```python
# packages
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud, STOPWORDS, ImageColorGenerator
```

19] ✓ 0.0s Python

```python
# read file
wildfireData = pd.read_csv('/Users/Anshu7/Downloads/California_Fire_Incidents.csv')
wildfireData.head()
```

20] ✓ 0.0s Python

	AcresBurned	Active	AdminUnit	AirTankers	ArchiveYear	CalFireIncident	CanonicalUrl	ConditionStatement	ControlStatement	Counties	...	SearchKeywords	Started	Status	Structu
0	257314.0	False	Stanislaus National Forest/Yosemite National Park	NaN	2013	True	/incidents/2013/8/17/rim-fire/	NaN	NaN	Tuolumne	...	Rim Fire, Stanislaus National Forest, Yosemite...	2013-08-17T15:25:00Z	Finalized	
1	30274.0	False	USFS Angeles National Forest/Los Angeles Count...	NaN	2013	True	/incidents/2013/5/30/powerhouse-fire/	NaN	NaN	Los Angeles	...	Powerhouse Fire, May 2013, June 2013, Angeles ...	2013-05-30T15:28:00Z	Finalized	
2	27531.0	False	CAL FIRE Riverside Unit / San Bernardino Natio...	NaN	2013	True	/incidents/2013/7/15/mountain-fire/	NaN	NaN	Riverside	...	Mountain Fire, July 2013, Highway 243, Highway...	2013-07-15T13:43:00Z	Finalized	
3	27440.0	False	Tahoe National Forest	NaN	2013	False	/incidents/2013/8/10/american-fire/	NaN	NaN	Placer	...	American Fire, August 2013, Deadwood Ridge, Fo...	2013-08-10T16:30:00Z	Finalized	
4	24251.0	False	Ventura County Fire/CAL FIRE	NaN	2013	True	/incidents/2013/5/2/springs-fire/	Acreage has been reduced based upon more accur...	NaN	Ventura	...	Springs Fire, May 2013, Highway 101, Camarillo...	2013-05-02T07:01:00Z	Finalized	

5 rows × 40 columns

Summary statistics

```python
# show all column names
wildfireData.columns
```

21] ✓ 0.0s Python

```
Index(['AcresBurned', 'Active', 'AdminUnit', 'AirTankers', 'ArchiveYear',
       'CalFireIncident', 'CanonicalUrl', 'ConditionStatement',
       'ControlStatement', 'Counties', 'CountyIds', 'CrewsInvolved', 'Dozers',
       'Engines', 'Extinguished', 'Fatalities', 'Featured', 'Final',
       'FuelType', 'Helicopters', 'Injuries', 'Latitude', 'Location',
       'Longitude', 'MajorIncident', 'Name', 'PercentContained',
       'PersonnelInvolved', 'Public', 'SearchDescription', 'SearchKeywords',
       'Started', 'Status', 'StructuresDamaged', 'StructuresDestroyed',
       'StructuresEvacuated', 'StructuresThreatened', 'UniqueId', 'Updated',
       'WaterTenders'],
      dtype='object')
```

```
1  # prints info about the dataframe
2  wildfireData.info()
```

22] ✓ 0.0s

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1636 entries, 0 to 1635
Data columns (total 40 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   AcresBurned           1633 non-null   float64
 1   Active                1636 non-null   bool
 2   AdminUnit             1636 non-null   object
 3   AirTankers            28 non-null     float64
 4   ArchiveYear           1636 non-null   int64
 5   CalFireIncident       1636 non-null   bool
 6   CanonicalUrl          1636 non-null   object
 7   ConditionStatement    284 non-null    object
 8   ControlStatement      111 non-null    object
 9   Counties              1636 non-null   object
 10  CountyIds             1636 non-null   object
 11  CrewsInvolved         171 non-null    float64
 12  Dozers                123 non-null    float64
 13  Engines               191 non-null    float64
 14  Extinguished          1577 non-null   object
 15  Fatalities            21 non-null     float64
 16  Featured              1636 non-null   bool
 17  Final                 1636 non-null   bool
 18  FuelType              12 non-null     object
 19  Helicopters           84 non-null     float64
 20  Injuries              120 non-null    float64
 21  Latitude              1636 non-null   float64
 22  Location              1636 non-null   object
 23  Longitude             1636 non-null   float64
 24  MajorIncident         1636 non-null   bool
 25  Name                  1636 non-null   object
 26  PercentContained      1633 non-null   float64
 27  PersonnelInvolved     204 non-null    float64
 28  Public                1636 non-null   bool
 29  SearchDescription     1619 non-null   object
 30  SearchKeywords        1433 non-null   object
 31  Started               1636 non-null   object
 32  Status                1636 non-null   object
 33  StructuresDamaged     67 non-null     float64
 34  StructuresDestroyed   175 non-null    float64
 35  StructuresEvacuated   0 non-null      float64
 36  StructuresThreatened  30 non-null     float64
 37  UniqueId              1636 non-null   object
 38  Updated               1636 non-null   object
 39  WaterTenders          146 non-null    float64
dtypes: bool(6), float64(17), int64(1), object(16)
memory usage: 444.3+ KB
```

```
1  # summary statistics
2  wildfireData.describe()
```

24] ✓ 0.0s Python

	AcresBurned	AirTankers	ArchiveYear	CrewsInvolved	Dozers	Engines	Fatalities	Helicopters	Injuries	Latitude	Longitude	PercentContained	PersonnelInvolved	StructuresDamaged	Stru
count	1633.000000	28.000000	1636.000000	171.000000	123.000000	191.000000	21.000000	84.000000	120.000000	1636.000000	1636.000000	1633.0	204.000000	67.000000	
mean	4589.443968	4.071429	2016.608802	11.561404	7.585366	23.565445	8.619048	5.357143	3.500000	37.203975	-108.082642	100.0	328.553922	67.970149	
std	27266.337722	6.399818	1.845340	14.455633	14.028616	41.004424	18.529642	7.265437	3.806231	135.401380	37.006927	0.0	521.138789	155.771975	
min	0.000000	0.000000	2013.000000	0.000000	0.000000	0.000000	1.000000	0.000000	0.000000	-120.258000	-124.196290	100.0	0.000000	0.000000	
25%	35.000000	2.000000	2015.000000	2.500000	1.000000	5.000000	1.000000	1.000000	1.000000	34.165890	-121.768358	100.0	55.000000	1.000000	
50%	100.000000	2.000000	2017.000000	6.000000	2.000000	11.000000	3.000000	2.000000	3.000000	37.104065	-120.461560	100.0	151.500000	6.000000	
75%	422.000000	4.000000	2018.000000	13.500000	5.000000	24.000000	6.000000	5.000000	4.000000	39.086808	-117.474072	100.0	350.000000	49.500000	
max	410203.000000	27.000000	2019.000000	82.000000	76.000000	256.000000	85.000000	29.000000	26.000000	5487.000000	118.908200	100.0	3100.000000	783.000000	

```
1  # no. of duplicates
2  wildfireData.loc[wildfireData.duplicated(),:]
```

25] ✓ 0.0s Python

AcresBurned	Active	AdminUnit	AirTankers	ArchiveYear	CalFireIncident	CanonicalUrl	ConditionStatement	ControlStatement	Counties	...	SearchKeywords	Started	Status	StructuresDamaged	StructuresDestro

0 rows × 40 columns

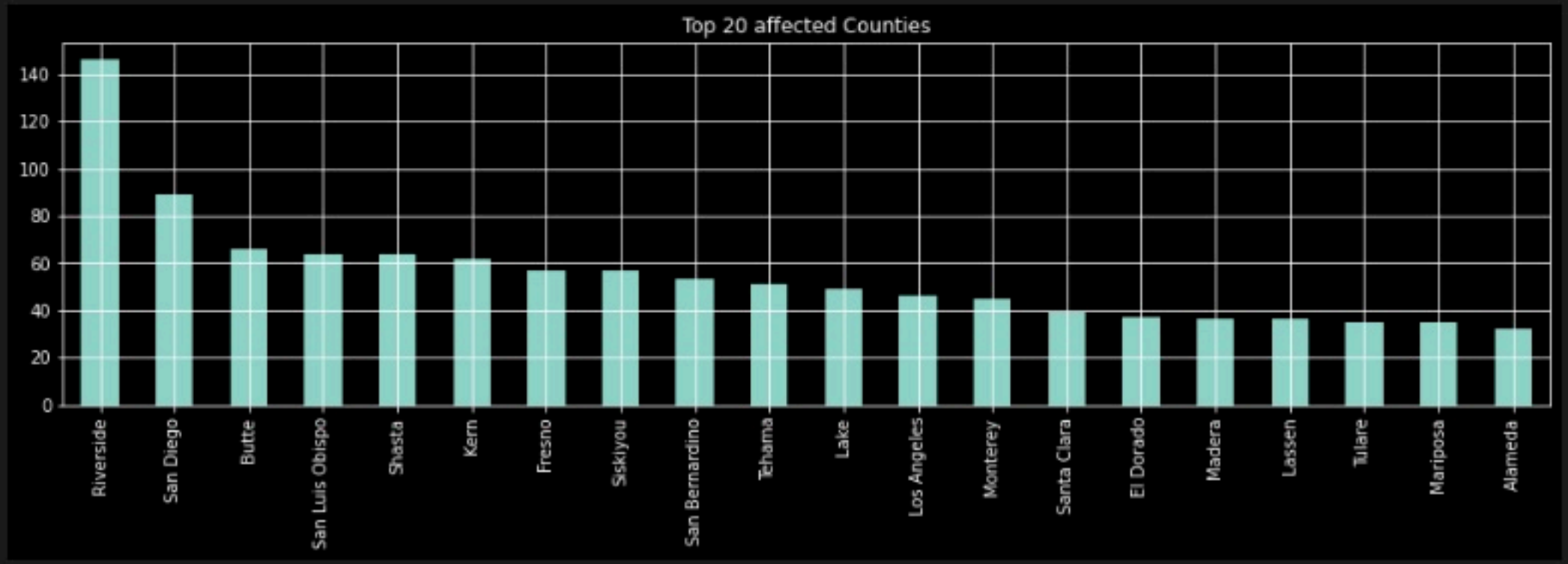
```
1  # counties (show top 10 only)
2  wildfireData.Counties.value_counts()[0:10]
```

26] ✓ 0.0s Python

```
Riverside          146
San Diego           89
Butte               66
San Luis Obispo     64
Shasta              64
Kern                62
Fresno              57
Siskiyou            57
San Bernardino      53
Tehama              51
Name: Counties, dtype: int64
```

```
1  plt.figure(figsize=(16,4))
2  wildfireData.Counties.value_counts()[0:20].plot(kind='bar')
3  plt.title('Top 20 affected Counties')
4  plt.grid()
5  plt.show()
```

] ✓ 0.1s

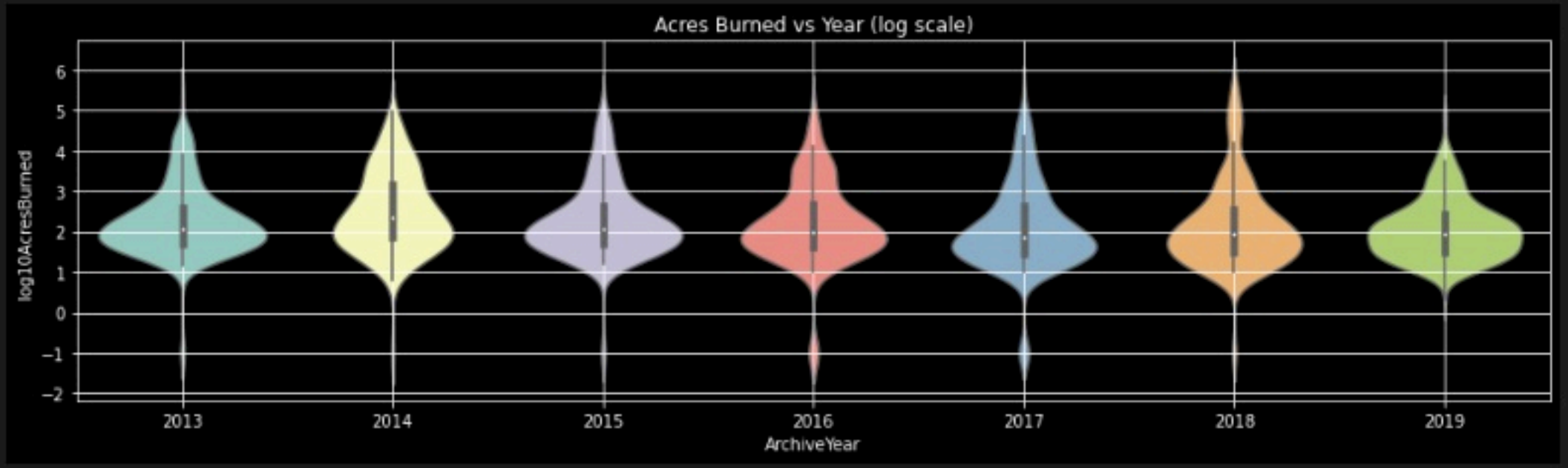


```
# log transformation
wildfireData['log10AcresBurned'] = np.log10(wildfireData.AcresBurned+0.1) # add 0.1 to avoid problems with log10(0)
```

✓ 0.0s

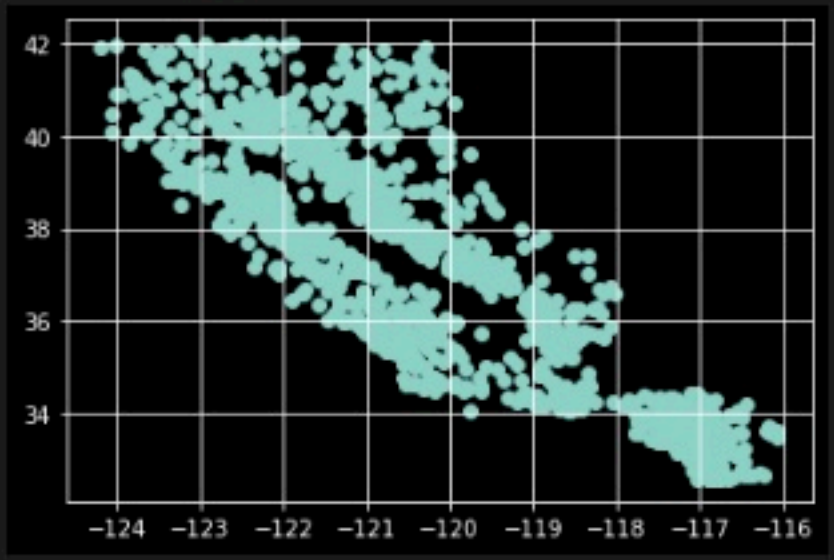
```
# violin plot in log coordinates
plt.figure(figsize=(16,4))
sns.violinplot(x='ArchiveYear', y='log10AcresBurned', data=wildfireData)
plt.grid()
plt.title('Acres Burned vs Year (log scale)')
plt.show()
```

✓ 0.1s



```
1  df_select = df_select[df_select.CanonicalUrl!='/incidents/2013/8/6/tram-fire/']
2  plt.scatter(df_select.Longitude, df_select.Latitude)
3  plt.grid()
4  plt.show()
```

[42] ✓ 0.1s



```python
import folium
import numpy as np

# Define map parameters
zoom_factor = 5  # Initial map size
radius_scaling = 50  # Scaling of bubbles

# Create the base map
my_map_1 = folium.Map(location=[36, -120], zoom_start=zoom_factor)

# Add circles to the map
for i in range(0, df_select.shape[0]):
    folium.Circle(
        location=[df_select.iloc[i]['Latitude'], df_select.iloc[i]['Longitude']],
        radius=np.sqrt(df_select.iloc[i]['AcresBurned']) * radius_scaling,
        color='red',
        popup='CanonicalUrl:' + df_select.iloc[i]['CanonicalUrl'] + ' - Year:' + str(int(df_select.iloc[i]['ArchiveYear'])) + ' - Acres Burned:' +
              str(df_select.iloc[i]['AcresBurned']),
        fill=True,
        fill_color='red'
    ).add_to(my_map_1)

# Add a legend to the map
legend_html = '''
     <div style="position: fixed;
                 bottom: 10px; left: 10px; width: 180px; height: 110px;
                 border:2px solid grey; z-index:9999; font-size:14px;
                 background-color:white;
                 ">  <b>Legend</b> <br>
                    Circle Size:<br>
                    Acres Burned > 0<br>
                  </div>
'''

my_map_1.get_root().html.add_child(folium.Element(legend_html))

# Add a title to the map
title_html = '''
             <h3 align="center" style="font-size:20px"><b>Fire Incident Locations</b></h3>
'''

my_map_1.get_root().html.add_child(folium.Element(title_html))

# Display the map
my_map_1

```