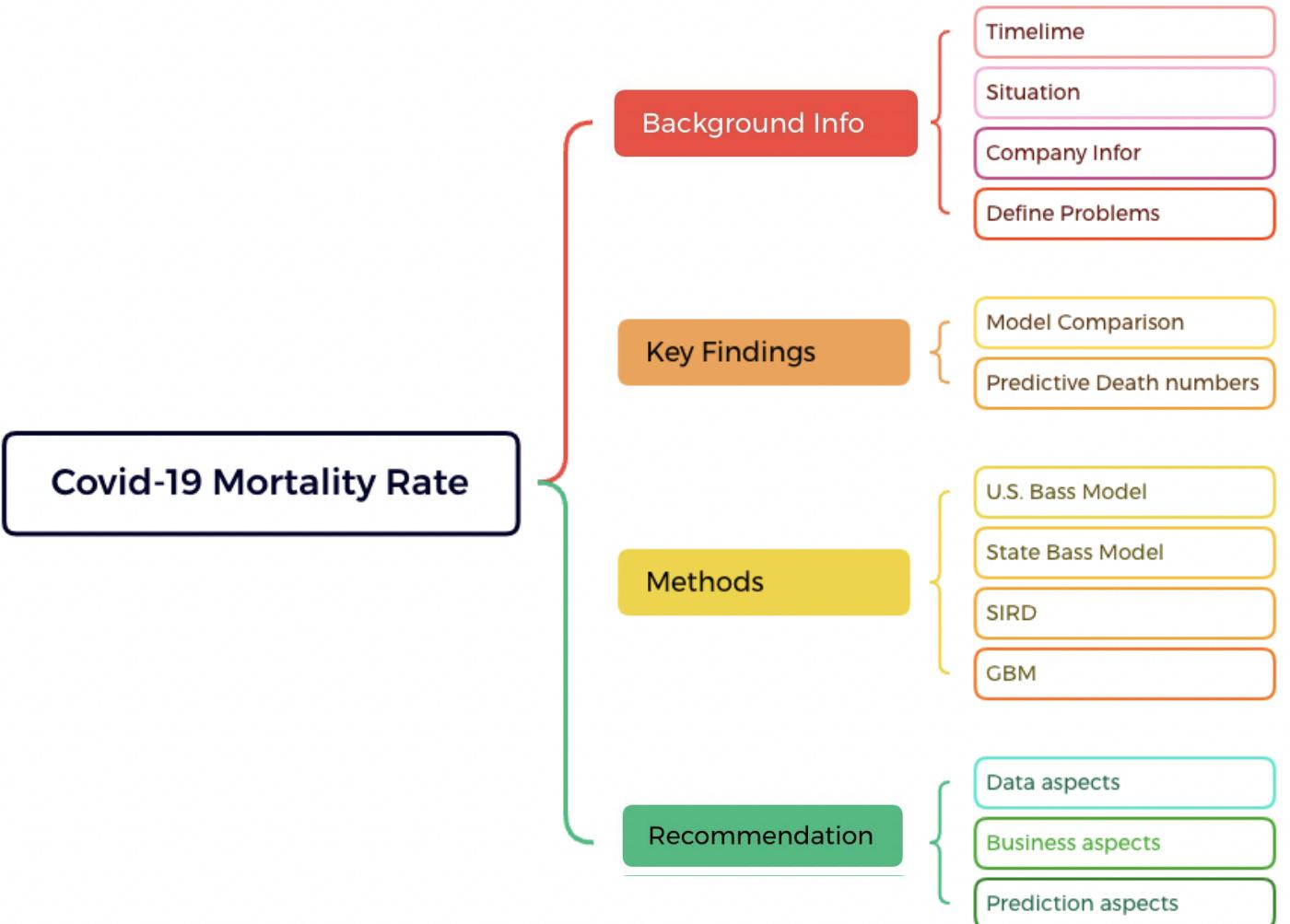
Winter Period:
Dec - Feb

Around peak time

Summer Period:
May - Aug

Slowing death rates

Q&A



Appendix—Bass Model Daily

Appendix—Bass Model Weekly

D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S								
Week #	Week	Actual Weekly Death	A(t-1)	A(t-1)^2	N(t)^A	A(t)^A	A(t-1)^A	Difference between actual and estimated	SUMMARY OUTPUT														
1	1/22/2020	1	0	0	4532.9865	4532.98649	0	20538901.55	Regression Statistics														
2	1/29/2020	0	1	1	4636.85	9169.83646	4532.986491	21500377.68															
3	2/5/2020	0	1	1	4742.1042	13911.9406	9169.836464	22487551.99	Multiple R	0.453803603													
4	2/12/2020	0	1	1	4848.7133	18760.6539	13911.94064	23510020.77	R Square	0.20593771													
5	2/19/2020	0	1	1	4956.6381	23717.292	18760.65395	24568261.25	Adjusted R Square	0.194909067													
6	2/26/2020	7	1	1	5065.8356	28783.1277	23717.29205	25591818.08	Standard Error	4975.41458													
7	3/4/2020	21	8	64	5176.2593	33959.387	28783.12769	26576698.92	Observations	147													
8	3/11/2020	104	29	841	5287.8588	39247.2458	33959.38704	26872392.17															
9	3/18/2020	900	133	17689	5400.5798	44647.8256	39247.24585	20255218.22	ANOVA														
10	3/25/2020	4326	1033	1067089	5514.364	50162.1896	44647.82562	1412208.952	df	SS	MS	F	Significance F										
11	4/1/2020	12226	5359	28718881	5629.1492	55791.3388	50162.1896	43518440.43	Regression	2	924490277.8	462245138.9	18.67298738	6.15927E-08									
12	4/8/2020	15091	17585	309232225	5744.8691	61536.2079	55791.3388	87350162.96	Residual	144	3564684034	24754750.24											
13	4/15/2020	15563	32676	1067720976	5861.4532	67397.6611	61536.20789	94120011.01	Total	146	4489174312												
14	4/22/2020	13830	48239	2327001121	5978.8268	73376.4878	67397.66106	61640921.18															
15	4/29/2020	12941	62069	3852560761	6096.911	79473.3988	73376.48782	46841553.95	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%									
16	5/6/2020	11086	75010	5626500100	6215.6228	85689.0217	79473.39884	23720573.67	Intercept	4526.817725	1018.762038	4.443449556	1.75388E-05	2513.158083	6540.4773								
17	5/13/2020	9726	86096	7412521216	6334.8749	92023.8966	85689.02168	11499729.49	A(t-1)	0.023637532	0.00453254	5.215074251	6.27086E-07	0.014678627	0.0325964								
18	5/20/2020	6826	95822	9181855684	6454.5756	98478.4722	92023.89658	137956.0664	A(t-1)^2	-2.35286E-08	4.02309E-09	-5.848382037	3.19747E-08	-3.14805E-08	-1.55766E-								
19	5/27/2020	6936	102648	10536611904	6574.6293	105053.101	98478.4722	130588.7983															
20	6/3/2020	5377	109584	12008653056	6694.9359	111748.037	105053.1015	1736955.126	p														
21	6/10/2020	4720	114961	13216031521	6815.3916	118563.429	111748.0374	4390665.789	q														
22	6/17/2020	4060	119681	14323541761	6935.8881	125499.317	118563.429	8270732.279	M														
23	6/24/2020	3677	123741	15311835081	7056.3135	132555.631	125499.3171	11419759.64															
24	7/1/2020	4282	127418	16235346724	7176.5519	139732.182	132555.6305	8378430.658	M														
25	7/8/2020	5251	131700	17344890000	7296.4837	147028.666	139732.1824	4184003.613	p														
26	7/15/2020	5855	136951	18755576401	7415.9858	154444.652	147028.6661	2436676.575	q														
27	7/22/2020	7220	142806	20393553636	7534.9315	161979.583	154444.6519	99181.83715															
28	7/29/2020	7295	150026	22507800676	7653.191	169632.774	161979.5834	128300.8037															
29	8/5/2020	7170	157321	24749897041	7770.6315	177403.406	169632.7744	360758.2208	3606013932														
30	8/12/2020	7205	164491	27057289081	7887.1173	185290.523	177403.4059	465284.0334															
31	8/19/2020	5007	171606	29470516416	8002.5102	192202.022	185290.522	4022021.021															

Appendix—Bass Model Monthly

	A	B	C	D	E	F	G	H	I	J	K
1	time	t	N(t)	A(t)	A(t-1)	p-N(t)	p-A(t)	p-A(t-1)	difference		
2	1/1/2020	1		0	0	18242.22	18242.221		0 3.33E+08		
3	2/1/2020	2		1		0 20243.7	38485.924	18242.22	4.1E+08		
4	3/1/2020	3		3326	3327	1 22368.31	60854.238	38485.92	3.63E+08		
5	4/1/2020	4		57134	60461	3327 24598.69	85452.927	60854.24	1.06E+09		
6	5/1/2020	5		44864	105325	60461 26910.44	112363.36	85452.93	3.22E+08		
7	6/1/2020	6		22925	128250	105325 29271.63	141634.99	112363.4	40279686		
8	7/1/2020	7		28717	156967	128250 31642.68	173277.67	141635	8559581		
9	8/1/2020	8		31831	188798	156967 33976.75	207254.41	173277.7	4604225		
10	9/1/2020	9		21361	210159	188798 36220.81	243475.22	207254.4	2.21E+08		
11	10/1/2020	10		23448	233607	210159 38317.42	281792.64	243475.2	2.21E+08		
12	11/1/2020	11		41105	274712	233607 40207.18	321999.82	281792.6	806087.9		
13	12/1/2020	12		77292	352004	274712 41831.82	363831.64	321999.8	1.26E+09	p	0.014924983
14	1/1/2021	13		100119	452123	352004 43137.69	406969.33	363831.6	3.25E+09	q	0.122974788
15	2/1/2021	14		67730	519853	452123 44079.24	451048.56	406969.3	5.59E+08	m	1158164.508
16	3/1/2021	15		33397	553250	519853 44622.31	495670.88	451048.6	1.26E+08		
17	4/1/2021	16		20768	574018	553250 44746.72	540417.59	495670.9	5.75E+08		
18	5/1/2021	17		17442	591460	574018 44447.76	584865.35	540417.6	7.29E+08	Sum-difference	15170206754
19	6/1/2021	18		9459	600919	591460 43736.66	628602.01	584865.4	1.17E+09		
20	7/1/2021	19		9045	609964	600919 42639.64	671241.66	628602	1.13E+09		
21	8/1/2021	20		31496	641460	609964 41195.93	712437.59	671241.7	94088673		
22	9/1/2021	21		54691	696151	641460 39454.84	751892.43	712437.6	2.32E+08		
23	10/1/2021	22		48578	744729	696151 37472.36	789364.79	751892.4	1.23E+08		
24	11/1/2021	23		31309	776038	744729 35307.64	824672.43	789364.8	15989152		
25	12/1/2021	24		43017	819055	776038 33019.61	857692.04	824672.4	99947800		
26	1/1/2022	25		65001	884056	819055 30664.15	888356.19	857692	1.18E+09		
27	2/1/2022	26		63682	947738	884056 28291.89	916648.08	888356.2	1.25E+09		
28	3/1/2022	27		27637	975375	947738 25946.77	942594.85	916648.1	2856863		
29	4/1/2022	28		13263	988638	975375 23665.23	966260.09	942594.9	1.08E+08		
30	5/1/2022	29		9953	998591	988638 21476.04	987736.13	966260.1	1.33E+08		
31	6/1/2022	30		9885	1008476	998591 19400.6	1007136.7	987736.1	90546646		
32	7/1/2022	31		13117	1021593	1008476 17453.59	1024590.3	1007137	18806055		
33	8/1/2022	32		13716	1035309	1021593 15643.82	1040234.1	1024590	3716490		
34	9/1/2022	33		12795	1048104	1035309 13975.13	1054209.3	1040234	1392700		

Appendix—Bass Model State Level

```
url = "https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_US.csv"
df = pd.read_csv(url, index_col=0)
df.head()
```

	iso2	iso3	code3	FIPS	Admin2	Province_State	Country_Region	Lat
UID								
84001001	US	USA	840	1001.0	Autauga	Alabama	US	32.539527 -86.6
84001003	US	USA	840	1003.0	Baldwin	Alabama	US	30.727750 -87.7
84001005	US	USA	840	1005.0	Barbour	Alabama	US	31.868263 -85.3
84001007	US	USA	840	1007.0	Bibb	Alabama	US	32.996421 -87.1
84001009	US	USA	840	1009.0	Blount	Alabama	US	33.982109 -86.5

5 rows × 1059 columns

← →

```
death_by_state = df.groupby('Province_State')[df.columns[11:]].sum()
death_by_state.head(10)
```

1/22/20 1/23/20 1/24/20 1/25/20 1/26/20 1/27/20 1/28/20 1/29/20 1/30/

Province_State	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	1/28/20	1/29/20	1/30/
Alabama	0	0	0	0	0	0	0	0	0
Alaska	0	0	0	0	0	0	0	0	0
American Samoa	0	0	0	0	0	0	0	0	0
Arizona	0	0	0	0	0	0	0	0	0
Arkansas	0	0	0	0	0	0	0	0	0
California	0	0	0	0	0	0	0	0	0
Colorado	0	0	0	0	0	0	0	0	0
Connecticut	0	0	0	0	0	0	0	0	0
Delaware	0	0	0	0	0	0	0	0	0
Diamond Princess	0	0	0	0	0	0	0	0	0

```
test = pd.DataFrame(death_by_state.iloc[1])
current_state = test.columns[0]
current_state
##### WARNING: is the data already counting cumulative or is it daily cases?
test.columns=['cum_death']
test['new_death']=test['cum_death'].diff().fillna(test['cum_death'])
#print(df)
#test.reset_index()
#test.head()
```

```
from math import exp
def model(params, time_step):
    p,q,m = params
    N=[]
    for t in range(1,time_step):
```

Appendix—Bass Model State Level

```
current=m*(1-exp(-(p+q)*t))/(1+(q/p)*exp(-(p+q)*t))
last=m*(1-exp(-(p+q)*(t-1)))/(1+(q/p)*exp(-(p+q)*(t-1)))
N.append(current-last)
return N

def error(params, time_step, actual_N):
    pred_N = model(params,time_step)
    N_error = []
    for t in range(0,time_step-1):
        N_error.append((actual_N[t+1]-pred_N[t])**2)
    total_error = sum(N_error)
    return total_error

import numpy as np
from scipy.optimize import minimize
import pandas as pd

initial_value = [0.027,0.003,0.03]
bnds = [(0, None), (0, None), (0, None)]
res = minimize(error,initial_value, args = (1028,test.new_death), bounds = bnds, tol=1e-8, method = 'Powell')
```

```
print(res)

direc: array([[-1.20012006e-05,  7.47837393e-04,  9.60104503e+01],
   [-3.04291749e-05,  1.12323411e-03, -1.39193406e+02],
   [-4.78326155e-06,  5.60700835e-04, -6.38933555e+01]])
fun: 28422.787858961932
message: 'Optimization terminated successfully.'
nfev: 1198
nit: 16
status: 0
success: True
x: array([1.08686633e-04, 6.25513198e-03, 1.75989179e+03])

def prediction(params,current_time, time_range):
    p,q,m = params
    N_t = []
    for t in range(current_time,current_time+time_range):
        current=m*(1-exp(-(p+q)*t))/(1+(q/p)*exp(-(p+q)*t))
        last=m*(1-exp(-(p+q)*(t-1)))/(1+(q/p)*exp(-(p+q)*(t-1)))
        N_t.append(current-last)
    return N_t
```

len(death_by_state)

58

Appendix—Bass Model State Level

```
final_prediction_df = pd.DataFrame()
for state in range(0,58):
    test = pd.DataFrame(death_by_state.iloc[state])
    current_state = test.columns[0]
    ##### WARNING: is the data already counting cumulative or is it daily cases?
    test.columns=['cum_death']
    test['new_death']=test['cum_death'].diff().fillna(test['cum_death'])
    initial_value = [0.027,0.03,100]
    bnds = [(0, None), (0, None), (10, None)]
    res = minimize(error,initial_value, args = (1028,test.new_death), bounds = bnds, tol=1e-8, method = 'Powell')
    print(res.x)
    final_prediction_df[current_state] = prediction(res.x,1018,30)

[6.85348468e-05 7.45512874e-02 7.15048376e+02]
[1.95890521e-06 1.32754303e-02 1.19923394e+03]
[4.74359040e-05 5.64467686e-03 6.25494210e+01]
[4.93288970e-04 4.34575555e-03 3.43540333e+04]
[1.73888509e-04 1.01622182e-01 4.71895508e+01]
[4.80542406e-06 2.29601548e-02 6.39887423e+04]
[2.93835533e-03 2.68151879e-02 2.15148859e+03]
[2.12348097e-03 2.85751395e-02 5.42844123e+03]
[1.71509248e-03 2.69908748e-02 7.61501051e+02]
[1.51278758e-17 5.02591791e-09 1.00000000e+01]
[1.68958999e-03 2.78090743e-02 7.32408180e+02]
```

```
final_prediction_df.sum(axis = 1)
0    276.415553
1    275.252001
2    274.892548
3    272.937188
4    271.785914
5    270.638720
6    269.495599
7    268.356543
8    267.221547
9    266.099604
10   264.963706
11   263.848846
12   262.722018
13   261.607214
14   260.496427
15   259.389649
16   258.286874
17   257.188094
18   256.093301
19   255.002488
20   253.915648
21   252.832772
22   251.753853
23   250.678883
24   249.607854
25   248.540759
26   247.477589
27   246.418336
28   245.362993
29   244.311552
dtype: float64
10      0.0 0.112722 0.066762 0.007445 0.0 0.000320 2.728484e- 3.728928e- 4
weekly_df = death_by_state[death_by_state.columns[:, :7]]
weekly_df.head()

1/22/20 1/29/20 2/5/20 2/12/20 2/19/20 2/26/20 3/4/20 3/11/20 3/18/20
Province_State
Alabama      0 0 0 0 0 0 0 0 0
Alaska        0 0 0 0 0 0 0 0 0
American Samoa 0 0 0 0 0 0 0 0 0
Arizona       0 0 0 0 0 0 0 0 0
Arkansas      0 0 0 0 0 0 0 0 0
5 rows x 150 columns
```

Appendix—Bass Model State Level

```

weekly_final_prediction_df = pd.DataFrame()
for state in range(0,58):
    test = pd.DataFrame(weekly_df.iloc[state])
    current_state = test.columns[0]
    ##### WARNING: is the data already counting cumulative or is it daily cases?
    test.columns=['cum_death']
    test['new_death']=test['cum_death'].diff().fillna(test['cum_death'])

initial_value = [0.027,0.03,500]
bnds = [(0, None), (0, None), (0, None)]
res = minimize(error,initial_value, args = (149,test.new_death), bounds = bnds, tol=1e-8, method = 'Powell')
print(res.x)
weekly_final_prediction_df[current_state] = prediction(res.x,147,4)

[2.53662352e-03 3.92312856e-02 2.17136281e+04]
[4.52641064e-04 5.09960391e-02 1.47285756e+03]
[3.57395004e-03 1.33183661e-02 4.48146283e+01]
[3.43372743e-03 3.05269579e-02 3.43702857e+04]
[2.66275774e-03 3.12050306e-02 1.38800931e+04]
[2.06263090e-04 1.21178910e-01 7.24063712e+04]

weekly_final_prediction_df

```

	Alabama	Alaska	American Samoa	Arizona	Arkansas	California	Colorado	Connecticut	Deli:
0	30.737555	4.084786	0.174536	70.778633	35.755428	0.097930	1.543645	0.0	0.2:
1	29.559719	3.902083	0.172993	68.680416	34.739994	0.086736	1.494937	0.0	0.2:
2	28.423980	3.726545	0.171447	66.636268	33.748125	0.076822	1.447766	0.0	0.2:
3	27.329070	3.557984	0.169897	64.645307	32.779600	0.068041	1.402083	0.0	0.2:

4 rows × 58 columns

```
weekly_final_prediction_df.sum(axis=1)
```

```

0    1966.177380
1    1908.668365
2    1852.592750
3    1797.931651
dtype: float64

```

```

monthly_df = death_by_state[death_by_state.columns[::30]]
monthly_df.head()

```

	1/22/20	2/21/20	3/22/20	4/21/20	5/21/20	6/20/20	7/20/20	8/19/20	9/18/
--	---------	---------	---------	---------	---------	---------	---------	---------	-------

Province_State	Alabama	0	0	0	183	529	838	1291	1944	24
Alaska	0	0	0	9	10	12	18	29		
American Samoa	0	0	0	0	0	0	0	0	0	
Arizona	0	0	2	208	764	1346	2784	4634	54	
Arkansas	0	0	0	42	110	224	363	631	11	

5 rows × 35 columns

```

monthly_final_prediction_df = pd.DataFrame()
for state in range(0,58):
    test = pd.DataFrame(monthly_df.iloc[state])
    current_state = test.columns[0]
    ##### WARNING: is the data already counting cumulative or is it daily cases?
    test.columns=['cum_death']
    test['new_death']=test['cum_death'].diff().fillna(test['cum_death'])
    initial_value = [0.027,0.03,500]
    bnds = [(0, None), (0, None), (0, None)]
    res = minimize(error,initial_value, args = (34,test.new_death), bounds = bnds, tol=1e-8, method = 'Powell')
    print(res.x)
    monthly_final_prediction_df[current_state] = prediction(res.x,33,2)

```

```

[1.09724254e-02 1.66298404e-01 2.17996097e+04]
[ 0.82684564 1.71448828 14.37953792]
[1.1994817e-03 1.68337786e-01 7.01835424e+01]
[1.48915765e-02 1.27672369e-01 3.46636125e+04]
[1.14451831e-02 1.34034738e-01 1.38623322e+04]
[5.35237879e-02 1.04530323e+00 7.39875900e+03]
[1.20972417e-01 9.13257263e-18 6.07894933e+03]
[4.10707613e-02 1.01075066e+00 6.31964760e+03]
[1.34604120e-08 2.12809506e+01 1.14047808e+02]

```

Appendix--SIRD

```
import pandas as pd

url = "https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_US.csv"
df = pd.read_csv(url,index_col=0)
df.head()

C iso2 iso3 code3 FIPS Admin2 Province_State Country_Region Lat
UID
84001001 US USA 840 1001.0 Autauga Alabama US 32.539527 -86.6
84001003 US USA 840 1003.0 Baldwin Alabama US 30.727750 -87.7
84001005 US USA 840 1005.0 Barbour Alabama US 31.868263 -85.3
84001007 US USA 840 1007.0 Bibb Alabama US 32.996421 -87.1
84001009 US USA 840 1009.0 Blount Alabama US 33.982109 -86.5

5 rows x 1064 columns
<   >
```

```
daily_death = pd.DataFrame(df[df.columns[11:]].sum())
daily_death.columns=['cum_death']
daily_death['new_death']=daily_death['cum_death'].diff().fillna(daily_death['cum_death'])

df_cases = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_US.csv")
df_cases.head()
```

```
import pandas as pd

url = "https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_US.csv"
df = pd.read_csv(url,index_col=0)
df.head()

C iso2 iso3 code3 FIPS Admin2 Province_State Country_Region Lat
UID
84001001 US USA 840 1001.0 Autauga Alabama US 32.539527 -86.6
84001003 US USA 840 1003.0 Baldwin Alabama US 30.727750 -87.7
84001005 US USA 840 1005.0 Barbour Alabama US 31.868263 -85.3
84001007 US USA 840 1007.0 Bibb Alabama US 32.996421 -87.1
84001009 US USA 840 1009.0 Blount Alabama US 33.982109 -86.5

5 rows x 1064 columns
<   >
```

```
daily_death = pd.DataFrame(df[df.columns[11:]].sum())
daily_death.columns=['cum_death']
daily_death['new_death']=daily_death['cum_death'].diff().fillna(daily_death['cum_death'])

df_cases = pd.read_csv("https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_US.csv")
df_cases.head()
```

Appendix--SIRD

```
daily_cases = pd.DataFrame(df_cases[df_cases.columns[11:]].sum())
daily_cases.columns=['cum_cases']
daily_cases['new_cases']=daily_cases['cum_cases'].diff().fillna(daily_cases['cum_cases'])
daily_cases.head(10)
```

	cum_cases	new_cases
1/23/20	1	1.0
1/24/20	2	1.0
1/25/20	2	0.0
1/26/20	5	3.0
1/27/20	5	0.0
1/28/20	5	0.0
1/29/20	6	1.0
1/30/20	6	0.0
1/31/20	8	2.0
2/1/20	8	0.0

```
import numpy as np
from scipy.optimize import minimize, curve_fit
from scipy import integrate, optimize
import pandas as pd
from scipy.integrate import odeint
import matplotlib.pyplot as plt

def model(params, time_step):
    beta, gamma, S0, I0, d = params
    N = S0 + I0
    S = [S0]
    I = [I0]
    R = [0]
    D = [0]
    dS = []
    dI = []
    dR = []
    dD = []
    for t in range(0,time_step):
        dS.append(-beta*I[t]*S[t]/N)
        dI.append(beta*I[t]*S[t]/N- gamma*I[t] - d*I[t])
        dR.append(gamma*I[t])
        dD.append(d*I[t])
        S.append(S[t] + dS[t])
        I.append(I[t] + dI[t])
        R.append(R[t] + dR[t])
        D.append(D[t] + dD[t])
    return I, D

def error(params, time_step, actual_I, actual_D):
    I_pred, D_pred = model(params, time_step)
    beta, gamma, S0, I0, d = params
    I_error = 0
    D_error = 0
    for t in range(0,time_step):
        I_error+=abs(I_pred[t]-actual_I[t])
        D_error+=abs(D_pred[t]-actual_D[t])
    total_error = I_error + D_error
    return total_error
```

Appendix--SIRD

```

actual_cum_cases = daily_cases["cum_cases"]
actual_cum_death = daily_death["cum_death"]
actual_cum_case = [i/10000000 for i in actual_cum_cases.values.tolist()]
actual_cum_death = [i/10000000 for i in actual_cum_death.values.tolist()]

initial_value = [8.63e-1,1.53e-4,18.3,6.78e-5,8.6e-1]
bnds = [(0, 1), (0, 1), (0, 50), (0, 1), (0, 1)]
res = minimize(error,initial_value, args = (1028, actual_cum_cases, actual_cum_death), bounds = bnds, tol=1e-7, method = 'Nelder-Mead', options={'maxiter':1000})
print(res)

final_simplex: (array([[8.63201345e-01, 2.48787398e-04, 1.26890726e+01, 6.77213074e-05,
   8.59425852e-01],
 [8.63201345e-01, 2.48787398e-04, 1.26890726e+01, 6.77213074e-05,
  8.59425852e-01],
 [8.63201345e-01, 2.48787397e-04, 1.26890726e+01, 6.77213073e-05,
  8.59425852e-01],
 [8.63201345e-01, 2.48787395e-04, 1.26890727e+01, 6.77213075e-05,
  8.59425852e-01],
 [8.63201345e-01, 2.48787395e-04, 1.26890727e+01, 6.77213075e-05,
  8.59425852e-01],
 [8.63201345e-01, 2.48787397e-04, 1.26890726e+01, 6.77213074e-05,
  8.59425852e-01],
 [8.63201345e-01, 2.48787398e-04, 1.26890726e+01, 6.77213074e-05,
  8.59425852e-01]]), array([4.22762574e+10, 4.22762574e+10, 4.22762574e+10, 4.22762574e+10,
 4.22762574e+10, 4.22762574e+10]))  

fun: 42276257373.543175  

message: 'Optimization terminated successfully.'  

nfev: 361  

nit: 157  

status: 0  

success: True

```

```

x: array([8.63201345e-01, 2.48787398e-04, 1.26890726e+01, 6.77213074e-05,
 8.59425852e-01])

from scipy.optimize import brute

grid = ((8.63e-1,8.63e-1),(1.53e-4,1.53e-4),slice(1,40, 0.1),(6.78e-5, 6.78e-5),(0.86,0.86))
resbrute = brute(error, grid, args = (1028, actual_cum_cases, actual_cum_death), Ns = 1,full_output=True, disp=True, finish=None)
print("Best coeffs: {}".format(resbrute[0]))
print("Score with best coeffs: {}".format(resbrute[1]))
print("Grid: {}".format(resbrute[2].tolist()))
print("Scores for grid: {}".format(resbrute[3].tolist()))

Best coeffs: [8.63e-01 1.53e-04 1.83e+01 6.78e-05 8.60e-01]
Score with best coeffs: 42276257374.381165
Grid: [[[0.863], [[0.863]], [[0.863]], [[0.863]], [[0.863]], [[0.863]], [[0.863]], [[0.863]], [[0.863]], [[0.863]]]
Scores for grid: [[[42276257416.695599], [[42276257415.98136]], [[42276257415.295685]], [[42276257414.63543]], [[42276257413.99786]], [[42276257412.543175]]]]  

>   

params = res.x
#params = np.array([0.022,0.00165,33,1e-7,7.5e-6])

```

Appendix--SIRD

```
# predictions
I, D = model(params, 1450)
D = pd.DataFrame(D)
D.columns = ['cum_death']
D['new_death'] = D['cum_death'].diff().fillna(D['cum_death'])
new_death = [i*1000000 for i in D['new_death']]
print(new_death[1018:])

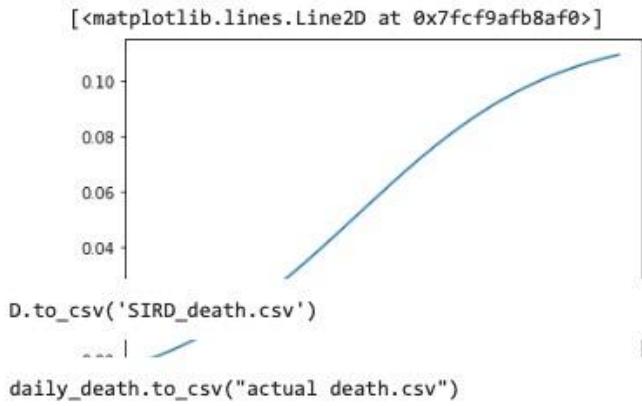
plt.plot(new_death)

[<matplotlib.lines.Line2D at 0x7fcf9aee9e80>
 0 200 400 600 800 1000 1200 1400
 0 600 1200 1400
 0 200 400 600 800 1000 1200 1400]
<matplotlib.lines.Line2D at 0x7fcf9aee9e80>

plt.plot(actual_cum_death[1018:])

[<matplotlib.lines.Line2D at 0x7fcf9aaff550>
 0 5 10 15 20 25 30 35
 0.1074 0.1076 0.1078 0.1080 0.1082 0.1084
 0.1074 0.1076 0.1078 0.1080 0.1082 0.1084]

plt.plot(D['cum_death'])
```



```
beta, gamma, S0, I0, d = params
beta * S0/gamma

44026.44433003097
```

Appendix--GBM

t*	Ahat(t)	Ahat(t-1)	Nhat(t)	SE(t)
A2	\$K\$3*(1-EXP(-(\$K\$1+\$K\$2)*D2))/(1+(\$K\$2/\$K\$1)*EXP(-(\$K\$1+\$K\$2)*D2))	\$K\$3*(1-EXP(-(\$K\$1+\$K\$2)*(D2-1)))/(1+(\$K\$2/\$K\$1)*EXP(-(\$K\$1+\$K\$2)*(D2-1)))	E2-F2	(B2-G2)^2

t	Death	Vaccine Administered	t*	Ahat(t)	Ahat(t-1)	Nhat(t)	SE(t)	Date	p	0.000652209
1.00	1.00	0.00	1.0000	437.8806478	0	437.8806478	190864.7004	2020/1/22	q	0.007680922
2.00	0.00	0.00	2.0000	878.8454359	437.8806478	440.9647881	194449.9444	2020/1/23	M	669025.1021
3.00	0.00	0.00	3.0000	1322.911591	878.8454359	444.0661547	197194.7498	2020/1/24	b	-34.63850985
4.00	0.00	0.00	4.0000	1770.096365	1322.911591	447.1847739	199974.222	2020/1/25		
5.00	0.00	0.00	5.0000	2220.417035	1770.096365	450.3206708	202788.7066	2020/1/26		
6.00	0.00	0.00	6.0000	2673.890905	2220.417035	453.4738693	205638.5501	2020/1/27		
7.00	0.00	0.00	7.0000	3130.535297	2673.890905	456.644392	208524.1007	2020/1/28		
8.00	0.00	0.00	8.0000	3590.367557	3130.535297	459.8322603	211445.7076	2020/1/29		
9.00	0.00	0.00	9.0000	4053.405051	3590.367557	463.0374943	214403.7211	2020/1/30	762167676	
10.00	0.00	0.00	10.0000	4519.665164	4053.405051	466.2601129	217398.4929	2020/1/31		
11.00	0.00	0.00	11.0000	4989.165298	4519.665164	469.5001335	220430.3753	2020/2/1		
12.00	0.00	0.00	12.0000	5461.92287	4989.165298	472.7575722	223499.7221	2020/2/2		
13.00	0.00	0.00	13.0000	5937.955314	5461.92287	476.0324439	226606.8876	2020/2/3		
14.00	0.00	0.00	14.0000	6417.280076	5937.955314	479.3247619	229752.2273	2020/2/4		
15.00	0.00	0.00	15.0000	6899.914614	6417.280076	482.6345381	232936.0974	2020/2/5		

Appendix--GBM

1	550.00	400.00	220300190.00	657.5050	029154.5541	020019.2221	315.512055	1112.051052	2022/10/1			
2	991.00	50.00	226590460.00	638.3050	629446.073	629133.0803	312.9927096	69165.16533	2022/10/8			
3	992.00	9.00	226607620.00	639.3023	629755.9387	629445.2551	310.6836076		2022/10/9	9.00		
4	993.00	134.00	226660597.00	640.2943	630061.8342	629753.4323	308.4019033		2022/10/10	134.00		
5	994.00	608.00	226714380.00	641.2860	630365.4442	630059.3091	306.1350996		2022/10/11	608.00		
6	995.00	769.00	226772393.00	642.2772	630666.6254	630362.7411	303.8843143		2022/10/12	769.00		9162.91599
7	996.00	659.00	226827925.00	643.2687	630965.7047	630664.0576	301.6471524		2022/10/13	659.00		
8	997.00	262.00	226891622.00	644.2590	631262.209	630962.7818	299.4272021		2022/10/14	262.00		
9	998.00	41.00	226919762.00	645.2547	631558.1371	631260.9275	297.2095308		2022/10/15	41.00		10787.00
10	999.00	9.00	226935897.00	646.2522	631852.4101	631557.4078	295.002246		2022/10/16	9.00		
11	1000.00	335.00	226989536.00	647.2440	632142.826	631850.004	292.821921		2022/10/17	335.00		
12	1001.00	434.00	227041882.00	648.2360	632431.1512	632140.4958	290.6553544		2022/10/18	434.00		
13	1002.00	757.00	227097267.00	649.2276	632717.2086	632428.7046	288.5039217		2022/10/19	757.00		
14	1003.00	557.00	227150610.00	650.2194	633001.2361	632714.8702	286.3658754		2022/10/20	557.00		
15	1004.00	577.00	227210823.00	651.2103	633282.8609	632998.6168	284.2440664		2022/10/21	577.00		
16	1005.00	22.00	227238204.00	652.2061	633563.8044	633281.6788	282.1255548		2022/10/22	22.00		
17	1006.00	8.00	227253469.00	653.2038	633843.1676	633563.1504	280.0171431		2022/10/23	8.00		
18	1007.00	204.00	227304956.00	654.1959	634118.9127	633840.9785	277.9342595		2022/10/24	204.00		
19	1008.00	429.00	227356976.00	655.1880	634392.5833	634116.718	275.8653014		2022/10/25	429.00		
20	1009.00	1127.00	227413256.00	656.1794	634664.0384	634390.227	273.8113738		2022/10/26	1127.00		
21	1010.00	549.00	227465632.00	657.1714	634933.6322	634661.8624	271.7698352		2022/10/27	549.00		
22	1011.00	187.00	227524078.00	658.1625	635200.9665	634931.2228	269.743741		2022/10/28	187.00		
23	1012.00	33.00	227549600.00	659.1587	635467.6434	635199.9224	267.7209755		2022/10/29	33.00		
24	1013.00	5.00	227562998.00	660.1566	635732.8075	635467.0995	265.7080471		2022/10/30	5.00		
25	1014.00	127.00	227600597.00	661.1509	635995.0087	635731.2927	263.7160055		2022/10/31	127.00		