R PROJECT

By Ryan Sierra, Meissane Mena and Charlotte Schoonheere

Main question:

What education to choose to succeed as a farmer?

Auxiliary questions:

Does your **level of formation** increase your **revenue** ?
Does having **employees** increase revenue?
Does your level of formation increase **the number** of employees you have?

CONTEXT OF THE PRESENTATION:

- Opportunity to apply our skills in statistics and data analysis in a **real context** that interests us.

- **Personal interest** in agriculture and finance that leaded to a choice of a study to understand how the level of agricultural training can have an impact on the financial management of farms and its overall sucess.

- On a larger scale, this type of research could have a **direct impact** on the agricultural sector and on **farmers' livelihoods.**

PROBLEMATIC OF THE PRESENTATION:

Understanding the impact of agricultural education on the farmers profitability.

→ highlight the **effectiveness of agricultural education** in improving **financial management practices** within the agricultural sector.

This research is **relevant** as it addresses the **practical challenges faced by farmers** in effectively managing their **financial obligations and number of employees**, thereby contributing to the sustainability and economic viability of farms.

DATASET PRESENTATION

Dataset subject: The dataset looks on the relationship between farm operators' financial management skills and their degree of agricultural education.

This structured dataset includes both **category and quantitative characteristics** pertaining to financial aspects, agricultural productivity indicators, and agricultural education.

The dataset originates from a real dataset from the **RICA Data Base**

Pre-processing steps involved **data cleaning** to address missing values, outlier detection and treatment, normalization of financial parameters, and standardization of educational categories for uniform analysis.

The **main characteristics** of the dataset include variables such as **the level of agricultural education** (primary agricultural training, secondary agricultural training long cycle...), **financial metrics** (revenue, equipment maintenance), and **number of employees**(from: 1 to 2, 3 to 4, 5 to 6,...).

The dataset offers insights into the **financial performance, operational activities and emplyment of a farm**

Difficulties can include: outliers, duplicates and errors in data entry or recording. Data cleaning avoids these problems but it takes time.

Most **important variables** in the dataset for our study: Level of Agricultural Training, Financial Metrics

Most relevant columns we choose to work with based on which ones where good indicators for our statistical analysis

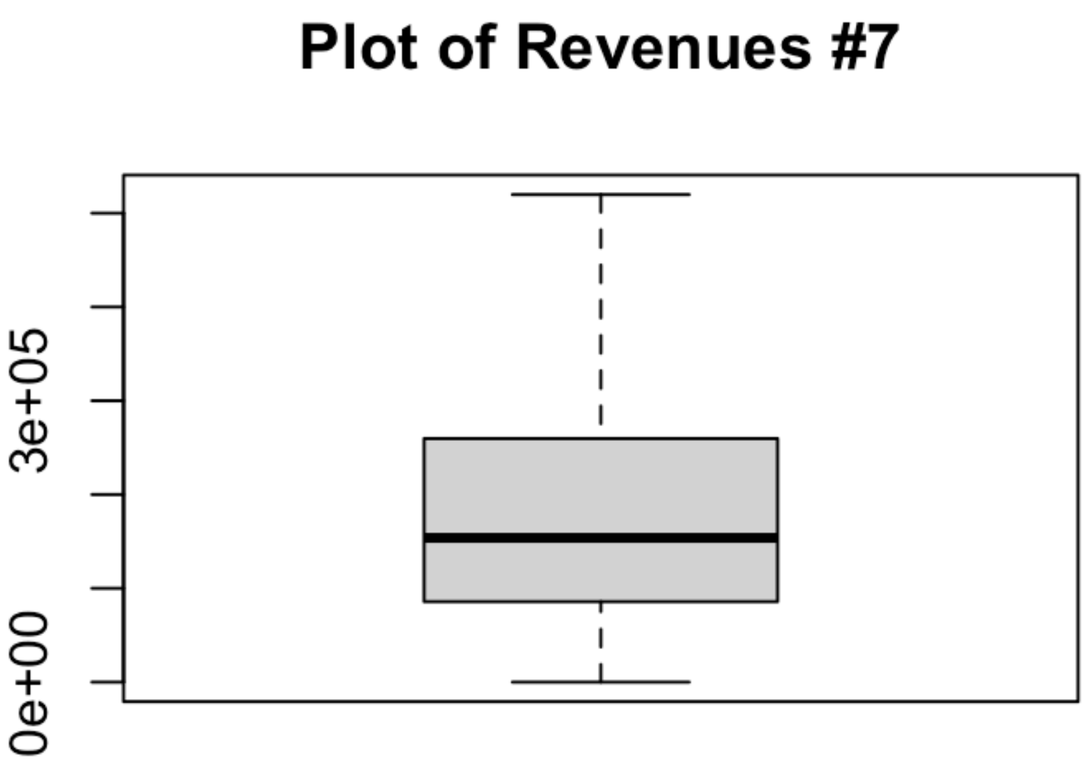
(example: to get relationship between number of employees and financial results):

Selected columns: revenues, number of permanent salaries, equipment maintenance expenses (in euros) and code agricultural training of the farm manager

code	EN		Yes
CHAFF	Revenues (in euros)		Y
EFF10	Number of permanent salaried employees other than the operating manager (portion)		Y
ENTMT	Equipment maintenance expenses (in euros)		Y
FOAGR	Code Agricultural training of the farm manager		Y

SHAPiRO TEST FOR CHAFF

shapiro=Anderson-Darling

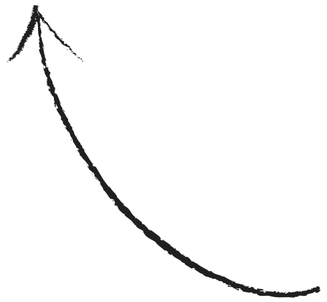



```
        Anderson-Darling normality test

data:  filter2_rev$CHAFF
A = 161.49, p-value < 2.2e-16

> boxplot(filter2_rev)
> boxplot(filter7_rev)
> ad_result <- ad.test(filter7_rev$CHAFF)
>
> # Print the result of the Anderson-Darling test
> print(ad_result)

        Anderson-Darling normality test

data:  filter7_rev$CHAFF
A = 142.26, p-value < 2.2e-16
```

To do the Shapito test we had to remove the outliers: here we filtered seven times because there were a lot of outlets. The aim of this test is to test the normality of the data from the revenues before going any further.

SHAPIRO TEST FOR ENTMT

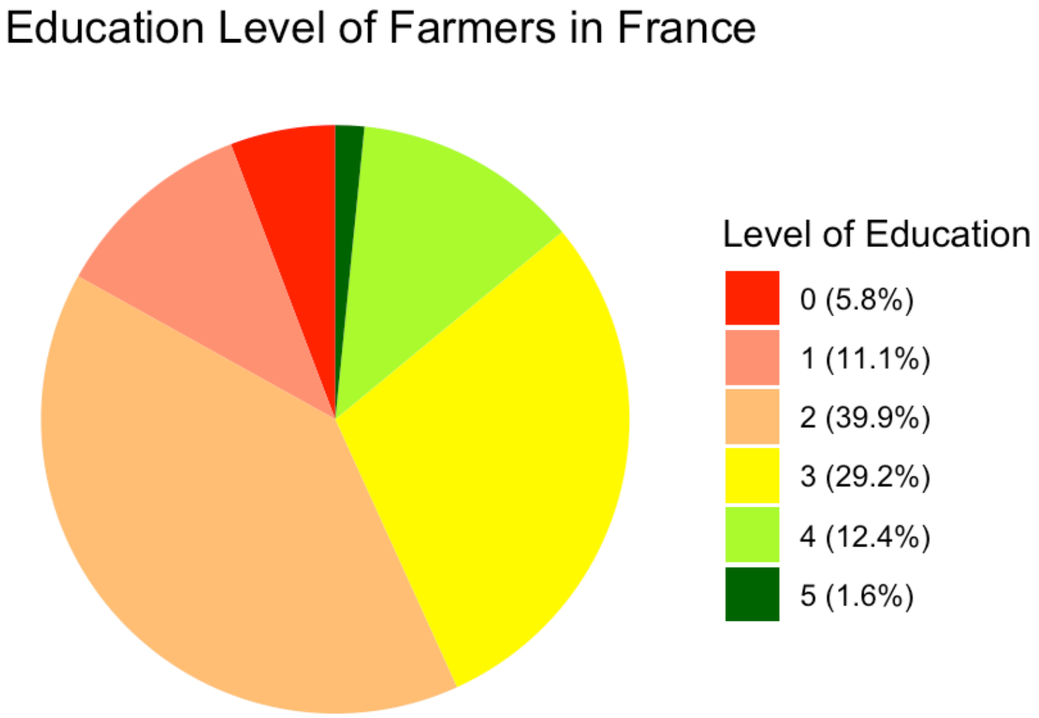
```
>
> ad_result <- ad.test(filter_Emain$ENTMT)
>
> # Print the result of the Anderson-Darling test
> print(ad_result)

        Anderson-Darling normality test

data:  filter_Emain$ENTMT
A = 168.53, p-value < 2.2e-16
```



Data is not normal

To do the Shapito test you had to remove the outliers: here we filtered twice because there were a lot of outliers. The aim is to test the normality of the equipment maintenance expenses before continuing with the tests.

DETAILS OF THE DIFFERENT LEVELS OF FORMATION



0. No adult training or internship of less than 120 hours
1. Primary agricultural training (BAA, CAPA, 200 or 320-hour type internship)
2. Secondary short cycle agricultural training (BEA, BPA, BEPA)
3. Secondary long cycle agricultural training (BTA, BATA, Bac series D')
4. Short higher agricultural education (BTSA, DUT)
5. Long higher agricultural education (Engineer, Bachelor's degree, Master's degree)

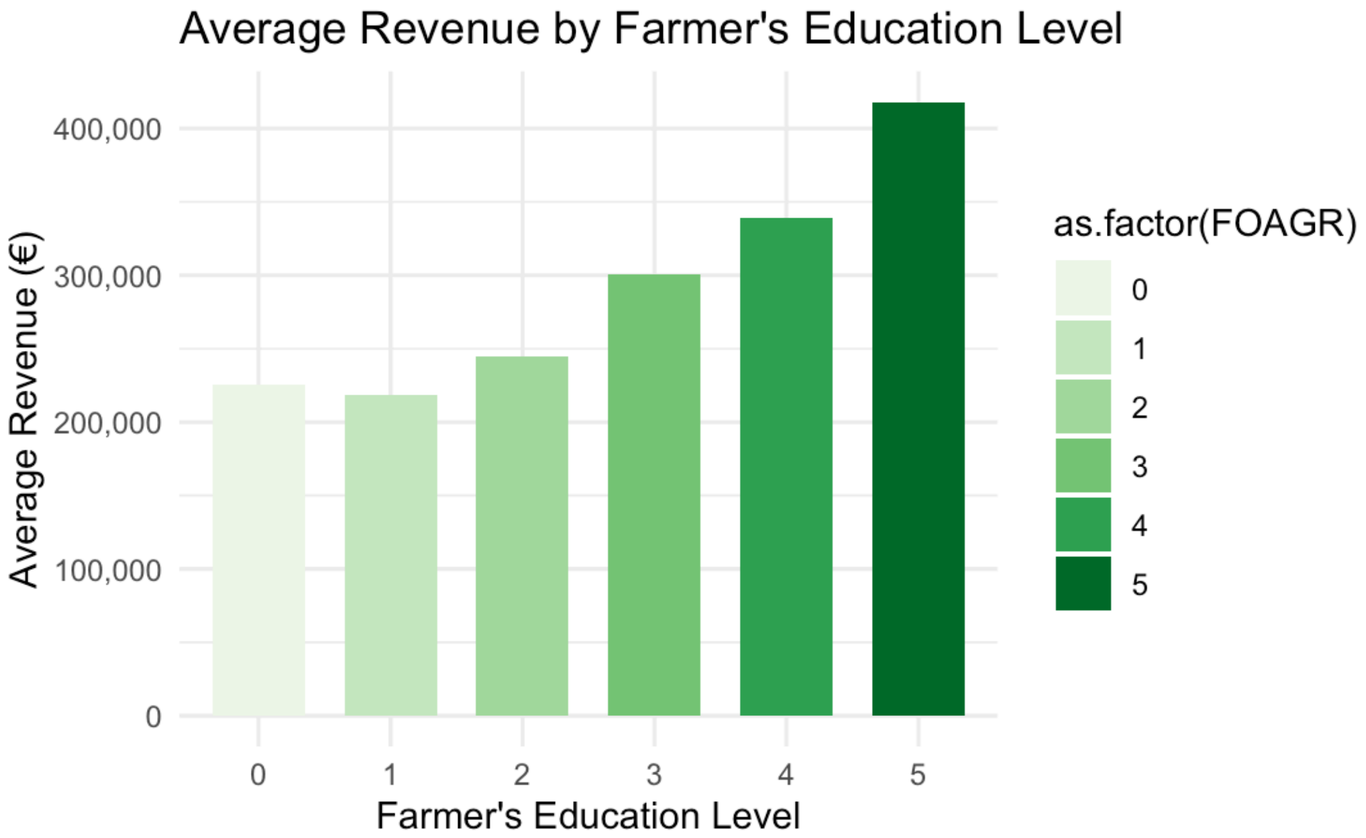
We can notice that in France, the most followed education level is **level 2, at around 40%**, followed by **level 3 at approximately 30%**. Few farmers pursue education up to level 5 (1.6%).

Interpretation/ what was the importance of perfoming this test:

Farmers with higher education levels, such as level 5 (long higher agricultural education), may have **better career advancement opportunities** compared to those with lower education levels. They are more likely to pursue managerial positions, research roles, or specialized agricultural practices.
They are also likely to be **more innovative** and **adaptable to changes in the agricultural industry.**
They may be better equipped to implement **sustainable farming practices,** adopt **new technologies**, and respond to **market dynamics effectively**.

Overall, the education level of farmers plays a significant role in shaping their **career trajectories**, influencing their skills, knowledge, and opportunities for advancement in the agricultural sector.

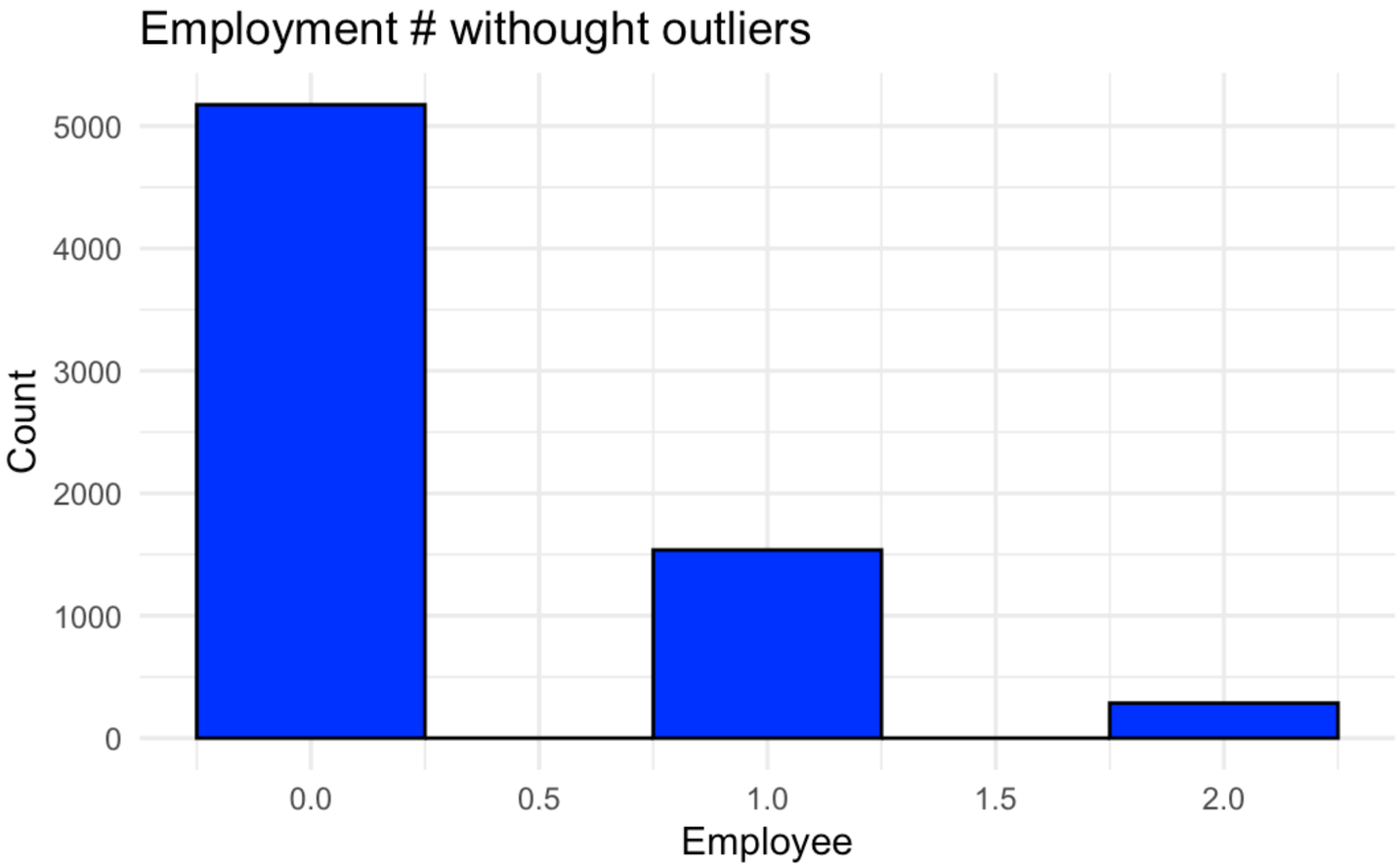
DOES YOUR LEVEL OF FORMATION INCREASE YOUR REVENUE?

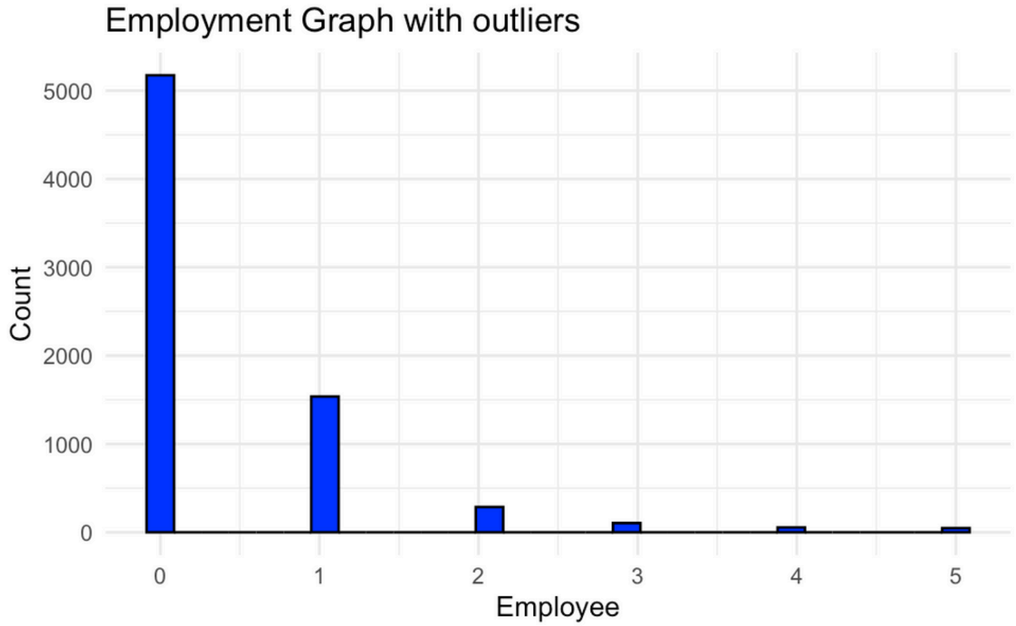


This graph illustrates the relationship between the farmer's **education level and their average revenue**. Indeed, we can observe that as the education level increases, the revenue also increases, with a significant revenue gap between level 0 and level 5. It's also noteworthy that there is no difference, or even a decrease in revenue, between level 0 and 1.

We saw in the first graph that higher education levels lead to greater agricultural advancement opportunities, which is also reflected in financial aspects (revenue).

HISTOGRAM OF EMPLOYMENT





The graph with outliers not being precise enough, we chose to redo it by removing the outliers.

X0: O employees
X1: between 1 and 2 employees
X2: between 2 and 3 employees

Majority of farmers have O employment: about 3 times more than the ones who have between 1 and 2 employees. The farmers are even fewer to have between 2 and 3 employees.

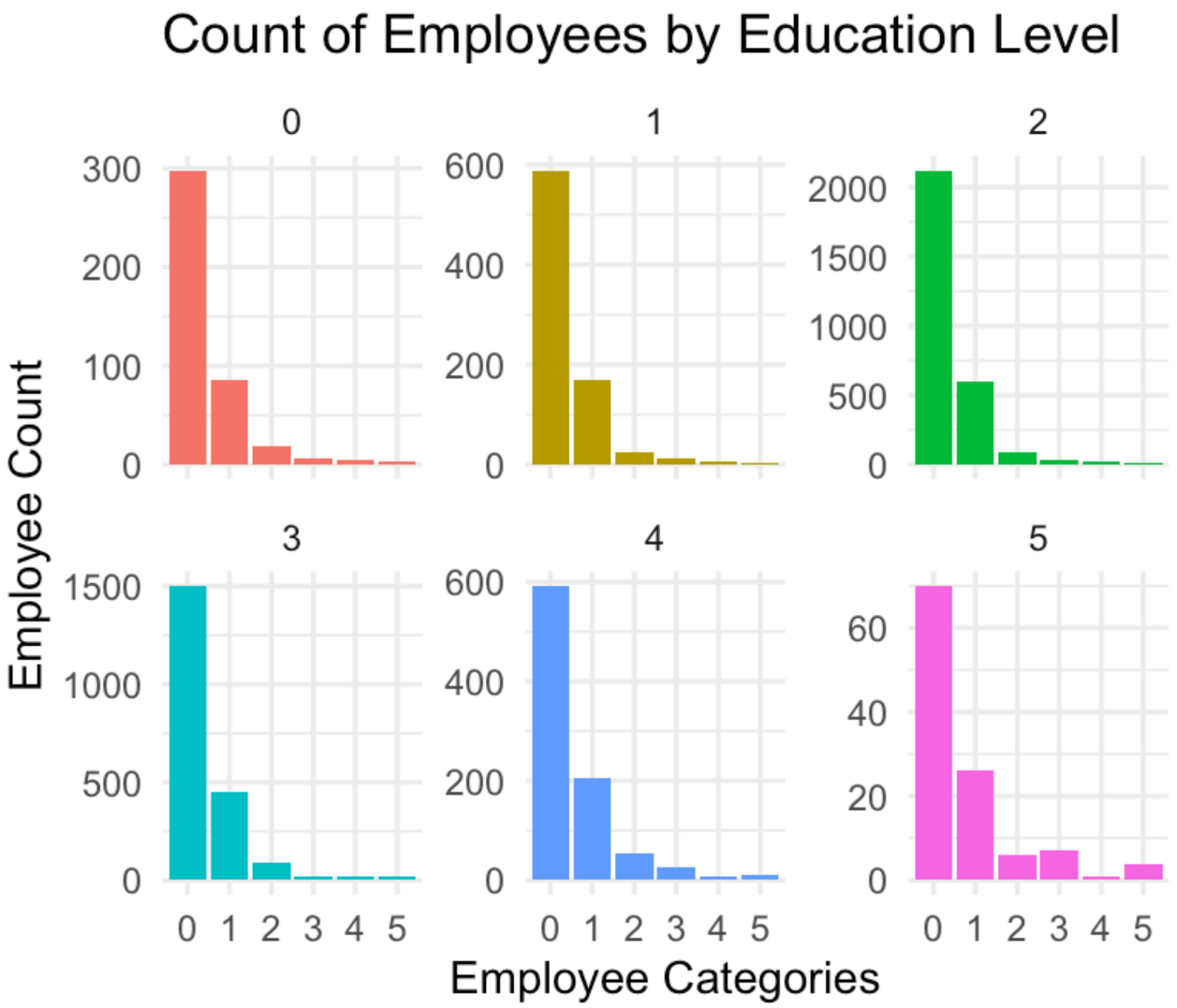
INTERPRETATION/IMPORTANCE OF THIS GRPAH

Farmers with a **limited number of employees** usually run a **small to medium-sized farm, and use multi-skilled expertise**. They master various aspects of farming, from production to financial management.

Investing in **specialized training and educational programs** can help them stay competitive and adapt to agricultural changes. Innovation and adaptation are essential for long-term profitability and sustainability of their farm.

In summary, the shortage of farmers on small to medium-sized farms highlights the **importance of versatility, effective resource managemen**t that is optained thnaks to a **complete education.**

DOES YOUR LEVEL OF EDUCATION INCREASE THE NUMBER OF EMPLOYEES YOU HAVE?



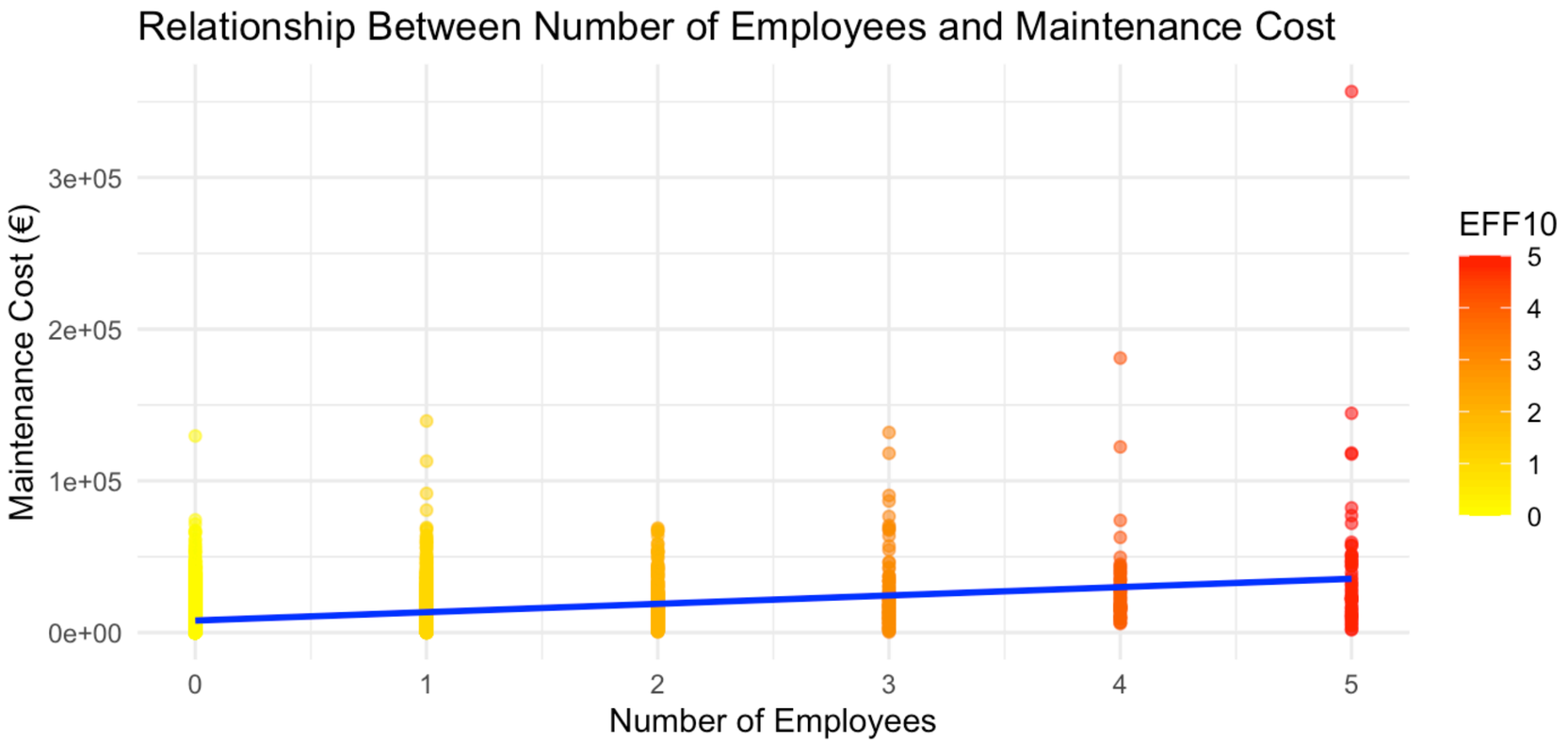
0. No adult training or internship of less than 120 hours
1. Primary agricultural training (BAA, CAPA, 200 or 320-hour type internship)
2. Secondary short cycle agricultural training (BEA, BPA, BEPA)
3. Secondary long cycle agricultural training (BTA, BATA, Bac series D')
4. Short higher agricultural education (BTSA, DUT)
5. Long higher agricultural education (Engineer, Bachelor's degree, Master's degree)

CORRELATION TEST BETWEEN THE EDUCATION LEVEL AND NUMBER OF EMPLOYEES

```
        Spearman's rank correlation rho

data:  RICA$EFF10 and RICA$NumericFOAGR
S = 5.9048e+10, p-value = 1.016e-05
alternative hypothesis: true rho is not equal to 0
sample estimates:
       rho
0.05197855
```

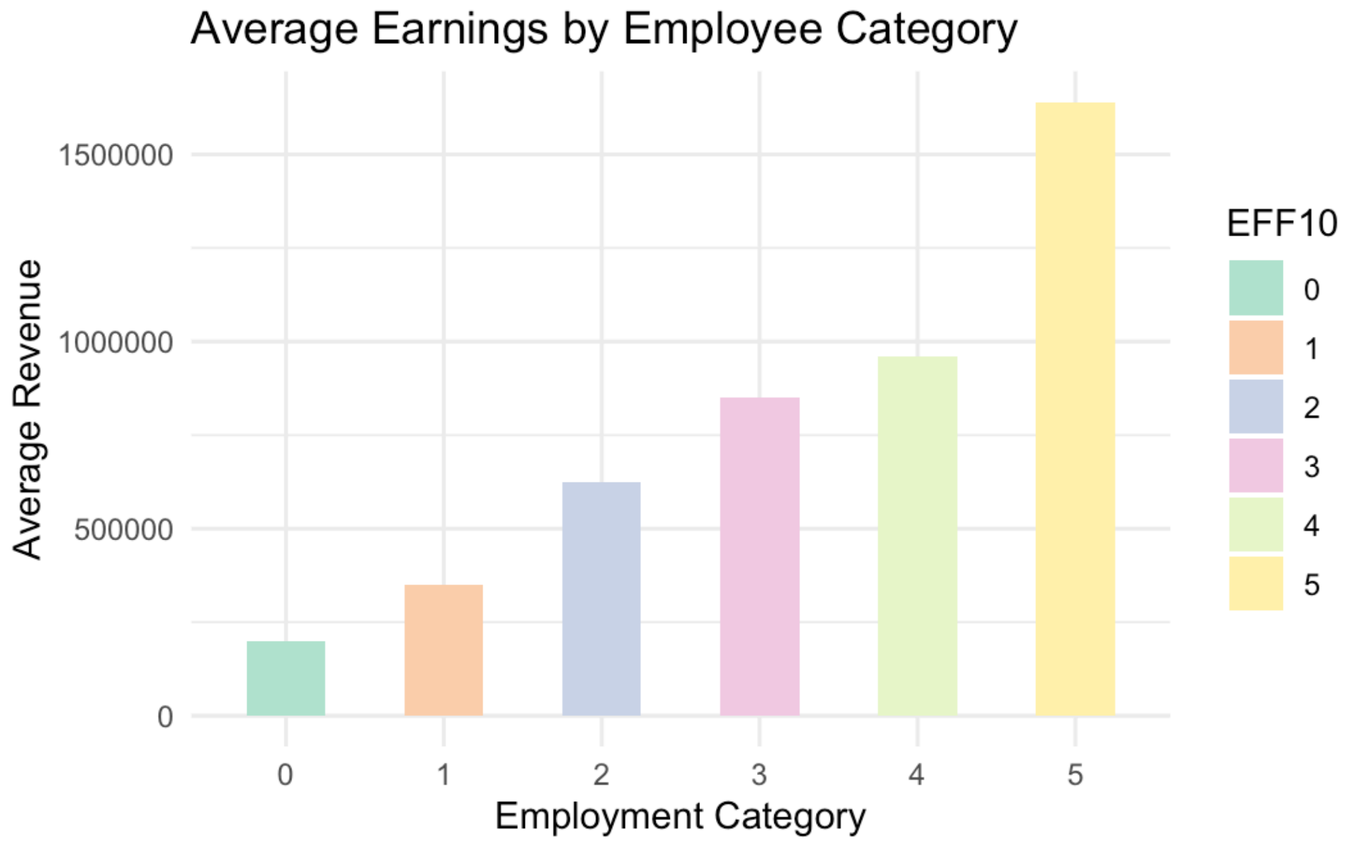
The weak correlation between farmers' education level and employee numbers suggests a limited direct relationship. To better understand dynamics, consider other variables, multivariate analysis, farm size, location, and type of farming. Revisit data collection and processing stages.

RELATIONSHIP BETWEEN NUMBER OF EMPLOYEES AND MAINTENANCE COSTS:



With a larger number of employees, farms may require **additional expenditure** on human resources management, such as s**alaries, benefits, training and recruitment.** This also includes costs associated with regulatory compliance and workplace safety.

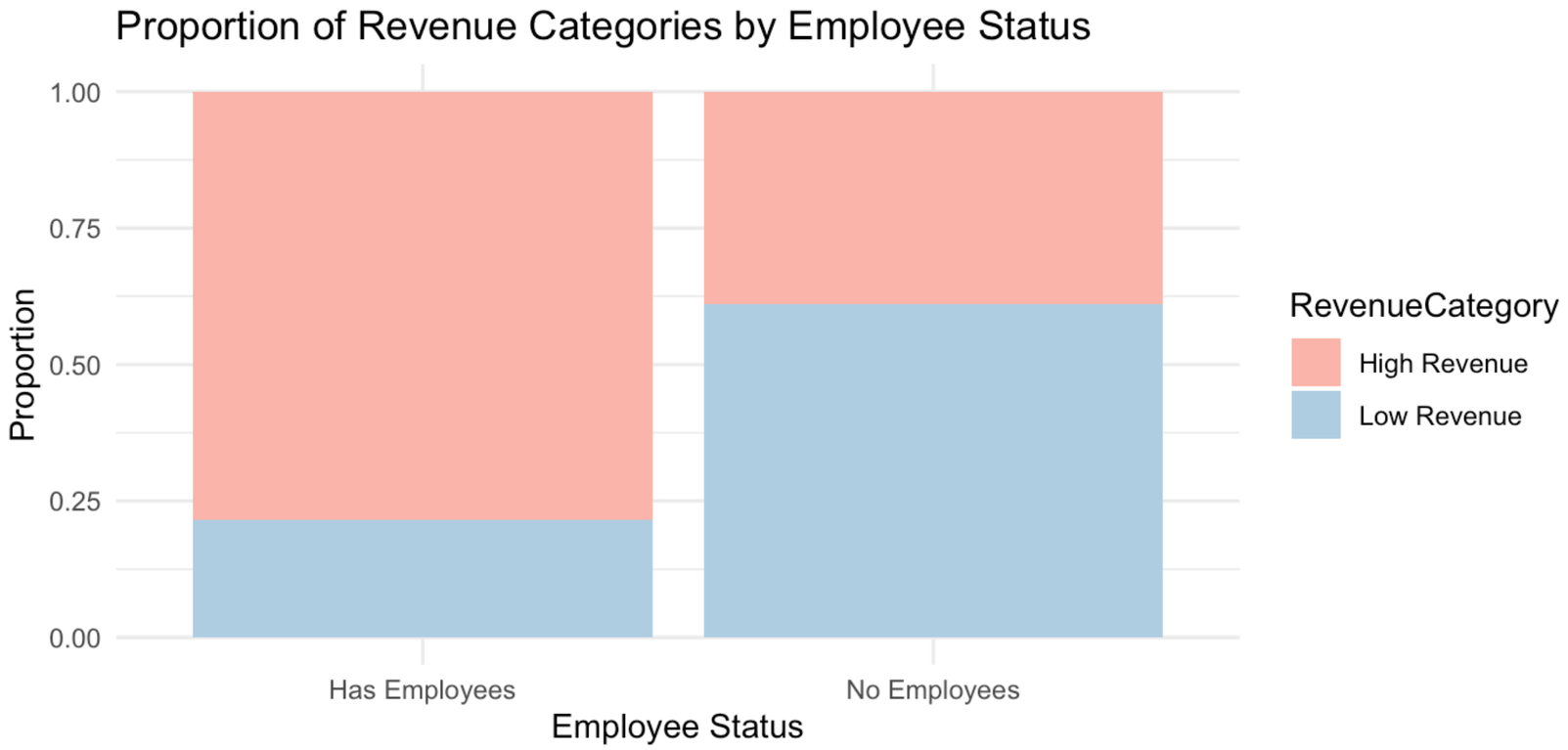
DOES HAVING EMPLOYEES INCREASE REVENUE?



Having more employees often translates to **increased productivity and efficiency** in farm operations. With a larger workforce, farmers can accomplish tasks more **quickly and effectively**, leading to **higher output and ultimately higher revenues.**

CHi SQUARE FOR THE ASSOCiATiON BETWEEN EMPLOYEES PRESENT AND REVENUE

```
        Pearson's Chi-squared test with Yates' continuity
        correction

data:  contingency_table
X-squared = 907.85, df = 1, p-value < 2.2e-16
```



THE RESULTS OF THE ANALYSIS STRONGLY REJECT THE NULL HYPOTHESIS, INDICATING A STATISTICALLY SIGNIFICANT ASSOCIATION BETWEEN HAVING EMPLOYEES AND GENERATING HIGHER OR LOWER REVENUE. IN PRACTICAL TERMS, THIS SUGGESTS THAT COMPANIES WITH EMPLOYEES TEND TO HAVE DIFFERENT REVENUE OUTCOMES COMPARED TO THOSE WITHOUT EMPLOYEES. THIS COULD IMPLY THAT THE PRESENCE OF EMPLOYEES IS A MEANINGFUL FACTOR INFLUENCING REVENUE LEVELS, POTENTIALLY INDICATING HIGHER BUSINESS ACTIVITY OR CAPACITY AMONG COMPANIES WITH EMPLOYEES. THEREFORE, THE PRESENCE OF EMPLOYEES COULD BE CORRELATED WITH HIGHER REVENUE, REFLECTING GREATER CAPACITY FOR PRODUCTION, SERVICE PROVISION, OR SALES ACTIVITIES.

CONCLUSION

The choice of education for success as an agricultural entrepreneur is not unique, but certain trends emerge. **High education levels generally correlate positively with economic performance**, leading to higher incomes and better financial management. **Specialized training,** such as agronomy can provide the necessary skills and knowledge for managing agricultural operations in modern commercial environments.

A higher number of employees can **optimize production**, leading to higher incomes. However, additional costs, such as maintenance costs, also play a role in financial success. Therefore, the choice of employee size and education depends on the **individual goals, size of the exploitation, available resources, and market conditions.**