**Analysis Tutorial Overview 5202**

Tareque Muzahid Zamee

**Develop functional R code for getting correlation & interaction patterns among different parameters of heart disease patients**

Every year, about 805,000 Americans have a heart attack and about half of all Americans (47%) have at least 1 of 3 key risk factors for heart disease namely high blood pressure, high cholesterol, and diabetes. By seeing this trend, I formulated some research questions that addressed interaction and correlation patterns among different parameters (such as age, sex, blood pressure, serum cholesterol, fasting blood sugar, etc.) of heart disease patients. I collected the heart disease patient dataset as a CSV file from <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset?resource=download> which is publicly available. I imported the .csv dataset into the R console and installed all the required packages (“tidyverse”, “dplyr”, “ggplot2”, “readr”). After that, I looked through different research papers regarding heart disease patients having the above-mentioned distinct parameters. I gathered some expedient ideas to navigate through my research questions. I successfully made bar graphs, scatter plot, histogram, and logistic regression plot by generating functional R codes from the heart disease dataset. From my plots and graph, I showed correlation between cholesterol and blood pressure based on age and sex, heart disease frequency based on age and sex, distribution of resting blood pressure and resting ECG (Electrocardiography) based on age and sex, and survival rate as a heart disease patient based on age and sex. I concluded that men are more vulnerable than women as Heart Disease Patients (who are experienced with different other diseases) and mid-age (40-60) heart disease patients are more prone to experience different other diseases. I proposed in my case study that managing stress and a controlled diet are the most important factors for maintaining a healthy life as a heart disease patient. Finally, all the functional R codes are saved into .R file and all the required files are shared at <https://github.com/tmzamee/mytutorial5202> in a public format.

**References**

1. Zhang, Y., Vittinghoff, E., Pletcher, M. J., Allen, N. B., Zeki Al Hazzouri, A., Yaffe, K., Balte, P. P., Alonso, A., Newman, A. B., Ives, D. G., Rana, J. S., Lloyd-Jones, D., Vasan, R. S., Bibbins-Domingo, K., Gooding, H. C., de Ferranti, S. D., Oelsner, E. C., & Moran, A. E. (2019). Associations of Blood Pressure and Cholesterol Levels During Young Adulthood With Later Cardiovascular Events. Journal of the American College of Cardiology, 74(3), 330–341. <https://doi.org/10.1016/j.jacc.2019.03.529>

2. Danja Sarink, Lee Nedkoff, Tom Briffa, Jonathan E Shaw, Dianna J Magliano, Christopher Stevenson, Haider Mannan, Matthew Knuiman, Joseph Hung, Graeme J Hankey, Paul Norman, Anna Peeters, Trends in age- and sex-specific prevalence and incidence of cardiovascular disease in Western Australia, European Journal of Preventive Cardiology, Volume 25, Issue 12, 1 August 2018, Pages 1280-1290, <https://doi.org/10.1177/2047487318786585>

3. Špinar, J. (2012). Hypertension and ischemic heart disease. Cor et Vasa, 54(6), e433-e438. <https://doi.org/10.1016/j.crvasa.2012.11.002>

4. Blokh, D., & Stambler, I. (2015). Information theoretical analysis of aging as a risk factor for heart disease. Aging and disease, 6(3), 196–207. <https://doi.org/10.14336/AD.2014.0623>

5. Vaartjes, I., Hoes, A., Reitsma, J. et al. Age- and gender-specific risk of death after first hospitalization for heart failure. BMC Public Health 10, 637 (2010). <https://doi.org/10.1186/1471-2458-10-637>

6. Bots, S. H., Peters, S. A. E., & Woodward, M. (2017). Sex differences in coronary heart disease and stroke mortality: a global assessment of the effect of ageing between 1980 and 2010. BMJ global health, 2(2), e000298. <https://doi.org/10.1136/bmjgh-2017-000298>

7. Masenga, S. K., & Kirabo, A. (2023). Hypertensive heart disease: risk factors, complications and mechanisms. Frontiers in cardiovascular medicine, 10, 1205475. <https://doi.org/10.3389/fcvm.2023.1205475>

8. Rodgers, J. L., Jones, J., Bolleddu, S. I., Vanthenapalli, S., Rodgers, L. E., Shah, K., Karia, K., & Panguluri, S. K. (2019). Cardiovascular Risks Associated with Gender and Aging. Journal of cardiovascular development and disease, 6(2), 19. <https://doi.org/10.3390/jcdd6020019>

9. Jung, E., Kong, S. Y., Ro, Y. S., Ryu, H. H., & Shin, S. D. (2022). Serum Cholesterol Levels and Risk of Cardiovascular Death: A Systematic Review and a Dose-Response Meta-Analysis of Prospective Cohort Studies. International journal of environmental research and public health, 19(14), 8272. <https://doi.org/10.3390/ijerph19148272>

10. <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset?resource=download>

11. R for Data Science. Wickham and Grolemund. 2023 (<https://r4ds.hadley.nz/>)

12. R Core Team. 2024. R: A Language and Environment for Statistical Computing.  
<https://www.r-project.org/>