# **Describe all features in the depression dataset**

1. Q1 🡺 Q42  
   survey questions and stored in multiple phases
   1. QxA 🡺 is the stored answer from 1 to 4 [**important**]
   2. QxE 🡺 the time taken until the user answered the question [not important]
   3. QxI 🡺 the question number in the survey [from 1 to 42] [not important]
2. Introelapse 🡺 server side duration [not important]
3. Testelapse 🡺 server side duration [not important]
4. Surveyelapse 🡺 server side duration [not important]
5. TIPI1 🡺 Extraverted, enthusiastic. [**important**]
6. TIPI2 🡺 Critical, quarrelsome. [**important**]
7. TIPI3 🡺 Dependable, self-disciplined. [**important**]
8. TIPI4 🡺 Anxious, easily upset. [**important**]
9. TIPI5 🡺 Open to new experiences, complex. [**important**]
10. TIPI6 🡺 Reserved, quiet. [**important**]
11. TIPI7 🡺 Sympathetic, warm. [**important**]
12. TIPI8 🡺 Disorganized, careless. [**important**]
13. TIPI9 🡺 Calm, emotionally stable. [**important**]
14. TIPI10 🡺 Conventional, uncreative. [**important**]

**NOTE: The TIPI items were rated "I see myself as:" \_\_\_\_\_ such that**

1 = Disagree strongly

2 = Disagree moderately

3 = Disagree a little

4 = Neither agree nor disagree

5 = Agree a little

6 = Agree moderately

1. = Agree strongly
2. The following items were presented as a check-list and subjects were instructed “In the grid below, check all the words whose definitions you are sure you know":
3. VCL1 🡺 boat [not important]
4. VCL2 🡺 incoherent [not important]
5. VCL3 🡺 pallid [not important]
6. VCL4 🡺 robot [not important]
7. VCL5 🡺 audible [not important]
8. VCL6 🡺 cuivocal [not important]
9. VCL7 🡺 paucity [not important]
10. VCL8 🡺 epistemology [not important]
11. VCL9 🡺 florted [not important]
12. VCL10 🡺 decide [not important]
13. VCL11 🡺 pastiche [not important]
14. VCL12 🡺 verdid [not important]
15. VCL13 🡺 abysmal [not important]
16. VCL14 🡺 lucid [not important]
17. VCL15 🡺 betray [not important]
18. VCL16 🡺 funny [not important]
19. Education 🡺 scale from 1 to 4 [**important**]
20. Urban 🡺 "What type of area did you live when you were a child?" [**important**]
21. gender 🡺 "What is your gender?", 1=Male, 2=Female, 3=Other [**important**]
22. engnat "Is English your native language?", 1=Yes, 2=No [**considerable**]
23. age 🡺 "How many years old are you?" [**important**]
24. hand 🡺 "What hand do you use to write with?", 1=Right, 2=Left, 3=Both [not important]
25. religion 🡺 "What is your religion?", 1=Agnostic, 2=Atheist, 3=Buddhist, 4=Christian (Catholic), 5=Christian (Mormon), 6=Christian (Protestant), 7=Christian (Other), 8=Hindu, 9=Jewish, 10=Muslim, 11=Sikh, 12=Other [**important**]
26. orientation 🡺 "What is your sexual orientation?", 1=Heterosexual, 2=Bisexual, 3=Homosexual, 4=Asexual, 5=Other [**considerable**]
27. race 🡺 "What is your race?", 10=Asian, 20=Arab, 30=Black, 40=Indigenous Australian, 50=Native American, 60=White, 70=Other [**important**]
28. voted 🡺 "Have you voted in a national election in the past year?", 1=Yes, 2=No [not important]
29. married 🡺 "What is your marital status?", 1=Never married, 2=Currently married, 3=Previously married [**important**]
30. familysize 🡺 "Including you, how many children did your mother have?" [**important**]
31. major 🡺 "If you attended a university, what was your major (e.g. "psychology", "English", "civil engineering")?" [**important**]
32. country 🡺 [not important]
33. screensize 🡺 [not important]
34. uniquenetworklocation 🡺 [not important]
35. source 🡺 [not important]

# **Input Format for features to predict inside model**

## The features will be explained by order they enter the model

1. From Q1A to Q42A
   1. Values [1, 2, 3, 4]
      1. 1 = Did not apply to me at all
      2. 2 = Applied to me to some degree, or some of the time
      3. 3 = Applied to me to a considerable degree, or a good part of the time
      4. 4 = Applied to me very much, or most of the time
2. From TIPI1 to TIPI10
   1. Values [1, 2, 3, 4, 5, 6, 7]
      1. 1 = Disagree strongly
      2. 2 = Disagree moderately
      3. 3 = Disagree a little
      4. 4 = Neither agree nor disagree
      5. 5 = Agree a little
      6. 6 = Agree moderately
      7. 7 = Agree strongly
3. Education
   1. Values [1, 2, 3, 4]
      1. 1 = Less than high school,
      2. 2 = High school,
      3. 3 = University degree,
      4. 4 = Graduate degree
4. Urban
   1. Values [1, 2, 3]
      1. 1=Rural (country side),
      2. 2=Suburban,
      3. 3=Urban (town, city)
5. Gender
   1. Values [1, 2, 3]
      1. 1=Male,
      2. 2=Female,
      3. 3=Other
6. Religion
   1. Values [1, 2, 3, …, 12]
      1. 1=Agnostic,
      2. 2=Atheist,
      3. 3=Buddhist,
      4. 4=Christian (Catholic),
      5. 5=Christian (Mormon),
      6. 6=Christian (Protestant),
      7. 7=Christian (Other),
      8. 8=Hindu,
      9. 9=Jewish,
      10. 10=Muslim,
      11. 11=Sikh,
      12. 12=Other
7. Race
   1. Values [1, 2, 3, 4, 5, 6, 7]
      1. 1=Asian,
      2. 2=Arab,
      3. 3=Black,
      4. 4=Indigenous Australian,
      5. 5=Native American,
      6. 6=White,
      7. 7=Other
8. Married
   1. Values [1, 2, 3]
      1. 1=Never married,
      2. 2=Currently married,
      3. 3=Previously married
9. Familysize
   1. Values [1, 2, 3, 4, …., 20]
      1. Family size at most is 20 person
10. age\_group
    1. Values [1, 2, 3, 4, 5, 6]
       1. value <= 10: 'Under 10' 🡺 1
       2. 10 <= value <= 16: 'Primary Children' 🡺 2
       3. 17 <= value <= 21: 'Secondary Children' 🡺 3
       4. 21 <= value <= 35: 'Adults' 🡺 4
       5. 36 <= value <= 48: 'Elder Adults' 🡺 5
       6. value >= 49: 'Older People' 🡺 6