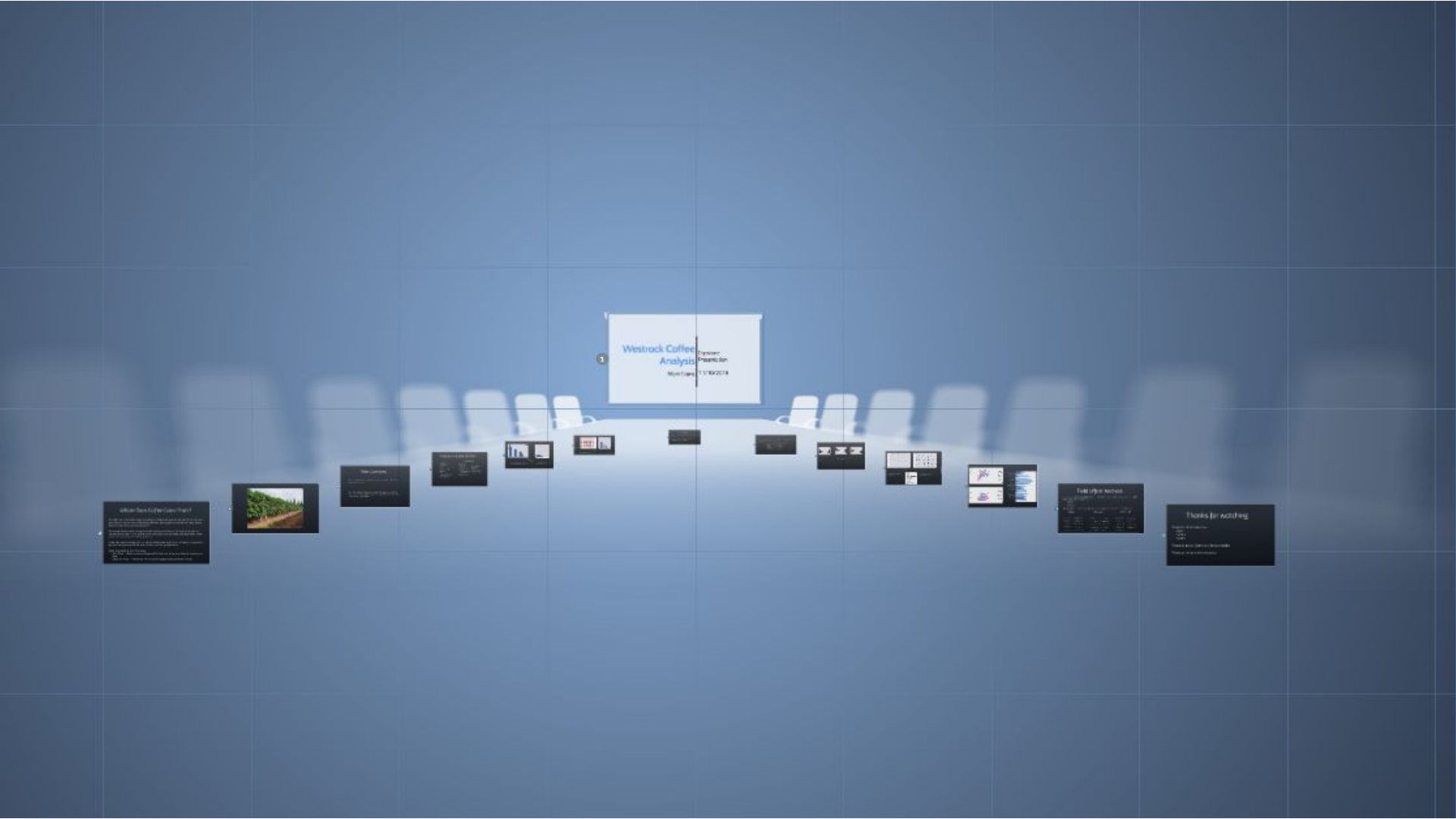
Westrock Coffee Analysis

Capstone Presentation

Mark Evans

11/10/2016


Westrock Coffee Analysis
Exploratory Presentation
Thanks for watching

Where Does Coffee Come From?

The coffee tree is a tropical evergreen shrub (genus Coffea) and grows between the Tropics of Cancer and Capricorn. The two most commercially important species grown are varieties of Coffea arabica (Arabicas) and Coffea canephora (Robustas).

The average Arabica plant is a large bush with dark-green oval leaves. The fruits, or cherries, are rounded and mature in 7 to 9 months; they usually contain two flat seeds, the coffee beans. When only one bean develops it is called a peaberry.

Coffee trees yield an average of 2 to 4 kilos of cherries and a good picker can harvest 45 to 90 kilos of coffee cherry per day; this will produce nine to 18 kilos of coffee beans.

Coffee is harvested in one of two ways:

- Strip Picked – all the cherries are stripped off of the branch at one time, either by machine or by hand.
- Selectively Picked – only the ripe cherries are harvested and they are picked by hand.

Main Questions

Question 1: Which courses (when adopted by a farmer) boost the farmer's coffee yield the most?

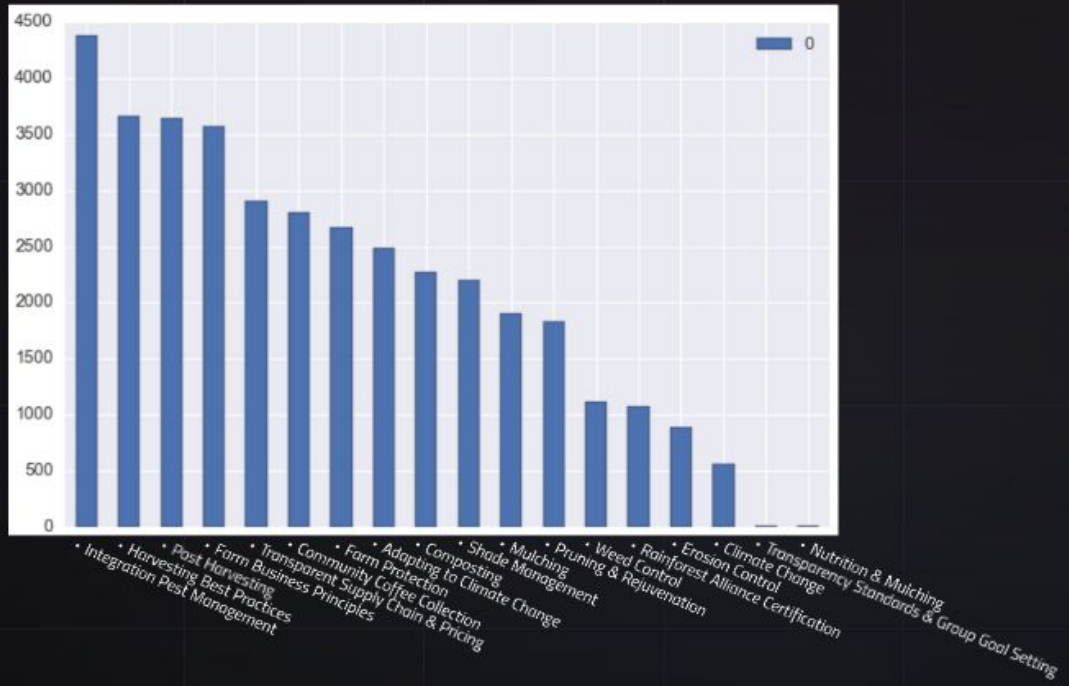
Question 2: How do the field officers ratings differ from one another? Are there some that are more lenient or more stringent? Can we do some sort of calibration across field officers?

What does the data look like?

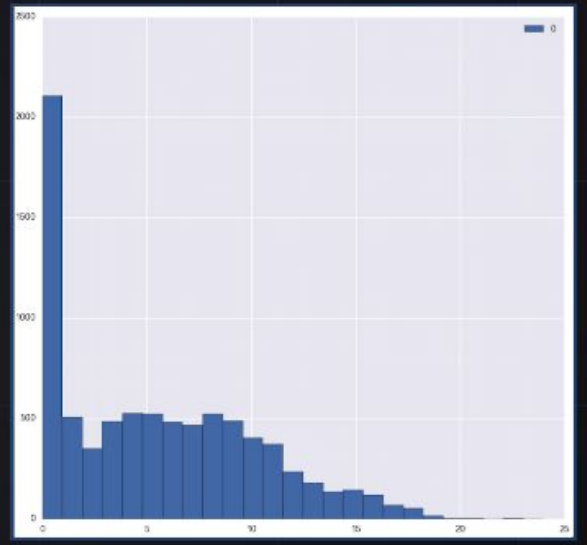
Station: Region
Staff: Field Officer
Group ID: ID of group of farmers
Household ID: ID of farmer
Farmer Name
Attendance: Percent that farmer attends courses
Adoption: Percent that farmer applies courses
2015 Yield
2016 Yield
Courses: Courses that farmer took
Trees: Number of trees on farm
Trees Producting: Number of trees that are producing yield

Courses Offered:

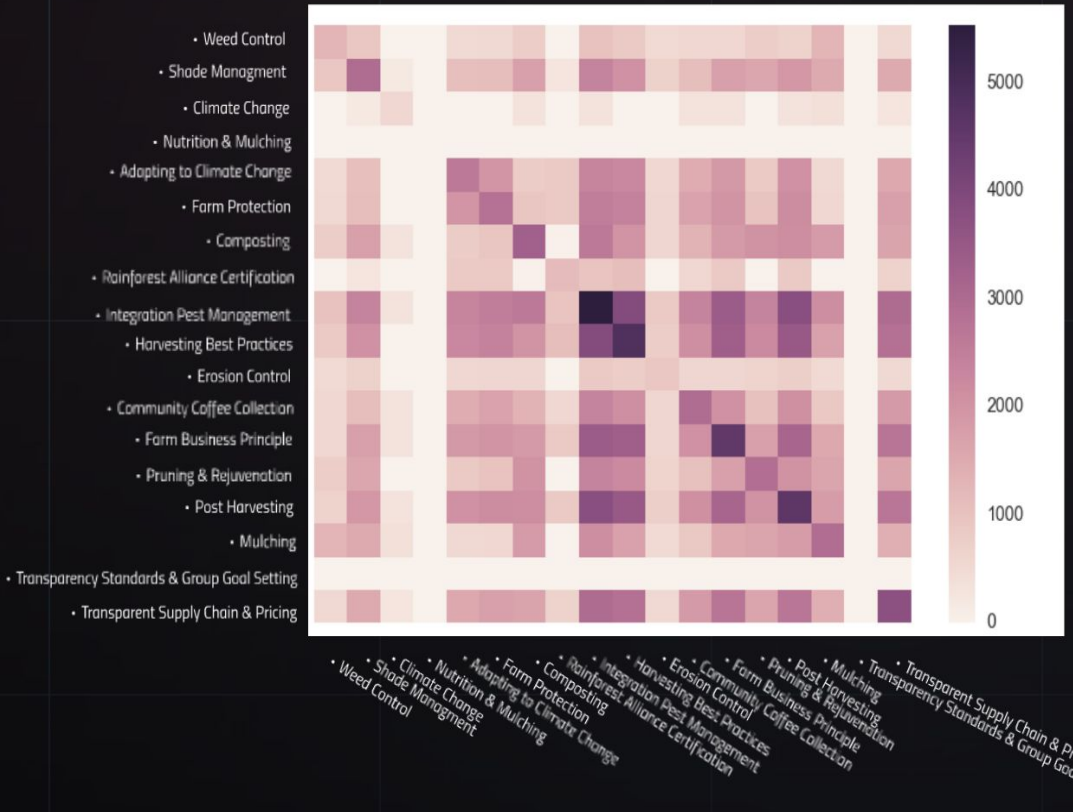
- Weed Control
- Shade Managment
- Climate Change
- Nutrition & Mulching
- Adapting to Climate Change
- Farm Protection
- Composting
- Rainforest Alliance Certification
- Integration Pest Management
- Harvesting Best Practices
- Erosion Control
- Community Coffee Collection
- Farm Business Principle
- Pruning & Rejuvenation
- Post Harvesting
- Mulching
- Transparency Standards & Group Goal Setting
- Transparent Supply Chain & Pricing



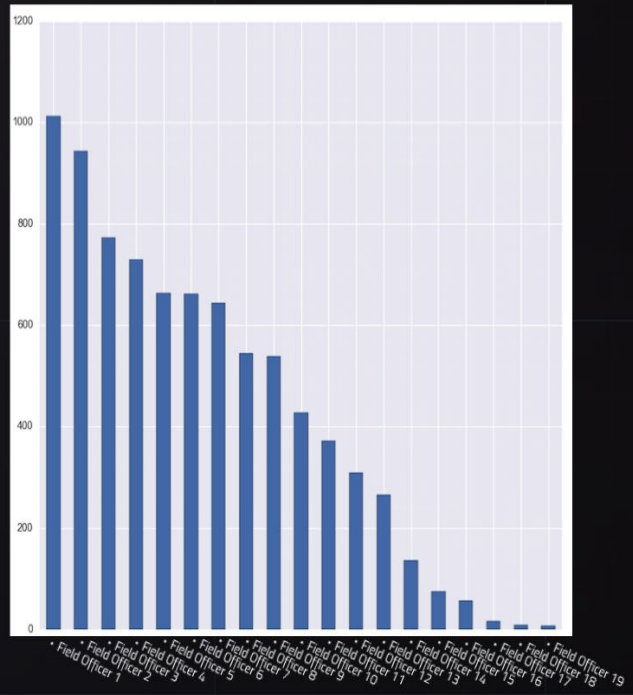
The number of farmers who took the course



The number of times farmers took x number of courses



Number of times courses are paired together



The number of farmers under each field officer

Course Analysis

Created Dummy Variables for each Field Officer, Group ID, and Courses taken once, and Courses taken 2 or more times

Linear Regression to model features vs. Average Yield Between 2015 and 2016

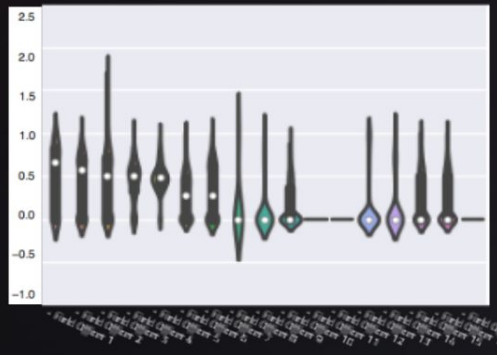
- K-Fold Cross Validation to get training, test set
- Backwards Stepwise Selection to remove non-significant features
- Intercept as 1 in the dataframe

Results

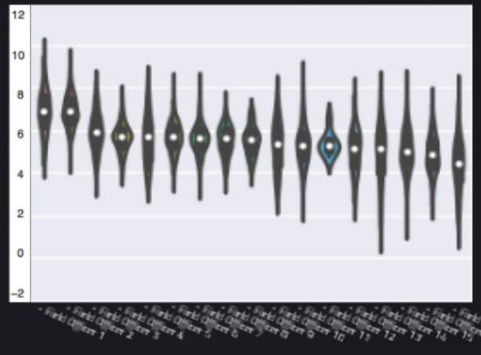
The intercept coefficient represents an increase of 231. The coefficient for adoption is 28, which represents an increase per farmers of 28 x the rate for a particular farmer.

There are 4 courses, when adopted, that had a significant positive impact on the average yield:

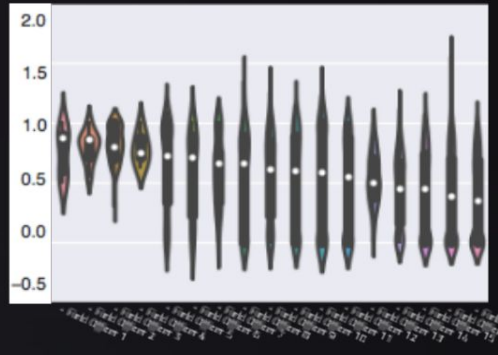
- Shade Managment 2 or more times = 593 x rate - 244
- Composting = 360 x rate - 49
- Weed Control = 268 x rate - 12
- Integrated Pest Management
    - 1 times = 205 x rate + 26
    - 2 or more times = 301 x rate - 100



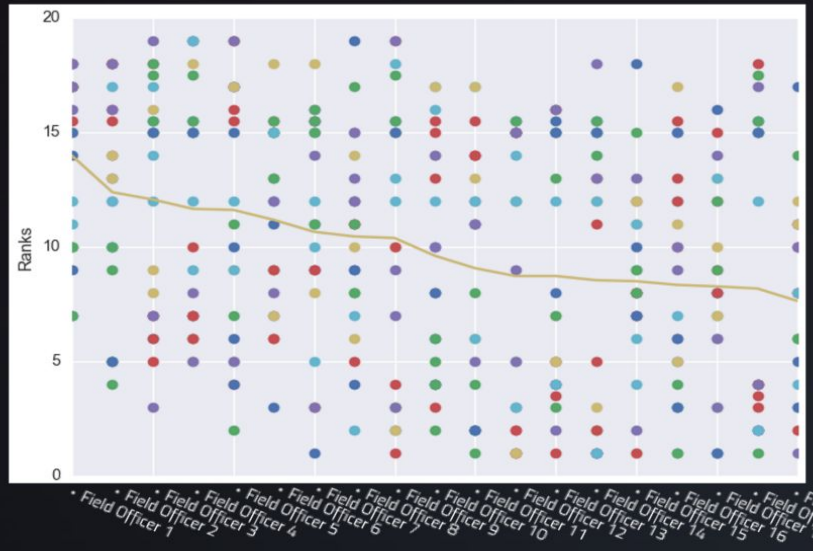
Adoption

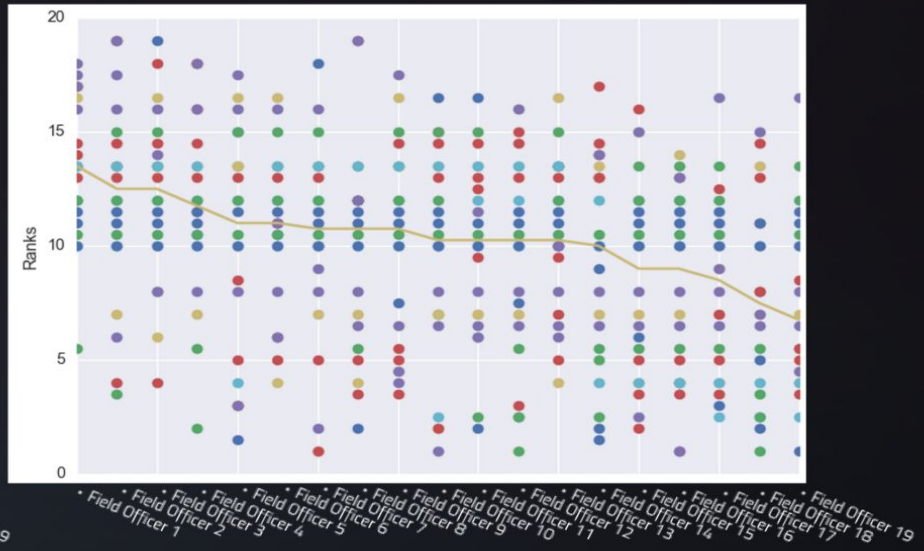


Average Yield (Log)



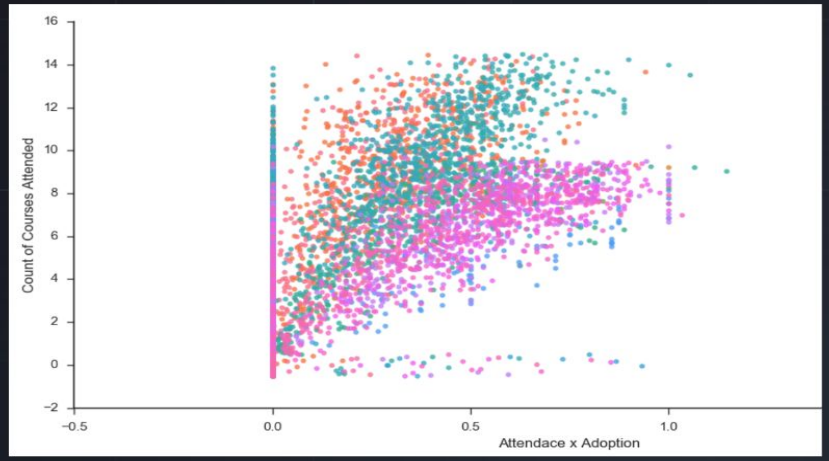
Attendace

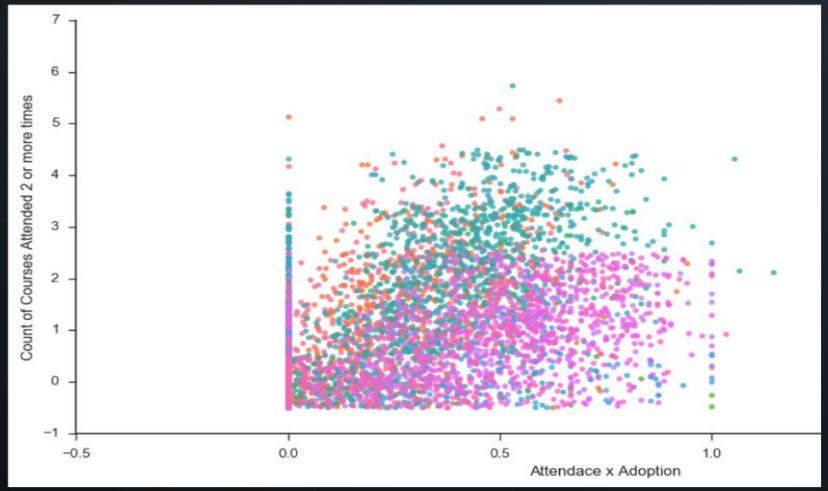


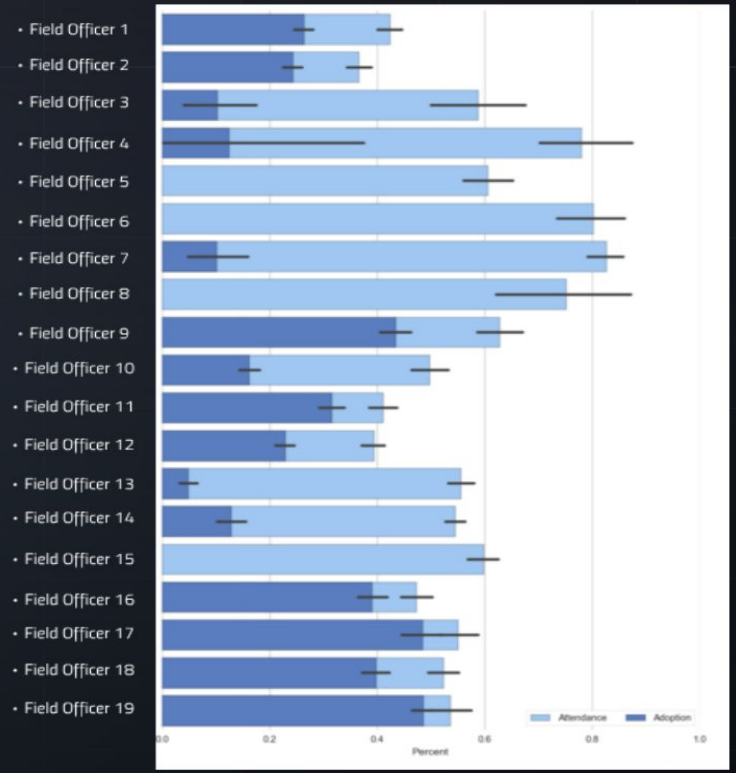


The rankings of every mean column with the average

The rankings of every median column with the median


Count of Courses Attended
Attendace x Adoption


Count of Courses Attended 2 or more times
Attendace x Adoption


Field Officer 1
Field Officer 2
Field Officer 3
Field Officer 4
Field Officer 5
Field Officer 6
Field Officer 7
Field Officer 8
Field Officer 9
Field Officer 10
Field Officer 11
Field Officer 12
Field Officer 13
Field Officer 14
Field Officer 15
Field Officer 16
Field Officer 17
Field Officer 18
Field Officer 19
Attendance
Adoption
Percent

Field Officer Analysis

Grouped all farmers by their field officer and calculated the mean and median values for each field officer in the following categories:

- Attendance
- Adoption
- Average Yield

The top 5 field officers who have the highes mean and median values for each category are:

	Adoption		Attendance		Average Yield	
	Mean	Median	Mean	Median	Mean	Median
1.	H. V.	U. R.	K. N.	N. S.	U. J.	U. J.
2.	M. E.	M. E.	N. S.	H. A.	U. I.	U. I.
3.	G. G.	H. V.	U. F.	U. F.	N. C.	N. C.
4.	U. R.	G. G.	H. A.	A. U.	U. R.	U. R.
5.	U. F.	U. F.	A. U.	K. N.	M. C.	M. F.

Thanks for watching

Thanks to all the instructors

- Ryan
- Scott S
- Scott C

Thanks to Brooke Cantrell and Westrock Coffee

Thanks to everyone who showed up