Hypothesis Testings

Climate and Geological Events

Overview

For the following hypothesis testing, we set a value of $\alpha = 0.05$

We set up three pairs of null/alternative hypotheses. We conducted 7 hypothesis testings. Including 3 one-way ANOVA tests and 3 Tukey HSD (honestly significant difference) tests, and 1 t-test

Data Collection

We collected annual average temperatures of various countries from the World Bank API.

We web scraped historical earthquake and volcanic eruption data from National Oceanic and Atmospheric Administration.

And we collected historical data on natural disasters from the FEMA API.

All data are inserted into AWS mysql database before analyses were conducted.

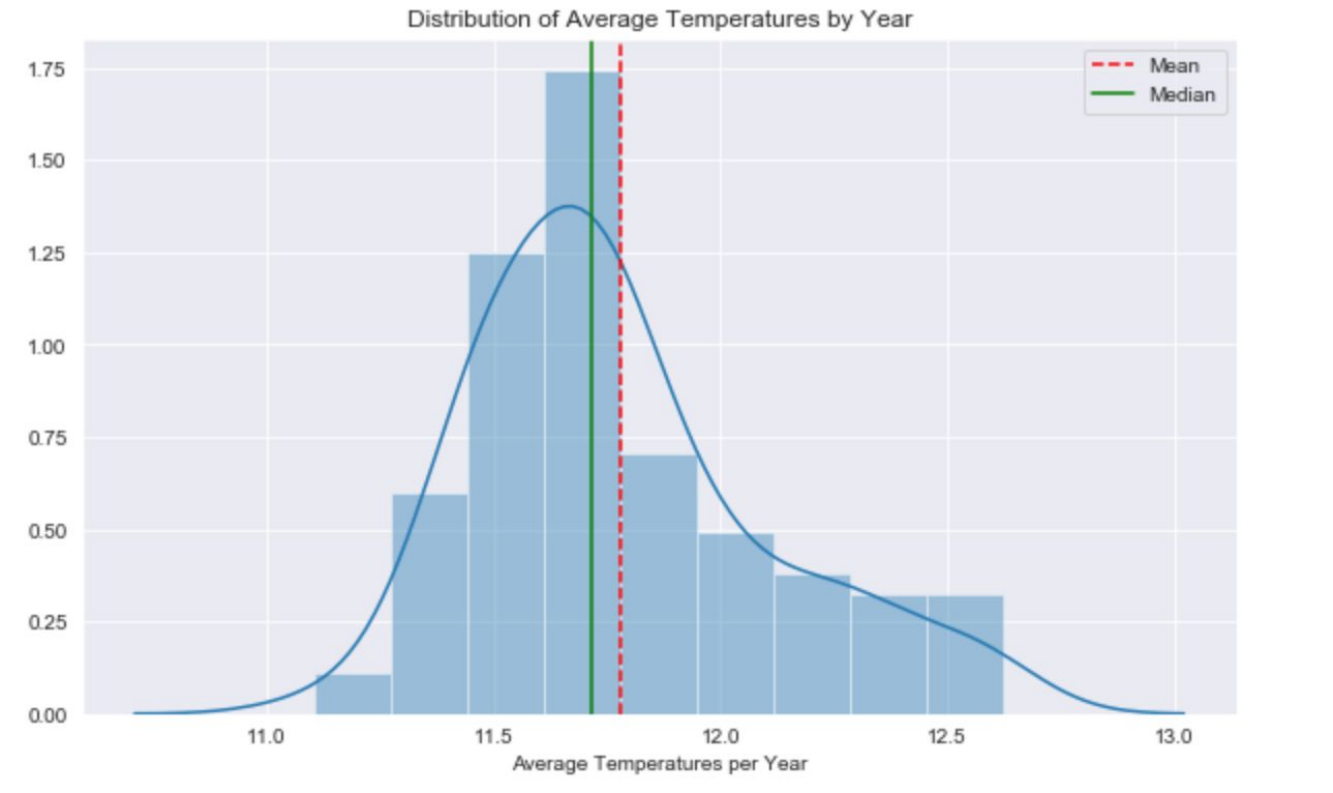
Hypothesis 1: Average Global Temperature

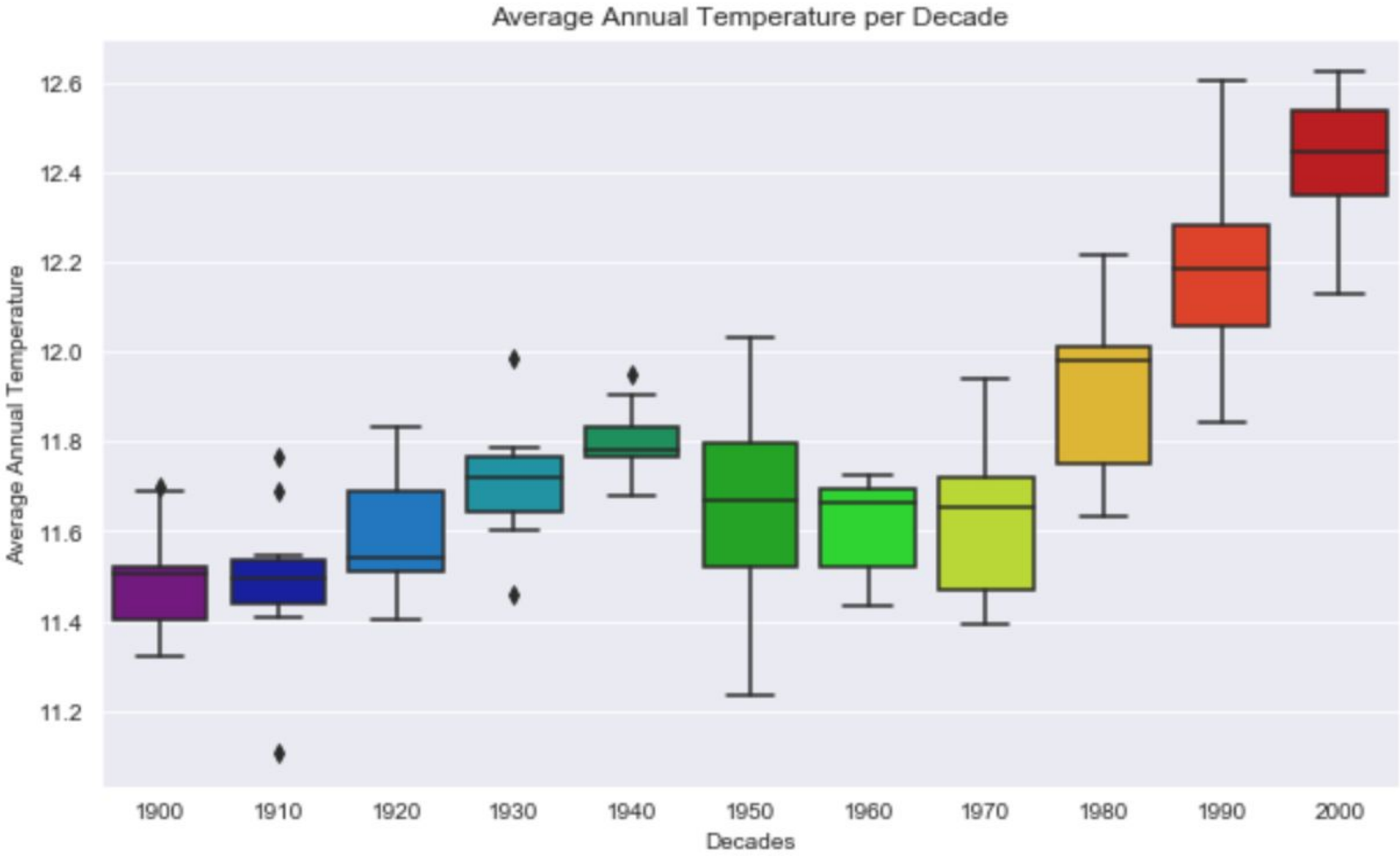
Null hypothesis: Global annual average temperatures are the same across decades.

Alternative hypothesis: Global annual average temperatures are different across decades.

Mean = 11.77772
Variance = 0.10542

Kurtosis = 2.817
Skew = -0.22




Average Annual Temperature per Decade
Average Annual Temperature
12.6
12.4
12.2
12.0
11.8
11.6
11.4
11.2
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
Decades

Hypothesis Testings: Average Global Temperature

Null hypothesis: Global annual average temperature is the same across decades.

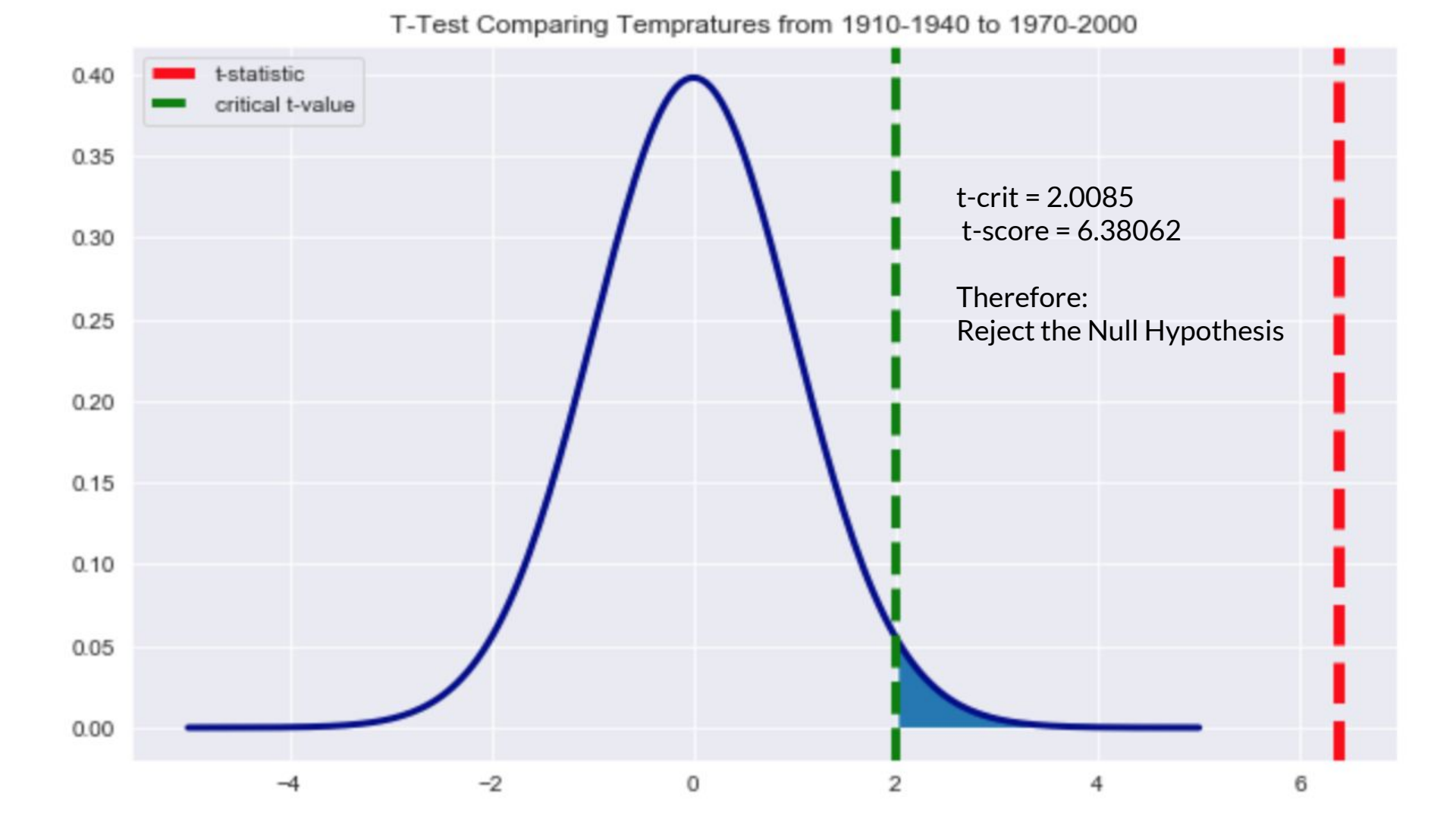
Alternative hypothesis: Global annual average temperature is different across decades.

f-critical value = 3.93

Given the results: Reject the Null Hypothesis

ANOVA test for Average Global Temperature by Decade

	df	sum_sq	mean_sq	F	PR(>F)
decade	1.0	5.925145	5.925145	116.101657	8.945390e-19
Residual	107.0	5.460650	0.051034	NaN	NaN


T-Test Comparing Tempratures from 1910-1940 to 1970-2000
t-statistic
critical t-value
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
−4
−2
0
2
4
6
t-crit = 2.0085
t-score = 6.38062
Therefore:
Reject the Null Hypothesis

Pairwise Tukey HSD: Global Temperatures

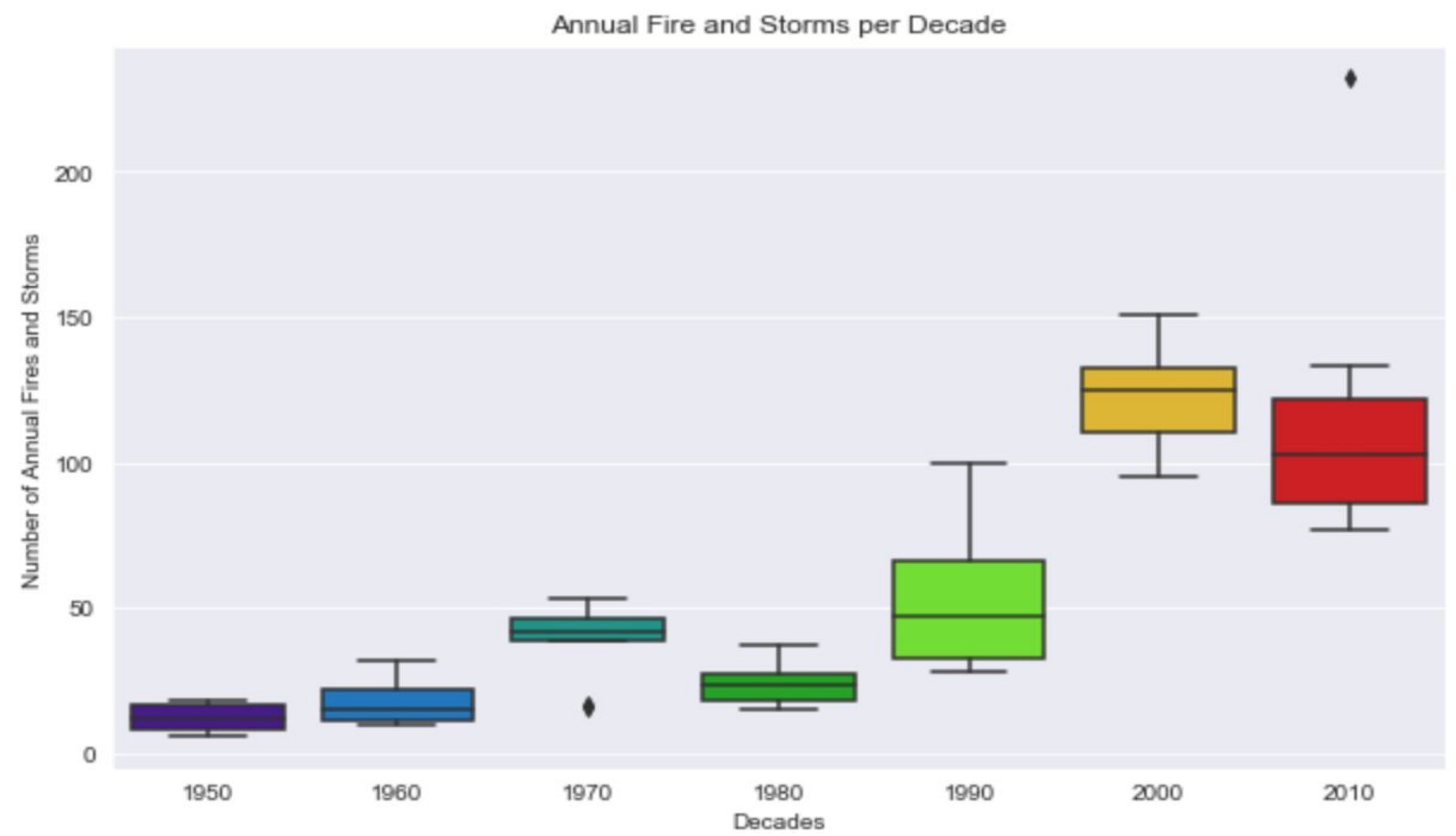
Pairwise Tukey HSD for Average Global Temperature by Decade
Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
1900	1910	-0.0024	-0.2538	0.249	False
1900	1920	0.0951	-0.1563	0.3465	False
1900	1930	0.217	-0.0344	0.4684	False
1900	1940	0.3088	0.0573	0.5602	True
1900	1950	0.1573	-0.0941	0.4087	False
1900	1960	0.1196	-0.1318	0.371	False
1900	1970	0.1407	-0.1107	0.3921	False
1900	1980	0.4389	0.1875	0.6903	True
1900	1990	0.6901	0.4387	0.9415	True
1900	2000	0.9326	0.6811	1.184	True
1910	1920	0.0975	-0.1472	0.3422	False
1910	1930	0.2194	-0.0253	0.4641	False
1910	1940	0.3111	0.0664	0.5558	True
1910	1950	0.1597	-0.085	0.4044	False
1910	1960	0.122	-0.1227	0.3667	False
1910	1970	0.1431	-0.1016	0.3878	False
1910	1980	0.4413	0.1965	0.686	True
1910	1990	0.6924	0.4477	0.9371	True
1910	2000	0.9349	0.6902	1.1796	True
1920	1930	0.1219	-0.1228	0.3666	False
1920	1940	0.2136	-0.0311	0.4583	False
1920	1950	0.0622	-0.1825	0.3069	False
1920	1960	0.0245	-0.2202	0.2692	False
1920	1970	0.0456	-0.1991	0.2903	False
1920	1980	0.3438	0.0991	0.5885	True
1920	1990	0.595	0.3503	0.8397	True
1920	2000	0.8374	0.5927	1.0821	True
1930	1940	0.0917	-0.153	0.3364	False
1930	1950	-0.0597	-0.3044	0.185	False
1930	1960	-0.0974	-0.3421	0.1473	False
1930	1970	-0.0763	-0.321	0.1684	False
1930	1980	0.2219	-0.0228	0.4666	False
1930	1990	0.4731	0.2284	0.7178	True
1930	2000	0.7155	0.4708	0.9602	True
1940	1950	-0.1515	-0.3962	0.0932	False
1940	1960	-0.1891	-0.4338	0.0556	False
1940	1970	-0.168	-0.4127	0.0767	False
1940	1980	0.1301	-0.1146	0.3748	False
1940	1990	0.3813	0.1366	0.626	True
1940	2000	0.6238	0.3791	0.8685	True
1950	1960	-0.0377	-0.2824	0.207	False
1950	1970	-0.0165	-0.2612	0.2282	False
1950	1980	0.2816	0.0369	0.5263	True
1950	1990	0.5328	0.2881	0.7775	True
1950	2000	0.7753	0.5306	1.02	True
1960	1970	0.0211	-0.2236	0.2658	False
1960	1980	0.3193	0.0746	0.564	True
1960	1990	0.5705	0.3258	0.8152	True
1960	2000	0.8129	0.5682	1.0576	True
1970	1980	0.2981	0.0534	0.5428	True
1970	1990	0.5493	0.3046	0.794	True
1970	2000	0.7918	0.5471	1.0365	True
1980	1990	0.2512	0.0065	0.4959	True
1980	2000	0.4937	0.249	0.7384	True
1990	2000	0.2425	-0.0022	0.4872	False

Hypothesis 2: Fires and Storms

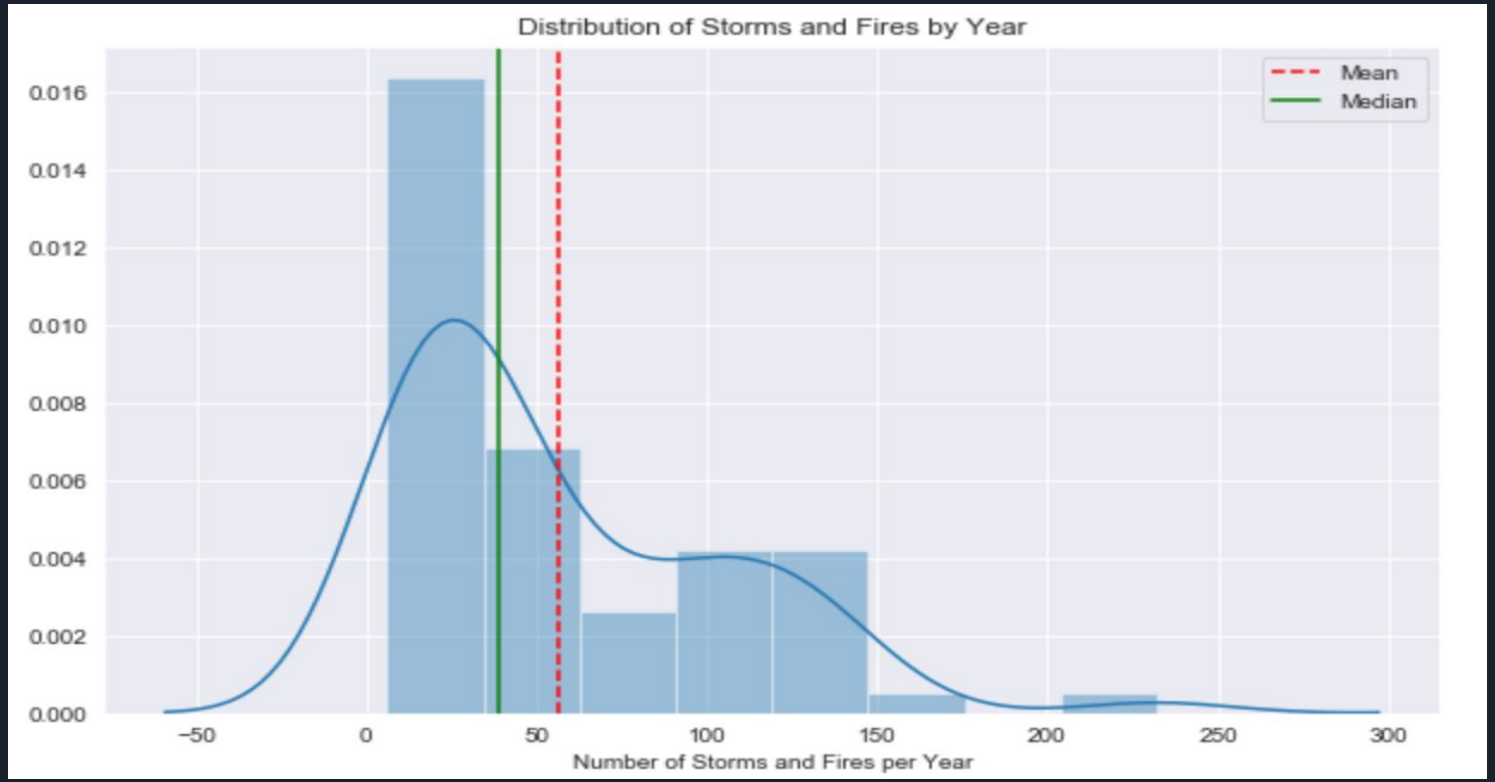
Null hypothesis: There is no difference on the occurrences of fires and storms across decades.

Alternative hypothesis: There is a difference on the occurrences of fires and storms across decades.


Annual Fire and Storms per Decade
Number of Annual Fires and Storms
0
50
100
150
200
1950
1960
1970
1980
1990
2000
2010
Decades

Mean = 56.4179
Variance = 2266.73

Kurtosis = 5.986
Skew = 1.022



Hypothesis Testing: Fires and Storms

Null hypothesis: There is no difference on the occurrences of fires and storms across decades.

Alternative hypothesis: There is a difference on the occurrences of fires and storms across decades.

f-critical value = 3.93

Given the results: Reject the Null Hypothesis

ANOVA Test for Storms and Fires by Decade

	df	sum_sq	mean_sq	F	PR(>F)
decade	1.0	5.925145	5.925145	116.101657	8.945390e-19
Residual	107.0	5.460650	0.051034	NaN	NaN

Pairwise Tukey HSD: Storms and Fires in US by Decade

Pairwise Tukey HSD for Storms and Fires by Decade
Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
1950	1960	5.3143	-27.5283	38.1568	False
1950	1970	26.4143	-6.4283	59.2568	False
1950	1980	11.2143	-21.6283	44.0568	False
1950	1990	39.8143	6.9717	72.6568	True
1950	2000	110.9143	78.0717	143.7568	True
1950	2010	102.0143	69.1717	134.8568	True
1960	1970	21.1	-8.7041	50.9041	False
1960	1980	5.9	-23.9041	35.7041	False
1960	1990	34.5	4.6959	64.3041	True
1960	2000	105.6	75.7959	135.4041	True
1960	2010	96.7	66.8959	126.5041	True
1970	1980	-15.2	-45.0041	14.6041	False
1970	1990	13.4	-16.4041	43.2041	False
1970	2000	84.5	54.6959	114.3041	True
1970	2010	75.6	45.7959	105.4041	True
1980	1990	28.6	-1.2041	58.4041	False
1980	2000	99.7	69.8959	129.5041	True
1980	2010	90.8	60.9959	120.6041	True
1990	2000	71.1	41.2959	100.9041	True
1990	2010	62.2	32.3959	92.0041	True
2000	2010	-8.9	-38.7041	20.9041	False

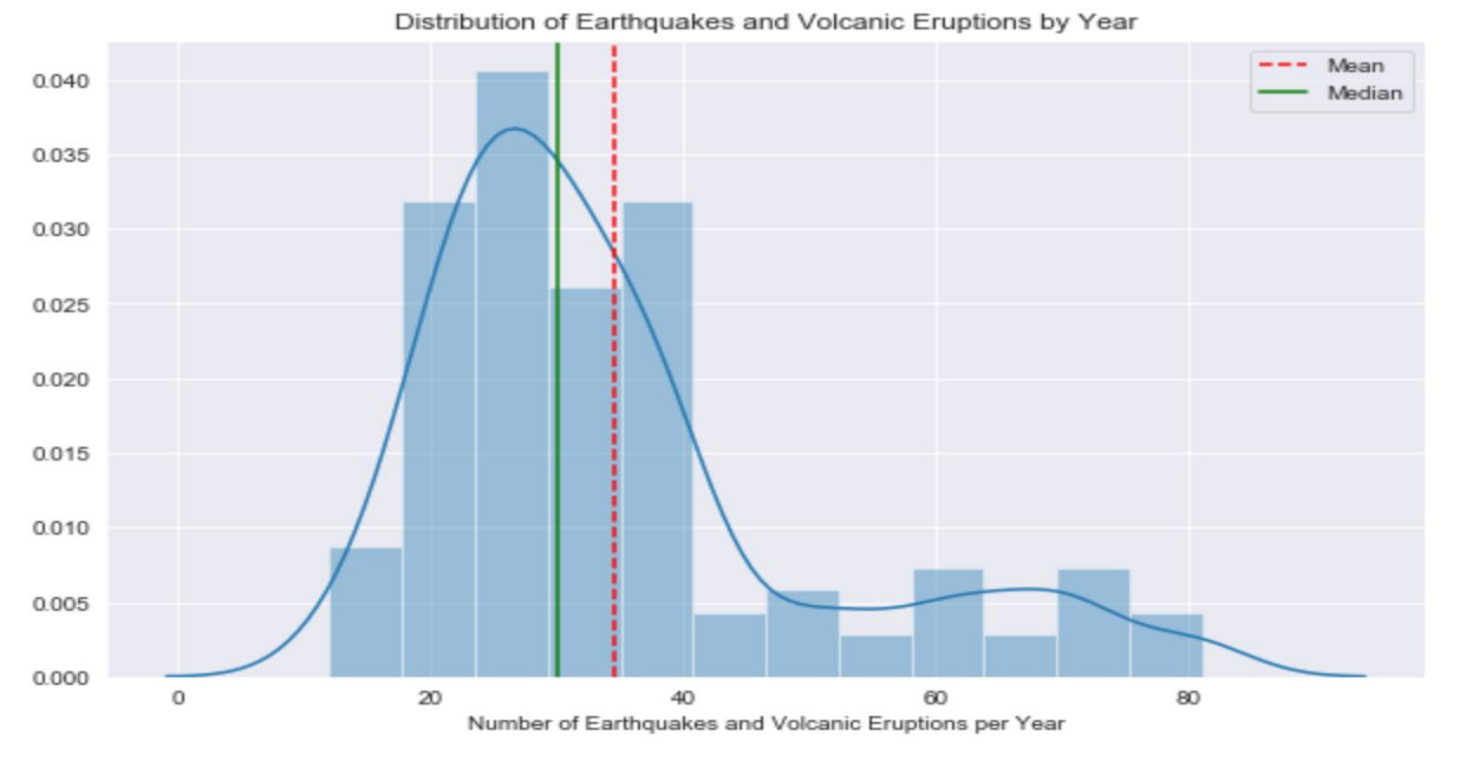
Hypothesis 3: Earthquakes and Volcano Eruptions

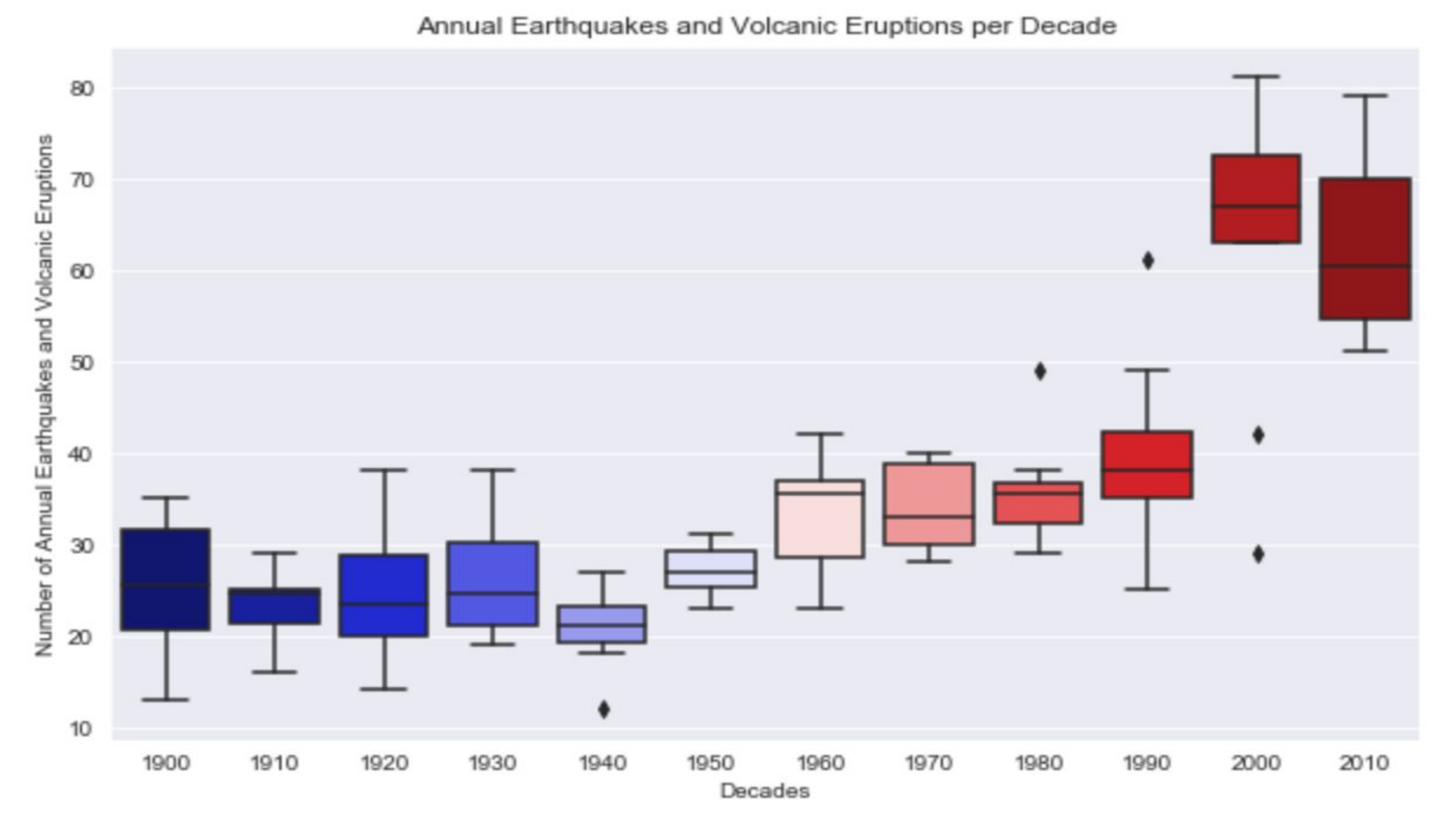
Null hypothesis: There is no difference in the number of occurrences of earthquakes and volcano eruptions across decades.

Alternative hypothesis: Number of occurrences of earthquakes and volcano eruptions is different across decades.

Mean = 34.5833
Variance = 248.6148

Kurtosis = 3.311
Skew = 0.0734




Annual Earthquakes and Volcanic Eruptions per Decade
Number of Annual Earthquakes and Volcanic Eruptions
80
70
60
50
40
30
20
10
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Decades

Hypothesis Testing: Earthquakes

Null hypothesis: There is no difference in the number of occurrences of earthquakes and volcano eruptions across decades.

Alternative hypothesis: Number of occurrences of earthquakes and volcano eruptions is different across decades.

f-critical value = 3.92

Given the results: Reject the Null Hypothesis

ANOVA Test for Earthquakes and Volcanoes by Decade

	df	sum_sq	mean_sq	F	PR(>F)
decade	1.0	16544.402797	16544.402797	149.702851	9.953915e-23
Residual	118.0	13040.763869	110.514948	NaN	NaN

Pairwise Tukey HSD: Global Volcanic Eruptions And Earthquakes

Pairwise Tukey HSD for Earthquakes and Volcanoes by Decade
Multiple Comparison of Means - Tukey HSD,FWER=0.05

group1	group2	meandiff	lower	upper	reject
1900	1910	-2.5	-14.2373	9.2373	False
1900	1920	-1.2	-12.9373	10.5373	False
1900	1930	0.4	-11.3373	12.1373	False
1900	1940	-4.6	-16.3373	7.1373	False
1900	1950	1.6	-10.1373	13.3373	False
1900	1960	7.6	-4.1373	19.3373	False
1900	1970	8.5	-3.2373	20.2373	False
1900	1980	9.9	-1.8373	21.6373	False
1900	1990	14.2	2.4627	25.9373	True
1900	2000	38.2	26.4627	49.9373	True
1900	2010	36.9	25.1627	48.6373	True
1910	1920	1.3	-10.4373	13.0373	False
1910	1930	2.9	-8.8373	14.6373	False
1910	1940	-2.1	-13.8373	9.6373	False
1910	1950	4.1	-7.6373	15.8373	False
1910	1960	10.1	-1.6373	21.8373	False
1910	1970	11.0	-0.7373	22.7373	False
1910	1980	12.4	0.6627	24.1373	True
1910	1990	16.7	4.9627	28.4373	True
1910	2000	40.7	28.9627	52.4373	True
1910	2010	39.4	27.6627	51.1373	True
1920	1930	1.6	-10.1373	13.3373	False
1920	1940	-3.4	-15.1373	8.3373	False
1920	1950	2.8	-8.9373	14.5373	False
1920	1960	8.8	-2.9373	20.5373	False
1920	1970	9.7	-2.0373	21.4373	False
1920	1980	11.1	-0.6373	22.8373	False
1920	1990	15.4	3.6627	27.1373	True
1920	2000	39.4	27.6627	51.1373	True
1920	2010	38.1	26.3627	49.8373	True
1930	1940	-5.0	-16.7373	6.7373	False
1930	1950	1.2	-10.5373	12.9373	False
1930	1960	7.2	-4.5373	18.9373	False
1930	1970	8.1	-3.6373	19.8373	False
1930	1980	9.5	-2.2373	21.2373	False
1930	1990	13.8	2.0627	25.5373	True
1930	2000	37.8	26.0627	49.5373	True
1930	2010	36.5	24.7627	48.2373	True
1940	1950	6.2	-5.5373	17.9373	False
1940	1960	12.2	0.4627	23.9373	True
1940	1970	13.1	1.3627	24.8373	True
1940	1980	14.5	2.7627	26.2373	True
1940	1990	18.8	7.0627	30.5373	True
1940	2000	42.8	31.0627	54.5373	True
1940	2010	41.5	29.7627	53.2373	True
1950	1960	6.0	-5.7373	17.7373	False
1950	1970	6.9	-4.8373	18.6373	False
1950	1980	8.3	-3.4373	20.0373	False
1950	1990	12.6	0.8627	24.3373	True
1950	2000	36.6	24.8627	48.3373	True
1950	2010	35.3	23.5627	47.0373	True
1960	1970	0.9	-10.8373	12.6373	False
1960	1980	2.3	-9.4373	14.0373	False
1960	1990	6.6	-5.1373	18.3373	False
1960	2000	30.6	18.8627	42.3373	True
1960	2010	29.3	17.5627	41.0373	True
1970	1980	1.4	-10.3373	13.1373	False
1970	1990	5.7	-6.0373	17.4373	False
1970	2000	29.7	17.9627	41.4373	True
1970	2010	28.4	16.6627	40.1373	True
1980	1990	4.3	-7.4373	16.0373	False
1980	2000	28.3	16.5627	40.0373	True
1980	2010	27.0	15.2627	38.7373	True
1990	2000	24.0	12.2627	35.7373	True
1990	2010	22.7	10.9627	34.4373	True
2000	2010	-1.3	-13.0373	10.4373	False

Conclusion

Geological events significantly increase in more recent decades, probably due to lack of seismic data collect technology in the early eras.

Annual average global temperature and number of extreme weather events are increasing in more recently decades. Climate change has significantly impacted the global ecosystem. It's time to act now.