

Data Visualization II



Maps:

Maps are used in a variety of fields to express data in an appealing and interpretive way. R provides a many of tools for creating different types of maps containing statistical information.

- Some of R graphics packages are used to create maps such as, ggplot2, leaflet, maps, mapdata, tmap, mapdeck, choroplethr, etc.



ggplot2 Package:

The `map_data()` function in the ggplot2 package which pull data for maps at different levels (“world”, “county”, “state”,).

The data give locations, longitude and latitude, of the borders of geographic polygons like counties and states.



Example 1:

1. Install `ggplot2` package.
2. Install `maps` package.
2. Display the data under `map_data("world")`

```
install.packages("ggplot2")  
library("ggplot2")
```

```
install.packages("maps")  
library("maps")
```

```
> map_data("world")
```

	long	lat	group	order	region	subregion
1	-6.989912e+01	1.245200e+01	1	1	Aruba	<NA>
2	-6.989571e+01	1.242300e+01	1	2	Aruba	<NA>
3	-6.994219e+01	1.243853e+01	1	3	Aruba	<NA>
4	-7.000415e+01	1.250049e+01	1	4	Aruba	<NA>



Example 1:

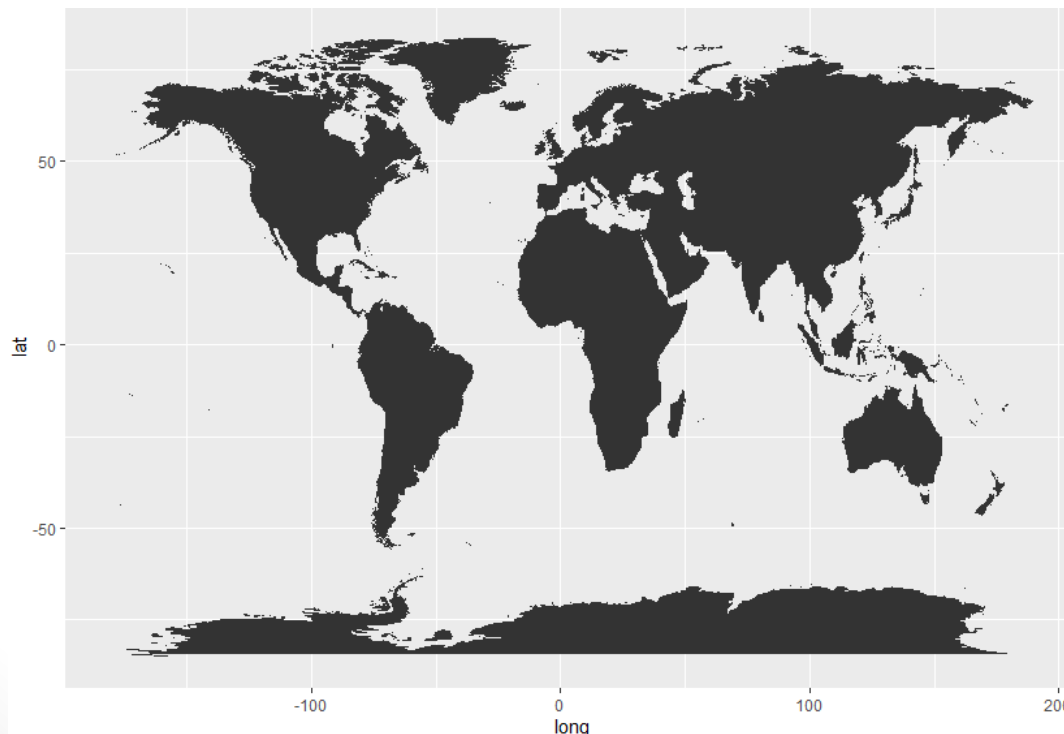
3. Display the location data for the world countries.

4. Graph the world map.

```
> world = map_data("world")  
> head(world)
```

	long	lat	group	order	region	subregion
1	-69.89912	12.45200	1	1	Aruba	<NA>
2	-69.89571	12.42300	1	2	Aruba	<NA>
3	-69.94219	12.43853	1	3	Aruba	<NA>
4	-70.00415	12.50049	1	4	Aruba	<NA>
5	-70.06612	12.54697	1	5	Aruba	<NA>
6	-70.05088	12.59707	1	6	Aruba	<NA>

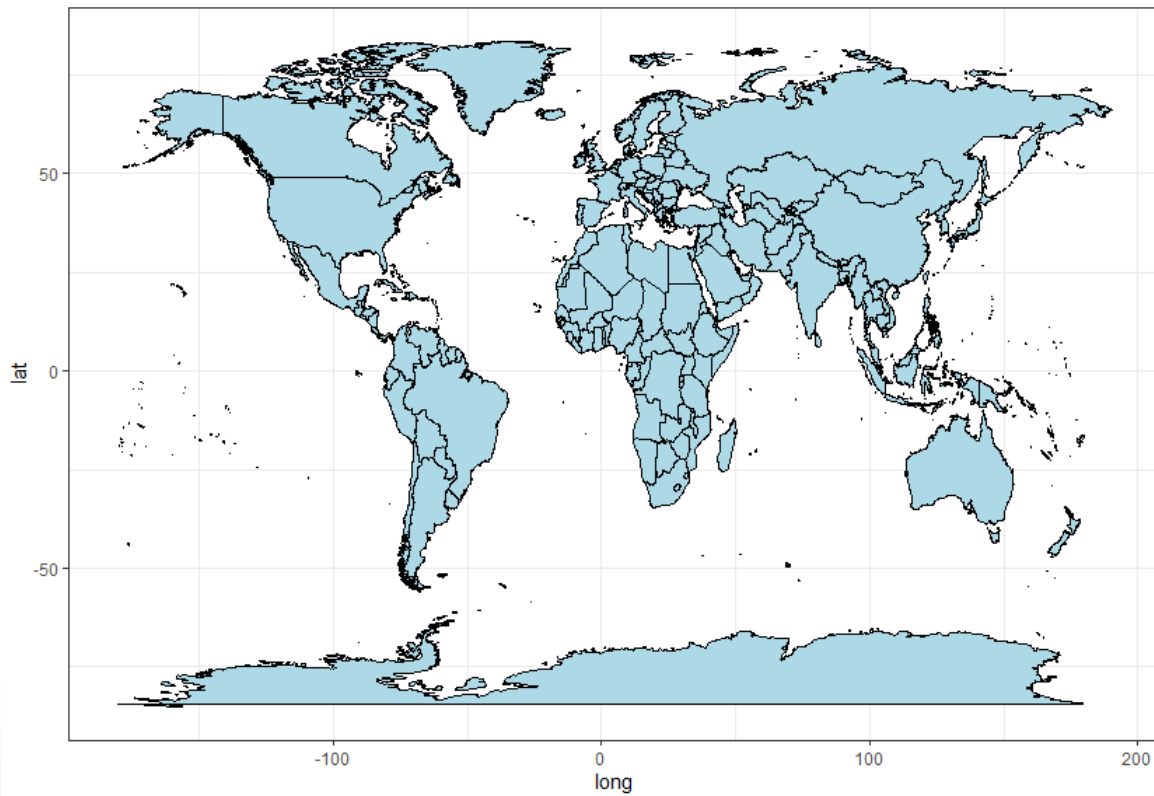
```
> ggplot(data = world, aes(x = long, y = lat, group = group)) +  
+   geom_polygon()  
.
```



Example 1:

5. Let's improve the quality of the world map by setting white background, present the countries border (in black), and fill the counties area with light blue.

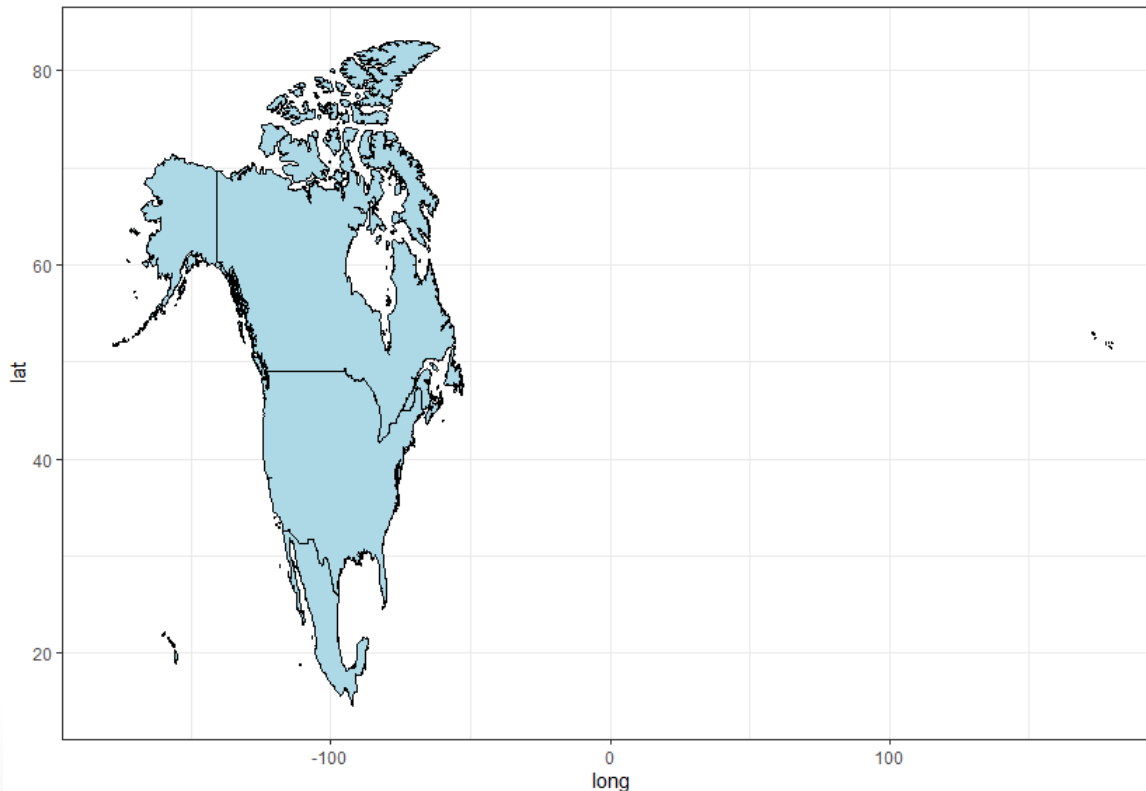
```
> ggplot(world, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_bw()
```



Example 1:

6. Graph the countries of North America.

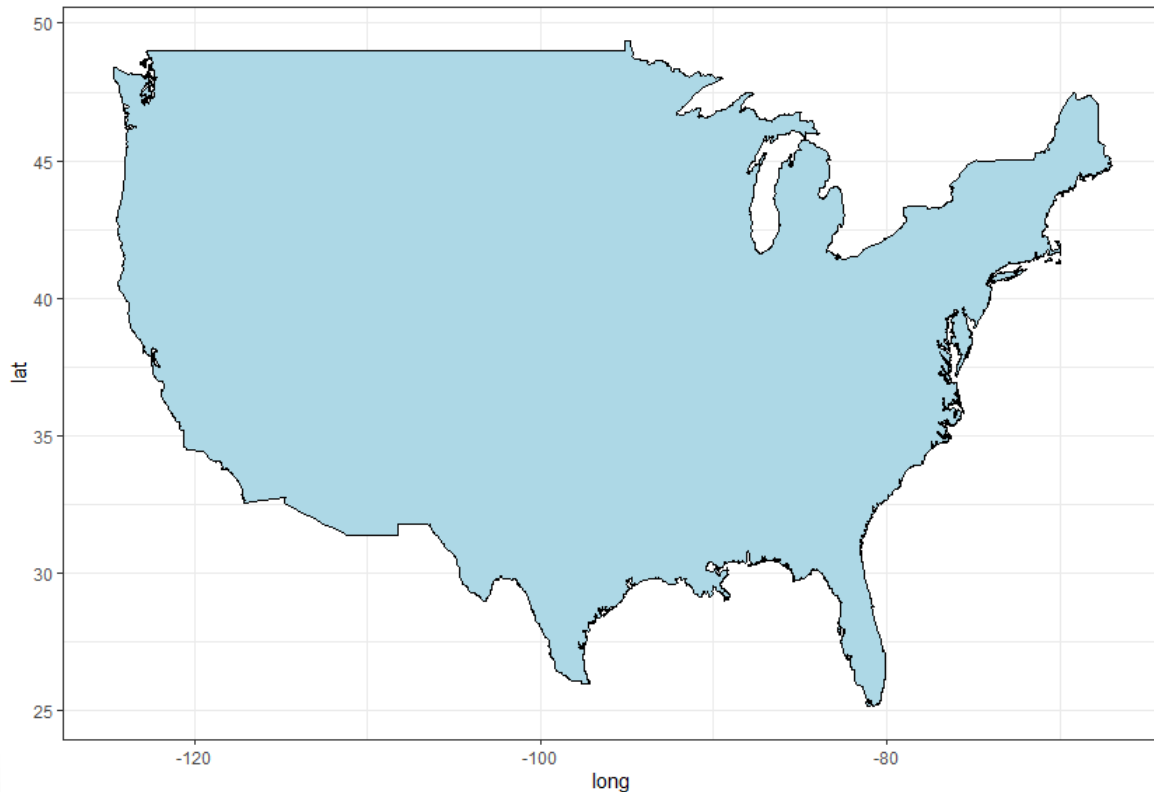
```
> Namerica <- map_data("world", region = c("canada", "usa", "mexico"))  
>  
> ggplot(data = Namerica, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_bw()
```



Example 1:

7. Graph the US map.

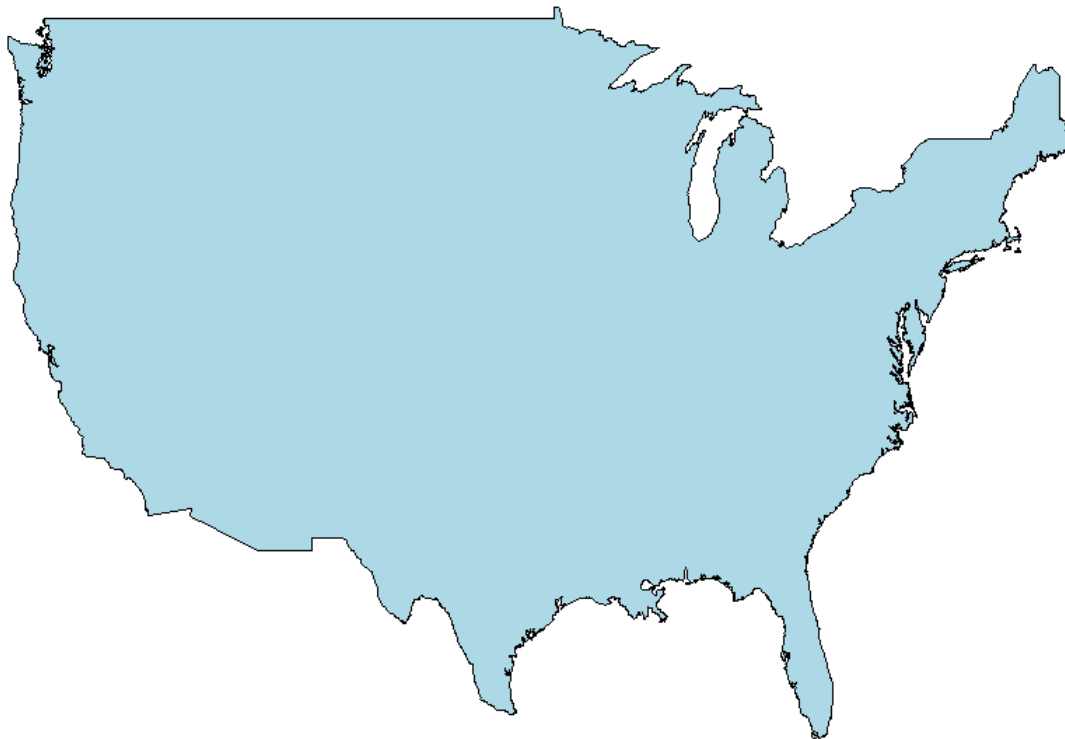
```
> usa = map_data("usa")  
>  
> ggplot(usa, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_bw()
```



Example 1:

8. We can remove all axis lines by using `theme_void()` function.

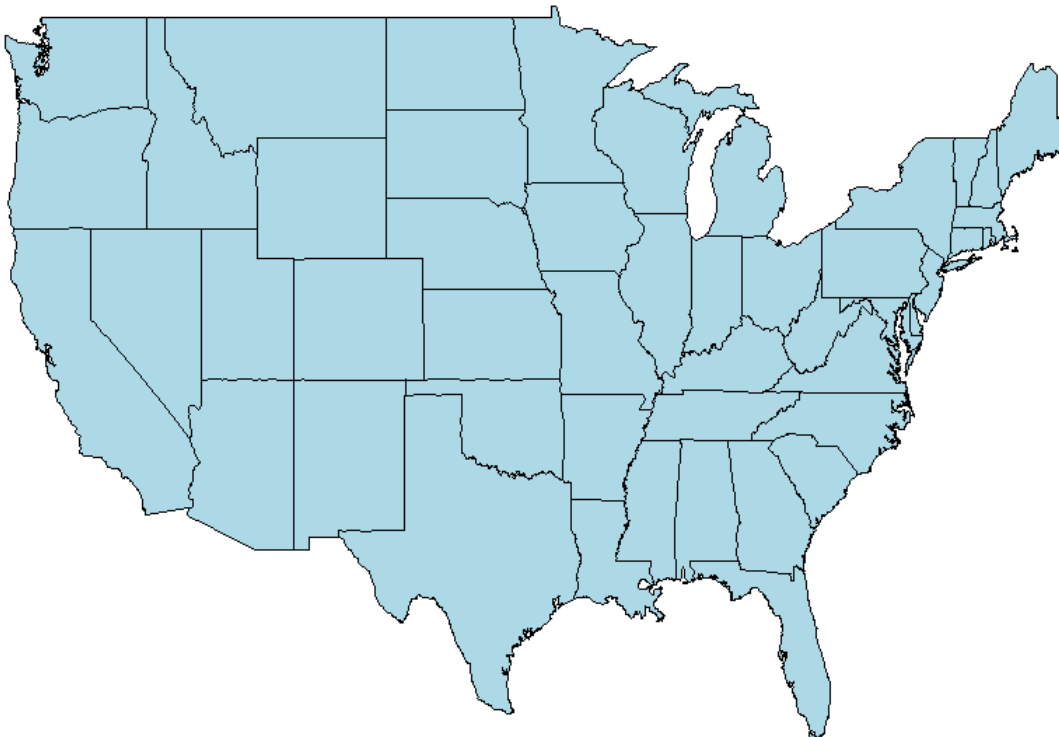
```
> ggplot(usa, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_void()
```



Example 1:

9. Graph map of the USA states.

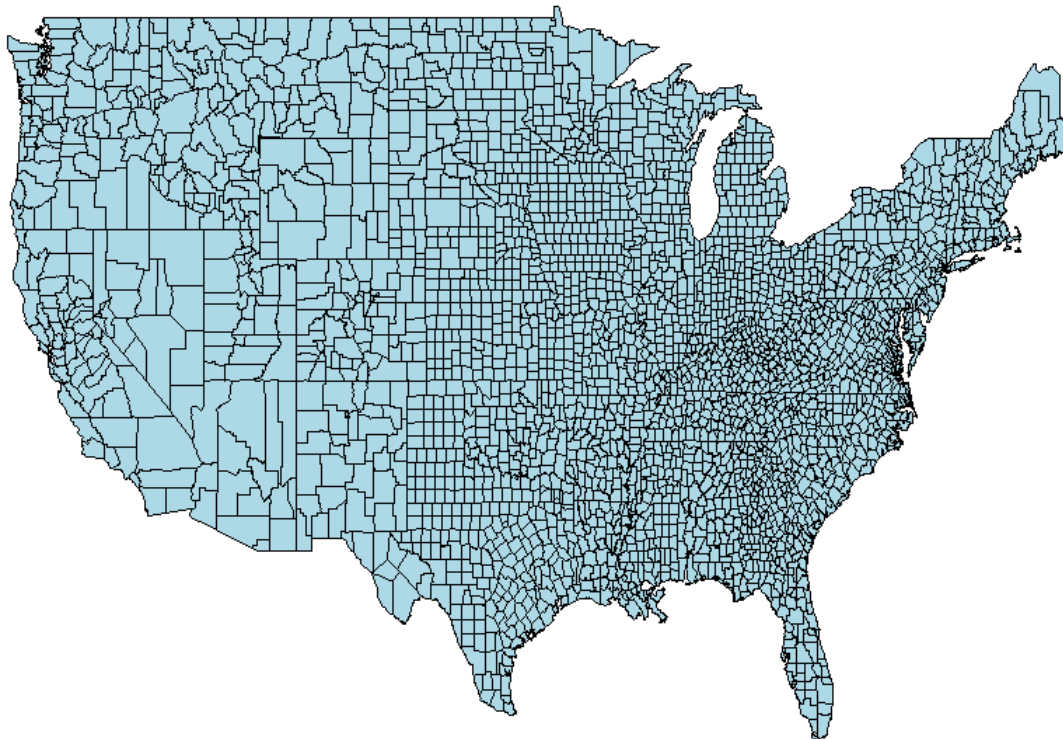
```
> us_states <- map_data("state")  
>  
> ggplot(us_states, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_void()
```



Example 1:

10. Graph map of the USA counties.

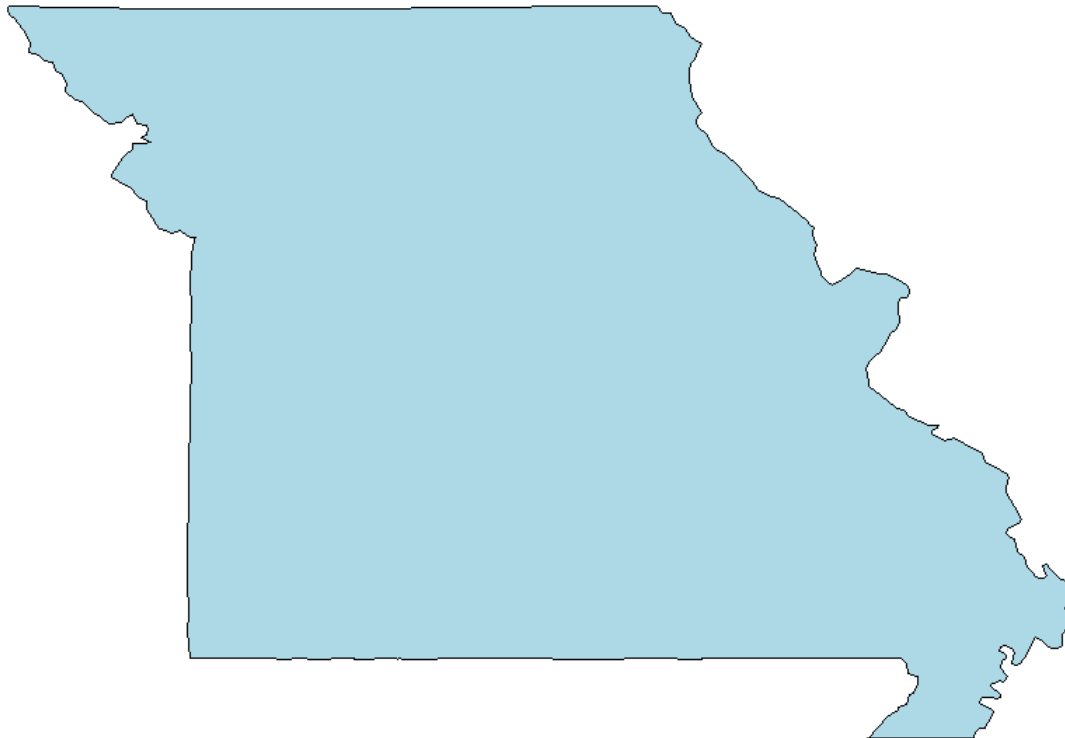
```
> US_counties <- map_data("county")  
>  
> ggplot(US_counties, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_void()
```



Example 1:

11. Graph the Missouri map (use `subset()` function).

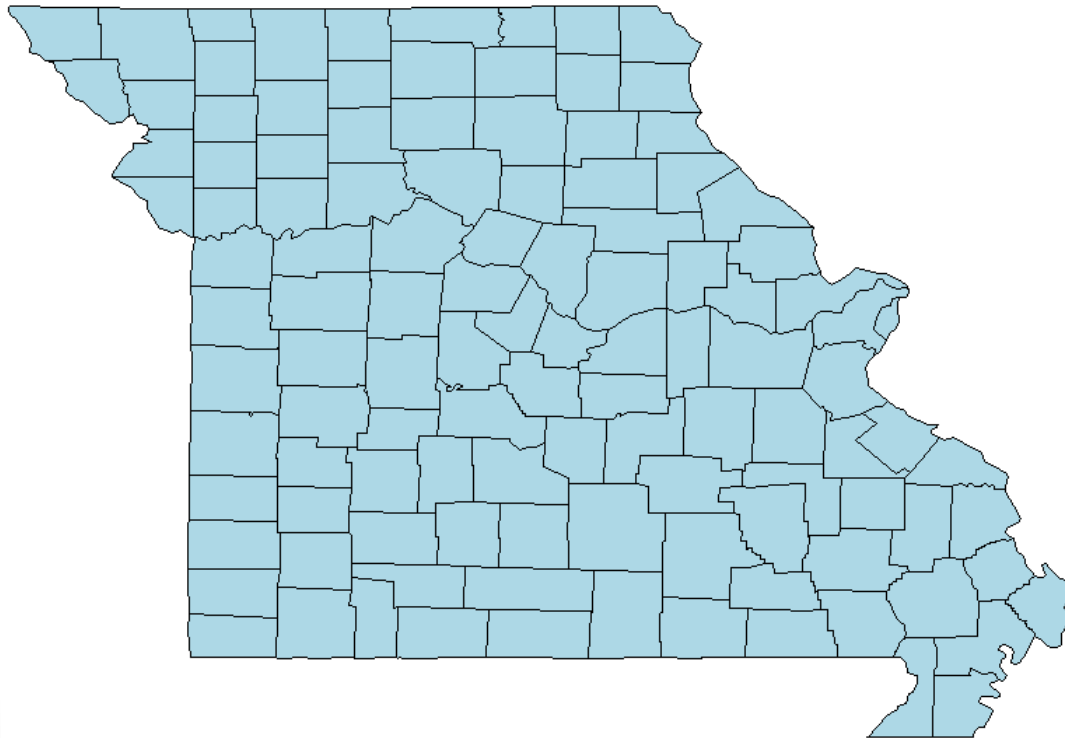
```
> Missouri <- subset(us_states, region=="missouri")  
>  
> ggplot(Missouri, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_void()
```



Example 1:

12. Graph the map of Missouri counties.

```
> Missouri_counties <- subset(US_counties, region=="missouri")  
>  
> ggplot(Missouri_counties, aes(x = long, y = lat, group = group)) +  
+   geom_polygon(fill = "lightblue", color = "black")+  
+   theme_void()
```



Example 2: (USArrests)

USArrests dataset contains statistics, in arrests per 100,000 residents for assault, murder, and rape in each of the 50 US states in 1973. Also given is the percent of the population living in urban areas.

1. Display the data dimension and head.

```
> dim(USArrests)
[1] 50  4
```

```
> head(USArrests)
```

	Murder	Assault	UrbanPop	Rape
Alabama	13.2	236	58	21.2
Alaska	10.0	263	48	44.5
Arizona	8.1	294	80	31.0
Arkansas	8.8	190	50	19.5
California	9.0	276	91	40.6
Colorado	7.9	204	78	38.7

- The first column is called row names, it doesn't have a label and doesn't count as an attribute.



Example 2: (USArrests)

2. Create a new dataset and add a new variable for the states.
3. Display the new data dimension and head.

```
> USArrests1 <-data.frame(state = tolower(rownames(USArrests)), USArrests)
```

```
> dim(USArrests1)  
[1] 50  5
```

```
> head(USArrests1)
```

	state	Murder	Assault	UrbanPop	Rape
Alabama	alabama	13.2	236	58	21.2
Alaska	alaska	10.0	263	48	44.5
Arizona	arizona	8.1	294	80	31.0
Arkansas	arkansas	8.8	190	50	19.5
California	california	9.0	276	91	40.6
Colorado	colorado	7.9	204	78	38.7



Example 2: (USArrests)

4. Create a new dataset which contains map details of the USA states.

```
> us_states <- map_data("state")
> head(us_states)
```

	long	lat	group	order	region	subregion
1	-87.46201	30.38968	1	1	alabama	<NA>
2	-87.48493	30.37249	1	2	alabama	<NA>
3	-87.52503	30.37249	1	3	alabama	<NA>
4	-87.53076	30.33239	1	4	alabama	<NA>
5	-87.57087	30.32665	1	5	alabama	<NA>
6	-87.58806	30.32665	1	6	alabama	<NA>



Example 2: (USArrests)

4. Graph the Choropleth map of the percent urban population (UrbanPop).

```
> ggplot(USArrests1, aes(map_id = state)) +  
+   geom_map(aes(fill = UrbanPop), map = us_states, col = "black") +  
+   expand_limits(x = us_states$long, y = us_states$lat)+  
+   theme_void()
```

