STA160-Final Project_Plots_Cindy

May 17, 2022

[185]:
```
import pandas as pd
import seaborn as sns
import plotnine as p9
import matplotlib.pyplot as plt
import folium
from folium.plugins import MarkerCluster
import sqlite3
```

[186]:
```
# connect to db
#con = sqlite3.connect('FPA_FOD_20170508.sqlite')
#cursorObj = con.cursor()
#cursorObj.execute("SELECT␣
 ↪FPA_ID,FIRE_CODE,FIRE_NAME,FIRE_YEAR,date(DISCOVERY_DATE),DISCOVERY_TIME,STAT_CAUSE_CODE,ST
 ↪FROM Fires where STATE='CA' LIMIT 10 ")
#cursorObj.execute('PRAGMA table_info(Fires)')
#cursorObj.fetchall()
```

[186]:
```
<sqlite3.Cursor at 0x7fc3747b7960>
```

[20]:
```
#convert to csv
#db_df = pd.read_sql_query("SELECT␣
 ↪FPA_ID,FIRE_CODE,FIRE_NAME,FIRE_YEAR,date(DISCOVERY_DATE),DISCOVERY_TIME,STAT_CAUSE_CODE,ST
 ↪FROM Fires where STATE='CA' ", con)
#db_df.to_csv('Fire_db.csv', index=False)
```

Variables:

FPA_ID = Unique identifier that contains information necessary to track back to the original record in the source dataset.

FIRE_CODE = Code used within the interagency wildland fire community to track and compile cost information for emergency fire.

FIRE_NAME = Name of the incident, from the fire report (primary) or ICS-209 report (secondary).

FIRE_YEAR = Calendar year in which the fire was discovered or confirmed to exist.

DISCOVERY_DATE = Date on which the fire was discovered or confirmed to exist.

DISCOVERY_TIME = Time of day that the fire was discovered or confirmed to exist.

STAT_CAUSE_CODE = Code for the (statistical) cause of the fire.

STAT_CAUSE_DESCR = Description of the (statistical) cause of the fire.

CONT_DATE= Date on which the fire was declared contained or otherwise controlled (mm/dd/yyyy where mm=month, dd=day, and yyyy=year).

FIRE_SIZE =Estimate of acres within the final perimeter of the fire.

FIRE_SIZE_CLASS= Code for fire size based on the number of acres within the final fire perimeter expenditures (A=greater than 0 but less than or equal to 0.25 acres, B=0.26-9.9 acres, C=10.0-99.9 acres, D=100-299 acres, E=300 to 999 acres, F=1000 to 4999 acres, and G=5000+ acres).

LATITUDE =Latitude (NAD83) for point location of the fire (decimal degrees).

LONGITUDE =Longitude (NAD83) for point location of the fire (decimal degrees).

FIPS_CODE =Three-digit code from the Federal Information Process Standards (FIPS) publication 6-4 for representation of counties and equivalent entities.

FIPS_NAME =County name from the FIPS publication 6-4 for representation of counties and equivalent entities.

State = Two-letter code for the state in which the unit is located (or primarily affiliated).

[174]:
```
#load dataset
firedb=pd.read_csv('Fire_db.csv')

#drop all na values
df_CA=firedb.dropna()
df_CA
```

```
/Users/chang/opt/anaconda3/lib/python3.8/site-
packages/IPython/core/interactiveshell.py:3165: DtypeWarning: Columns (1,8,14)
have mixed types.Specify dtype option on import or set low_memory=False.
```

[174]:

	FPA_ID	FIRE_CODE	FIRE_NAME	FIRE_YEAR	\
0	FS-1418826	BJ8K	FOUNTAIN	2005	
1	FS-1418827	AAC0	PIGEON	2004	
2	FS-1418835	A32W	SLACK	2004	
17	FS-1418881	BHA3	FREDS	2004	
23	FS-1419089	BEZ8	HOT SPRINGS	2005	
…	…	…	…	…	
177324	FWS-2015USCASWRJ42X	J42X	POOL 11	2015	
177325	FWS-2015USCASJRJ31F	J31F	ALFALFA	2015	
177326	FWS-2015USCALURJH8F	JH8F	PARKING LOT 3	2015	
177329	FWS-2015USCAMDRJ3UT	J3UT	REFUGE	2015	
177330	FWS-2015USCATKRJP41	JP41	HILL	2015	

	date(DISCOVERY_DATE)	DISCOVERY_TIME	STAT_CAUSE_CODE	\
0	2005-02-02	1300.0	9.0	

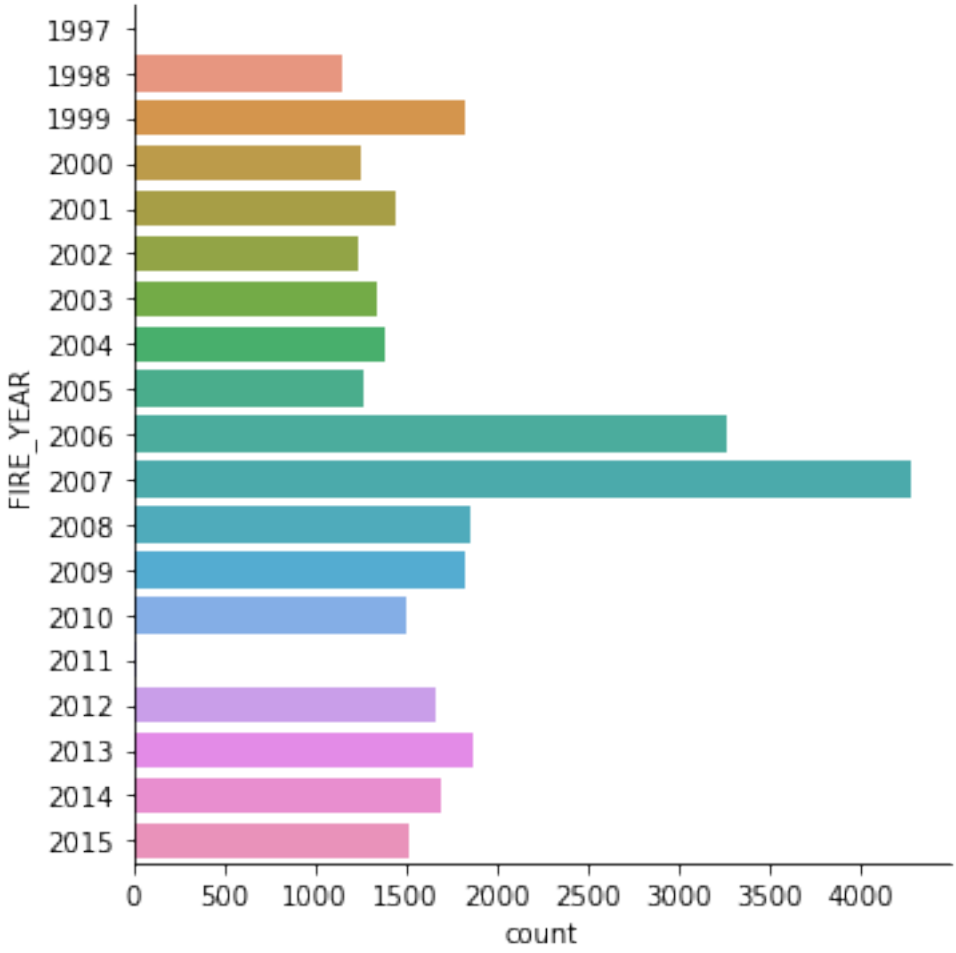
```
1                2004-05-12           845.0              1.0
2                2004-05-31          1921.0              5.0
17               2004-10-13          1618.0              2.0
23               2005-05-06          1145.0              5.0
…                     …                …                 …
177324           2015-10-23          1000.0              5.0
177325           2015-09-24           900.0             13.0
177326           2015-02-06          1300.0              9.0
177329           2015-09-10          1700.0              9.0
177330           2015-05-07          1730.0              5.0

         STAT_CAUSE_DESCR date(CONT_DATE)  FIRE_SIZE FIRE_SIZE_CLASS  \
0           Miscellaneous      2005-02-02       0.10               A
1               Lightning      2004-05-12       0.25               A
2          Debris Burning      2004-05-31       0.10               A
17          Equipment Use      2004-10-17    7700.00               G
23         Debris Burning      2005-05-09       0.10               A
…                     …              …         …                …
177324     Debris Burning      2015-10-23       0.50               B
177325  Missing/Undefined      2015-09-24       0.10               A
177326      Miscellaneous      2015-02-06       0.10               A
177329      Miscellaneous      2015-09-10       0.20               A
177330     Debris Burning      2015-05-07       0.20               A

          LATITUDE   LONGITUDE  FIPS_CODE  FIPS_NAME STATE
0       40.036944 -121.005833       63.0     Plumas    CA
1       38.933056 -120.404444       61.0     Placer    CA
2       38.984167 -120.735556       17.0  El Dorado    CA
17      38.780000 -120.260000       17.0  El Dorado    CA
23      38.700278 -119.840556        3.0     Alpine    CA
…             …         …         …         …   …
177324  39.371051 -122.139002       11.0     Colusa    CA
177325  37.193556 -120.841750       47.0     Merced    CA
177326  37.192558 -120.824639       47.0     Merced    CA
177329  41.461000 -120.506434       49.0      Modoc    CA
177330  41.943937 -121.564483       93.0   Siskiyou    CA

[30448 rows x 16 columns]
```

[178]:
```
#correlation plot (only for cts variables)
corr = df_CA.corr()
corr
corr.style.background_gradient(cmap='coolwarm')
```

[178]: <pandas.io.formats.style.Styler at 0x7fc318661e20>

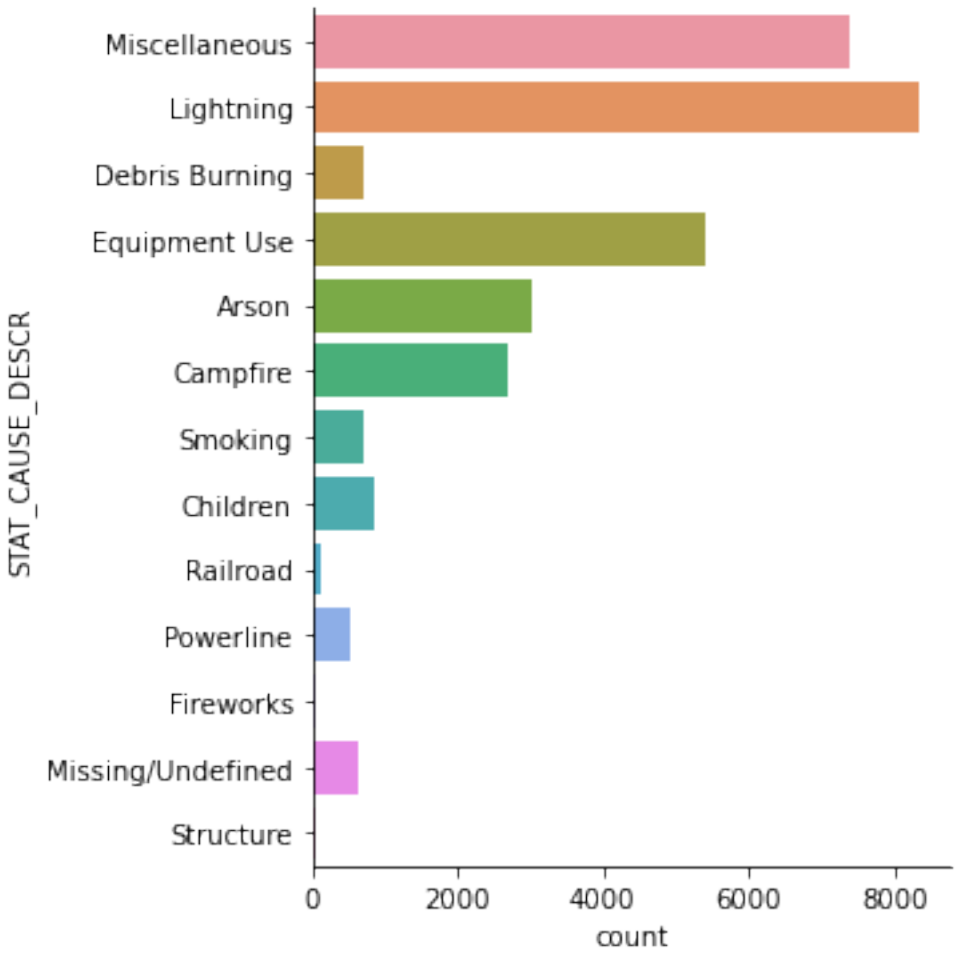
[95]:
```
#bar chart of fire year
sns.catplot(y="FIRE_YEAR", kind="count",  data=df_CA)
```

[95]: <seaborn.axisgrid.FacetGrid at 0x7fa134d6e580>



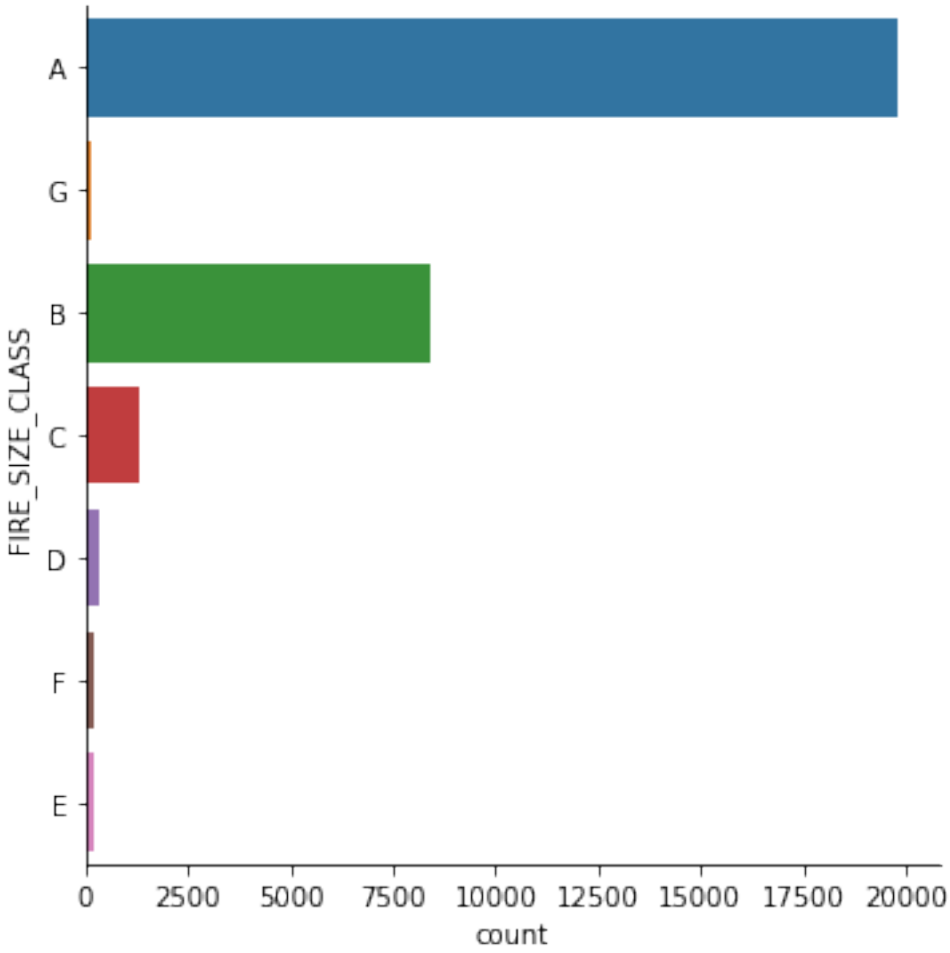
[94]:
```
#bar chart of fire cause descr
sns.catplot(y="STAT_CAUSE_DESCR", kind="count",  data=df_CA)
```

[94]: <seaborn.axisgrid.FacetGrid at 0x7fa134f02a60>



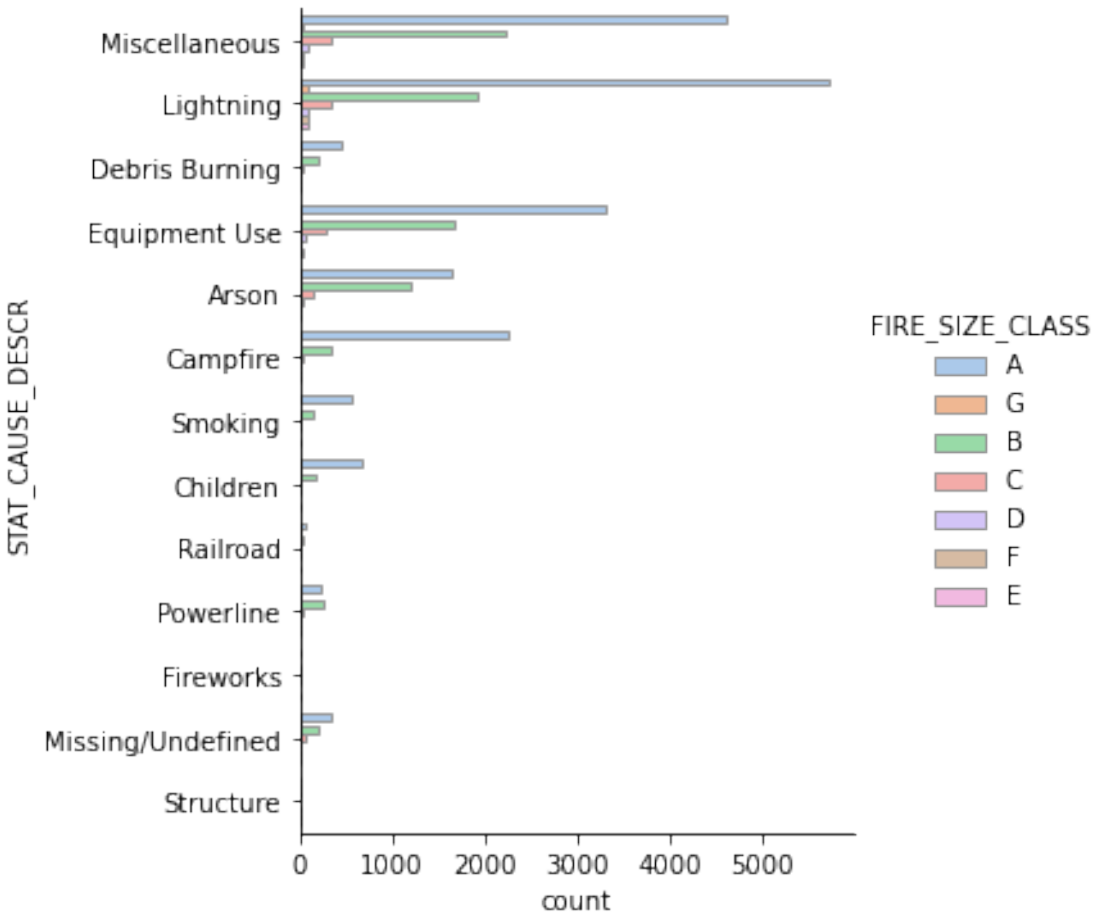
[96]:
```
#bar chart of fire size
sns.catplot(y="FIRE_SIZE_CLASS", kind="count",  data=df_CA)
```

[96]: <seaborn.axisgrid.FacetGrid at 0x7fa139738ee0>

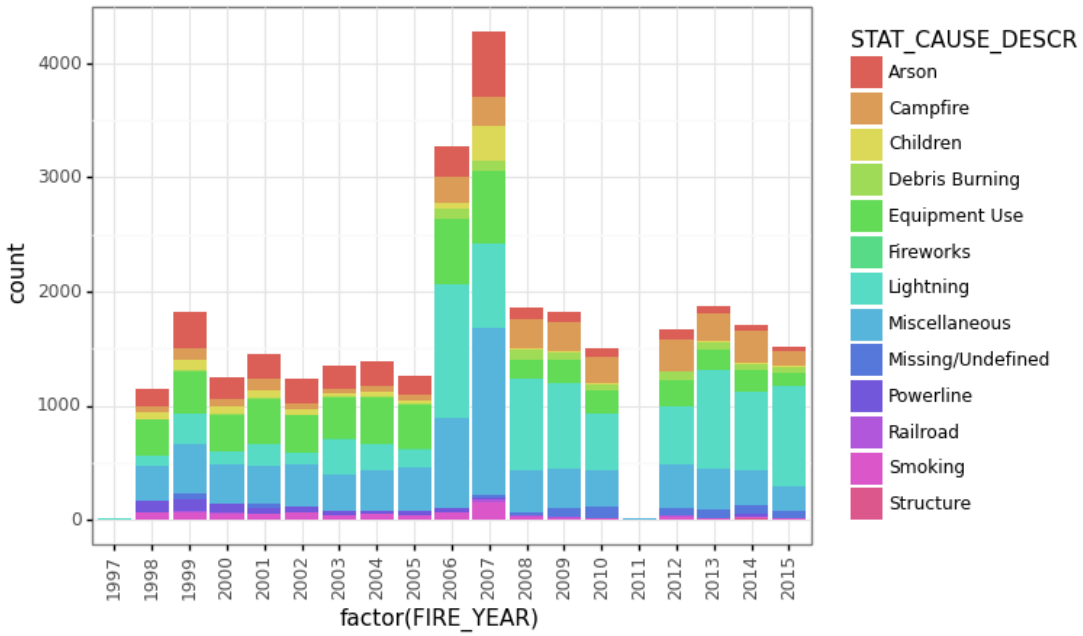


[98]:
```
#bar chart of fire size class and STAT_CAUSE_DESCR
sns.catplot(y="STAT_CAUSE_DESCR", hue="FIRE_SIZE_CLASS", kind="count",
            palette="pastel", edgecolor=".6",
            data=df_CA)
```

[98]: <seaborn.axisgrid.FacetGrid at 0x7fa1397533a0>

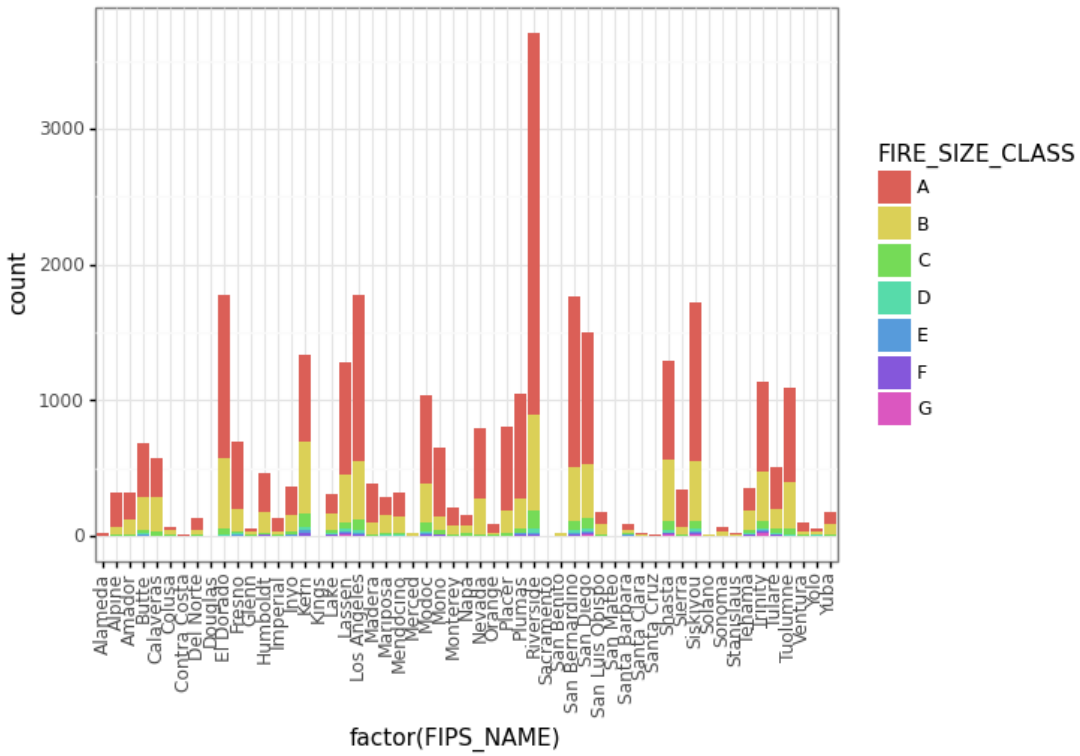


[104]:
```
#histgram of STAT_CAUSE_DESCR and fire year

(p9.ggplot(data=df_CA,
           mapping=p9.aes(x='factor(FIRE_YEAR)',fill = 'STAT_CAUSE_DESCR'))
    + p9.geom_bar()
    + p9.theme_bw()
    + p9.theme(axis_text_x = p9.element_text(angle=90))
)
```



[104]: <ggplot: (8770654033065)>

[107]:
```
#histgram of fire size class and fips name

(p9.ggplot(data=df_CA,
           mapping=p9.aes(x='factor(FIPS_NAME)',fill = 'FIRE_SIZE_CLASS'))
    + p9.geom_bar()
    + p9.theme_bw()
    + p9.theme(axis_text_x = p9.element_text(angle=90))
)
```



[107]: <ggplot: (8770654985204)>

[102]:
```
# map for all STAT_CAUSE_DESCR in CA
#empty map
world_map= folium.Map(tiles="cartodbpositron")
marker_cluster = MarkerCluster().add_to(world_map)
#for each coordinate, create circlemarker of host name
for i in range(len(df_CA)):
        lat = df_CA.iloc[i]['LATITUDE']
        long = df_CA.iloc[i]['LONGITUDE']
        radius=5
        popup_text = """host : {} """
        popup_text = popup_text.format(df_CA.iloc[i]['STAT_CAUSE_DESCR'])
        folium.CircleMarker(location = [lat, long], radius=radius, popup=␣
 ↪popup_text, fill =True).add_to(marker_cluster)
#show the map

display(world_map)
```

<folium.folium.Map at 0x7fc3728a8d90>

[184]:
```
#select only Lightning in CA
light_df = df_CA[df_CA['STAT_CAUSE_DESCR'] == 'Lightning']
```

```
#map only for Lightning in CA
import plotly.express as px

fig = px.scatter_mapbox(light_df, lat="LATITUDE", lon="LONGITUDE",␣
 ↪hover_data=["STAT_CAUSE_DESCR"],
                        color_discrete_sequence=["fuchsia"], zoom=3, height=300)
fig.update_layout(mapbox_style="open-street-map")
fig.update_layout(margin={"r":0,"t":0,"l":0,"b":0})
fig.show()
```

[ ]: