Analysis of P&C Claim Loss Data

Objective

The objective of this project is to evaluate whether P&C claim loss severities differ between Arizona and Florida across Fire, Wind, and Water perils using statistical hypothesis testing, and to assess the actuarial implications of any significant differences.

1.Compare Losses by State and Peril (t-tests)

A two-sample t-test was conducted to compare the losses between Arizona and Florida from fire, wind, and water damage each.

- For the null hypothesis (H_0) we suppose that the difference in the estimated mean for the losses from a specific type of peril between Arizona and Florida is equal to zero.
- For the alternative hypothesis (H_1) we suppose that the difference is not equal to zero.

Summary Table for t-test Results

Statistics	Water	Wind	Fire
t-value	-8.00	0.32	-7.90
Degrees of Freedom	244.51	195.57	226.26
p-values	4.97×10^{-14}	0.751	1.19×10^{-13}
Confidence Interval	(-3624.47, -2192.39)	(-8303.33, 11490.44)	(-4464.72, -2682.43)
Significant?	Yes	No	Yes

Interpretation:

- Independent two-sample t-tests were conducted to compare mean losses between Florida and Arizona for Fire, Wind, and Water perils.
- The t-test for Fire losses shows a statistically significant difference between the two states ($t = -7.90$, $p < 0.001$).
- The t-test for Water losses also shows a statistically significant difference between Florida and Arizona ($t = -8.00$, $p < 0.001$).
- In contrast, the t-test for Wind losses does not show a statistically significant difference between the states ($t = 0.32$, $p = 0.751$).
- The confidence intervals for Fire and Water losses do not include zero, reinforcing that the differences are statistically meaningful.
- The confidence interval for Wind losses includes zero, indicating that any observed difference may be due to random variation.
- Overall, the results suggest that state location significantly affects Fire and Water losses, but does not significantly affect Wind losses.

2. Analyze Fire Losses by State (ANOVA)

I conducted a one-way ANOVA test to see whether fire losses differ significantly between Arizona and Florida.

Summary table for ANOVA Results

Statistic	**Fire Loss**
F-value	62.92
p-value	4.09×10^{-14}
Between-Groups DF	1
Within-Groups DF	306
Significant?	Yes

Interpretation:

- The ANOVA results show a strong statistical difference in mean fire losses between Arizona and Florida
- The ANOVA test produced $F(1, 306) = 62.92$, $p = 4.09 \times 10^{-14}$, indicating the difference is extremely unlikely to be due to chance.
- Florida has much higher Fire losses compared to Arizona.
- The F-value shows that the difference between state mean is much greater than the variation within each states.
- The results confirm fire losses vary by state, meaning geographical location affects how severe the fire losses are.

3. Interpretation & Actuarial Implications

According to the results of the data, insurers should use state/location as a key rating variable, primarily for the losses of fire and water. While the t-test and ANOVA show the average fire and and water loss in Florida are quite higher and statistically different than that of Arizona. All the while wind loss is not significantly different. This can imply the underwriting guidelines and base rates for the perils fire and water in Florida should be higher than that of Arizona to represent the greater expected severity. In the eyes of pricing, Florida would require higher premiums, deductibles and tighter eligibility such as that of construction and protection for the building's standards. In the case of Portfolio risk management, Florida's high severities show that insurers in Florida will have a greater tail risk from fire and water related events and may need more reinsurance, stricter limits and even diversification possibly into lower risk areas like Arizona to insure not too much money loss.

Multiple limitations should be accounted for though.

- This data only account for 2 states, so other regions may not have similar results

- T-tests and ANOVA go off reasonably normal residuals and independent observations, so if there is a large skew in the data which is common is insurance data, or if there's clustering of events, the assumptions of the tests may be violated.
- This data only takes the average loss of severity, while things such as the frequency of claims and exposure aren't included in the data.
- The sample size is also a limitation as while it's not exactly small it's still limited and there could be more uncontrolled rating variables that might differ between the states.

Because of said limitations, this data should be used more as a directional guide for those seeking to get involved in insurance. Those who choose to invest in Florida should be ready to support higher pricing and tighter risk management in Florida, but not telling investors that the information provided is for production-ready rating factors.