

# HTML Doc

Analysis Setup

Analysis

Import Our Vaccine Provider and Supply Data

Transform our Vaccine Data

Visualize our Vaccine Data

Export our Transformed Dataset and Visualization

Bonus

Regression Example

Regression Table

Regression Chart

Regression Equation

## Analyzing Vaccine Supply in Texas

Matt Worthington

The LBJ School of Public Affairs

2021-03-29

### Analysis Setup

Before we start building out our reproducible analysis, let's go ahead and make sure any R packages are loaded and installed properly. The code to install necessary packages and load them can be viewed by clicking on the "Show Code" arrow.

```
knitr::opts_chunk$set(warning = FALSE, message = FALSE)
# In case these aren't installed, uncomment this and run it.
# install.packages("janitor", "tidyverse", "gt")
# devtools::install_github("utexas-lbjp-data/lbjdata")

library(janitor)      # Package with useful + convenient data cleaning functions
library(tidyverse)    # Core Set of R Data Science Tools (dplyr, ggplot2, tidyr, readr, etc.)
```

### Analysis

### Import Our Vaccine Provider and Supply Data

This data comes from the Texas Department of State Health Services and contains the list of vaccine providers across the state of Texas, which can be found on [this page](#). They use it for their own interactive mapping application of vaccine provider sites.<sup>1</sup> Each provider is assigned a type and has a report of how much vaccine supply they have for each of the three approved vaccines. We'll use the `read_csv()` function to read in the data straight from the DSHS website. This will help make sure our analysis is "living", meaning any chart we make will update whenever the feed from DSHS gets updated, and "reproducible", meaning anyone who takes this R Markdown document can run it in their RStudio IDE and get the exact same thing you did.

The `read_csv()` comes from the `readr` package that was loaded when we ran `library(tidyverse)` in the setup chunk above (lines 18:30 in the RMarkdown document).

```
provider_data_raw <- readr::read_csv("https://genesis.soc.texas.gov/files/accessibility/vaccineprovideraccessabilitydata.csv") %>%
  janitor::clean_names() # This function makes column headers machine readable

dplyr::glimpse(provider_data_raw) # glimpse() lets you preview a data object
```

```
## Rows: 3,381
## Columns: 17
```

# PDF Doc

Analyzing Vaccine Supply in Texas

true

2021-03-29

### Analysis Setup

Before we start building out our reproducible analysis, let's go ahead and make sure any R packages are loaded and installed properly. The code to install necessary packages and load them can be viewed by clicking on the "Show Code" arrow.

```
knitr::opts_chunk$set(warning = FALSE, message = FALSE)
# In case these aren't installed, uncomment this and run it.
# install.packages("janitor", "tidyverse", "gt")
# devtools::install_github("utexas-lbjp-data/lbjdata")

library(janitor)      # Package with useful + convenient data cleaning functions
library(tidyverse)    # Core Set of R Data Science Tools (dplyr, ggplot2, tidyr, readr, etc.)
```

### Analysis

### Import Our Vaccine Provider and Supply Data

This data comes from the Texas Department of State Health Services and contains the list of vaccine providers across the state of Texas, which can be found on this page. They use it for their own interactive mapping application of vaccine provider sites.<sup>1</sup> Each provider is assigned a type and has a report of how much vaccine supply they have for each of the three approved vaccines. We'll use the `read_csv()` function to read in the data straight from the DSHS website. This will help make sure our analysis is "living", meaning any chart we make will update whenever the feed from DSHS gets updated, and "reproducible", meaning anyone who takes this R Markdown document can run it in their RStudio IDE and get the exact same thing you did.

The `read_csv()` comes from the `readr` package that was loaded when we ran `library(tidyverse)` in the setup chunk above (lines 18:30 in the RMarkdown document).

```
provider_data_raw <- readr::read_csv("https://genesis.soc.texas.gov/files/accessibility/vaccineproviderjanitor::clean_names() # This function makes column headers machine readable

dplyr::glimpse(provider_data_raw) # glimpse() lets you preview a data object

## Rows: 3,380
## Columns: 17
## $ name          <chr> "Hamarsyl Pharmacy LLC", "Baylor Scott & White Cl-
## $ type          <chr> "Pharmacy", "Other", "Pharmacy", "Pharmacy", "Com-
## $ tsa           <chr> NA, NA, NA, NA, NA, "C", NA, NA, NA, NA, NA, NA, -
## $ street        <chr> NA, "2608 BROCKTON DR", "Bill Mack Rd", "1305 N M-
## $ city          <chr> "McAllen", "AUSTIN", "Shamrock", "Vider", "Pearla-
## $ county        <chr> NA, NA, "Wheeler", "Orange", "Brazoria", "Wilbarg-
```

<sup>1</sup>The link for this map is google.com

1

# Shifting How You Work

## Source vs. Output