CAL Fire

Data Analysis with Python Vibe Coding

Justin Merlin
Fall 2025

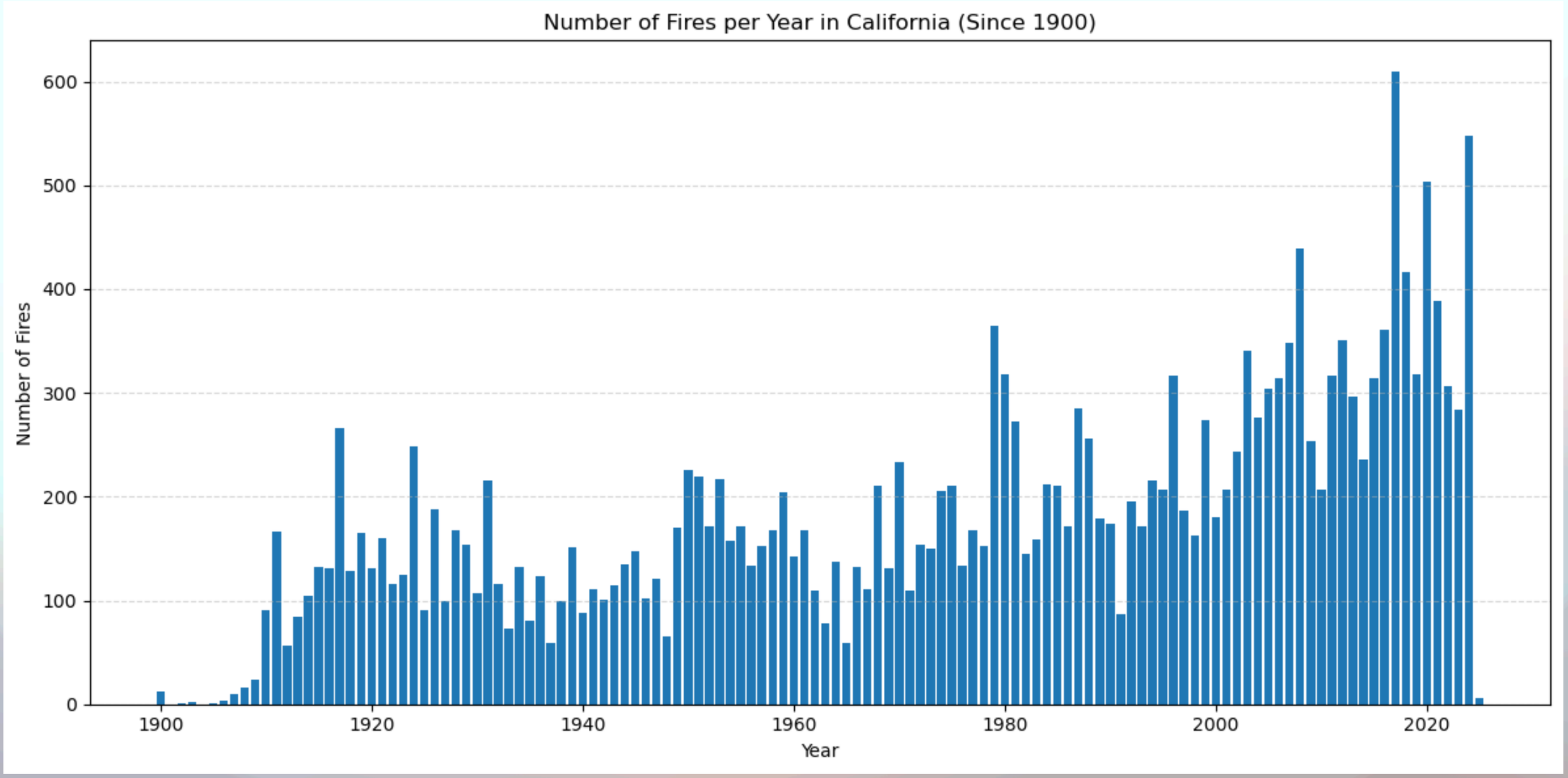
California Voxel-wise Prediction of Wildfire (time)

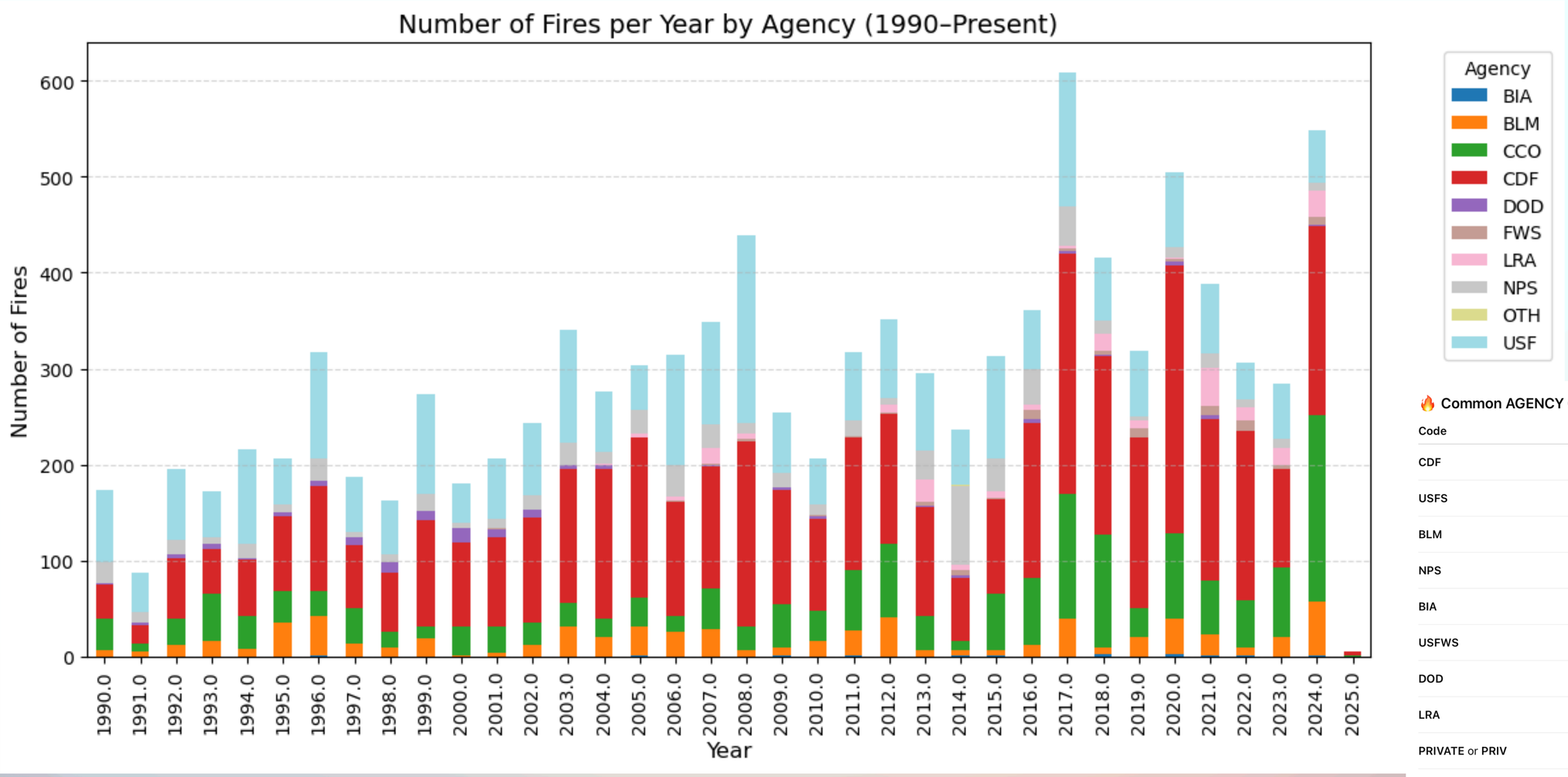
Wildfire Risk function (X, Y, time, factors) -> [0, 1]

Factors:

- **Historical fires** (ignition, **perimeters**)
- Climate and weather (Prism)
  - Max Temp
  - Precipitation
  - Max vapor pressure deficit
- Topography
  - Elevation
  - Slope
  - Aspect
- Fuel and vegetation (cover, type, etc.)
- Human activity and infrastructure
- Additional
  - Drought severity
  - Historical wind
  - Etc.

Source:
California_Fire_Perimeters Geopackage


Number of Fires per Year in California (Since 1900)
Number of Fires
0
100
200
300
400
500
600
1900
1920
1940
1960
1980
2000
2020
Year

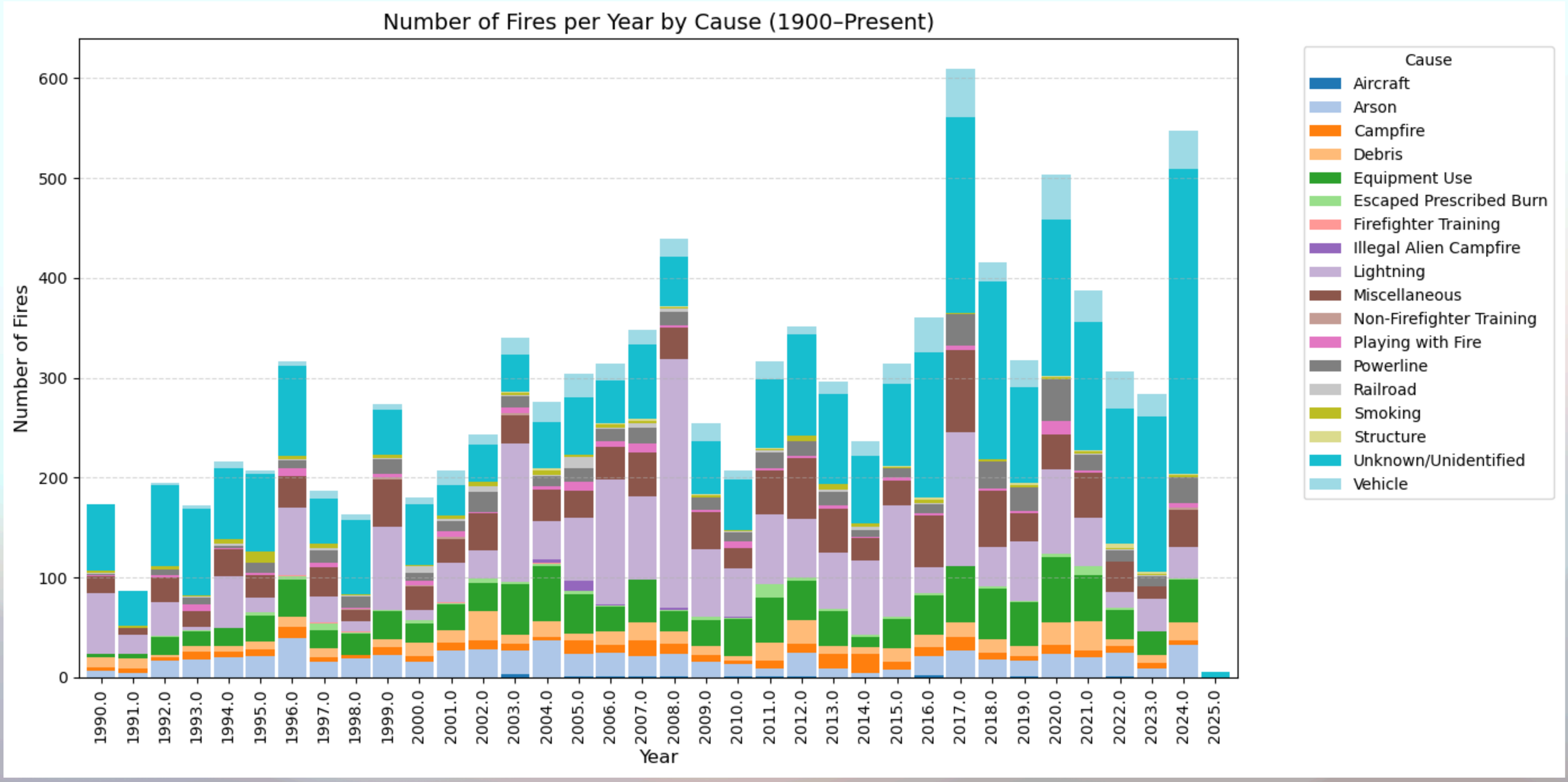


🔥 Common AGENCY Codes and Full Names

Code	Full Agency Name
CDF	California Department of Forestry and Fire Protection (**CAL FIRE**)
USFS	United States Forest Service
BLM	Bureau of Land Management
NPS	National Park Service
BIA	Bureau of Indian Affairs
USFWS	United States Fish and Wildlife Service
DOD	Department of Defense
LRA	Local Responsibility Area (local/county fire departments)
PRIVATE or **PRIV**	Private Land or Private Agency
OTHER	Other or Unknown Agency
UNK	Unknown or Unclassified

🔥 Top Large Wildfires in California (by Burned Area)

Year	Fire Name	Agency	Acres Burned (GIS_ACRES)
2020	August Complex	USF	1,032,700
2021	Dixie	CDF	963,405
2002	Biscuit	USF	501,082
2024	Park	CDF	429,603
2018	Ranch	CDF	410,203
2020	SCU Lightning Complex	CDF	396,825
2020	Creek	USF	379,842
2020	North Complex	USF	318,797
2012	Rush	BLM	315,511
2020	Hennessey	CDF	305,352
2017	Thomas	USF	281,791
2003	Cedar	USF	270,686
2013	Rim	USF	256,176
2007	Zaca	CDF	240,359
2018	Carr	CDF	229,651
2021	Monument	USF	223,108
2021	Caldor	USF	221,786
2021	River Complex	USF	199,354
2020	Castle	USF	170,648

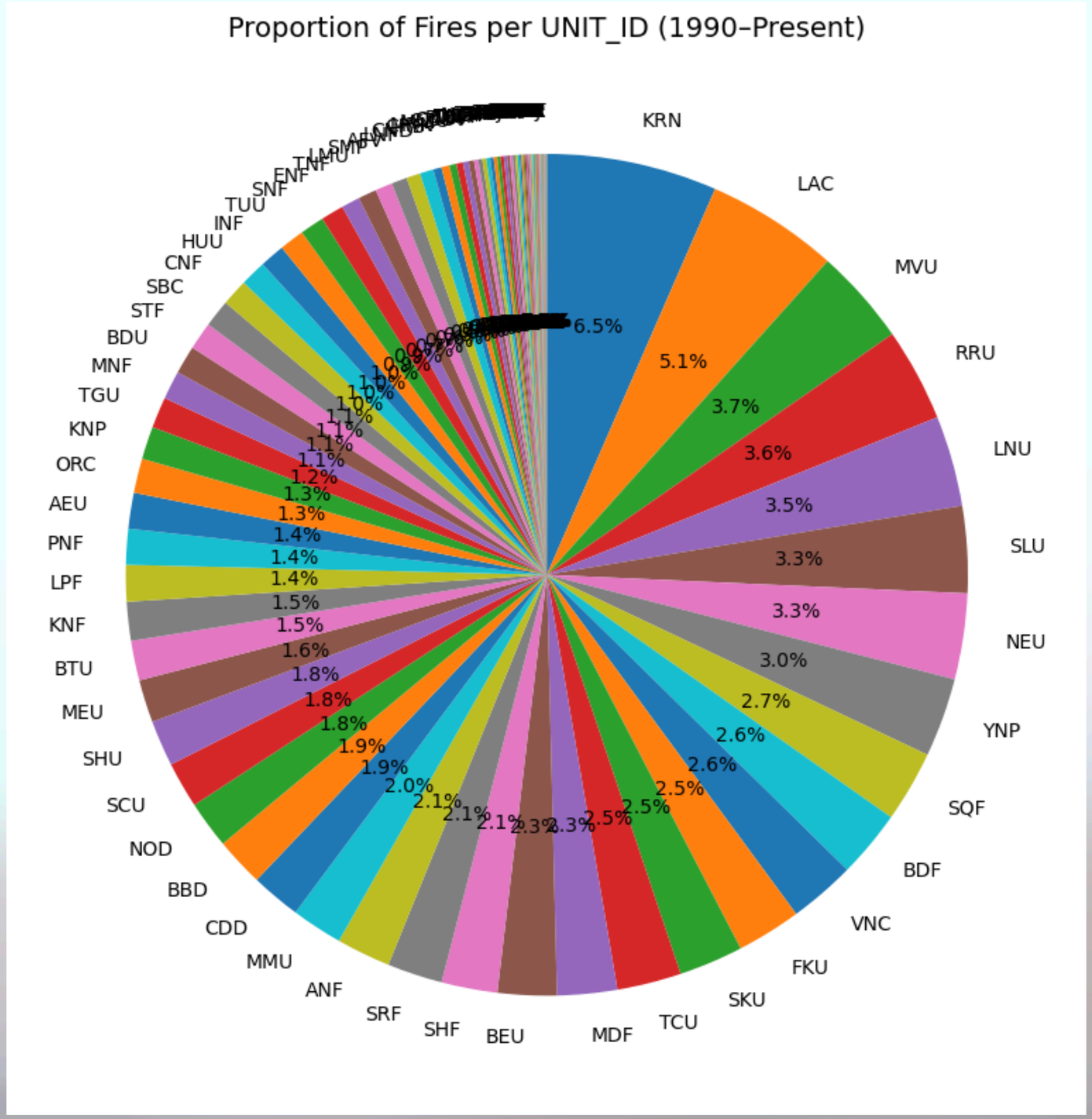

Number of Fires per Year by Cause (1900–Present)
Number of Fires
Year
0
100
200
300
400
500
600
1990.0
1991.0
1992.0
1993.0
1994.0
1995.0
1996.0
1997.0
1998.0
1999.0
2000.0
2001.0
2002.0
2003.0
2004.0
2005.0
2006.0
2007.0
2008.0
2009.0
2010.0
2011.0
2012.0
2013.0
2014.0
2015.0
2016.0
2017.0
2018.0
2019.0
2020.0
2021.0
2022.0
2023.0
2024.0
2025.0
Cause
Aircraft
Arson
Campfire
Debris
Equipment Use
Escaped Prescribed Burn
Firefighter Training
Illegal Alien Campfire
Lightning
Miscellaneous
Non-Firefighter Training
Playing with Fire
Powerline
Railroad
Smoking
Structure
Unknown/Unidentified
Vehicle

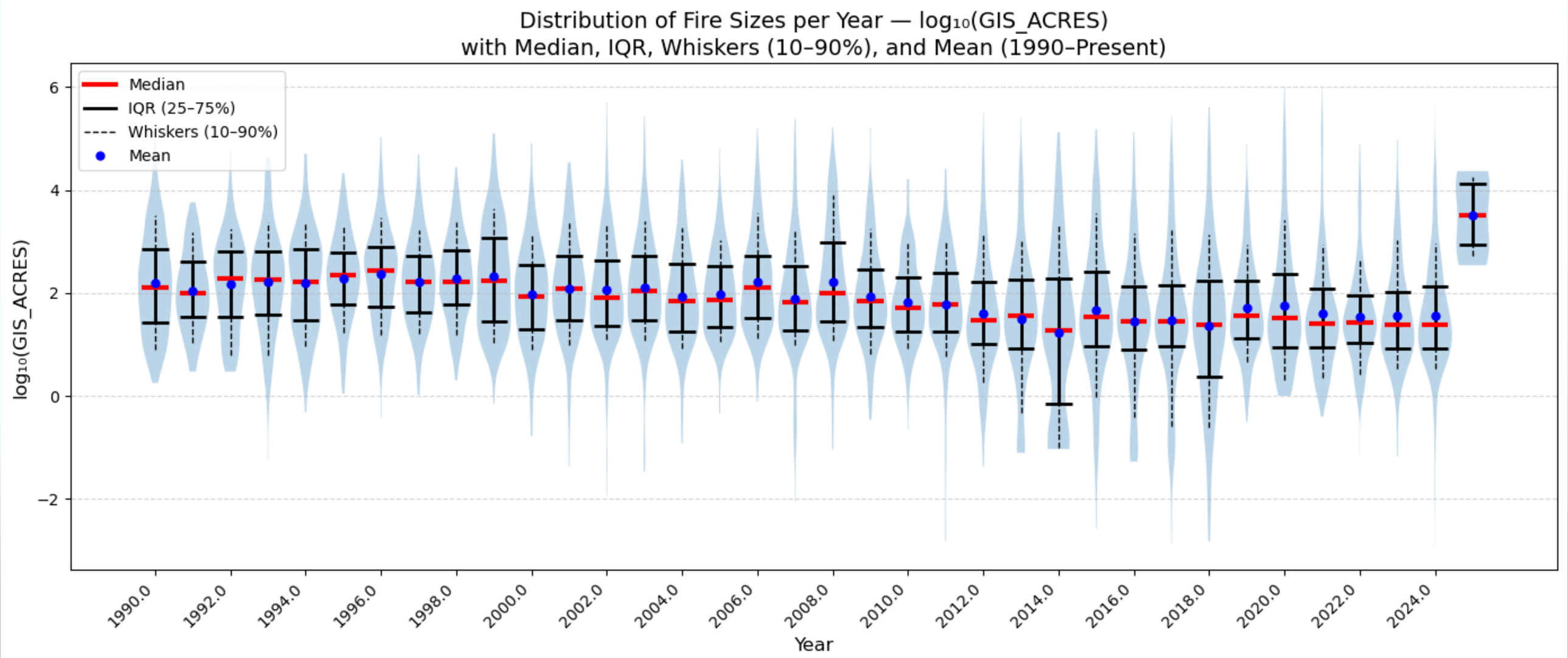
Top 10 Responding Units Since 1990:

	Unit_ID	count
0	KRN	677
1	LAC	525
2	MVU	382
3	RRU	371
4	LNU	360
5	SLU	344
6	NEU	341
7	YNP	314
8	SQF	283
9	BDF	268

🔥 Top 10 Responding Units by Total Burned Area (1990–Present):

	UNIT_ID	GIS_ACRES
12	BTU	1.756823e+06
51	LNU	1.751738e+06
78	SHF	1.404817e+06
60	MNF	1.337420e+06
39	KNF	1.294882e+06
52	LPF	1.256431e+06
89	SQF	8.865796e+05
69	PNF	8.562016e+05
90	SRF	8.458106e+05
86	SNF	7.361541e+05

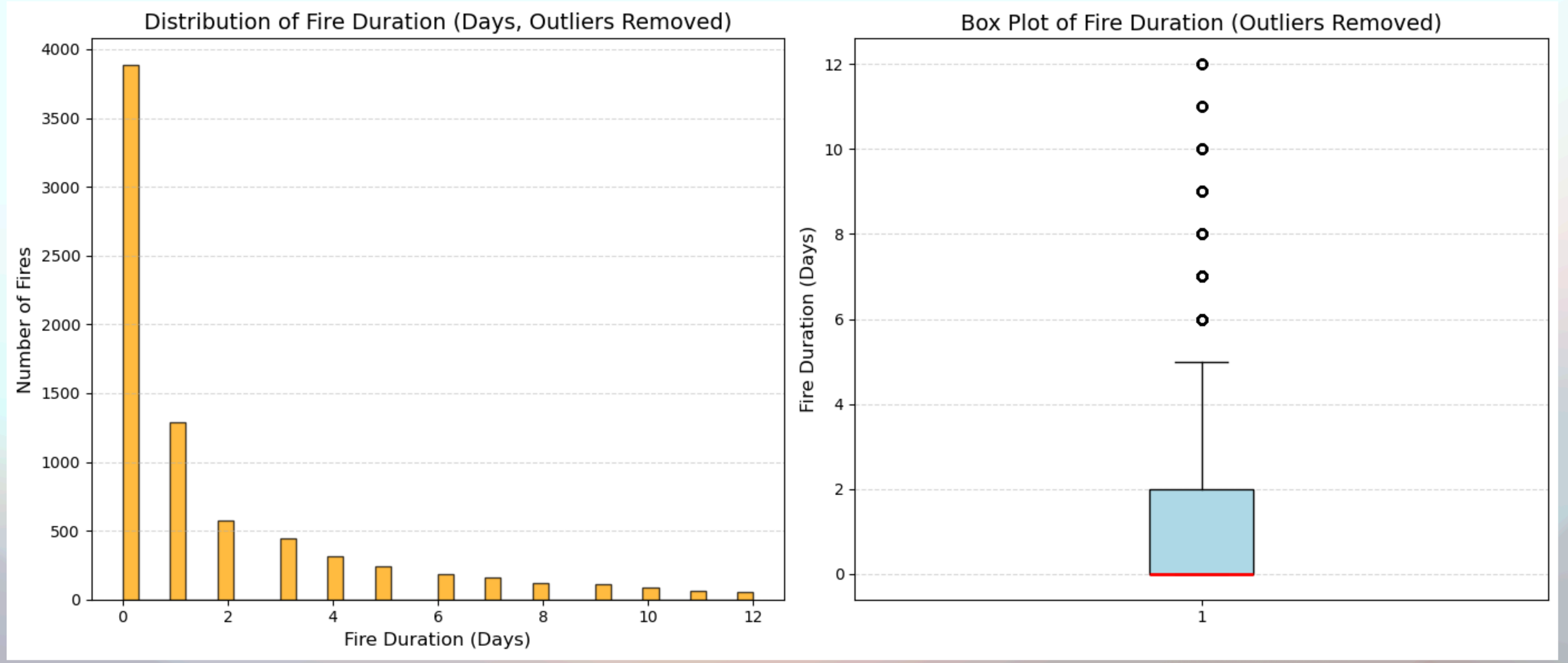

Proportion of Fires per UNIT_ID (1990–Present)
KRN
LAC
MVU
RRU
LNU
SLU
NEU
YNP
SQF
BDF
VNC
FKU
SKU
TCU
MDF
BEU
SHF
SRF
ANF
MMU
CDD
BBD
NOD
SCU
SHU
MEU
BTU
KNF
LPF
PNF
AEU
ORC
KNP
TGU
MNF
BDU
STF
SBC
CNF
HUU
INF
TUU
SNF
ENF
6.5%
5.1%
3.7%
3.6%
3.5%
3.3%
3.3%
3.0%
2.7%
2.6%
2.6%
2.5%
2.5%
2.3%
2.3%
2.1%
2.1%
2.1%
2.0%
1.9%
1.9%
1.8%
1.8%
1.8%
1.6%
1.5%
1.5%
1.4%
1.4%
1.4%
1.3%
1.3%
1.2%
1.1%
1.1%
1.1%
1.1%
1.0%
1.0%
1.0%

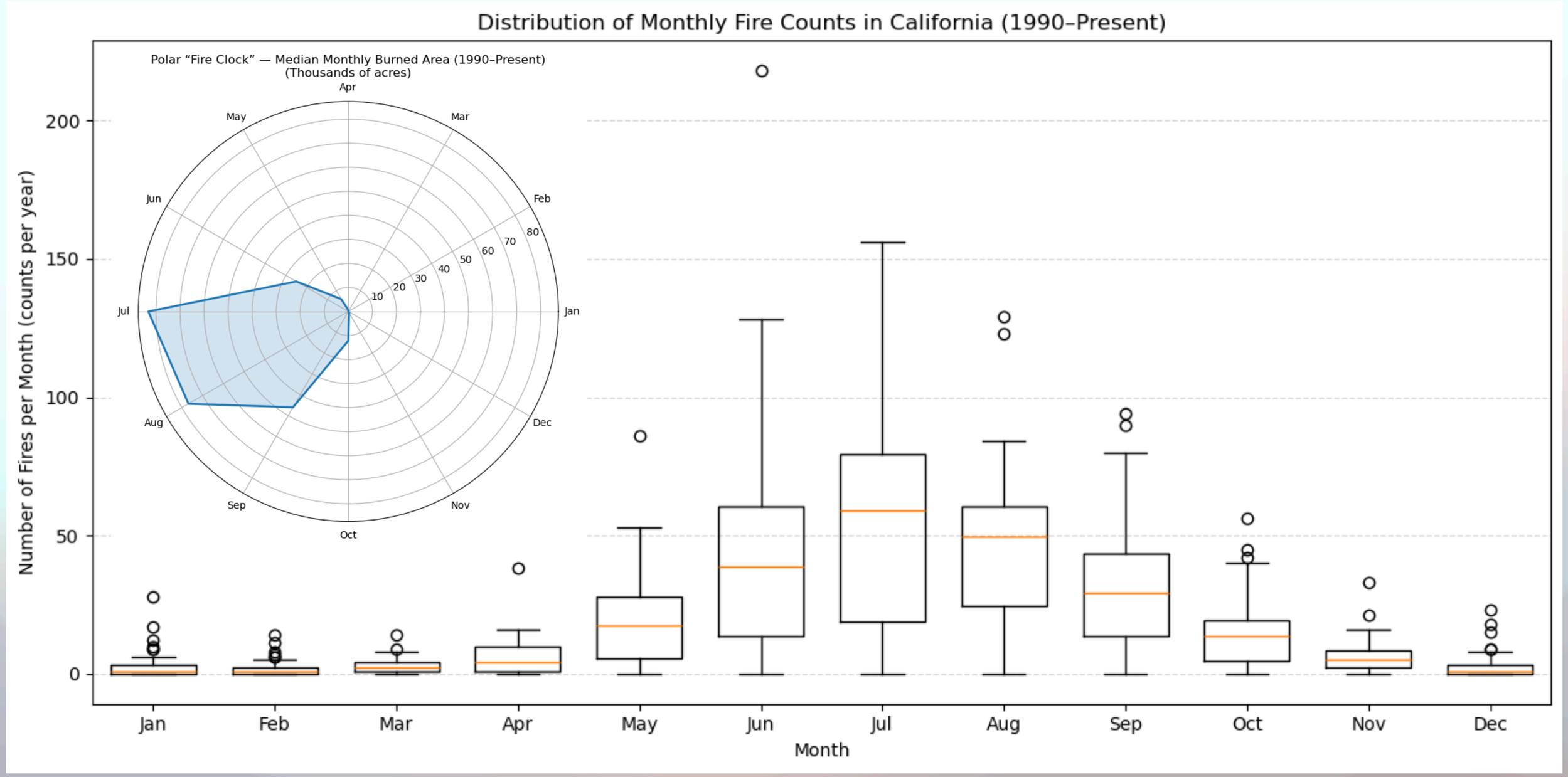

Distribution of Fire Sizes per Year — log₁₀(GIS_ACRES)
with Median, IQR, Whiskers (10–90%), and Mean (1990–Present)
Median
IQR (25–75%)
Whiskers (10–90%)
Mean
log₁₀(GIS_ACRES)
6
4
2
0
−2
1990.0
1992.0
1994.0
1996.0
1998.0
2000.0
2002.0
2004.0
2006.0
2008.0
2010.0
2012.0
2014.0
2016.0
2018.0
2020.0
2022.0
2024.0
Year

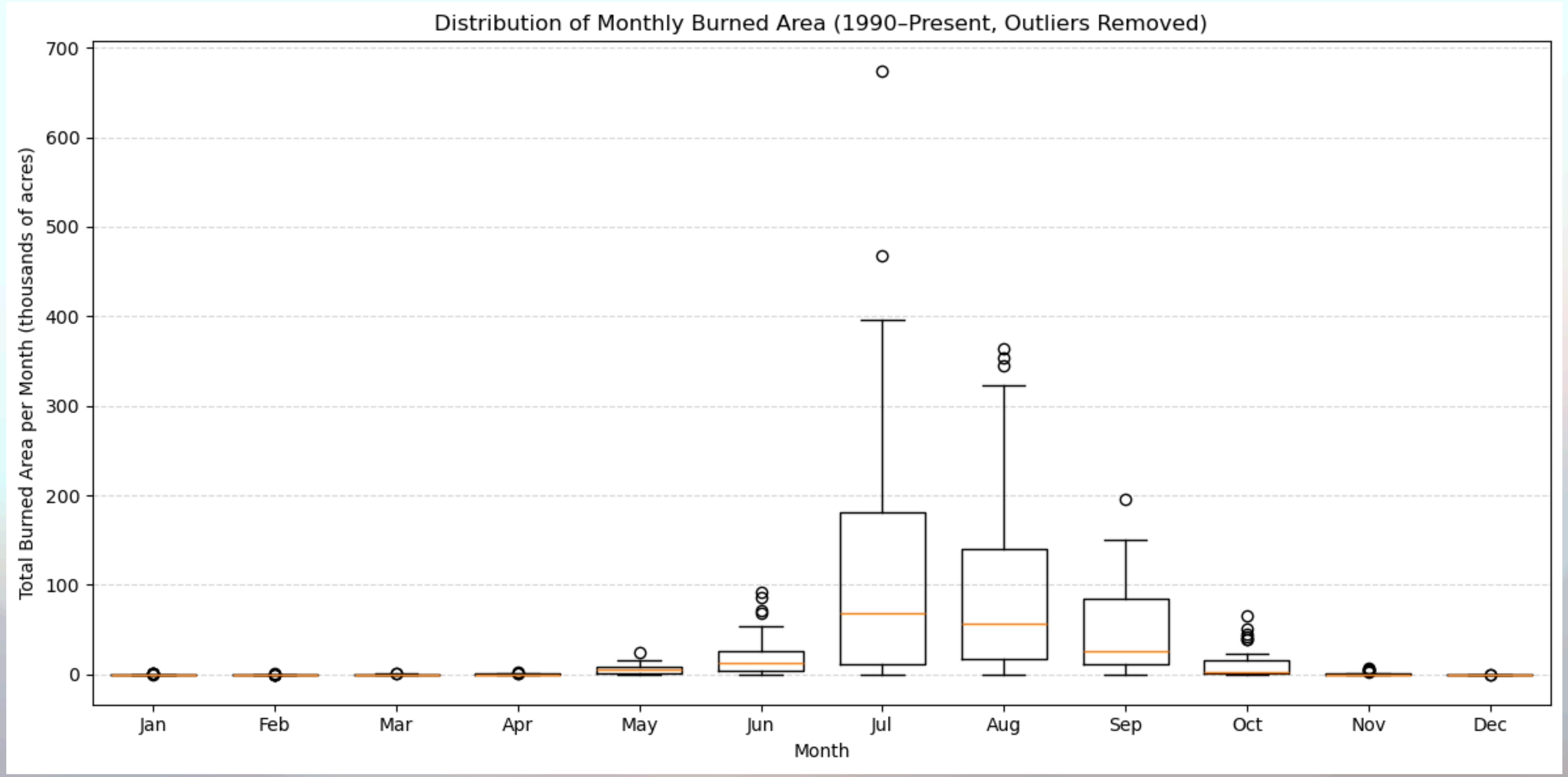
(1990-Now)

🔥 Fire Duration Summary (in Days)

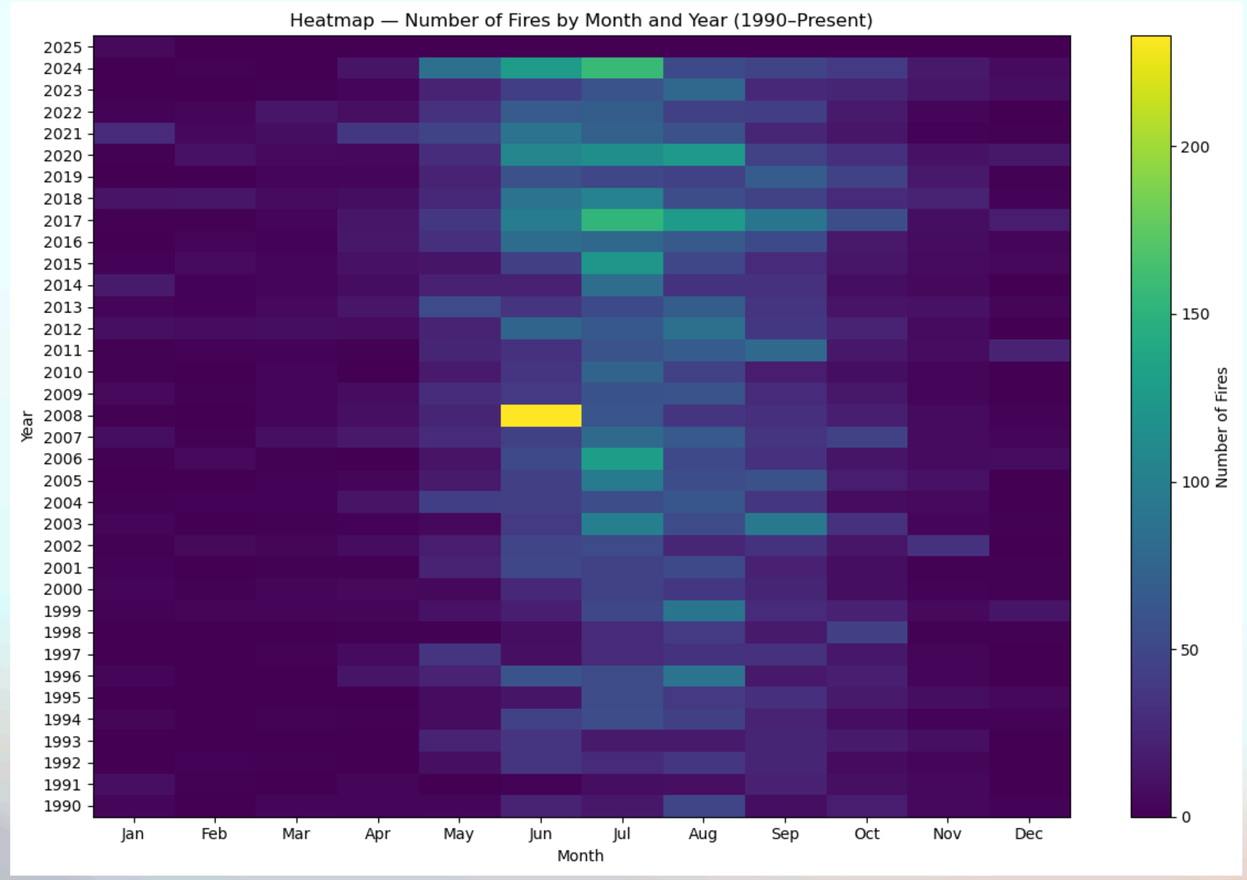
Statistic	All Fires	After Outlier Removal
Count	8,740	7,519
Mean	9.21	1.68
Standard Deviation (Std)	42.93	2.65
Minimum	0.00	0.00
25th Percentile (Q1)	0.00	0.00
Median (Q2)	1.00	0.00
75th Percentile (Q3)	5.00	2.00
Maximum	3,319.00	12.00

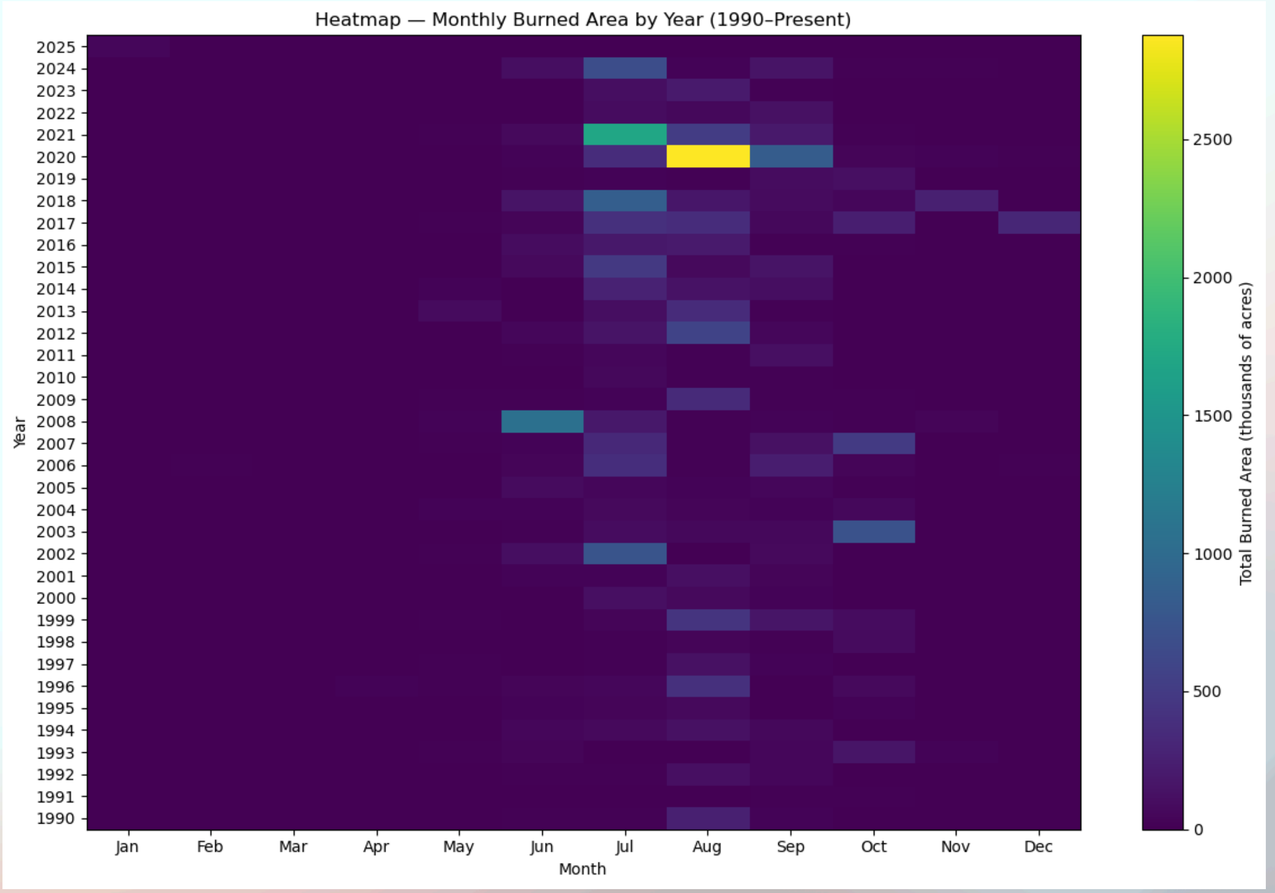

Distribution of Fire Duration (Days, Outliers Removed)
Number of Fires
Fire Duration (Days)
0
500
1000
1500
2000
2500
3000
3500
4000
0
2
4
6
8
10
12
Box Plot of Fire Duration (Outliers Removed)
Fire Duration (Days)
0
2
4
6
8
10
12
1

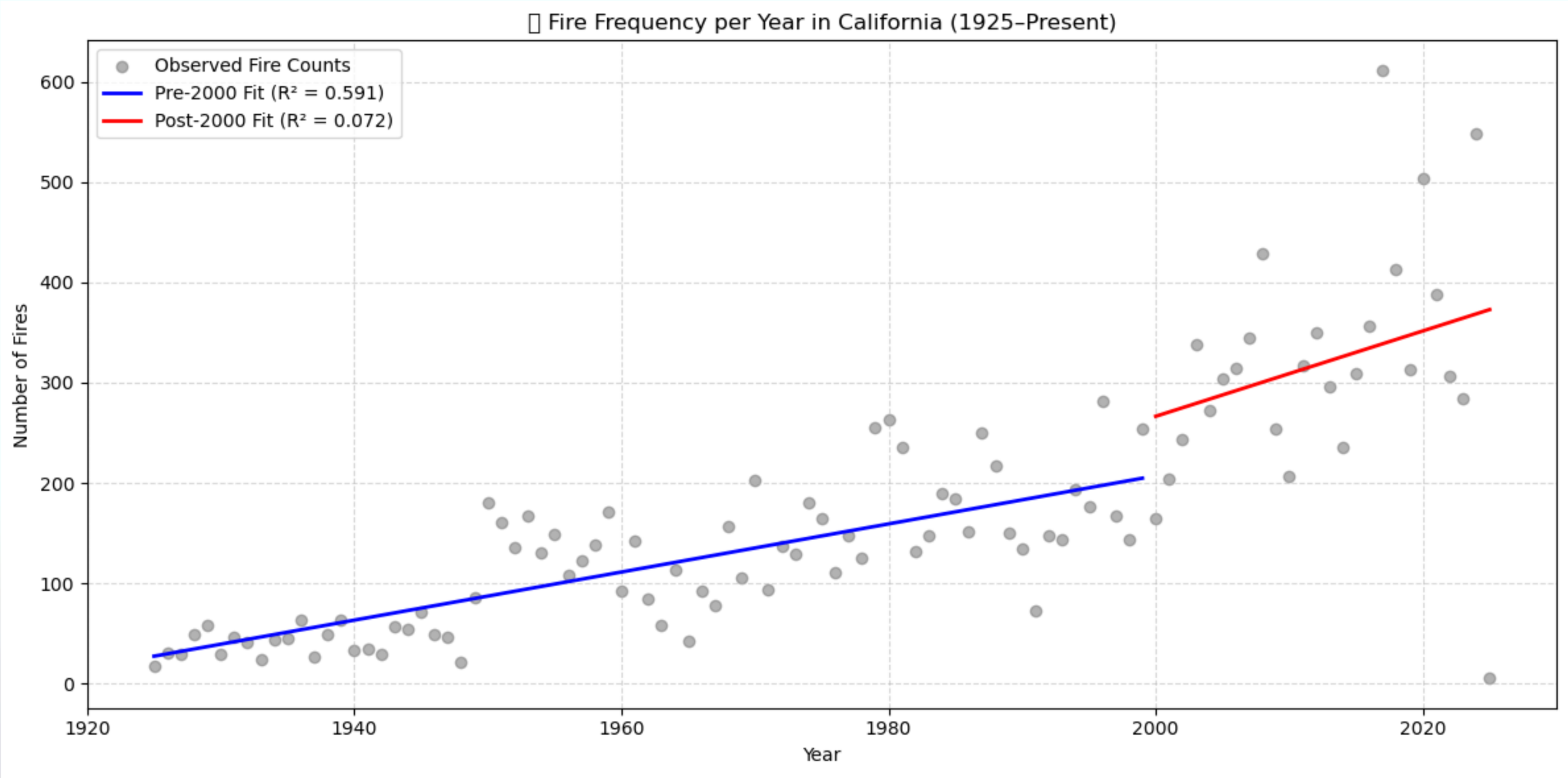

Distribution of Monthly Fire Counts in California (1990–Present)
Polar “Fire Clock” — Median Monthly Burned Area (1990–Present)
(Thousands of acres)
Apr
Mar
Feb
Jan
Dec
Nov
Oct
Sep
Aug
Jul
Jun
May
10
20
30
40
50
60
70
80
Number of Fires per Month (counts per year)
0
50
100
150
200
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month


Distribution of Monthly Burned Area (1990–Present, Outliers Removed)
Total Burned Area per Month (thousands of acres)
700
600
500
400
300
200
100
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month

SEASONALITY






🔥 Fire Frequency per Year in California (1925–Present)
Observed Fire Counts
Pre-2000 Fit (R² = 0.591)
Post-2000 Fit (R² = 0.072)
Number of Fires
Year
0
100
200
300
400
500
600
1920
1940
1960
1980
2000
2020