Final Project

Jingwen Zhang

1. Project Area: Residential Fires with Fatalities

2. Project Goals:

A. Creating data that can provide information about whether there are any changes in the number of fire incidents for each states during the 4 years. Make the comparison between states and years possible

B. Creating tables that shows the address and causes for each fire incidents with fatalities, that may help identify what is the most common cause for the residential fires during the 4 years and in what states.

C. Provide graphic visualizations for the residential fires with fatalities in the U.S. And group them by causes.

3. Source from: Diane Mclaughlin's Google Drive.

4. Files (Data from 2012-2015):

basicincident.txt
causes.txt
Incidentaddress.txt

5. Folder contains:

1. **Project Summary**
2. Code 1.R
   (Code I used to generate new csv files that I will use in my **Code 2** and **Code 3.** It takes a really long time to run)
3. **Code 2.R** (Code for bar graph and table1)
4. **Code 3.R** (Code for map and table2)
5. **new data** folder (all csv files that are created by myself)
6. **plot** folder (all plots and tables created by the codes)
7. 2012-2015 folders. (the original data)

* Code and csv files that are used to produce shiny deliverables are also in the **final 1** folder, which is in the **new data** folder.

6. Final project **URL**:

https://zjwlucy.shinyapps.io/final_project2/

In my final project, I only use the data from 2012 to 2015 because start from 2012, the way how they record the data changes. Also, the address and other things may not be so accurate from 2006 to 2011. So comparison between years among 2006-2011 and 2012-2015 may be misleading. Only years from

2012 to 2015 are taken into consideration in this project. Starting from 2012, they have an additional text file called “causes,” which records the cause for each fire incident. It might be helpful to plot the incidents on the map and by the causes so that we can identify the most common cause for residential fires in a specific area.

Code 1

In code 1, I read all the text files and select the variables that I need, which include: STATE, FDID, INC_DATE, EXP_NO, INC_NO, PROP_USE, OTH_DEATH, FF_DEATH. According to the guide for the data sets, the first 4 variables are the unique features I can use when I combine the tables. I combined “CAUSE” and “incidentaddress” to the “basicincident” table for each year. Then I combine the main tables of 4 years together and get addresstable.csv for Code 3. I also create 4 csv files, which I will use in code 2. More details are in the Code 1‘s comments.

Code 2

Code 2 contains the code that creates **plot 1** and **table 1** on the URL

In code 2, I check whether the count for number of death is 0. Then, in order to find the incidents that are only related to **residential fires with fatalities**, I use the sqldf package. From the Guide for the data, I know that the residential fires can be identified by the "PROP_USE" column. The "PROP_USE" for residential fires are the 400 series, which means the value for "PROP_USE" is between 400 and 500. Also, I need the incidents with non-zero death for both “FF_DEATH” and “OTH_DEATH.” So for multiple conditional selection, sqldf works better than filter function in dplyr package. I also divided the states into 4 regions of the US. More details are on the URL.

Code 3

Code 3 contains the code that creates **plot 2** and **table 2** on the URL.

In Code 3, I replace the cause numbers by their corresponding causes for the fire incidents. Also, I rearrange the columns then create the table and save it as table 2 for my shiny app. I only show the first 15 rows of my table because the table is too long. Another thing I did is to get the longitudes and latitudes for each of the observations. There are 6200 observations in my data but geocode() has a limit of 2500 everyday. I run out of my limit so I only get the locations for 2298 of them. I commented the code about how I get the longitudes and latitudes because it takes a long time to run. After I get the locations, I add them to table 2 and use this new table to plot on my map.

The points on the map is the count for the number of fire incidents with fatalities in that specific city. The larger the size of the points, the bigger the number is. In Code 3 I choose a sample size of 500. On the shiny URL, other sample sizes are possible by moving the slider on the left side.

Plots and **tables** created by the code. (More details are on the URL.)

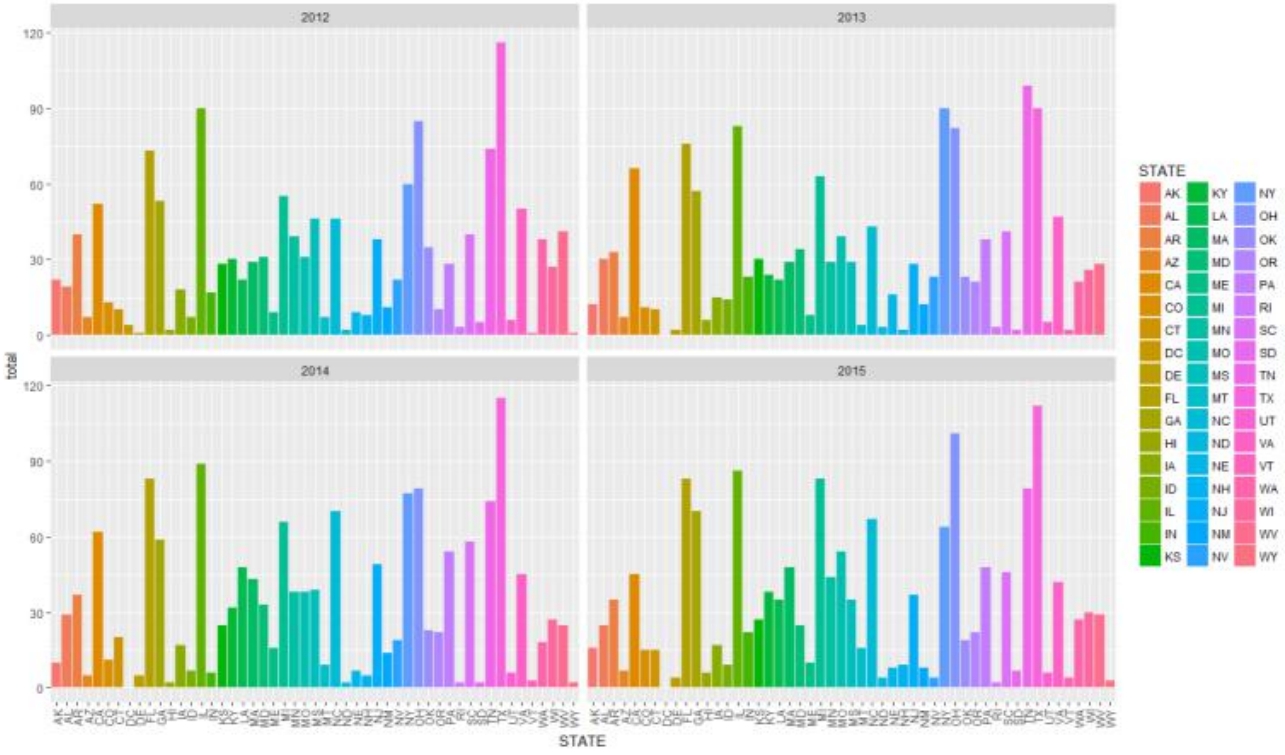
Table 1

year	Other	Firefighter	Total
2012	1507	4	1511
2013	1501	0	1501
2014	1627	0	1627
2015	1642	6	1648

Table 2

IncidentNO	year	State	City	ZIP	Street	Other	Firefighter	Cause
0000001	2012	AK	Saint Marys	99658	Airguard	1	0	Open Flame
0000001	2012	AK	Alakanuk	99554	Angela Smiths Residence	1	0	Intentional
0001719	2012	AK	Anchorage	99508	42nd	1	0	Investigation with Arson Module
0016530	2012	AK	Anchorage	99508	Mc Carrey	1	0	Intentional
0000001	2012	AK	Copper Center	99573	Richardson	1	0	Unknown
0002137	2012	AK	WASILLA	99654	N COLONIAL DR	1	0	Other Heat
0043841	2012	AK	TRAPPER CREEK	99683	E PETERSVILLE RD	1	0	Unknown
0012260	2012	AK	WILLOW	99688	N DESHKA LOOP	1	0	Unknown
0000451	2012	AK	Soldotna	99669	King Salmon	1	0	Smoking
0000001	2012	AK	Hope	99605	Ressurection	1	0	Unknown
SFM-001	2012	AK	Dillingham	99576	Emperor	1	0	Unknown
0000002	2012	AK	Naknek	99633	Russian Orthadox Church	1	0	Other Equipment
0100506	2012	AK	Fairbanks	99701	EIGHTEENTH	2	0	Other Unintentional, Careless
0100725	2012	AK	Fairbanks	99701	LAKEVIEW TERRACE	1	0	Smoking
0000455	2012	AK	NORTH POLE	99705	ROZAK	1	0	Unknown

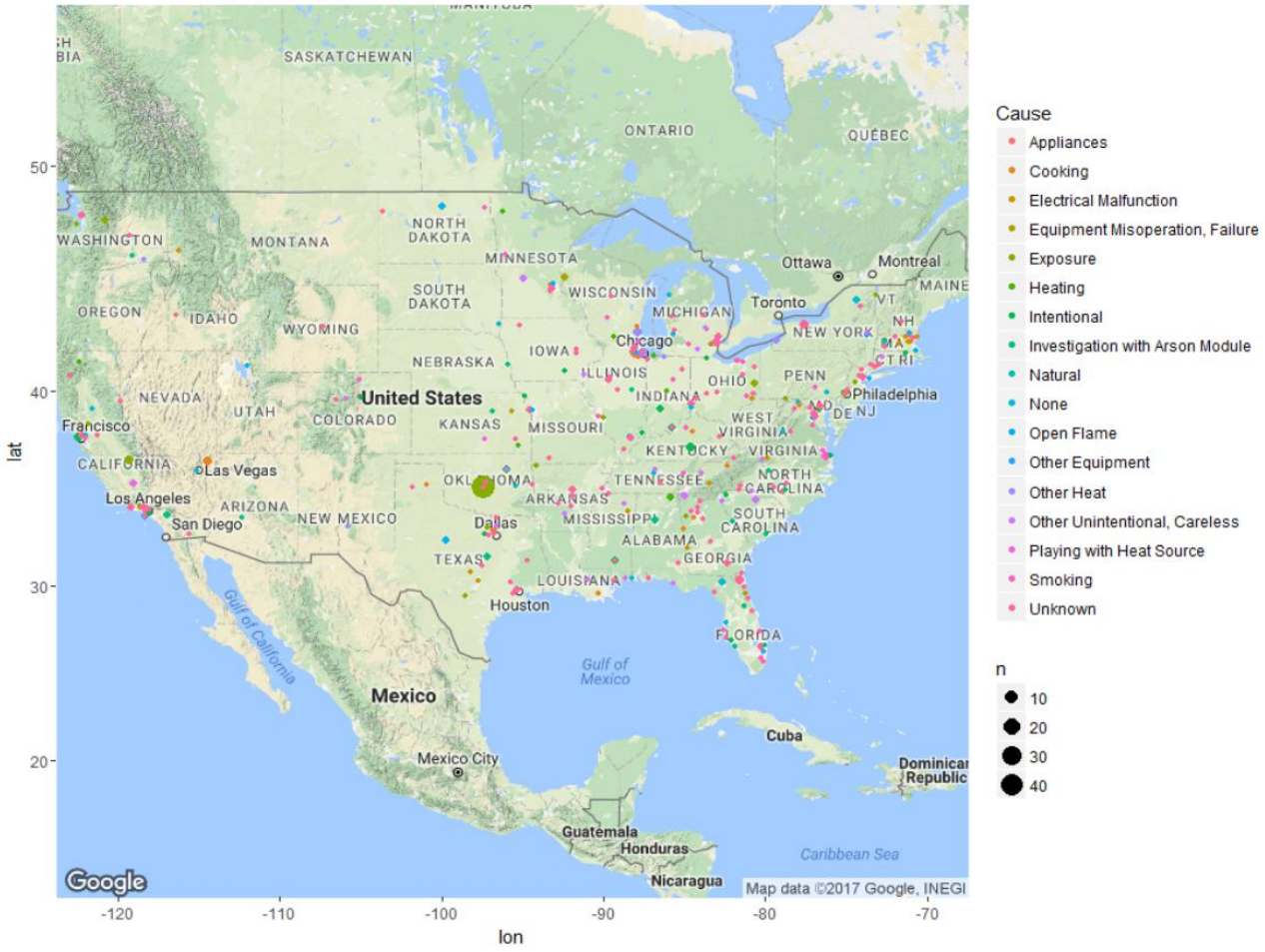
The Plot that is not on the URL but might be helpful



Plot 1


Number of Residential Fires with Fatalities
region
midwest
northeast
south
west
2012
2013
2014
2015
STATE

Plot 2


Cause
Appliances
Cooking
Electrical Malfunction
Equipment Misoperation, Failure
Exposure
Heating
Intentional
Investigation with Arson Module
Natural
None
Open Flame
Other Equipment
Other Heat
Other Unintentional, Careless
Playing with Heat Source
Smoking
Unknown
n
10
20
30
40
lat
lon
United States
Mexico
Google
Map data ©2017 Google, INEGI

Plot 2 (Watercolor Version)