Supplementary Material: Analysis of COVID-19 Vaccination Trends Between 2021 and 2023

1 Overview

Analysis of COVID-19 Vaccination Trends: Distribution and Administration

This document provides supplementary material for the manuscript "Analysis of COVID-19 Vaccination Trends: Distribution and Administration". It includes additional details on the methods used in the analysis, as well as additional results that were not included in the main manuscript.

2 Code and file information

- "Cann-MADA-project.Rproj": Establishes relative file paths for project
- "README.md": Provides brief order of scripts for reproducing and summarizes the folders within the project
- "code" folder: Contains all code for processing, exploratory data analysis, and modeling analysis
  - "processing-code" subfolder:
    - "processing.qmd": Contains code for processing the raw data into the processed data
  - "eda-code" subfolder:
    - "eda.qmd": Contains code for exploratory data analysis
  - "analysis-code" subfolder:
    - "analysis.qmd": Contains code for modeling analysis
- "data" folder:

  - "raw-data" subfolder: Contains the raw COVID-19 Vaccine data
  - "processed-data" subfolder: Contains the processed data used in the analysis

- "results" folder: Contains all results from the analysis
  - "figures" subfolder: Contains all figures generated from eda and analysis
  - "tables" subfolder: Contains all tables generated from eda and analysis

- "assets" folder:
  - Contains workflow schematic image
  - Contains the CDC U.S. Regions image
  - Contains american journal of epidemiology reference style and vancouver reference style (.csl files)
  - "references" subfolder:
    * "project-citations.bib": Contains the references used in the manuscript

- "products" folder:
  - "manuscript" subfolder: Contains manuscript.qmd file to create project manuscript
    * "supplement" subfolder: Contains this file and the supplementary figures and tables

3 Reproducing Results

Reproducing this project requires R, RStudio, and Microsoft Word. Files should be run in the following order.

1) In the code > processing-code folder: processing.qmd
2) In the code > eda-code folder: exploratoryanalysis.qmd
3) In the code > analysis-code folder: analysis.Rmd
4) In the products > manuscript folder: manuscript.qmd
5) In the products > manuscript > supplement folder: Supplementary-Material.qmd

4 Supplementary Results

Table one displays a summary of each variable of the COVID-19 vaccine dataset.

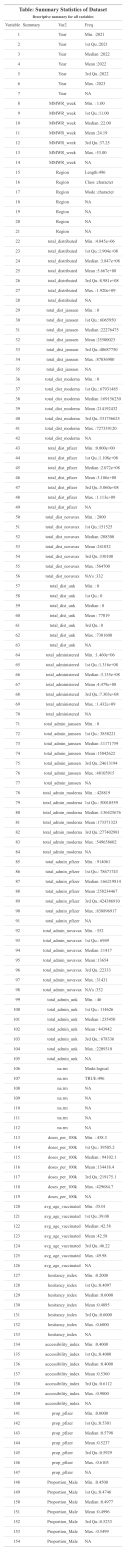

Table: Summary Statistics of Dataset

Descriptive summary for all variables

Variable	Summary	Var2	Freq
1		Year	Min. :2021
2		Year	1st Qu.:2021
3		Year	Median :2022
4		Year	Mean :2022
5		Year	3rd Qu.:2022
6		Year	Max. :2023
7		Year	NA
8		MMWR_week	Min. : 1.00
9		MMWR_week	1st Qu.:11.00
10		MMWR_week	Median :22.00
11		MMWR_week	Mean :24.19
12		MMWR_week	3rd Qu.:37.33
13		MMWR_week	Max. :53.00
14		MMWR_week	NA
15		Region	Length:496
16		Region	Class :character
17		Region	Mode :character
18		Region	NA
19		Region	NA
20		Region	NA
21		Region	NA
22		total_distributed	Min. :4.845e+06
23		total_distributed	1st Qu.:1.994e+08
24		total_distributed	Median :2.847e+08
25		total_distributed	Mean :5.667e+08
26		total_distributed	3rd Qu.:8.981e+08
27		total_distributed	Max. :1.820e+09
28		total_distributed	NA
29		total_dist_janssen	Min. : 0
30		total_dist_janssen	1st Qu.: 6665850
31		total_dist_janssen	Median :22276475
32		total_dist_janssen	Mean :35500025
33		total_dist_janssen	3rd Qu.:48005790
34		total_dist_janssen	Max. :97836900
35		total_dist_janssen	NA
36		total_dist_moderna	Min. : 0
37		total_dist_moderna	1st Qu.: 6755140
38		total_dist_moderna	Median :18938210
39		total_dist_moderna	Mean :31491032
40		total_dist_moderna	3rd Qu.:51778025
41		total_dist_moderna	Max. :72733120
42		total_dist_moderna	NA
43		total_dist_pfizer	Min. : 0.000e+00
44		total_dist_pfizer	1st Qu.: 1.108e+08
45		total_dist_pfizer	Median :2.072e+08
46		total_dist_pfizer	Mean :3.186e+08
47		total_dist_pfizer	3rd Qu.:5.069e+08
48		total_dist_pfizer	Max. :1.113e+09
49		total_dist_pfizer	NA
50		total_dist_novavax	Min. : 2800
51		total_dist_novavax	1st Qu.: 151525
52		total_dist_novavax	Median :280388
53		total_dist_novavax	Mean :310182
54		total_dist_novavax	3rd Qu.:438108
55		total_dist_novavax	Max. :764500
56		total_dist_novavax	NA's :332
57		total_dist_unk	Min. : 0
58		total_dist_unk	1st Qu.: 0
59		total_dist_unk	Median : 0
60		total_dist_unk	Mean : 77819
61		total_dist_unk	3rd Qu.: 0
62		total_dist_unk	Max. :7381600
63		total_dist_unk	NA
64		total_administered	Min. : 1.400e+05
65		total_administered	1st Qu.:1.131e+08
66		total_administered	Median :1.131e+08
67		total_administered	Mean :4.479e+08
68		total_administered	3rd Qu.:7.303e+08
69		total_administered	Max. :1.432e+09
70		total_administered	NA
71		total_admin_janssen	Min. : 0
72		total_admin_janssen	1st Qu.: 3858221
73		total_admin_janssen	Median :13177799
74		total_admin_janssen	Mean :15845632
75		total_admin_janssen	3rd Qu.:24611194
76		total_admin_janssen	Max. :48405653
77		total_admin_janssen	NA
78		total_admin_moderna	Min. : 42603
79		total_admin_moderna	1st Qu.: 8681859
80		total_admin_moderna	Median :136429076
81		total_admin_moderna	Mean :175971323
82		total_admin_moderna	3rd Qu.:277462981
83		total_admin_moderna	Max. :548634842
84		total_admin_moderna	NA
85		total_admin_pfizer	Min. : 91404
86		total_admin_pfizer	1st Qu.: 78657703
87		total_admin_pfizer	Median :186239814
88		total_admin_pfizer	Mean :258234467
89		total_admin_pfizer	3rd Qu.:424168810
90		total_admin_pfizer	Max. :930956917
91		total_admin_pfizer	NA
92		total_admin_novavax	Min. : 552
93		total_admin_novavax	1st Qu.: 6949
94		total_admin_novavax	Median :11417
95		total_admin_novavax	Mean :13454
96		total_admin_novavax	3rd Qu.:22333
97		total_admin_novavax	Max. :31431
98		total_admin_novavax	NA's :332
99		total_admin_unk	Min. : 46
100		total_admin_unk	1st Qu.: 116626
101		total_admin_unk	Median : 215458
102		total_admin_unk	Mean : 642942
103		total_admin_unk	3rd Qu.: 678334
104		total_admin_unk	Max. :2289518
105		total_admin_unk	NA
106		na.m	Mode:logical
107		na.m	TRUE:496
108		na.m	NA
109		na.m	NA
110		na.m	NA
111		na.m	NA
112		na.m	NA
113		doses_per_100k	Min. : 458.1
114		doses_per_100k	1st Qu.: 85985.2
115		doses_per_100k	Median : 94102.1
116		doses_per_100k	Mean :134414.4
117		doses_per_100k	3rd Qu.:219171.1
118		doses_per_100k	Max. :429654.7
119		doses_per_100k	NA
120		avg_age_vaccinated	Min. :0.01
121		avg_age_vaccinated	1st Qu.:39.08
122		avg_age_vaccinated	Median :42.38
123		avg_age_vaccinated	Mean :42.58
124		avg_age_vaccinated	3rd Qu.:46.22
125		avg_age_vaccinated	Max. :49.98
126		avg_age_vaccinated	NA
127		hesitancy_index	Min. :0.2000
128		hesitancy_index	1st Qu.:0.4097
129		hesitancy_index	Median :0.6000
130		hesitancy_index	Mean :0.4895
131		hesitancy_index	3rd Qu.:0.6000
132		hesitancy_index	Max. :0.6000
133		hesitancy_index	NA
134		accessibility_index	Min. :0.4000
135		accessibility_index	1st Qu.:0.4000
136		accessibility_index	Median :0.4000
137		accessibility_index	Mean :0.5360
138		accessibility_index	3rd Qu.:0.6112
139		accessibility_index	Max. :0.9000
140		accessibility_index	NA
141		prop_pfizer	Min. :0.0000
142		prop_pfizer	1st Qu.:0.3381
143		prop_pfizer	Median :0.5758
144		prop_pfizer	Mean :0.5237
145		prop_pfizer	3rd Qu.:0.9919
146		prop_pfizer	Max. :0.9305
147		prop_pfizer	NA
148		Proportion_Male	Min. :0.4500
149		Proportion_Male	1st Qu.:0.4746
150		Proportion_Male	Median :0.4975
151		Proportion_Male	Mean :0.4996
152		Proportion_Male	3rd Qu.:0.5253
153		Proportion_Male	Max. :0.5490
154		Proportion_Male	NA

Figure 1: Table One: Summary Statistics of the Vaccination Data

Figure one displays the correlatins between all variables within the COVID-19 vaccine dataset.

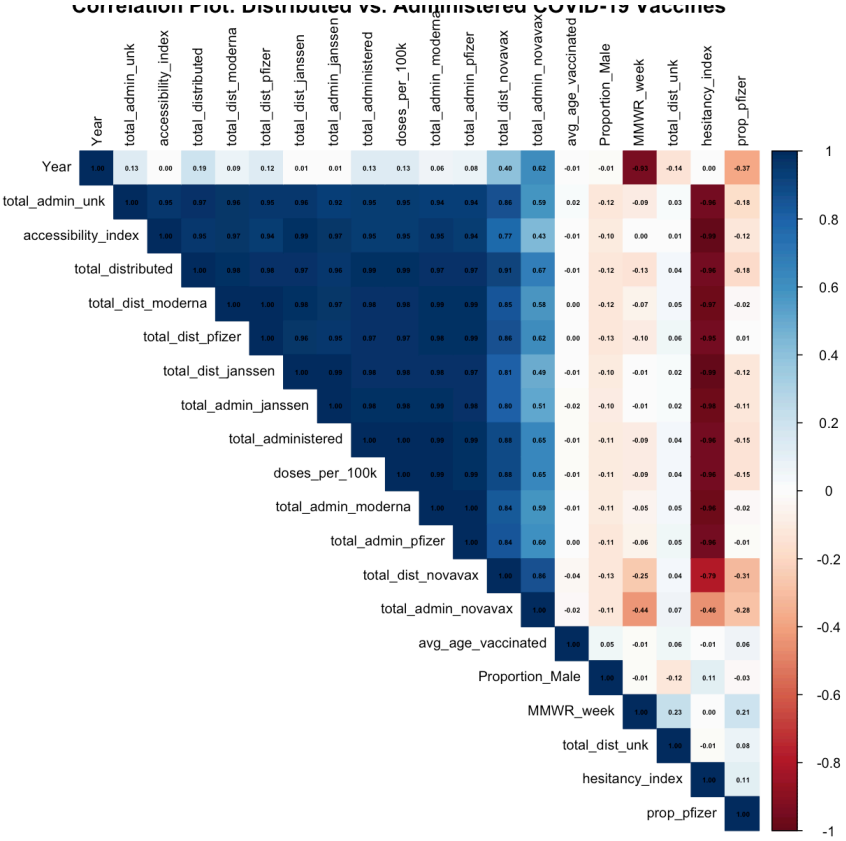

Figure 2: Figure One: Overall Correlation Plot of COVID-19 Vaccine Data

Figure two shows the correlations of specifically the distribution and administration variables in the dataset. The correlation between total_administered and total_distributed is 0.99; the correlation between total_admin_janssen and total_dist_janssen is 0.99; the correlation between total_admin_moderna and total_dist_moderna is 0.99; the correlation between total_admin_pfizer and total_dist_pfizer is 0.99; the correlation between total_admin_novavax and total_dist_novavax is 0.86, which is the lowest of all manufacturers; and the correlation between total_admin_unk and total_dist_unk is 0.03.

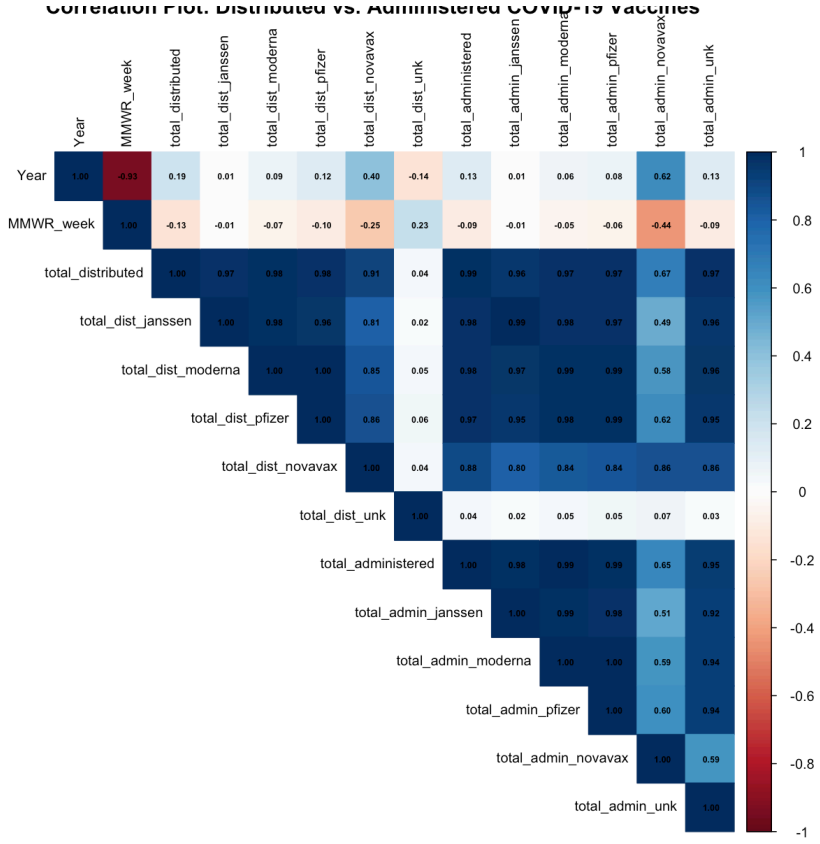

Figure 3: Figure Two: Correlation Plot of Distributed vs. Administrated COVID-19 Vaccine Data

Below, you will several scatterplots. Figure three is an overall scatterplot of the relationship between administered and distributed doses. The points closely follow the diagonal line, indicating a strong relationship between the two variables (confirmed by the correlation coeffecient of 0.99).

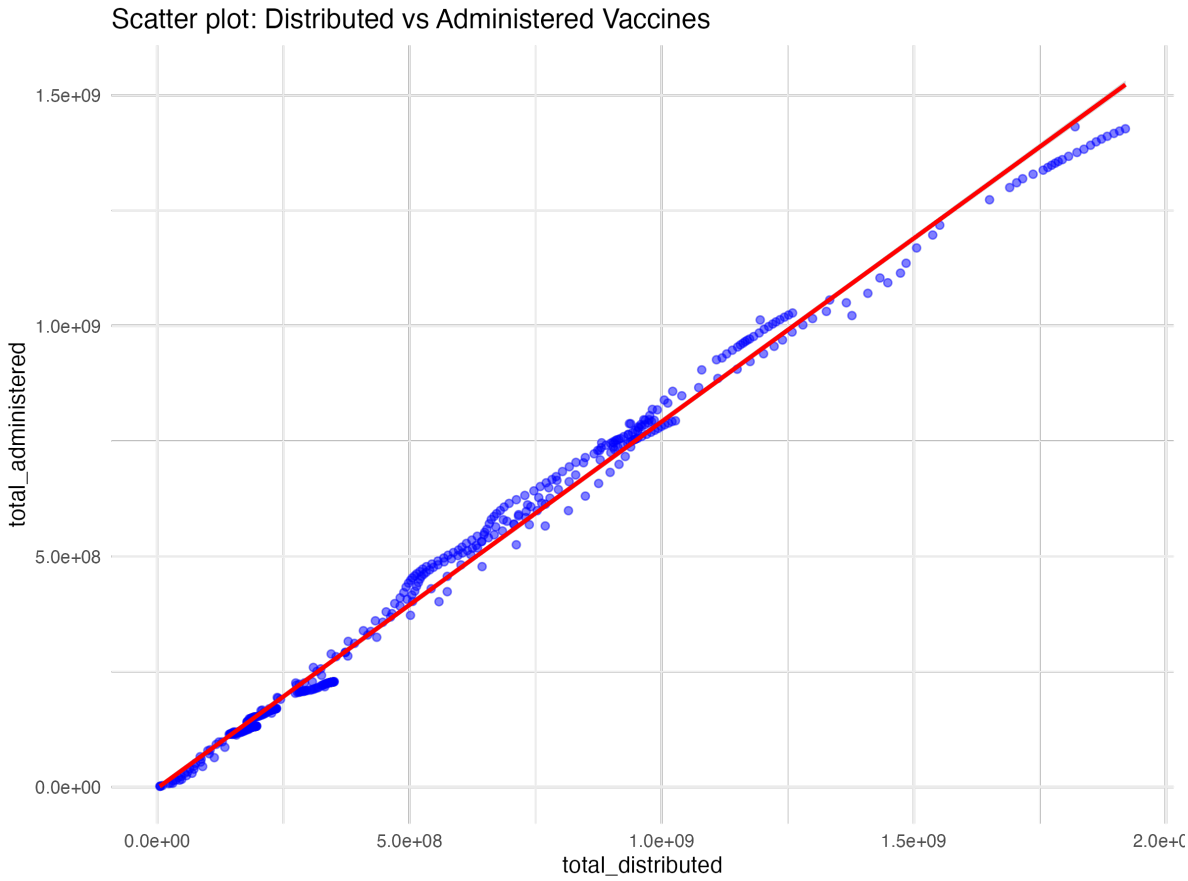

Figure 4: Figure Three: Scatterplot of Distributed vs Administered Doses

Figures 4.1, 4.2, 4.3, and 4.2 show the scatterplots of the relationship between administered and distributed doses for each vaccine manufacturer. It appears as though all regions have a strong positive correlation between the number of vaccines distributed and administered. However, Pfizer and Moderna's points still follow the diagonal line the closest.

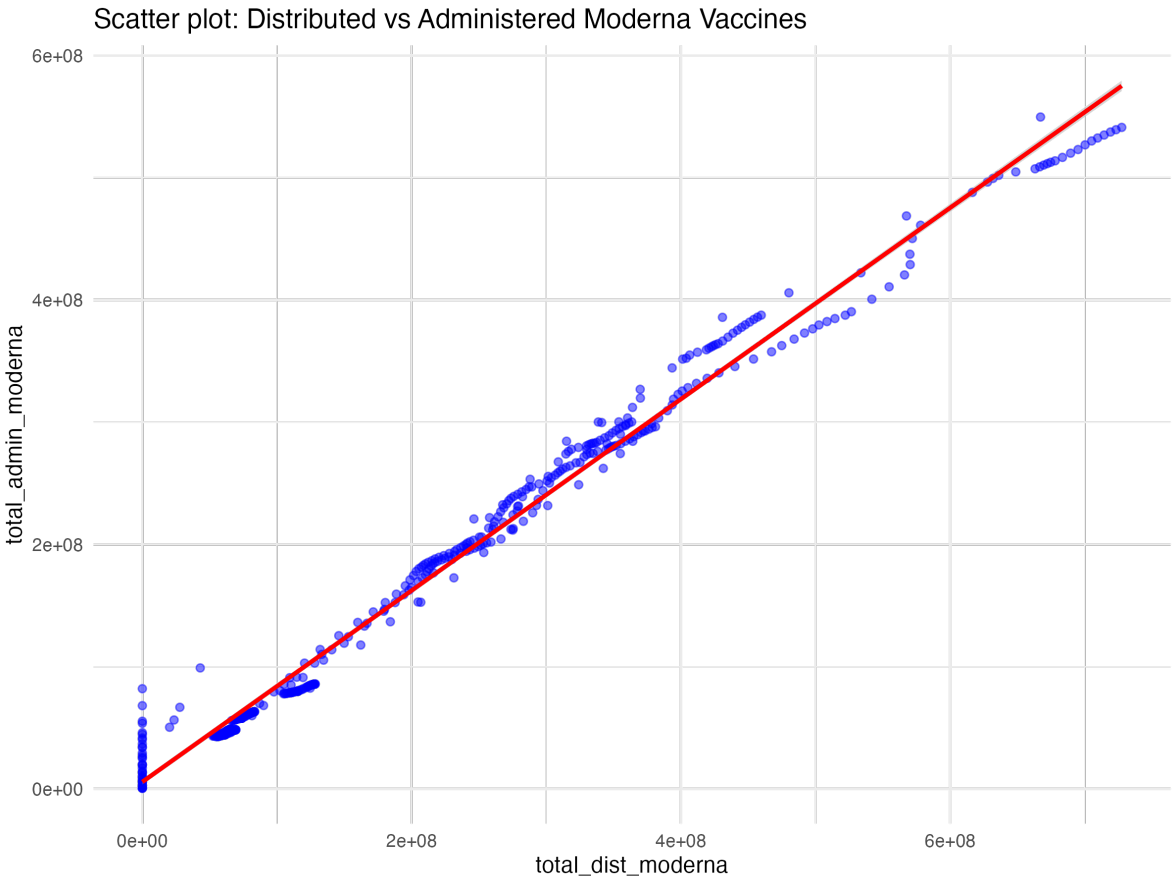

Figure 5: Figure 4.1: Scatterplot of Distributed vs Administered Moderna Doses

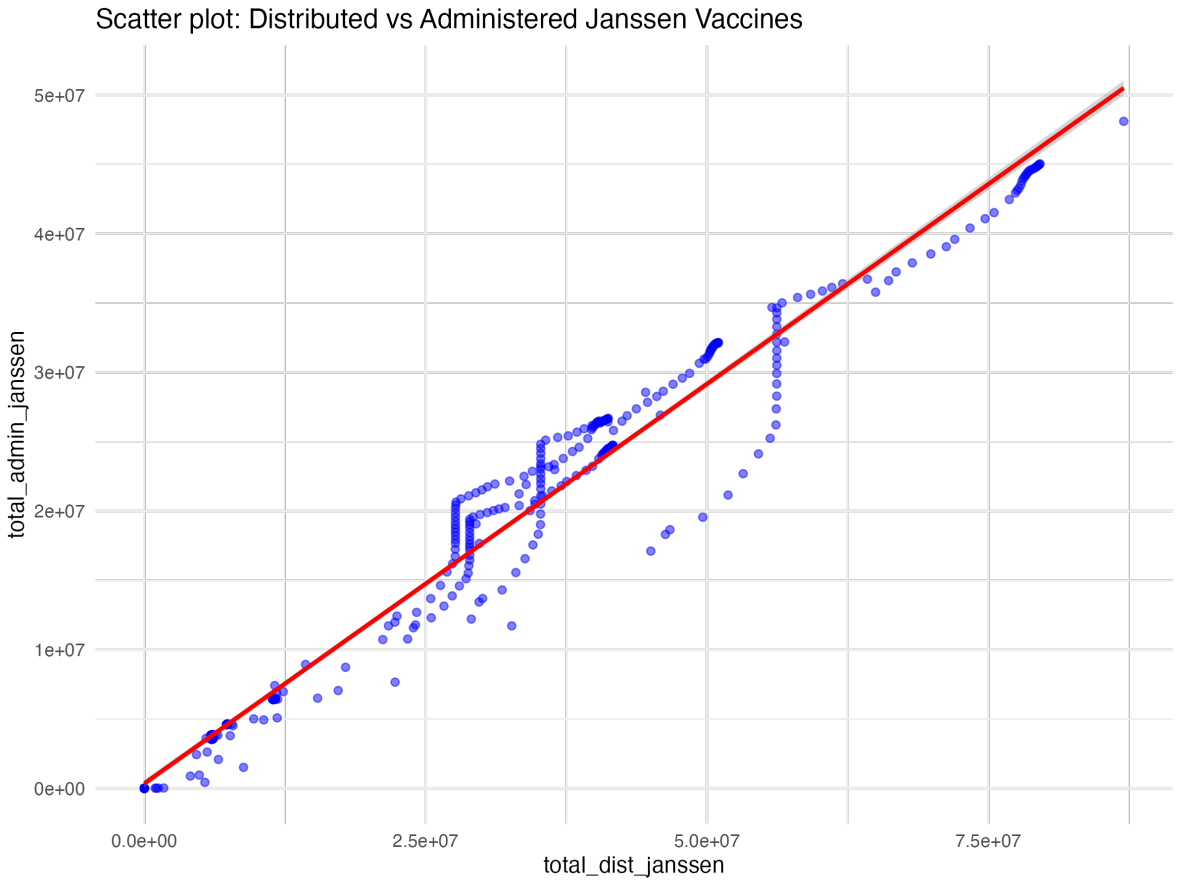

Figure 6: Figure 4.2: Scatterplot of Distributed vs Administered Janssen Doses

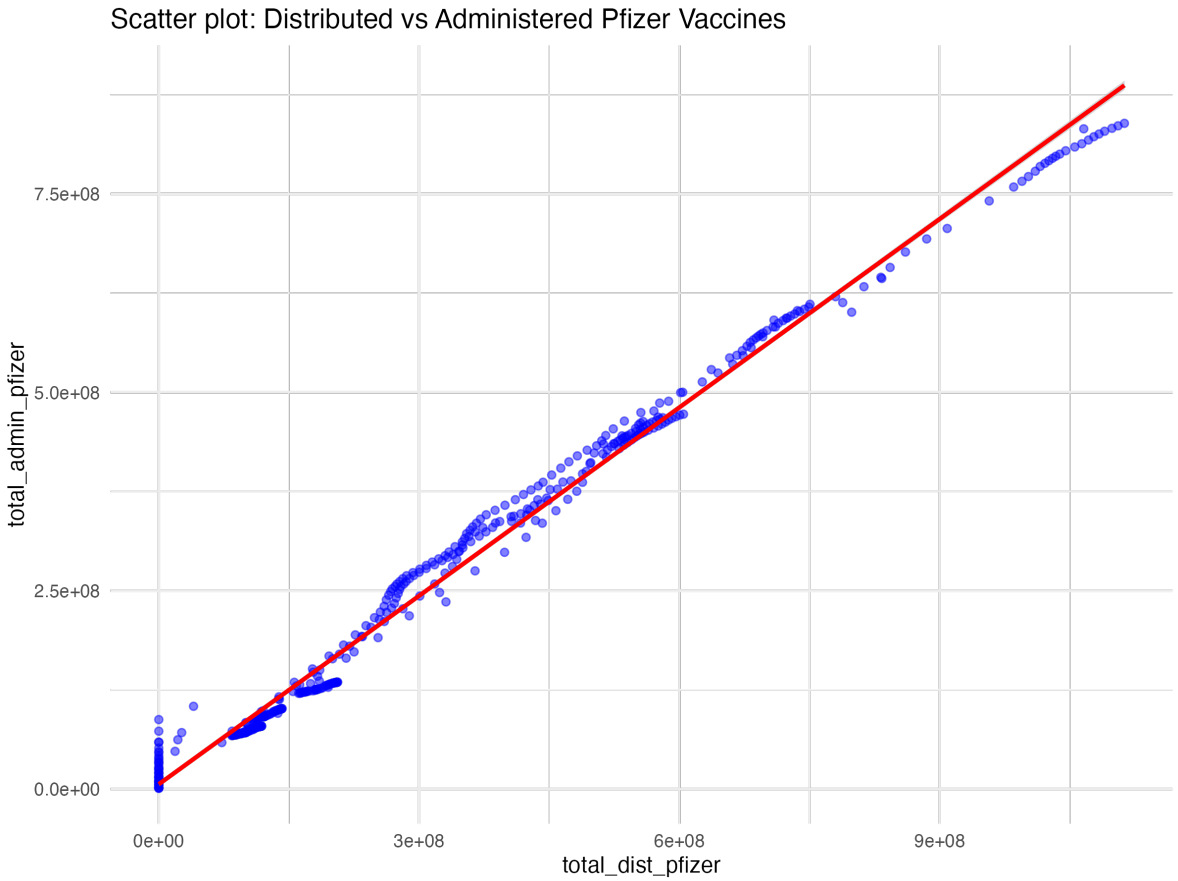

Figure 7: Figure 4.3: Scatterplot of Distributed vs Administered Pfizer Doses

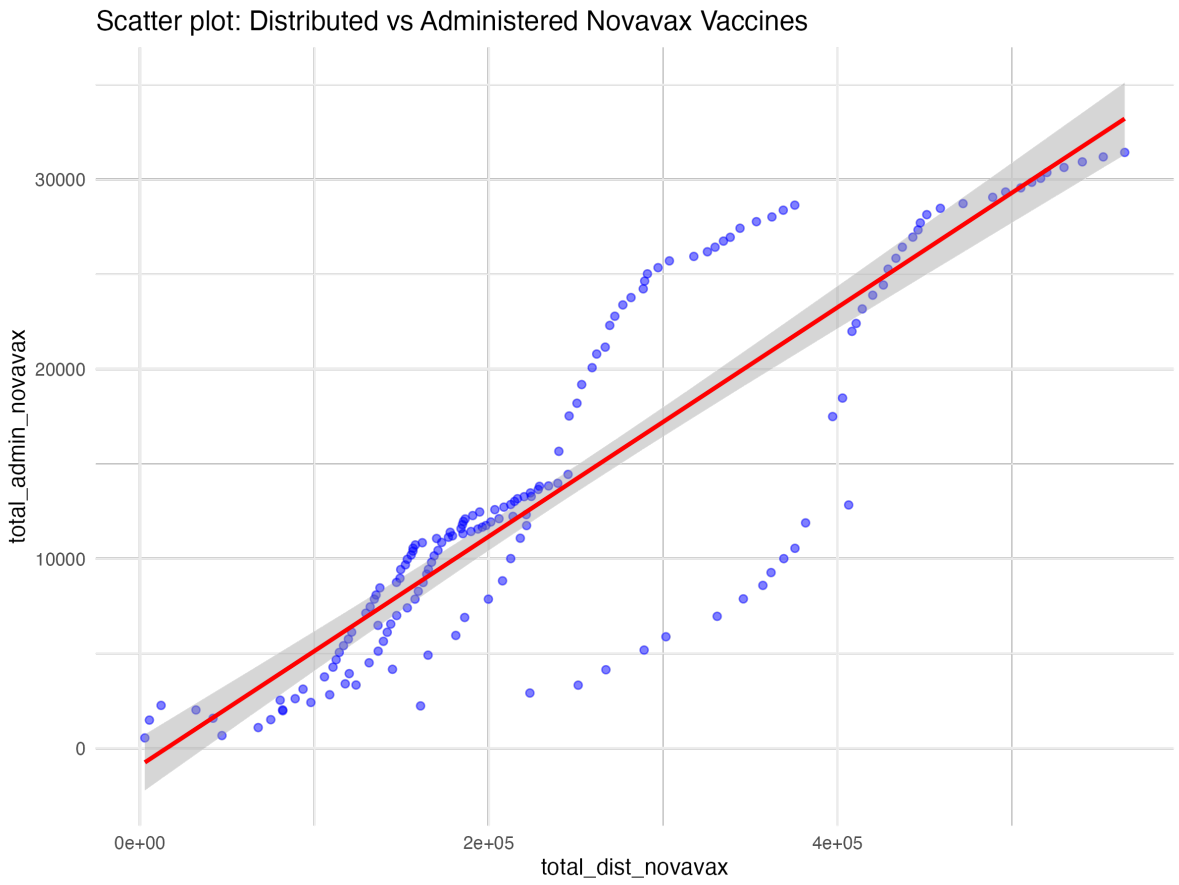

Figure 8: Figure 4.4: Scatterplot of Distributed vs Administered Novavax Doses

Figures 5.1, 5.2, 5.3, and 5.4 show the scatterplots of the relationship between administered and distributed doses for each region of the U.S..

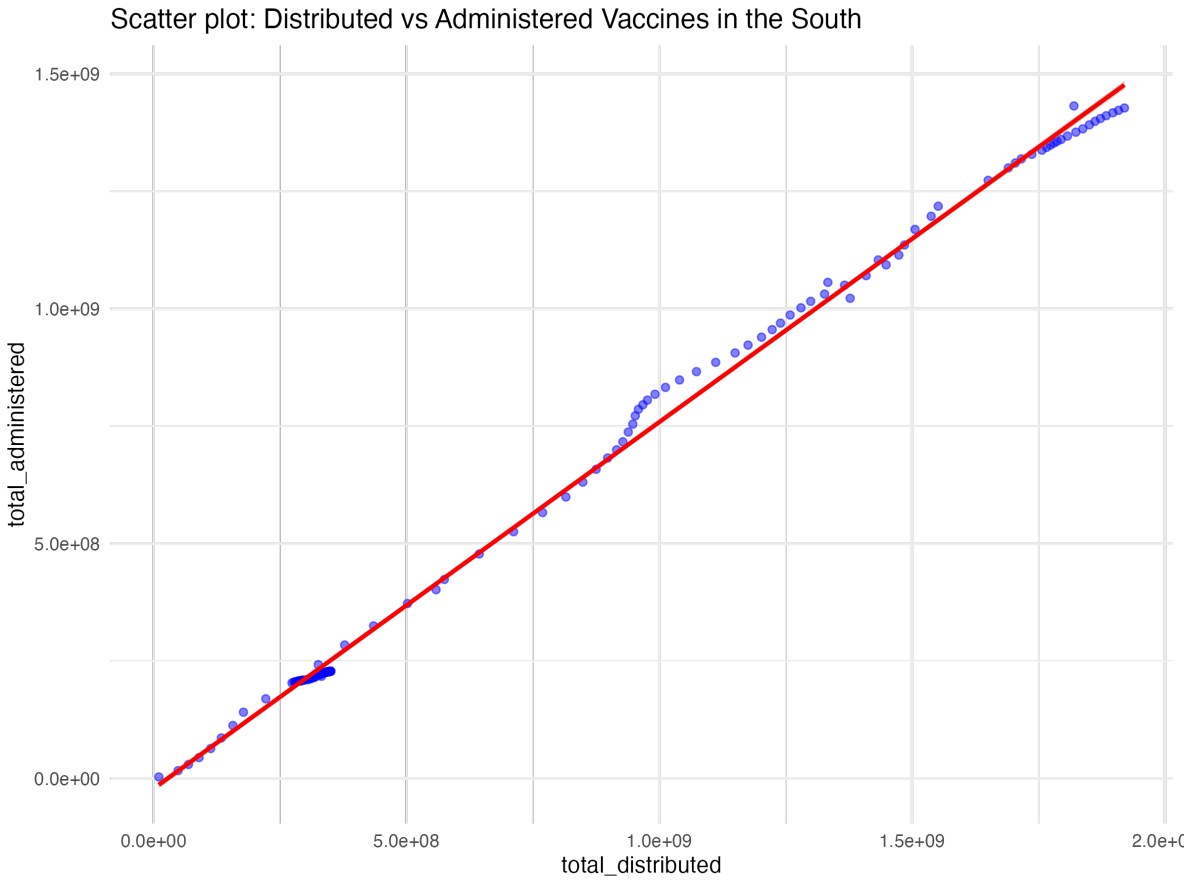

Figure 9: Figure 5.1: Scatterplot of Distributed vs Administered Doses in the South

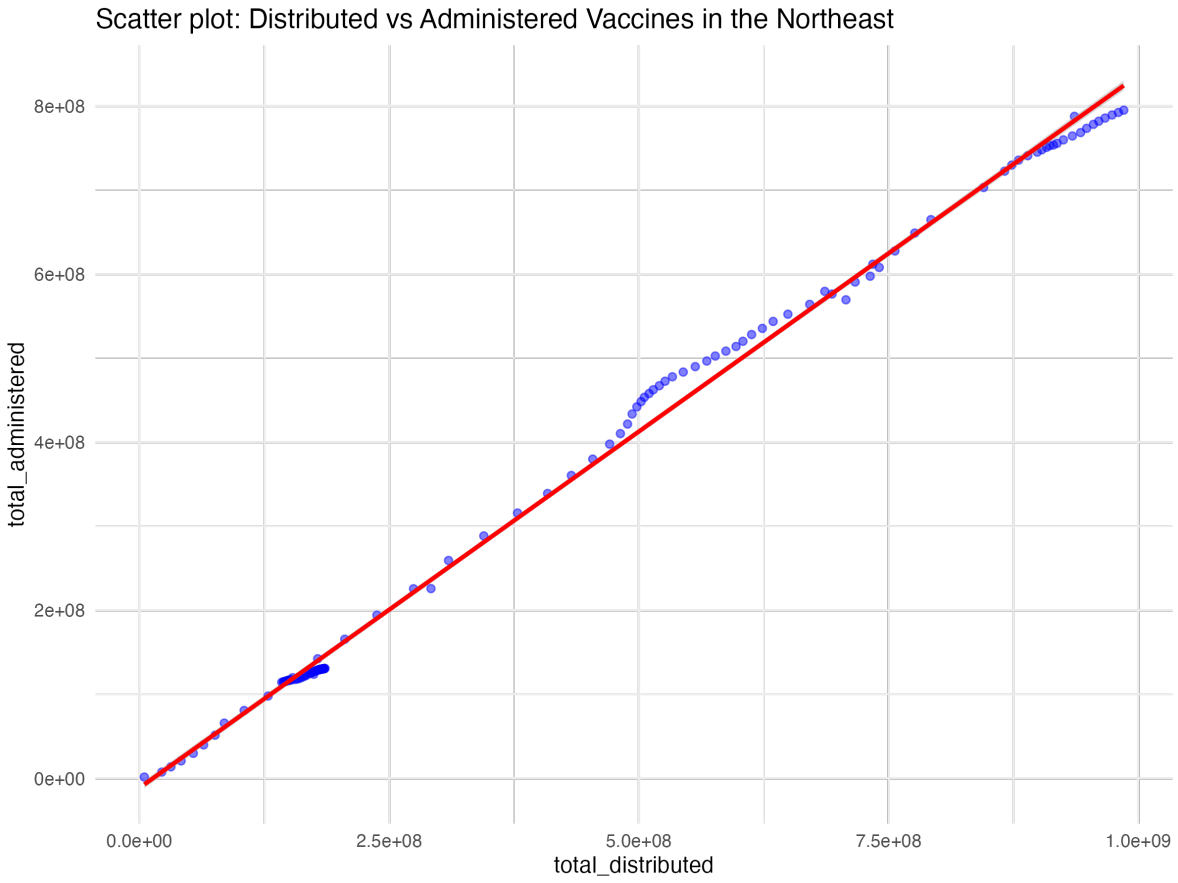

Figure 10: Figure 5.2: Scatterplot of Distributed vs Administered Doses in the Northeast

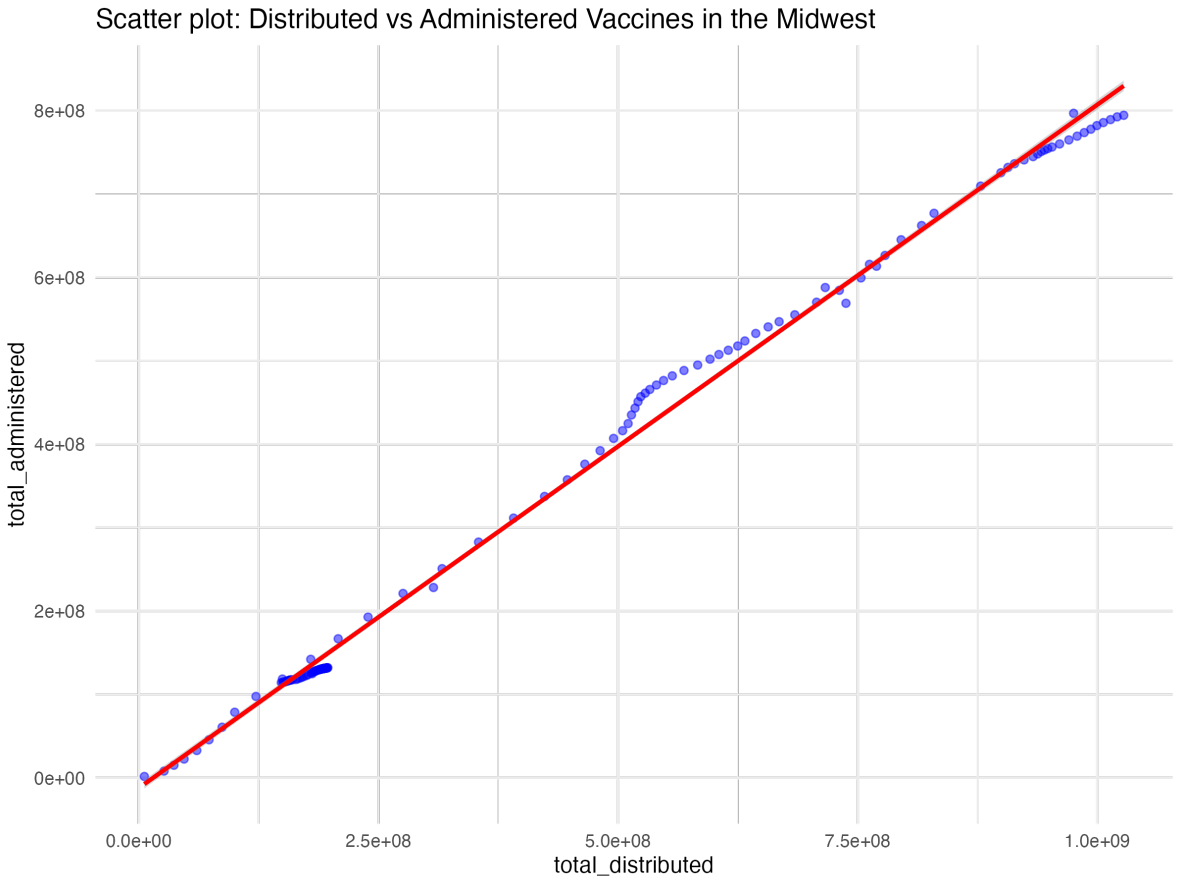

Figure 11: Figure 5.3: Scatterplot of Distributed vs Administered Doses in the Midwest

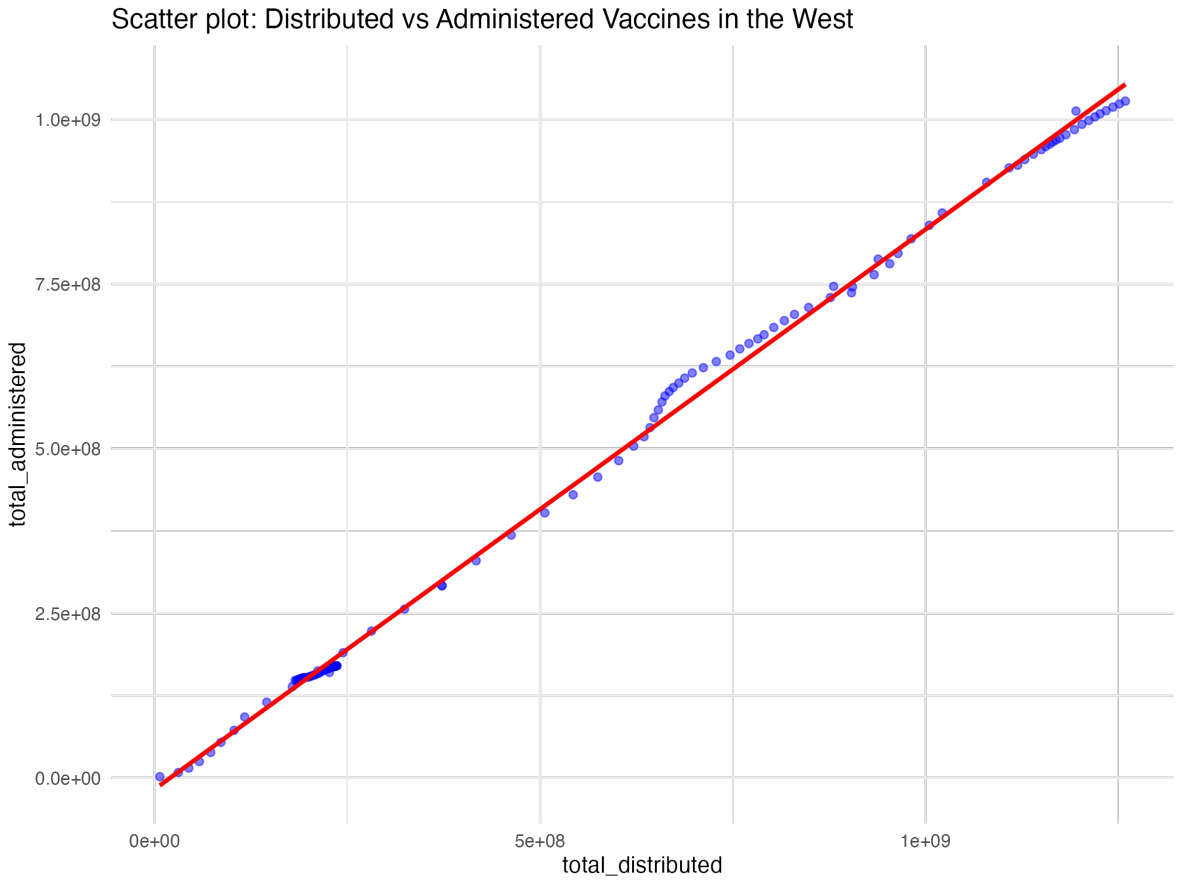

Figure 12: Figure 5.4: Scatterplot of Distributed vs Administered Doses in the West

Table two shows the correlations between doses administered and distributed in each region of the U.S.. All four regions have high correlations, however the west is the highest.

Region	cor
Midwest	0.9980132
Northeast	0.9983011
South	0.9987066
West	0.9990827

Table Two: Regional Correlations Between Doses Administered and Distributed

Table three shows the percent rate change in distribution and administration of the COVID-19 vaccine across time in each region of the U.S.. The drops depicted in figures 3 and 4 in the

manuscript can be seen in this table by the large percent drops in doses administered and distributed.

Year	Region	total_distributed	total_administered	pct_change_distributed	pct_change_administered
2021	Midwest	24555140480	20187443217	NA	NA
2021	Northeast	23150152060	19674659552	NA	NA
2021	South	46149547900	35805677422	NA	NA
2021	West	30421086245	25387091552	NA	NA
2022	Midwest	27431307710	21402228470	11.71310	6.017529
2022	Northeast	26339708355	21361831614	13.77769	8.575356
2022	South	51282910065	38338925179	11.12332	7.074989
2022	West	33682007550	27511629525	10.71928	8.368576
2023	Midwest	3671695345	2491217667	-86.61495	-88.360008
2023	Northeast	3460208875	2467206770	-86.86315	-88.450397
2023	South	6551831030	4312412188	-87.22414	-88.751870
2023	West	4408413035	3209234256	-86.91167	-88.334990

Table Three: Percent Rate Change in Distribution and Administration of COVID-19 Vaccine with time in each Region

Table 4.1 and 4.2 show the percent rate change in distribution (4.1) and administration (4.2) of the COVID-19 vaccine across time for each manufacturer. The drops depicted in figures 5 and 6 in the manuscript can be seen in this table by the large percent drops in doses administered and distributed.

Manufacturer	Year	Total Doses Distributed	Rate of Change
Janssen	2021	6175408550	-
Janssen	2022	5881340100	-4.76%
Janssen	2023	591262800	-89.95%
Moderna	2021	48667627880	-
Moderna	2022	51287641510	5.38%
Moderna	2023	6284176640	-87.75%
Novavax	2021	0	-
Novavax	2022	17008700	Inf%
Novavax	2023	22520600	32.41%
Pfizer	2021	66166755725	-
Pfizer	2022	81547873670	23.25%
Pfizer	2023	10313691355	-87.35%

Table 4.1: Percent Rate Change in Distribution of COVID-19 Vaccine with time for each Manufacturer

Manufacturer	Year	Total Doses Administered	Rate of Change
Janssen	2021	3580220161	-
Janssen	2022	3529276616	-1.42%
Janssen	2023	351643618	-90.04%
Moderna	2021	40522029628	-
Moderna	2022	40942762461	1.04%
Moderna	2023	4527384067	-88.94%
Novavax	2021	0	-
Novavax	2022	753416	Inf%
Novavax	2023	1485876	97.22%
Pfizer	2021	56854386805	-
Pfizer	2022	64035402590	12.63%
Pfizer	2023	7194506084	-88.76%

Table 4.2: Percent Rate Change in Administration of COVID-19 Vaccine with time for each Manufacturer

Table 5 shows the RMSE value of the null model (with no predictors) from the modeling analysis.

.metric	.estimator	.estimate
rmse	standard	0.1693047
rsq	standard	NA

Table 5: Null Model RMSE