

Python Case Study

Descriptive Statistics Question

Cardio Good Fitness

# Data Details:

* **Product:** Self Explanatory **(**string)
* **Age**: Self Explanatory **(**integer)
* **Gender:** Self Explanatory **(**string)
* **Education:** Self Explanatory **(**integer)
* **MaritalStatus:** Self Explanatory **(**string)
* **Usage:** Self Explanatory **(**integer)
* **Fitness:** Self Explanatory **(**integer)
* **Income:** Self Explanatory (integer)
* **Miles:** Self Explanatory(integer)

# Part-1: Descriptive Analysis Questions

1. Which product is most used?
2. What is the average age of customers, is it same as median? Justify your answer by plotting distribution of age.
3. How many males have income greater than average income? Also check whether income is positive or negatively skewed?
4. Find ratio of males and females
5. What percent of unmarried men have fitness level less than or equal to 3?
6. What percent of married women have fitness level greater than 3? And what may be the reason behind that?
7. How many unmarried men have used treadmill less than 5 times in a week?
8. Give your suggestion, if there were null values in feature “Miles” then using which central tendency we can impute it? Justify your answer.
9. Which product is least used by unmarried customers?
10. If there were null values in feature “Education” using which measure of center can we impute it?
11. Fitness is most related to which feature?
12. Find gender-wise average age
13. Does Education have high correlation with Income? If yes how can you practically relate it?
14. Is there any negatively skewed feature?