Uni Variant Data Analysis Project

Bushra Naeemi

The American University of Afghanistan
Division of Science, Technology, and Mathematics
Instructor: Asadullah Jawid

November 23, 2019

Contents

1 Summary

The data is about the farmers from 3 provinces. farmers has been asked about their personal informations like age and gender, and their family members. Also about the income, accupation, crop net income, and rainfall.

2 Introduction

The data set is from 3 provinces and. It has been taken from the farmers about their farms. We have both numeric and categorical data, each has been categorized in columns.

2.1 Numerical Variable

Work.on.farm.FAMILY: Shows the number of family members working in the farm
Work.off.farm.Family: Shows the number of family members not working in the farm.
X.Not..employed...: number of net employed.
Female..work.on.farm: Number of Female working in farm.
Crop.Net.Income: Crop net income of farmer.
t4: Shows the temperature in the specific area in April.
r4: Shows the rainfall in the specific area in April.

2.2 Sample

Data has been collected from three provinces.

3 Methods

To analyze data we used a combination of both numerical measures and graphical methods. more specifically:

Numerical Measures:
- Mean(Measures of Central tendency)
- Standard Deviation and Range(Dispersion)
- skewness and kurtosis(Distribution)

Graphical Methods:

- Piechart(categorical variables)
- Density plot(numerical variables)
- histogram(numerical variables)
- boxplot(numerical varibale.

4 Results and Discussions

4.1 Numerical Variable

Statistics of Numerical Variables:

Variable	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis
$X.Not..employed...$	5.286	6.188	−191.0	33.0	−21.09	671.57
$Female..work.on.farm$	1.718	1.353	0.000	10.00	1.286	3.452
$Crop.Net.Income$	78369.5	287632.5	−951500	7089500	14.147	279.70
$t4$	4.726	2.997	0.217	11.514	0.2628	−1.1044
$r4$	60.021	27.843	28.932	123.623	0.853	−0.682

Table 1: Descriptive Statistics of Numerical Variables

numeric values from table
Tabel 1 The table above shows the comparison between the numeric values based of mean, standard deviation, minimum maximum, skewness, and kurtosis.

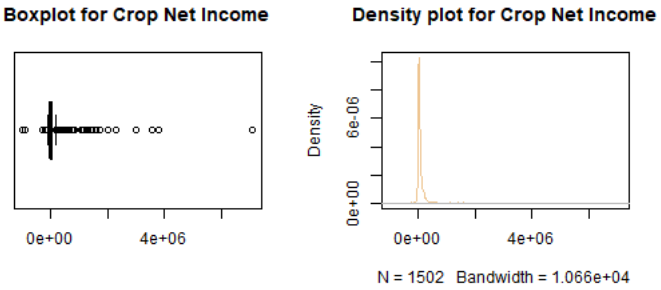

Figure 1: The box plot of Crop net income with the outliers with causes problems

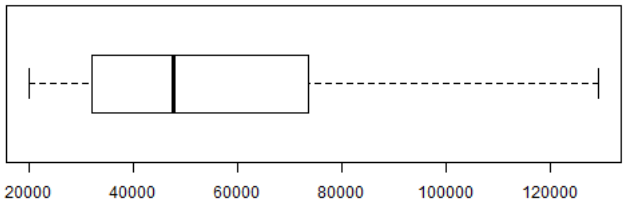

Figure 2: The box plot of Crop net income without outliers

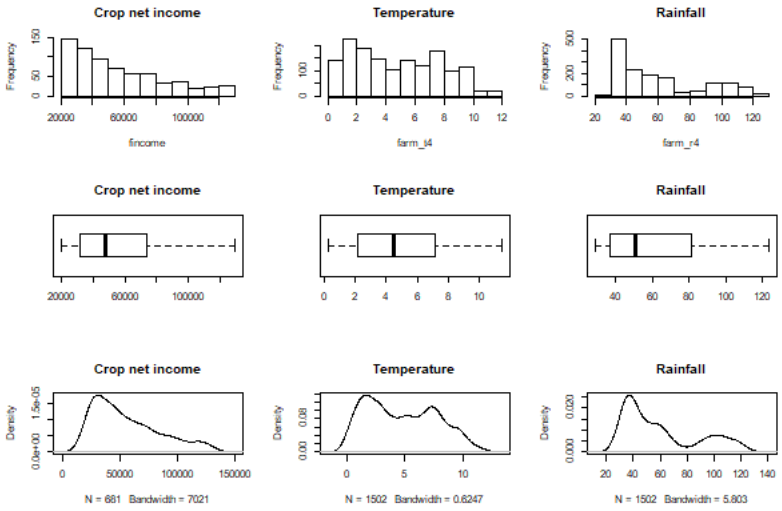

Figure 3: The comparison between Income rainfall and temperature by graphs

4.2 Categorical Variable

Importance.of.farming.0.not.atll.Impornt.1.importnt.2.Very.impornt: Shows the importance of farming for the farmers.

Table 2: Summary Statistics of Categorical Variables

Variable	Categories
Importance.of.farming.0.not.atll.Impornt.1.importnt.2.Very.impornt	not.atall[0]important[1]Ve

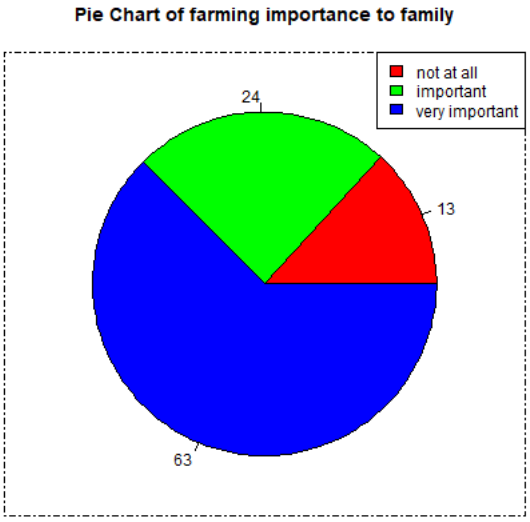

Figure 4: The pie chart of importance of farming for farmers

importance of farming
The pie chart shows that 13 percent of farmers think that farming is not important at all, 24 percent think that farming is important, and 63 percent of farmers think that farming is very important.

4.3 Sub-Sample Analysis

Crop Net Income, Temperature, Rainfall in 3 provinces
In this section we show the Crop Net Income in province 1, province 2, and province 3. The temperature in province 1, province 2, and province 3. The rainfall in province 1, province 2, and province 3.

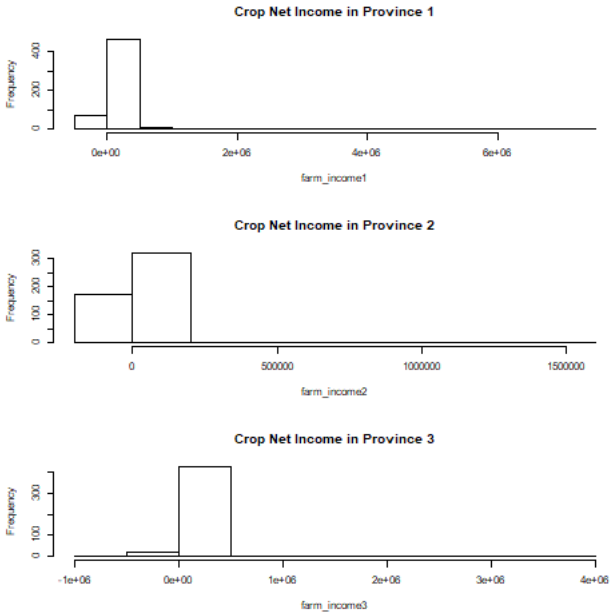

Figure 5: The Net income of 3 provinces

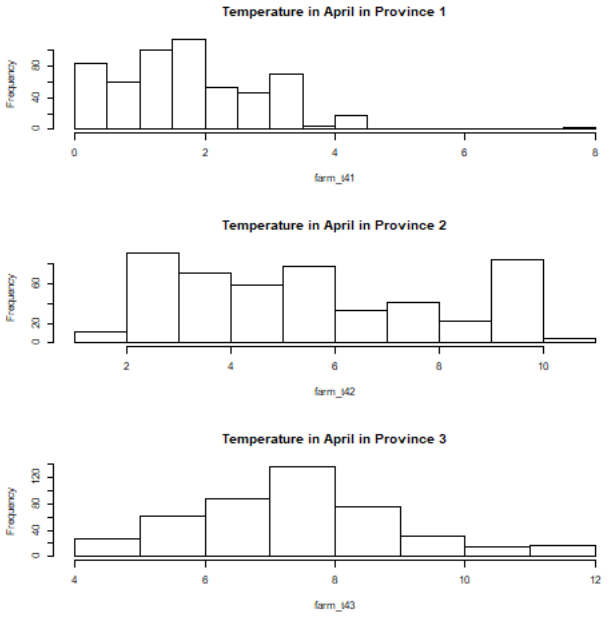

Figure 6: The temperature of 3 provinces

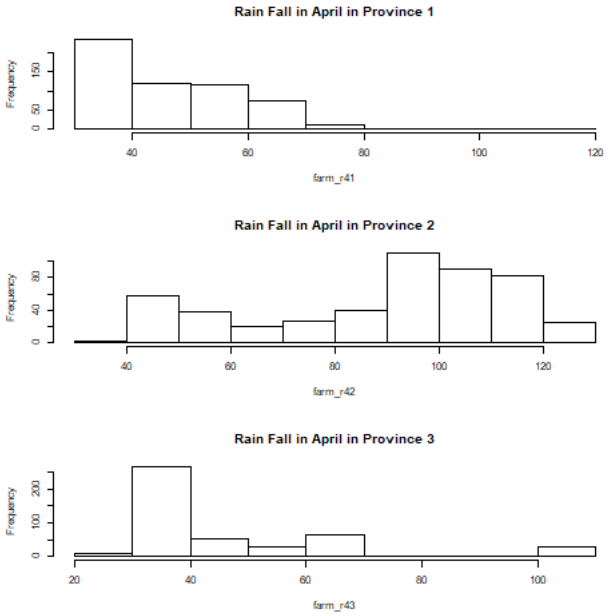

Figure 7: The rainfall of 3 provinces