**Capstone Project Proposal: (Organic Produce – Price prediction)**

In the current environment, Food is an important attribute of our everyday lives.  The type of food intake will have a direct impact on our health. The world is leaning towards Organic food eventually.

http://www.ecowatch.com/10-reasons-consumers-buy-organic-1881899943.html

There are two types of vegetables and fruits available to us – “Organic” or “Conventional”. This is **based** on demand **and** creates resources. The price that a person requires to pay can vary by different factors such as:

1. Availability
2. Logistics
3. Location

I will work on a **generic dataset of vegetables and fruits with their respective pricing.**

1. What is the **problem you want to solve**?

**I want to solve following problems:**

1. What are the **price ranges** of organic vegetables vs Conventional vegetables?
2. What is the **probability of the increase** in price using predictive Analysis?
3. Which **market location has more price difference**?
4. Which **vegetable/fruit is costlier**?
5. Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn’t have otherwise?

My client will be **Individual Wholesale Consumers** who are interested in buying Organic and conventional produce.

The List of Consumers are:

1. Restaurant/Hotel
2. Grocery Retailers
3. Individuals /Home
4. Culinary schools

If they have the predictive prices of the produce, they can decide when to buy **produce.** (seasonal availability/ month prices based)

1. What data are you going to use for this? How will you acquire this data?

**Type of data:** Secondary Data.

**Data-source:**  <http://www.ers.usda.gov/data-products/organic-prices.aspx>

1. In brief, outline your **approach to solving this problem** (knowing that this might change later).
   1. I will first clean the dataset. Do the required data wrangling.
   2. Resolve and conclude statistical methods that be applied.
   3. Model and visualize the dataset and use predictive Analysis.
   4. Analyze the findings and solve the problems that have been formulated.
   5. Present and stipulate the findings.

5. What are your **deliverables**? (Typically, this would include code, along with a paper and/or a slide deck)

1) I will formulate the dataset with R programming

2) I will provide Statistical Findings.

3) I will present Solutions in a slide deck (PowerPoint presentation).