6ZG-01

Understanding of Clinical Growth Charecterestics of Males and Females from Remote Healthcare Data

Mehdi Hasan[1], Fumihiko Yokota[2], Akira Fukuda[1] and Ashir Ahmed[1]

[1]Faculty of Information Science and Electrical Engineering, Kyushu University, Japan

[2]Institute of Decision Science for a Sustainable Society, Kyushu University, Japan

{mehdi, fukuda}@f.ait.kyushu-u.ac.jp, yokota.fumihiko.785@m.kyushu -u.ac.jp,

ashir@ait.kyushu u.ac.jp

Abstract

CDC (Center for Disease Control and Prevention) reports a clinical growth chart for boys and girls. However, it limits until the age of 2-20. The chart shows a clear uptrend of three anthropometric data (height, weight and BMI). How about people with >20 years of age? In this paper, we investigate from our portable health clinic data to observe the other anthropometric data (height, weight, BMI, waist, hip) for the age >20 years. Our chart shows the clinical growth characteristics depending on the age and gender. This paper reports the findings analyzed from 40,391 checkup data collected by a remote healthcare system (we call it portable health clinic) from more than 54 locations for more than 8 years of observations in both rural and urban areas in Bangladesh.

Keywords

Remote healthcare system, clinical growth characteristics.

1. Introduction

We observed the anthropometric data (height, weight, BMI, waist, hip). It was not an easy task to analyze the raw data. The data was collected from remote areas and was digitized by human being. There were unintentional errors, some fields were empty and some fields had impossible data. We first removed the incomplete records. Based on our common sense, we detected the impossible data and also removed them from our potential records [3]. The number of records of the raw data was around 40,391. The data has been collected from 54 locations of both rural and urban areas in Bangladesh since 2010. After eliminating the incomplete, doubtful data, inconsistent. We found 22,721 records.

2. Methodology

In this research, first, we selected our experimental data. From the raw data, eliminated the Incorrect format data, Missing data, Out of normal range and Inconsistent data. After that, we calculated maximum, minimum, average and standard deviation for both gender.

2.1 Selection criteria of data

In portable health clinic, we can measure the health status by collecting 23 anthropometric items. In this study we considered 40,391 anthropometric data (male=25,304 and female=15,087) and 6 anthropometric.

2.2 Clean unusual data

Data were screened carefully to get the processed data which has been statistically to ensure the clusters by the value of Pearson correlation coefficient. In particular, 17,670 unprocessed data have been removed from the data set by the concept of existing acceptance range. The study also followed level of data entry errors to get the characteristics of removed data were — Incorrect format data, Missing data, Out of normal range and Inconsistent data. From these data, we cleaned the unusual data based on the following criteria,

Criteria 1: Remove the data from age, where age<20 & age>100

Criteria 2: Remove the data from

height, where height<140 & height>185

 Criteria 1: Remove the data from weight, where weight<40 & weight>145

 Criteria 1: Remove the data from BMI, where bmi<11 & bmi>55

 Criteria 1: Remove the data from waist, where waist<30 & waist>160

 Criteria 1: Remove the data from hip, where hip<35 & hip>170

3. Apply Pearson Correlation coefficient

 The Pearson correlation coefficient measures the strength between variables and relationships.

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - ([\sum x)^2][N\sum (y)^2 - ([(y)^2]}}$$

where:

N = number of pairs of scores

$\sum xy$ = sum of the products pf paired scores

$\sum x$ = sum of x scores

$\sum y$ = sum of y scores

$\sum x^2$ = sum of squared x scores

$\sum y^2$ = sum of squared y scores

 In order to determine, the relationship is between two variables, formula must be followed to produce what is referred to as the coefficient value. The coefficient value can range between -1.00 and 1.00. If the coefficient value is in the negative range, then that means the relationship between the variables is negatively correlated, or as one value increases, the other decreases.

4. Findings

 Remote healthcare involves physicians, patients, academics, health service organization, and industries. All these stakeholders have different expectations about remote healthcare development. Governments want to cut costs, industries want to make business, patients want better healthcare, and physicians want an easier and faster way to get.

TABLE 1: Clinical growth chart for male and female

	Maximum	Minimum	Mean	SD
Age (M)	20	100	41.6	12.7
Age (F)	20	96	41.8	12.3
Height(M)	135	185	163.5	6.3
Height(F)	135	184	151.5	5.6
Weight(M)	40	120	60.7	11.2
Weight(F)	40	112	54.3	10..1
BMI (F)	14	42	22.68	3.6
BMI (F)	15	46	23.8	4.0

1. Conclusion

 This paper reports the findings analyzed from 40,391 checkup data. We took the statistical approach to find the patterns. We took the mean for every single age from 20 to 100 and plotted them. The plots were not smooth to draw a conclusion. From our eyeball measurements, we classified the patterns into three different age groups.

Acknowledgement

References

1. Data Cleaning: Detecting, Diagnosing, and Editing Data Abnormalities. Available at, http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020267 (last accessed on 14.4.2018).

2. Hasan M, Fukuda A, Maruf R, Yokota F, Ahmed A. Errors in remote healthcare system: Where, how and by whom?. TENCON 2017 - 2017 IEEE Region 10 Conference. 2017. doi:10.1109/tencon.2017.8227856.

3. Pearson correlation coefficient: formula, example & significance: https://study.com/academy/lesson/pearson-correlation-coefficient-formula-example-significance.html ((last accessed on 1.12.2018).