

## **기상 기후 빅데이터 보고서**

### **기상 빅데이터 분석을 통한 국내 산불 발생 위험 평가 및 예측**

강원대학교 AI SW / 기상 기후 빅데이터

201721196 장준영

# 목 차

Ch1. Abstract

Ch2. Introduction

Ch3. Method

- 1) 데이터 수집
- 2) 기상 요인과 산불 발생 간의 상관성 분석
- 3) 지역별 산불 발생 패턴 분석
- 4) 산불 위험 예측을 위한 통합 지수

Ch4. Results

Ch5. Conclusion

## Ch1. Abstract

지구온난화에 따른 이상기후 현상은 전 세계적으로 다양한 형태의 자연재해 빈도와 강도를 증가시키고 있으며,

특히 산불 발생 빈도의 증가와 그 심각성은 점점 뚜렷해지고 있다.

한국 또한 최근 들어 건조한 날씨와 강한 바람 등의 기상 조건이 지속되면서 산불 발생 빈도와 피해 규모가 급격히 늘어나고 있다.

특히, 대형 산불은 산림 생태계 파괴 뿐 아니라 인명과 재산 피해까지 초래하고 있어 이에 대한 효과적인 대응이 절실한 상황이다.

본 분석은 이러한 산불 발생을 사전에 효과적으로 예측하기 위해 빅데이터 분석을 활용한 예측 모델을 구축하고 평가한다.

이를 위해 기상청, 산림청에서 제공하는 기후 데이터(온도, 습도, 강수량, 풍속) 및 과거 산불 데이터를 결합한 대규모 데이터베이스를 제공한다.

본 분석은 기상 요인 중 산불 발생과 가장 밀접한 관계를 가지는 요소들을 선별하고 이들의 영향력을 정량적으로 평가한다.

이를 통해 온도, 습도, 강수량, 풍속과 같은 기상 요인들이 산불 발생에 미치는 영향을 분석하고, 각 지역의 산불 발생 위험도를 정확히 평가한다.

본 분석의 결과는 향후 산불 예측과 조기 경고 시스템 개발에 직접적으로 활용될 수 있으며, 이를 통해 산불 발생 시 신속한 대응 및 예방 전략 수립이 가능해질 것으로 기대된다.

궁극적으로 본 분석의 성과는 산불로 인한 피해를 최소화하고 산림 자원을 보호하는 데 크게 기여할 것으로 기대한다.

## Ch2. Introduction

기후 변화는 오늘날 인류가 직면한 가장 시급한 환경 위협 중 하나이다. 전 세계적으로 평균 기온이 상승하고 강수 패턴이 변화하면서 태풍, 흥수, 폭염, 가뭄 등과 같은 극단적인 기상 현상의 빈도와 강도가 증가하고 있다. 이러한 현상은 환경 파괴 뿐 아니라 경제적, 사회적 손실로 이어지고 있으며, 특히 최근 몇 년 동안 산불 발생 빈도가 급격히 증가하여 많은 국가에서 큰 피해를 야기하고 있다.

한국 역시 예외는 아니다. 한국은 지리적으로 사계절이 뚜렷하고 건조한 계절이 존재하여 산불 발생의 환경적 조건이 자연스럽게 형성되어 있다. 하지만 최근 들어 봄철과 초여름 기간 동안 대기가 장기간 건조하고 강풍이 지속되는 등 기상 조건이 악화되면서 산불 발생이 빈번해지고 피해 규모도 크게 확대되고 있다. 대표적인 사례로는 2022년 동해안 산불과 2023년 강릉과 경북 지역 산불, 가장 최근에 발생한 2025년 3월 전국 동시다발적으로 발생한 산불이 있다. 이 산불들은 막대한 산림 손실과 생태계 파괴를 초래했을 뿐 아니라 수많은 인명 피해와 수십억 원의 경제적 손실을 가져왔다.

그러나 기존의 산불 예측 및 대응 시스템은 주로 역사적인 사례와 경험적 통계 방법에 의존하고 있어 빠르게 변화하는 기상 환경에서 정확한 예측과 효과적인 대응이 어려운 실정이다.

특히 기후 변화로 인해 기상 패턴이 이전과는 다른 형태로 나타나고 있어, 기존의 예측 모델로는 한계가 명확히 드러나고 있다.

이에 따라 새로운 접근 방식으로 빅데이터 기반의 정밀하고 정량적인 분석 기법이 필요하게 되었다.

본 분석은 이러한 문제의식을 바탕으로 기상 데이터와 역사적 산불 발생 데이터를 종합적으로 분석하여 산불 발생 위험을 평가하고 효과적으로 예측할 수 있는 새로운 분석 방법을 찾고자 한다.

특히 산불 발생에 직접적으로 영향을 미치는 기상 조건을 면밀히 분석하고, 각 기상 요소들이 산불 발생에 미치는 영향의 강도를 평가하여 지역별 산불 발생 가능성 을 보다 정확하게 예측하는 것이 본 분석의 목표다.

## Ch3. Method

### 1. 데이터 수집

본 분석에서는 한국 기상청에서 제공하는 산불이 자주 발생하는 지역인 강원 영동 지방, 경북 북부 지방의 2016년부터 2025년까지 3, 4월의 일일 기상 데이터를 수집한다. 기상 데이터는 온도, 습도, 강수량, 풍속 등 산불 발생에 영향을 미치는 주요 요소들을 포함하고 있다.

또한, 산림청으로부터 동일 기간 동안 발생한 전국의 산불 발생 기록을 수집하여 지역별, 시기별 산불 발생 현황을 파악할 수 있도록 한다.

이 데이터는 정확한 시공간적 기준을 통해 결합하여 하나의 통합된 빅데이터로 구축한다.

### 2. 기상 요인과 산불 발생 간의 상관성 분석

본 분석은 산불 발생과 기상 요소 간의 상관성을 규명하기 위해 통계적 접근을 채택한다. 온도, 습도, 강수량, 풍속 등 주요 기상 요소와 산불 발생 가능성 간의 관계를 다양한 분석을 통해 정량적으로 평가한다.

이를 통해 산불 발생과 가장 높은 상관성을 보이는 기상 요소를 선별하고, 이들 요인이 산불 위험 평가에 중요한 변수로 활용될 수 있는지를 심도 있게 탐색한다.

온도, 습도, 강수량, 풍속이 산불 발생 가능성을 Spearman correlation 분석을 진행 결과 온도의  $r$  값 : 0.885로 매우 강한 양의 상관관계를 보인다.

습도의  $r$  값 : -0.626으로 강한 음의 상관관계를 보인다.

풍속의  $r$  값 : 0.097으로 거의 무의미한 상관관계를 보인다.

\*\* "기상청 API에서 제공된 강수량 데이터가 -9.0으로 통일된 값으로 입력되어 있어 결측치로 처리  
온도, 습도, 풍속 모두 p-value 값은 0에 수렴하여, 유의미한 요소들임을 확인했다.

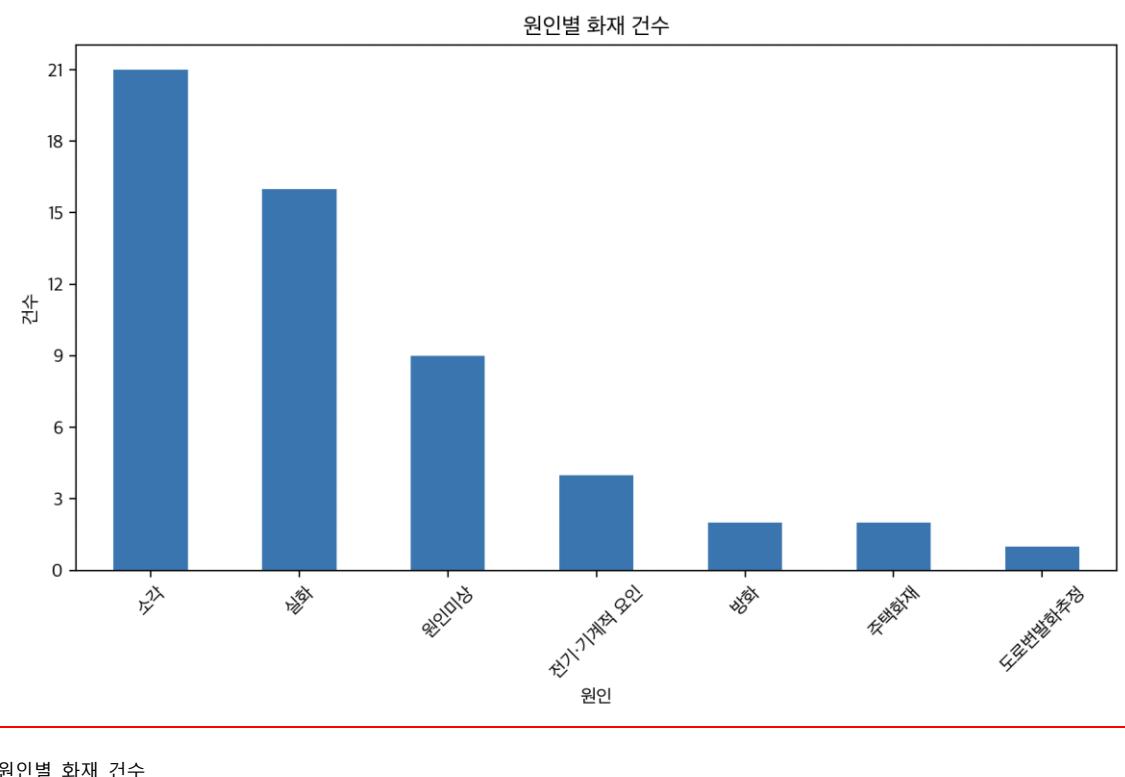
### 3. 지역별 산불 발생 패턴 분석

공간적 분석을 통해 지역별 산불 발생 빈도와 강도를 시각화하고,

이를 기상 조건과 결합하여 지역별 위험도를 평가한다.

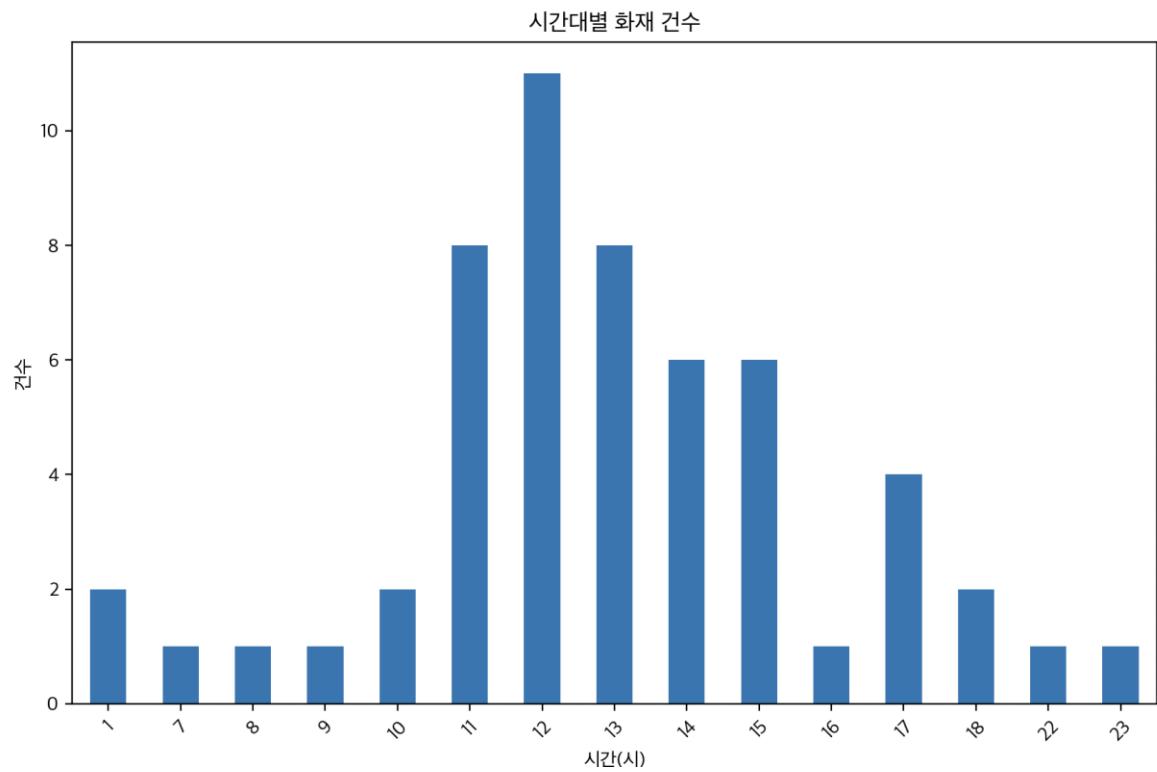
산불 발생 위험이 높은 지역을 명확하게 시각화하고,

각 지역별로 산불 발생 가능성이 높은 기상 조건을 식별한다.



원인 별 화재 건수로 보면 소각, 실화 같은 일상적인 사유가 제일 주된 원인이다.

또한 원인들을 살펴보면, 자연적인 원인 혹은 원인 미상과 같은 사유보다 인위적으로 발생한 사건이 주된 요인이다.



시간대별 화재 건수

화재가 주로 발생한 시간대는 11 ~ 13 시로 낮 시간이며,  
인위적인 원인으로 발생한 사건이 많기 때문에 낮 시간에 화재가 많이 발생한다.

## 4. 산불 위험 예측을 위한 통합 지수

분석된 기상 요소 간 상관성과 지역적 특성을 종합하여 산불 위험도를 나타내는 통합 지수로 계산한다.

이 지수는 다양한 기상 조건과 지역별 특성을 종합적으로 반영하여 산불 발생 가능성을 수치적으로 표현하는 데 목적이 있다.

산출된 통합 지수는 향후 국가적 산불 예방과 대응 체계 수립에 효과적으로 활용될 수 있을 것이다.

산불 발생에 영향을 주는 기상요소는 습도, 풍속, 온도, 강수량 순이다.

\* Fire Weather Index (FWI : 산불기상지수) : 기상요소(온도, 습도, 강수량, 풍속 등)를 이용해 산불 위험도를 정량적으로 나타내는 지수로, 캐나다 산불관리청에서 개발한 시스템이다. (이를 조금 더 간단하고 쉽게 바꿔 조정했다.)

\* 이하 편의성을 위해 FWI 지수라 칭한다.

FWI 지수를 이용하여 각 지역의 산불 위험 예측한다.

지역은 주로 산불이 발생하는 지역인 강원 영동지방과 경북 북부지방의 데이터를 이용한다.

### FWI 지수 공식

$$FWI = (0.1 * 평균기온 + 0.2 * 평균풍속) * (1 - (상대습도 / 100))$$

강원 영동 지역, 경북 북부 지역의 관측소는 약 14 곳이였으나  
데이터가 존재하는 관측소는 8 곳이였다.

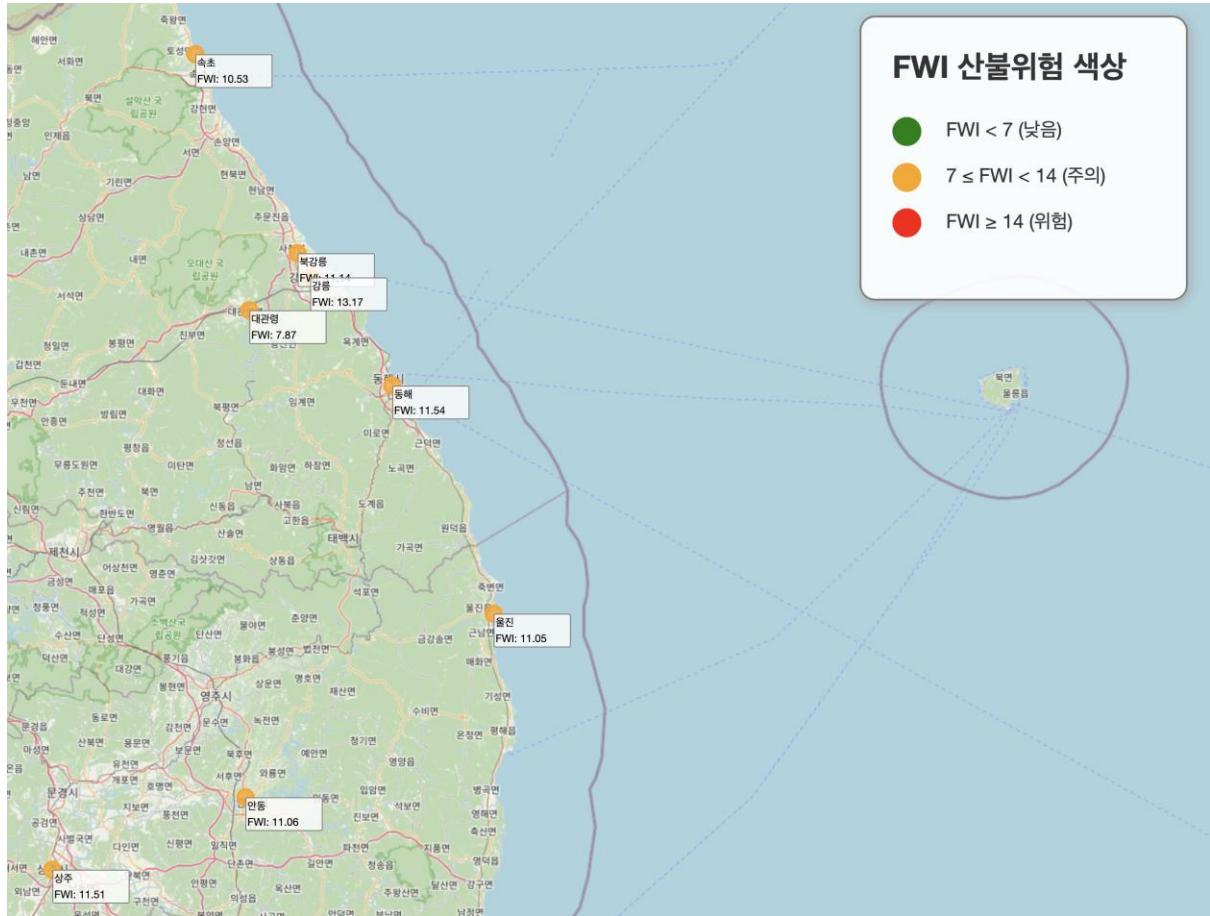
2016년 ~ 2025년의 3, 4월 데이터에서 관측소의 데이터를 평균을 구하고,  
기상청 API 데이터에서 제공하는 관측소의 위도, 경도를 이용해 위치를 구했다.  
이를 Python의 library인 folium을 이용하여 지도위에 시각화를 진행했다.

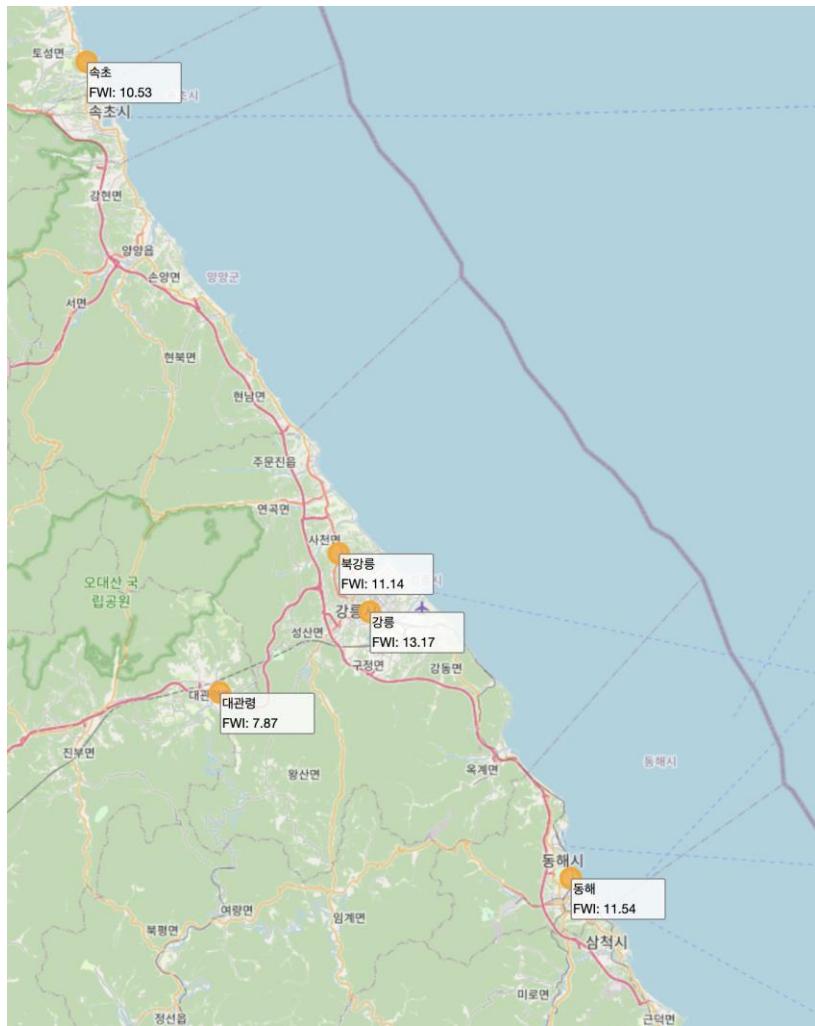
추가적으로 FWI 지수와 실제 화재 발생 건수를 이용하여 산불 발생 확률과 실제 화재 발생을 비교한다.

FWI 지수가 7 미만은 낮음, 7 이상 14 미만은 주의, 14 이상은 위험

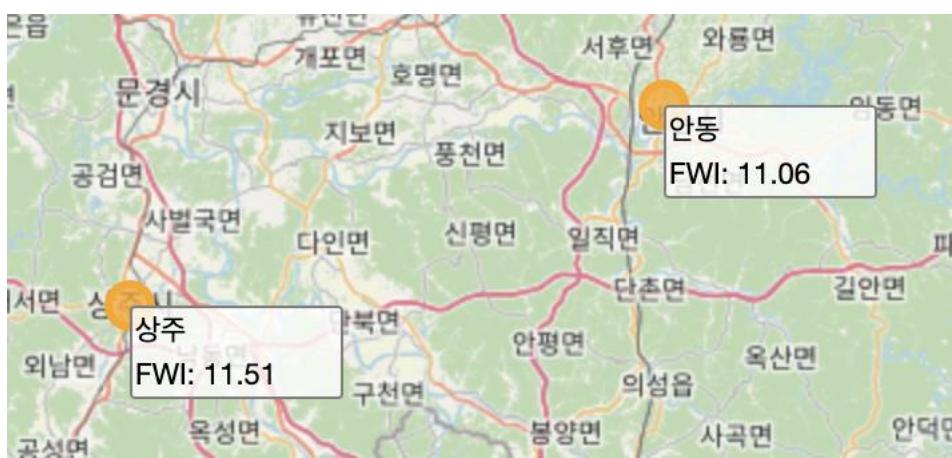
## Ch4. Results

하단의 이미지에 관측소들에서 FWI 지수 확인 가능

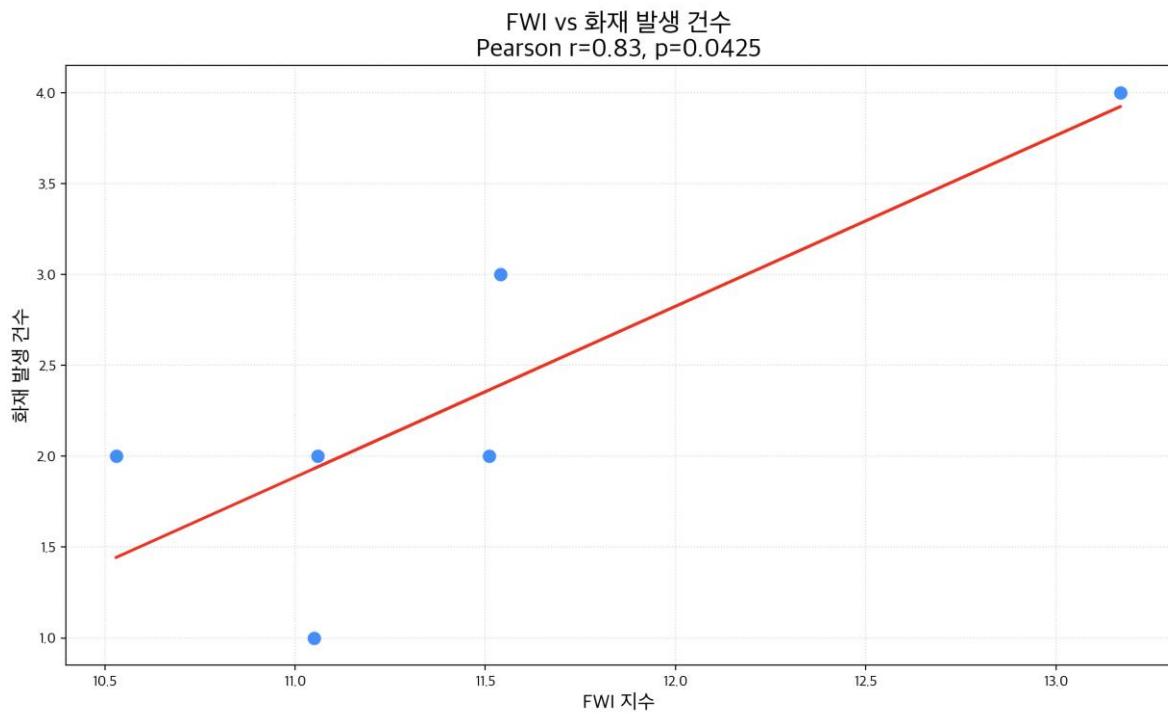




시인성을 위해 확대한 강원 영동 지방의 FWI 지수



시인성을 위해 확대한 강원 영동 지방의 FWI 지수



FWI 와 화재 발생 건수의 피어슨 상관분석 그래프

피어슨 상관계수  $r = 0.83$  으로 강한 양의 상관관계를 보인다.

이는 FWI 지수가 높을수록 실제 화재 발생 건수가 증가하는 경향 존재함을 보여준다.

또한, 유의확률도 통계적으로 유의미함을 보여준다.

결과적으로 FWI 지수가 높아지면, 실제 산불 발생도가 증가한다.

이를 통해, FWI 지수는 실제 발생과 관련이 있다고 본다.

## Ch.5 Conclusion

본 연구는 최근 증가하고 있는 국내 산불 발생 문제에 효과적으로 대응하기 위해 기상 빅데이터를 활용한 산불 발생 위험 평가 및 예측 방법을 제안하고 분석하였다. 이를 위해 2016년부터 2025년까지 강원 영동 지역과 경북 북부 지역의 주요 관측소 데이터를 바탕으로 기온, 습도, 풍속과 같은 기상 요소들이 산불 발생에 미치는 영향을 체계적으로 평가하였다.

분석 결과, 산불 발생과 가장 강한 상관관계를 보이는 기상 요소는 온도와 습도였으며, 풍속은 상대적으로 영향력이 낮았다. 특히 온도는 산불 발생과 매우 강한 양의 상관성을 보였고, 습도는 강한 음의 상관성을 나타냈다. 이를 바탕으로 구축된 통합 지수인 FWI 지수는 산불 발생 위험을 정량적으로 평가하는 데 매우 효과적이었다. 실제 산불 발생 데이터를 이용한 피어슨 상관 분석 결과, FWI 지수와 산불 발생 건수 간의 상관계수가 0.83으로 매우 높은 상관성을 보였다. 이는 FWI 지수가 높을수록 실제 산불 발생 위험도 함께 증가하는 경향을 잘 반영하고 있음을 시사한다.

또한 지역별 산불 발생 빈도와 강도를 GIS를 통해 시각적으로 표현하여, 각 지역의 산불 위험성을 명확히 파악할 수 있도록 하였다. 이를 통해 지역별로 맞춤화된 산불 예방 및 대응 전략 수립에 실질적인 도움을 줄 것으로 기대된다.

본 연구의 결과는 향후 실질적인 산불 예방과 조기 경고 시스템 구축에 유용하게 활용될 수 있다. 특히, 기상 데이터를 기반으로 한 정확한 산불 위험 예측은 인명과 재산 피해를 최소화하고 산림 자원 보호에 큰 기여를 할 수 있을 것으로 전망된다. 다만 본 연구는 일부 데이터의 결측치나 한정된 관측소 데이터라는 한계점을 가지므로, 향후 보다 많은 관측 데이터를 확보하고, 추가적인 기상 및 인위적 요인까지 종합적으로 고려한 확장된 분석이 필요할 것으로 판단된다.