

# Datathon

Data People

2023-02-19

What trends do you notice with respect to the total inpatient beds for COVID by specific age groups and time?

```
# Loading Libraries and reading the data
library(tidyverse)
fullData <- read_delim("COVID-19_Reported_Patient_Impact_and_Hospital_Capacity_by_Facility.csv")
# Below a table of the first few rows of the full data frame of all the data called fullData is shown (note that not all columns are shown here)
knitr::kable(head(fullData))
```

hospital_pk	collection_week	state	ccn	hospital_name	address	city	zip	hospital_subtype	fips_code	is_metro_micro	total_beds
052052	2020-05-29	CA	052052	KINDRED HOSPITAL RIVERSIDE	2224 MEDICAL CENTER DRIVE	PERRIS	92571	Long Term	06065	TRUE	100
070038	2020-04-24	CT	070038	CONNECTICUT HOSPICE INC,THE	100 DOUBLE BEACH ROAD	BRANFORD	06405	Short Term	09009	TRUE	100
100034	2020-05-01	FL	100034	MOUNT SINAI MEDICAL CENTER OF FLORIDA, INC	4300 ALTON RD	MIAMI BEACH	33140	Short Term	12086	TRUE	100
100069	2020-05-15	FL	100069	ADVENTHEALTH CARROLLWOOD	7171 N DALE MABRY HWY	TAMPA	33614	Short Term	12057	TRUE	100
100070	2020-04-24	FL	100070	SHOREPOINT HEALTH VENICE	540 THE RIALTO	VENICE	34285	Short Term	12115	TRUE	100
100156	2020-06-19	FL	100156	HCA FLORIDA LAKE CITY HOSPITAL	340 NW COMMERCE DR	LAKE CITY	32055	Short Term	12023	TRUE	100

```
# size10000 - smaller data set of only 10000 rows
size10000 <- sample_n(fullData, 10000)
# Below a table of the first few rows of the data frame called size10000 is shown (note that not
# all columns are shown here)
knitr::kable(head(size10000))
```

hospital_pk	collection_week	state	ccn	hospital_name	address	city	zip	hospital_subtype	fips_code	is_metro_micro	total_beds
390039	2021-06-11	PA	390039	UPMC SOMERSET	225 SOUTH CENTER AVENUE	SOMERSET	15501	Short Term	42111	TRUE	100
031319	2022-11-04	AZ	031319	MT. GRAHAM REGIONAL MEDICAL CENTER	1600 SOUTH 20TH AVENUE	SAFFORD	85546	Critical Access Hospitals	04009	TRUE	100
420009	2021-11-26	SC	420009	PRISMA HEALTH OCONEE MEMORIAL HOSPITAL	298 MEMORIAL DR	SENECA	29672	Short Term	45073	TRUE	100
170182	2021-03-19	KS	170182	MENORAH MEDICAL CENTER	5721 WEST 119TH STREET	OVERLAND PARK	66209	Short Term	20091	TRUE	100

hospital_pk	collection_week	state	ccn	hospital_name	address	city	zip	hospital_subtype	fips_code	is_metro_micro	total_b
240104	2021-10-08	MN	240104	ST FRANCIS REGIONAL MEDICAL CENTER	1455 ST FRANCIS AVENUE	SHAKOPEE	55379	Short Term	27139	TRUE	
390028	2021-04-23	PA	390028	UPMC MERCY	1400 LOCUST STREET	PITTSBURGH	15219	Short Term	42003	TRUE	

```
# final - the smaller data set of 10000 called size10000 with any averages and sums of inpatient beds
# with values less than 4 replace with a value of 2. The data is also arranged by date (collection week)
# and is shown from oldest to newest week
final <- size10000 %>%
  # -999,999.0 is any average or sum value that is 4 or less. The mutated columns below replace any value of -999999.0 as 2
  # since 2 is the average between 0 and 4, and will better represent the data than replacing -999,999.0 with 0
  mutate(total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg = replace(total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg, total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg == -999999.0, 2.0)) %>%
  mutate(total_adult_patients_hospitalized_confirmed_covid_7_day_avg = replace(total_adult_patients_hospitalized_confirmed_covid_7_day_avg, total_adult_patients_hospitalized_confirmed_covid_7_day_avg == -999999.0, 2.0)) %>%
  mutate(inpatient_beds_used_covid_7_day_avg = replace(inpatient_beds_used_covid_7_day_avg, inpatient_beds_used_covid_7_day_avg == -999999.0, 2.0)) %>%
  mutate(total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum = replace(total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum, total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum == -999999.0, 2.0)) %>%
  mutate(total_adult_patients_hospitalized_confirmed_covid_7_day_sum = replace(total_adult_patients_hospitalized_confirmed_covid_7_day_sum, total_adult_patients_hospitalized_confirmed_covid_7_day_sum == -999999.0, 2.0)) %>%
  mutate(inpatient_beds_used_covid_7_day_sum = replace(inpatient_beds_used_covid_7_day_sum, inpatient_beds_used_covid_7_day_sum == -999999.0, 2.0)) %>%
  select(collection_week, hospital_subtype, state, city,
total_pediatric_patients_hospitalized_confirmed_covid_7_day_coverage, total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum, total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg,
total_adult_patients_hospitalized_confirmed_covid_7_day_coverage, total_adult_patients_hospitalized_confirmed_covid_7_day_sum, total_adult_patients_hospitalized_confirmed_covid_7_day_avg,
inpatient_beds_used_covid_7_day_coverage, inpatient_beds_used_covid_7_day_sum, inpatient_beds_used_covid_7_day_avg) %>%
  arrange(collection_week)

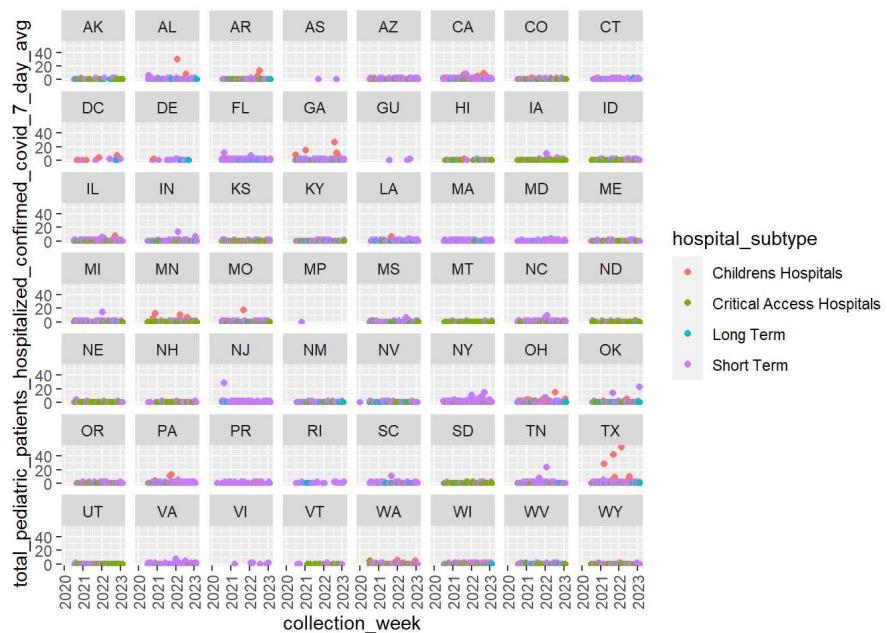
# Below a table of the first few rows of the data frame called final is shown (note that not
# all columns are shown here)
knitr::kable(head(final))
```

collection_week	hospital_subtype	state	city	total_pediatric_patients_hospitalized_confirmed_covid_7_day_coverage	total_pediatric_j
2019-12-27	Short Term	NV	LAS VEGAS		1
2020-02-07	Short Term	TX	LUBBOCK		0
2020-02-28	Critical Access Hospitals	MS	RIPLEY		0
2020-02-28	Critical Access Hospitals	IA	SHENANDOAH		0
2020-03-13	Short Term	KY	RUSSELLVILLE		0
2020-03-13	Critical Access Hospitals	MN	INTERNATIONAL FALLS		0

```
####----- below is for average

# The dot plots below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))

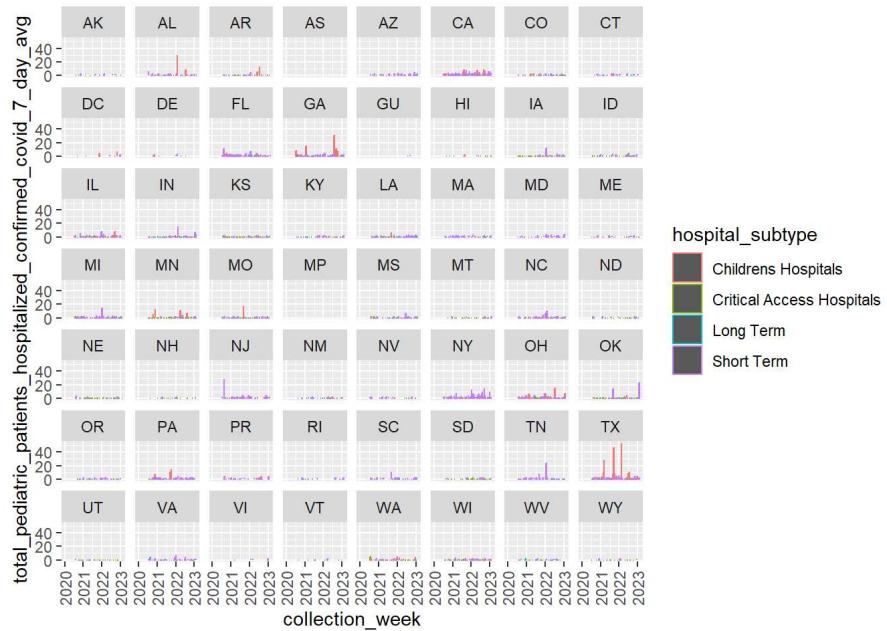
## Warning: Removed 1008 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

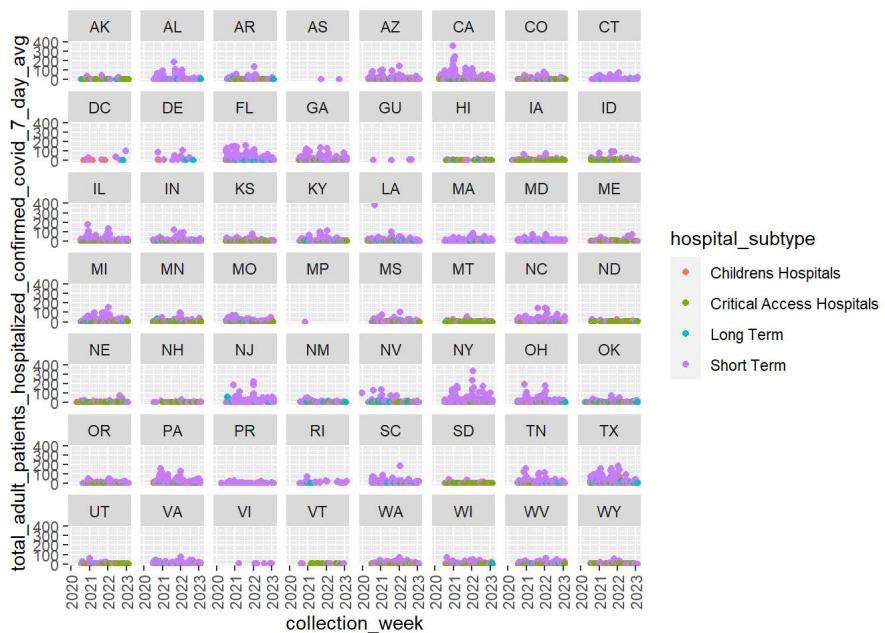
```
## Warning: Removed 1008 rows containing missing values (`position_stack()`).
```



```
# The dot plots below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

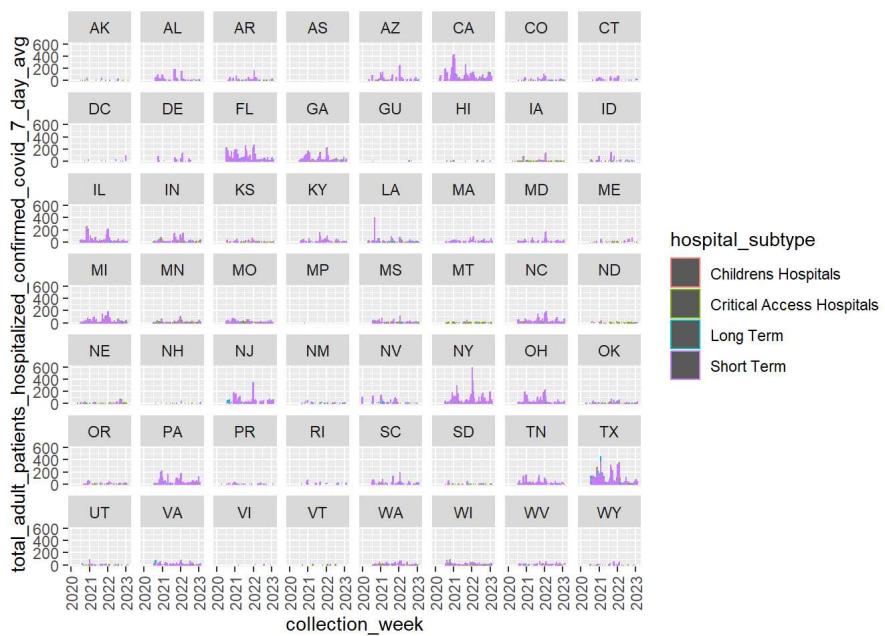
```
## Warning: Removed 927 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

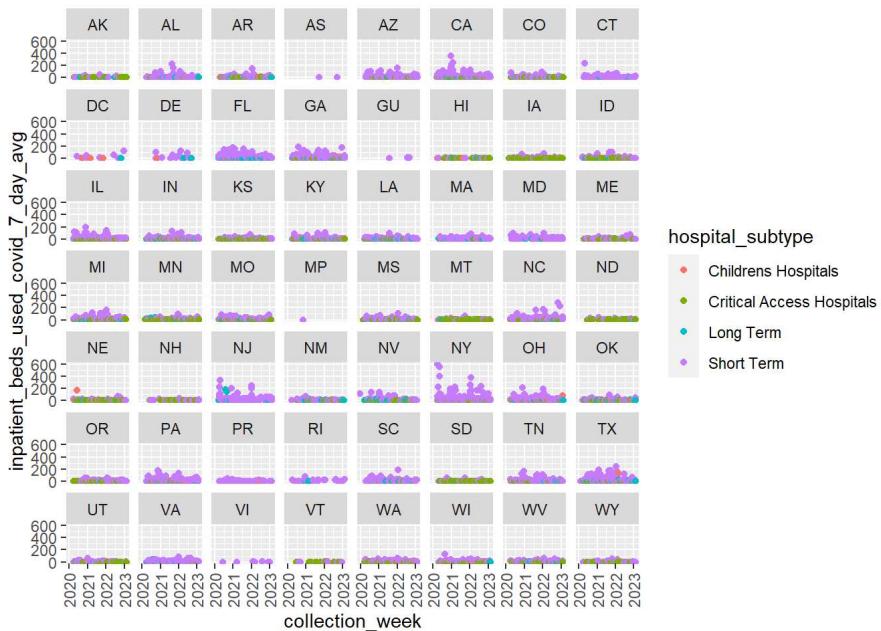
```
## Warning: Removed 927 rows containing missing values (`position_stack()`).
```



```
# The dot plots below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total amount of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

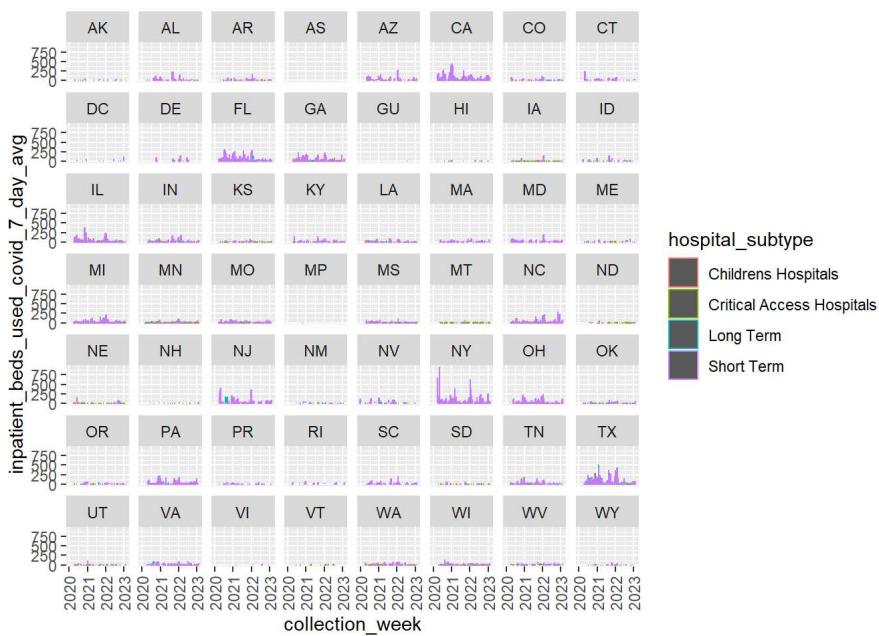
```
## Warning: Removed 57 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day average each week from 2020-04-10 to 2023-02-10
# for total amountn of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_avg)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

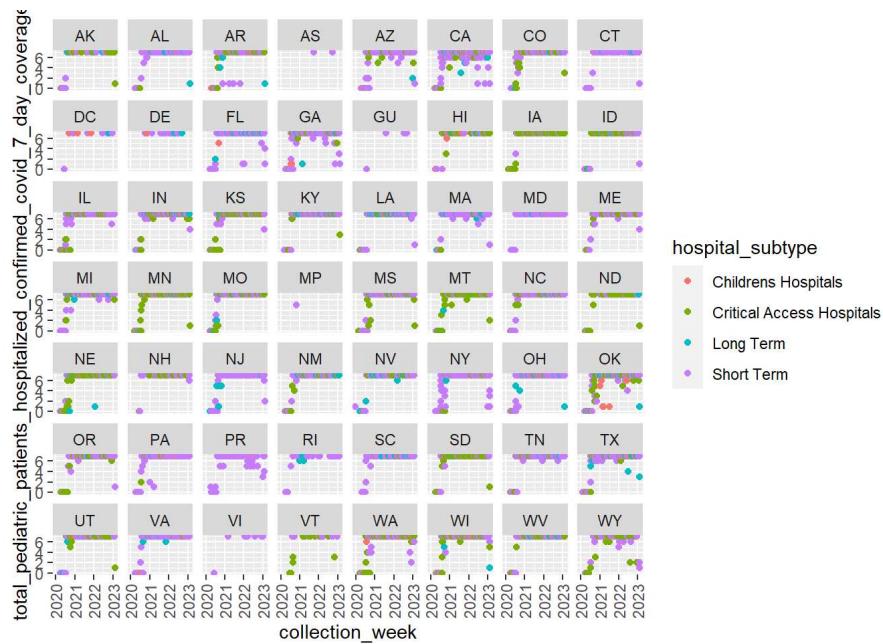
```
## Warning: Removed 57 rows containing missing values (`position_stack()`).
```



```
####----- below is for coverage
```

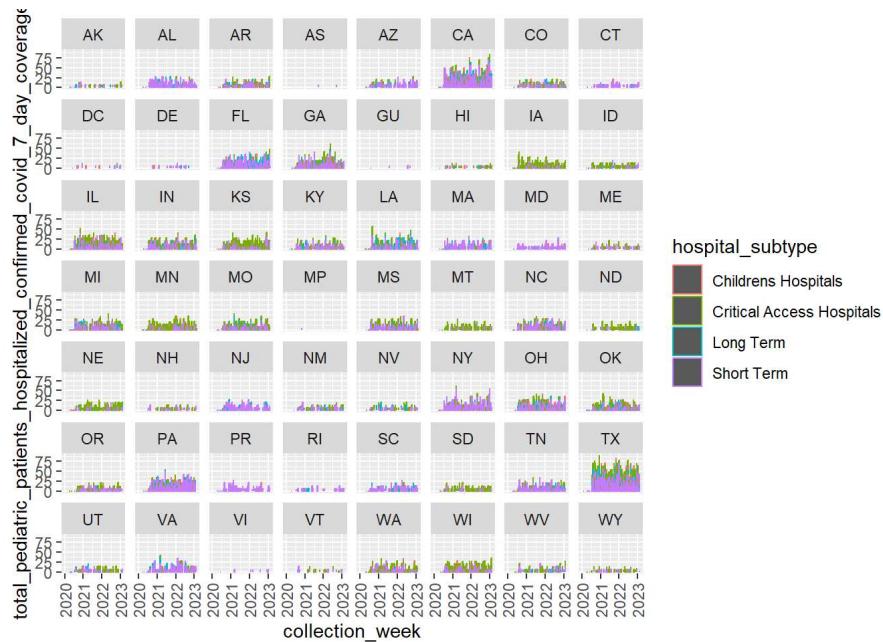
```
# The dot plots below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



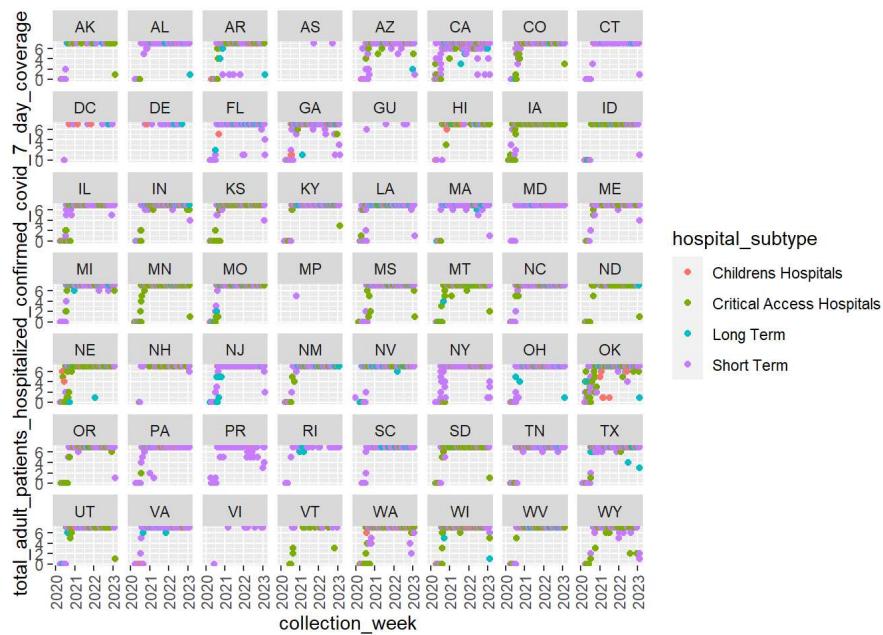
```
# The bar graphs below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



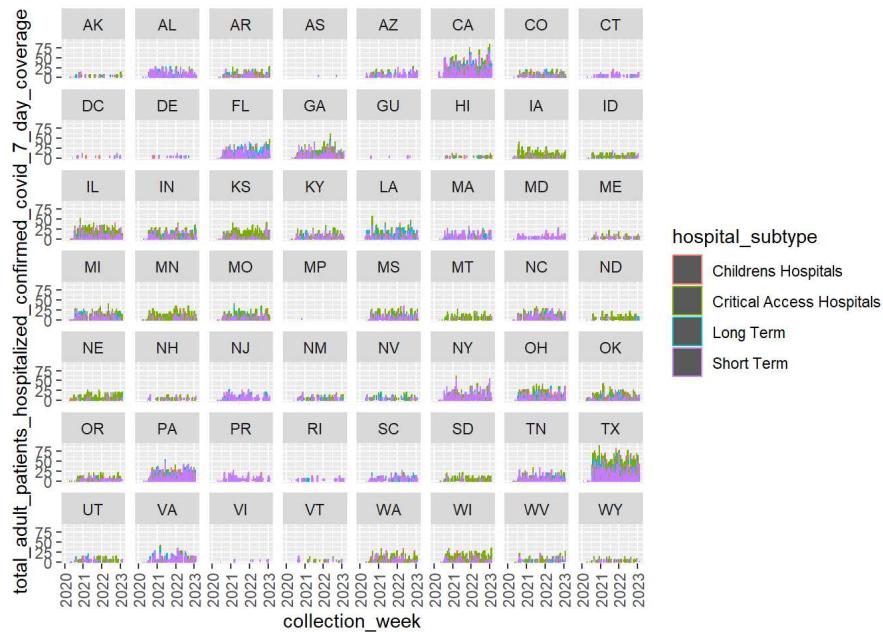
```
# The dot plots below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



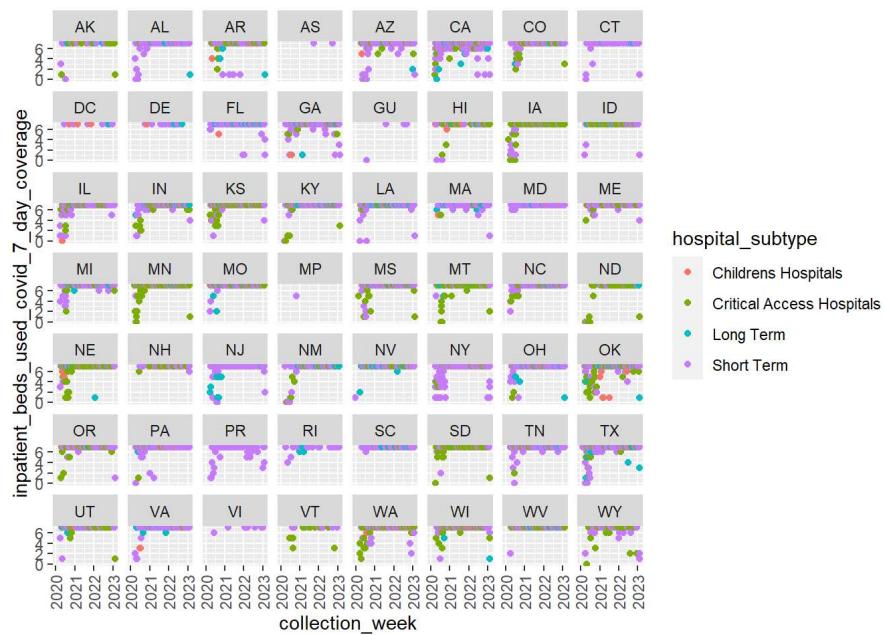
```
# The bar graphs below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



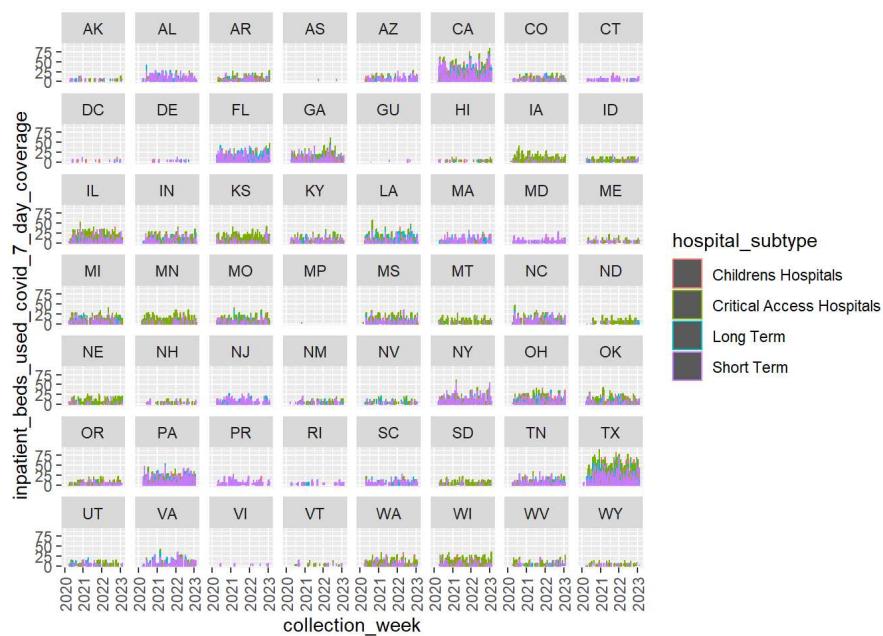
```
# The dot plots below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total amountn of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



```
# The bar graphs below are of each state for the 7 day coverage each week from 2020-04-10 to 2023-02-10
# for total amountn of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_coverage)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

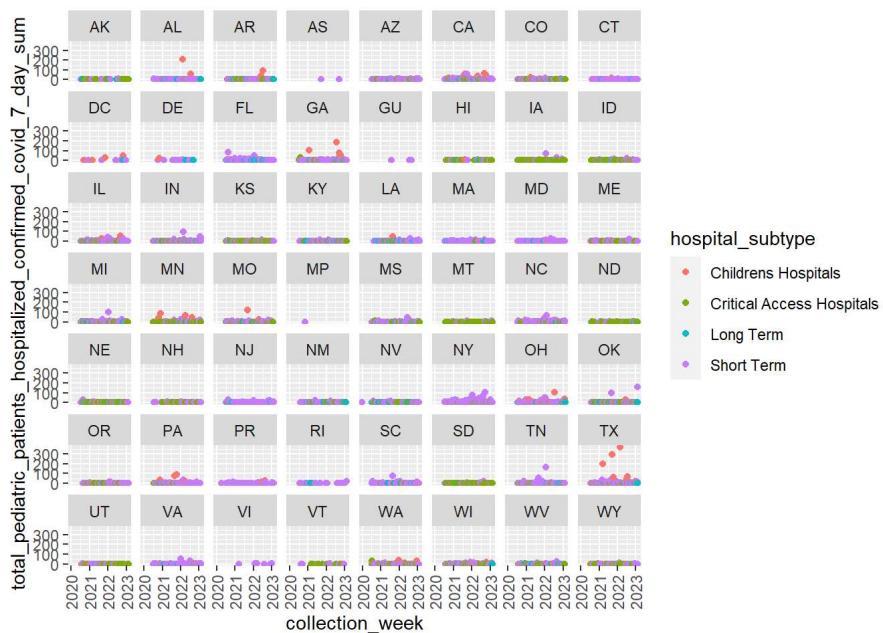


```
####----- below is for sum
```

```
# The dot plots below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

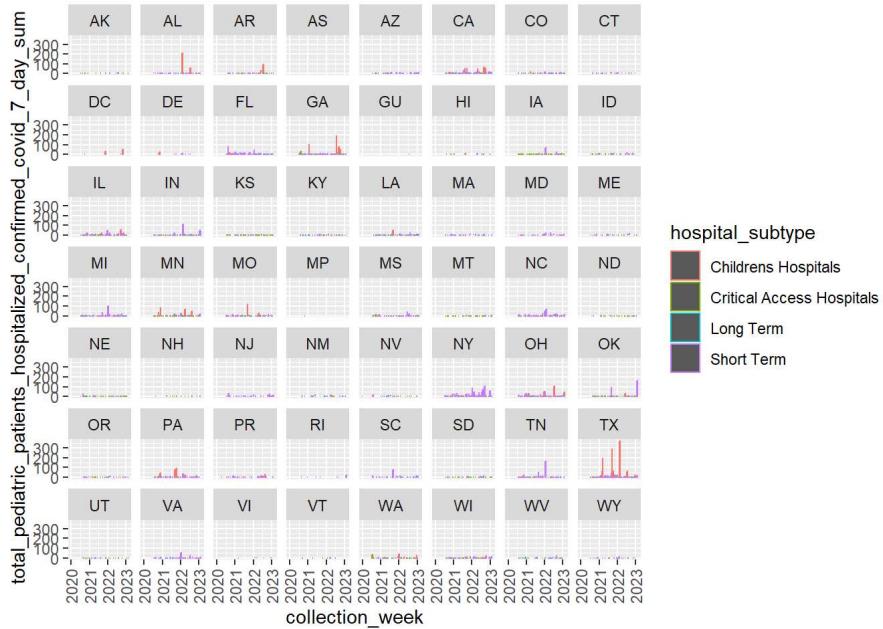
```
## Warning: Removed 1008 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total pediatric patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_pediatric_patients_hospitalized_confirmed_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

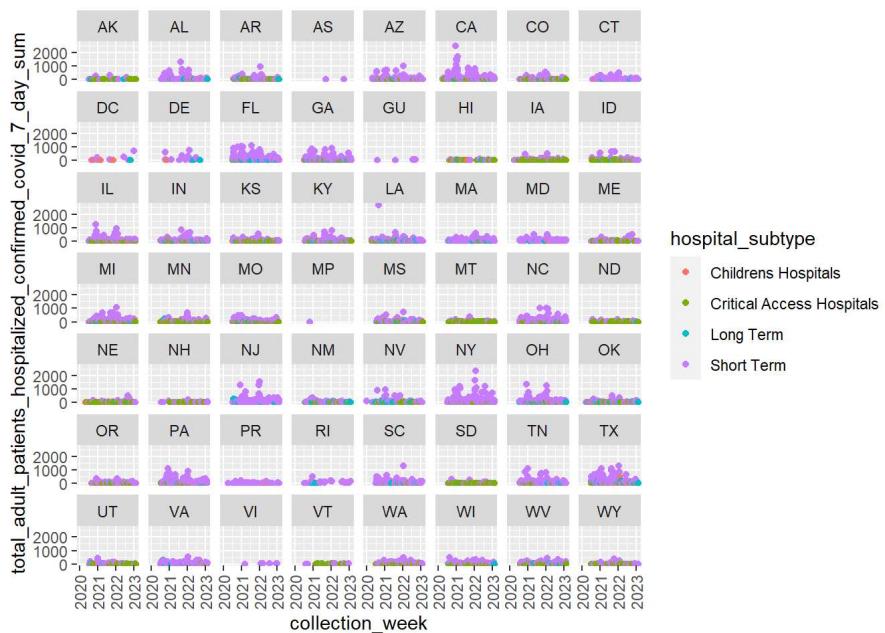
```
## Warning: Removed 1008 rows containing missing values (`position_stack()`).
```



```
# The dot plots below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

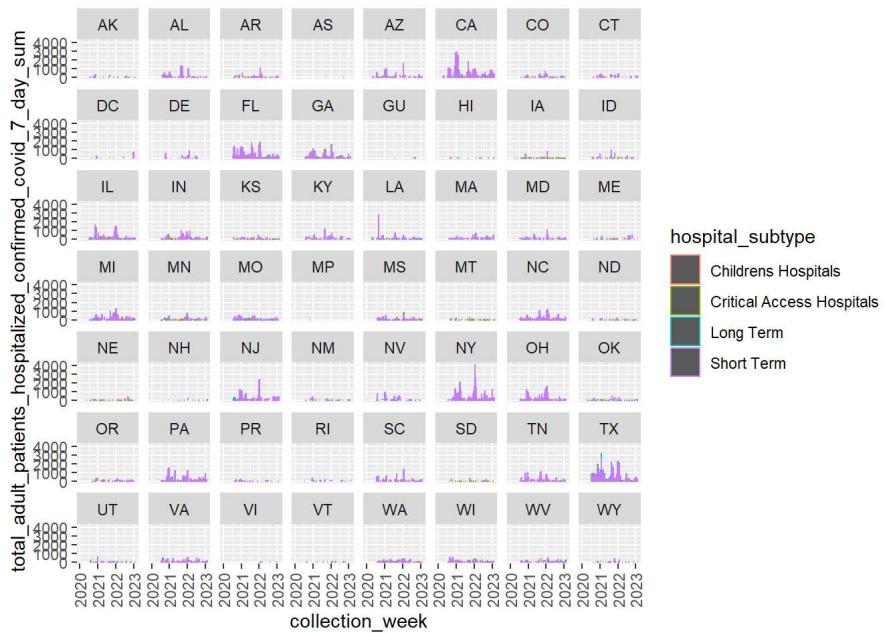
```
## Warning: Removed 927 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total adult patients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = total_adult_patients_hospitalized_confirmed_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

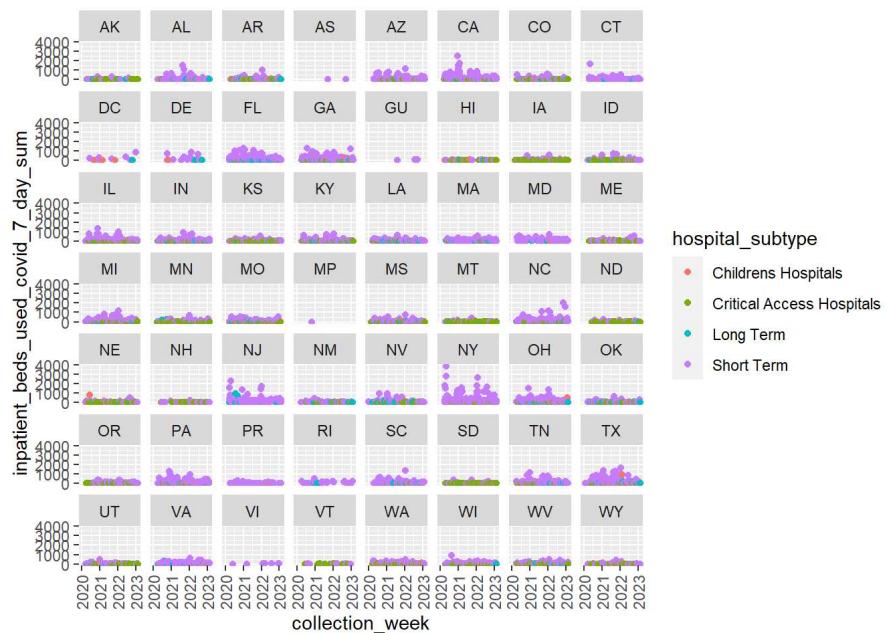
```
## Warning: Removed 927 rows containing missing values (`position_stack()`).
```



```
# The dot plots below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total amount of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_point(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

```
## Warning: Removed 57 rows containing missing values (`geom_point()`).
```



```
# The bar graphs below are of each state for the 7 day sum each week from 2020-04-10 to 2023-02-10
# for total amountn of inpatients with COVID hospitalized.
```

```
ggplot(data = final, mapping = aes(col = hospital_subtype)) +
  geom_col(mapping = aes(x = collection_week, y = inpatient_beds_used_covid_7_day_sum)) +
  facet_wrap(~state) + theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```

```
## Warning: Removed 57 rows containing missing values (`position_stack()`).
```

