Assignment: Esoph and Young People Survey

MEF - BDA 503

Nov 21, 2017

Questions below ask for your insight about the data. There is no single and constructed way to the right answer. Objective of this assignment is to help you help yourself to get data from outside resources, analyze the data, validate your model and convey your conclusions in a clear and reproducible way. You are expected to use R Markdown outputs on your Progress Journals to show your work.

1. Use `esoph` data set to see if (o)esophageal cancer is related to alcohol consumption, age or tobacco consumption. (Just write `esoph` to your R console to get the data.)

2. Use the Young People Survey's Hobbies & Interests category answers to infer "meaning" from the data. You are expected to use methods described in the lecture notes. But also you are welcome to use different methods as well. You can get the data from Kaggle.

    1. History: Not interested 1-2-3-4-5 Very interested (integer)
    2. Psychology: Not interested 1-2-3-4-5 Very interested (integer)
    3. Politics: Not interested 1-2-3-4-5 Very interested (integer)
    4. Mathematics: Not interested 1-2-3-4-5 Very interested (integer)
    5. Physics: Not interested 1-2-3-4-5 Very interested (integer)
    6. Internet: Not interested 1-2-3-4-5 Very interested (integer)
    7. PC Software, Hardware: Not interested 1-2-3-4-5 Very interested (integer)
    8. Economy, Management: Not interested 1-2-3-4-5 Very interested (integer)
    9. Biology: Not interested 1-2-3-4-5 Very interested (integer)
    10. Chemistry: Not interested 1-2-3-4-5 Very interested (integer)
    11. Poetry reading: Not interested 1-2-3-4-5 Very interested (integer)
    12. Geography: Not interested 1-2-3-4-5 Very interested (integer)
    13. Foreign languages: Not interested 1-2-3-4-5 Very interested (integer)
    14. Medicine: Not interested 1-2-3-4-5 Very interested (integer)
    15. Law: Not interested 1-2-3-4-5 Very interested (integer)
    16. Cars: Not interested 1-2-3-4-5 Very interested (integer)
    17. Art: Not interested 1-2-3-4-5 Very interested (integer)
    18. Religion: Not interested 1-2-3-4-5 Very interested (integer)
    19. Outdoor activities: Not interested 1-2-3-4-5 Very interested (integer)
    20. Dancing: Not interested 1-2-3-4-5 Very interested (integer)
    21. Playing musical instruments: Not interested 1-2-3-4-5 Very interested (integer)
    22. Poetry writing: Not interested 1-2-3-4-5 Very interested (integer)
    23. Sport and leisure activities: Not interested 1-2-3-4-5 Very interested (integer)
    24. Sport at competitive level: Not interested 1-2-3-4-5 Very interested (integer)
    25. Gardening: Not interested 1-2-3-4-5 Very interested (integer)
    26. Celebrity lifestyle: Not interested 1-2-3-4-5 Very interested (integer)
    27. Shopping: Not interested 1-2-3-4-5 Very interested (integer)
    28. Science and technology: Not interested 1-2-3-4-5 Very interested (integer)
    29. Theatre: Not interested 1-2-3-4-5 Very interested (integer)
    30. Socializing: Not interested 1-2-3-4-5 Very interested (integer)
    31. Adrenaline sports: Not interested 1-2-3-4-5 Very interested (integer)
    32. Pets: Not interested 1-2-3-4-5 Very interested (integer)