Report for Vital Signs Diagnosis data

Bangcale, Kurt Andrei A.[1]
Era, Chloe Josephine F.[2]

1Bangcale, Kurt Andrei A.,2Era, Chloe Josephine F., Data Science for the Health Sciences/CON29, School of Health Sciences, Mapúa University

INTRODUCTION

The dataset contains medical datadata from 1000 patients. This dataset contains the vital signs of individual patients, health. It also includes the gender, lifestyle, age, etc.Introduction to the dataset and Methods used for analysis.

The dataset being analyzed also investigates the connection between age and range of lifestyle and demographic characteristics. The key variables in the dataset include age, which is the age of the individual, heart rate, or the individual's heart rate in beats per minute, and stress level, an ordinal variable representing different levels of stress.

An individual's age can have an impact on their vital signs.Right now, the objective is to find if the age of patients is associated with their vital signs such as heart rate and stress level. The procedure of the methods will analyze if there is really an influence to it from the two variables, where heart rate and stress level are influenced by a patient's age.

METHODS

The researchers used a personal laptop with RStudio installed and internet access for downloading the necessary dataset. The dataset file, *1_Vital_signs_diagnosis_data_Group_017.csv*, was downloaded first from online and then loaded into the program using the command typed in the Rscript after creating a new project.

The dataset was cleaned first, and then after that, plot information for age, heart rate, and stress levels data were applied.

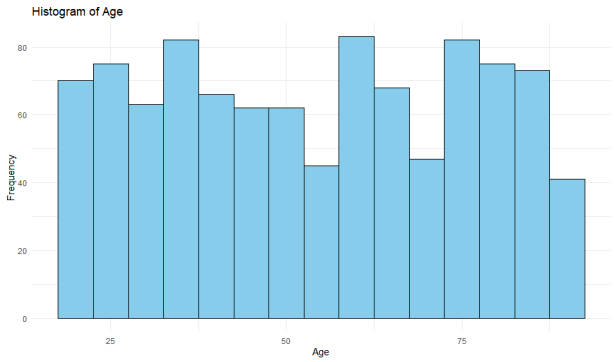


Figure 1: Age Histogram

The histogram shows the age of all 1000 patients in the dataset, where "Frequency" is the number of individuals. It clearly shows here that the majority of patients from the dataset have an age range between 50+ years old.

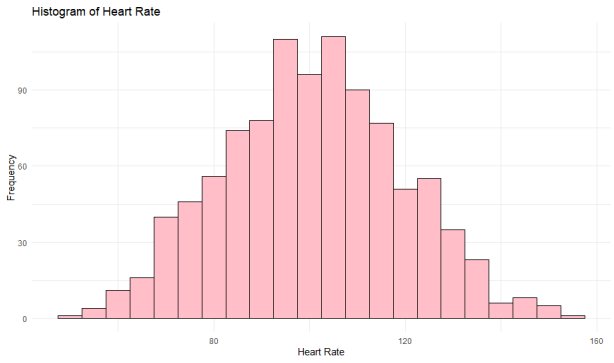


Figure 2:Heart rate Plot

It created a histogram for the patients' heart rate. Majority of the patients' heart rate in the dataset ranges between 90-120 BPM (beats per minute).

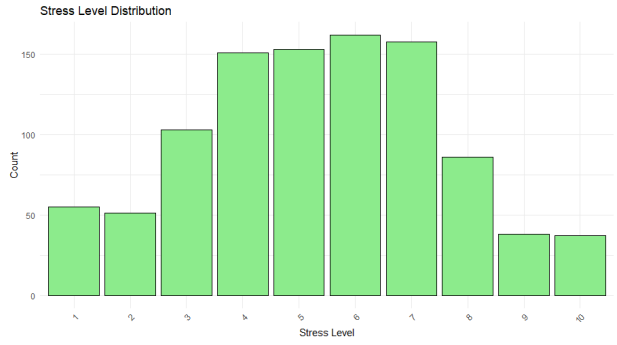


Figure 3: Stress Level plot

This shows the plot for the patients' stress levels, where 1 indicating the lowest stress experienced and 10 being the highest. The plot shows that the majority of patients experience stress levels between 4-7.

After this analysis, the researchers used the code for plotting Age vs heart rate/stress level. Age vs heart rate relationship is compatible for pearson's correlational test due to that being straightforward and linear, while spearman correlational coefficient tests for the age vs stress level, since none of the variables are ordinal, and it is also a non-parametric test for monotonic relationship strength measure.

The Age-heart rate relationship is also possible for evaluation of simple linear regression to model the sum of squared residuals, wherein the slope provides the estimate of the change in the health literacy score for each unit increase in age.

The age-stress level relationship however, was measured with logistic regression, because oth variables are in binary.

RESULTS

After inputting and running the codes, the researchers found the results.

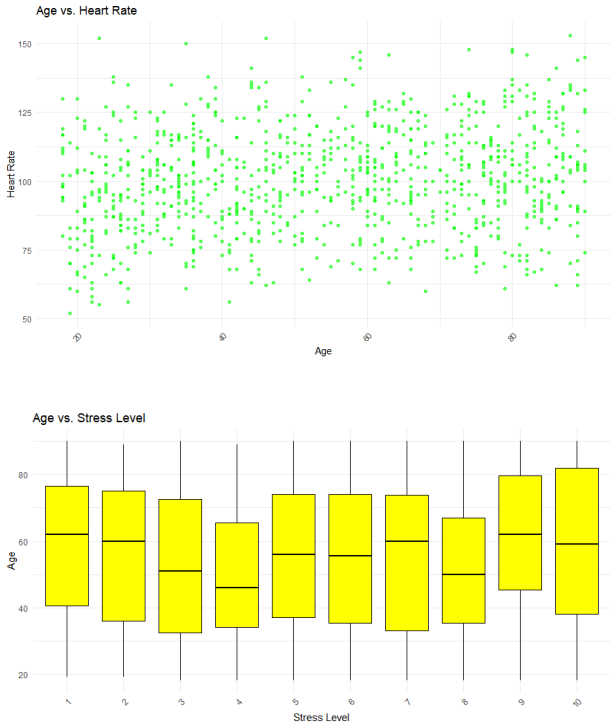


Figure 4. Age vs Heart rate/stress level plots

The scatterplot for age-heart rate relationship indicates that

The Pearson's Rank Correlation for Age vs Heart rate values would be:

t-value: 5.4225

p-value: 7.383×10^{-8}

Correlation Coefficient: 0.1697

95% Confidence Interval: [0.1086, 0.2294]

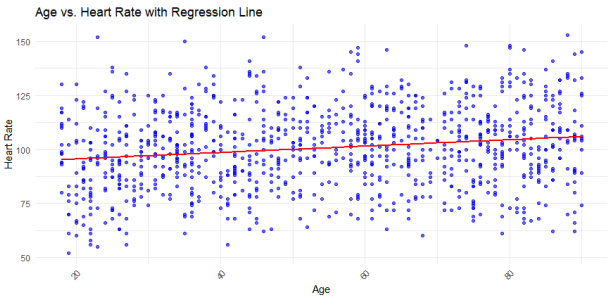
This signifies that the relationship age-heart rate has a very weak positive correlation, with a person's coefficient value of 0.1697 . Also, the p-value is less than 0.05, indicating the correlation is not significant. In other words,

age is hardly a factor in the heart rate of the patients in the dataset.

For the simple linear regression plot results for Age-heart rate plot (Figure 5), the R-squared is 0.02879, and the Slope (β_1 for Age) is 0.14742. This means that for each additional year of age, the heart rate indicates that this relationship is also not significant. The r-squared for this is very low too, meaning that the variance in heart rates is barely even affected by age, but it has a positive correlation.

For the spearman correlational result of the age-stress levels, it has a correlation of ρ=0.0286, which shows a very weak but positive relationship between the two variables. It also has a p-value of 0.3672, greater than 0.05, meaning that this relationship between age and stress level is not even significant.

In terms of logistic regression result, the p-value of 0.3853 is greater than 0.05, still indicating that age is not not significant in predicting stress level. Although, there is a very small positive effect of age on stress level but the effect is weak and is barely even noticeable,



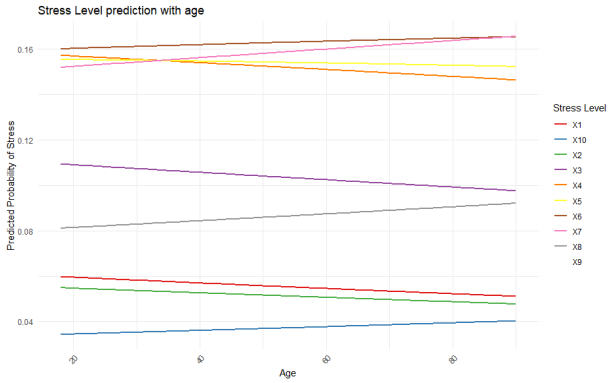


Figure 5.age vs heart rate/stress level linear/logistic regression plots

CONCLUSION

While the relationships are actually significant, the low r-squared values suggests that ae only explains a small portion of the variability in Heart Rate and a very weak influence on a person's stress level.

From the results of varying interpretations about the effects of age to both heart rate and stress levels, one thing for sure is that there are many other factors that influence an individual's vital signs. Each variable's effect on another variable varies too, and also the different levels of statistically significant relationship ranges.