

Thank you for accepting the rules.



\$500 • 604 teams

## How Much Did It Rain? II

Thu 17 Sep 2015

Merger and 1st Submission Deadline

Mon 7 Dec 2015 (7.4 days to go)

### Dashboard

Home

Data

Make a submission

Information

Description

Evaluation

Rules

Prizes

Timeline

Forum

Scripts

New Script

New Notebook

Leaderboard

My Team

My Submissions

### Leaderboard

1. PuPa
2. Li-Der
3. mind.cool
4. cydonia
5. JamesGoodchild
6. S & A
7. Pedro Lima
8. Mikhail
9. Mike Kim
10. junfeng

### 399 Scripts

Beware of Outliers !!  
24 Votes / 2 months ago / Python

Rain Rate estimation from KDP  
Cut f  
2 Votes / 2 days ago / R

[Competition Details](#) » [Get the Data](#) » [Make a submission](#)

### Data Files

| File Name           | Available Formats                |
|---------------------|----------------------------------|
| train               | <a href="#">.zip (240.42 mb)</a> |
| test                | <a href="#">.zip (134.52 mb)</a> |
| sample_solution.csv | <a href="#">.zip (4.40 mb)</a>   |
| sample_dask         | <a href="#">.py (1.96 kb)</a>    |

The training data consists of NEXRAD and MADIS data collected on 20 days between Apr and Aug 2014 over midwestern corn-growing states. Time and location information have been censored, and the data have been shuffled so that they are not ordered by time or place. The test data consists of data from the same radars and gauges over the remaining days in that month. Please see this page to understand more about polarimetric radar measurements.

## File descriptions

- **train.zip** - the training set. This consists of radar observations at gauges in the Midwestern US over 20 days each month during the corn growing season. You are also provided the gauge observation at the end of each hour.
- **test.zip** - the test set. This consists of radar observations at gauges in the Midwestern US over the remaining 10/11 days each month of the same year(s) as the training set. You are required to predict the gauge observation at the end of each hour.
- **sample\_solution.zip** - a sample submission file in the correct format
- **sample\_dask.py** - Example program in Python that will produce the sample submission file. This program applies the [Marshall-Palmer](#) relationship to the radar observations to predict the gauge observation.

Marshall-Palmer in R  
10 Votes / 57 days ago / R

Exploring the how much does it  
rain data  
2 Votes / 10 days ago / Python

Pattern in Distribution of  
Outcomes  
5 Votes / 41 days ago / R

gbm\_inches\_only  
6 Votes / 47 days ago / R

## Forum (52 topics)

scoring  
2 hours ago

Last minute team up!  
4 hours ago

Rain Rate estimation from KDP  
Cut f  
2 days ago

Ignored Ids  
4 days ago

What else besides Marshall-  
Palmer?  
5 days ago

H2ORF cleaned, cut 69 pur vars 3  
n 3  
6 days ago

teams

players

entries

## Data columns

To understand the data, you have to realize that there are multiple radar observations over the course of an hour, and only one gauge observation (the 'Expected'). That is why there are multiple rows with the same 'Id'.

The columns in the datasets are:

- Id: A unique number for the set of observations over an hour at a gauge.
- minutes\_past: For each set of radar observations, the minutes past the top of the hour that the radar observations were carried out. Radar observations are snapshots at that point in time.
- radardist\_km: Distance of gauge from the radar whose observations are being reported.
- Ref: Radar reflectivity in km
- Ref\_5x5\_10th: 10th percentile of reflectivity values in 5x5 neighborhood around the gauge.
- Ref\_5x5\_50th: 50th percentile
- Ref\_5x5\_90th: 90th percentile
- RefComposite: Maximum reflectivity in the vertical column above gauge. In dBZ.
- RefComposite\_5x5\_10th
- RefComposite\_5x5\_50th
- RefComposite\_5x5\_90th
- RhoHV: Correlation coefficient (unitless)
- RhoHV\_5x5\_10th
- RhoHV\_5x5\_50th
- RhoHV\_5x5\_90th
- Zdr: Differential reflectivity in dB
- Zdr\_5x5\_10th
- Zdr\_5x5\_50th
- Zdr\_5x5\_90th
- Kdp: Specific differential phase (deg/km)
- Kdp\_5x5\_10th
- Kdp\_5x5\_50th
- Kdp\_5x5\_90th
- Expected: Actual gauge observation in mm at the end of the hour.

## Referencing this data

To reference this dataset in scientific publications, please use the following citation:

Lakshmanan, V, A. Kleeman, J. Boshard, R. Minkowsky, A. Pasch, 2015. The AMS-AI 2015-2016 Contest: Probabilistic estimate of hourly rainfall from radar. 13th Conference on Artificial Intelligence, American Meteorological Society, Phoenix, AZ

