

DATA ANALYSIS FOR RELEVANCE OF THE PHYSICAL TRAINING APPLICATION

Ruzhitskyi S. V.

Academic supervisor – Prof. Dr. Sc. Igor Grebennik

Kharkiv National University of Radio Electronics, System Engineering Dept
Kharkiv, Ukraine

e-mail: stanislav.ruzhitskyi@nure.ua

Main goal of this work is to explore the data on people's physical activity, as a result of the research we should draw a conclusion about the relevance of developing a system for collecting and providing access to physical training programs. The dataset which was been used during the research was given from the Behavioral Risk Factor Surveillance System. From dataset was taken all necessary data to the research and was used to make some analyses and visualization of this data. According to the results of the research was given a conclusion about expediency of creation collecting and providing system.

Before developing the web application to collect and provide access to physical training programs was decided to analyze field of gyms and physical activity of people to decide and answer on questions about relevance of the application.

For the research was found the open datasets on Kaggle resource, which provide the information from the Behavioral Risk Factor Surveillance System, which based on the people's questioner. These datasets contain information about nutrition, physical activity, sleep, obese level, and some information about people.

To conduct the research was decided to use Python programming language and relative libraries. Python was selected due to its greats libraries well known for the data analysis and data visualization, such as pandas, matplotlib, seaborn. Also, research was conduct using Jupyter Notebook as a convenient environment for data analysis.

First of all, data in the datasets was cleaned from the unnecessary columns and fields, which will not usable during the research, also some columns were renamed to more comfortable usage.

Returning to the purpose of the study, let's try to find out is the physical activity of people relates to the body mass index (BMI). So, from the dataset was took the information about the job title of person and BMI category.

	Occupation	Accountant	Doctor	Engineer	Lawyer	Manager	Nurse	Sales Representative	Salesperson	Scientist	Software Engineer	Teacher
BMI Category												
Normal	26.0	65.0	56.0	42.0	NaN	NaN		NaN	NaN	NaN	NaN	6.0
Normal Weight	5.0	2.0	4.0	1.0	NaN	7.0		NaN	NaN	NaN	2.0	NaN
Obese	NaN	4.0	NaN	2.0	NaN	NaN		2.0	NaN	NaN	1.0	1.0
Overweight	6.0	NaN	3.0	2.0	1.0	66.0		NaN	32.0	4.0	1.0	33.0

Figure 1 – Pivot table for occupation and BMI category

From the figure above we see different job some of them are sedentary another movable. But, from the figure can not say anything, because for example doctors and nurses have same style of job, but by the BMI they are opposite each other.

Then let's see the ratio between physical activity in minutes, daily steps count and BMI category of people, plot below.

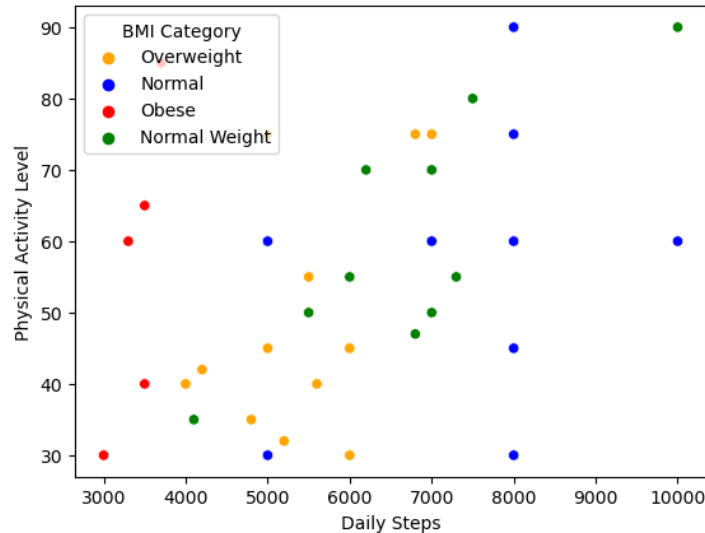


Figure 2 – Plot of activity

From the plot we can see that obese can depend on the people's daily steps count, we clearly see that people with obese have less than 4000 steps, also with more counts of steps people have normal weight, and people with normal weight have more physical activity and steps, with some exceptions, so we need more research.

After this let's calculate mean, median for physical activity in minutes, daily steps count and age of participants and divide it on BMI categories and also output the min and max physical activity and daily steps.

BMI Category	mean	median	min	max	mean	median	min	max	mean	median
Normal	57.692308	60.0	30	90	6887.179487	8000.0	5000	10000	38.482051	37.0
Normal Weight	60.333333	60.0	35	90	6766.666667	7000.0	4000	10000	38.380952	37.0
Obese	55.000000	60.0	30	85	3350.000000	3400.0	3000	3700	38.000000	36.5
Overweight	61.236486	45.0	30	90	6965.540541	6000.0	4000	10000	47.885135	45.0

Figure 3 – Calculation results

According to calculation we can see that people with obese have less average activity per day compared to other also they have less steps count, but we can highlight that people with overweight have mean activity and steps even more than the people with normal weight, but for this we also calculated the median, which point that median time of activity for this people is less than

other, and their median steps count less than people’s with normal weight. Also, it can be due to people with overweight average age older than other participants.

For now, we can say that physical activity is important for people to avoid overweight, and for now we need to answer if the application for training will be popular among the people that want to train.

To answer on this question let’s look at another dataset, which is represent results of the survey of people in the USA in the different years, that demonstrate their weight level and activity.

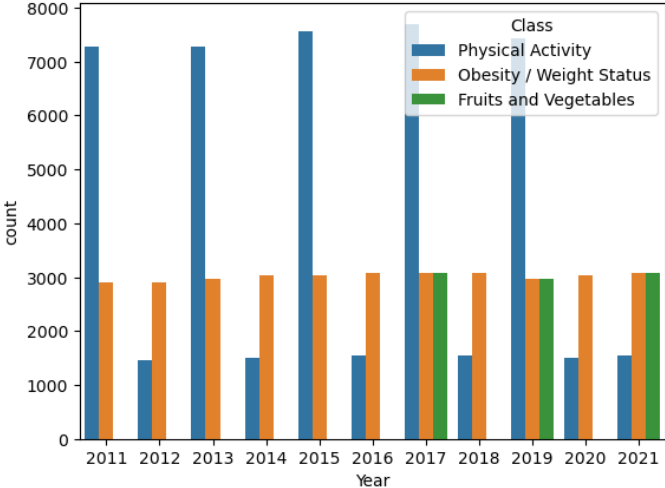


Figure 4 – Plot represent count of people, their obesity and activity

According to plot below we can see that count of people with obesity was less than people that have healthy lifestyle and high level of physical activity and this trend was in one year, but then these people again was more, but we see that in 2020 and 2021 two years in a row was more people with obesity, and the cause of this can be COVID pandemic and lockdown of gyms.

So making a conclusion, we can say that trainings and high level of physical activity is important and helpful to people to avoid overweight, also we can say that the web application that allows people to train at home will be popular through the people, because a lot of gyms still don’t recover their activity after lockdown, but people need something instead, also the application will be useful, because people will not need to waste a time on a trip to gym.

Resources:

1. Nutrition, Physical Activity, and Obesity. [Dataset]. <https://www.kaggle.com/datasets/mattop/nutrition-physical-activity-and-obesity/data>. 2022.
2. Sleep, Health, and Lifestyle Dataset. [Dataset]. <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>. 2022.
3. Petrova R., Syrotenko O. Comparative analysis of application programming interface architectural solutions.