Employees

Zeybel, Ahmet Oğuz*

Abstract

To determine whether the number of workers employed in routine jobs has decreased or not in countries with the development of technology and to interpret and predict what the future of routine work will be as technology evolves, we used dataset included 16 variables deriving from ILOSTAT. When we did a literature review, we observed that as the need for manpower decreased with the development of technology, the workforce shifted from the agricultural sector to the industrial sector, and productivity and economic growth increased. However, when performing data analysis, having too many NA values and their importance in dataset prevented us from interpreting the large data set and pushed us to examine a smaller sample.

1 Introduction

The research question: The research question of this study is to determine whether the number of workers employed in routine jobs has decreased or not in countries with the development of technology. Data Source: This dataset has been downloaded from ILOSTAT. Observation and Variables: In the dataset, the observation section includes only countries. The variables section includes reference area, source, economic activity, time, total, employees. Employees is also divided into 11 for itself. They are managers, professionals, technicians and associate professional, clerical support workers, service and sales workers, skilled agricultural,forestry and fishery workers, craft and related workers, plant and machine operators, and assemblers, elementary occupations, armed forces occupations, not elsewhere classified. Employees are all those workers who hold paid employment jobs, which are those where the incumbents hold employment contract which give them a basic remuneration not directly dependent upon the revenue of the unit for which they work. Reference means countries. The Importance and Purpose of Study: As technology advances, new job groups will emerge while some jobs groups will disappear. This research will help to make accurate predictions about the workforce in the coming years.

*20080509, Github Repo

1.1 Literature Review

With the development of technology, it has emerged over the years that the number of people working in routine jobs and the number of employed people have decreased and their productivity has increased. Especially in the field of financial transactions, information technology can reduce the workforce required for each transaction, while also greatly increasing the overall volume of financial markets.(Fung, 2006). Occupations in the middle of the wage distribution in high-income countries tend to involve predominantly routine tasks - tasks that are easily codiable and therefore particularly susceptible to automation.(Cortes & Morris, 2019). The development of technology and the machines produced require less manpower, while at the same time leading to an increase in production and efficiency.(Gallardo & Sauer, 2018) With this situation begins an employment transition from agriculture to industry, which in turn helps economic growth.(Mueller et al., 2019)

2 Data

```
library(tidyverse)
library(here)
library(readxl)
library(dplyr)

employees <- read_excel("C:/Users/rstudio/Desktop/final/data/EES_TEES_ECO_OCU_NB_A_EN.xl
    col_types = c("text", "skip", "text",
        "numeric", "numeric", "numeric",
        "numeric", "numeric", "numeric",
        "numeric", "numeric", "numeric",
        "numeric", "numeric", "numeric",
        "numeric", "text"), skip = 5)
```

I change the name of data to something more easly recallable.

With the `echo=FALSE` option, prevent the codes from appearing in the derived pdf file and report your results in tables.

3 Methods and Data Analysis

I selected 20 of the High Income countries using subset.

```
employees %>%
  ggplot(aes(y = `Total`, x = Time,color=`Economic activity`)) +
  geom_point()
```

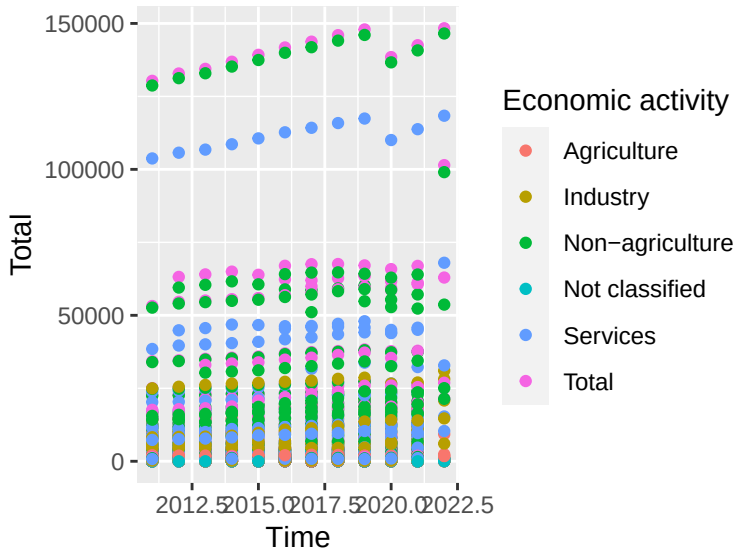
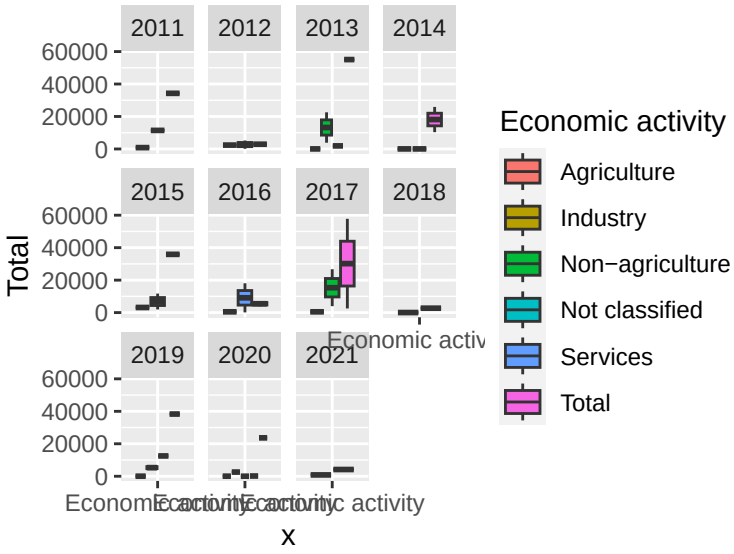


Figure 1: An Awesome Plot

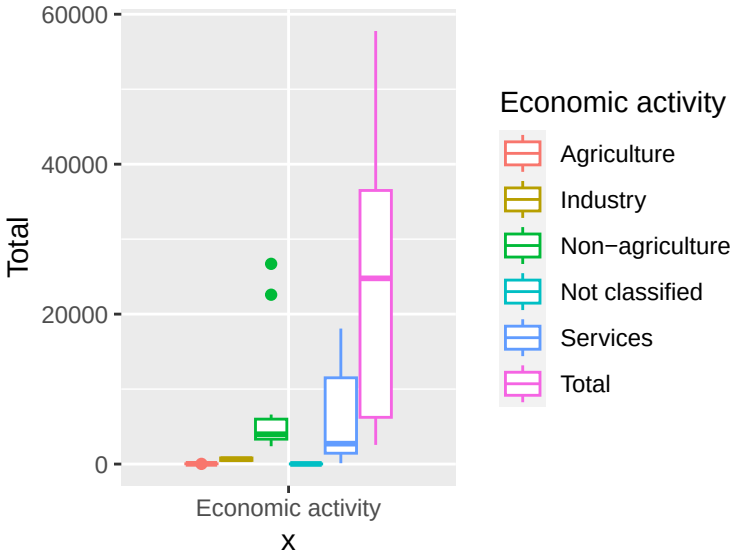
I selected 10 of the High Income countries using subset. I called it highinc short for High Income

```
highinc <- employees[ which(employees$`Reference area`==c('Switzerland','Japan','Canada'



View(highinc)
```

```
p1 <- ggplot(highinc, aes(x='Economic activity', y=Total, fill=`Economic activity`)) +
    geom_boxplot() +
    facet_wrap(~`Time`)

p1
```




```
highinc %>%
  ggplot(aes(y = Total, x = 'Economic activity',color=`Economic activity`,fill)) +
  geom_boxplot()
```



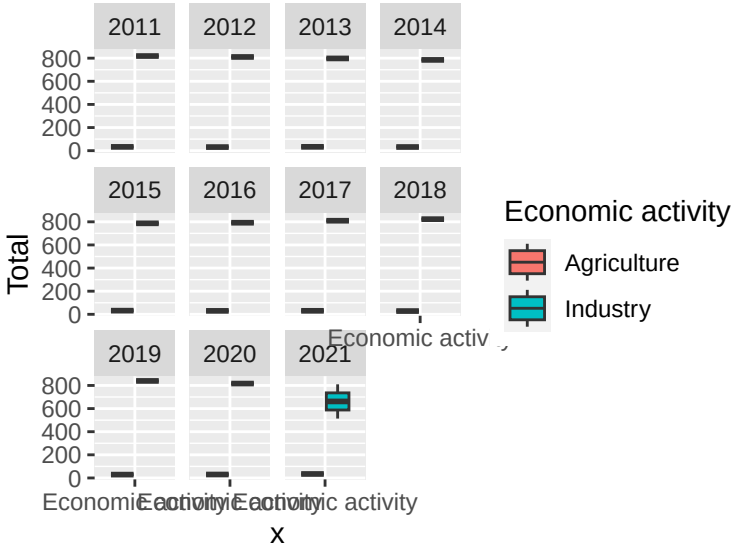
This is Economic Activity of Developed Countries years combined.

There is too much Na in data which I cannot handle without causing a distortion to the analysis. I realised that if I should check the relationship between agriculture and Industry through the years for developed countries It might show me yearly decline or increase. However When I subset the data using "&" function (For developed countries and Economic Activity classified by Industry and Agriculture) NA values and data's merged qualities gave me really hard time. So I have decided to look for developed countries which has non NA values for Industry and Agriculture and It left me a really small sample size.

```
highinc1 <- employees[ which(employees$`Reference area`==c('Norway','Finland','Sweden')


naless1=na.omit(employees)
```

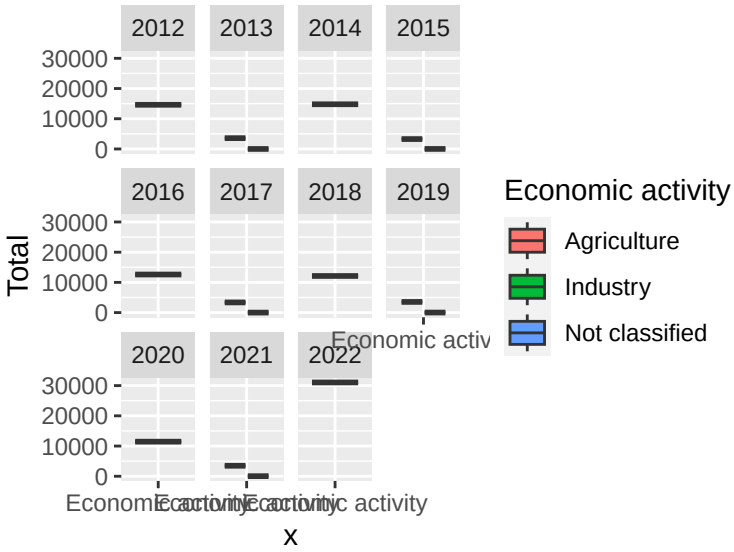
I tried Na omit function but since there was too much Na in important places. It did not work. I also tried changing Na values to median or mean but since there are too much confounding factors, it distorted the data even more.

```
p1 <- ggplot(highinc1, aes(x='Economic activity', y=Total, fill=`Economic activity`)) +
    geom_boxplot() +
    facet_wrap(~`Time`)

p1
```

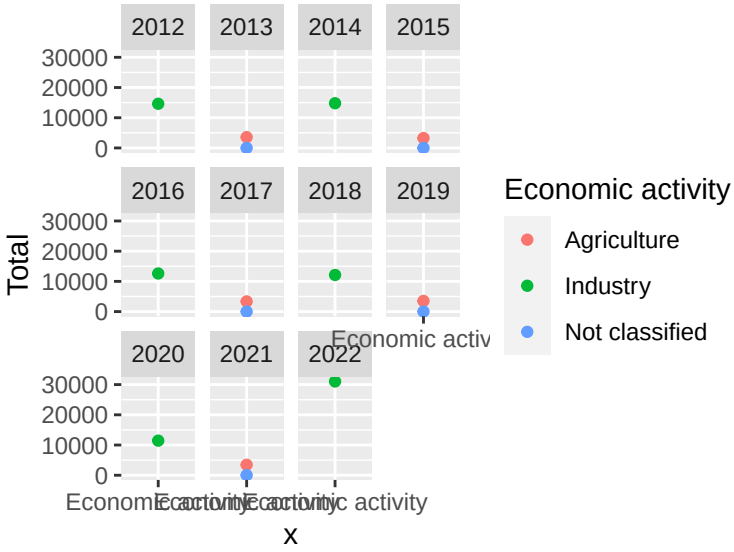


And trouble with this specific subset that it is already well developed so in analysing last 5 years of data it is already seen Industry is greater part of their economy with subtle changes through the years. Also Geographically and culturally they are very close societies. So There is probably an "overfitting" error.

```
developing <- employees[ which(employees$`Reference area`==c('India','South Korea','Braz

p1 <- ggplot(developing, aes(x='Economic activity', y=Total, fill=`Economic activity`))
    geom_boxplot() +
    facet_wrap(~`Time`)

p1
```




```
p3 <- ggplot(developing, aes(x='Economic activity', y=Total, color=`Economic activity`)) +
    geom_point() +
    facet_wrap(~`Time`)

p3
```



4 Conclusion

In conclusion. I have used data set gathered from the ILOSTAT It has been merged from few different reliable sources, but since the merging applications can be chaotic, it resulted in many NA values in key parts of the data set.I tried best to my abilities however it was

not enough to change the Na values or completely omiting them. It is clear to me more advanced methods should be in better use. Which i don't know about.

I was searching for a relationship between time related relation in Economic activities classified as: Industry or Agriculture. Statistical calculations has been inconclusive because said road blocks. Still I try my best as to use methods I learned in this class. Using dplyr package to clean the data and ggplot2 package to draw graphs. However it didn't yield any significant results.

In order for this analysis to be significant, we need to go to the data collection step and start over with a more specific data that I could process.

5 References

Cortes, G. M., & Morris, D. M. (2019). Are routine jobs moving south? Evidence from changes in the occupational structure of employment in the US and mexico. *Evidence from Changes in the Occupational Structure of Employment in the US And Mexico*.

Fung, M. K. (2006). Are labor-saving technologies lowering employment in the banking industry? *Journal of Banking & Finance*, *30*(1), 179–198.

Gallardo, R. K., & Sauer, J. (2018). Adoption of labor-saving technologies in agriculture. *Annual Review of Resource Economics*, *10*, 185–206.

Mueller, V., Masias, I., & Vallury, S. (2019). Labor-saving technologies and structural transformation in northern ghana. *Agricultural Economics*, *50*(5), 581–594.