Hurricane Public Assistance

Jingwen Xu

2020/11/9

Data cleaning and organization

We only consider about the hurricane disaster during 2009-2018. So before mapping, I need to simply filter and clean the data.

```
disaster <- read.csv("PublicAssistanceFundedProjectsDetails.csv")

## Filter the data about hurricane
hurricane <- disaster %>% filter(incidentType=="Hurricane")
## unique(hurricane$incidentType)
hurricane %<>% select(-3)

## We only study the hurricane declared during 2009-2018.
hurricane$year <- substr(hurricane$declarationDate,1,4)
hurricane %<>% filter(year=="2009"|year=="2010"|year=="2011"|year=="2012"|year=="2013"|year=="2014"|yea

## summary(hurricane$totalObligated-hurricane$federalShareObligated) These two columns are totally the
## summary(hurricane)
## We found that there are negative values in "projectAmount", "federalShareObligated", "totalObligated
hurricane %<>% filter(projectAmount>=0&federalShareObligated>=0)
kable(head(hurricane)) %>% kable_styling(font_size=12)
```

disasterNumber	declarationDate	pwNumber	applicationTitle	applicantId	damageCa
1866	2009-12-22T05:00:00.000Z	1	DIW-097-02F	097-U15P3-00	F - Public
1866	2009-12-22T05:00:00.000Z	2	DIW-097-01F	097-U15P3-00	F - Public
1866	2009-12-22T05:00:00.000Z	3	DIW-097-03F	097-U15P3-00	F - Public
1866	2009-12-22T05:00:00.000Z	4	DIW-097-04F	097-U15P3-00	F - Public
1866	2009-12-22T05:00:00.000Z	5	DIW-097-01B	097-U15P3-00	B - Protec
1866	2009-12-22T05:00:00.000Z	6	FOL-01B	003-26992-00	B - Protec

Mapping

I map the data using two methods – ggplot2 and tmap. And there are three data information that I want to show on the map – the count of hurricane for each county, the total project amount and total federal obligated per county.

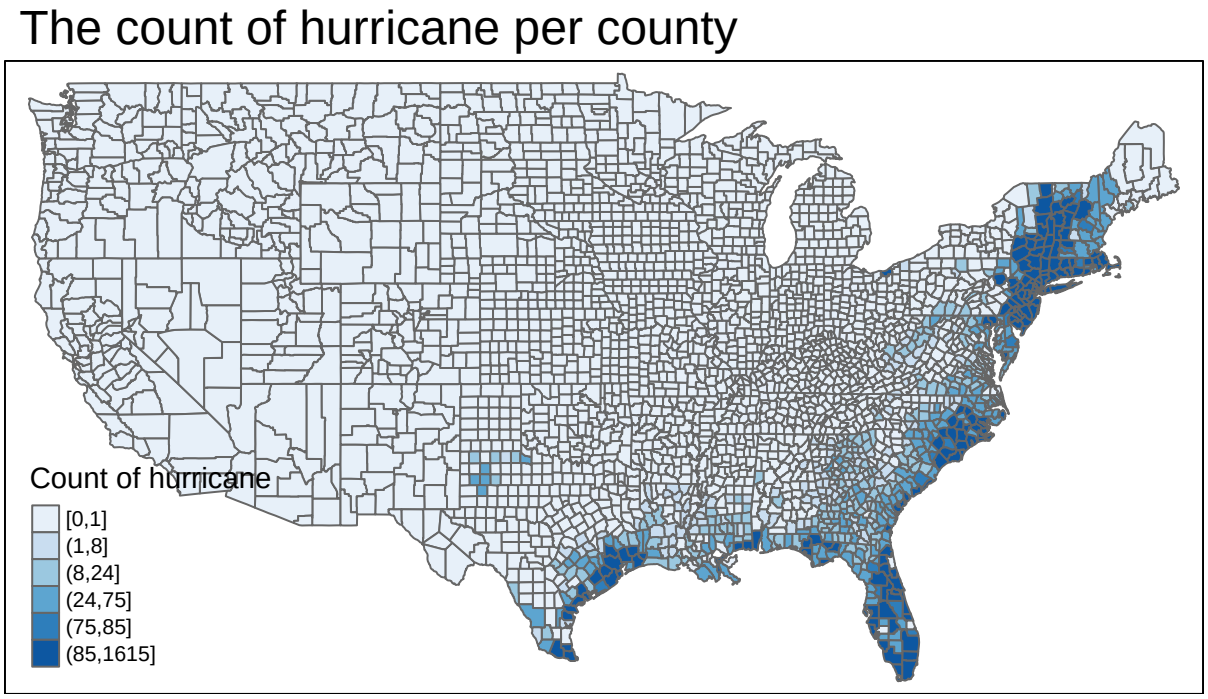
Mapping with tmap

```
hurricane$county <- tolower(hurricane$county)
hurricane$state <- tolower(hurricane$state)
```

```
Map <- st_as_sf(map('county',plot=F,fill=T))
Map%<>%separate(ID,c("state","county"),sep=",")
hurr_count <- hurricane %>% group_by(county,state) %>% summarise(Count=n())
```

```
## `summarise()` regrouping output by 'county' (override with `.groups` argument)
```

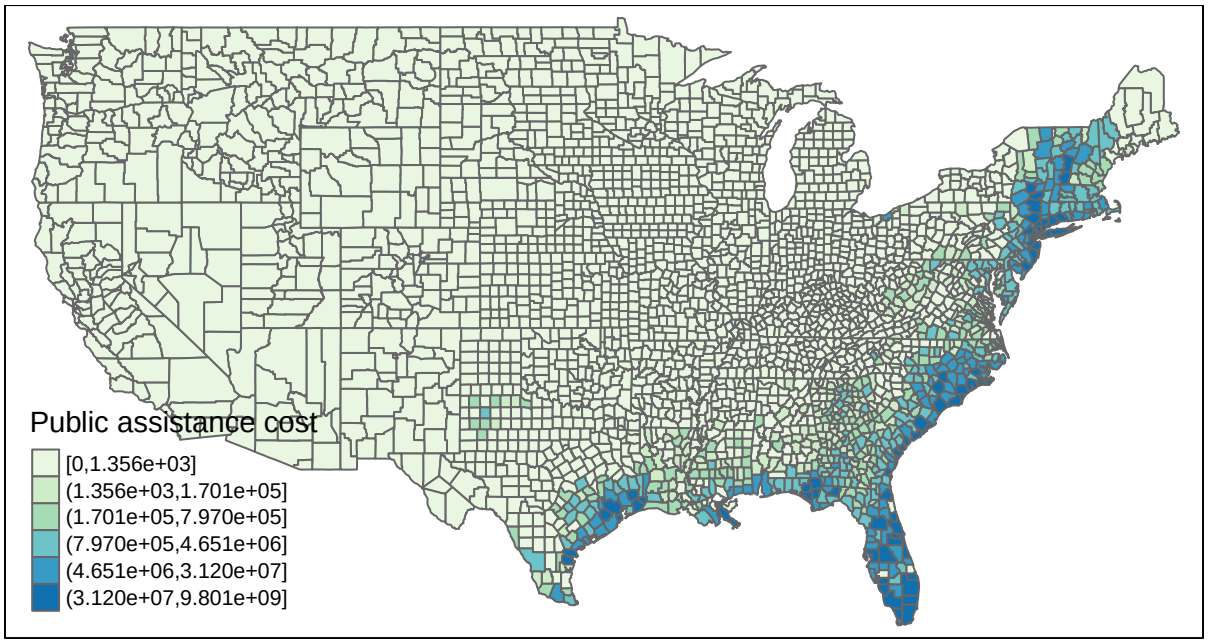
```
hurr_count_t <- left_join(Map,hurr_count,by=c("county","state"))
hurr_count_t$Count[is.na(hurr_count_t$Count)]=0.01
hurr_count_t$Count <- cut(hurr_count_t$Count,breaks=c(-1,1,8,24,75,85,1615),labels=c("[0,1]","(1,8]","(
tm_shape(hurr_count_t,title="The count of hurricane per county")+
  tm_polygons("Count",palette="Blues",title="Count of hurricane")+tm_layout(main.title='The count of hu
```




```
assist <- hurricane %>% group_by(county,state) %>% summarise(project_t=sum(projectAmount))
```

```
## `summarise()` regrouping output by 'county' (override with `.groups` argument)
```

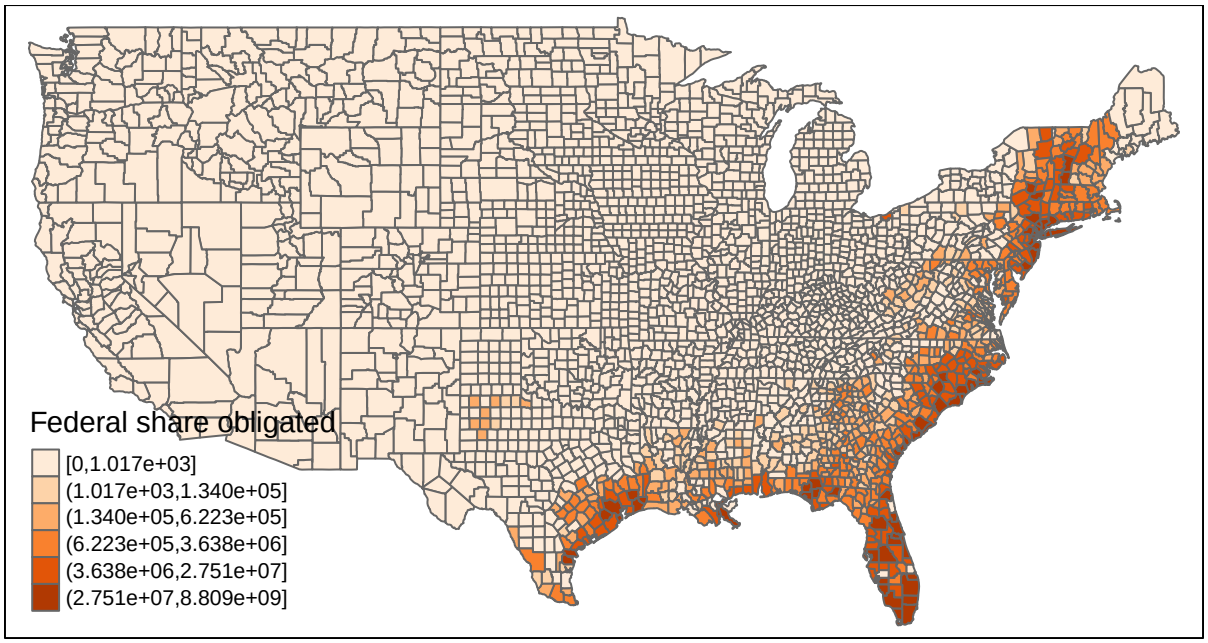
```
assist_t <- left_join(Map,assist,by=c("county","state"))
assist_t$project_t[is.na(assist_t$project_t)]=0.01
assist_t$project_t <- cut(assist_t$project_t,breaks=c(-1,1.356e+03,1.701e+05,7.970e+05,4.651e+06,3.120e
tm_shape(assist_t,title="The total public assistance cost per county")+tm_polygons("project_t",palette=
```




```
federal <- hurricane %>% group_by(county,state) %>% summarise(federal_total=sum(federalShareObligated))
```

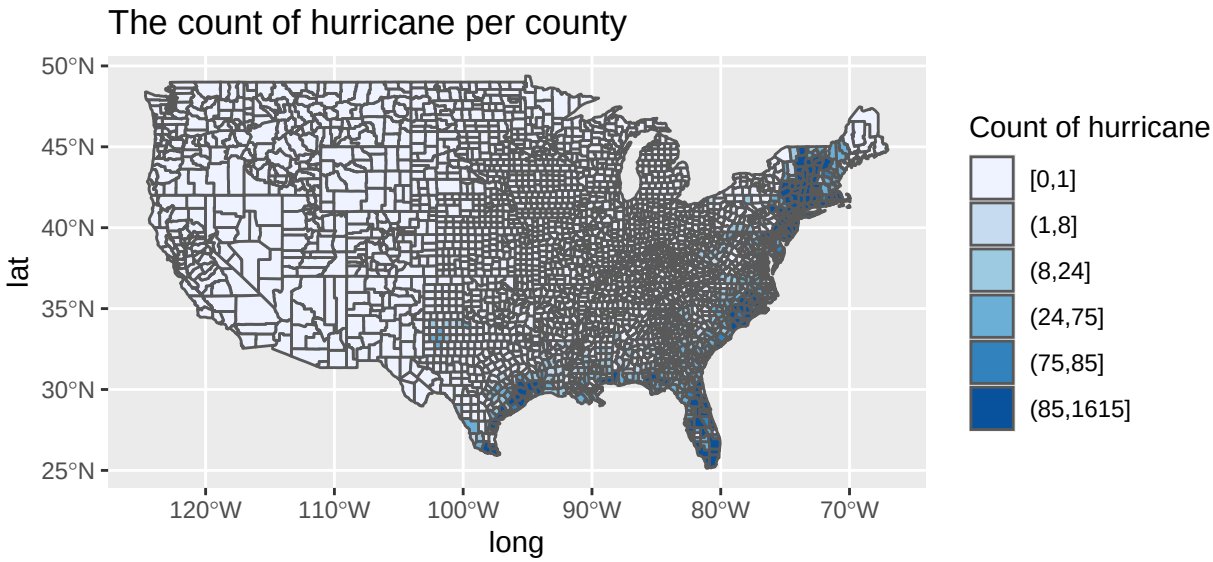
```
## `summarise()` regrouping output by 'county' (override with `.groups` argument)
```

```
federal_t <- left_join(Map,federal,by=c("county","state"))
federal_t$federal_total[is.na(federal_t$federal_total)]=0.01
federal_t$federal_total <- cut(federal_t$federal_total,breaks=c(-1,1.017e+03,1.340e+05,6.223e+05,3.638e
tm_shape(federal_t,title="The total Federal Share Obligated per county")+tm_polygons("federal_total",pal
```

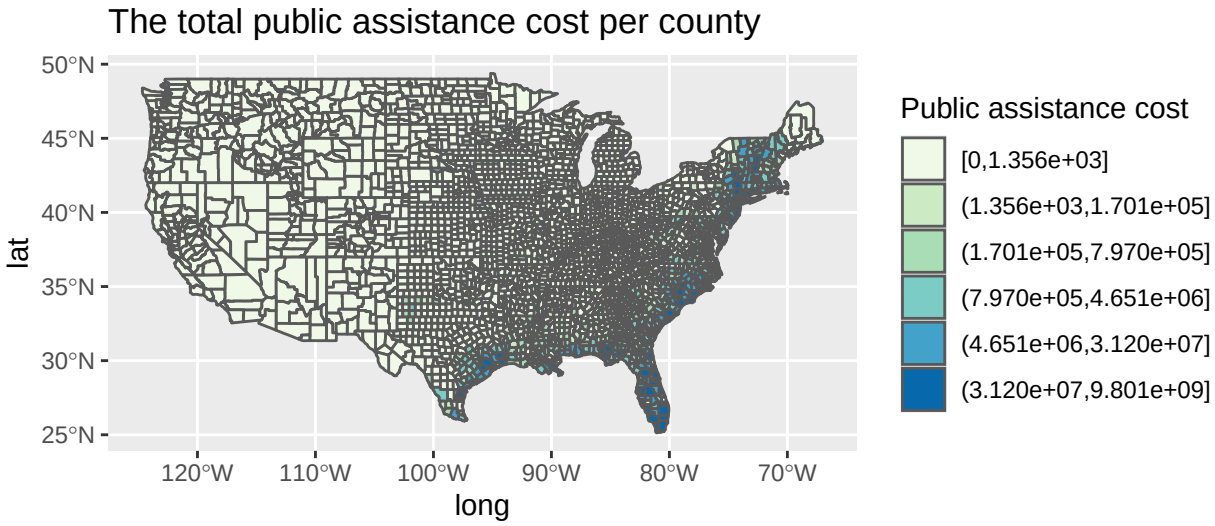


Mapping with ggplot

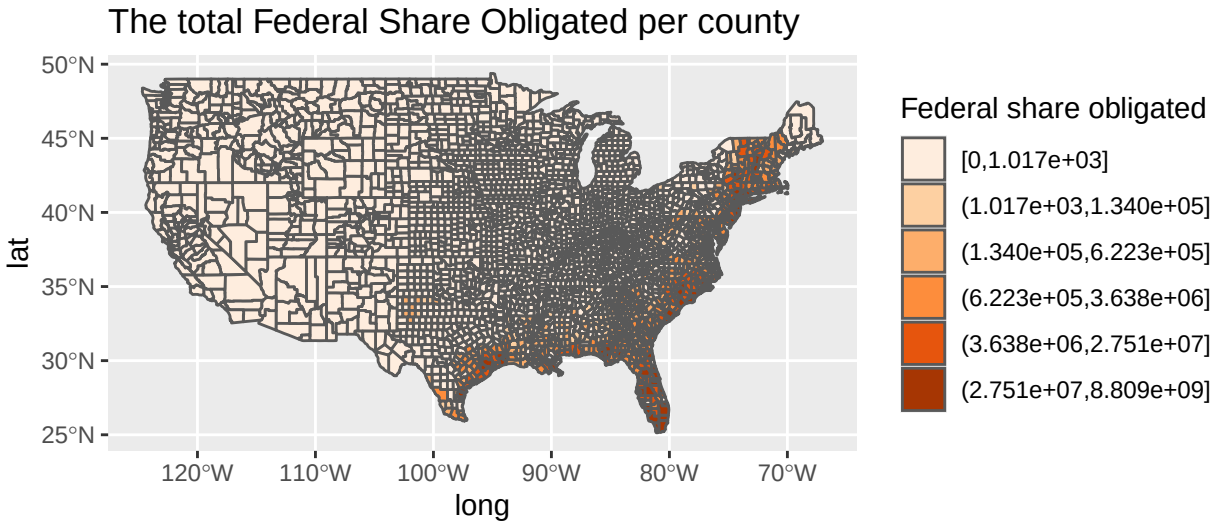
```
county <- map_data("county")
state <- map_data("state")
ggplot() + ggtitle("The count of hurricane per county")+
  geom_polygon(data=county, aes(x=long, y=lat, group=group),
               color="grey", fill="white", size = .2, alpha=0.9) +
  geom_polygon(data=state, aes(x=long, y=lat, group=group),
               color="black", fill="white", size = .2, alpha = .1)+
  geom_sf(data=hurr_count_t,mapping=aes(fill = Count))+
  scale_fill_brewer(name="Count of hurricane")
```




```
ggplot() + ggtitle("The total public assistance cost per county")+
  geom_polygon(data=county, aes(x=long, y=lat, group=group),
               color="grey", fill="white", size = .2, alpha=0.9) +
  geom_polygon(data=state, aes(x=long, y=lat, group=group),
               color="black", fill="white", size = .2, alpha = .1)+
  geom_sf(data=assist_t,mapping=aes(fill = project_t))+
  scale_fill_brewer(name="Public assistance cost",type="seq",palette=4)
```




```
ggplot() + ggtitle("The total Federal Share Obligated per county")+
  geom_polygon(data=county, aes(x=long, y=lat, group=group),
               color="grey", fill="white", size = .2, alpha=0.9) +
  geom_polygon(data=state, aes(x=long, y=lat, group=group),
               color="black", fill="white", size = .2, alpha = .1)+
  geom_sf(data=federal_t,mapping=aes(fill = federal_total))+
  scale_fill_brewer(name="Federal share obligated",type="seq",palette=7)
```



Conclusion

From the mappings for three data information, we can find that the distribution of count of hurricane, total project amount and total federal obligated are similar which is reasonable. The county with more hurricane need more public assistance and federal obligated.