ACME Ghana Agricultural Profit Analysis
OMSBA-5112 Summer 2020

By: Kaitlyn Jakubek, Gabe Boucaud & Brittney Cherry

Memorandum

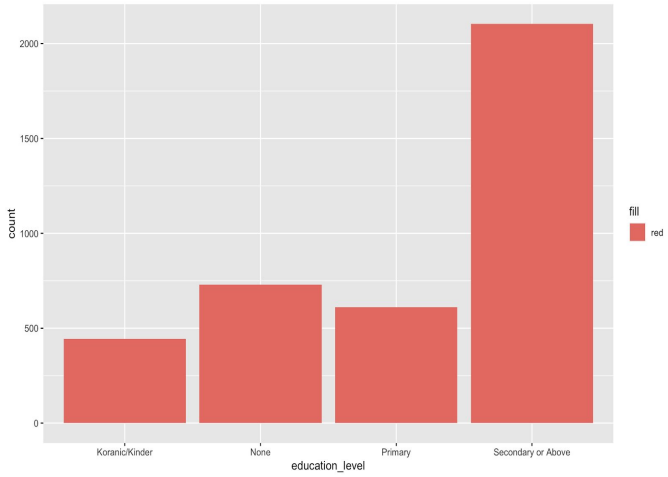
To: Board Members of ACME
From: Brittney Cherry, Kaitlyn Jakubek & Gabe Boucaud
Date: September 3,2020
Subject: Analysis of Agricultural profit in Ghana

Purpose and Scope:
The purpose of this analysis is to determine the most profitable way for ACME to enter Ghana's agricultural market. Our team focused on the effects location, education, and community attributes have on agricultural profits. Specifically, this report focuses on the location where ACME should invest and focus further research.

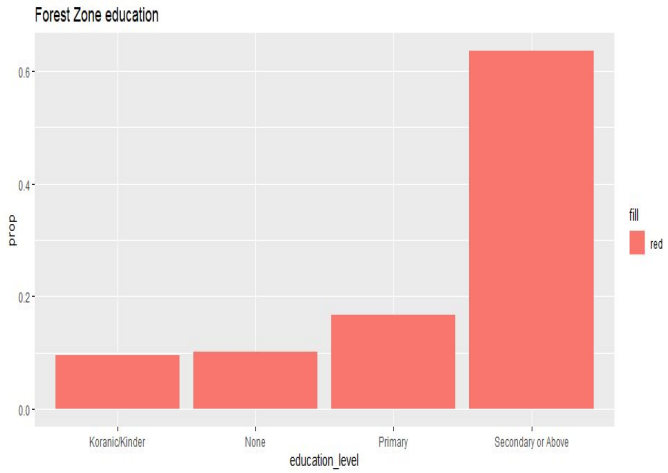
This report uses data collected through the Ghana Living Standard Survey 4 (GLSS 4) from 1998 to 1999 to illustrate the effects household education and community demographics have on household agricultural profits in Ghana. The analysis and summary below identify educational and community characteristics which are helpful in determining how and where ACME can successfully enter and expand into Ghana's agricultural markets.

Analysis:
The following recommendations have significant impact on profitability, measured by total agricultural profits by area of land (ropes) in Ghanaian households.

In terms of education, an additional level of education from having completed no school to completing secondary school or above increases agricultural profits by 30%. Of interest, is how many households have achieved an educational level high enough to see increased profits. The following chart displays the number of households, within the survey, who have achieved each maximum level of education.

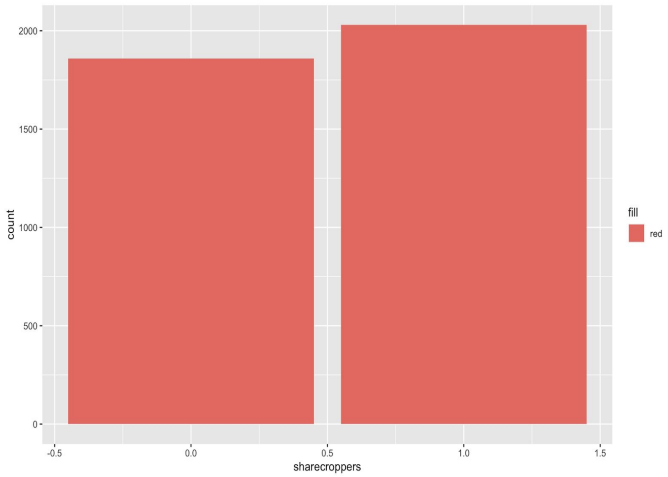


In terms of the type of land to purchase, on average land in forested ecological zones yield the most consistent profits and tend to be the most educated in the entire country. The following graph breaks down household maximum education level based on households located in forested ecological zones. The comparison between the education levels of all those surveyed and those surveyed who live in forested ecological zones shows distinct differences in educational household maximum educational attainment.



Looking deeper into community characteristics, households in communities where there is access to a bank, community coop and where there is sharecropping see an increase in agricultural profits. With households involved in sharecropping having a 8.6% increase in the level of profit per rope.Additionally, landowners who hire others to work on their land earn a higher profit than those who only work the land themselves. This presents a key opportunity for ACME, if land is purchased in communities that have high levels of education and we allow members of the community to use it for sharecropping it is likely that we will see profits increase.

With sharecropping leading to increased profits the number of households within the survey that are involved in sharecropping is displayed below. The households with a value of 1 are involved in sharecropping.Whereas, those households that show a 0 aren't.



Conclusion:

In summary, we strongly recommended ACME consider the availability of forested ecological communities whose household members have a primary education and cultivating an environment where sharecropping is achievable within a community. However, these results are limited and more research is needed to determine which crops and methods lead to higher profits. ACME will also want to consider entering Ghana's agricultural market in a manner that unites people towards a common goal and assists with the existing community focused mindset.

Statistical Analysis

Abstract

The purpose of this analysis is to understand the effects location, education, and community attributes have on a household's agricultural profit. From this analysis we are able to provide suggestions and insights on what ACME should consider when entering Ghana's agricultural market.

Theory & Hypothesis

The survey includes information from 5,998 sampled households. We narrowed this down to include only those households with income and or expenses from agriculture. This encompasses estimated income from agricultural products grown and consumed by the same household. We also focused on households that own, rent, or sharecrop farmland in known units including acres, poles, and ropes. These land areas are combined with agricultural profit to allow comparison of households with both large and small quantities of land. After these filters are applied, our sample size includes the profits per rope of 3,927 households.

Our analysis focuses on how community attributes and education, within households, affect agricultural profit. With regards to education, we based our analysis on the highest educational level achieved per household. We also looked at the differences between types of schooling that household members received. Within community attributes, we focused on the following characteristics of households located in rural areas: available schools, agricultural extension centers, farm cooperatives, sharecropping, and farming mutual aid. We expect the following:

H1.

Null Hypothesis: The effect of higher education on household profits per rope isn't statistically significant at a 1% level of confidence.

Alternative hypothesis: The effect of higher education on household profits per rope is statistically significant at a 1% level of confidence.

H2.

Null Hypothesis: The effect of community trade on profits per rope isn't statistically significant at a 1% level of confidence.

Alternative Hypothesis:Households with access to community trade in agricultural goods will have a statistically significant increase in profits per rope at a 1% level of confidence.

H3.

Null Hypothesis: The effect of a favorable eco system on profits per rope isn't statistically significant at a 1% level of confidence.

Alternative Hypothesis:The effect of a favorable eco system on profits per rope is statistically significant at a 1% level of confidence.

Method

The following analysis was completed using R programming language. The source code can be found on GitHub, although the datafiles will need to be loaded separately: https://github.com/kjakubek/Data-Translation-Challenge.git

Steps taken towards analysis follow:

Step 1: Upload relevant libraries and datasets

Our team found the *tidyverse* and *haven* libraries encompassed the tools needed to load, analyze, and visualize the data. Next we loaded the original data files gathered from the survey information. We included information from the following sections related to education, agriculture, and community attributes.

At the individual and household level:

- Section 0a: General Survey Information
- Section 2a: General Education
- Section 2c: Literacy Education
- Section 8b: Agricultural Plot Details
- Aggregate 2: Agricultural income and farm equipment depreciation

At the community level:

- Community Section 2: Economy and Infrastructure
- Community Section 3: Education
- Community Section 5b: Agriculture

Step 2: Sort Survey Information

While the data originally comes with all variables numerically entered, we found it more useful to use the category labels such as "urban" and "rural" instead of 1 and 2. For this section we focused on the ecological zone (ez), urban vs. rural (loc2).

Step 3: Sort Agriculture data

Multiple data sets account for land areas in Ghana related to households. We used the data set from Section 8 part B of the GLSS 4 questionnaire which includes all farm lands owned or operated by household members during the prior 12 months. The main challenge with the agricultural data was converting all plots of land to a common unit. We decided to keep ropes since it was the smallest land unit commonly found in Ghana. Conveniently, poles and acres are both equivalent to nine ropes.

Step 4: Sort General Education data

Section 2 part A of the GLSS 4 questionnaire includes general information on an individual's education. This section captures the highest level obtained by an individual which is what we were most interested in. Since the data collected did not include a way to examine “years of education”, we grouped education into four levels: no education, kinder, primary, and secondary or above, from here we looked at the highest level of education completed for a household.

Step 5: Sort Literacy Education data

Section 2 Part A of the GLSS 4 questionnaire also included general information on the literacy, reading, and arithmetic abilities of respondents. The data provided us with yes/no answers of reading/writing English and Ghanian, and another yes/no answers regarding arithmetic.English is the the official language of Ghana, so respondents that spoke or wrote English and Ghanian were given a 3, English only a 2, Ghanian only a 1, and neither a 0. We broke this up into 2 variables for reading and writing and took whichever individual had the maximum level in the household to represent the household.

Step 6: Sort Income, Expenses, and Profit data

The income and expense was reported in the agg2 section of the GLSS 4 questionnaire. The data includes households with agricultural profit from crops, expenses from crops, agricultural profit from crops combined with aquaculture and livestock, expenses from crops combined with aquaculture and livestock, and depreciation. First, the households who don’t have any agricultural profits and or expenses are removed from the data by removing all households with a 0 value in all of the profit, expense and depreciation categories of the data set. Once those are removed, a new column is created that takes the profits less expenses from crops combined with aquaculture and livestock and subtracts out depreciation. This is a new column that shows the household’s profit. The next step was to delete the scientific notation and divide total household profit by the number of ropes in the household. This leaves you with profit_per_rope.

Step 7: Sort Community data

The community data was reported in it’s own section of the GLSS 4 questionnaire. with data sets with information pertaining to community demographics, economy, infrastructure, education, agriculture, and health. We primarily drew information from the economy, infrastructure, education, and agriculture datasets. Almost all of these community characteristics were coded with a 1 for said characteristic existing in the community and 0 otherwise. We were

able to group the communities by area enumeration number, because communities within the number were walking distance from one another.

Step 8: Join All data sets

All data sets were joined on the profits data because it allowed us to exclude households who did report any agricultural income because they would be irrelevant to the topic. The data was joined by the “clust” or area enumeration number shown in the documentation.

Step 9: Outlier removal

From our sample, profit per rope averaged 91,650 Cedi with a median value of 33,708 Cedi. This shows there are a few higher profit-per-area plots of land which bring the average above the median value. After excluding only the 20 most extreme values (positive and negative for 40 outliers total), the mean profit-per-area decreased to 76,616 and median stayed at 33,708. Clearly just those 40 extreme values removed had a significant effect on the mean, although the data maintained the original right skew.

The effect was more pronounced in the changes to the linear model, as some variables that were significant before removing the outliers were no longer significant after removing the 40 most extreme values. We considered putting the outliers back into the data, because the log model showed that even with the outliers removed there was still significance in the forest ecological zone, having a primary school and community coop.

In making the decision on if outliers should ultimately be removed we considered the fact that it didn’t appear that any of the outliers were obviously inaccurate. Next, we considered how the outlier changed the results. With the outlier in the linear model the forest, rural and coastal ecological zones were significant to the level of household profit per rope. There was also significance in the level of household profit per rope when a community has a road, primary school, secondary school and community coop. To be conservative in our analysis we decided to drop the outliers. This resulted in these variables becoming insignificant in the linear model.However, the forest ecological zone, primary school and community coop remained significant in the log model, even when the outliers were removed.

Interestingly, with the outliers included in the linear model the maximum educational level was not significant to household profit per rope. However, once the outliers were removed the maximum education per household became significant. Whereas, in the log model, with and without the outliers the maximum level of education was significant. Overall, there were more variables that became insignificant in the linear model with the outliers removed. This led to us deciding to remove the 40 most extreme values from both the linear and log data.

Step 10: Data Analysis

Part of data analysis is determining outliers within the data. Once these were removed, we were able to look at a summary of our data and begin identifying which variables were statistically significant in the linear models. We started with all variables included and reduced to

those statistically significant at the 1% level. We also looked at the correlation between variables whether or not they were statistically significant in the linear models.

Step 11: Graphs

The linear model and log model graphs were made with the statistically significant (at a level of 1%) variables on the x axis. The linear models were used to create bar charts that showed the number of households at each of the levels of maximum education.There was also a bar chart made of the number of households that had access to sharecropping. The normal QQ and residuals vs. fitted for the linear and log models were also included based on the relationships between the data. We also included a box plot visually showing the relationship between education level and profits per rope.

Results & Discussion

Based on the results of our analysis we can now interpret if our original 3 hypotheses are correct.

H1.

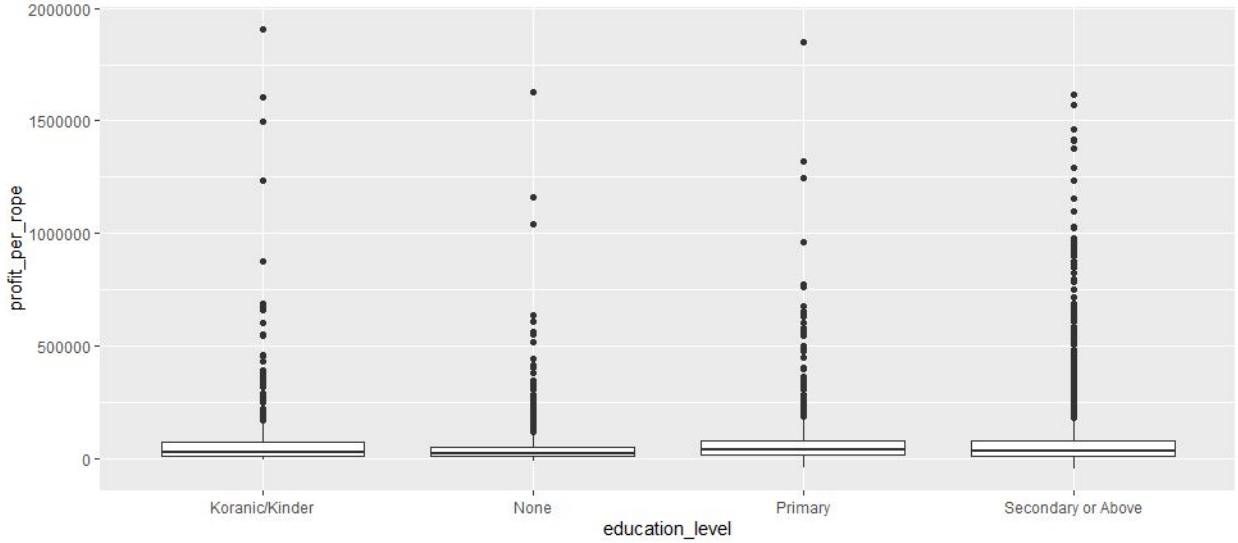
Null Hypothesis: The effect of higher education on household profits per rope isn't statistically significant at a 1% level of confidence.

Alternative hypothesis: The effect of higher education on household profits per rope is statistically significant at a 1% level of confidence.

From our analysis profit per rope in Cedi has a positive relationship with education level, but it is not conclusive. It is clear that households with no education typically see less agricultural profits, but after the koranic or kindergarten level of education is completed, the profits will be similar for the additional educational levels. It is important to note that of the three ecological zones in Ghana, the most educated was the forest zone, this zone did show the greatest median profit per rope. Our logarithmic model showed that there is a positive and statistically significant relationship with profit per rope, so we should not reject this null hypothesis, but there is reason to be skeptical due to the extremely large range in profits. The results of our log model show that an additional educational level leads to a 30% increase in profit per rope, but with a range of profits this large it is difficult to interpret what a 30% increase in profit actually does. It would be smart to review these model results if we looked at a specific income range with more variability and possibly more detail in educational levels.

Additionally, literacy levels were not statistically significant when put into the linear or the log model.

For maximum education level per household we reject the null hypothesis. However, for maximum literacy per household we fail to reject the null hypothesis.



H2.

Null Hypothesis: The effect of community trade on profits per rope isn't statistically significant at a 1% level of confidence.

Alternative Hypothesis:Households with access to community trade in agricultural goods will have a statistically significant increase in profits per rope at a 1% level of confidence.

The analysis on how community trade in agriculture impacts profit per rope found that for those households involved in sharecropping there is a 8.6% increase in profit per rope. There is also an increase in the effect when the household is located in the forest ecological zone. Farm mutual aid also showed positive impacts on profit per rope at an increase of 5% for households that are engaged in this activity. However, while sharecropping is found to be statistically significant at the 1% level, farm mutual aid isn't statistically significant. Being involved in a community coop had little impact on profits per rope with an increase of only .2% for households involved in this type of trade.

This leads to mixed results on if community trade in agriculture leads to higher profits. While all of the trade types (sharecropping,farm mutual aid and community coop) did show a positive increase in profits, only trade through sharecropping is statistically significant. For share cropping we reject the null hypothesis and for community coop and farm mutual aid we fail to reject the null hypothesis.

H3

Null Hypothesis: The effect of a favorable eco system on profits per rope isn't statistically significant at a 1% level of confidence.

Alternative Hypothesis:The effect of a favorable eco system on profits per rope is statistically significant at a 1% level of confidence.

> The three ecosystems that were provided in the data from the survey were "Coastal", "Savannah" and "Forest". From our initial statistics it seemed that coastal ecosystems yielded the greatest profits of all 3 of these, but after removing the 40 most extreme values it was clear that this category was heavily right skewed due to some of the largest instances of profit per rope. For example the mean profit for coastal households was 122,642 Cedi, but the median 21,603 was over 100,000 Cedi less. After removing 20 of the largest values this mean profit decreased to 70680, while households in forests had a mean of 8097. The median profit per rope from these forest ecosystems was also 42906. Forest ecosystems are the safest ecosystem in Ghana to realize consistent agricultural profits for long-term revenue. These forested communities also boast the greatest average maximum level of education among households in all of these ecological zones.

Summary Statistics

Prior to removing the outliers talked about above:

Entire dataset
Household mean: 91,650
Household median: 33,708
Min: -646,912
Max: 7,031,725
Mean, maximum household education level: 2.05

By Ecological Zone:
Coastal households (20% of dataset):
- Mean profit: 122642
- Median profit: 21603
- Max: 7,031,725 (largest in entire dataset)
- Min: -58228
- Mean education level: 1.93

Forest households (53% of dataset):
- Mean profit: 89012
- Median profit: 43023
- Max: 3,236,801
- Min: -79776

- Mean max education level: 2.34

Savannah households (27% of dataset)
- Mean profit: 74015
- Median profit: 26545
- Max: 2,352,761
- Min: -646912 (smallest value in dataset)
- Mean max education level: 1.58

<u>Urban vs Rural:</u>
Urban Households (16% of dataset):
- Mean profit: 67579
- Median profit: 21134
- Max: 3236801
- Min: -193930 (smallest value in dataset)
- Mean max education level: 2.25

Rural Households (83% of dataset):
- Mean profit: 96442
- Median profit: 35941
- Max: 7031725
- Min: -646912 (smallest value in dataset)
- Mean max education level: 2.01

<u>Education Levels:</u>
No education (18% of dataset):
- Mean profit: 74033
- Median profit: 24738
- Max: 3729090
- Min: -193930 (smallest value in dataset)

Koranic/Kindergarten (11.4% of dataset)
- Mean profit: 100680
- Median profit: 30178
- Max: 3484467
- Min: -262471

Primary (15% of dataset):
- Mean profit: 107756
- Median profit: 38636
- Max: 7031725
- Min: --646912

Secondary or above (54% of dataset)
- Mean profit: 91164
- Median profit: 37055
- Max: 4723940
- Min: -58228

After Removing Outliers:

Entire dataset:

- Household mean: 76,616
- Household median: 33708
- Min: -42821
- Max: 1906764

By Ecological Zone:

Coastal households (20% of dataset):

- Mean profit: 70680
- Median profit: 21449
- Max: 1906746
- Min: -38299

Forest households (53% of dataset):

- Mean profit: 80097
- Median profit: 42906
- Max: 137741
- Min: -4281

Savannah households (27% of dataset)

- Mean profit: 74102
- Median profit: 27083
- Max: 1630514
- Min: -36655 (smallest value in dataset)

Education Levels:

No education:

- Mean profit: 74033
- Median profit: 24738
- Max: 3729090
- Min: -193930

Koranic/Kindergarten: (no change after removing outliers)

- Mean profit: 100680
- Median profit: 30178
- Max: 3484467
- Min: -262471

Primary:

- Mean profit: 80656
- Median profit: 38636
- Max: 1848194
- Min: --38299

Secondary or above:

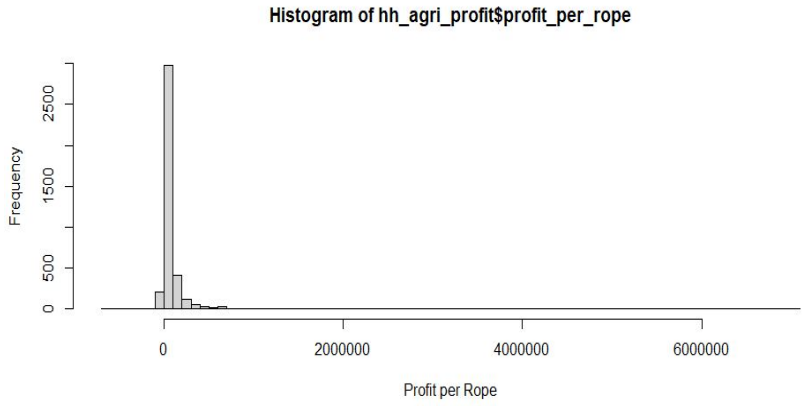
- Mean profit: 81392

- Median profit: 37099
- Max: 1617234
- Min: -42821

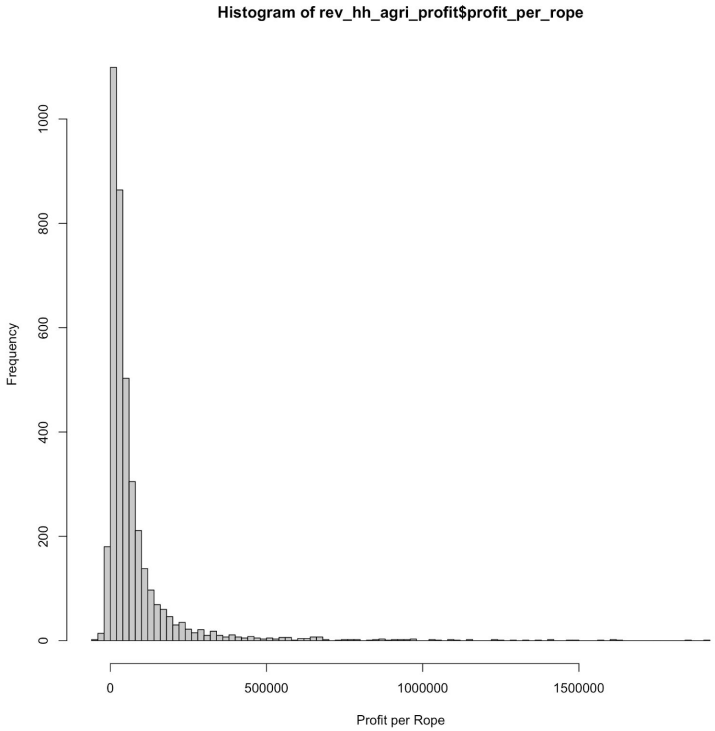
Explanation of summary statistics:

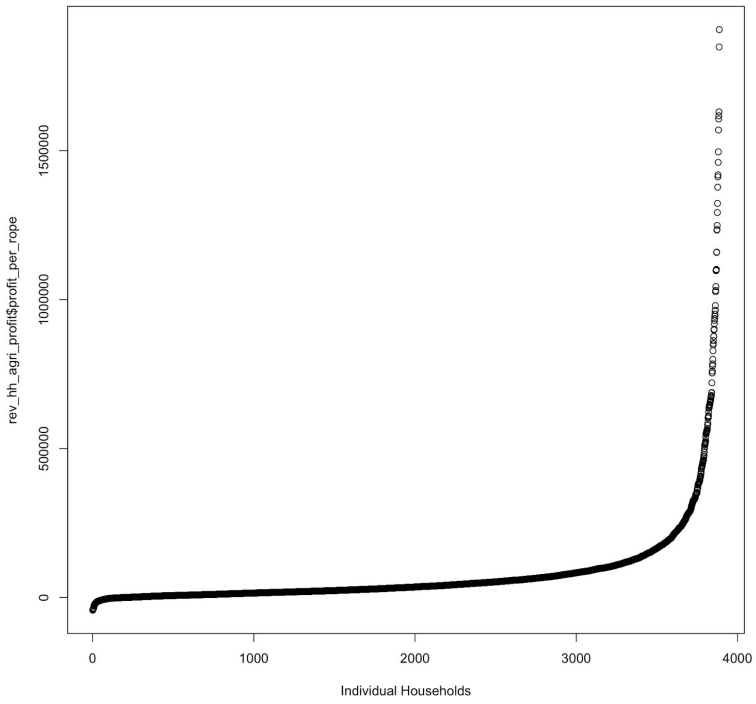
From this information we would assume that forest ecological zones realize the greatest median profits which is more relevant than the mean in this situation due to the skew of data. After removing outliers households in the forest zones actually showed the greatest median and mean of all ecological zones. This likely occurs because forested ecological zones receive more rainfall and possibly because forests showed the highest max level of education for each household.. Rural households showed notable greater profit per rope than urban households which is expected. It is also important to note the average and median profit per rope for the koranic/kindergarten education level through the secondary and above level hovered between a similar range. This may be showing that after the first level of education where basic reading, writing, and arithmetic are taught, additional education is not vital to realizing agricultural profits.

Distribution of profit per rope before removing 40 outliers.



Distribution of Profit per rope after removing 40 outliers:



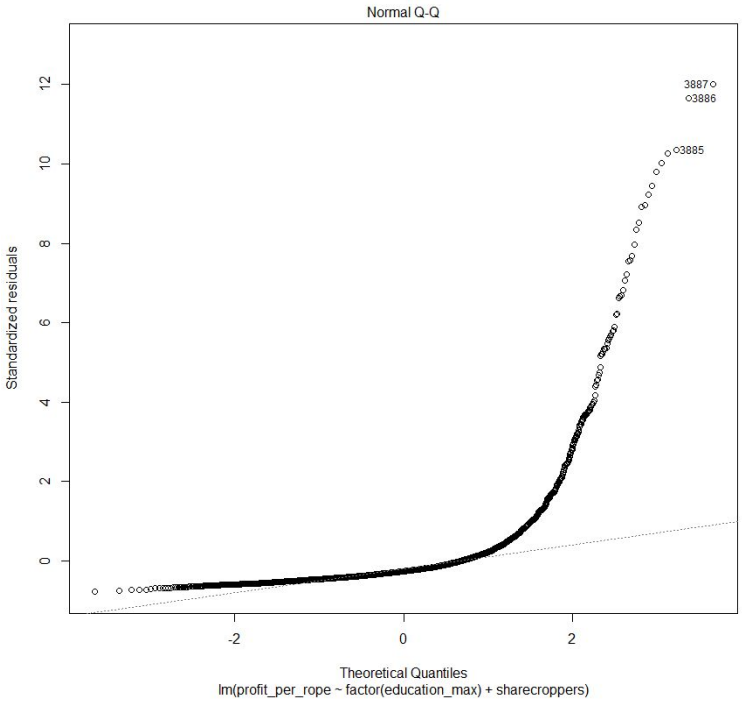


Linear Model Analysis:

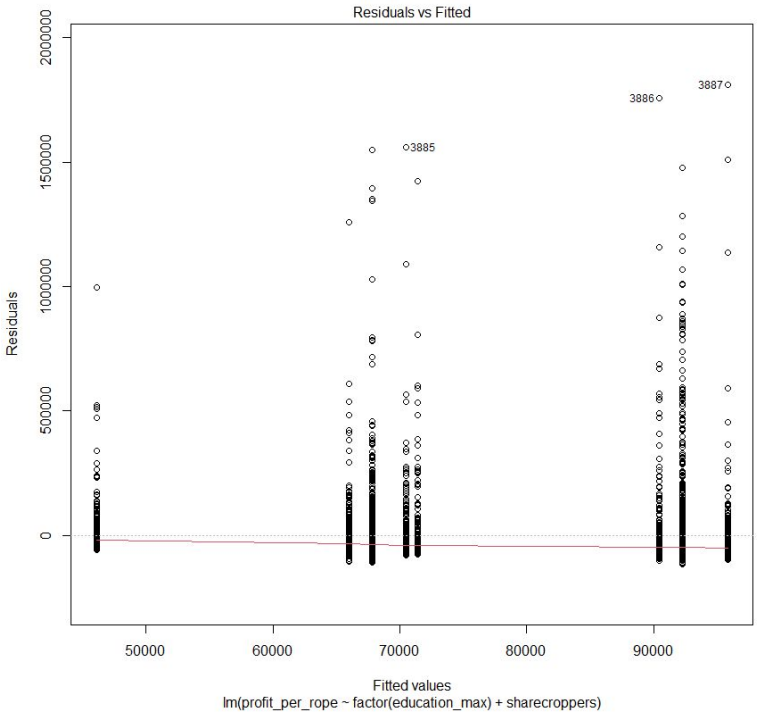
The linear model for profit per rope includes two statistically significant variables at the 1% level: the maximum education of a household and if the household participated in sharecropping.

```
Call:
lm(formula = profit_per_rope ~ factor(education_max) + sharecroppers,
    data = rev_hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-115568  -61799  -39104    -416 1810923

Coefficients:
                     Estimate Std. Error t value           Pr(>|t|)
(Intercept)             46070       5906   7.800 0.0000000000000079 ***
factor(education_max)1  25340       9091   2.787           0.005339 **
factor(education_max)2  19952       8338   2.393           0.016759 *
factor(education_max)3  21741       6534   3.327           0.000885 ***
sharecroppers           24432       4894   4.992 0.0000006246994798 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 150900 on 3882 degrees of freedom
Multiple R-squared:  0.01079,   Adjusted R-squared:  0.009773
F-statistic: 10.59 on 4 and 3882 DF,  p-value: 0.00000001565
```

Normal Q-Q plot:



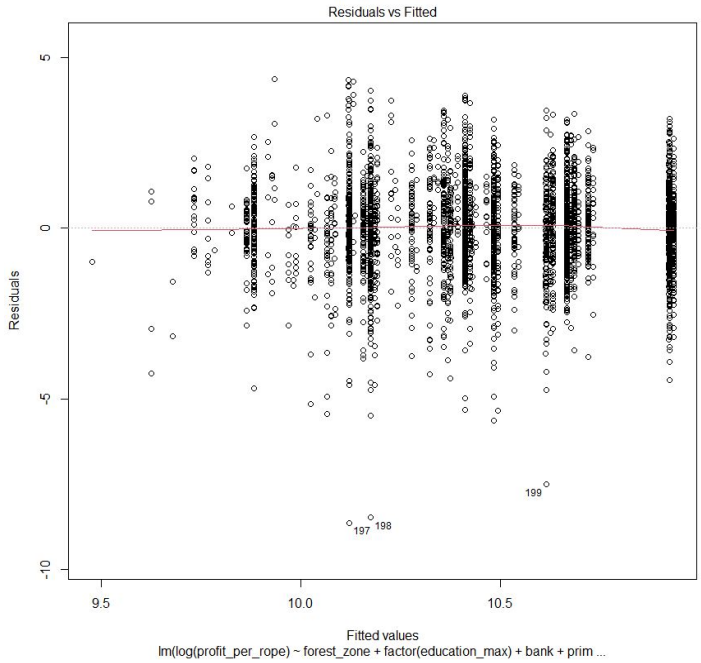
Residuals vs. Fitted:



We found a better linear model when analyzing how different variables affect the percentage change in profit per rope. Below is the statistical summary of this model.

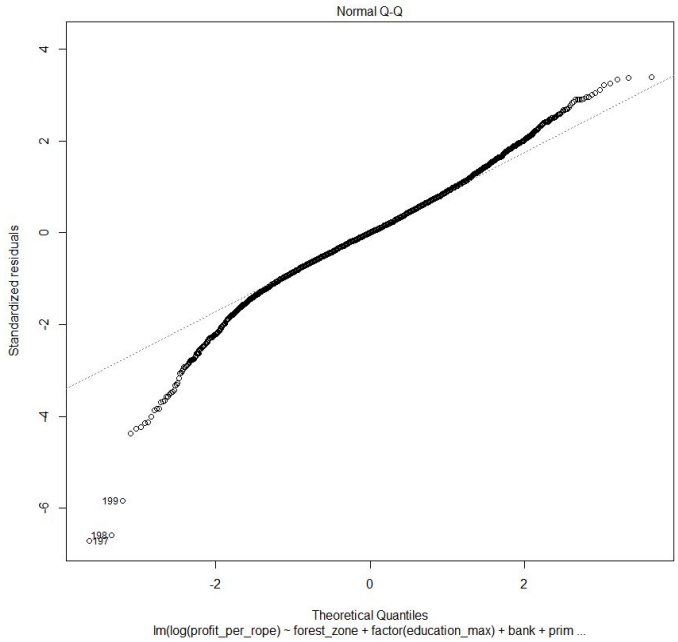
```
Call:
lm(formula = log(profit_per_rope) ~ forest_zone + factor(education_max) +
    bank + prim_school + community_coop + sharecroppers, data = rev_hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-8.6390 -0.7320  0.0069  0.7742  4.3693

Coefficients:
                     Estimate Std. Error t value            Pr(>|t|)
(Intercept)           9.88284    0.06256 157.984 < 0.0000000000000002 ***
forest_zone           0.30813    0.04999   6.163       0.000000000788 ***
factor(education_max)1  0.29099    0.07968   3.652             0.000264 ***
factor(education_max)2  0.30221    0.07341   4.117       0.000039274493 ***
factor(education_max)3  0.29124    0.05890   4.945       0.000000795236 ***
bank                 -0.38733    0.08345  -4.642       0.000003576514 ***
prim_school           0.23778    0.05491   4.331       0.000015260297 ***
community_coop       -0.25635    0.05139  -4.989       0.000000636163 ***
sharecroppers         0.20180    0.05365   3.761             0.000172 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.29 on 3682 degrees of freedom
  (196 observations deleted due to missingness)
Multiple R-squared:  0.05338,   Adjusted R-squared:  0.05132
F-statistic: 25.95 on 8 and 3682 DF,  p-value: < 0.00000000000000022
```

Residuals vs. Fitted:



- This plot of the residual vs fitted values in shows that our log model has much more constant variance than the original linear model.

Normal Q-Q:



Additionally we can note a high level of correlation between the forest ecological zone and other variables. See the correlation tables in the appendix.

Appendix

Correlation Table:

	clust	coastal_zone	forest_zone	savannah_zone	rural	profit_per_rope	education_max
clust	1.000000000	-0.77039984	-0.028671026	0.720967679	-0.170753713	-0.038116611	-0.083678552
coastal_zone	-0.770399838	1.00000000	-0.527532004	-0.302130016	-0.029308274	-0.019383434	-0.048382655
forest_zone	-0.028671026	-0.52753200	1.000000000	-0.650450347	0.004882275	0.024466642	0.260800569
savannah_zone	0.720967679	-0.30213002	-0.650450347	1.000000000	0.020725397	-0.010123856	-0.249387582
rural	-0.170753713	-0.02930827	0.004882275	0.020725397	1.000000000	0.037605988	-0.076787083
profit_per_rope	-0.038116611	-0.01938343	0.024466642	-0.010123856	0.037605988	1.000000000	0.054248641
education_max	-0.083678552	-0.04838265	0.260800569	-0.249387582	-0.076787083	0.054248641	1.000000000
calc_max	-0.093161022	-0.02273211	0.220243873	-0.226812282	-0.053981277	0.034856908	0.761917844
road	-0.130246774	0.01366484	0.027736150	-0.043340241	0.599543836	0.006799709	0.026056186
bank	0.009768333	0.02184633	0.070333531	-0.098453803	0.047045248	-0.031180672	0.050644230
daily_market	-0.035014149	0.05806758	0.051247419	-0.109421745	0.104760123	-0.022473199	0.048937030
periodic_market	0.161649691	-0.14297598	-0.070813835	0.207291836	0.186402037	0.031199737	-0.002344294
prim_school	-0.015245315	-0.09508609	0.019601903	0.063018797	0.642414012	0.025175517	-0.010149733
jss_school	-0.172093816	0.09610570	0.040074081	-0.130893057	0.383514898	0.007485145	0.127438929
agg_ext_center	0.079811087	-0.06515904	0.060224624	-0.009321172	0.182209383	-0.033798367	0.029988707
community_coop	-0.104463732	0.05454922	0.054820032	-0.110284906	0.234129535	0.002598827	0.053618980
irrigated_fields	0.053975392	-0.09411949	0.135911790	-0.068357202	0.116047011	-0.026128413	0.081432700
sharecroppers	-0.248812229	-0.09239723	0.431475121	-0.401549802	0.400136070	0.086998318	0.123967012
farm_mutual_aid	0.097358585	-0.21118650	0.035953520	0.148473142	0.625409828	0.050485816	-0.101885746

	calc_max	road	bank	daily_market	periodic_market	prim_school	jss_school
clust	-0.09316102	-0.130246774	0.009768333	-0.03501415	0.161649691	-0.01524532	-0.172093816
coastal_zone	-0.02273211	0.013664836	0.021846328	0.05806758	-0.142975978	-0.09508609	0.096105698
forest_zone	0.22024387	0.027736150	0.070333531	0.05124742	-0.070813835	0.01960190	0.040074081
savannah_zone	-0.22681228	-0.043340241	-0.098453803	-0.10942175	0.207291836	0.06301880	-0.130893057
rural	-0.05398128	0.599543836	0.047045248	0.10476012	0.186402037	0.64241401	0.383514898
profit_per_rope	0.03485691	0.006799709	-0.031180672	-0.02247320	0.031199737	0.02517552	0.007485145
education_max	0.76191784	0.026056186	0.050644230	0.04893703	-0.002344294	-0.01014973	0.127438929
calc_max	1.00000000	0.028978891	0.040500754	0.05149545	-0.012727844	-0.01264346	0.119451738
road	0.02897889	1.000000000	0.183583878	0.24301576	0.111340837	0.67301938	0.556026185
bank	0.04050075	0.183583878	1.000000000	0.44785989	0.027894811	0.17229776	0.271001815
daily_market	0.05149545	0.243015759	0.447859893	1.00000000	-0.158385949	0.22807597	0.299682900
periodic_market	-0.01272784	0.111340837	0.027894811	-0.15838595	1.000000000	0.25102137	0.257125145
prim_school	-0.01264346	0.673019379	0.172297762	0.22807597	0.251021374	1.00000000	0.579765606
jss_school	0.11945174	0.556026185	0.271001815	0.29968290	0.257125145	0.57976561	1.000000000
agg_ext_center	0.03928643	0.229068385	0.391184922	0.36627202	0.164665352	0.24537527	0.356482749
community_coop	0.04173664	0.241864347	0.132415397	0.14686573	0.166628682	0.27014838	0.296309919
irrigated_fields	0.07325250	0.166513293	0.013286526	0.09893803	-0.007480604	0.15627662	0.203471519
sharecroppers	0.13105684	0.480073755	0.048641127	0.20460423	-0.035044440	0.39385756	0.376421343
farm_mutual_aid	-0.09042690	0.481042163	-0.080686466	-0.01209508	0.250999854	0.55445180	0.226821261

	agg_ext_center	community_coop	irrigated_fields	sharecroppers	farm_mutual_aid
clust	0.079811087	-0.104463732	0.053975392	-0.24881223	0.09735858
coastal_zone	-0.065159037	0.054549219	-0.094119487	-0.09239723	-0.21118650
forest_zone	0.060224624	0.054820032	0.135911790	0.43147512	0.03595352
savannah_zone	-0.009321172	-0.110284906	-0.068357202	-0.40154980	0.14847314
rural	0.182209383	0.234129535	0.116047011	0.40013607	0.62540983
profit_per_rope	-0.033798367	0.002598827	-0.026128413	0.08699832	0.05048582
education_max	0.029988707	0.053618980	0.081432700	0.12396701	-0.10188575
calc_max	0.039286429	0.041736641	0.073252496	0.13105684	-0.09042690
road	0.229068385	0.241864347	0.166513293	0.48007376	0.48104216
bank	0.391184922	0.132415397	0.013286526	0.04864113	-0.08068647
daily_market	0.366272022	0.146865729	0.098938030	0.20460423	-0.01209508
periodic_market	0.164665352	0.166628682	-0.007480604	-0.03504444	0.25099985
prim_school	0.245375268	0.270148385	0.156276619	0.39385756	0.55445180
jss_school	0.356482749	0.296309919	0.203471519	0.37642134	0.22682126
agg_ext_center	1.000000000	0.319543821	0.165734901	0.12850680	0.15967688
community_coop	0.319543821	1.000000000	0.095468117	0.27739381	0.25859269
irrigated_fields	0.165734901	0.095468117	1.000000000	0.18950772	0.09844417
sharecroppers	0.128506801	0.277393810	0.189507720	1.00000000	0.41137508
farm_mutual_aid	0.159676875	0.258592692	0.098444172	0.41137508	1.00000000

Savannah Zone Linear Model w/ All Variables:

```
Call:
lm(formula = profit_per_rope ~ +savannah_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    +road + bank + daily_market + periodic_market + +prim_school +
    jss_school + +agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-809105  -88750  -46435   11211 6674436

Coefficients:
                                  Estimate Std. Error t value      Pr(>|t|)
(Intercept)                        17553.4    51228.2   0.343      0.731879
savannah_zone                      16372.5    12472.9   1.313      0.189380
rural                              65292.3    18177.5   3.592      0.000332 ***
factor(education_max)1             15639.9    18056.7   0.866      0.386459
factor(education_max)2             25063.1    19411.1   1.291      0.196720
factor(education_max)3             16082.4    18859.3   0.853      0.393846
factor(read_max)English and Ghanian   -9783.4    89217.9  -0.110      0.912687
factor(read_max)Ghanian            45735.3    99827.0   0.458      0.646873
factor(read_max)None               31726.8    94939.0   0.334      0.738261
factor(write_max)English and Ghanian  14056.4    85716.6   0.164      0.869750
factor(write_max)Ghanian          -35896.4    98249.5  -0.365      0.714862
factor(write_max)None               7954.4    91813.3   0.087      0.930965
factor(calc_max)1                    452.7    16133.0   0.028      0.977616
road                              -89923.0    15244.2  -5.899 0.00000000397 ***
bank                                8312.6    20511.0   0.405      0.685299
daily_market                      -12195.0    16845.2  -0.724      0.469142
periodic_market                   -16971.1    14686.3  -1.156      0.247926
prim_school                      -100342.6    16814.6  -5.968 0.00000000262 ***
jss_school                         70072.6    12939.3   5.415 0.00000006480 ***
agg_ext_center                    -25970.9    15697.1  -1.655      0.098106 .
community_coop                     25859.8    11470.7   2.254      0.024225 *
irrigated_fields                  -37556.3    19362.3  -1.940      0.052493 .
sharecroppers                      67842.5    12737.9   5.326 0.00000010604 ***
farm_mutual_aid                    45471.8    14740.6   3.085      0.002051 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 283200 on 3903 degrees of freedom
Multiple R-squared:  0.0405,    Adjusted R-squared:  0.03485
F-statistic: 7.163 on 23 and 3903 DF,  p-value: < 0.00000000000000022
```

- The results of this model show that households in the Savannah ecological zone do not see much of an effect from an increase in education
- It is interesting to note that a community with a road has a negative relationship with profits and is statistically significant. Having a primary school also seems to have a negative relationship with profits, but secondary schools show a positive relationship. Consistent with the rest of our data communities that take part in sharecropping remains significant with a positive relationship with profits.

Savannah Zone Log Model w/ All Variables:

```
Call:
lm(formula = log(profit_per_rope) ~ savannah_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    road + bank + daily_market + periodic_market + prim_school +
    jss_school + agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-8.6008 -0.7346 -0.0194  0.7685  5.0224

Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                       9.815185   0.251258  39.064 < 0.0000000000000002 ***
savannah_zone                     0.018874   0.060534   0.312             0.755221
rural                             0.183715   0.088444   2.077             0.037852 *
factor(education_max)1            0.313840   0.086542   3.626             0.000291 ***
factor(education_max)2            0.386482   0.092808   4.164    0.000031940890768 ***
factor(education_max)3            0.422249   0.090329   4.675    0.000003050843006 ***
factor(read_max)English and Ghanian   0.015155   0.421765   0.036             0.971338
factor(read_max)Ghanian           0.242190   0.472787   0.512             0.608500
factor(read_max)None              0.127008   0.449754   0.282             0.777657
factor(write_max)English and Ghanian -0.197718   0.403576  -0.490             0.624223
factor(write_max)Ghanian          -0.140450   0.463961  -0.303             0.762121
factor(write_max)None             -0.023102   0.434104  -0.053             0.957561
factor(calc_max)1                  0.008652   0.077023   0.112             0.910568
road                              -0.290391   0.072421  -4.010    0.000061988628458 ***
bank                              -0.241178   0.100164  -2.408             0.016096 *
daily_market                       0.050546   0.081026   0.624             0.532782
periodic_market                    0.069173   0.069893   0.990             0.322393
prim_school                        0.136671   0.080284   1.702             0.088777 .
jss_school                        -0.102236   0.061764  -1.655             0.097955 .
agg_ext_center                     0.040381   0.075238   0.537             0.591505
community_coop                    -0.249232   0.055347  -4.503    0.000006904152776 ***
irrigated_fields                  -0.163911   0.092658  -1.769             0.076976 .
sharecroppers                      0.463721   0.061558   7.533    0.000000000000062 ***
farm_mutual_aid                    0.133339   0.070987   1.878             0.060412 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.329 on 3687 degrees of freedom
  (216 observations deleted due to missingness)
Multiple R-squared:  0.05335,   Adjusted R-squared:  0.04744
F-statistic: 9.034 on 23 and 3687 DF,  p-value: < 0.00000000000000022
```

- After running a log on profit per_rope this model shows that an increase in educational level does lead to an increase in profit in the Savannah ecological zones.
- Communities with roads continue to hold a negative and significant relationship with profit per rope in this ecological zone.
- Sharecropping retains the positive and significant relationship with profit per rope

Coastal Zone Linear Model w/ All Variables:

```
call:
lm(formula = profit_per_rope ~ +coastal_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    +road + bank + daily_market + periodic_market + +prim_school +
    jss_school + +agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-801029  -87373  -45088   12168 6656696

Coefficients:
                                  Estimate Std. Error t value      Pr(>|t|)
(Intercept)                        19881.9    50866.6   0.391       0.69592
coastal_zone                       38417.5    12322.2   3.118       0.00184 **
rural                              57698.3    18245.2   3.162       0.00158 **
factor(education_max)1             14328.0    17993.7   0.796       0.42592
factor(education_max)2             23636.7    19309.2   1.224       0.22098
factor(education_max)3             16585.7    18761.5   0.884       0.37674
factor(read_max)English and Ghanian -15796.7    89108.0  -0.177       0.85930
factor(read_max)Ghanian            40215.8    99727.7   0.403       0.68678
factor(read_max)None               23477.1    94868.2   0.247       0.80456
factor(write_max)English and Ghanian  13948.3    85621.5   0.163       0.87060
factor(write_max)Ghanian           -36854.3    98146.6  -0.376       0.70731
factor(write_max)None              10008.6    91722.0   0.109       0.91311
factor(calc_max)1                   -257.1    16115.6  -0.016       0.98727
road                               -92457.1    15252.1  -6.062 0.00000000147 ***
bank                                8230.8    20436.0   0.403       0.68715
daily_market                       -12242.3    16811.2  -0.728       0.46652
periodic_market                     -7423.4    14669.8  -0.506       0.61287
prim_school                        -90900.4    16851.9  -5.394 0.00000007297 ***
jss_school                          60437.4    13160.7   4.592 0.00000452283 ***
agg_ext_center                     -21010.5    15746.6  -1.334       0.18219
community_coop                      19681.1    11542.1   1.705       0.08824 .
irrigated_fields                   -31356.3    19440.3  -1.613       0.10684
sharecroppers                       64747.2    11540.6   5.610 0.00000002159 ***
farm_mutual_aid                     57300.8    14581.9   3.930 0.00008657008 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 282900 on 3903 degrees of freedom
Multiple R-squared:  0.04246,   Adjusted R-squared:  0.03682
F-statistic: 7.525 on 23 and 3903 DF,  p-value: < 0.00000000000000022
```

- The results of this linear model has similar results to those mentioned in the Savannah ecological zone. Communities with roads and primary schools maintain the negative and significant relationship, while sharecropping remains positive and significant at all levels. Although in these coastal zones having a mutual aid system between farms holds a positive and significant relationship with profit per rope.

Coastal Zone Log Model w/ All Variables:

```
Call:
lm(formula = log(profit_per_rope) ~ coastal_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    road + bank + daily_market + periodic_market + prim_school +
    jss_school + agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-8.3679 -0.7371 -0.0309  0.7672  5.2257

Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                       9.870753   0.248301  39.753 < 0.0000000000000002 ***
coastal_zone                     -0.357485   0.060607  -5.898   0.0000000040007128 ***
rural                             0.242777   0.088549   2.742             0.006141 **
factor(education_max)1            0.310376   0.085887   3.614             0.000306 ***
factor(education_max)2            0.379452   0.091912   4.128   0.0000373348669850 ***
factor(education_max)3            0.393028   0.089450   4.394   0.0000114500036755 ***
factor(read_max)English and Ghanian   0.039772   0.419661   0.095             0.924502
factor(read_max)Ghanian           0.276177   0.470586   0.587             0.557321
factor(read_max)None              0.189628   0.447764   0.424             0.671955
factor(write_max)English and Ghanian -0.177095   0.401658  -0.441             0.659304
factor(write_max)Ghanian         -0.140807   0.461772  -0.305             0.760438
factor(write_max)None            -0.038894   0.432080  -0.090             0.928280
factor(calc_max)1                 0.014301   0.076658   0.187             0.852021
road                             -0.257952   0.072177  -3.574             0.000356 ***
bank                             -0.267645   0.099509  -2.690             0.007185 **
daily_market                      0.063202   0.080591   0.784             0.432955
periodic_market                   0.018987   0.069482   0.273             0.784665
prim_school                       0.077726   0.080201   0.969             0.332539
jss_school                       -0.036892   0.062383  -0.591             0.554301
agg_ext_center                    0.004174   0.075132   0.056             0.955703
community_coop                   -0.203326   0.055535  -3.661             0.000255 ***
irrigated_fields                 -0.223880   0.092784  -2.413             0.015874 *
sharecroppers                     0.418407   0.055099   7.594   0.0000000000000391 ***
farm_mutual_aid                   0.061507   0.070152   0.877             0.380674
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.322 on 3687 degrees of freedom
  (216 observations deleted due to missingness)
Multiple R-squared:  0.06217,   Adjusted R-squared:  0.05632
F-statistic: 10.63 on 23 and 3687 DF,  p-value: < 0.00000000000000022
```

- The log of profit per rope in a coastal zone shows almost entirely similar results as when we ran a log model on the savannah zones.
- Sharecropping and education level still have positive relationships with profit per rope, but communities with roads or community coops still present a negative relationship.

Forest Zone Linear Model w/ All Variables:

```
Call:
lm(formula = profit_per_rope ~ +forest_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    +road + bank + daily_market + periodic_market + +prim_school +
    jss_school + +agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-814010  -90007  -45018   11980 6648110

Coefficients:
                                  Estimate Std. Error t value       Pr(>|t|)
(Intercept)                        41390.4    50974.5   0.812       0.416852
forest_zone                       -41546.8    10758.5  -3.862       0.000114 ***
rural                              60443.3    18144.1   3.331       0.000872 ***
factor(education_max)1             18599.4    18021.4   1.032       0.302103
factor(education_max)2             29741.5    19380.3   1.535       0.124958
factor(education_max)3             23464.1    18900.5   1.241       0.214512
factor(read_max)English and Ghanian   -9049.2    89047.4  -0.102       0.919062
factor(read_max)Ghanian            43491.6    99653.1   0.436       0.662549
factor(read_max)None               25266.8    94787.1   0.267       0.789819
factor(write_max)English and Ghanian   9581.9    85577.6   0.112       0.910855
factor(write_max)Ghanian          -34571.6    98083.2  -0.352       0.724504
factor(write_max)None              10345.2    91660.8   0.113       0.910145
factor(calc_max)1                    446.6    16104.5   0.028       0.977879
road                              -95226.0    15287.1  -6.229 0.000000000519 ***
bank                               13825.8    20507.7   0.674       0.500239
daily_market                      -15122.7    16828.4  -0.899       0.368898
periodic_market                   -14083.4    14501.6  -0.971       0.331527
prim_school                       -96374.5    16693.2  -5.773 0.000000008379 ***
jss_school                         63095.4    12961.9   4.868 0.000001173431 ***
agg_ext_center                    -21627.6    15701.3  -1.377       0.168455
community_coop                     22105.6    11440.2   1.932       0.053398 .
irrigated_fields                  -31169.2    19397.5  -1.607       0.108164
sharecroppers                      83507.6    12902.2   6.472 0.000000000108 ***
farm_mutual_aid                    47273.1    14378.8   3.288       0.001019 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 282700 on 3903 degrees of freedom
Multiple R-squared:  0.04373,   Adjusted R-squared:  0.0381
F-statistic: 7.761 on 23 and 3903 DF,  p-value: < 0.00000000000000022
```

- Similar results to the previous linear models in the different ecological zones.

Forest Zone Log Model w/ All Variables:

```
Call:
lm(formula = log(profit_per_rope) ~ forest_zone + rural + factor(education_max) +
    factor(read_max) + factor(write_max) + factor(calc_max) +
    road + bank + daily_market + periodic_market + prim_school +
    jss_school + agg_ext_center + community_coop + irrigated_fields +
    sharecroppers + farm_mutual_aid, data = hh_agri_profit)

Residuals:
    Min      1Q  Median      3Q     Max
-8.6003 -0.7435 -0.0244  0.7667  5.2048

Coefficients:
                                  Estimate Std. Error t value             Pr(>|t|)
(Intercept)                       9.733817   0.249284  39.047 < 0.0000000000000002 ***
forest_zone                       0.252674   0.052453   4.817          0.000001515 ***
rural                             0.210236   0.088283   2.381              0.017298 *
factor(education_max)1            0.283279   0.086224   3.285              0.001028 **
factor(education_max)2            0.341865   0.092459   3.697              0.000221 ***
factor(education_max)3            0.358428   0.090362   3.967          0.000074297 ***
factor(read_max)English and Ghanian -0.017291  0.420331  -0.041              0.967189
factor(read_max)Ghanian           0.243152   0.471282   0.516              0.605928
factor(read_max)None              0.153249   0.448368   0.342              0.732526
factor(write_max)English and Ghanian -0.157176 0.402352  -0.391              0.696083
factor(write_max)Ghanian          -0.159605  0.462507  -0.345              0.730049
factor(write_max)None             -0.041830  0.432763  -0.097              0.923003
factor(calc_max)1                 0.006666   0.076771   0.087              0.930808
road                              -0.251140  0.072526  -3.463              0.000541 ***
bank                              -0.291366  0.100075  -2.911              0.003619 **
daily_market                      0.074949   0.080837   0.927              0.353908
periodic_market                   0.075072   0.069004   1.088              0.276696
prim_school                       0.128662   0.079676   1.615              0.106436
jss_school                        -0.072572  0.061783  -1.175              0.240219
agg_ext_center                    0.020407   0.075111   0.272              0.785874
community_coop                    -0.232998  0.055163  -4.224          0.000024598 ***
irrigated_fields                  -0.202114  0.092709  -2.180              0.029313 *
sharecroppers                     0.313794   0.062186   5.046          0.000000473 ***
farm_mutual_aid                   0.148944   0.069081   2.156              0.031141 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.324 on 3687 degrees of freedom
  (216 observations deleted due to missingness)
Multiple R-squared:  0.05924,   Adjusted R-squared:  0.05338
F-statistic:  10.1 on 23 and 3687 DF,  p-value: < 0.00000000000000022
```

- Similar results to other log models ran on the three different ecological zones.