International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 02 | February - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Analysis of Fitness Based on Smart Watch Data

Bijoy Kumar Dey
Department of Statistics
University of Calcutta
Baruli, Dakshin Gobindapur,Kolkata: 700145, India
bijoykumardey017@gmail.com

Ritika Khan
Department of Statistics
University of Calcutta
73, B, Baghbazar Street, Kolkata:700003, India
ritikakhan26471@gmail.com

Mainak Kunai
Department of Statistics
University of Calcutta
Sreenagar, Ghola, Sodepur, Panihati(m), Kolkata: 700111, India
mainakunai98@gmail.com

***Abstract*— Smart watch, a new trend to many young as well as aged persons now a days, can serve as a fitness tracker and be far more accurate than a phone. Besides many technological advantages, a smartwatch can easily collect the data related to fitness and other movements also. Smart watch provides such an organized human health and activities dataset so that by this dataset a human activities analysis can be performed and by this analysis a conclusion can be drawn about their health. Accordingly, our Project aim is to analyze the data and examine whether the data collected through smart watches can describe the physical fitness of individuals on the basis of some of the factors related to daily activities.**

Keywords*— *Data Visualization, Correlation plot, Scatter plot, MANOVA, Fisher Exact Test, P-value

I. INTRODUCTION

Nowadays, the rise of smart watches is revolutionizing healthcare in an entirely new way by allowing us to take charge of our physical and mental well-being right from our wrists. Smart watches have the potential to support health in everyday living by enabling self-monitoring of personal activity, obtaining feedback based on activity measures and supporting bi-directional communication with health care providers and family members. However, the smart watches, an emerging technology and research, are at a nascent stage. These wearable devices are quickly becoming indispensable technologies in the healthcare industry. Smart watches can monitor sleep habits and physical activities in our context. These also encourage active and healthy lifestyle of every person. In this discussion, our intension is to verify how well the health related data of smart watches can explain fitness of a person in some extent.

II. DATA DESCRIPTION

The information of our interest has been extracted from the csv file named "pmdata_daily_activity_merged.csv".

The above file contains all of the participants' cleaned and properly formatted daily data from November 1, 2019 to March 31, 2020. This file includes daily activities, sleep data as well as personal information about each of the 16 participants, including age, gender and height.

Furthermore, another csv file has been taken under consideration named "pmdata_sleep_unedited.csv" which contains the data related to sleep because the raw sleep data contains a wealth of information that may be useful to anyone interested in studying sleep in greater depth, so unedited sleep data is included for all the participants.

The underlined dataset in this consideration consists of the following information:

For each of the 16 participants, participant id wise their daily data are given as;

- **Age:** Age of each participant is given to analyse the data age-group wise.
- **Steps:** Total steps, each participant can walk per day, is taken into consideration.
- **Distance (in km):** Total distance (in km), each participant can cover per day, is also included.
- **Calories Burn:** Measurement of Calories burnt by each participant on daily basis should be considered.
- **Very Active Minutes:** This indicates the time (in minutes) per day a participant is highly active.
- **Moderately Active Minutes:** This indicates the time (in minutes) per day a participant is moderately active.
- **Lightly Active Minutes:** This indicates the time (in minutes) per day a participant is lightly active.
- **Sedentary Minutes:** This specifies the time (in minutes) per day a participant is sitting, reclining or lying down.
- **Minutes Asleep:** This suggests the time (in minutes) per day a participant is asleep.
- **Time in Bed (in Minutes):** This specifies the time (in minutes) spent in bed by each of the participants per day.
- **Main Sleep:** This is a categorical variable which depicts a participant is having main sleep or not.

III. METHODOLOGY

Age is enormously associated with physical fitness. With the increase in the age, muscle strength, bone density starts to decrease continuously, besides that the body composition shifts from muscle to fat and these severely affect the overall health. Research has repeatedly shown that necessary precautions in older adults is linked with lower disease risk, reduced risk of falls and improved overall health. Therefore, it becomes essential to split the data with respect to different age groups namely '21-30', '31-40', '41-50' and '51-60'.

Graphic visual representation of information is a crucial component in understanding and identifying patterns and trends in data. It encompasses a wide range of techniques to clarify, interpret and analyze the data by drawing line segments or plotting several points on the graph. Accordingly, in this discussion several graphs are plotted viz. bar diagrams, line diagrams and scatter plots.

In addition to the above, several correlation plots have also been drawn to graphically display the correlation

matrices which also provide p-values and confidence intervals to help in determining the statistical significance of the correlations. Primarily, a correlation plot has been drawn among the variables viz. Height, Steps, Distance (in km), Calories Burn, Sedentary minutes, Minutes Asleep and Active minutes to identify which pair of variables are highly correlated with each other. Furthermore, several correlation plots among some specified variables viz. Steps, Distance (in km), Calories Burn, Very Active minutes, Moderately Active minutes, Lightly Active minutes, Sedentary minutes and Minutes Asleep for each of the four age-groups (21-30, 31-40, 41-50, 51-60) have been used to recognize the highly correlated variables for different age groups.

Sleep has a crucial role in physical fitness. Sleep gives our body enough time to recover, conserve energy, repair and build up the muscles worked during exercise. For that reason, it's become vital to analyze the sleep related data.

In the first place, each participant has been taken into consideration to check whether (s)he has Insomnia or not. According to some Articles, if a person has sleep efficiency less than or equals to 85, then (s)he is diagnosed with Insomnia. However, if the sleep efficiency is more than 85, (s)he will not be an insomnia patient. Comparing with this standard range, the occurrence of insomnia has been checked for each patient.

Moreover, Fisher's Exact test has been used to determine if there is any non-random association between the two categorical variables namely Insomnia (which depicts a participant has insomnia or not) and Main sleep.

Last but not the least Multivariate Analysis of Variance (MANOVA) has been performed to determine multiple levels of independent variables on their own or in combination with another have an effect on the dependent variables. In this discussion, our interest is to identify the effect of different age-groups on several dependent variables such as, Steps, Distance, Calories Burn, Active Minutes, Sedentary Minutes and Minutes Asleep. Furthermore, it is necessary to recognize how insomnia affects the same dependent variables viz. Steps, Distance, Calories Burn, Active Minutes, Sedentary Minutes and Minutes Asleep.

IV. Result

After implementing all the procedures mentioned in Methodology, the following consequences had arisen:

Before anything else several graphs, which were appropriate for the analysis, were drawn to identify the patterns of the data. The graphical representation began with a Line Diagram (Fig 1) which exhibited a nonlinear decreasing relationship between Age-group and Active Minutes.

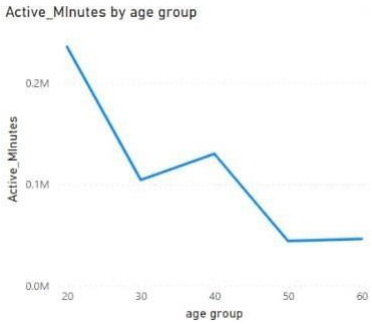

Figure 1: Line Diagram of age group with active minutes

Followed by this, a Scatter Plot (Fig 2) had also been drawn which depicted a linear increasing relation between Steps and Calories Burn.

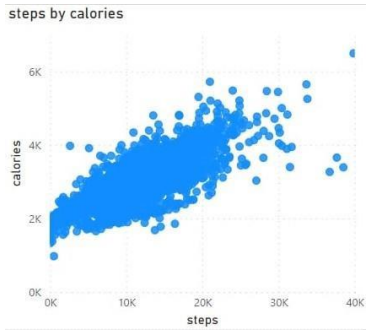

Figure 2: Scatter Plot of Steps vs Calories Burn

After that the tendency of Insomnia was displayed by different age-groups through a Bar Diagram (Fig 3).

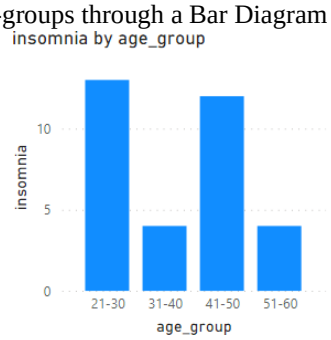

Figure 3: Bar diagram of Age Group vs Insomnia

To identify the relationships between Steps with Very Active minutes, Steps with Moderately Active minutes and Steps with Lightly Active minutes, three Scatter plots had been drawn.

Scatter plot (Figure 4) of Steps vs Very Active minutes depicted an increasing relationship between them.

Figure 4: Scatter Plot of steps vs very active minutes

Scatter plot (Fig 5) of Steps with Moderately Active minutes depicted a sort of increasing relationship between them but not exactly.

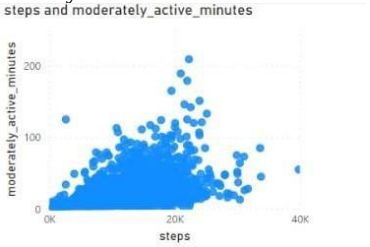

Figure 5: Scatter Plot of steps vs active minutes

Scatter plot (Fig 6) of Steps with Lightly Active minutes depicted an increasing relationship between them.

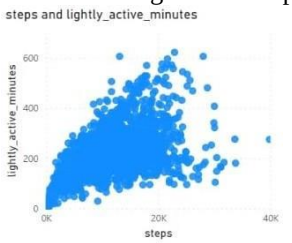

Figure 6: Scatter plot of Steps vs Lightly Minutes

Additionally, several correlation plots resulted in various interrelationships among the variables. Primarily, a correlation plot among Heights, Steps, Distance (in km), Calories Burn, Sedentary Minutes, Minutes Asleep, and Active Minutes was drawn to graphically interpret the correlations among the variables. The observations collected from the correlation plot (Fig 7) are as follows:

- The variable 'Steps' is highly positively correlated with 'Distance (in km)' as well as 'Calories Burn' and 'Active Minutes'.
- The variable 'Distance (in km)' is highly positively correlated with 'Calories Burn' as well as 'Active Minutes'.
- The variable 'Calories Burn' is highly positively correlated with 'Active Minutes'.
- On contrary, the variables 'Steps', 'Distance (in km)', 'Minutes Asleep' and 'Active Minutes' are highly negatively correlated with 'Sedentary Minutes'.

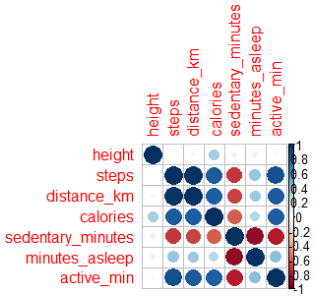

Figure 7: Correlation plot before classification of age

Besides that, 4 other correlation plots had been accounted for comparing the variables namely Steps, Distance (in km), Calories Burn, Very Active Minutes, Moderately Active Minutes, Lightly Active Minutes, Sedentary Minutes and Minutes Asleep for each of the 4 age-groups viz. '21-30', '31-40', '41-50', '51-60' which are as follows:

- For '21-30' age group (Fig 8), 'Steps' and 'Distance' are highly positively correlated, in the similar way, 'Very Active Minutes' is positively correlated with 'Steps', 'Distance (in km)' and 'Calories Burn'. On the other hand, 'Sedentary Minutes' and 'Minutes Asleep' are highly negatively correlated.

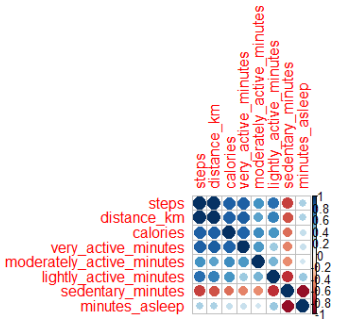

Figure 8: Correlation plot for 21-30 age group

- For '31-40' age group (Fig 9), 'Steps' is highly positively correlated with 'Distance (in km)' and 'Lightly Active Minutes'. Similarly, 'Distance (in km)' is highly positively correlated with 'Lightly Active Minutes'. In contrast, 'Sedentary Minutes' is highly negatively correlated with 'Lightly Active Minutes' and 'Minutes Asleep'.

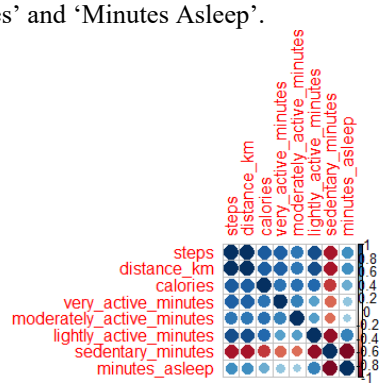

Figure 9: Correlation plot for 31-40 age group

- For '41-50' age group (Fig 10), 'Steps' and 'Distance (in km)' are highly positively correlated with each other. However, 'Sedentary Minutes' and 'Minutes Asleep' are negatively correlated with each other.

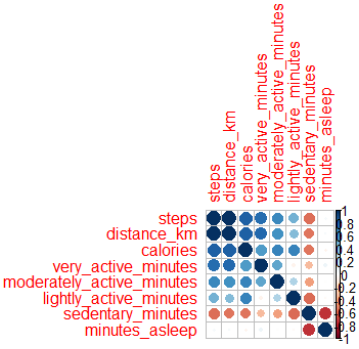

Figure 10: Correlation plot for 41-50 age group

- For '51-60' age group (Fig 11), 'Steps' and 'Distance (in km)' are highly positively correlated with 'Calories Burn', yet 'Sedentary Minutes' is highly negatively correlated with 'Minutes Asleep'.

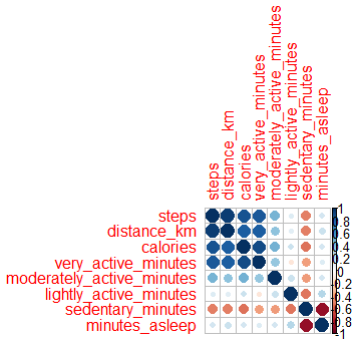

Figure 11: Correlation plot for 51-60 age group

Furthermore, the observation was that the age - group '21-30' had the highest insomnia level, the age-group of '41-50' had the second highest insomnia level and the next level was captured by both the age-groups '31-40' and '51-60'.

One more observation was that for '21-30' age group, with respect to Steps, initially Very Active Minutes increased and after reaching at a certain point it stopped to increase.

To examine whether there is any non-random relationship between Insomnia and Main Sleep, Fisher's Exact Test was performed which produced the p-value as 1.305e-6 (i.e., <0.05). This implies the null hypothesis is rejected at 95% level of significance i.e., the classifications are different which means someone doesn't have main sleep may not imply that (s)he has insomnia.

Additionally, to determine the effect of multiple levels of independent variables on the dependent variables, Multivariate Analysis of Variance (MANOVA) was performed by considering Steps, Distance (in km), Calories Burn, Active Minutes, Sedentary Minutes and Minutes Asleep as dependent variables, further Age-group as independent variable. This resulted in the p-value 2.2e-16 (i.e., <0.05) which implies that the null hypothesis is rejected i.e., age-group wise the difference among the dependent variables is significant at 95% level of significance. Similarly, the MANOVA considering Steps, Distance (in km), Calories Burn, Active Minutes, Sedentary Minutes and Minutes Asleep as dependent variables and insomnia as independent variable produced the p-value 2.6e-15 (i.e., <0.05) which indicates that the null hypothesis is rejected at 95% level of significance i.e., insomnia affects the dependent variables.

V. Conclusion

The motivation of this analysis is how the information related to daily activities of various individuals collected by smart watches can explain the physical fitness of those individuals with respect to some of the fitness indicators. According to the objective of the study, appropriate methods had been followed to reach at a conclusion. From several graphical figures it can be inferred that with the increase in age, the time when the participants are active has a decreasing tendency. In the similar manner, as steps of the participants increases the calories burn as well as the very active minutes both start to increase. In addition, irrespective of any age-group the physical fitness indicators viz. steps, calories burn, distance, active minutes, very active minutes, moderately active minutes and lightly active minutes have close positive interrelationship within themselves. In contrast, the relationship between sedentary minutes and the time of sleeping is negative. Afterwards focusing on insomnia, it is prominent that insomnia mostly affects the 21-30 and 41-50 age groups; therefore, it can be interpreted as age-group has an effect on insomnia, on contrary main sleep has no relation with insomnia. Besides that, age group has effect on the physical fitness related variables viz. steps, distance, calories burn, active (very, moderately, lightly) minutes. As a result, it is obvious that age group wise insomnia affects physical fitness. Summarizing all the interpretations, it can be concluded that according to age groups, the data collected by smart watches can explain the physical fitness of individuals in the perspective of some of the daily physical activities.

Author Contributions

B.K. Dey and R. Khan are mainly responsible for the Statistical Analysis and Interpretation. M. Kunai is responsible for the establishment of the Data Visualization portion. The writing of the paper is mainly completed by R. Khan. All the authors equally contributed to the article and approved the submitted version.

References

[1] Vicente Martínez-Vizcaínoa, Mairena Sánchez-Lópeza (2008), "Relationship Between Physical Activity and Physical Fitness in Children and Adolescents" (vol. 61, Issue 2), Revista Española de Cardiología, 108-111

[2] David R. Bassett, Jr.,corresponding author Lindsay P. Toth, Samuel R. LaMunion, Scott E. Crouter, "Step Counting: A Review of Measurement Considerations and Health-Related Applications", Sports Med. 2017; 47(7): 1303–1315.

[3] David L. Reed, William P. Sacco, "Measuring Sleep Efficiency: What Should the Denominator Be?", J Clin Sleep Med. 2016 Feb 15; 12(2): 263–266.

[4] Sergio Garbarino, Paola Lanteri, Nicola Luigi Bragazzi (2021), Nicola Magnavita & Egeria Scoditti, "Role of sleep deprivation in immune-related disease risk and outcomes", Commun Biol 4, 1304 (2021). https://doi.org/10.1038/s42003-021-02825-4

[5] Louisa G. Sylvia, Emily E. Bernstein, Jane L. Hubbard, Leigh Keating, Ellen J. Anderson, "A Practical Guide to Measuring Physical Activity" J Acad Nutr Diet. 2014 Feb; 114(2): 199–208.

[6] Kirstie N. Anderson, "Insomnia and cognitive behavioural therapy—how to assess your patient and why it should be a standard part of care", Vol 10, Supplement 1 (January 2018): Journal of Thoracic Disease (Dedicated to the 3rd Clinical Update Sleep: International Conference 2018, Endorsed by ERS)

[7] Maciej Kochman, Wojciech Kasperek, Agnieszka Guzik, Mariusz Drużbicki, "Body Composition and Physical Fitness: Does This Relationship Change in 4 Years in Young Adults?", Int J Environ Res Public Health. 2022 Feb; 19(3): 1579.

[8] Liu H, Liu Y and Li B (2021) Predictive Analysis of Health/Physical Fitness in Health-Promoting Lifestyle of Adolescents. Front. Public Health 9:691669. doi: 10.3389/fpubh.2021.691669