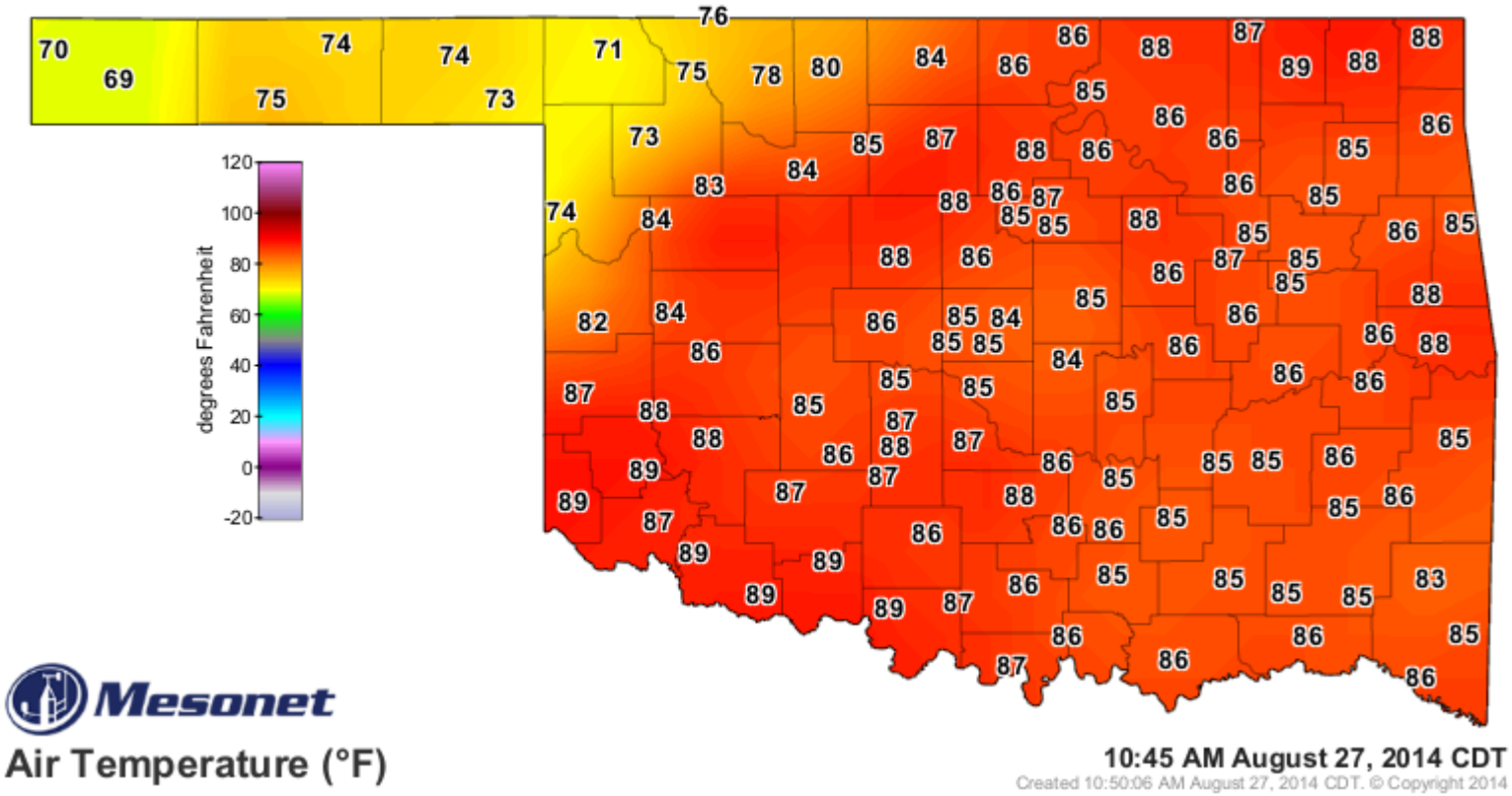
# CS 2334: Project 1
# Reading Data from Files

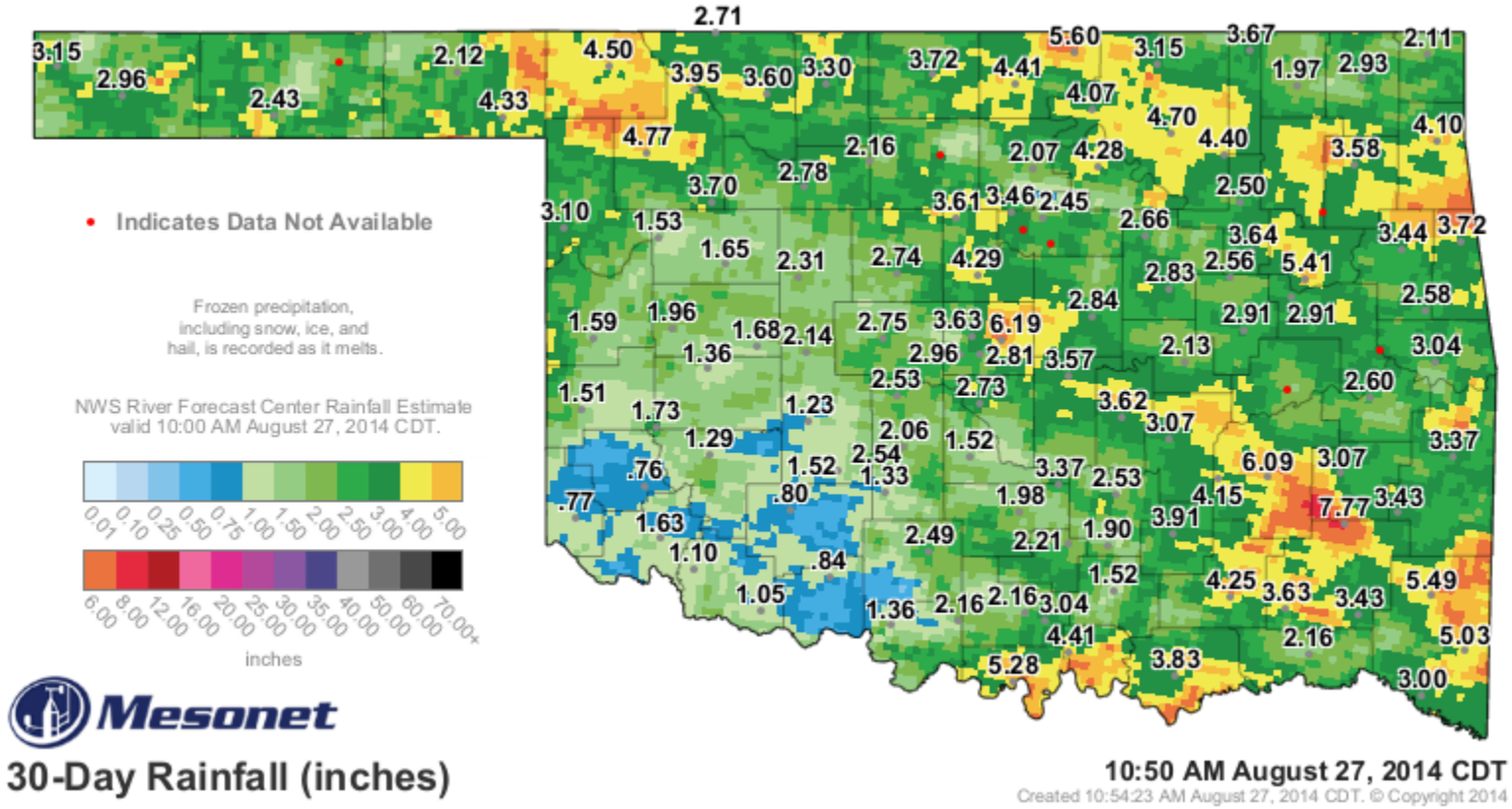
# Oklahoma Mesonet

- Stations in every county of Oklahoma
- Collect data every 5 minutes
  - Air and soil temperature
  - Rainfall
  - Wind speed and direction
  - …



# Project 1

- Read data from files for a set of months
  - Including daily temperature and rainfall
- Compute statistics across days in a month
- Compute statistics across the months



# Data Format

- One file per month
- Each row: data for a single day
- Represented as a CSV file
- Some values are invalid

| YEAR | MONTH | DAY | STID | TMAX | TMIN | TAVG | RAIN |
|---|---|---|---|---|---|---|---|
| 2002 | 8 | 1 | NRMN | 95.59 | 76.06 | 84.94 | 0 |
| 2002 | 8 | 2 | NRMN | 98.44 | 75.45 | 85.82 | 0 |
| 2002 | 8 | 3 | NRMN | 97.54 | 73.45 | 86.13 | 0 |
| 2002 | 8 | 4 | NRMN | 96.31 | 76.35 | 86.3 | 0 |
| 2002 | 8 | 5 | NRMN | 95.38 | 75.31 | 84.74 | 0 |
| 2002 | 8 | 6 | NRMN | 95.5 | 71.53 | 84.46 | 0 |
| 2002 | 8 | 7 | NRMN | 95.52 | 72.32 | 83.79 | 0 |
| 2002 | 8 | 8 | NRMN | 93.96 | 75.13 | 82.77 | 0 |
| 2002 | 8 | 9 | NRMN | 95.4 | 68.58 | 81.8 | 0 |
| 2002 | 8 | 10 | NRMN | 96.42 | 69.08 | 81.83 | 0 |
| 2002 | 8 | 11 | NRMN | 94.19 | 72 | 81.84 | 0 |
| 2002 | 8 | 12 | NRMN | 94.39 | 69.31 | 81.99 | 0.01 |
| 2002 | 8 | 13 | NRMN | 81.88 | 65.64 | 72.67 | 0.29 |
| 2002 | 8 | 14 | NRMN | 82.78 | 64.36 | 71.9 | 0.03 |
| 2002 | 8 | 15 | NRMN | 90.21 | 70.34 | 79.38 | 0 |
| 2002 | 8 | 16 | NRMN | 92.8 | 70.3 | 81.06 | 0 |
| 2002 | 8 | 17 | NRMN | 95.22 | 75.43 | 84.4 | 0 |
| 2002 | 8 | 18 | NRMN | -996 | -996 | -996 | -999 |
| 2002 | 8 | 19 | NRMN | -996 | -996 | -996 | -999 |
| 2002 | 8 | 20 | NRMN | 94.59 | 74.12 | 84.52 | 0 |

# Observation Class

Two instance variables:

- value (double): value to be represented (could be a temperature or a rainfall measurement)
- valid (boolean): indicates whether the value is valid or not

Class is immutable

# DailyData Class

| YEAR | MONTH | DAY | STID | TMAX | TMIN | TAVG | RAIN |
|---|---|---|---|---|---|---|---|
| 2002 | 8 | 1 | NRMN | 95.59 | 76.06 | 84.94 | 0 |

Captures all information for a single day (one row in the table):

- year, month, day (int)
- stationID (String)
- temperatureMax, temperatureMin, temperatureAverage and rainFall (Observations)

Class is immutable

# MonthlyData Class

Captures data for an entire month (from a single file):

- days (ArrayList<DailyData>): one entry for each day
- rainMax, rainMin, and rainAverage (doubles): computed over all days with valid Observations
- temperatureMax (double): max over all days of temperatureMax (valid samples only)
- temperatureMin (double): min over all days of temperatureMin
- temperatureAverage (double): average of temperatureAverage
- year and month (ints): copy of year and month from first day

# DataSet Class

Represents data from a set of months:

- months (ArrayList<MonthlyData>): one entry per month
- rainMax, rainMin, and rainAverage (doubles): computed over all months
- rainMaxMonth (MonthlyData): reference to the month with the maximum rainfall
- rainMinMonth (MonthlyData): reference to the month with the minimum rainfall

# DataSet Class (cont)

- temperatureMax (double): max over all days of temperatureMax (valid samples only)
- temperatureMaxMonth (MonthlyData): reference to the month with the maximum temperature
- temperatureMin (double): min over all days of temperatureMin
- temperatureMinMonth (MonthlyData): reference to the month with the minimum temperature
- temperatureAverage (double): average of temperatureAverage

# Provided Code

- We provide a starting point for project 1
- Keep all of the declared variables and methods
- Fill in missing implementation

- Example: constructor for MonthlyData

```
public MonthlyData(String fileName) throws IOException,
NumberFormatException, FileNotFoundException{…}
```

# Testing

Implement Junit tests for:

- Observation
- DailyData
- MonthlyData
- DataSet

# Strategies for Success

- Work with project partner in person
- Start early
- Implement and test incrementally
- Write documentation as you go

# Submission

- Submit only one file per group: project1.zip (casing matters)
- Due date: Wednesday, September 23rd @1:29pm (before class!)
- Submit to project1 dropbox on D2L