Hurricane Data Analysis

Donghao Xue

2020/11/9

Introduction

The data contains information about hurricane events in the USA from the 2009 to 2018. we will mainly focus on analyzing the count of hurricane events in each county, the total public assistance cost in each county and the total Federal Share Obligated in each county. We plan to conconstruct mapping for these three problems by using ggplot and tmap seperately.

Data cleaning and organization

```
disaster <- read.csv("C:/Users/Daniel Xue/Desktop/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 s
## summary(hurricane)
## We found that there are negative values in "projectAmount", "federalShareObligated", "totalObligated
hurricane %<>% filter(projectAmount>=0&federalShareObligated>=0)
kable(head(hurricane)[,c(1,2,10,12,15)]) %>% kable_styling(font_size=12)
```

disasterNumber	declarationDate	county	state	projectAmount
1866	2009-12-22T05:00:00.000Z	Mobile	Alabama	0.00
1866	2009-12-22T05:00:00.000Z	Mobile	Alabama	58425.34
1866	2009-12-22T05:00:00.000Z	Mobile	Alabama	0.00
1866	2009-12-22T05:00:00.000Z	Mobile	Alabama	12778.47
1866	2009-12-22T05:00:00.000Z	Mobile	Alabama	15290.26
1866	2009-12-22T05:00:00.000Z	Baldwin	Alabama	9820.02

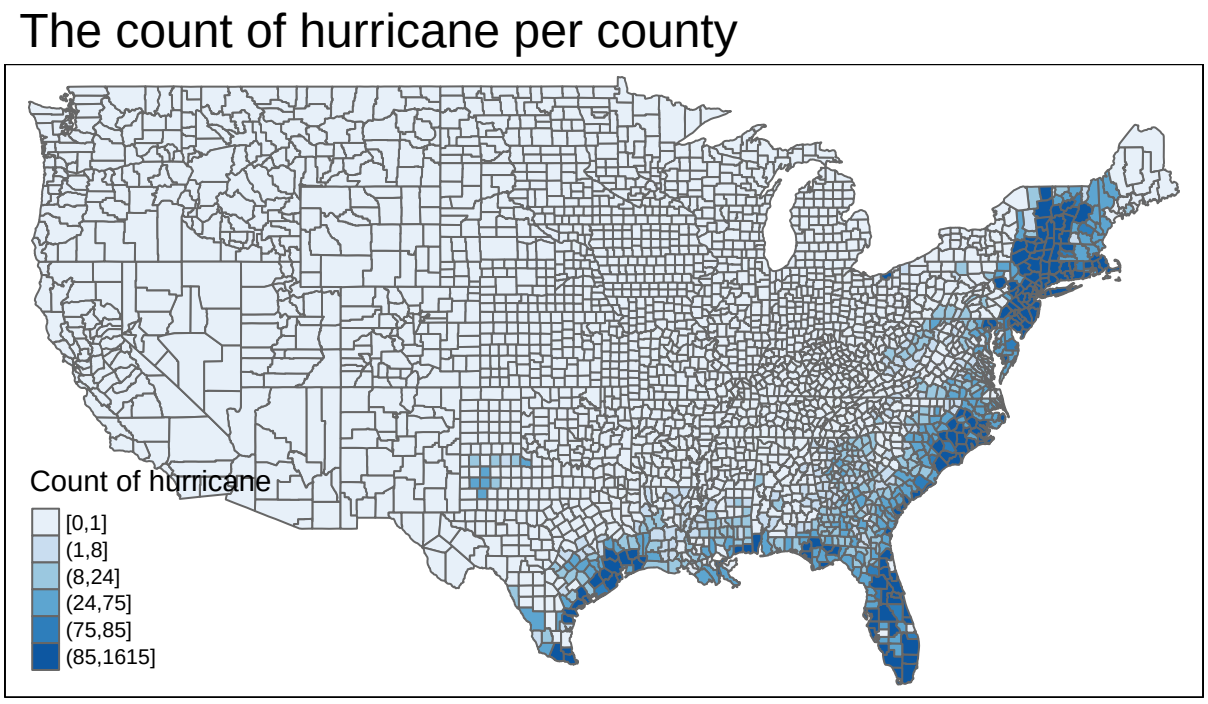
Mapping

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)
```

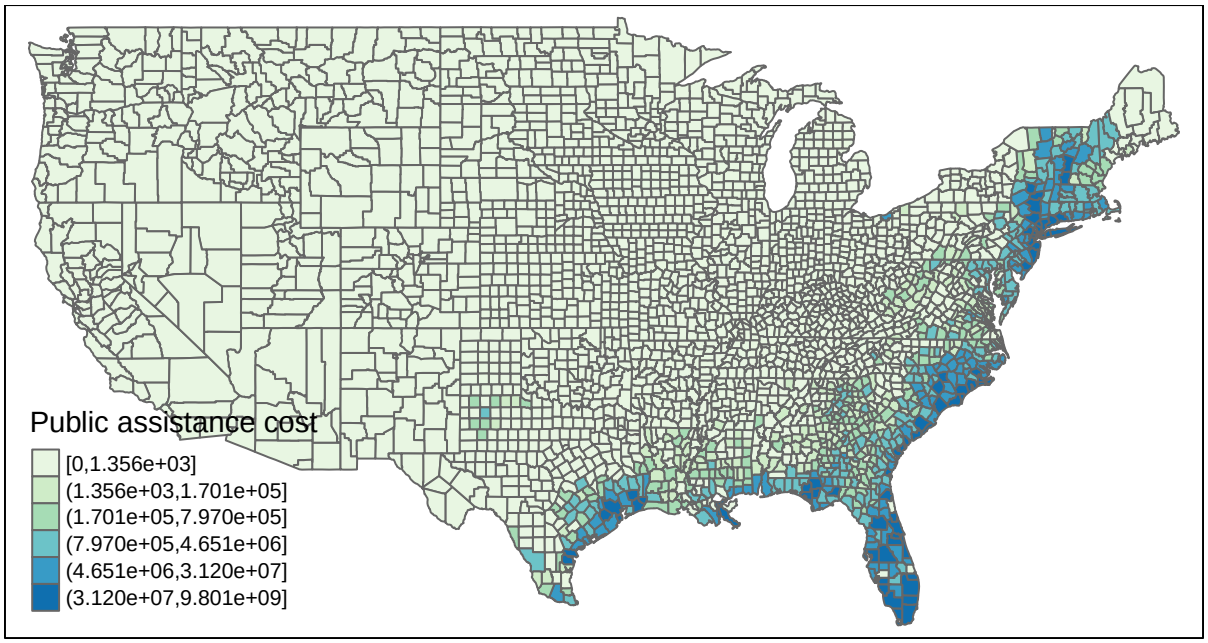
```
hurr_count_t <- left_join(Map,hurr_count,by=c("county","state"))
hurr_count_t$Count[is.na(hurr_count_t$Count)]=0.01
hurr_count_t$Count <- cut(hurr_count_t$Count,breaks=c(-1,1,8,24,75,85,1615),labels=c("[0,1]","(1,8]","(8
tm_shape(hurr_count_t,title="The count of hurricane per county")+
  tm_polygons("Count",palette="Blues",title="Count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

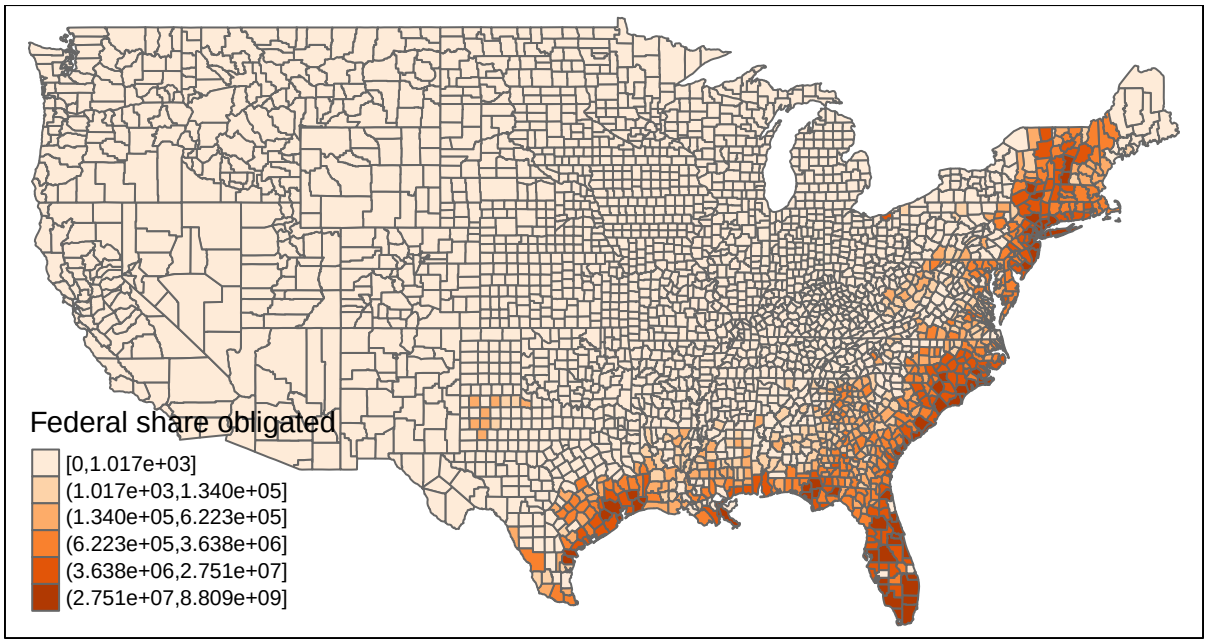
```
assist_t <- left_join(Map,assist,by=c("county","state"))
assist_t$project_t[is.na(assist_t$project_t)]=0.01
assist_t$project_t <- cut(assist_t$project_t,breaks=c(-1,1.356e+03,1.701e+05,7.970e+05,4.651e+06,3.120e
tm_shape(assist_t,title="The total public assistance cost per county")+tm_polygons("project_t",palette=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




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

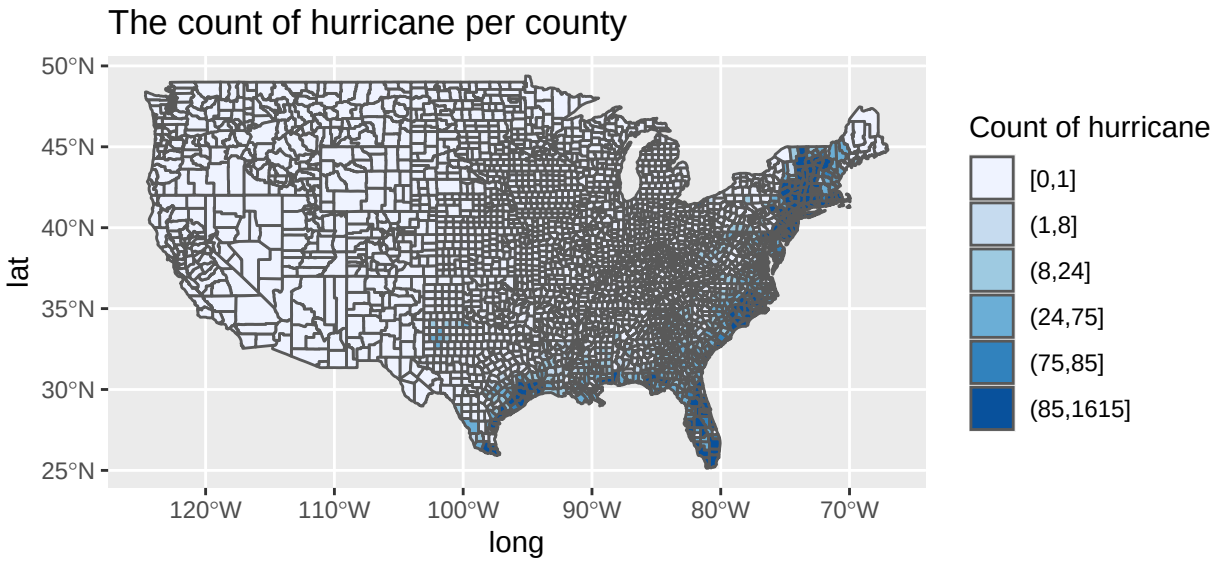
```
## `summarise()` regrouping output by 'county' (override with `.groups` argument)
```

```
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
labels=c("[0,1.017e+03]","(1.017e+03,1.340e+05]","(1.340e+05,6.223e+05]","(6.223e+05,3.638e+06]","(3.63
tm_shape(federal_t,title="The total Federal Share Obligated per county")+tm_polygons("federal_total",pa
```

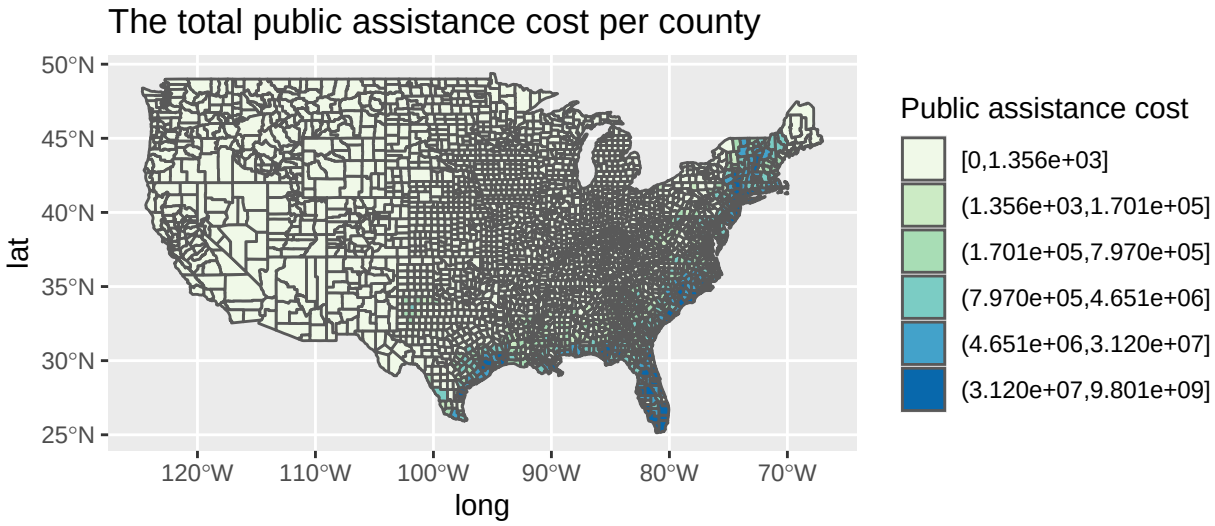


Mapping with ggplot

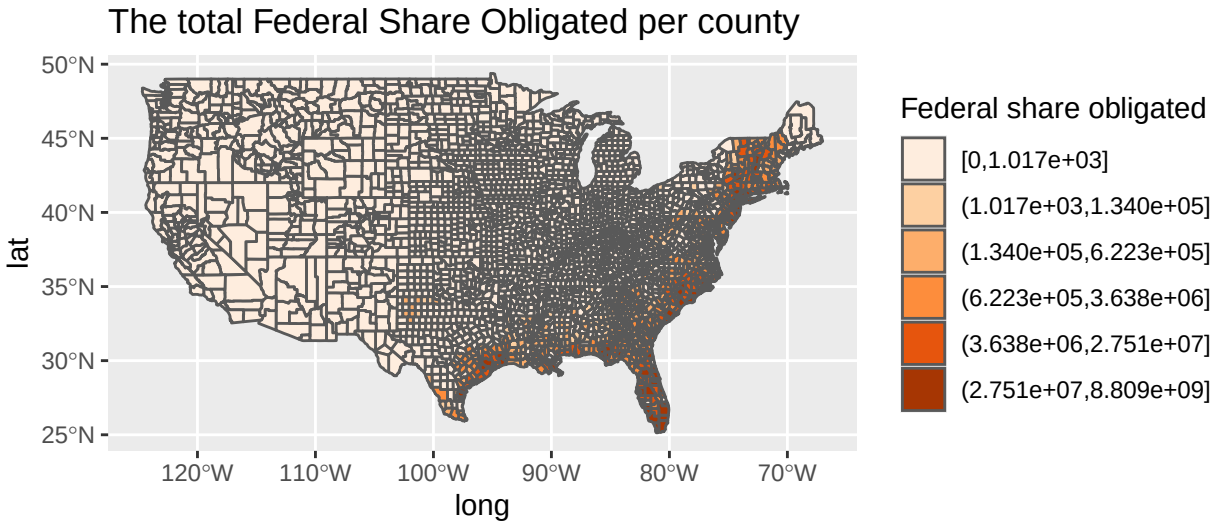
```
county <- map_data("county")
state <- map_data("state")
ggplot() + ggtitle("The count of hurricane per county")+
  geom_polygon(data=county, aes(x=long, y=lat, group=group),
               color="grey", fill="white", size = .2, alpha=0.9) +
  geom_polygon(data=state, aes(x=long, y=lat, group=group),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



Discussion

From the plots we can see that hurricane mostly happens in the east coast of USA, which leads to the cost of public assistance and Federal Share also focusing on the east coast.

Reference

The data is collected from: https://www.fema.gov/openfema-data-page/public-assistance-funded-projects-details-v1