Soil Analytics

Report on Initial Findings

Author: Nathalie C. Cochrane

Company: Regenerative Cloud

Date: December 3rd, 2020

Context

The farm covers an area of 1,516 acres. Overall, the farm's operation has been moving away from tillage since the 90s and efforts to move to regenerative practices have been implemented in the last 5-7 years.

For this analysis, 2 data sets were provided. The first data set, contains 64 columns and 640 entries. The second data set, contains 18 columns and 69 entries. The first data set will be referred to as DS1 and the second data set will be referred to as DS2.

After an initial discussion with the farmer, we have decided to focus this analysis on Farm A as this was the part of the operation that has been exposed to Hog waste between 1975 and 1996 which has resulted in beneficial impact on the soil. The goal of this analysis is to identify positive transformations in the soil that are a result of the regenerative practices.

DS1 & DS2 General Field Information:

- Soil Class: MIN: Mineral soil. Low percentage of humic matter. Target pH 6.0

- HM percent: Humic matter percent is a measure of the chemically active fraction of organic matter. The humic matter values are usually much lower than the actual organic matter content.

- W/V: The soil weight/volume is shown in grams/cubic centimeter and is used to determine the soil class. Soils high in sand have high W/V, while soils high in organic matter have low W/V. Loamy and clayey soils are intermediate.

- CEC: Cation exchange capacity is a measure of the soil's capacity to hold basic cations such as potassium, calcium, and magnesium, plus the acidic cations hydrogen and aluminum. CEC increases as soil organic matter, pH, and clay content increase. This calculation is given in milliequivalents per 100 grams of soil. Cations are positively charged ions such as calcium (Ca++), magnesium (Mg++), and potassium (K+). The larger the CEC value, the more cations the soil is able to hold against leaching.

- BS%: Base saturation percent is the percent of the CEC that is occupied by the basic cations [potassium (K), calcium (Ca), and magnesium (Mg)]. BS% indicates the pH and lime status of the soil. As pH increases, BS% also increases. On soils that are properly limed, BS% should range from 70 to 90. On acidic soils, BS% ranges from 50 to 60.

- Ac: Exchangeable acidity is the portion of the CEC that is occupied by acidic cations [Aluminum (Al), hydrogen (H)]. The amount of acidity decreases as soil pH increases.

- pH: Soil pH is a measure of the active acidity [hydrogen (H)] in the soil solution.

- P-1 and K-1: Phosphorus (P) and potassium (K) are shown as indexes used to evaluate nutrient availability to plants. Fertilizer recommendations for P and K decrease as the index increases. An index of 25 or lower is

considered too low for optimum plant growth. A range of 26 to 50 is medium, and an index of greater than 50 is high. Adding more phosphorus when the index is greater than 50 should not generate a response. Fertilizer rates are given as pounds of P2O5 and K20 per acre or per 1,000 square feet.

- Ca and Mg%: Both calcium (Ca) and magnesium (Mg) are shown as percentages of CEC. Soil calcium is seldom low enough to limit plant growth. In general, calcium is the most common cation in the soil. Calcium percentage is essential for calculating CEC and to evaluate the relationship between calcium, magnesium, and potash (K). If the magnesium percent is low, magnesium will be recommended in the form of dolomitic lime or of a fertilizer containing magnesium.

- S (sulfur), Mn (manganese), Zn (zinc), Cu (copper): An index is determined for each of these nutrients. An index of 25 or lower is considered too low for optimum plant growth. A range of 26 to 50 is medium, and a range of greater than 50 is high. Adding more nutrients should not generate a response when the index is greater than 50. Sulfur is difficult to interpret since, like nitrogen, it leaches readily from sandy soils.

- Na is sodium (https://ucanr.edu/blogs/blogcore/postdetail.cfm?postnum=12766)

- Fields with 100% null values or all = 0 values: Address2, LayerID, Last Crop, LIME MONTH, LIME YEAR, LIME IN TONES,Mn Avail Crop2,SS Result (soluble salts), NN Result, AM Result, comment crop 1, all crop 2 fields. All these fields have been dropped from the data set

- NIT CROP1- Nitrogen recommendation from soil analysis lab. This field has been dropped also as it has a singular value for all entries 50-70.

Analysis Findings and Visualizations

For this analysis, we are focusing only on Farm A entries and 4 main variables: Cation Exchange, pH, P and K

FARM A, DS1 Descriptive statistics:

	CATION EXCHANGE	pH	P	K
count	260.00	260.00	260.00	260.00
mean	7.54	6.16	155.04	72.27
std	1.45	0.19	82.73	15.10
min	4.90	5.70	40.00	33.00
25%	6.40	6.10	98.00	62.00
50%	7.20	6.20	127.00	70.00
75%	8.40	6.30	201.00	83.00
max	12.80	6.70	559.00	124.00

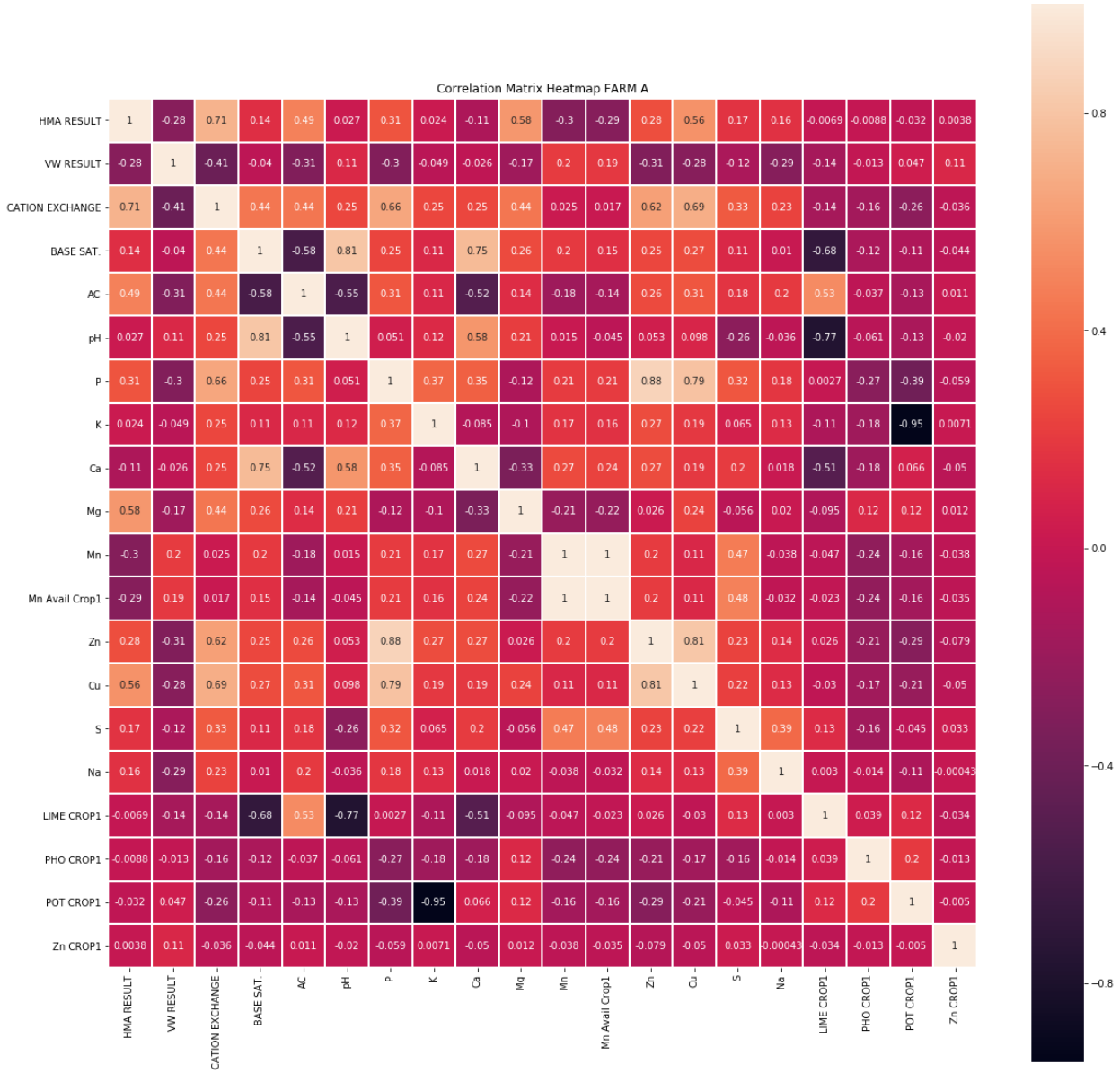
This chart shows the basic statistics on our 4 key variables:

- Cation Exchange's average is at 7.54
- pH average is 6.16
- P average is 155.04
- K average is 72.27

For example: the row "75%" reads that 75% of the observations of Cation Exchange are below 8.4.

We also see that there is a potential measurement error that occurred as there is a very high max for P and K, 559 and 124 respectively.

FARM A, Correlation Matrix of all variables:



This rich visualizations allows us to see the relationship between all variables present in the DS1 for Farm A.

When implementing Machine Learning algorithms, it is important to address highly correlated variables as those can have a strong influence on their predictive abilities. At this stage, we are only exploring so we won't remove any variables.

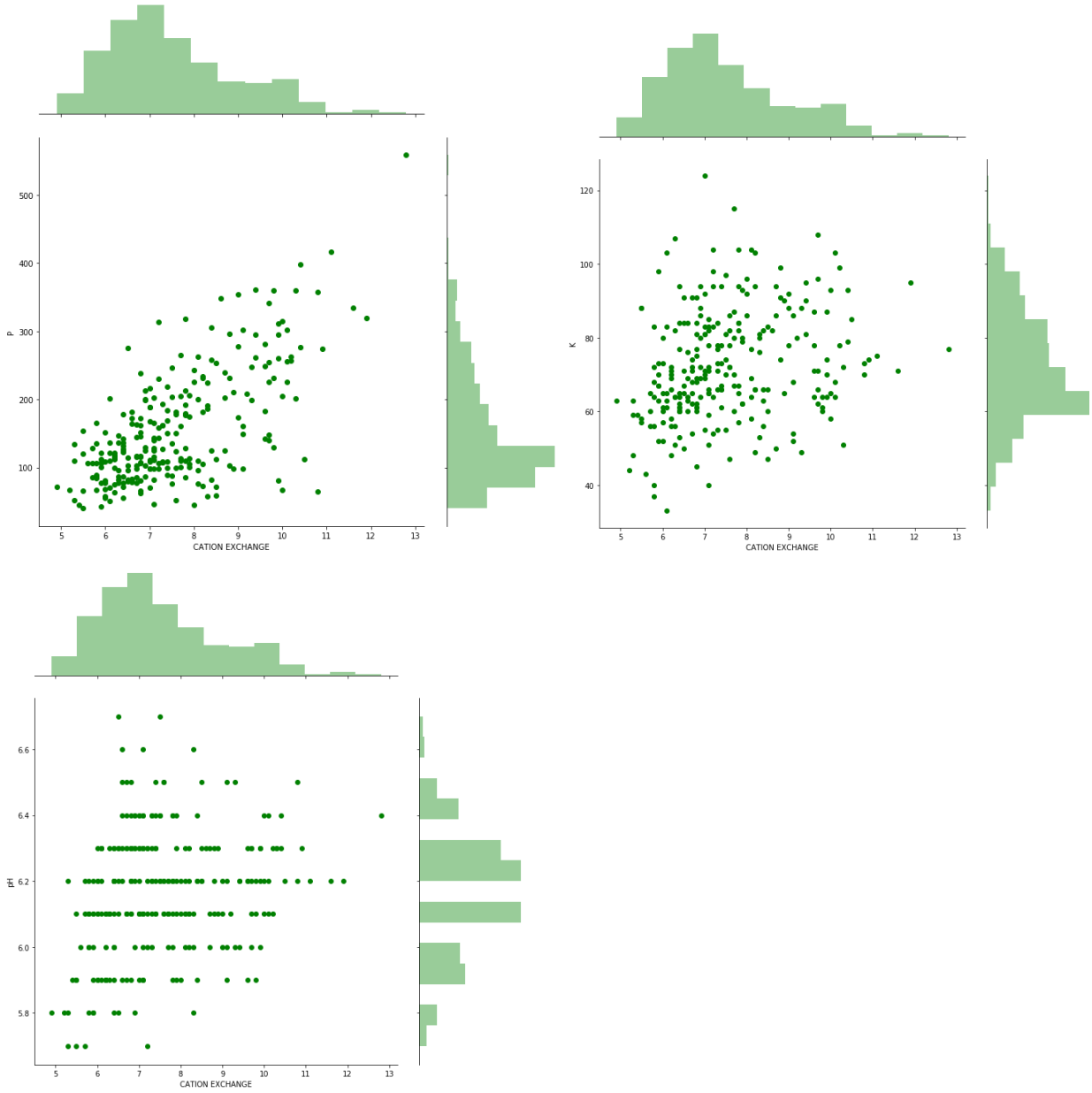
FARM A, Correlation Matrix Learnings:

- CATION EXCHANGE isn't showing strong correlation with any particular variables. The most meaningful correlations are with HMA results, P, Zn and Cu
- pH is positively correlated (0.81) with Base Sat and negatively correlated with Lime Crop 1 (-0.77)
- P is positively correlated with Zn and Cu (0.88 and 0.79 respectively)
- K is very negatively correlated with Pot Crop 1 (-0.95).
- Mn and Mn Avail crop 1 have a correlation of 1. This means that they are near identical fields.
- Zn and Cu are positively correlated (0.81).

FARM A Variables of Interests, Relationship Visualizations

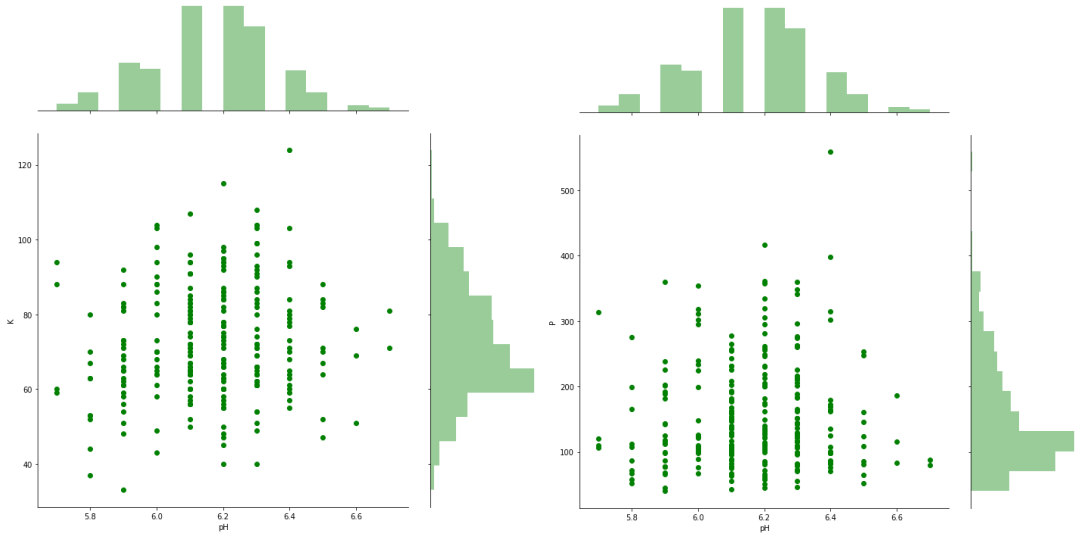
All the charts in this section show the relationship between 2 variables. The center of the plots are the values pairs between the 2 variables, the top and left sections are the distributions of each variables.

CATION EXCHANGE with pH, K and P:



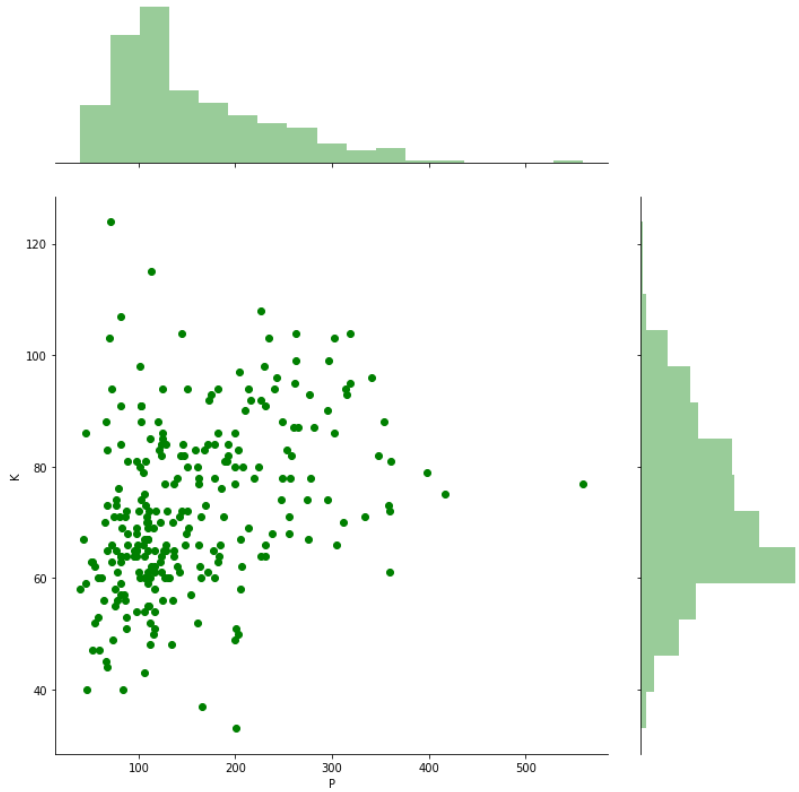
Here we see that Cation EXCHANGE has the strongest correlation with P. K and pH have minor influence on Cation Exchange.

pH with K and P:



Here we see that pH doesn't have strong correlation with neither K nor P

K and P:



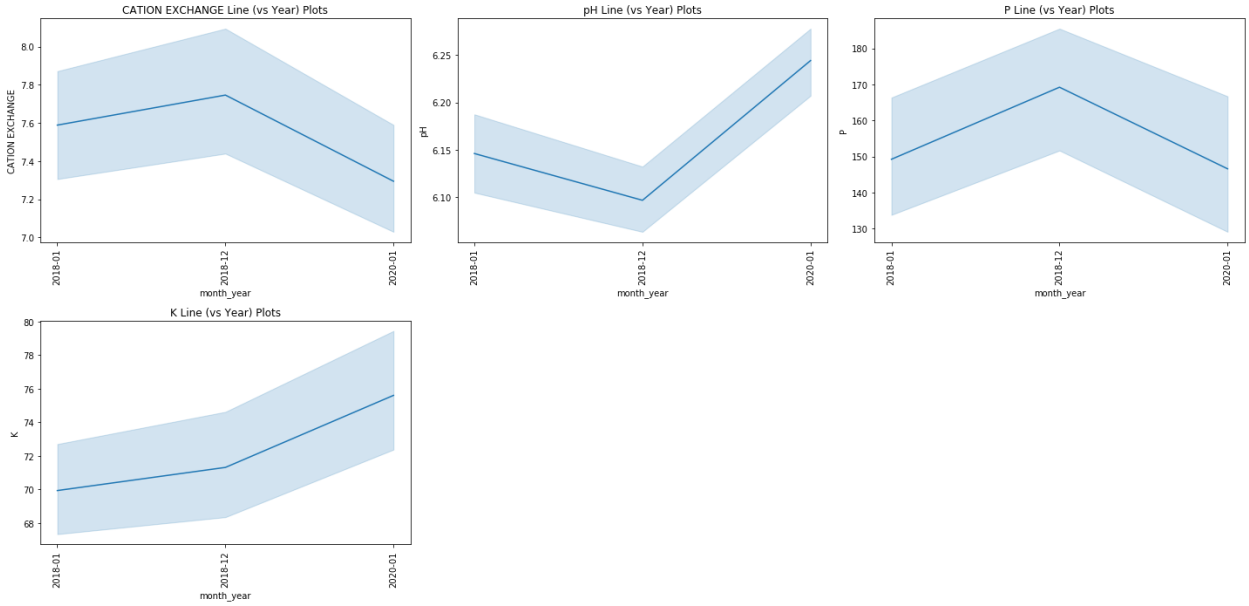
Here we see that K and P are weakly correlated.

FARM A Variables of Interests, Fluctuations Over Time

DS1:

Here we see the fluctuation of our 4 variables of interest over time. DS1 provides us with visibility in the last 2 years (2018- 2020) with 3 measurements taken : Jan 2018, Dec 2018 and Jan 2020. The blue line shows the average and the blue halo around the line shows the possible values across the samples.

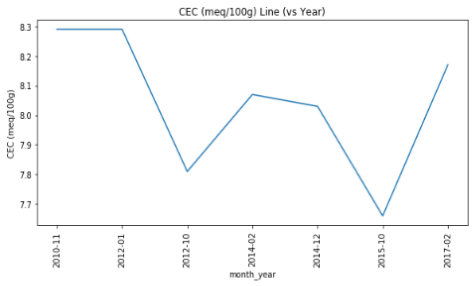
- Cation Exchange has overall decreased from 7.6 on average to 7.4
- pH has increased from 6.15 to 6.25
- P has remained similar at around 150
- K has increased from 70 to 75

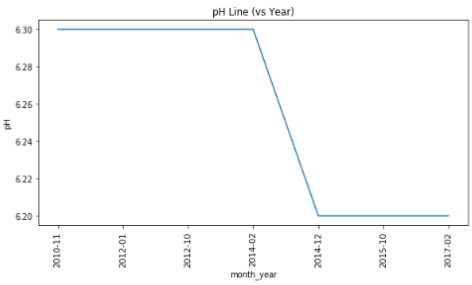


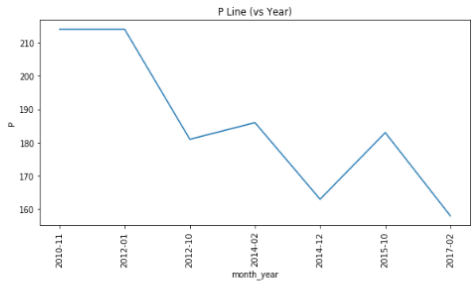
DS2:

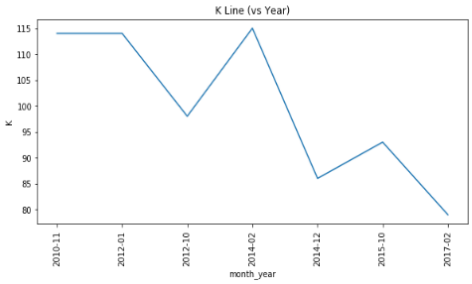
DS2 provides us with visibility in the 8 years (2010- 2017) with 7 measurements taken: Nov 2010, Jan 2012, Oct 2012, Feb 2014, Dec 2014, Oct 2015 and Feb 2017. The blue line shows the value for each period.

- Cation Exchange has overall decreased from 8.3 to 8.2 with strong fluctuations in between.
- pH has decreased from 6.3 to 6.2
- P has decreased from 210 to 160
- K has decreased from 115 to 80









Overall it seems that are 4 variables have been on a downward trend since 2010. pH and K have increased in the last few years but Cation Exchange and P have continued to decline.

Next Steps: Schedule a discussion with the farmer to understand his take aways from this analysis and possible further deep dive.