

## ▼ California Wildfire Data Exploration

[Link to Notebook](#)

```
import requests
import pandas as pd
import matplotlib.pyplot as plt
font = {'family' : 'normal',
        'size'   : 15}
```

```
plt.rc("font", **font)
```

```
from datetime import datetime
```

```
data = pd.read_csv("https://nathanpersonalbucket.s3-us-west-2.amazonaws.com/caliData.data")
data = data[data["YEAR_"] > 1970]
```

## ▼ Make date into python date time object

```
#pd.to_datetime(data["ALARM_DATE"], infer_datetime_format=True)
data = data.dropna(subset=["ALARM_DATE"])
#data[data["ALARM_DATE"] == NaN]
data["StartTime"] = data.apply(lambda x: datetime.strptime(str(x["ALARM_DATE"]), "%Y-
```

```
data
```

	OBJECTID	YEAR_	STATE	AGENCY	UNIT_ID	FIRE_NAME	INC_NUM	ALARM_DATE	C
0	1	2007.0	CA	CCO	LAC	OCTOBER	00246393	2007-10-21	2
1	2	2007.0	CA	CCO	LAC	MAGIC	00233077	2007-10-22	2
2	3	2007.0	CA	USF	ANF	RANCH	00000166	2007-10-20	2
3	4	2007.0	CA	CCO	LAC	EMMA	00201384	2007-09-11	2
4	5	2007.0	CA	CCO	LAC	CORRAL	00259483	2007-11-24	2
...	...	...	...	...	...	...	...	...	...
20808	21434	2019.0	CA	LRA	LDF	PALISADES	00000580	2019-10-21	2
20809	21435	2019.0	CA	CCO	LAC	MUREAU	NaN	2019-10-30	2

## ▼ Fires per Year

```
year = data[data["YEAR_"] > 1970].groupby("YEAR_")["YEAR_"].count()
year
```

```
YEAR_
1971.0    94
1972.0   137
1973.0   130
1974.0   181
1975.0   165
1976.0   112
1977.0   148
```

1978.0	126
1979.0	256
1980.0	263
1981.0	239
1982.0	132
1983.0	148
1984.0	198
1985.0	184
1986.0	151
1987.0	251
1988.0	219
1989.0	152
1990.0	136
1991.0	74
1992.0	150
1993.0	144
1994.0	194
1995.0	178
1996.0	285
1997.0	166
1998.0	143
1999.0	259
2000.0	176
2001.0	204
2002.0	243
2003.0	340
2004.0	274
2005.0	306
2006.0	315
2007.0	346
2008.0	431
2009.0	254
2010.0	209
2011.0	317
2012.0	350
2013.0	298
2014.0	232
2015.0	311
2016.0	347
2017.0	607
2018.0	412
2019.0	307

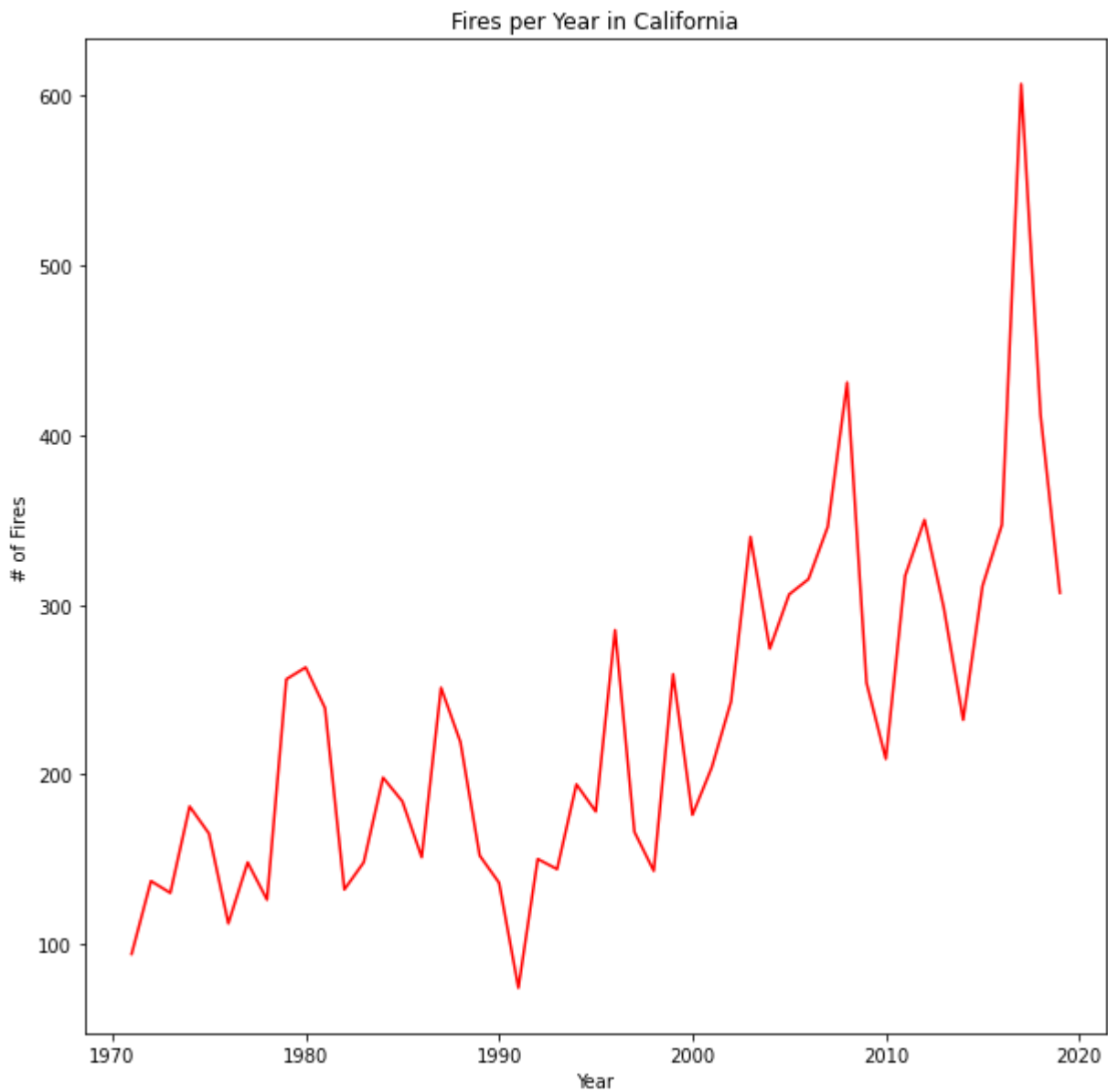
Name: YEAR\_, dtype: int64

```
fig, ax = plt.subplots (figsize = (10,10))

yearplot = year.plot.line(ax=ax, color="red")
plt.title = "Year"

yearplot.set_xlabel("Year")
yearplot.set_ylabel("# of Fires")
yearplot.title.set_text("Fires per Year in California")
```

findfont: Font family ['normal'] not found. Falling back to DejaVu Sans.  
findfont: Font family ['normal'] not found. Falling back to DejaVu Sans.



## ▼ Average Fire Size

```
data["GIS_ACRES"].mean()
```

1873.1779512474882

Fire Size per Year

```
size = data[data["YEAR_"] > 1970].groupby("YEAR_")["GIS_ACRES"].sum()  
size
```

YEAR_	
1971.0	5.851422e+04
1972.0	8.627648e+04
1973.0	2.213121e+05
1974.0	1.301359e+05
1975.0	1.874528e+05
1976.0	1.934319e+05
1977.0	4.576750e+05
1978.0	1.132783e+05
1979.0	3.261346e+05
1980.0	3.283649e+05
1981.0	3.001230e+05
1982.0	1.342519e+05
1983.0	8.247766e+04
1984.0	1.941268e+05
1985.0	5.347286e+05
1986.0	1.006874e+05
1987.0	8.468597e+05
1988.0	2.852651e+05
1989.0	1.204998e+05
1990.0	3.322944e+05
1991.0	3.596971e+04
1992.0	1.980371e+05
1993.0	3.092328e+05
1994.0	3.595380e+05
1995.0	1.830570e+05
1996.0	6.341339e+05
1997.0	2.106247e+05
1998.0	1.527535e+05
1999.0	7.772295e+05
2000.0	2.448704e+05
2001.0	2.395434e+05
2002.0	9.638985e+05
2003.0	9.677654e+05
2004.0	2.648496e+05
2005.0	2.555328e+05
2006.0	7.447646e+05
2007.0	1.039345e+06
2008.0	1.379773e+06
2009.0	4.358395e+05
2010.0	1.014744e+05
2011.0	2.024268e+05
2012.0	8.477077e+05
2013.0	5.698204e+05
2014.0	3.759496e+05
2015.0	7.891699e+05
2016.0	5.316685e+05
2017.0	1.424559e+06
2018.0	1.590150e+06
2019.0	2.808573e+05

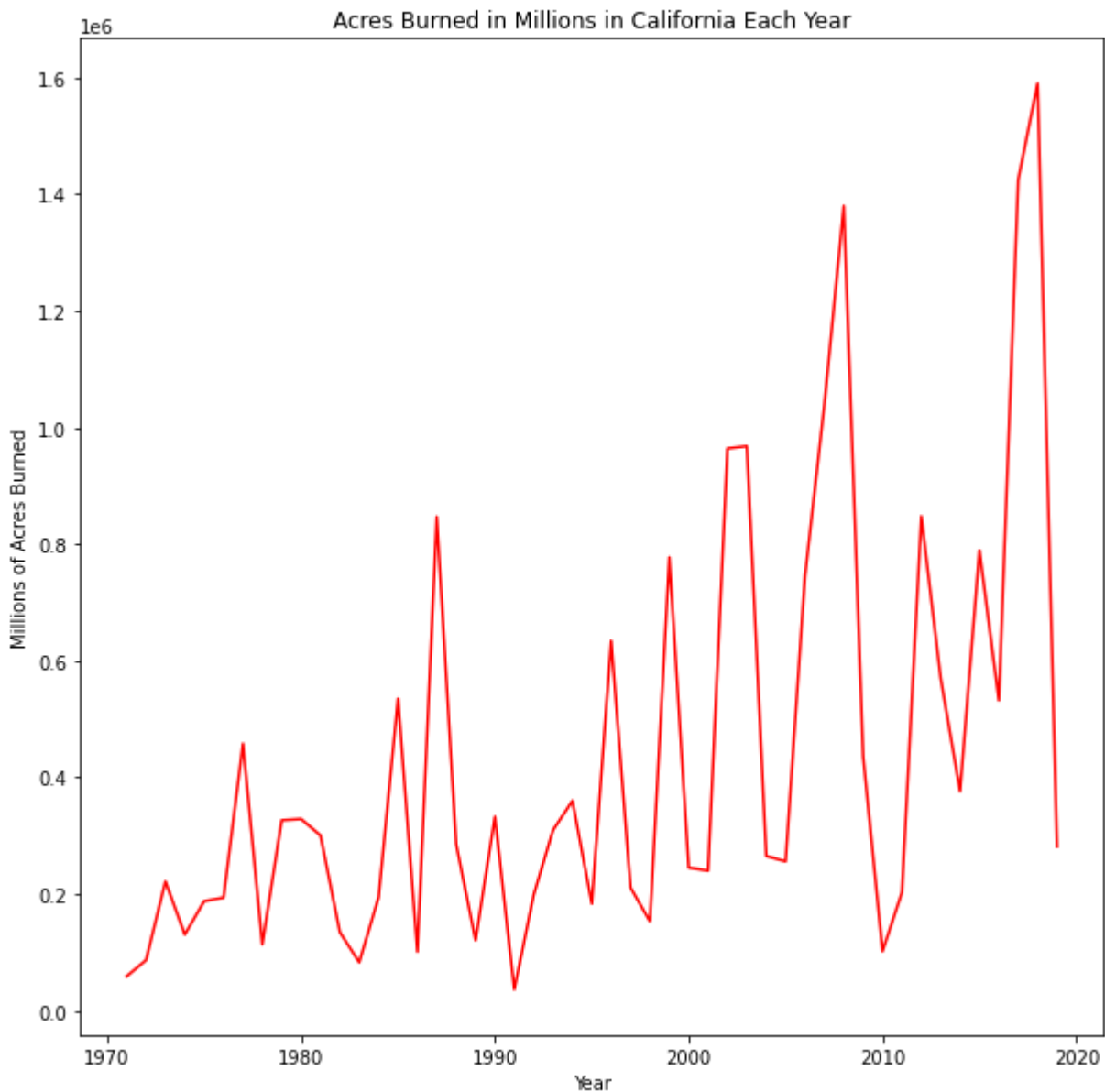
Name: GIS\_ACRES, dtype: float64

```
fig, ax = plt.subplots (figsize = (10,10))
```

```
ax.plot (year, GIS_ACRES, color = "red")
```

```
sizePlt = size.plot.line(ax=ax, color="red")
sizePlt.set_xlabel("Year")
sizePlt.set_ylabel("Millions of Acres Burned")
sizePlt.title.set_text("Acres Burned in Millions in California Each Year")
sizePlt
```

<matplotlib.axes.\_subplots.AxesSubplot at 0x7fe19a0f2b00>



## ▼ Fires by month

```
data["Month"] = data.apply(lambda x: x["StartTime"].strftime("%b"), axis=1)
#data["Year"] = data.apply(lambda x: x["StartTime"].strftime("%Y"), axis=1)
```

```
month = data.groupby("Month")["Month"].count()
print(month)
monthList = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']
month = month.reindex(monthList, axis=0)
```

month

```
Month
Apr      248
Aug     2369
Dec      155
Feb       82
Jan      117
Jul     2753
Jun     1979
Mar       10
May       79
Nov      342
Oct      810
Sep     1540
Name: Month, dtype: int64
Month
Jan      117
Feb       82
Mar       10
Apr      248
May       79
Jun     1979
Jul     2753
Aug     2369
Sep     1540
Oct      810
Nov      342
Dec      155
Name: Month, dtype: int64
```

```
fig, ax = plt.subplots (figsize = (10,10))

monthplt = month.plot.bar(ax=ax, color="#eb7c7c")

monthplt.set_ylabel("Total Fires")
monthplt.title.set_text("Total Fires per Month in California Since 1970")
```

Total Fires per Month in California Since 1970

