Predicting Labor Informality

Nelson Chacon Rendon

24 December, 2020

Contents

Introduction

Describes the dataset and variables, summarizes the goal of the project and key steps performed.

References like this (Fernández 2019).

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

International Labor Organization (ILO) states that. . . .

The term "informal sector" was first coined in 1972, at the outcome of a comprehensive ILO employment advisory mission to Kenya[33], and was later converted to "informal economy" to underline the fact that informality is not a "sector", but a certain way of carrying out economic activities. Interestingly, "The (Kenya mission) report acknowledges that the informal sector idea originated not with the high-level foreign"development experts" brought in for the mission but from the work and the staff of the Institute or Development Studies of the University of Nairobi, a fact which has been generally forgotten since then. In other words, it was not the ILO which invented the concept of the informal sector. It came out of the thinkers and analysts of the Third World"

The informal economy comprises more than half of the global labour force and more than 90 per cent of micro and small enterprises (MSEs) worldwide. Informality is an important characteristic of labour markets in the world with millions of economic units operating and hundreds of millions of workers pursuing their livelihoods in conditions of informality.

The expression "informal economy" encompasses a huge diversity of situations and phenomena. Indeed, the informal economy manifests itself in a variety of forms across and within economies. Formalization processes and measures aiming to facilitate transitions to formality need to be tailored to specific circumstances faced by different countries and categories of economic units or workers.

Work in the informal economy is often characterized by small or undefined work places, unsafe and unhealthy working conditions, low levels of skills and productivity, low or irregular incomes, long working hours and lack of access to information, markets, finance, training and technology. Workers in the informal economy are not recognized, registered, regulated or protected under labour legislation and social protection. The root causes of informality include elements related to the economic context, the legal, regulatory and policy frameworks and to some micro level determinants such as low level of education, discrimination, poverty and, as mentioned above, lack of access to economic resources, to property, to financial and other business services and to markets. The high incidence of the informal economy is a major challenge for the rights of workers and decent working conditions and has a negative impact on enterprises, public revenues, government's scope of action, soundness of institutions and fair competition.
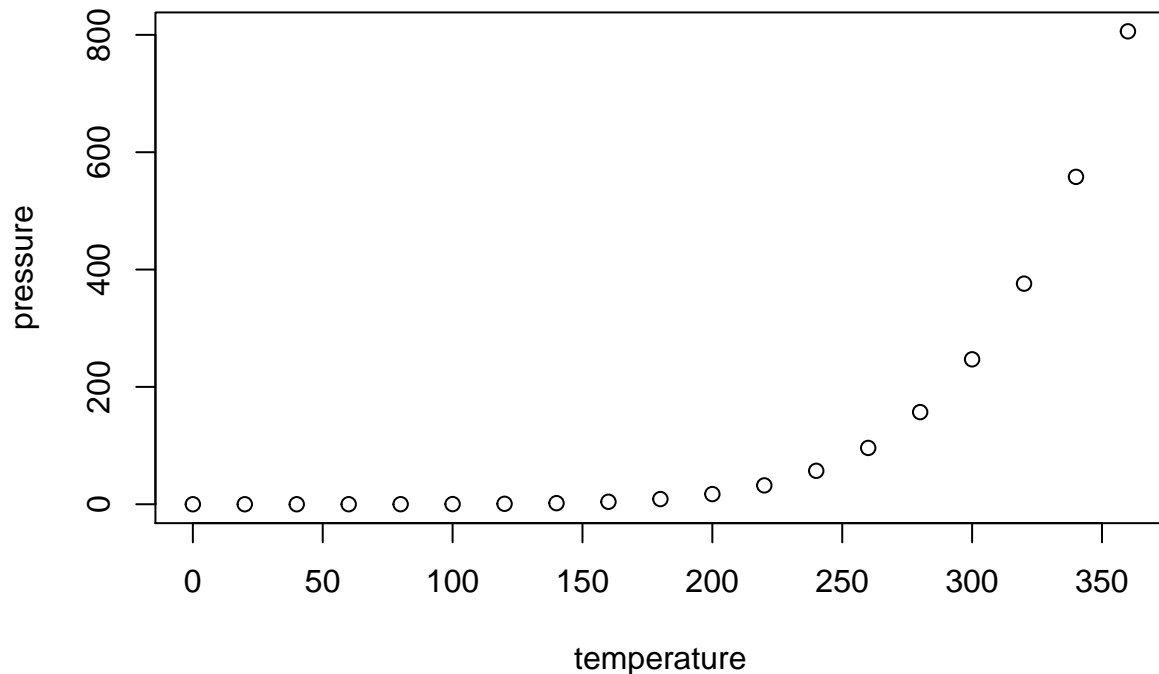
the term "informal economy" referred "to all economic activities by workers and economic units that are – in law or in practice – not covered or insufficiently covered by formal arrangements." This description was endorsed 13 years later which further clarified that the term does not cover illicit activities, and that the expression "economic units" in the definition refers to units that (a) employ hired labour; (b) individuals working on their own account, and (c) cooperatives and social and solidarity economy units. R.204 constitutes a historic landmark for the world of work as it is the first international standard focusing exclusively on the informal economy in its entirety.

To promote decent work, there needs to be a comprehensive and integrated strategy cutting across a range of policy areas and involving a range of institutional and civil society actors that eliminates the negative aspects of informality, while preserving the significant job creation and income generation potential of the informal economy. It should promote the protection and incorporation of workers and economic units in the informal economy into the mainstream economy.

Analysis

Explains the process and techniques used, including data cleaning, data exploration and visualization, any insights gained, and your modeling approach. At least two different models or algorithms must be used, with at least one being more advanced than simple linear regression for prediction problems.

You can also embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

describing data

Plot 2

Imputing missing data

Sometimes we have some problems with missing values in real-life datasets. Sometimes missing values is a real problem for research. How to recover missing values in our datasets? In this post we will consider one very interesting method of recovering missing values. It is method a bagging of regression trees. It provides the recovery of missing values for several variables at once, based on regression dependencies.

For bagging of regression trees we will use caret package for R programming language.

Data partition

The partitioning of the database into the training and testing sub-samples is done following the criteria of maintaining the balance of formal and informal workers from the original base, that is, we know that the original base presents a proportion of about 80% of informal workers and 20% of formal workers, for this reason, it is of utmost importance to maintain these proportions in the sub-samples that we will create and use for both the training and the validation of our models.

Results

This section presents the modeling results and discusses the model performance.
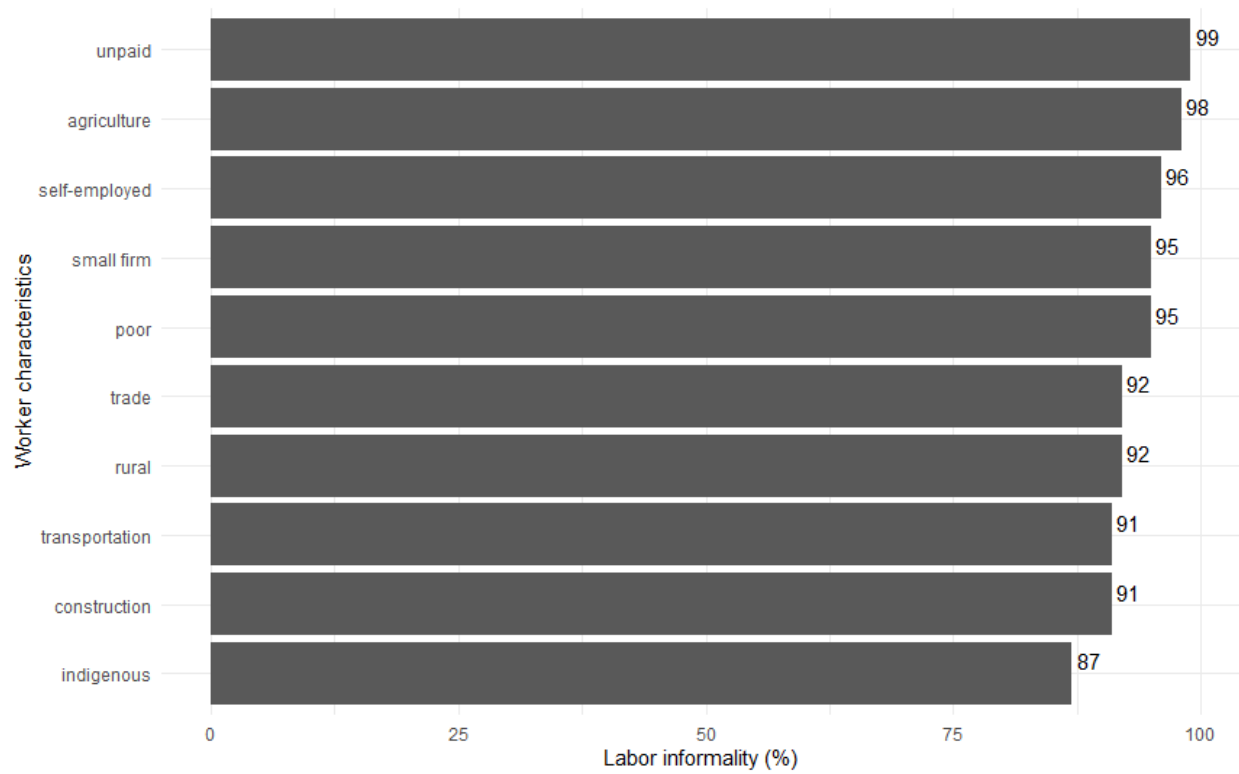
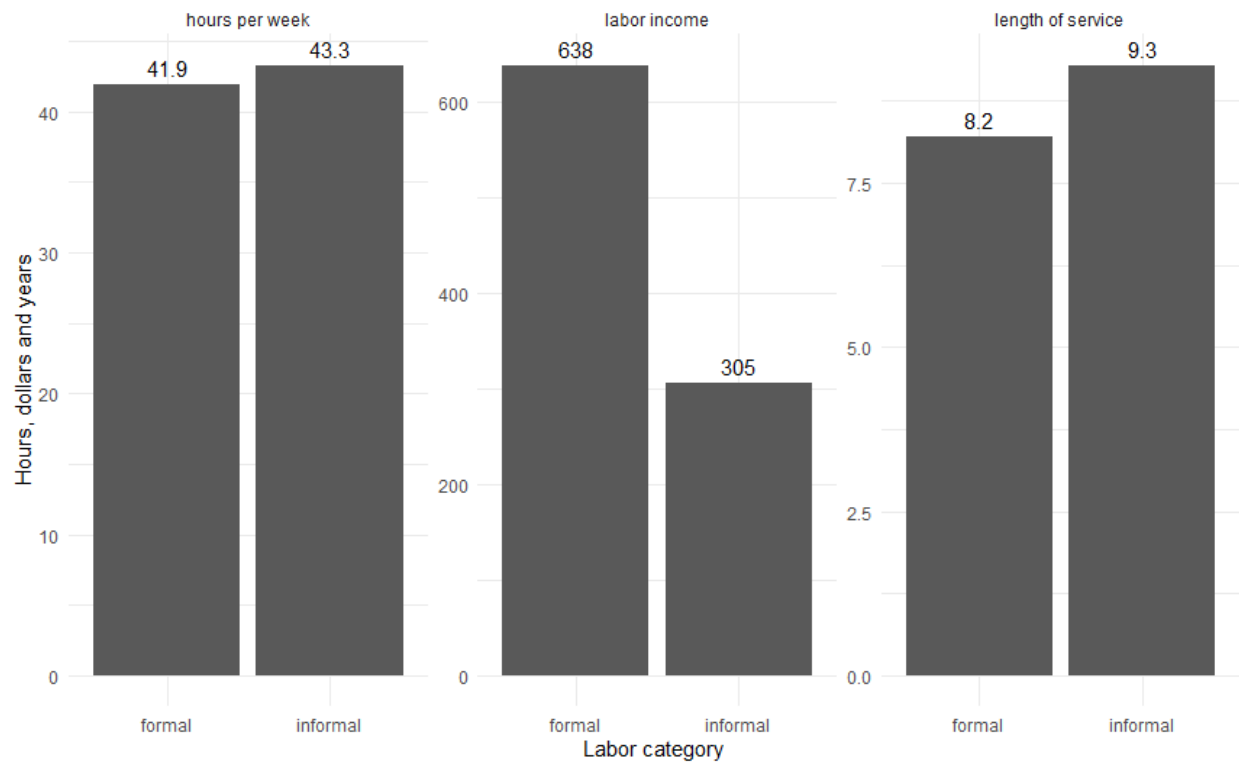Figure 1: Percentage of informal workers according to different criteria

Figure 2: Wage differences, seniority and hours worked between formal and informal workers

Conclusion

Brief summary of the report, its potential impact, its limitations, and future work.

- 5 points for 3 files

- 25 points a good report

- 20 points good R code script

- try to send it January 3rd

In a later paper I would like to distinguish informal workers between those who have voluntarily chosen to be part of this sector and those who are informal out of necessity, in this sense I think that in order to make this distinction it will be key to distinguish the monetary income of each sector, hoping that the voluntary informals earn much more money than the informals out of necessity, another group of variables could also help to make this distinction, perhaps elements such as the characteristics of the housing, the assets they possess or the degree of education of the workers could be variables that would allow for a relatively precise classification of both sectors. In this way, appropriate policies could be designed to advance the labor formalization of these people; it is very likely that the measures required to formalize some are not necessarily the same as those leading to the formalization of the others.

References

Fernández, José Miguel Molina. 2019. "The Dynamics of Labor Informality and International Trade in Bolivia." *Economia Coyuntural,Revista de Temas de Perspectivas Y Coyuntura* 4 (3): 9–34. https://ideas.repec.org/a/grm/ecoyun/201907.html.