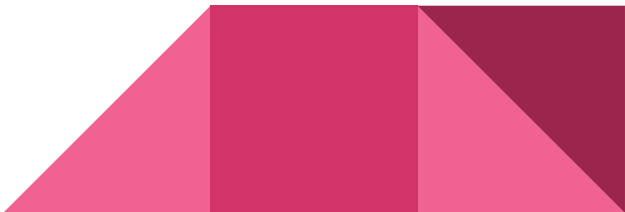


# Crop data analytics using image and non-image features

Team members : Asish Varanasi  
Dheeraj Raghavendra  
Purnachand Jaddu

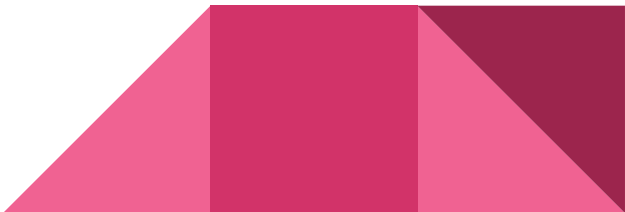
Mentors : Dr. Avinash Sharma  
Dr. P. Krishna Reddy

# Motivation

- Agriculture plays a vital role in economy for developing countries like India.
  - Rapid technological advancement - but not much focus on agricultural domain.
  - Several newly emerging problems in agriculture.
  - In 2012, the NCRB of India reported 13,754 farmer suicides.
  - Hence, a need for guidance to the farmers.
  - Esagu is one such platform.
- 

# Key pain points

## Recommendation of Diagnostics:

- Accuracy
  - Timely advice
  - Minimum required input
  - Ease of communication
  - Cost effectiveness
  - Feedback system
- 

# Insights mining

Periodic behaviour of diseases with respect to :

- Season
- Crop
- Pesticides/Fertilizers
- Location
- Pests



# Visualization

- Heatmaps showing the variation in the intensity of factors w.r.t:
  - Time
  - Location
- Plots like Bar charts, Pie diagrams etc.
- A dashboard for interactive visualization of the data.



# Data

- Input data primarily contains non-image features.
- Weather details, soil conditions and farm report constitute the major part.
- Pre-processing involving data cleaning, integration and normalization.
- Finally, analyzation after data transformation.



# Representation

PGM (Probabilistic Graphical Model)

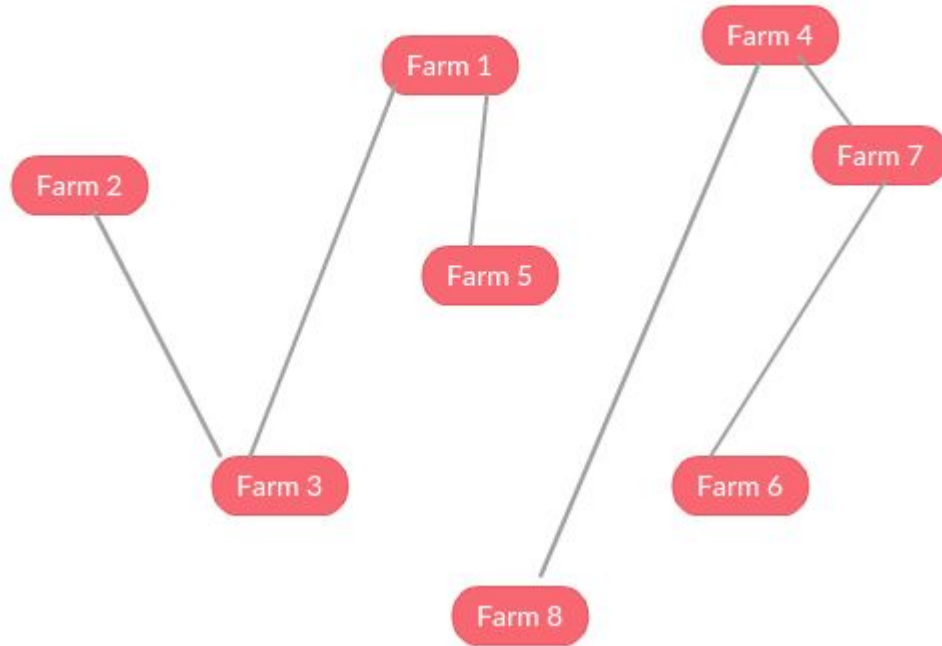
- Farms as vertices and edges indicating the dependency on various factors.

Graph signalling

- A plot of the condition of each farm with respect to time/season.

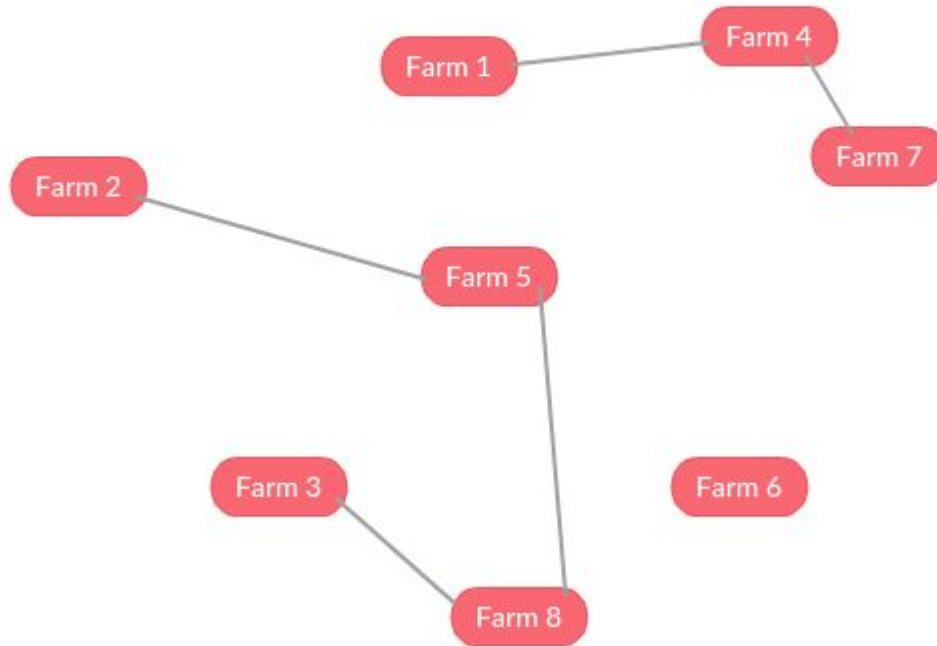


## Graph based on neighbourhood

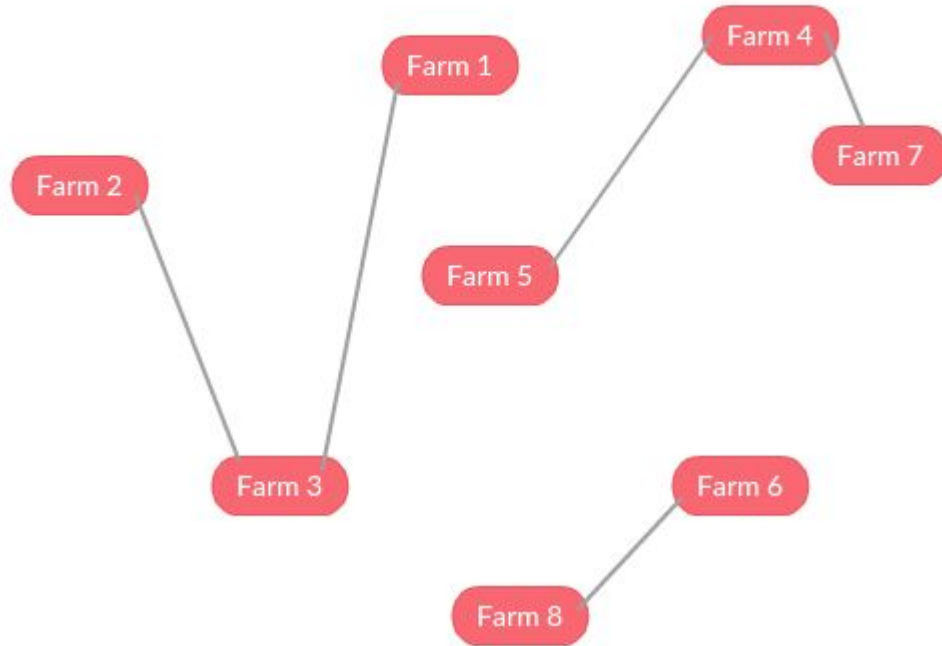




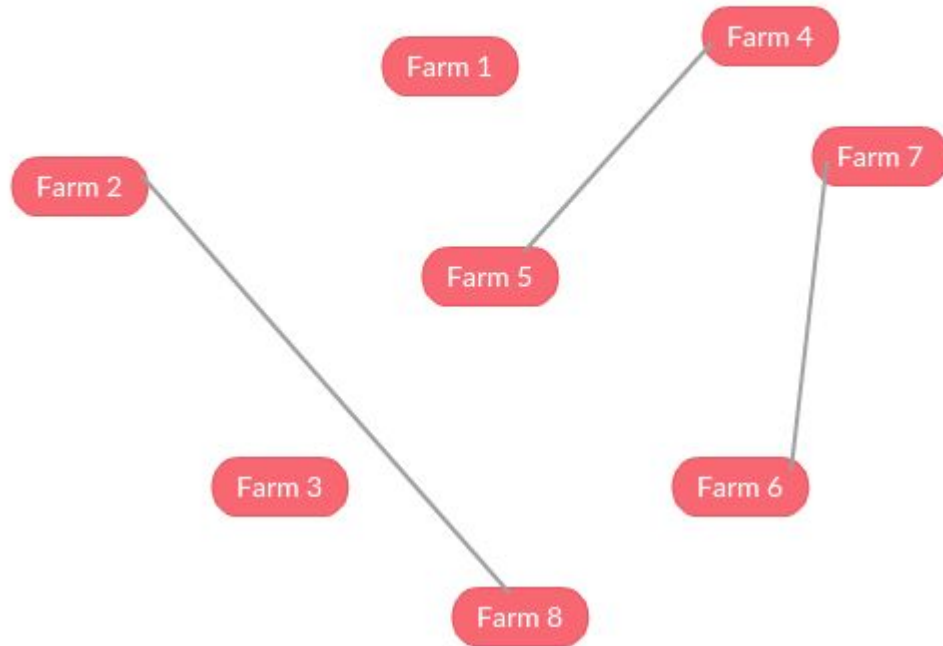
## Graph based on temperature

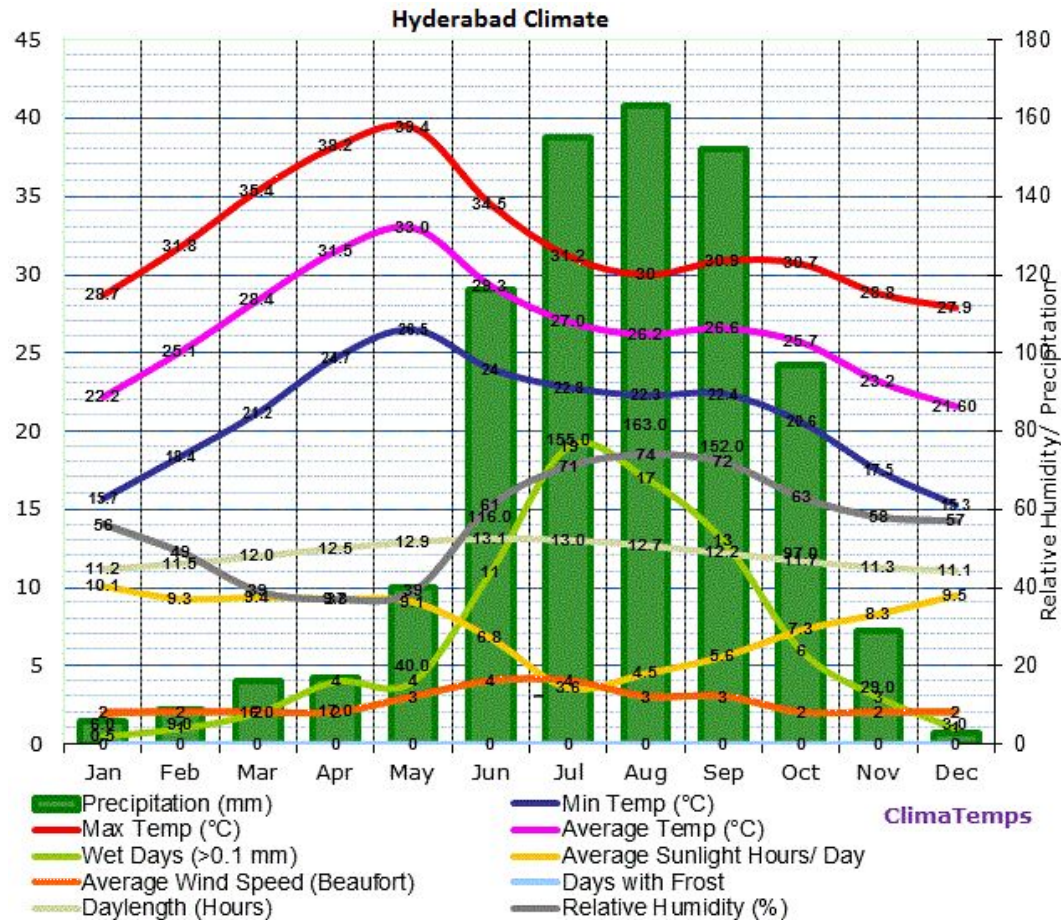


## Graph based on rainfall



## Graph based on crop species





# Models

## Graphical method

- Taking into account all the dependencies in the graphs, we model the system.

## Spectral analysis


- Considering timely behaviour of the respective feature.

## Topic modeling

- Discovering the abstract symptoms using keywords.

Predicting the solution using Multi View Learning techniques.



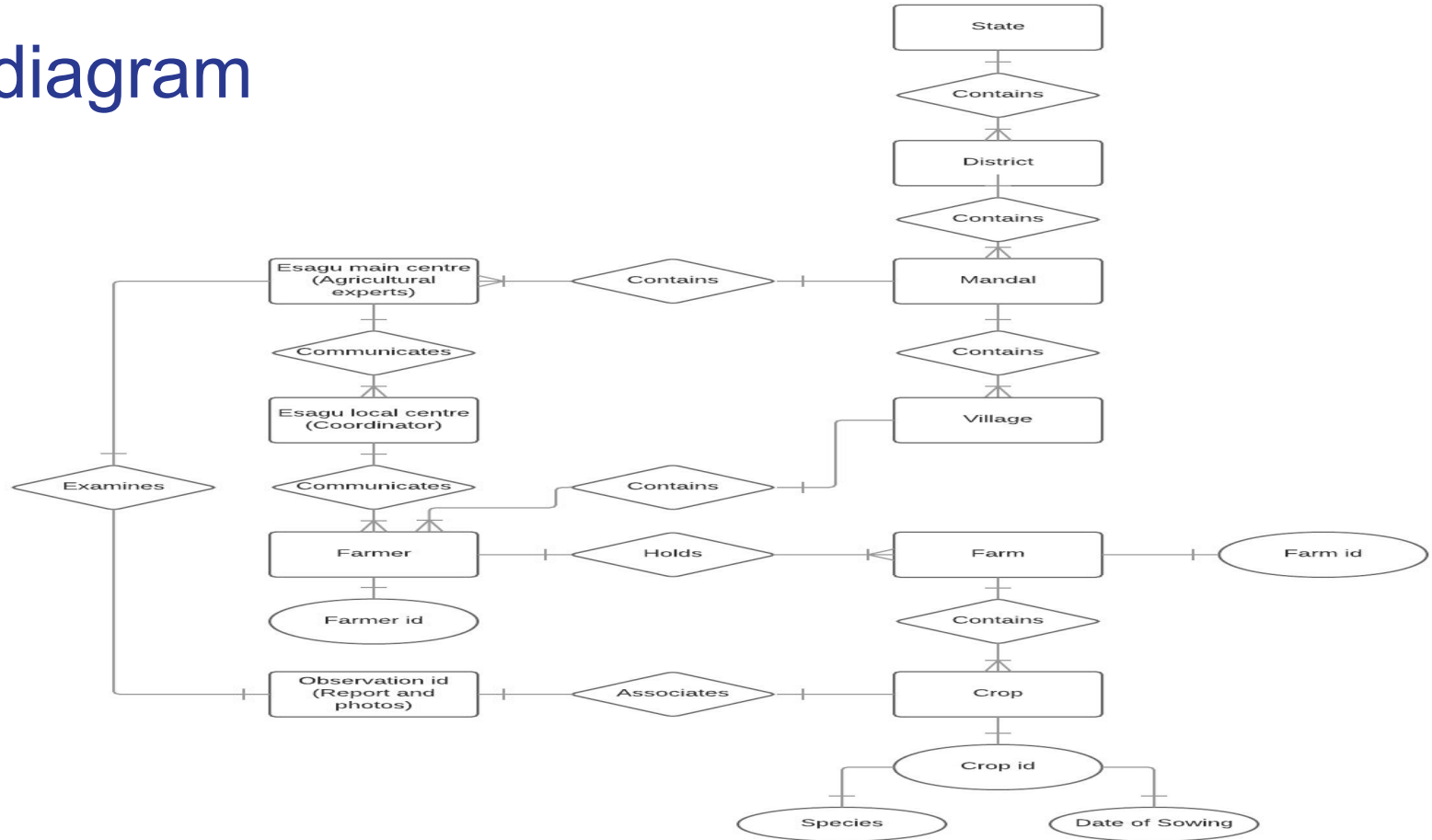
- Specific MVL solution will depend on the data.
  - If the dimension is large, then we have to use Canonical Correlation Analysis.
  - If the entire data is labeled, we can use multiview supervised learning.
  - If we have both labeled and unlabeled data, then we have to use multiview semi supervised learning.
  - Several such procedures like active learning, ensemble learning etc.
  - Finally, a feedback to improve the accuracy of the prediction algorithm.
- 

# Progress

- First, we were given the data regarding a single cotton crop for one season.
- Heatmap visualization was performed on that data.
- Later, text analysis was performed on 5 year data regarding several crops across several villages.
- Visualization was done using several plots like histograms and pie charts.

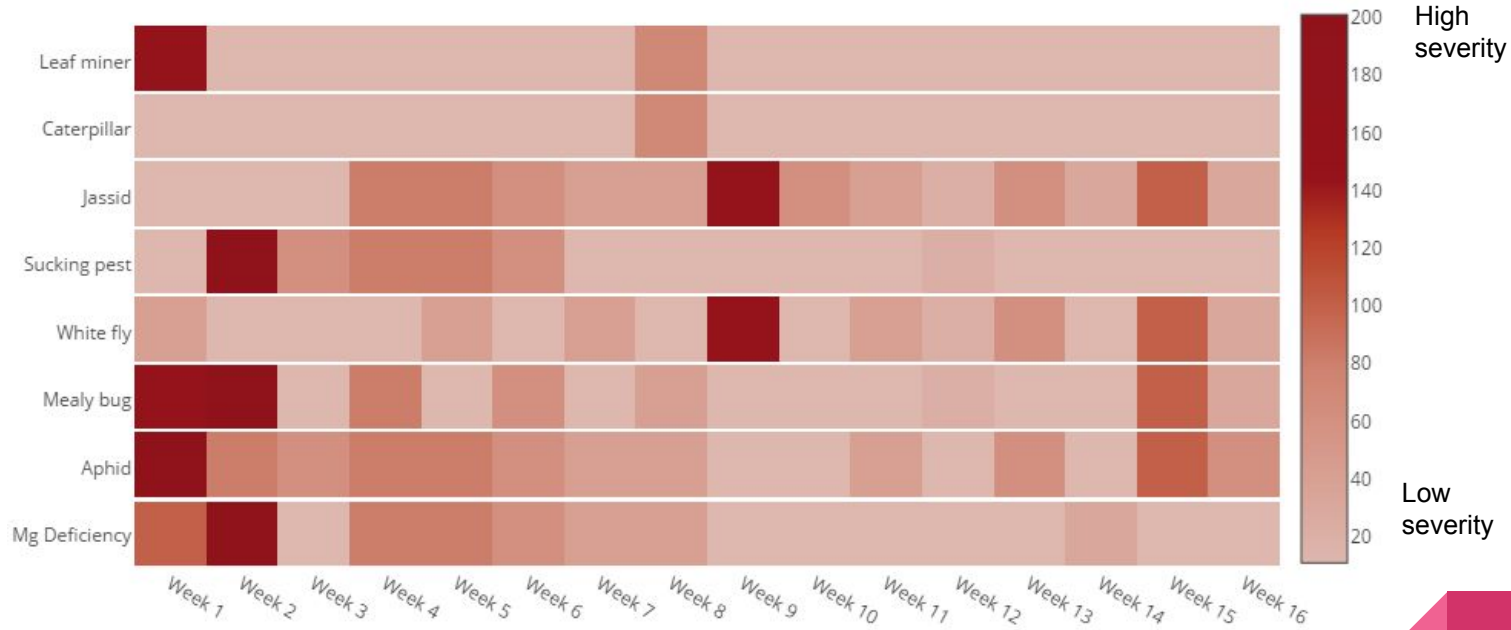


# ER diagram





# Dynamics of a single cotton crop (25-7-16 to 7-11-16)

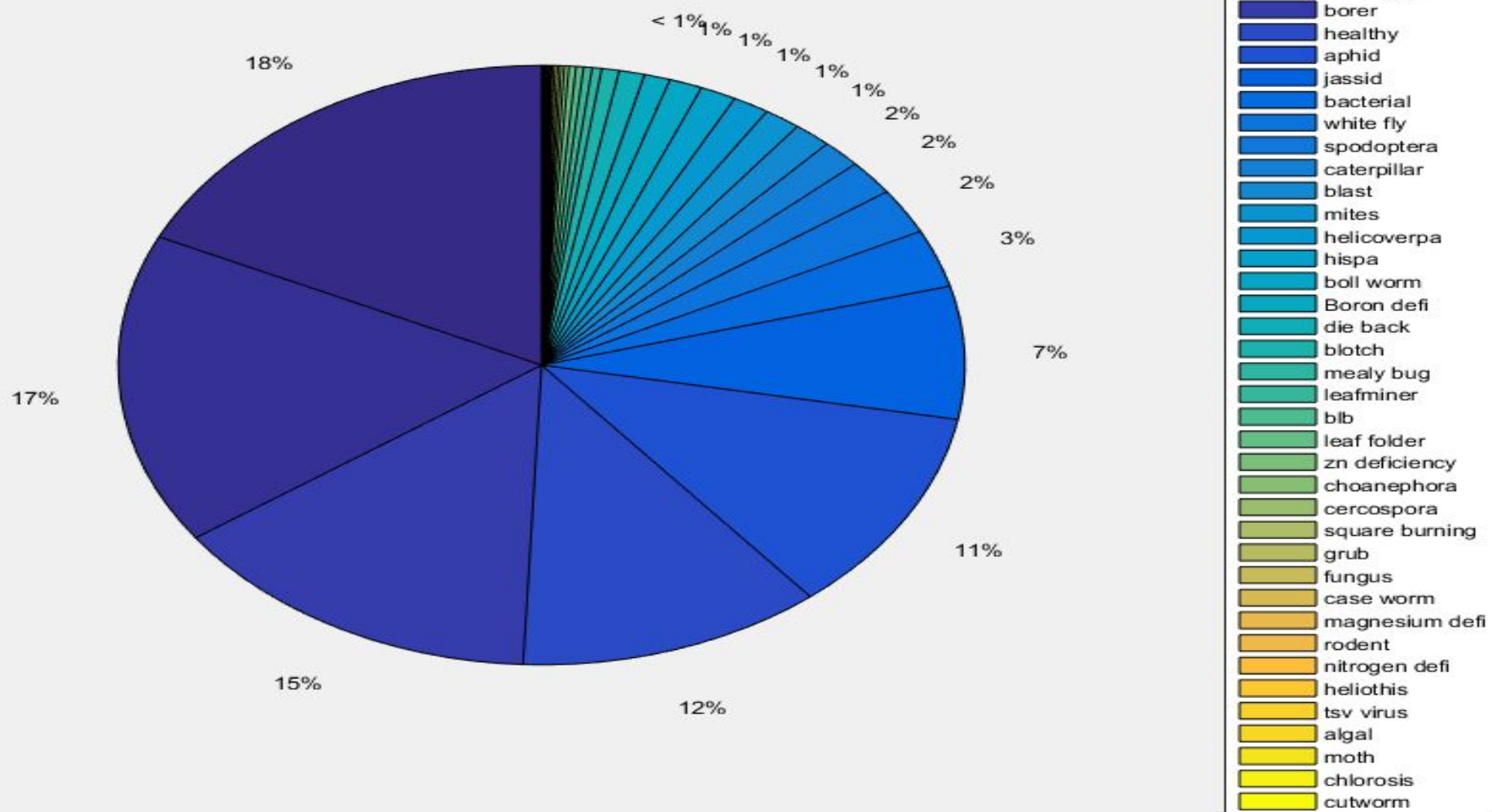


Visualization of provided data

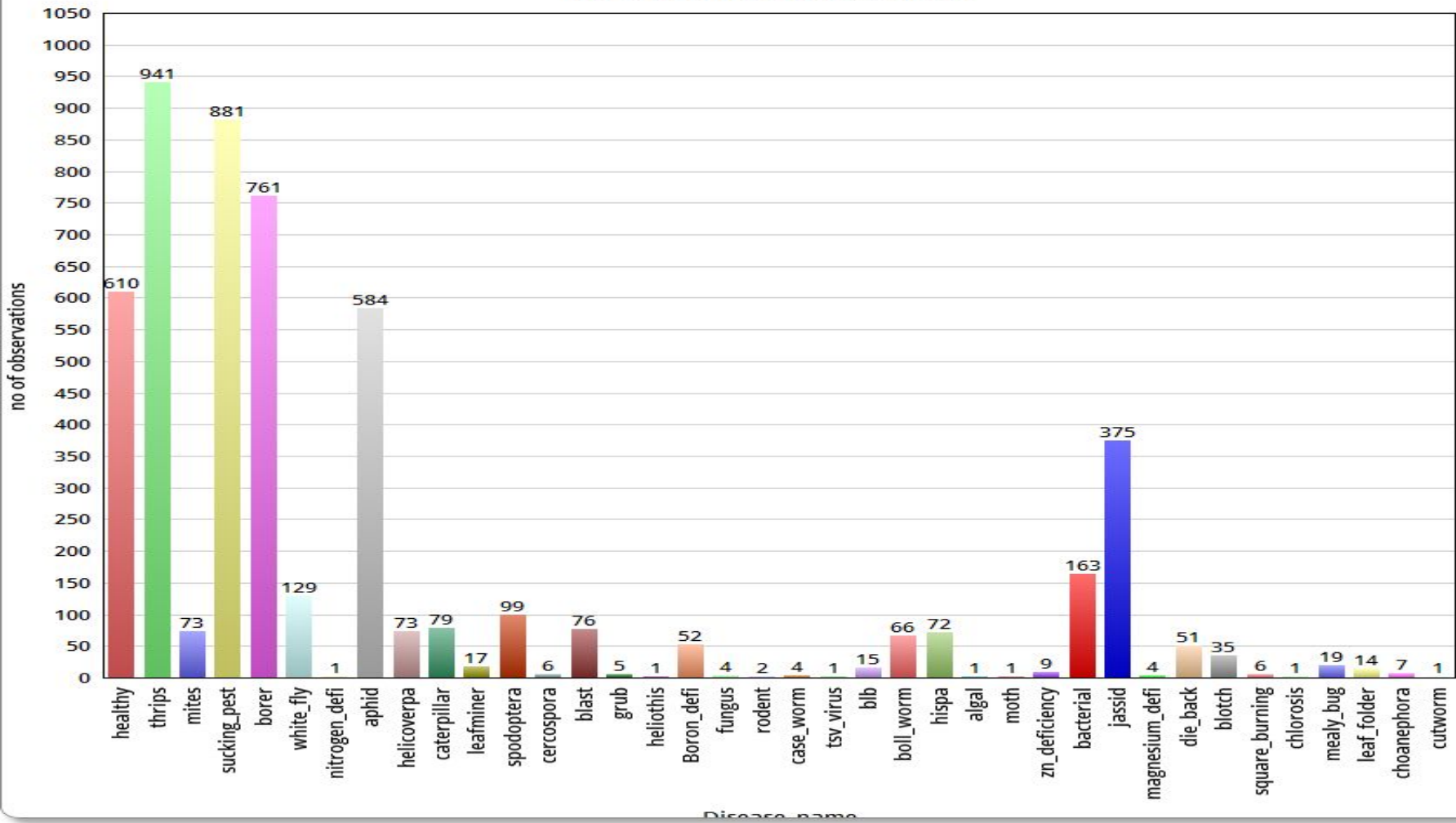
# Text analysis

- Performed on data corresponding to several farms for 5 years (2012 to 2016).
- Total number of advices = 6764.
- The occurrences of different problems were counted by identifying corresponding keywords.
- Advices containing more than one problem would increment the count of respective problems





**Disease\_count**  
No of obs having a particular disease

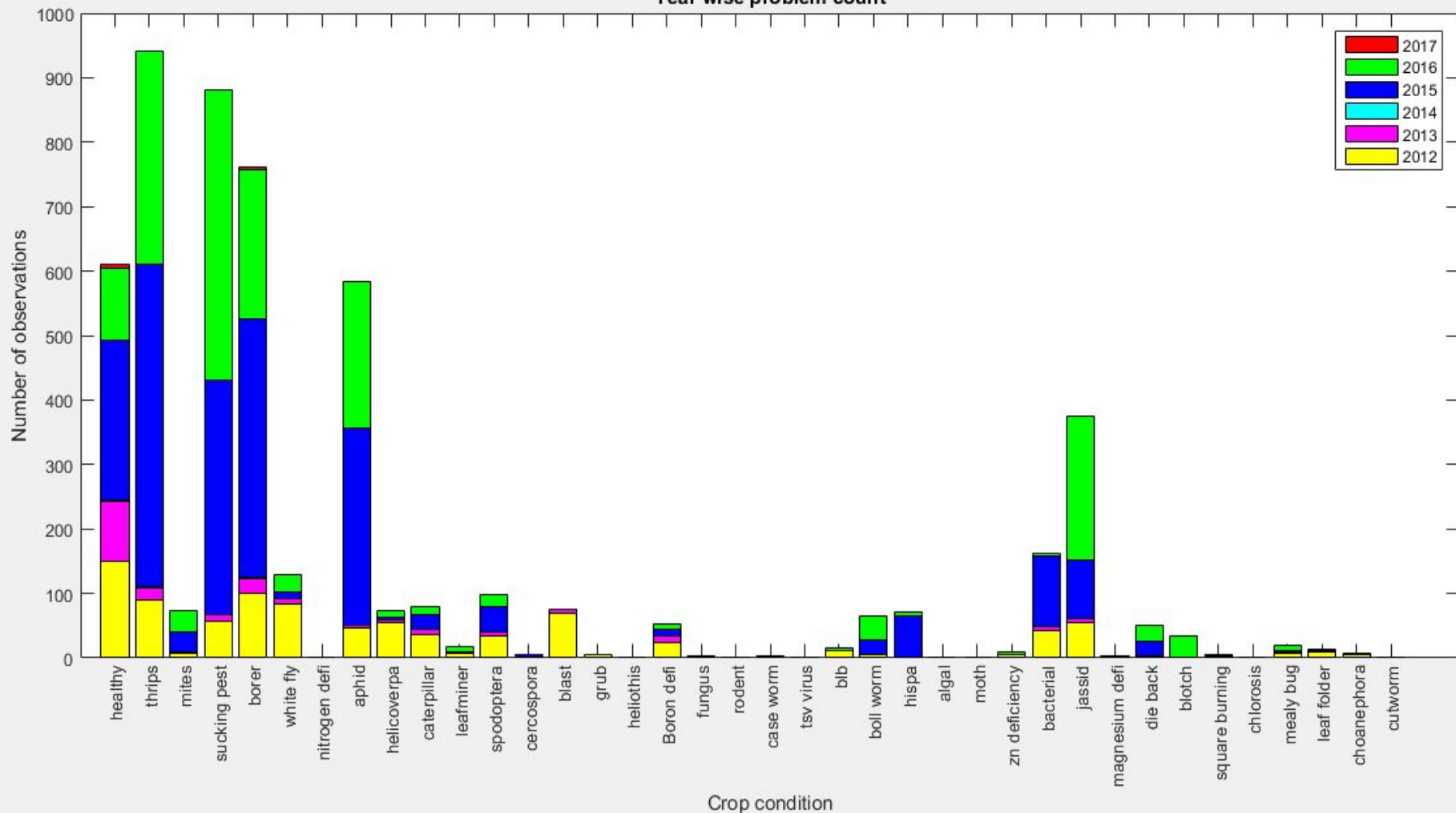


But here, the entries do not sum up to 6764. They sum up to a total of 5239. It is due to following reasons :

- Some of the advices are regarding fertilizers and do not contain the problem.
- Advices containing general preventive measures.



Year wise problem count



**Year** : **Number of advices**

2012 : 1834

2013 : 672

2014 : 6

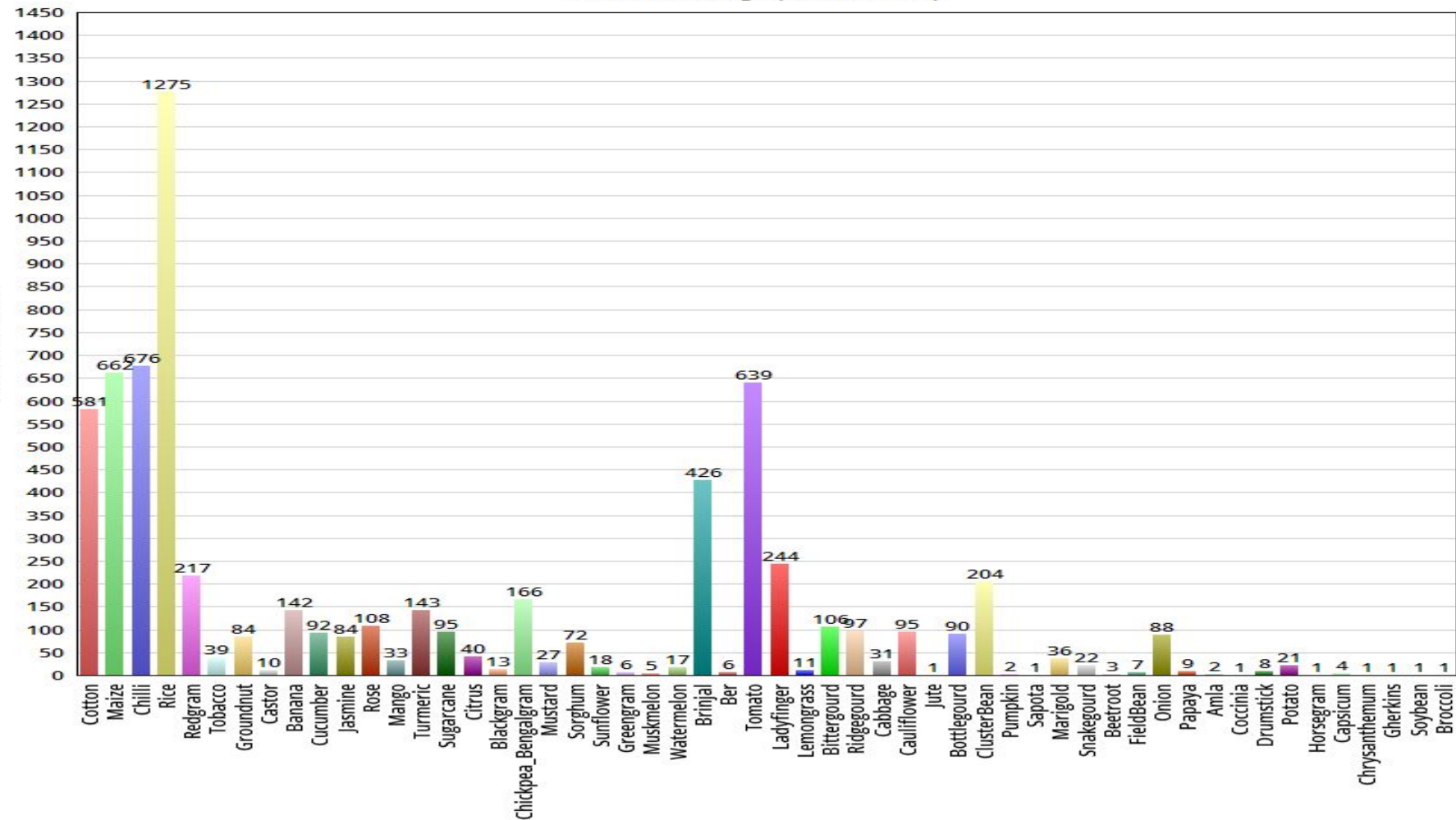
2015 : 2400

2016 : 1841

2017 : 11

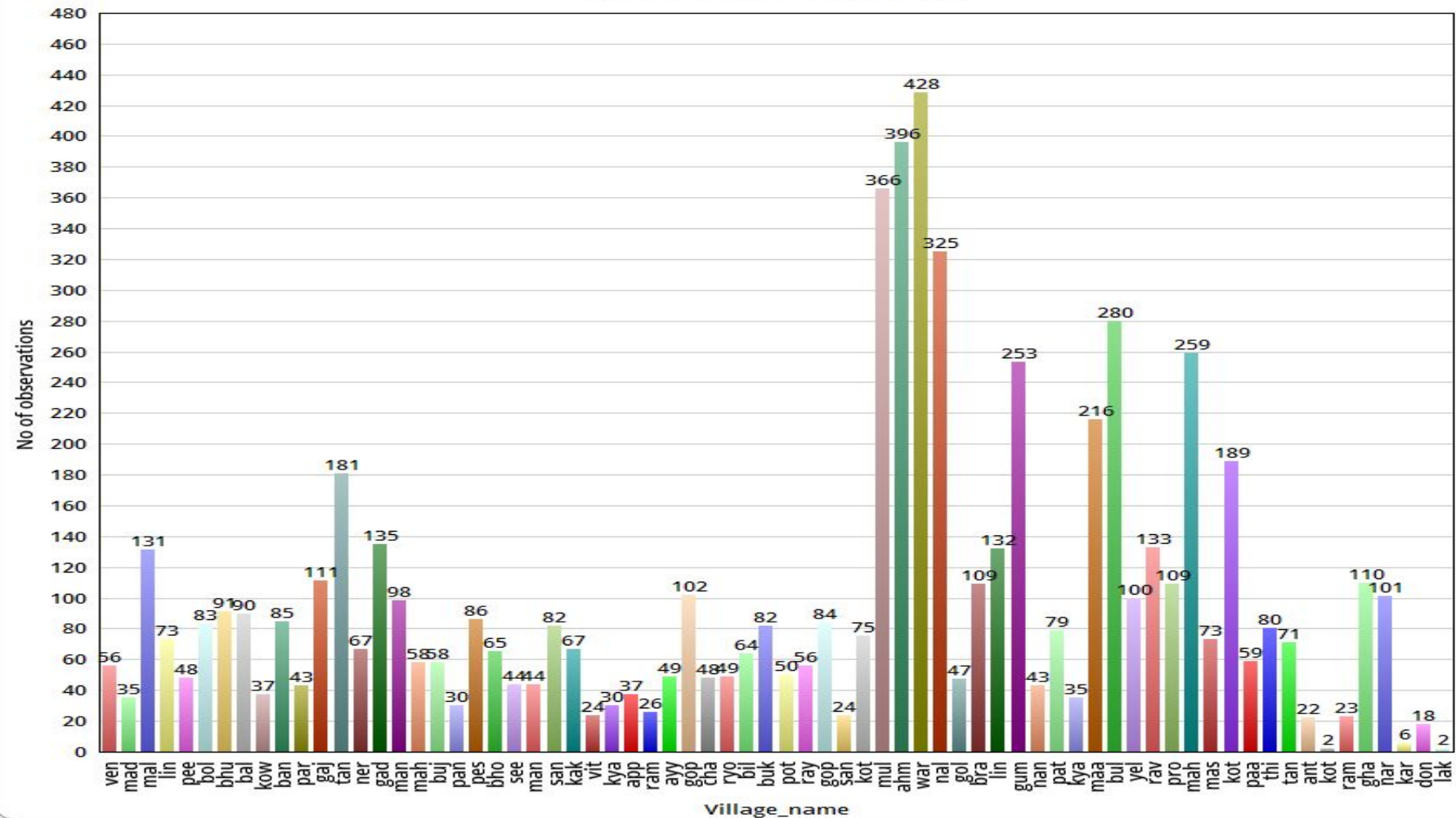


**Crop\_count**  
No of obs having a particular crop





**Village wise Distribution**  
No of obs from a particular village



# Timeline

- One week from now : Developing dashboard for interactive visualization of the data and prescriptive analytics service.
- One month from now : Model implementation.
- Next semester : recommender system.



# References

- <http://insait.in/AIPA2012/articles/009.pdf>
  - [http://www.saravananraj.net/wp-content/uploads/2014/12/45\\_AFITA\\_ICT-for-Agricultural-extension\\_India.pdf](http://www.saravananraj.net/wp-content/uploads/2014/12/45_AFITA_ICT-for-Agricultural-extension_India.pdf)
  - <http://insait.in/AIPA2012/articles/005.pdf>
  - [https://www.researchgate.net/publication/233572675\\_Analysing\\_dynamics\\_of\\_crop\\_problems\\_by\\_applying\\_text\\_analysis\\_methods\\_on\\_farm\\_advisory\\_data\\_of\\_eSaguTM.](https://www.researchgate.net/publication/233572675_Analysing_dynamics_of_crop_problems_by_applying_text_analysis_methods_on_farm_advisory_data_of_eSaguTM)
- 