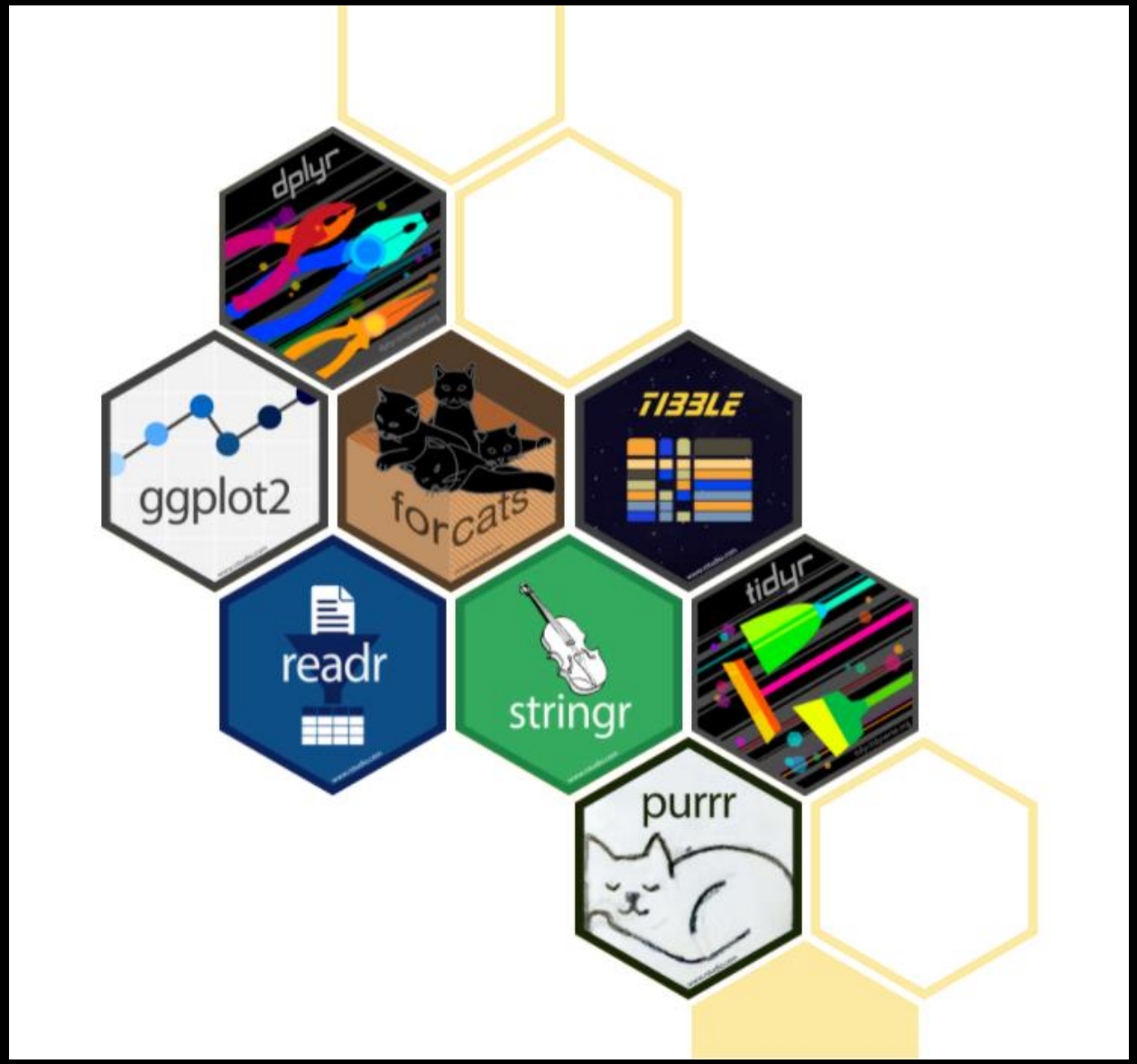
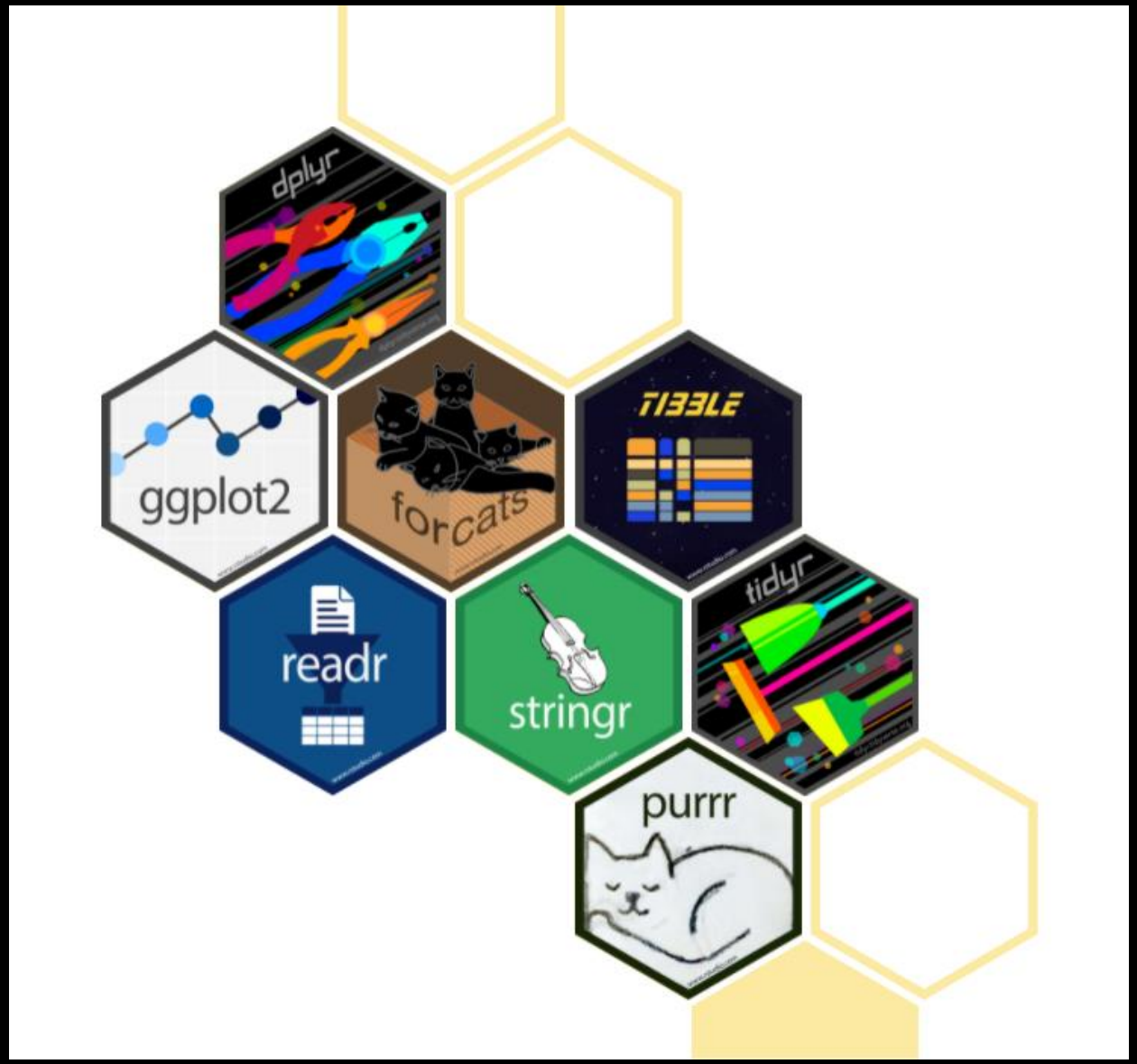
tidyverse

- Collection of packages that work with each other
- Intended for complete data process: import, cleaning, manipulation, visualization, modeling, etc.



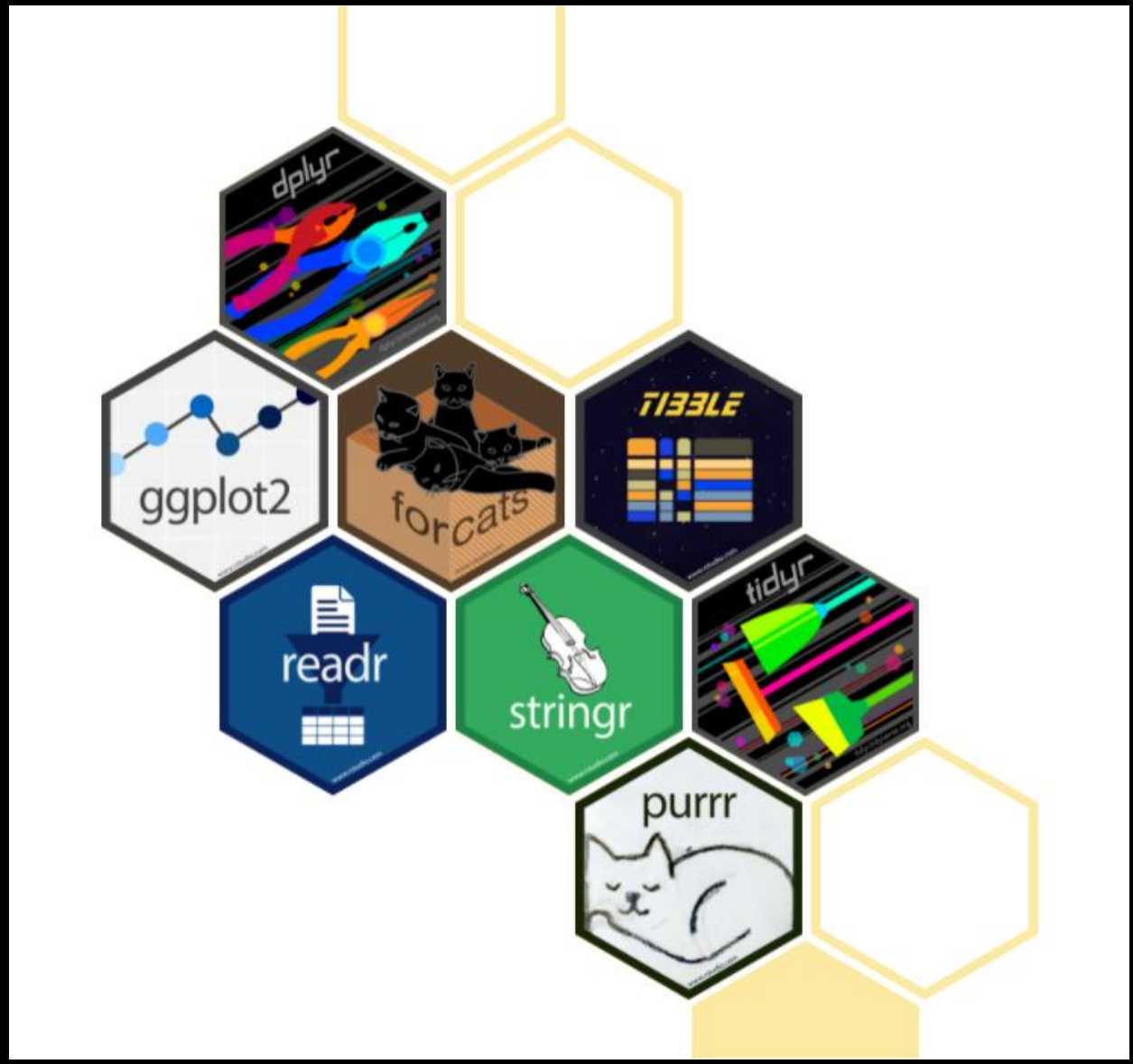
tidyverse

- Really good references/learning materials available online
- Very popular way to use R!



tidyverse

- We will use some tidyverse methods for data manipulation and data visualization



tidyverse data manipulation

- some common commands used in tidyverse for data manipulation
  - subset data
  - adding, deleting, changing data variables
  - rearranging data

- filter()
- mutate()
- select()
- spread()
- arrange()

tidyverse data manipulation

- pipe operator
  - %>%
  - take the output from left side and put it into the next command or argument
  - easy to set up “layers” of commands

tidyverse data manipulation

- pipe operator
  - %>%
  - take the output from left side and put it into the next command or argument
  - easy to set up “layers” of commands

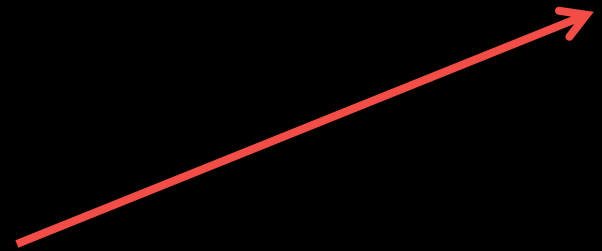
```
data <- covid19_df %>%
    filter(location == "Bangladesh") %>%
    spread(data_type, value) %>%
    arrange(date) %>%
    mutate(deaths = deaths_new)
```

tidyverse data visualization

- ggplot2 package
- “Grammar of Graphics”
- building plots in layers

ggplot data visualization

- Always start with a ggplot() command to create a graph grid/plot → ggplot(data = mpg) +
- plotting commands set up in rows with '+' between each



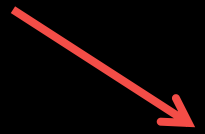
ggplot data visualization

- Tell ggplot what type of plot to make and which variables to use

```
ggplot(data = mpg) +
  geom_point(mapping = aes(x=displ,y=hwy)) +
```

ggplot data visualization

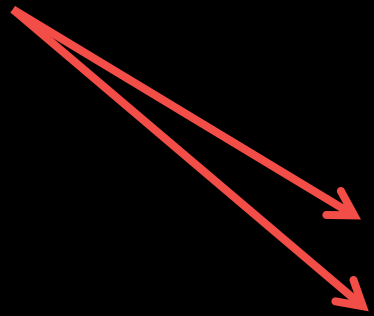
- Make adjustments to the scale (optional)



```
ggplot(data = mpg) +
    geom_point(mapping = aes(x=displ,y=hwy)) +
    scale_x_continuous(breaks=c(2,4,6))
```

ggplot data visualization

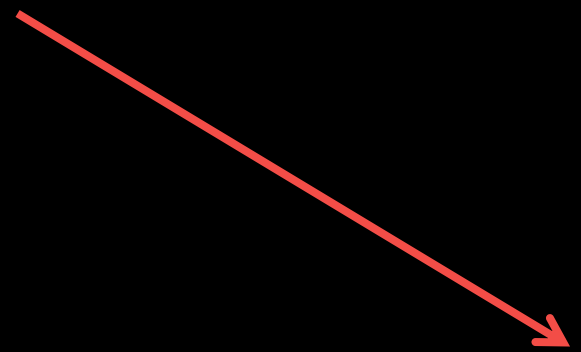
- Add titles and labels (optional)



```
ggplot(data = mpg) +
    geom_point(mapping = aes(x=displ,y=hwy)) +
    scale_x_continuous(breaks=c(2,4,6)) +
    ggtitle(“Car Data”) +
    labs(x=“Engine Size”, y=“Fuel Efficiency”)
```

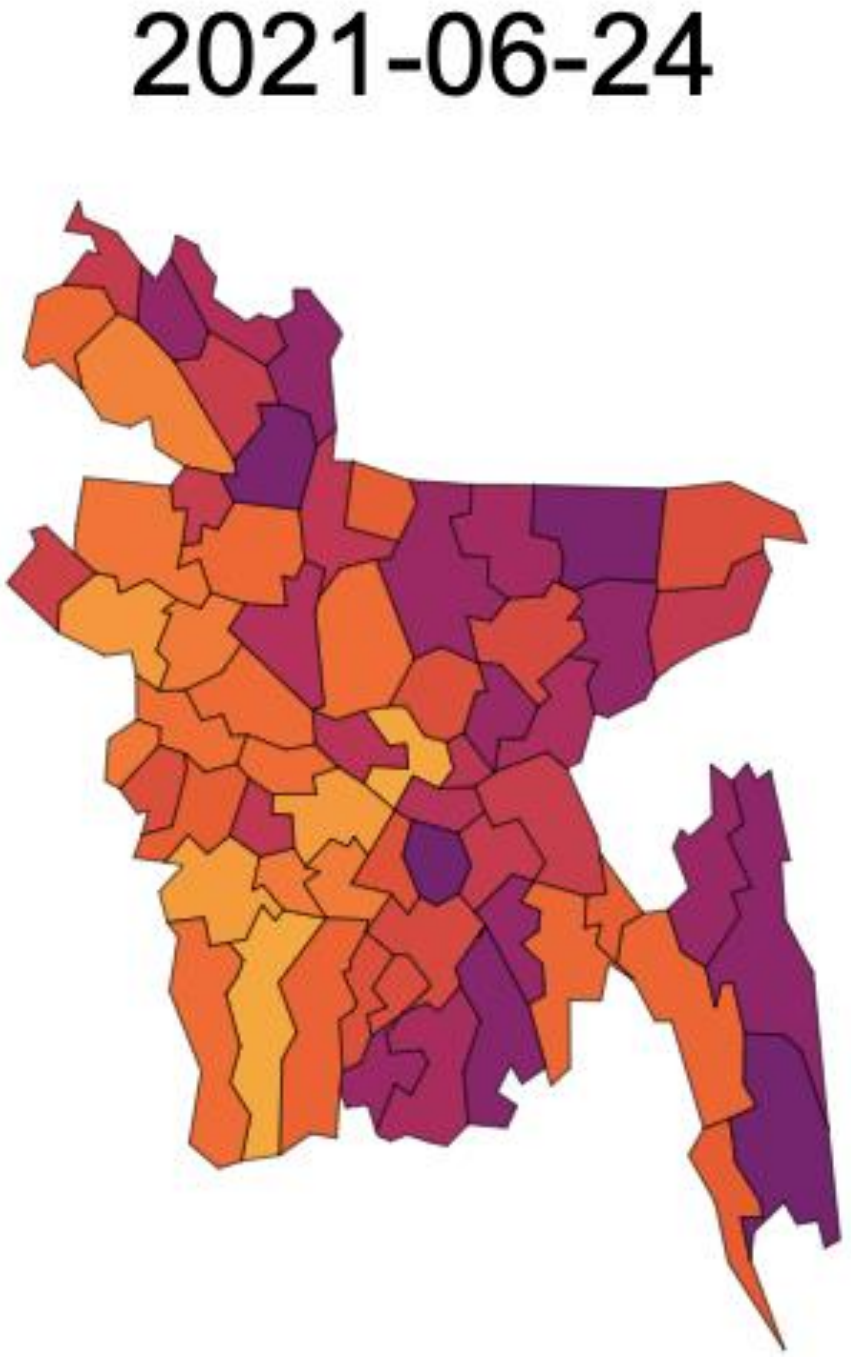
ggplot data visualization

- Use ggthemes to set up and standardize visual details



```
ggplot(data = mpg) +
    geom_point(mapping = aes(x=displ,y=hwy)) +
    scale_x_continuous(breaks=c(2,4,6)) +
    ggtitle(“Car Data”) +
    labs(x=“Engine Size”, y=“Fuel Efficiency”) +
    theme
```

Mapping in R

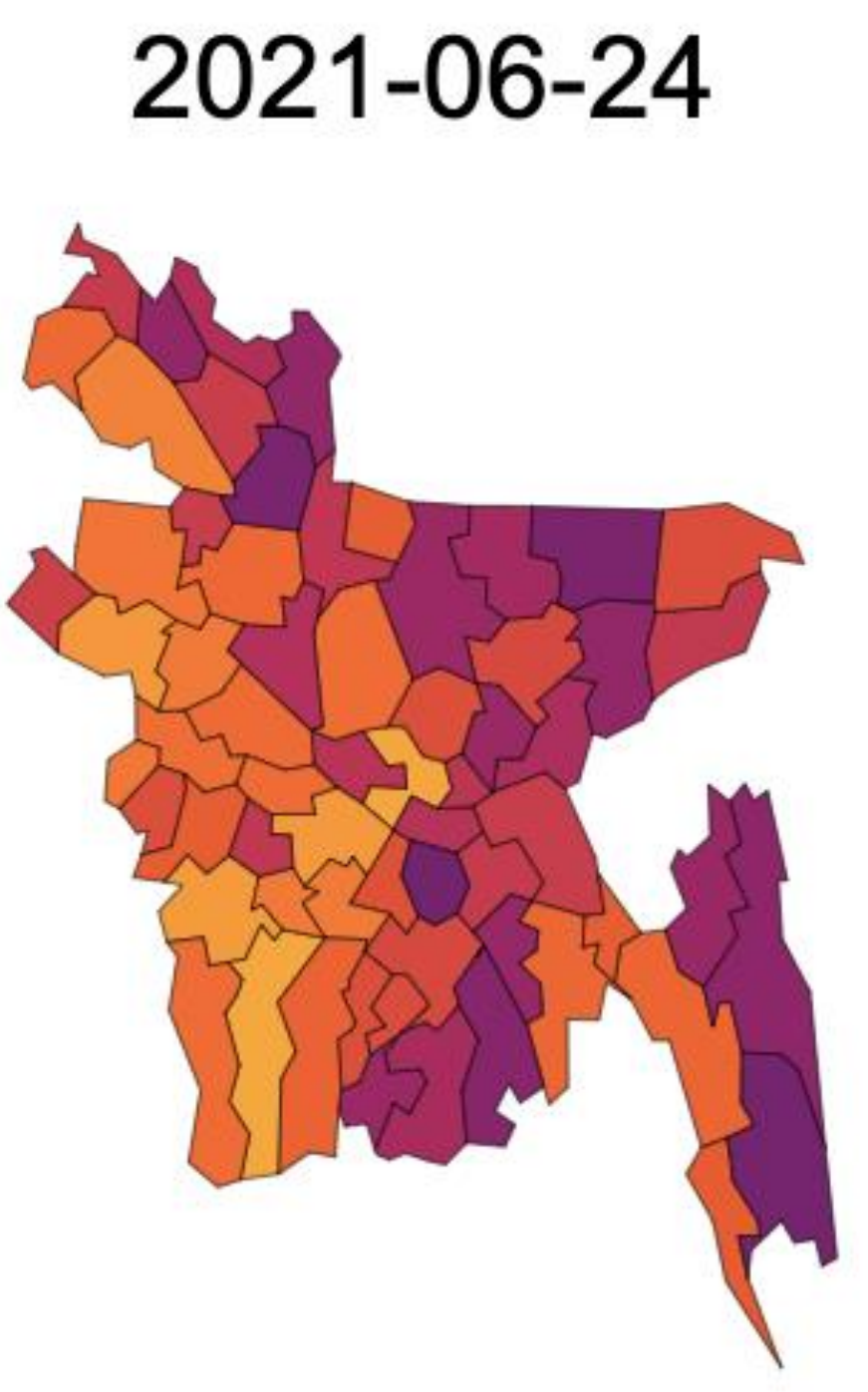


Spatial & Mapping Packages

- ggmap: use ggplot2 for mapping
- rgdal: spatial data processing
- rgeos: vector processing
- maptools: mapping features
- sf: simplifies spatial data

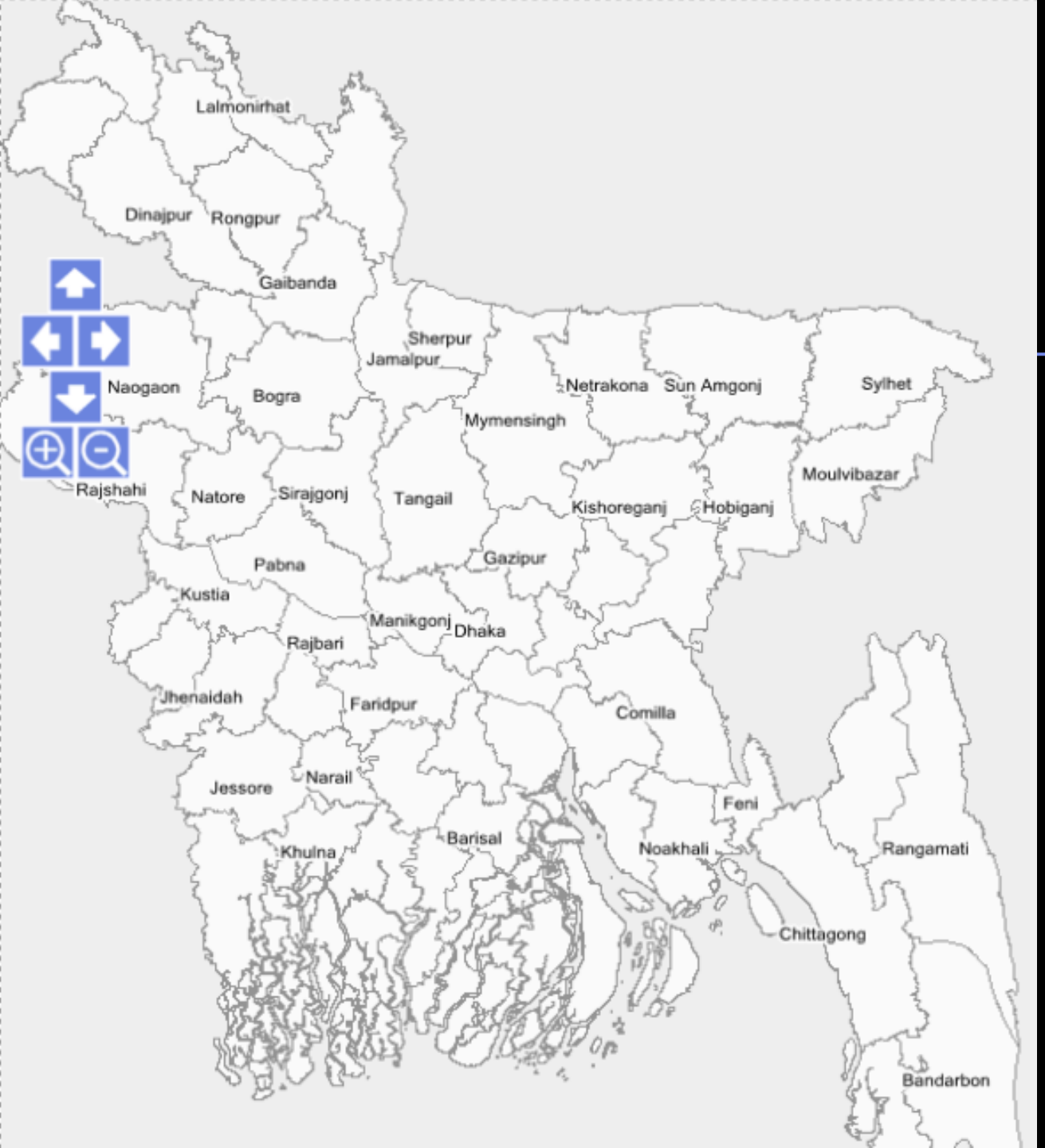
- RColorbrewer: color schemes for mapping
- viridis: additional color schemes

- many, many more!



Mapping Data

- Many types of maps!
- spatial visualization of data
  - choropleth map
    - color in each administrative area corresponds to our data of interest
    - areal data
    - popular visualization



Mapping Data

1. Spatial object: geographic representation of our area of interest
   - boundaries of administrative areas
   - where are they located? (CRS)
   - id/names for administrative areas
   - may be very large

Mapping Data

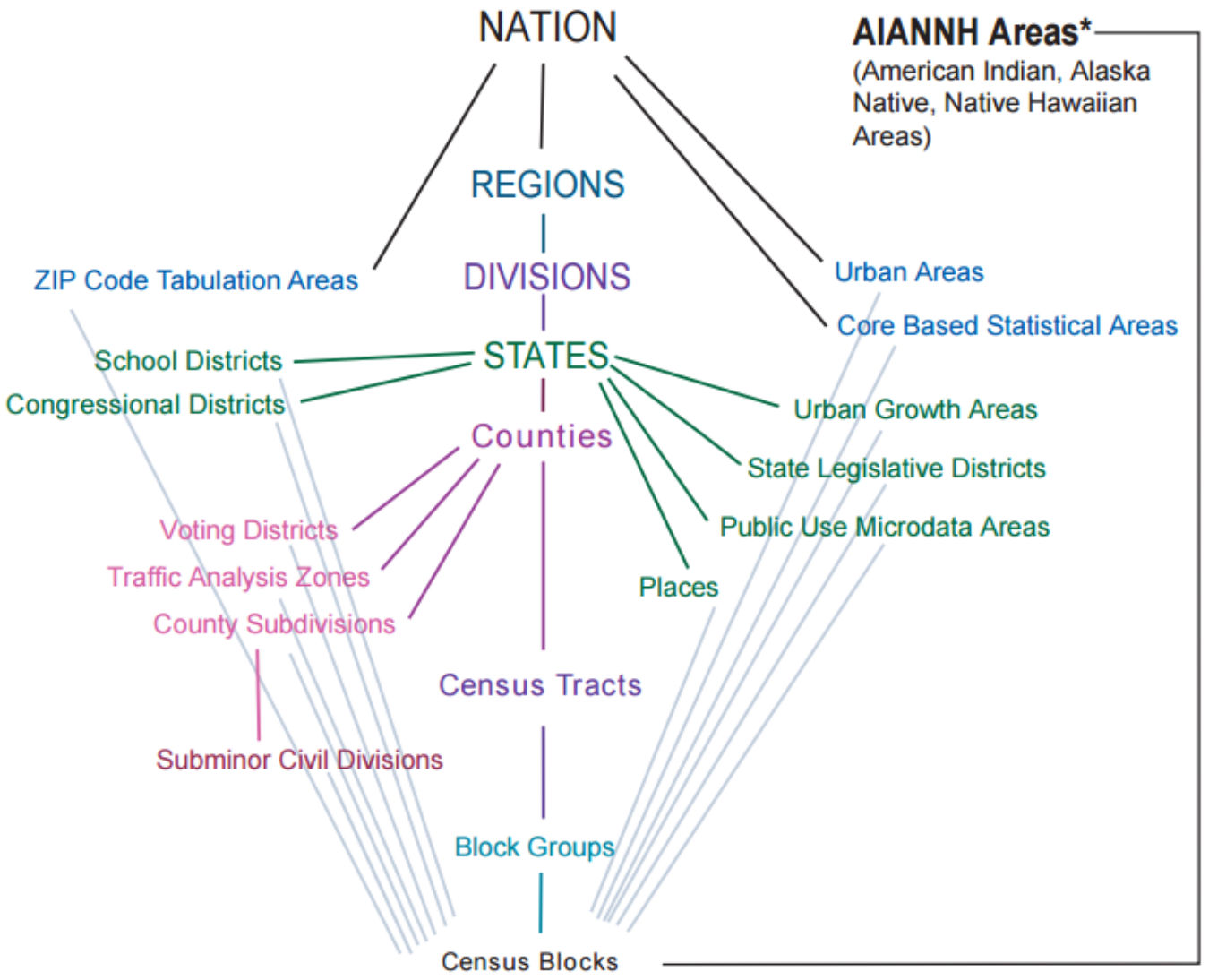
1. Spatial object: geographic representation of our area of interest
   - boundaries of administrative areas
   - where are they located? (CRS)
   - id/names for administrative areas
   - may be very large
2. Data to put on the map!
   - associated with each administrative area

Spatial Object

- Shapefile
- KML: from GoogleEarth/GoogleMaps
- GeoJSON
- and others!

- Create your own
- Download from online
  - GADM
  - Several levels of administrative areas

Standard Hierarchy of Census Geographic Entities



Spatial Object

- Most Important !!
  - Names of administrative areas must match your dataset!
  - This is how you will match your dataset to your spatial object
- Dataset gets merged to spatial object
  - when plotting, we assign a color for the area based on the value of the data variable you want to plot

Website References

- https://www.tidyverse.org/
- https://ggplot2.tidyverse.org/reference/
- https://www.r-graph-gallery.com/
- https://colorbrewer2.org/
- https://gadm.org/data.html
- https://www.rspatial.org/