

ICU Report

Leon Victor

29/04/2021

Summary

This project is part of EpimedSolutions challenge. Its description is showed below:

You have also received the file “dataset.RData” together with this document. The .RData file contains data on 10.000 patients that has been admitted to a certain Intensive Care Unit (ICU). It is your responsibility to create an R package to help to analyze this data.

Steps it must have:

The package should contain functions to:

- 1) make the appropriate treatment to the dataset.
- 2) produce some plots to understand the relationship between the variables and get some insights about what increases the chances of a patient to die in this ICU.
- 3) make predictions about the probability of a patient to die in this ICU.
- 4) produce a final report – in English! – using the other functions created to describe the results you got while analyzing data.

Variables descriptions:

1. UnitAdmissionId: unique identifier for each patient.
2. Age: age of patient (in years).
3. Gender: gender of patient (“F” for female and “M” for male).
4. UnitDischargeCode: “A” (alive) or “D” (dead) at the end of admission.
5. UnitLengthStay: number of days each patient has been in the ICU.
6. IsArterialHypertension: indicates if the patient has arterial hypertension.
7. IsDiabetesComplicated: indicates if the patient has diabetes.

Report

As part of the analysis, it is fundamental to the use of “tools_package.R” file which follows this report on the same fold.

Loading the dataset and package

The initial step is to load the dataset to start its analysis as well as some libraries required.

Data cleaning and treatment

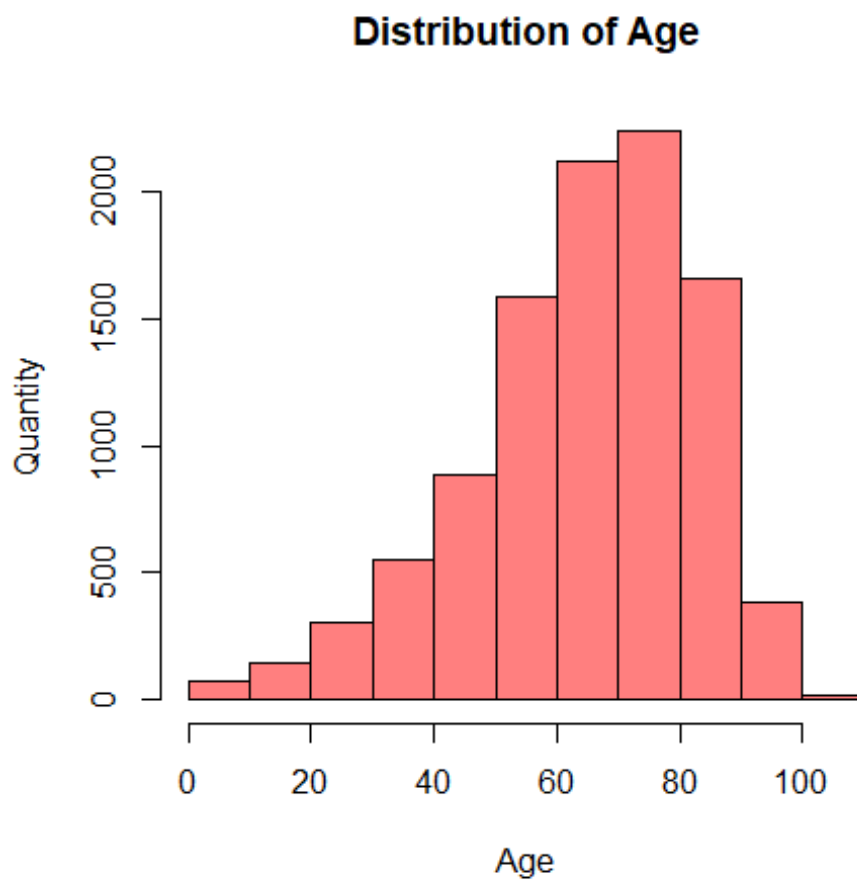
Then loaded, we must clean the dataset to provide concise elements for use in exploratory analysis. The package “tools_packages.R” has several functions which will guide any approaches on this report.

After the treatment, the new values for data is determined as following: Male = 1, Female = 0, Died = 1 ,and Alive = 0.

Exploratory Analysis

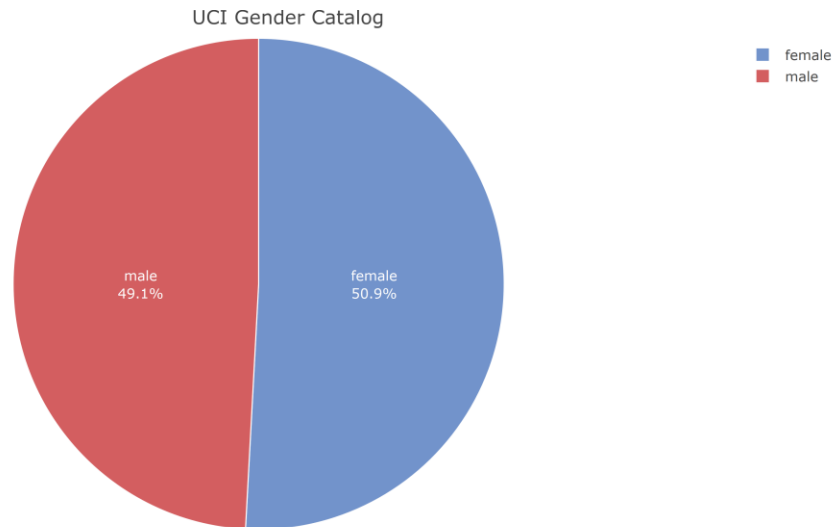
The first step in this segment is to understand the distribution of age in UCI. Making use of simple elements, it is possible to have in mind the type of the data and circumstances to identify any disequilibrium in the dataset.

Insight: The major ages are between 60 and 80 years.



In addition, here is showed how is distributed the gender in the UCI. It helps observe whether men or women have been getting into UCI more.

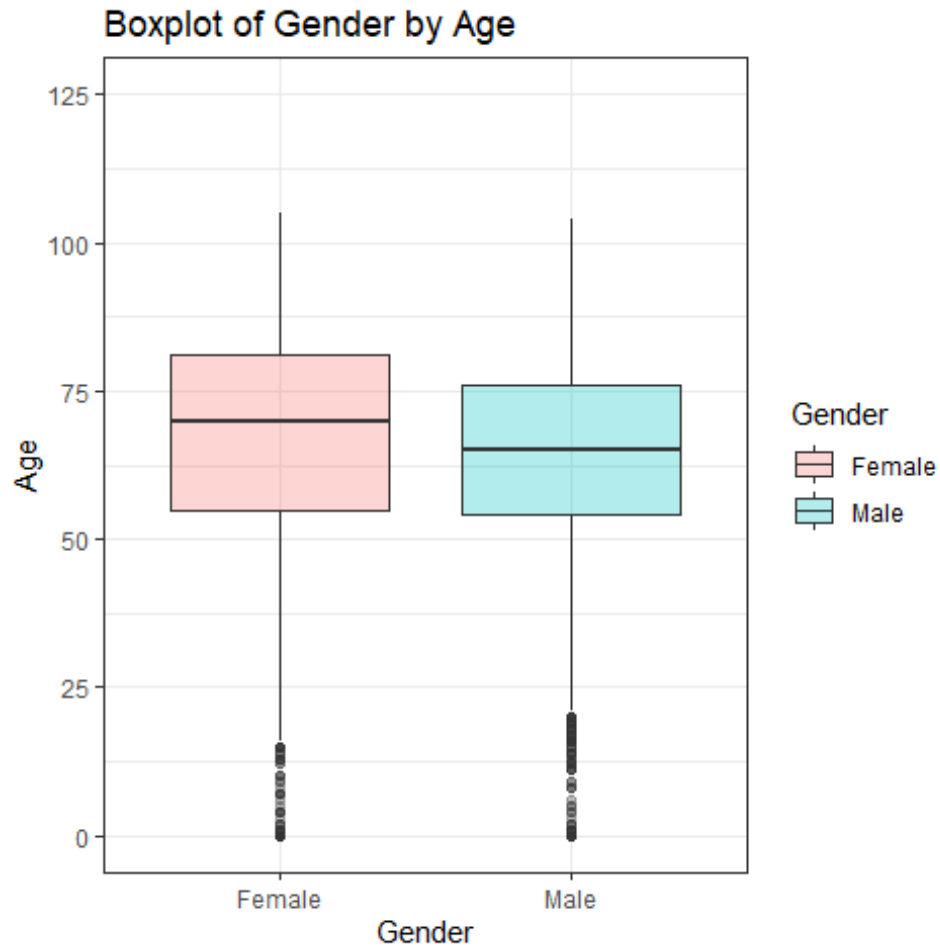
Insight: The allocation of both genders is the same.



Diving deeply into this subject yet, Using a boxplot we can see some important aspects related to gender and age from UCI patients.

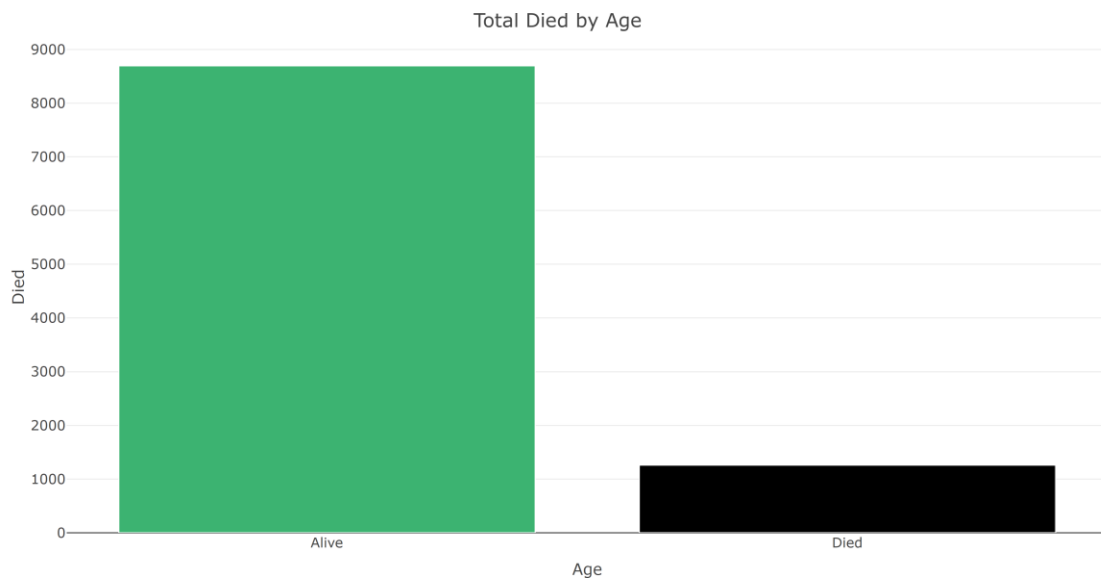
Insights:

1. Women age is more than the men age;
2. The concentration of 25 % of ages are around: 55 years old for both;
3. Median: Men is less than Women;
4. 75 % is around 80 years old for Women, while 75 years old for men;
5. Both have some youngs in UCI, the men years range is more than women though.



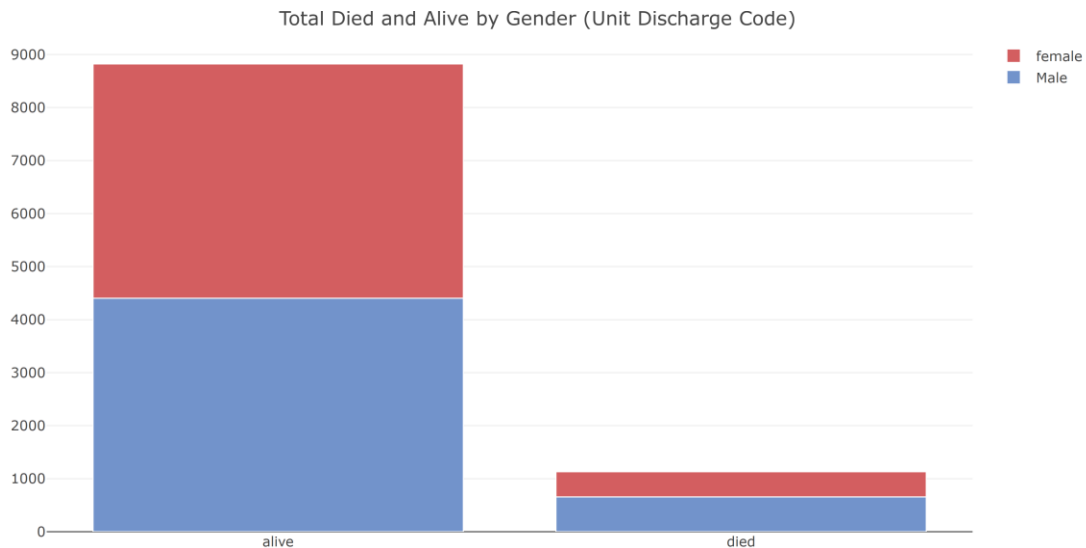
The total of dead and alive is a precious measurement that contributes to creating an overview of how many people did not survive.

Insight: The number of patients who survived is pretty superior for not survived.



In addition to the last plot, to have it more precisely and segmented by gender brings the concept of how it is distributed.

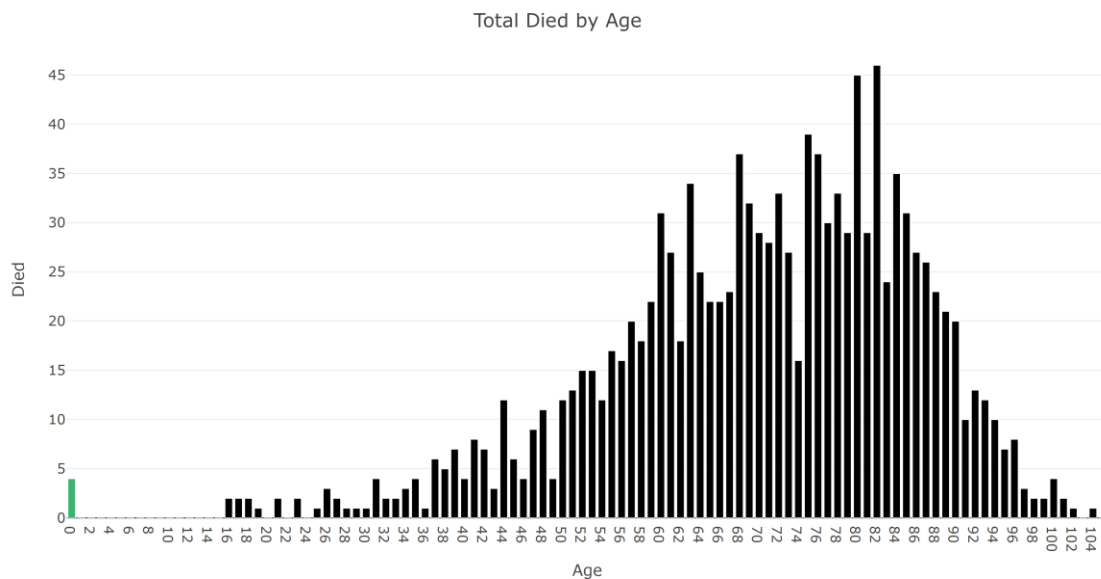
Insight: There is no difference between the total of women or men who died or survived.



As previously exhibited, there is not a massive difference in the number of lost lives for gender. In this subject yet, another observation is showed by Total Died by Age. and it is observed how many lost their lives segregated by age.

Insights:

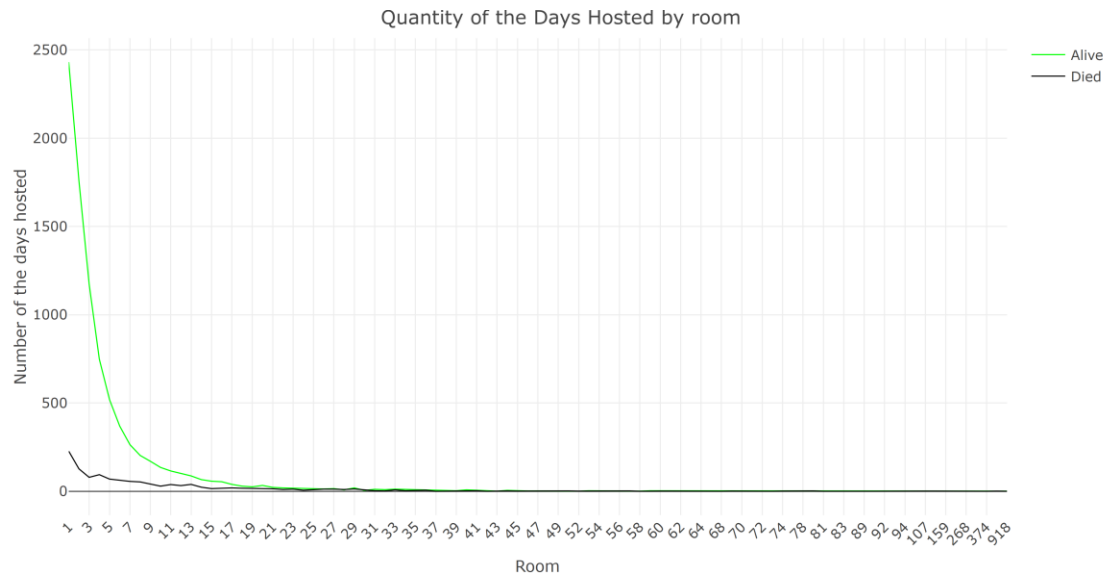
1. People with more than 60 (sixty) years are a substantial mass;
2. 04 (four) babies, 0 (zero) age, lost their lives as well.



Furthermore, How is the behavior of people in the rooms?

Insights:

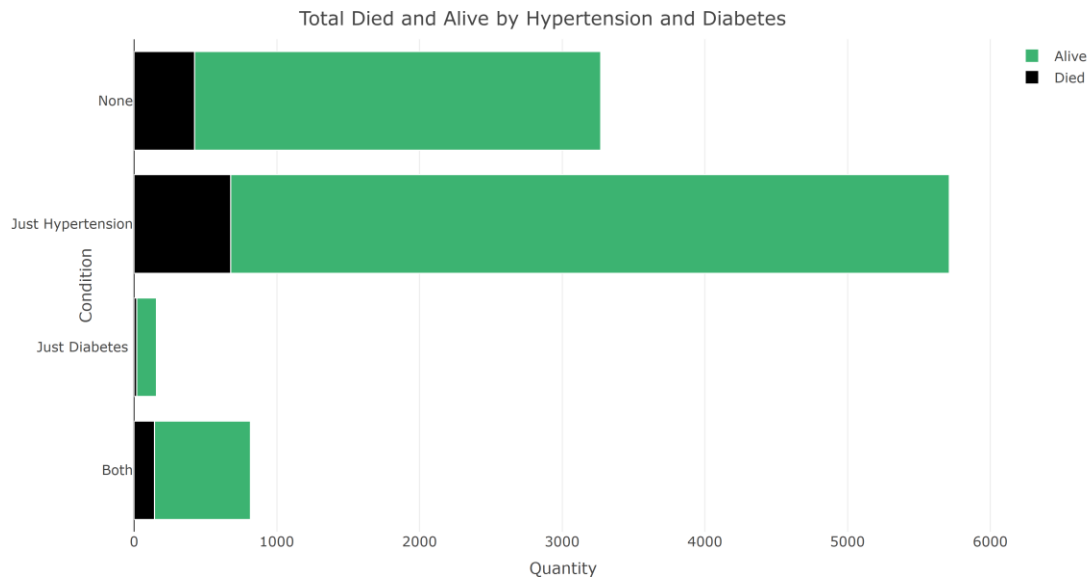
1. Most of the time, the rooms are occupied for 01 (one) day;
2. People who are alive concentrate more attend for rooms;
3. As long as the number of days decreases, the number of people making use of rooms is less.



The last analysis step is the relation between Arterial Hypertension and Diabetes. This insight is also related to people who have died.

Insights:

1. The result shows most parts of patients have Hypertension and no Diabetes;
2. Meanwhile, the second big part of patients has none;
3. More dead for Hypertension and none disease.



Conclusions

1. The most of people who use the UCI are between 60 to 80 years old;
2. Basically, there is no concentration of gender in UCI. Both have almost the same population;
3. Women age is more than the men age;
4. The concentration of 25 % of ages are around: 55 years old for both;
5. Median: Men is less than Women;
6. 75 % is around 80 years old for Women, while 75 years old for men;
7. Both have some youngs in UCI, the men years range is more than women though;
8. The number of patients who survived is pretty superior for not survived;
9. There is no difference in the number of deaths by gender. They are very similar;
10. About deaths: people with more than 60 years are the majority;
11. 04 (four) babies, 0 (zero) age, lost their lives;
12. Most of the time, the rooms are occupied for 01 (one) day;
13. People who alive concentrate more attend for rooms;
14. As long as the number of days decreases, the number of people making use of decreases;
15. Most of the patients have Hypertension and no Diabetes;

16. the second big part of patients do not have any disease;
17. Majority of deaths are from Hypertension and none disease.