Helping Farmers Harness the Power of Data

Team Members
Ben Cook
Charles Hornbaker
Ryan King
Conor Myhrvold

Background and Motivation

Unprecedented amounts of data are available on modern farms. From yield monitors to weather sensors to infrared imaging, farmers are able to keep track of every detail on their farms. However, most farmers are not taking advantage of this data. Much of the data is never reviewed after being collected. The data that is reviewed remains inaccessible, trapped in complicated legacy software, agronomist reports, and countless pages of spreadsheets.

Agricultural production lags far behind other industries in using data to make decisions, but data-based decision-making will be essential in the future. Farming is a challenging business. Reduced availability of water and increasing input costs are squeezing farmers' bottom lines during a time of extreme volatility in the commodities markets. New regulations on irrigation, pesticide use, and fertilizer application mean farmers must find a new way to boost yield. Soon, using data to optimize decision-making will no longer be a novel luxury -- it will be essential for farmers to stay in business.

The purpose of this project is to provide a tool for farmers to access and visualize their data with minimal effort by building a dashboard implemented in d3.js. Eventually, we imagine such a tool as a one-stop shop that allows farmers to see a farm not through the eyes of a 20th century agriculturalist, but also as the 21st century data scientist: the Bloomberg terminal for the farm, with relevant information about the farm itself in addition to other exogenous factors that could affect it.

Project Objectives

1. Show how the data that affects a farm changes over space and time.
2. Create modular visualizations for different aspects of farm-related data
3. Provide an intuitive, simple interface for exploring the data and conducting basic analysis.

Data

The data for this project is derived from several sources:

1. Crop yield from combines (geo-located data 6 times per second)

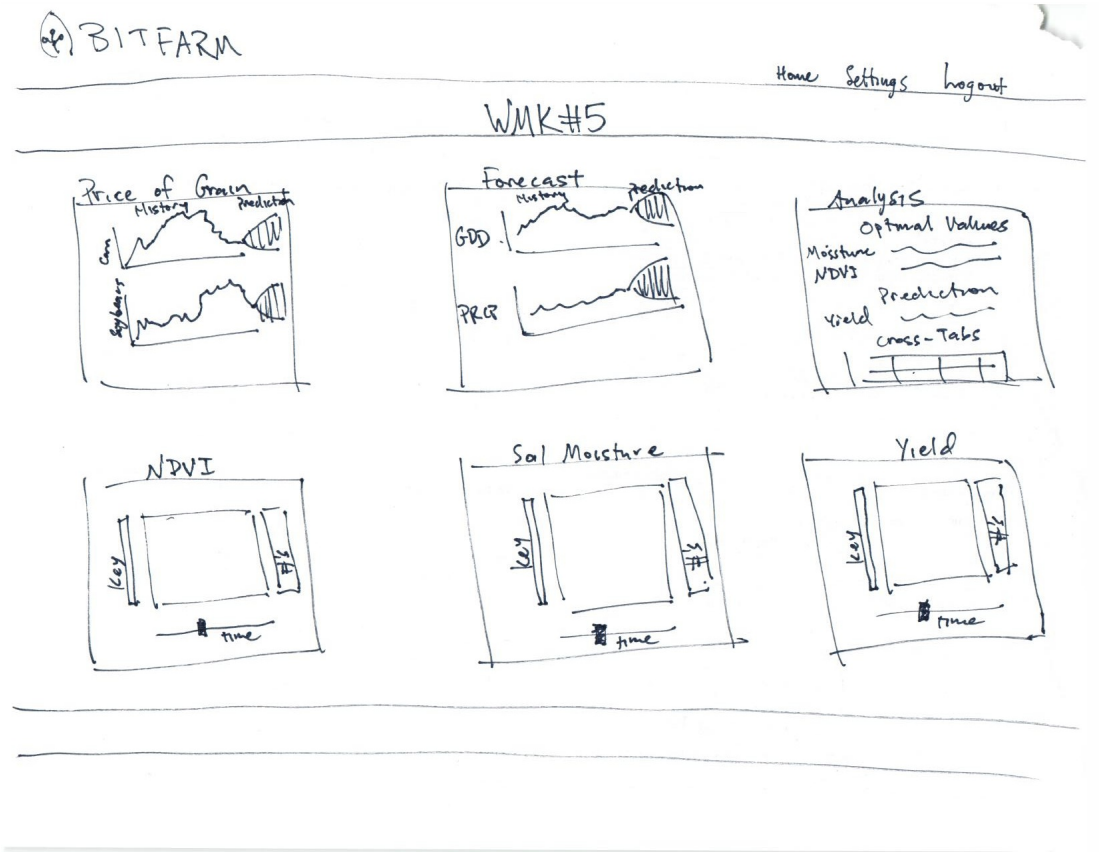
2. Crop yield at the county-level
3. Soil moisture throughout the United States since 1978
4. Daily precipitation, min. temperature, and max temperature throughout the United States
5. Soil type and drainage conditions throughout the United States
6. Grain prices over time

Data Processing

The data sources listed above have mostly already been collected and are in various stages of data cleanup. We may do a small amount of data analysis in order to, for example, show uncertainty about future grain prices with a Gaussian process.

Visualization

This project will consist of several visualizations accessed through a primary dashboard interface. Each module will display a different aspect or combination of the data related to the farm or farms. The sketch below depicts the general concept for the dashboard. Because there are four group members, we have the advantage of being able to produce several modular visualizations that cover a wider range of data for this project. The dashboard concept below will likely change over the course of the project, but it is included for initial planning purposes.



Yield: The yield module is one of the primary focuses of this project. This visualization will show the data for a given field collected from tractors. This makes up our most robust data set. This will be a geographical visualization that shows the level of yield and yield patterns for a field. The user can mouse over the area and see the actual values at any point, as well as highlighting areas of similar yield. Other data sources can be integrated to give a composite view of the existing features in the field and the collected data over space and time.

Price of Grain: This visualization will show the price of different types of grain over a chosen time period. Users can choose the type of grain and zoom in or out to see current prices and historical trends. It will be a multiline chart with brushing and detail areas. Additional metrics will be shown in an information box below.

Growing Degree Days (GDD) and Precipitation Forecast: This visualization shows precipitation and GDD in a single visualization. It will be a multiline graph that includes brushing and detail areas. (http://en.wikipedia.org/wiki/Growing_degree-day)

Analysis: This portion will provide various metrics from the available farm data for exploration and analysis. It will likely include a crosstab feature for selecting and comparing different data types.

Soil Moisture / NDVI: These modules will allow us to display data at the regional and county level. The visualization will be geographical over the entire country, with the option to compare similar regions and zoom to specific areas. It will also include options to explore the data over time.

Must-Have Features

The intent of this project is to provide a usable, simple interface to explore the data specifically applicable to a farmer. The must-have features below are based on the data that is already available and our main objectives:

- Display basic yield data and soil moisture levels
- Grain prices, GDD, and precipitation forecast

Optional Features

- Prediction capability for price of grain and other forecasts. We are implementing several predictive analysis methods in related projects for other courses, and will attempt to integrate these results into the visualization if possible.
- Other field-level measurements such as NDVI, and soil nutrients

Project Schedule

13 March: Proposal submitted
14 March: Individual team member tasks defined
Refine sketches

Identify additional data requirements / shortfalls
Create project repository / Post data / Progress Book

14-24 March: Individual component development

24 March: Team meeting / Progress report

31 March: Functional prototypes of individual components completed

7 April: Integrate components from all team members

10 April: Functional prototype completed

14 April: Project review with TFs

17 April: Team meeting / Progress report

22 April: Project website created / prototype visualization posted

28 April: Project screencast recorded

1 May: Project and Screencast submitted