**Diagrams and plots implemented by code**

1. Зображення, що містить текст, знімок екрана, Шрифт, дизайн

   Автоматично згенерований опис Frequency table: it divides all amount of people by age and shows us how many people are in each formed group:

Frequency table

1. Histogram of age frequencies: this histogram visualizes data from previous frequency table.

Зображення, що містить схема, квадрат, текст, Графік

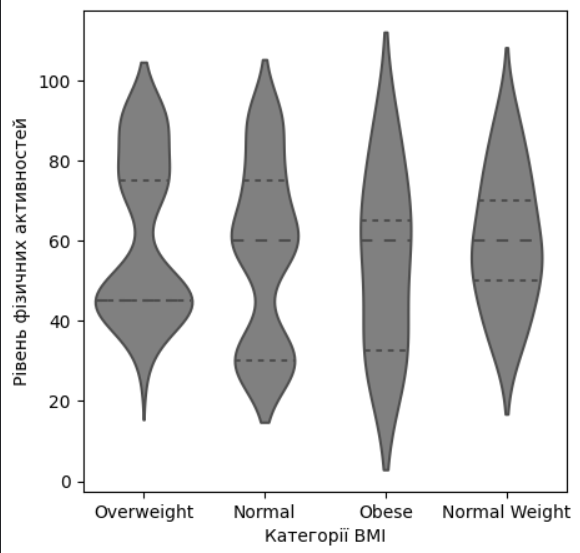
Автоматично згенерований опис

Histogram of age frequencies

1. Зображення, що містить знімок екрана, схема, дизайн

   Автоматично згенерований описContour plot: represent relationship between age of people and duration of sleep.

Contour plot

1. Violin plot: this type of plot combines aspects of a box plot and a density plot and shows the distribution of physical activity levels across different BMI categories.

Violin plot

1. Зображення, що містить схема, текст, знімок екрана, Прямокутник

   Автоматично згенерований описBox plot: this type of plot is used to display the distribution of sleep duration across different occupations based on five-number summary: minimum, first quartile, median, third quartile and maximum.
2. Correlation matrix: shows the correlation coefficients between multiple variables in a dataset: age, physical activity level, stress level, daily steps and sleep duration. Coefficient ranges from -1 to 1: if it`s positive, then as one variable increases, the other variable increases too; if it`s negative, then as one variable increases, the other tends to decrease. If coefficient is equal to 0, then there is no linear relationship between variables.

Зображення, що містить текст, знімок екрана, квадрат, схема

Автоматично згенерований опис

**A/B testing**

Conducted an A/B test based on the following hypotheses:

Null Hypothesis: The average sleep duration does not change depending on the physical activity level.  
Alternative Hypothesis: The average sleep duration differs depending on the physical activity level.

To estimate the test, I have used the physical activity level (hereinafter referred to as “pal”) scores ranging from 0 to 100 and have divided group of people into those who sleep less than 6 hours (low\_sleep\_group) and those who sleep 6 hours and more (high\_sleep\_group).

To visualize the data I have used a box plot, which shows that people who sleep less than 6 hours are less physically active than those who sleep more than 6 hours. Comparing the mean values, determined that the difference in pal between these two groups is 18.8.

Зображення, що містить текст, знімок екрана, ряд, схема

Автоматично згенерований опис

After that I conducted a permutation test where duration of sleep was not considered: the test involved multiple permutations of new groups and computing the differences in mean values foe each permutation. This resulted in distribution of permutation differences, from which the p-value was calculated. Based on this value, it was made a conclusion that the average level of physical activity varies depending on a person`s sleep duration.

p-value: 0.021 (Since it`s less than alpha-value (alpha-value = 0.05) than we reject H0 hypothesis).

Reject H0: The average level of physical activity varies depending on sleep duration.

Зображення, що містить знімок екрана, схема, Графік, ряд

Автоматично згенерований опис