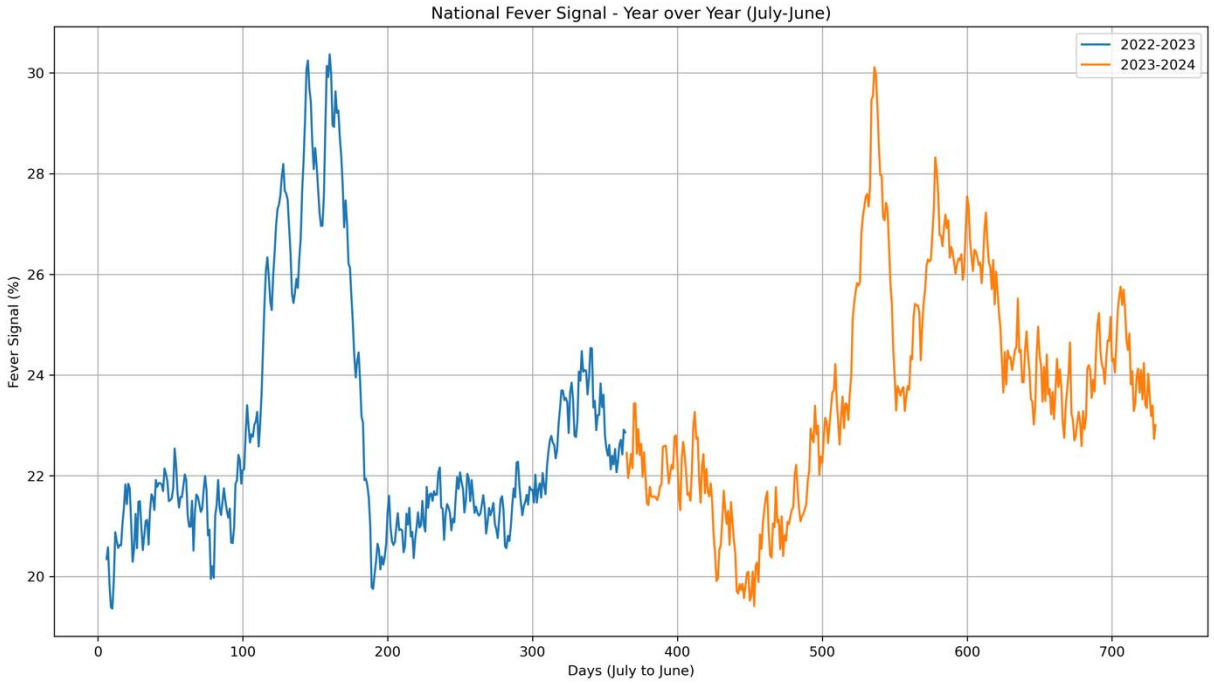
Fever Signal Analysis and Anomaly Detection:

A Spatio-temporal Study of National Health Patterns, 2022-2024

Joel Skaria

Executive Summary



- Key Findings: Dataset: 843,935 observations over 730 days (2022-2024)
- Identified 21 significant anomaly days (2.88% of timespan)
- Distinct seasonal patterns in fever signals
- Regional variations between states
- Different patterns between adult/pediatric populations

Part 1: Data Exploration

Dataset Overview

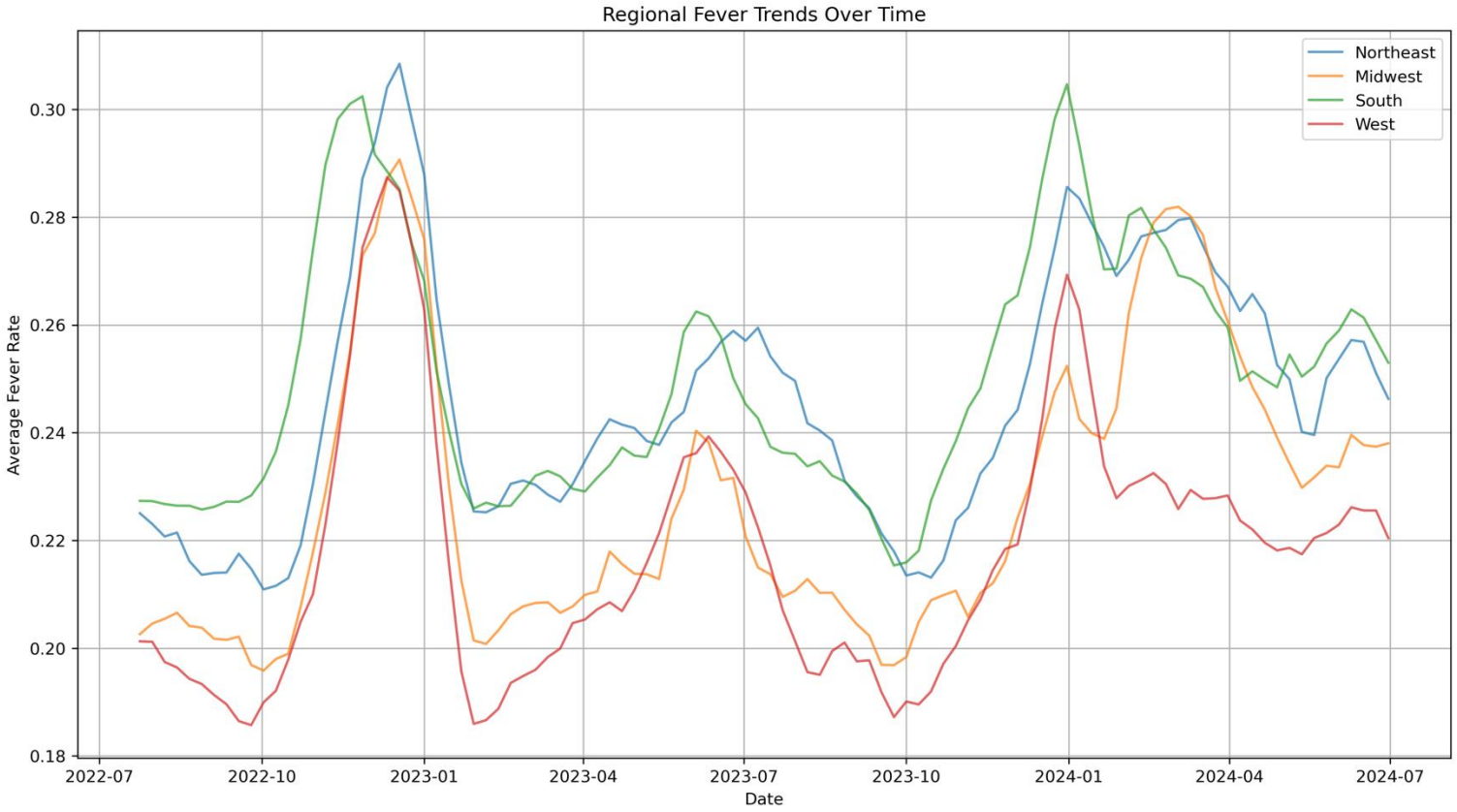
VOLUME: 843,935 OBSERVATIONS

TIMEFRAME: JULY 2022 - JUNE 2024

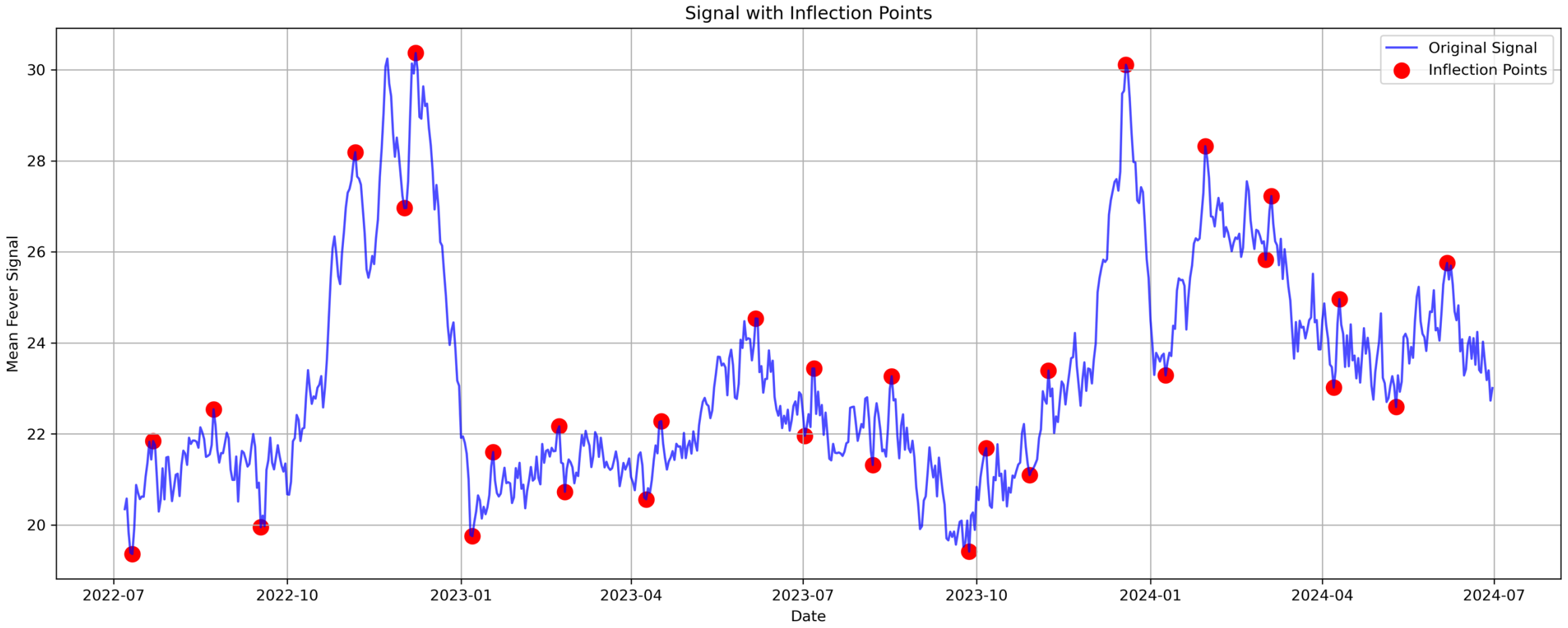
GEOGRAPHIC COVERAGE

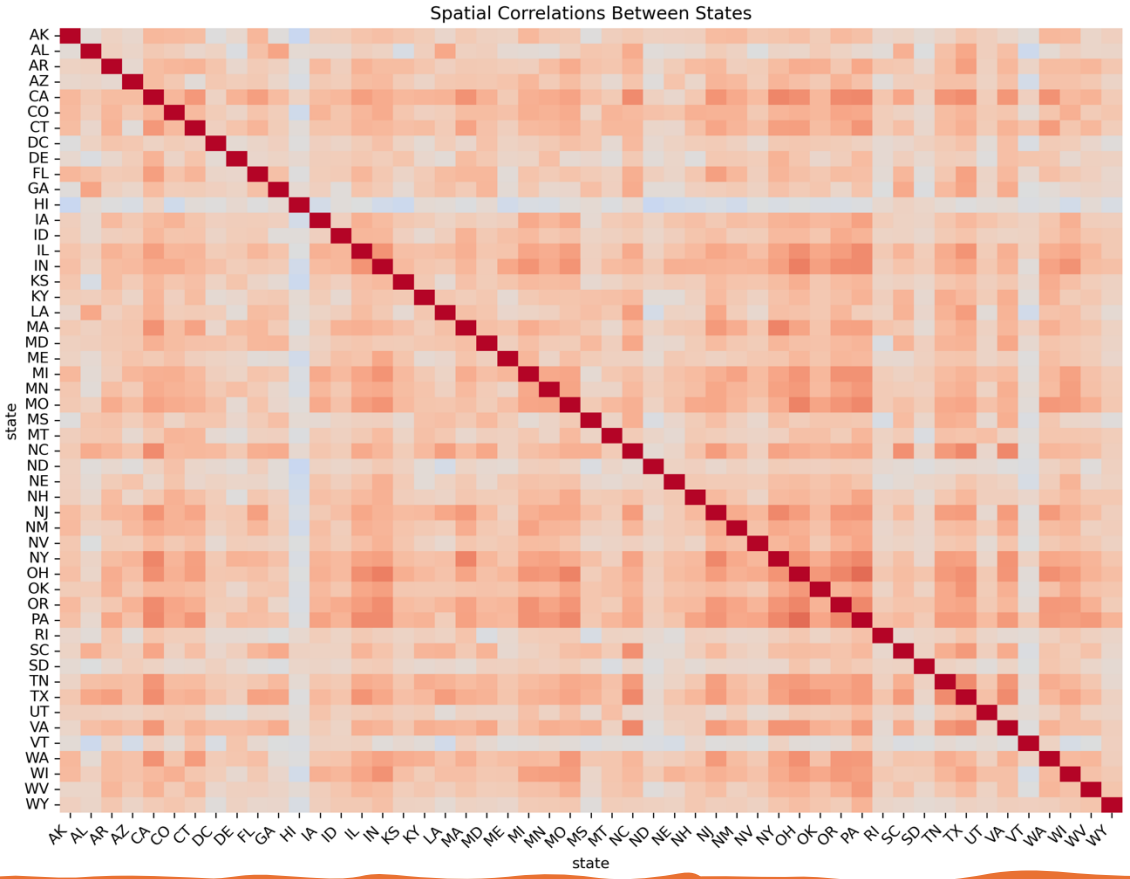
Demographic Analysis

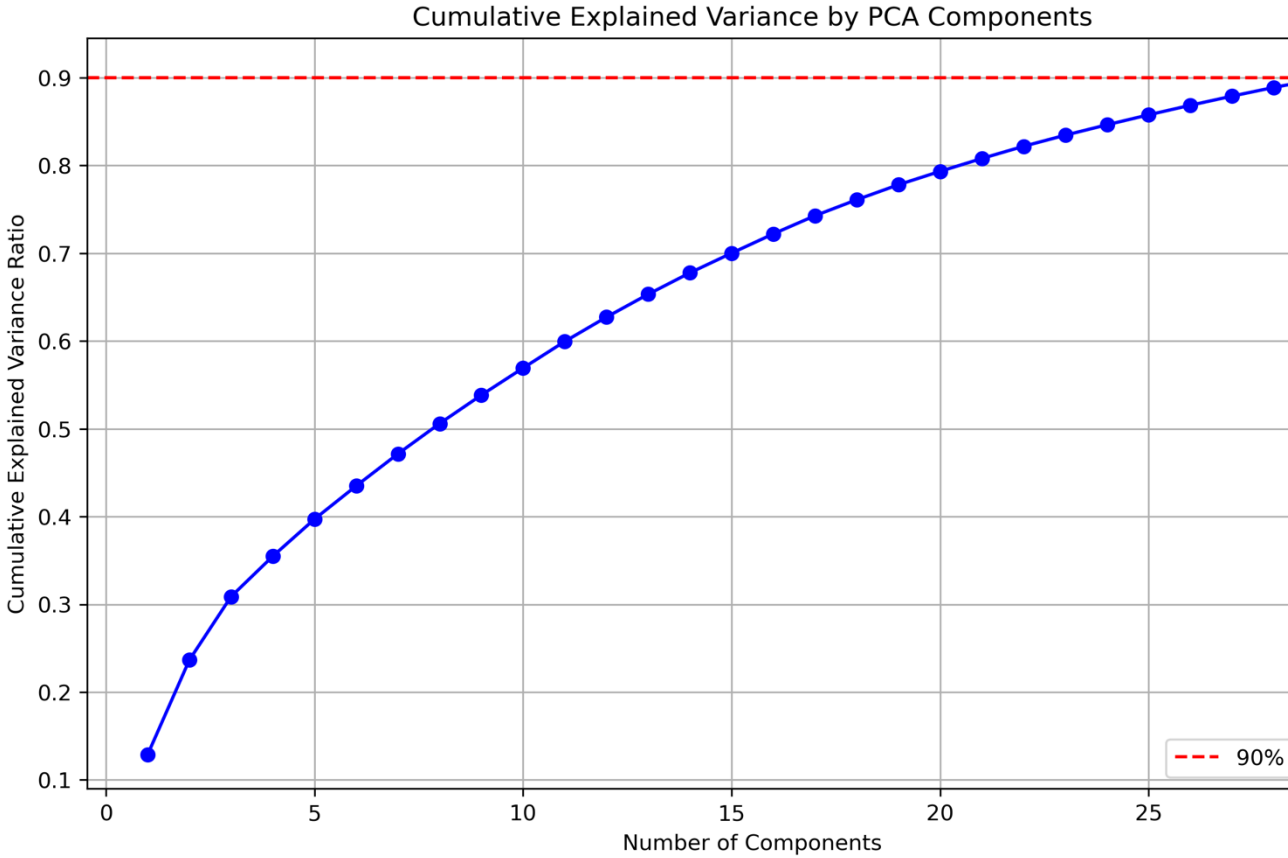
Regional representation



Signal Patterns and Inflection Points







Statistical Analysis

- PCA decomposition results (90% variance explained by 30 components)
- Spatial correlations between states

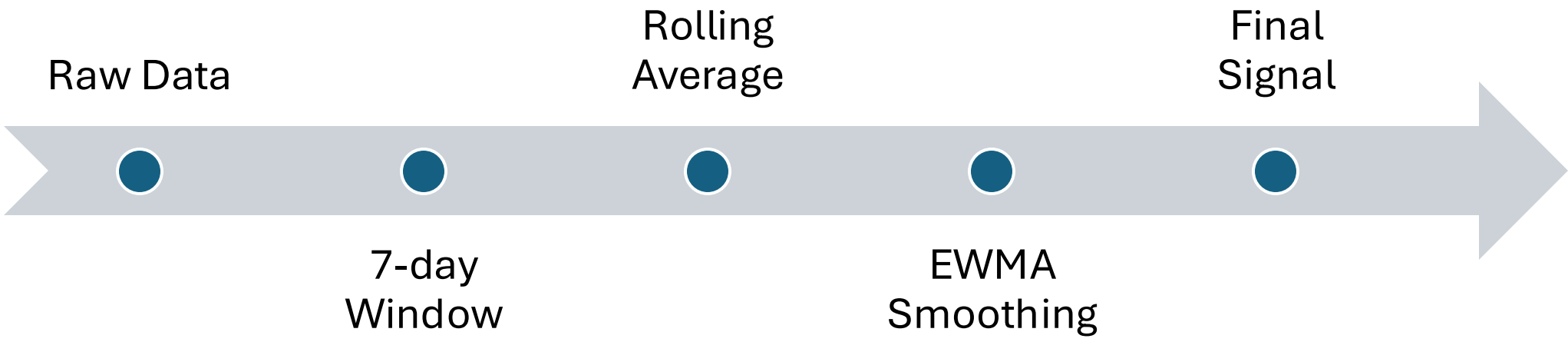
Part 2: Signal Creation and Analysis

Methodology

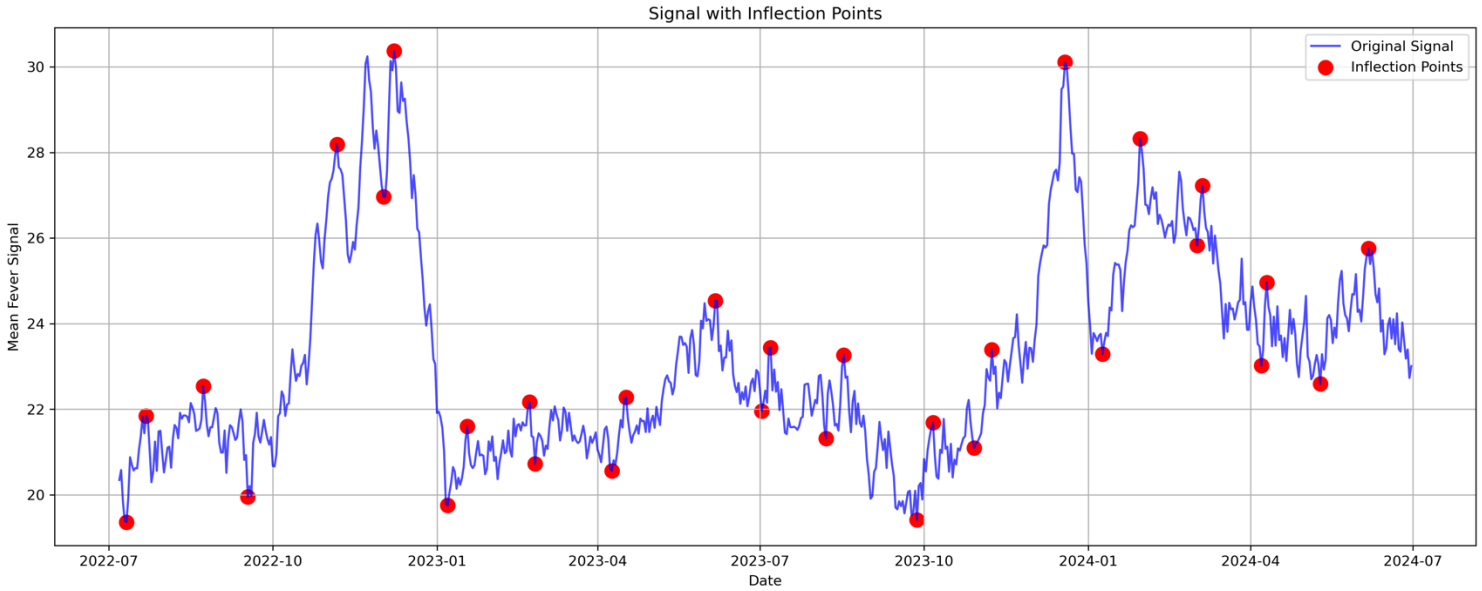
- Denominator Selection Rationale:
  - 1. 7-day Rolling Average - Smooths out daily reporting variations - Captures weekly cyclical patterns - Reduces impact of weekend/holiday effects
  - 2. EWMA Smoothing - Gives more weight to recent observations - Reduces noise while preserving trends - Better handles irregular reporting patterns

Methodology Explained

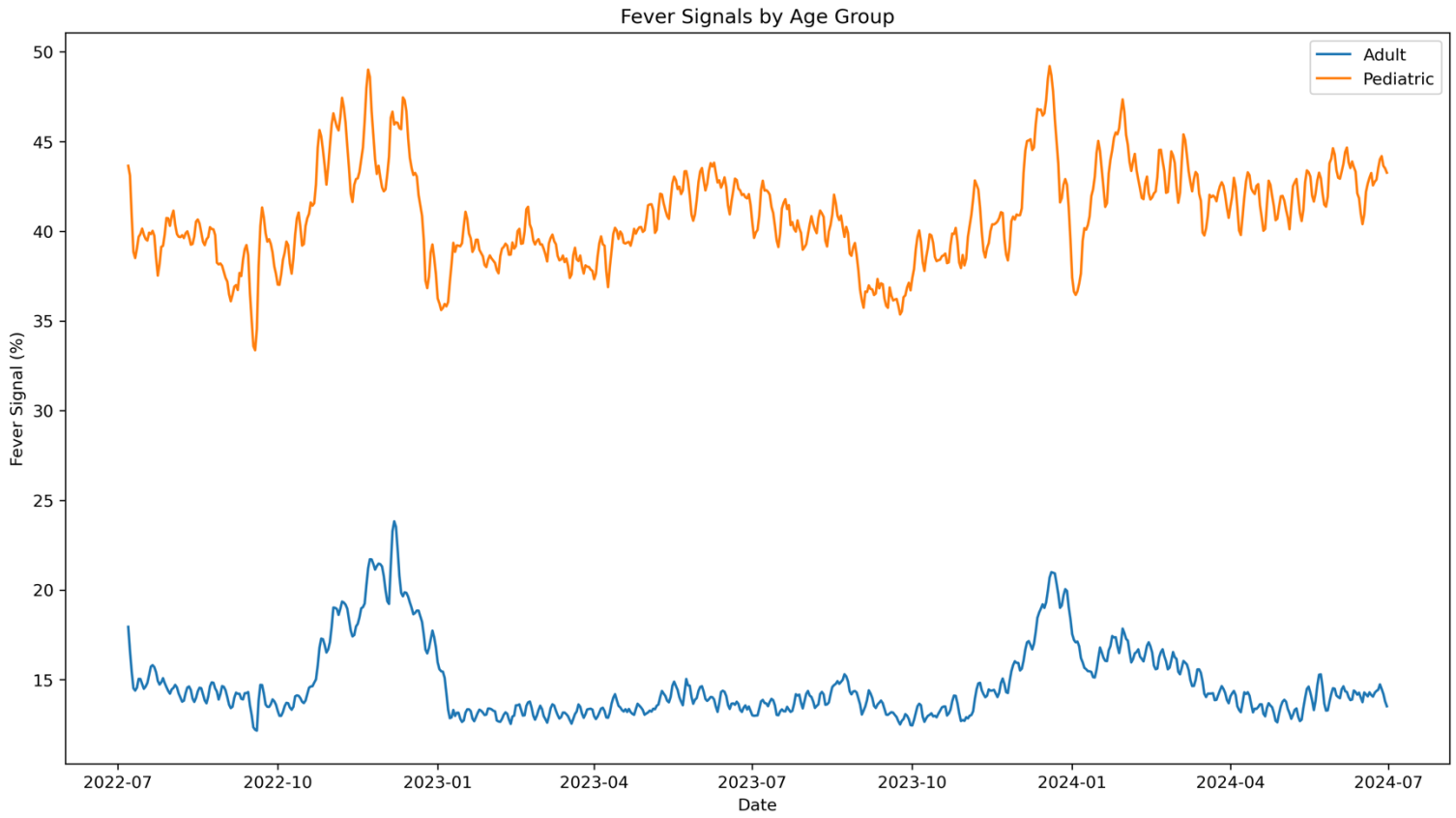
- 7-day window captures weekly patterns
- Rolling average smooths reporting variations
- EWMA gives more weight to recent observations
- Final signal balances noise reduction with trend preservation



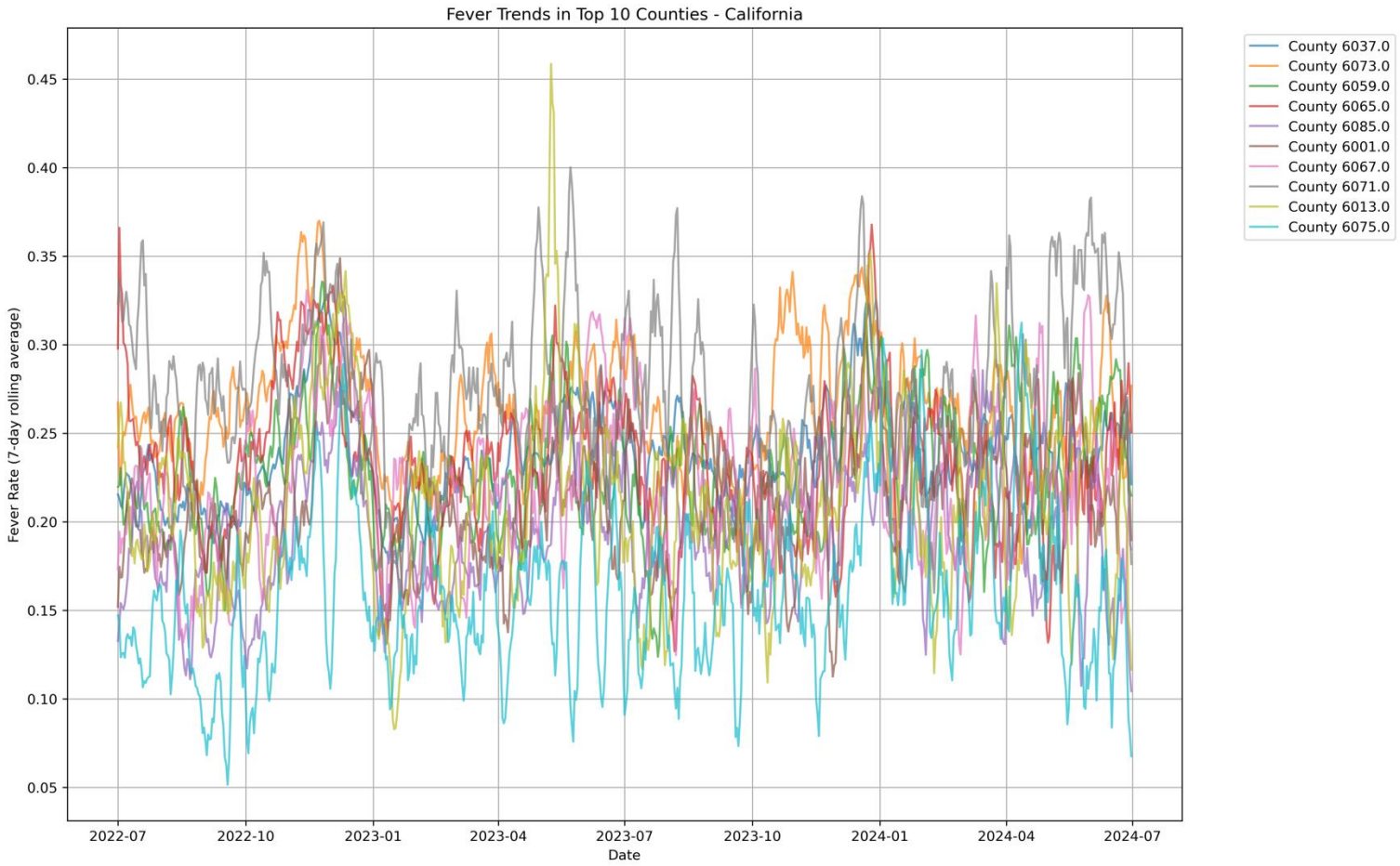
National Signal



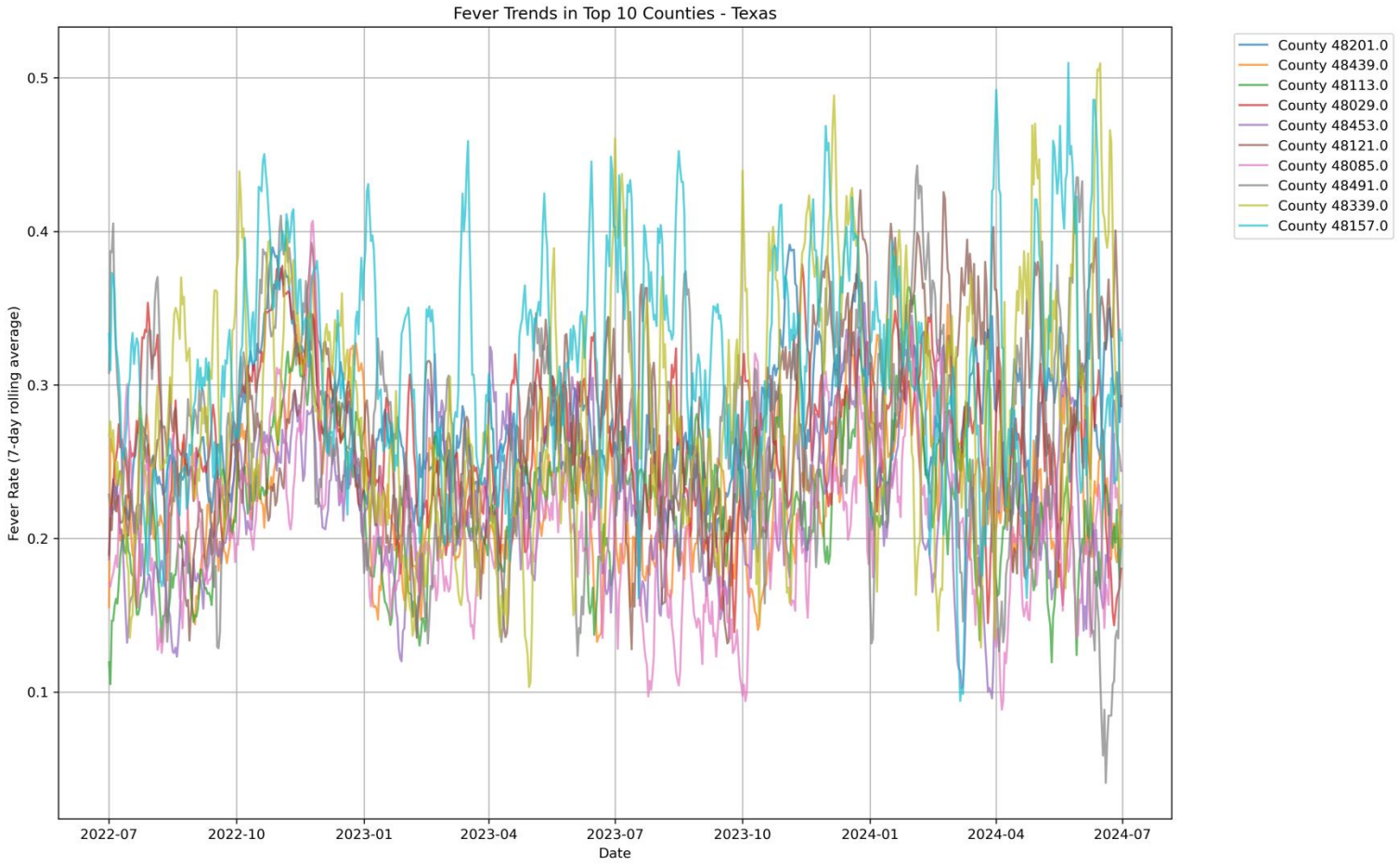
Age Group Comparison



Geographic Analysis - California

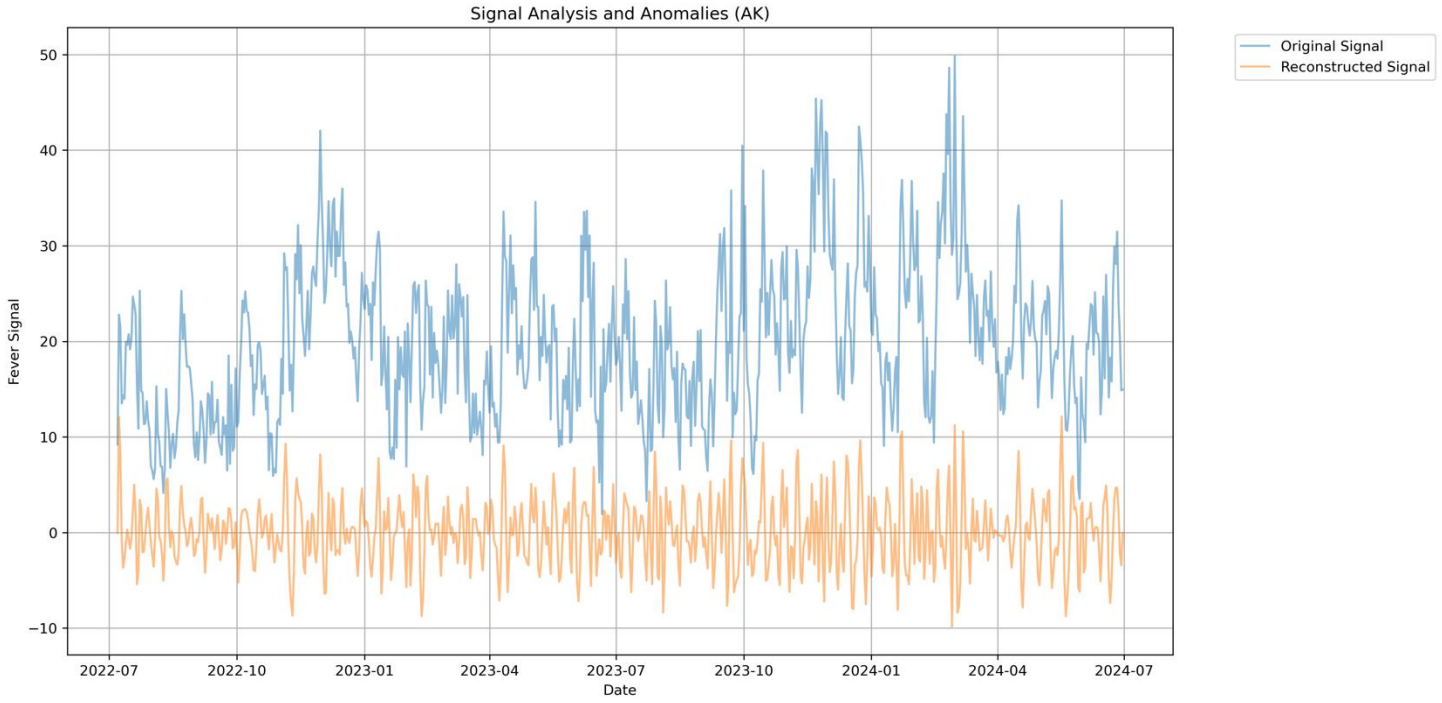


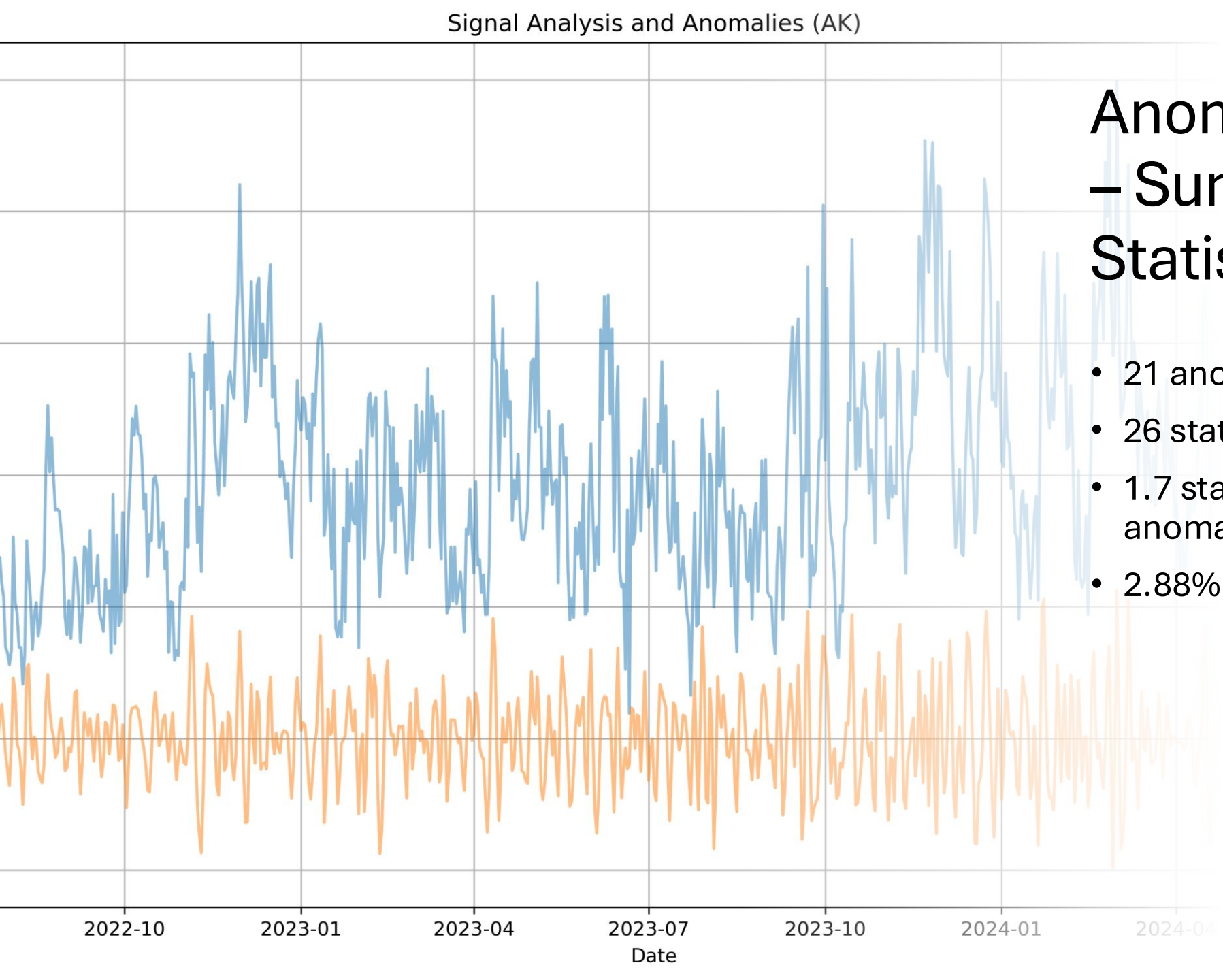
Geographic Analysis - Texas



Anomaly Detection Methodology

- Methodology: Signal decomposition using PCA (90% variance explained)
  - Bandpass filtering to isolate relevant frequencies
  - Statistical thresholds (3σ and 2.5σ)
  - Multi-state validation

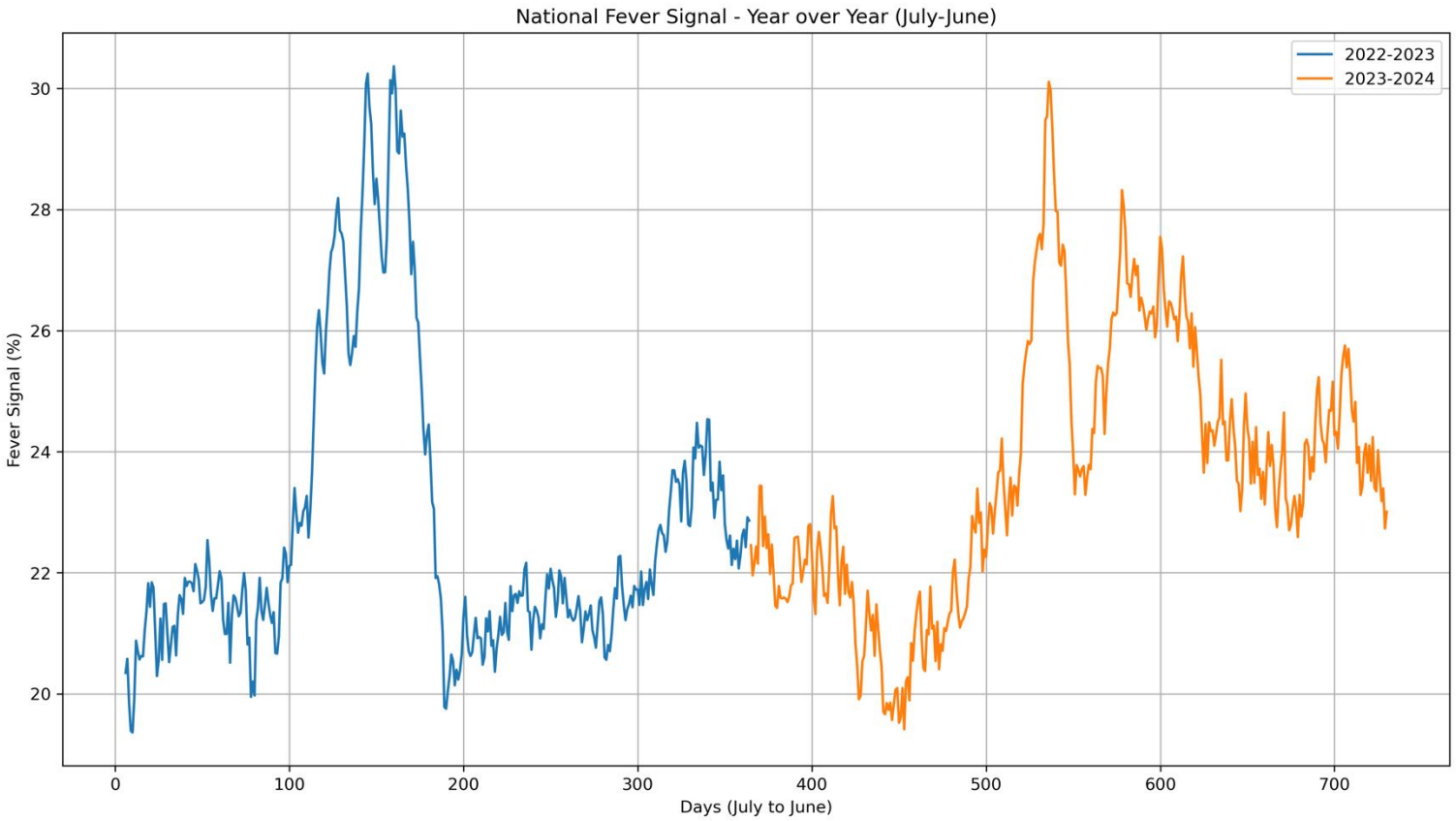




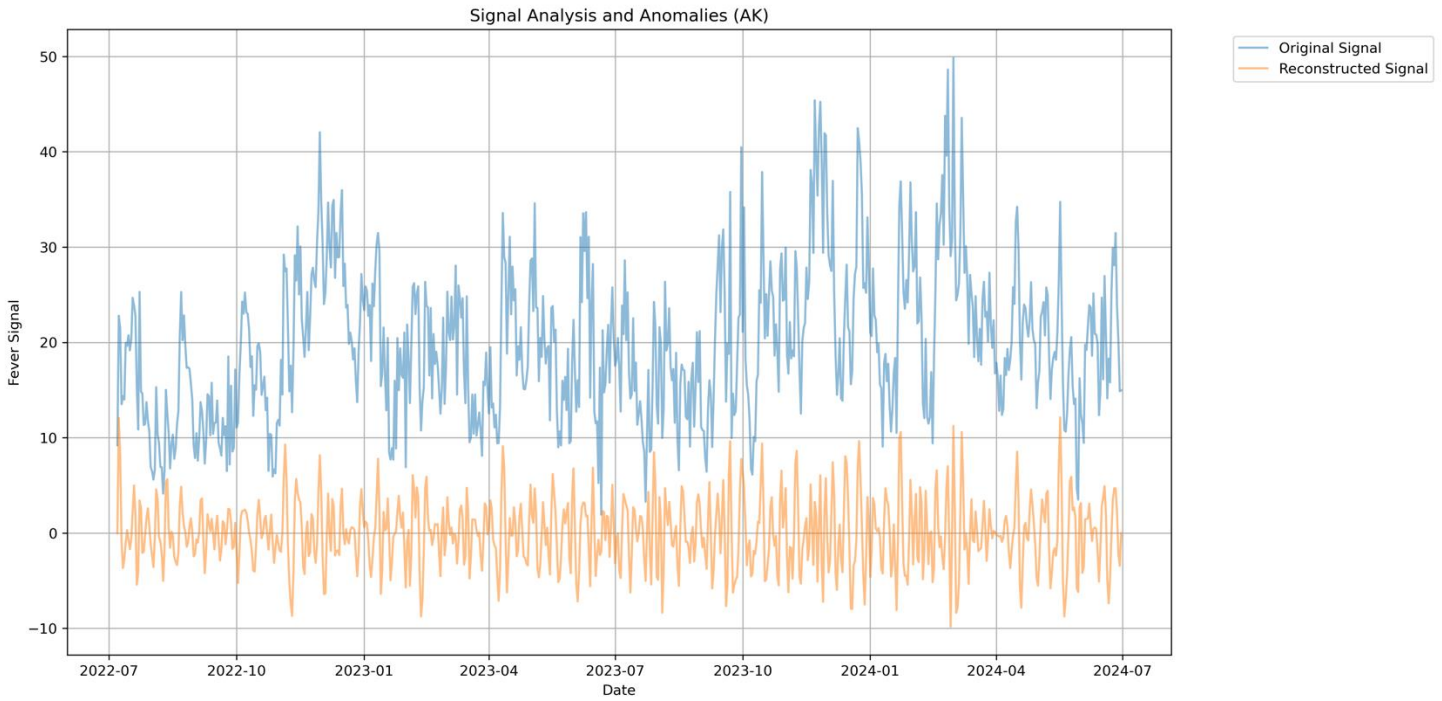
Anomaly Detection – Summary Statistics

- 21 anomalous days detected
- 26 states showing anomalies
- 1.7 states affected per anomaly day
- 2.88% of total timespan

Year-over-Year Comparison



Anomaly Detection



- 21 anomalous days
- 26 states showing anomalies
- 1.7 states per anomaly day (average)

Forecasting Approach

Key Components

- Seasonal decomposition
- Geographic factors
- Demographic considerations

Model Strategy

- Feature selection
- Temporal aspects
- Spatial correlations

Implementation Plan

DATA PIPELINE REQUIREMENTS

VALIDATION APPROACH

SUCCESS METRICS

Limitations and Future Work

- Current Limitations:
  - Rural state data sparsity
  - Smartphone dependency
  - Demographic representation
- Future Development:
  - Enhanced modeling incorporating:
    - Improved spatial correlation analysis
    - Better seasonality modeling
    - More granular demographic segmentation

Questions?

Contact: Joel Skaria, joel.skaria.js@gmail.com