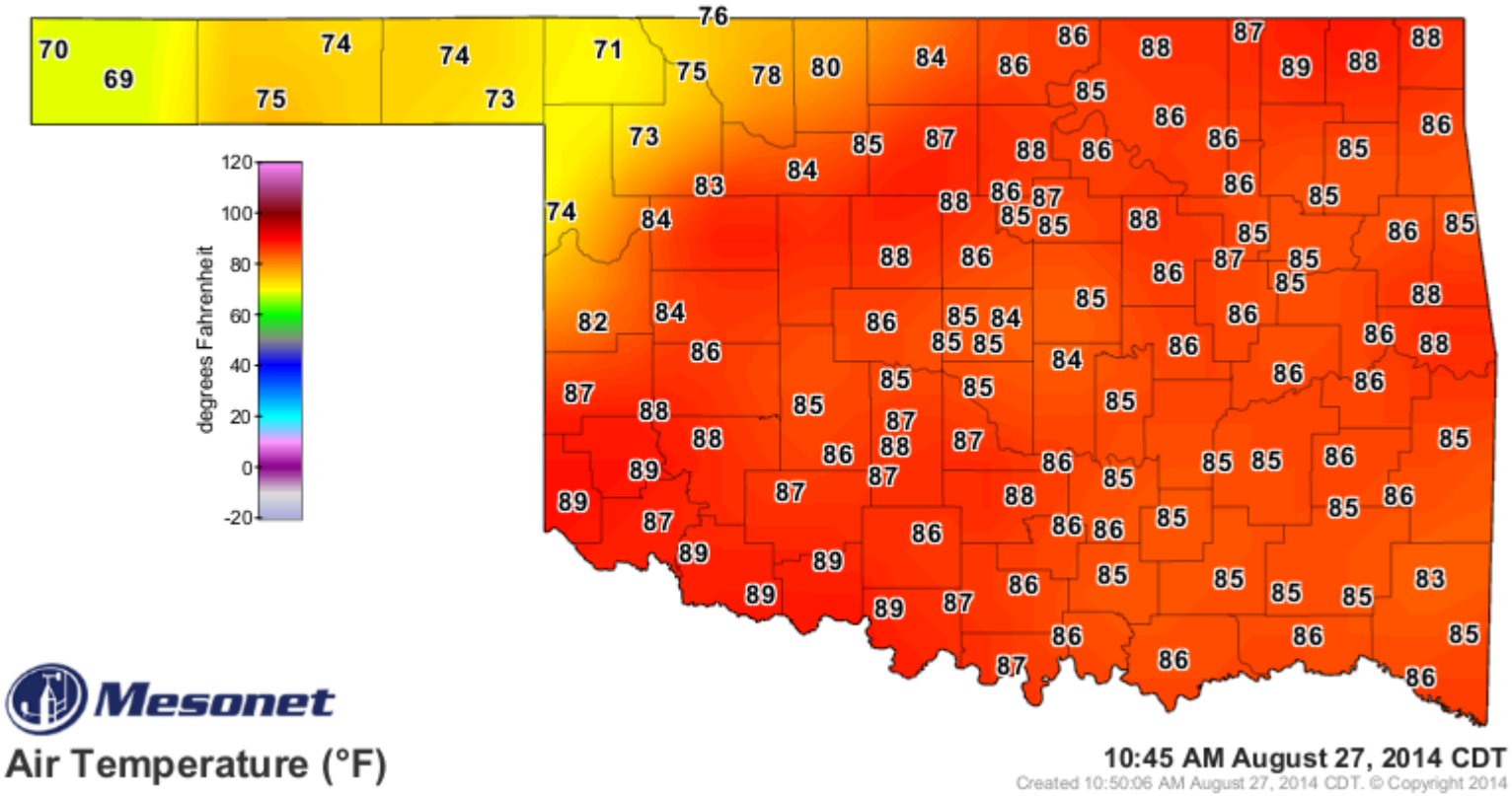
CS 2334: Project 1
Reading Data from Files

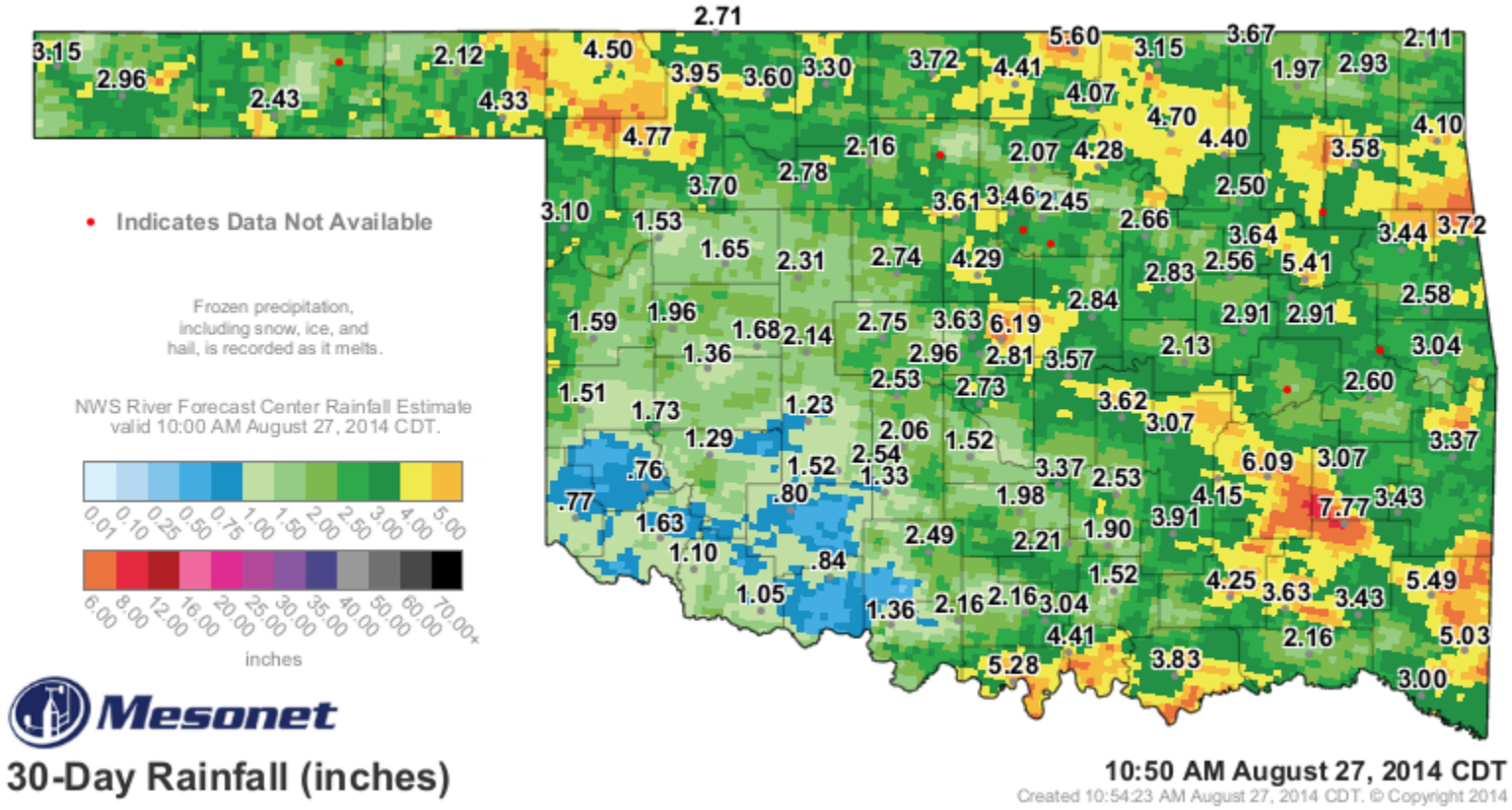
Oklahoma Mesonet

- 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 wind speed and solar radiation
- Compute statistics across days in a month



Data Format

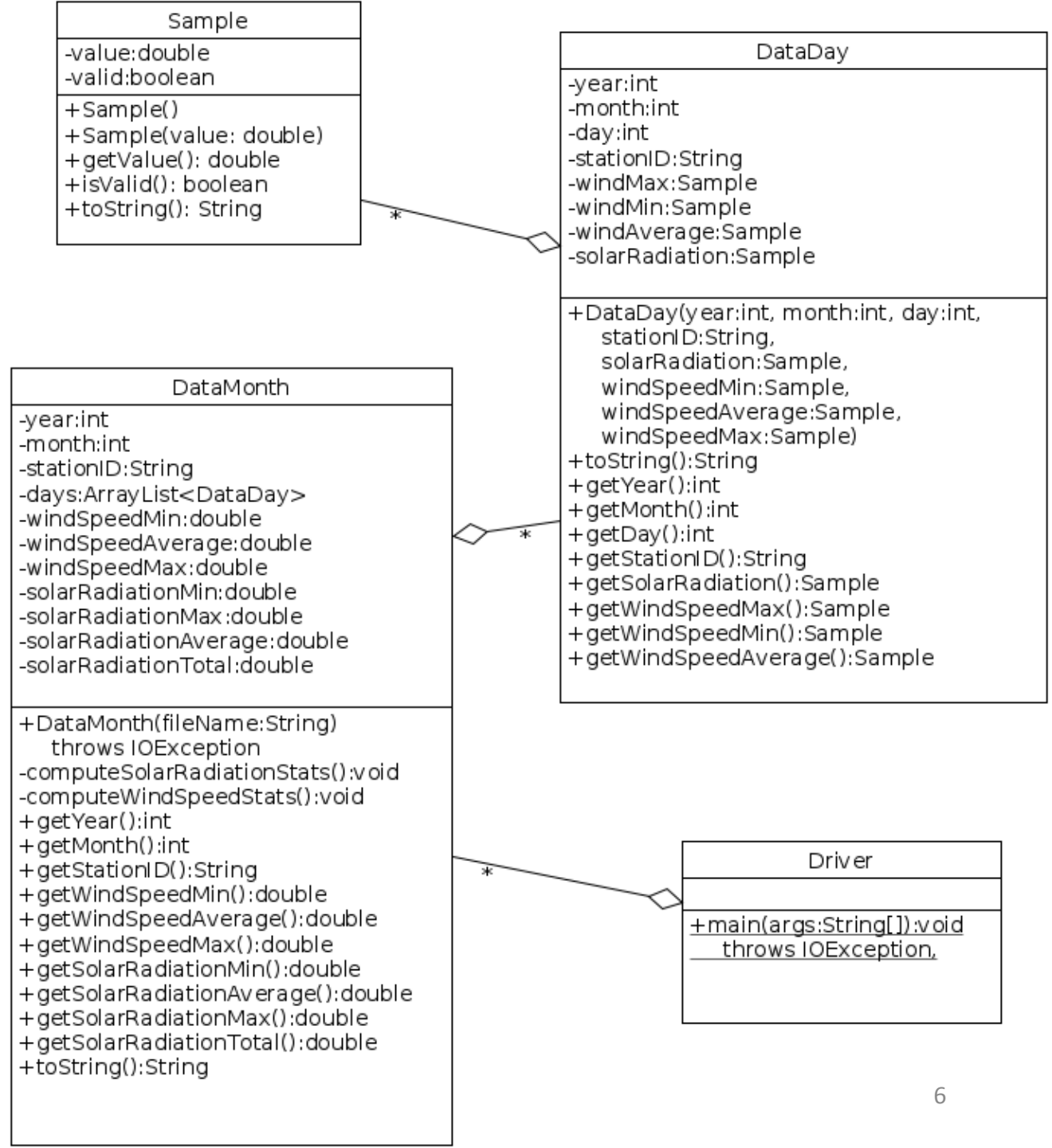
- One file per month
- Each row: data for a single day
- Represented as a CSV file
- Some values are invalid
- All CSV files have the same format

YEAR	MONTH	DAY	STID	ATOT	WSMX	WSMN	WSPD
2014	6	1	TISH	22.7	16.67	1.11	9
2014	6	2	TISH	-996	19.75	2.1	8.83
2014	6	3	TISH	27.67	16.7	2.41	9.69
2014	6	4	TISH	28.52	22.17	2.2	11.88
2014	6	5	TISH	25.12	21.07	2.29	10.07
2014	6	6	TISH	13.2	22.88	0	7.41
2014	6	7	TISH	24.61	16.86	0	8.31
2014	6	8	TISH	17.8	20.42	1.21	7.61
2014	6	9	TISH	19.62	18.45	3.12	8.83
2014	6	10	TISH	25.66	16.07	2.33	8.35
2014	6	11	TISH	29.92	11.92	1.27	5.99
2014	6	12	TISH	15.23	28.5	1.69	8.12
2014	6	13	TISH	30.52	13.2	1.36	6.35
2014	6	14	TISH	22.02	20.92	3.89	11.93
2014	6	15	TISH	23.85	20.31	6.38	12.79
2014	6	16	TISH	25.76	21.68	5.4	13.93
2014	6	17	TISH	23.72	21.25	8.03	14.52

Key Columns

- ATOT: total solar radiation
- WSMX: wind speed max
- WSMN: wind speed min
- WSPD: wind speed average

Solution Design



Sample Class

Two instance variables:

- value (double): value to be represented (could be a wind speed or solar radiation measurement)
- valid (boolean): indicates whether the value is valid or not

Class is immutable

DataDay Class

YEAR	MONTH	DAY	STID	ATOT	WSMX	WSMN	WSPD
2014	6	1	TISH	22.7	16.67	1.11	9

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

- year, month, day (int)
- stationID (String)
- windSpeedMax, windSpeedMin, windSpeedAverage and solarRadiation (all are Samples)

Class is immutable

DataMonth Class

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

- days (ArrayList<DataDay>): one entry for each day
- solarRadiationMax, solarRadiationMin, and solarRadiationAverage (doubles): computed over all days with valid Samples
- windSpeedMax (double): max over all days of windSpeedMax (valid Samples only)
- windSpeedMin (double): min over all days of windSpeedMin
- windSpeedAverage (double): average of windSpeedAverage
- year and month (ints): copy of year and month from first day
- stationID (String): station described by this month's data

Provided Materials

- Several CSV files
- Project specification
  - Details of the requirements, including the expected toString() values for the Sample, DataDay and DataMonth classes
  - Don’t deviate from the specification – we won’t be able to compile our tests against your code

Testing

Implement your own Junit tests for:

- Sample
- DataDay
- DataMonth

- Be thorough
- Derive from the provided CSV files or create your own

Strategies for Success

- Work with project partner in person
- Start early
- Implement and test incrementally
- Don’t deviate from our design
- Write documentation as you go

Submission

- Due date: Wednesday, September 21st @1:29pm (before class!)
- Submit to the Web-Cat server

Grading

- Grading criteria will be similar to what we are using in the labs
- Except …

Code Review

Each group must come in for a code review with me or one of the TAs

- 15-minutes
- Discuss documentation, implementation and performance on tests
- Both group members must be able to answer questions about all aspects of the code
- Reserve a time or walk-in
- If you complete your code review before the deadline, you may resubmit to address issues
- Code reviews must be completed by Wednesday, September 28th