Final Project

Temperature Volatility Aggregation Sandbox

Presentation by: Cole Davis

Brown University

April 22, 2024

Intro

- How should we aggregate granular spatial time series into larger regional statistics?
  - Specifically, how should we aggregate temperature volatility?

Intro

- How should we aggregate granular spatial time series into larger regional statistics?
  - Specifically, how should we aggregate temperature volatility?
- Applied Example: aggregating temperature data reported by weather stations into a regional measure of temperature volatility

Intro

- How should we aggregate granular spatial time series into larger regional statistics?
  - Specifically, how should we aggregate temperature volatility?
- Applied Example: aggregating temperature data reported by weather stations into a regional measure of temperature volatility
  - Should we compute temperature volatility at a weather station level and aggregate that?

Intro

- How should we aggregate granular spatial time series into larger regional statistics?
  - Specifically, how should we aggregate temperature volatility?
- Applied Example: aggregating temperature data reported by weather stations into a regional measure of temperature volatility
  - Should we compute temperature volatility at a weather station level and aggregate that?
  - How do we account for correlation between weather stations?

Intro

- How should we aggregate granular spatial time series into larger regional statistics?
  - Specifically, how should we aggregate temperature volatility?
- Applied Example: aggregating temperature data reported by weather stations into a regional measure of temperature volatility
  - Should we compute temperature volatility at a weather station level and aggregate that?
  - How do we account for correlation between weather stations?
  - At what stage should we *demean* the data?
  - Can we introduce a custom weighting scheme in the aggregation process?
  - How do we measure temperature volatility?

Simulation Sandbox

- To effectively test these questions, we need to run numerical simulations that effectively do the following:

Simulation Sandbox

- To effectively test these questions, we need to run numerical simulations that effectively do the following:
  1. Create a general region and split it into states

Simulation Sandbox

- To effectively test these questions, we need to run numerical simulations that effectively do the following:
  1. Create a general region and split it into states
  2. Create weather stations in each state

Simulation Sandbox

- To effectively test these questions, we need to run numerical simulations that effectively do the following:
  1. Create a general region and split it into states
  2. Create weather stations in each state
  3. Generate temperature data over a given period of time

Region and States

- Region size is based on the contiguous region formed by Iowa, Illinois, Nebraska, and Minnesota
  - Can be modified

Region and States

- Region size is based on the contiguous region formed by Iowa, Illinois, Nebraska, and Minnesota
  - Can be modified
- Region is rectangular
  - Cannot be (easily) modified

Region and States

- Region size is based on the contiguous region formed by Iowa, Illinois, Nebraska, and Minnesota
  - Can be modified
- Region is rectangular
  - Cannot be (easily) modified
- Region is randomly split into four rectangular states (A, B, C, and D)
  - Number of states cannot be (easily) modified
  - Splitting process could be modified as long as states remain rectangular

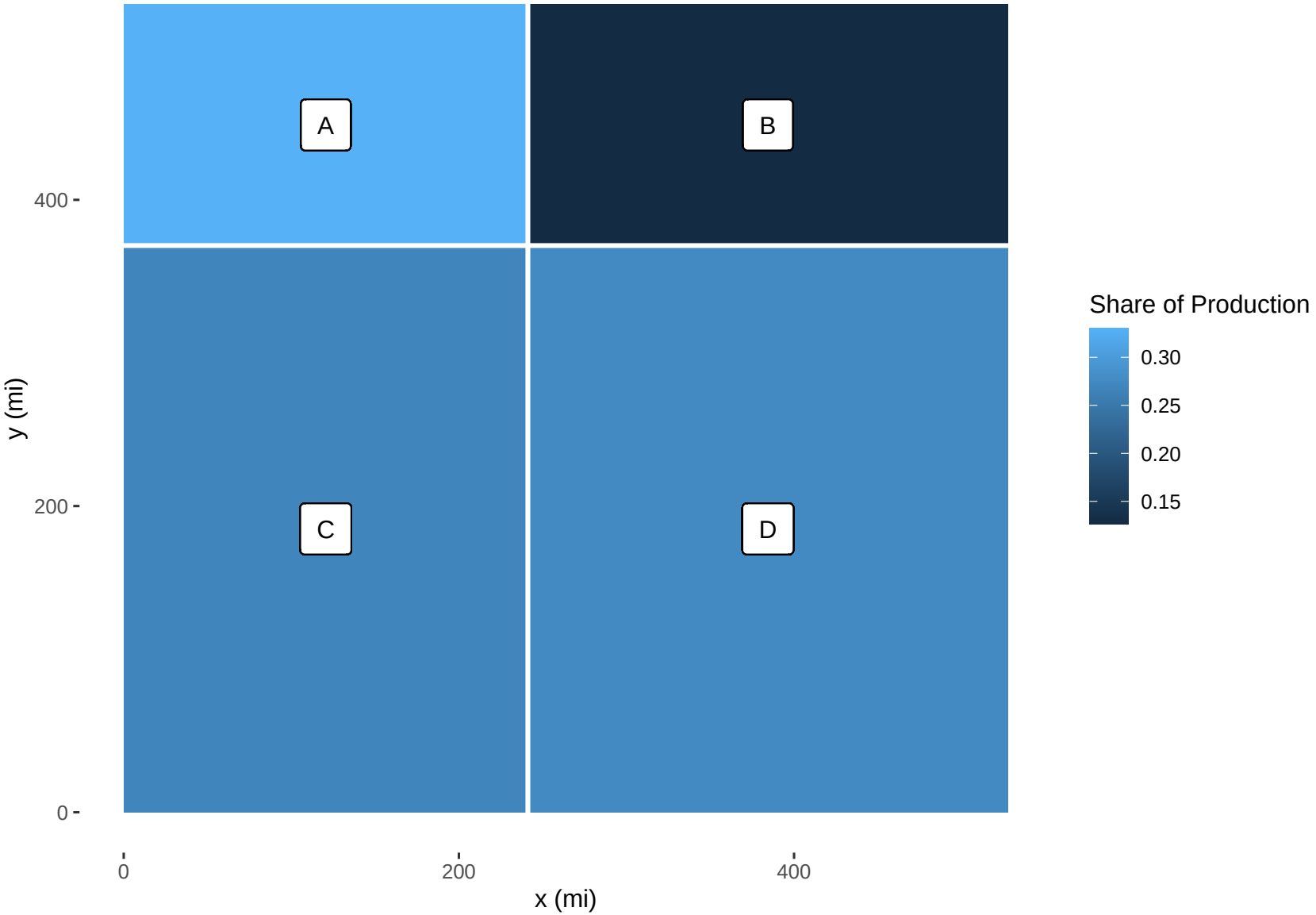
Region and States

- Region size is based on the contiguous region formed by Iowa, Illinois, Nebraska, and Minnesota
  - Can be modified
- Region is rectangular
  - Cannot be (easily) modified
- Region is randomly split into four rectangular states (A, B, C, and D)
  - Number of states cannot be (easily) modified
  - Splitting process could be modified as long as states remain rectangular
- States are randomly assigned a share (to be used in intermediate aggregation)

Region and States

- Region size is based on the contiguous region formed by Iowa, Illinois, Nebraska, and Minnesota
  - Can be modified
- Region is rectangular
  - Cannot be (easily) modified
- Region is randomly split into four rectangular states (A, B, C, and D)
  - Number of states cannot be (easily) modified
  - Splitting process could be modified as long as states remain rectangular
- States are randomly assigned a share (to be used in intermediate aggregation)
- (x, y) coordinates of state corners and shares are exported to a csv

Region Map


A
B
C
D
400
200
0
y (mi)
0
200
400
x (mi)
Share of Production
0.30
0.25
0.20
0.15

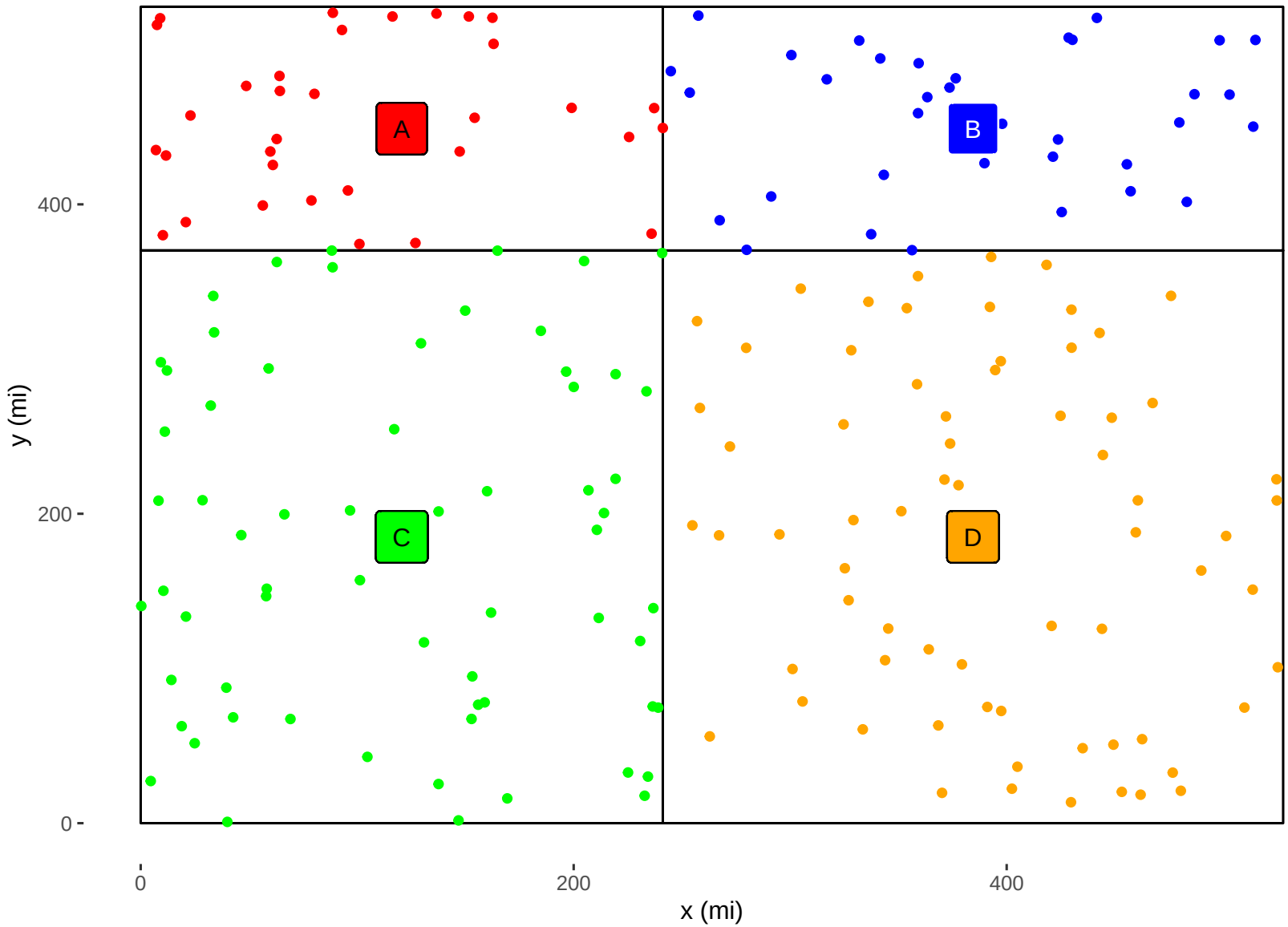
Weather Stations

- 200 weather stations are placed randomly throughout the region
  - Can be modified

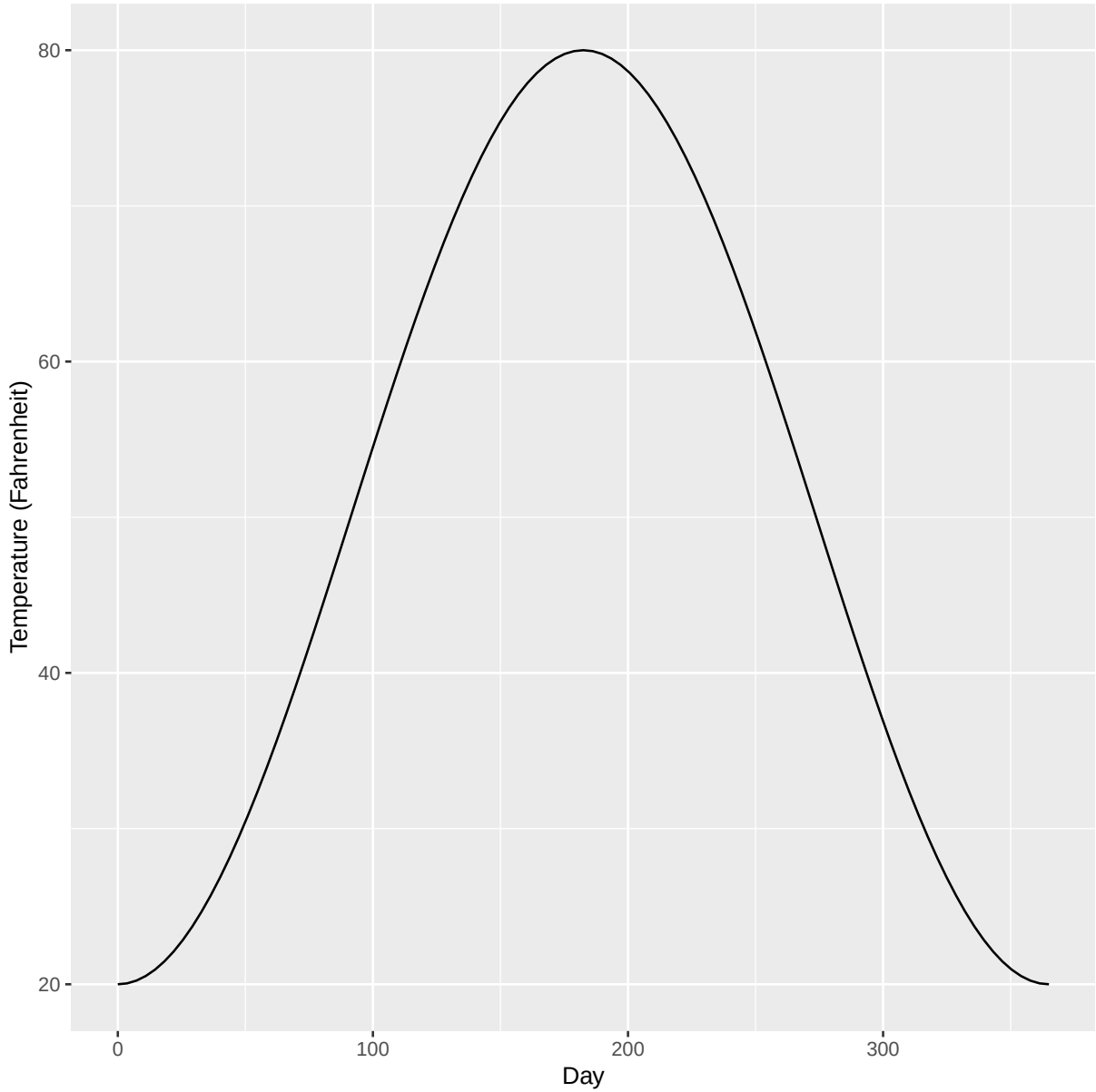
Weather Stations

- 200 weather stations are placed randomly throughout the region
  - Can be modified
- Weather stations are then sorted into their appropriate state
- (x, y) coordinates of weather stations in a state are exported to a csv

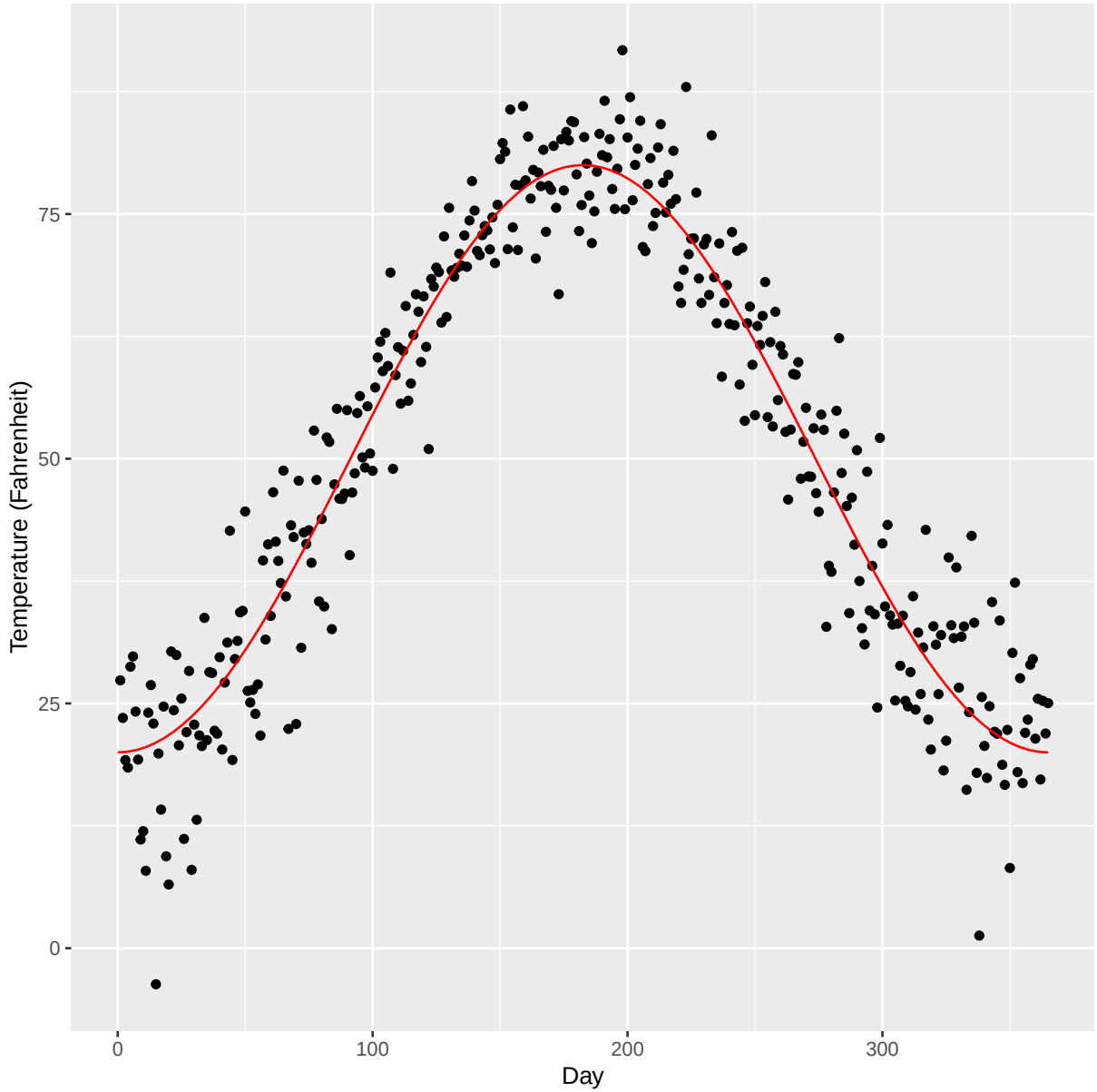
Weather Station Map



Temperature Data - Base



Temperature Data - Noisy



Temperature Data - Generation

- 4 Random points are chosen to be centroids of a cluster

Temperature Data - Generation

- 4 Random points are chosen to be centroids of a cluster
- Centroids are given a year of randomly generated temperature data

Temperature Data - Generation

- 4 Random points are chosen to be centroids of a cluster
- Centroids are given a year of randomly generated temperature data
- At time t, a non centroid weather station x has temperature distributed as

$$\mathcal{N}\left(x^*(t), 2.5d\left(\frac{x}{\|x\|}, \frac{x^*}{\|x^*\|}\right)\right)$$

in the winter and

$$\mathcal{N}\left(x^*(t), 2d\left(\frac{x}{\|x\|}, \frac{x^*}{\|x^*\|}\right)\right)$$

in the summer

Thank you!