

Machine Learning (online)

Coursework

2. Coursework resources

2.1. Part 1

As a reminder:

In the first part you will be given data collected from various individuals on several variables. The goal will be to use unsupervised learning techniques such as Principal Component Analysis or Cluster Analysis to summarise the information in the data by appropriate tables and/or plots.

Description for the EWCS dataset

Task:

Visualise and describe the data via unsupervised learning methods.

Source:

European Working Conditions Survey 2016.

Variable and value labels for EWCS 2016 dataset:

Q2a - Gender

Values 1: Male 2: Female

Q2b - Age

Values numeric

Q87a - I have felt cheerful and in good spirits [...which is the closest to how you have been feeling over the last two weeks]

Q87b - I have felt calm and relaxed [...which is the closest to how you have been feeling over the last two weeks]

Q87c - I have felt active and vigorous [...which is the closest to how you have been feeling over the last two weeks]

Q87d - I woke up feeling fresh and rested [...which is the closest to how you have been feeling over the last two weeks]

Q87e - My daily life has been filled with things that interest me [...which is the closest to how you have been feeling over the last two weeks]

Values for variables Q87a to Q87e

1. All of the time.
2. Most of the time
3. More than half of the time
4. Less than half of the time
5. Some of the time

6. At no time

Q90a - At my work I feel full of energy [Please tell me how often you feel this way...]

Q90b - I am enthusiastic about my job [Please tell me how often you feel this way...]

Q90c - Time flies when I am working [Please tell me how often you feel this way...]

Q90f - In my opinion, I am good at my job [Please tell me how often you feel this way...]

Values for variables Q90a to Q90f

1. Always.
2. Most of the time
3. Sometimes
4. Rarely
5. Never