

Untitled

Na SeungChan

2022-10-27

Problem 1

```
Dfp1 <- read.csv("./sample_df.csv")
```

(a)

```
P1sample <- Dfp1 %>%  
  select(s100_1)  
  
num_trials <- 1500  
  
P1sample %>%  
  slice_sample(n = 100, replace = TRUE)
```

```
##           s100_1  
## 1      1.02660975  
## 2     -0.40940346  
## 3     -0.20754711  
## 4     -1.00799375  
## 5     -0.76771876  
## 6     -0.74767084  
## 7     -0.01761712  
## 8       0.34586903  
## 9     -0.01155405  
## 10    -0.32126615  
## 11    -0.14877718  
## 12     1.35812386  
## 13    -0.03046694  
## 14     0.90983067  
## 15    -0.20480121  
## 16     0.25175439  
## 17    -0.94389671  
## 18    -1.67578028  
## 19     0.54860754  
## 20    -1.49404570  
## 21    -0.48979705  
## 22     1.17614715  
## 23     1.15335453  
## 24    -0.26670340  
## 25    -1.35354693  
## 26     0.73823416  
## 27     0.54860754
```

28 0.24427281
29 1.19900178
30 -0.14877718
31 1.35812386
32 -1.80154608
33 -0.01155405
34 -1.56512471
35 -0.64214631
36 0.34586903
37 -0.30536769
38 1.90296267
39 -1.67578028
40 0.74313608
41 0.93268270
42 0.54860754
43 0.15939577
44 -0.48979705
45 1.77920441
46 1.90296267
47 -0.34033350
48 0.31532570
49 -0.17976545
50 -0.10566903
51 -0.16609280
52 0.25175439
53 1.17614715
54 -0.41139019
55 -0.61973599
56 -0.40940346
57 1.15335453
58 0.25175439
59 0.72459175
60 -0.82265038
61 -0.01761712
62 -0.17976545
63 1.35812386
64 -1.22819634
65 -0.81207759
66 0.73823416
67 1.04329131
68 0.39360628
69 -0.77708731
70 -0.55121576
71 -0.40940346
72 -0.17112235
73 -0.25325830
74 -0.81626214
75 -0.76771876
76 -0.94716143
77 0.06261040
78 -0.16973990
79 -0.32126615
80 -0.77708731
81 -1.49404570

```
## 82 -1.22819634
## 83 -1.56512471
## 84  1.94094052
## 85 -1.35354693
## 86 -0.16609280
## 87 -1.56512471
## 88 -0.55121576
## 89 -0.76031206
## 90 -0.01155405
## 91 -0.64234476
## 92 -0.61973599
## 93 -0.20754711
## 94 -0.76031206
## 95  0.39360628
## 96 -0.14877718
## 97 -0.40940346
## 98 -0.03046694
## 99  0.73823416
## 100 -0.94389671
```

```
P1_bs <- 1:num_trials %>%
  map_dfr(
    ~P1sample %>%
      slice_sample(n = 100, replace = TRUE) %>%
      summarise(mean_mu = mean(s100_1))
  )

mumean <- mean(P1_bs$mean_mu)
semu <- sd(P1_bs$mean_mu)/sqrt(100)
CIP1 <- c(mumean - 1.96*semu, mumean + 1.96*semu)
CIP1
```

```
## [1] -0.08875917 -0.05646426
```

Problem 2

(a)

```
xi <- c("one two three", "four five six", "SEVEN EIGHT")
```

```
word(xi, 2)
```

```
## [1] "two" "five" "EIGHT"
```

```
mysecondwords <- function(x) {
  y <- x %>%
    str_remove("[A-z]+\\s") %>%
    str_extract("[A-z]+")
  return(y)
}
```

```
mysecondwords(xi)
```

```
## [1] "two" "five" "EIGHT"
```

Problem 3

```
hn19<- haven::read_sas("HN19_ALL.sas7bdat")
head(hn19)
```

```
## # A tibble: 6 x 831
##   mod_d   ID   ID_fam year region town_t apt_t psu   sex   age age_m~1 incm
##   <chr>  <chr> <chr>  <dbl> <dbl> <dbl> <dbl> <chr> <dbl> <dbl>  <dbl> <dbl>
## 1 2022.0~ A751~ A7512~ 2019     1     1     1 A751     1    61     NA     1
## 2 2022.0~ A751~ A7512~ 2019     1     1     1 A751     1    28     NA     1
## 3 2022.0~ A751~ A7512~ 2019     1     1     1 A751     1    53     NA     1
## 4 2022.0~ A751~ A7512~ 2019     1     1     1 A751     2    50     NA     1
## 5 2022.0~ A751~ A7512~ 2019     1     1     1 A751     1    16     NA     1
## 6 2022.0~ A751~ A7512~ 2019     1     1     1 A751     2    71     NA     2
## # ... with 819 more variables: ho_incm <dbl>, incm5 <dbl>, ho_incm5 <dbl>,
## #   edu <dbl>, occp <dbl>, wt_hs <dbl>, wt_itvex <dbl>, wt_oe <dbl>,
## #   wt_pft <dbl>, wt_ex1 <dbl>, wt_ige <dbl>, wt_ntr <dbl>, wt_tot <dbl>,
## #   wt_oent <dbl>, wt_pfnt <dbl>, wt_exint <dbl>, wt_igent <dbl>,
## #   kstrata <dbl>, cfam <dbl>, genertn <dbl>, allownc <dbl>, house <dbl>,
## #   live_t <dbl>, ainc_unit1 <dbl>, ainc_1 <dbl>, ainc <dbl>, marri_1 <dbl>,
## #   marri_2 <dbl>, fam_rela <dbl>, tins <dbl>, npins <dbl>, id_f <chr>, ...
```

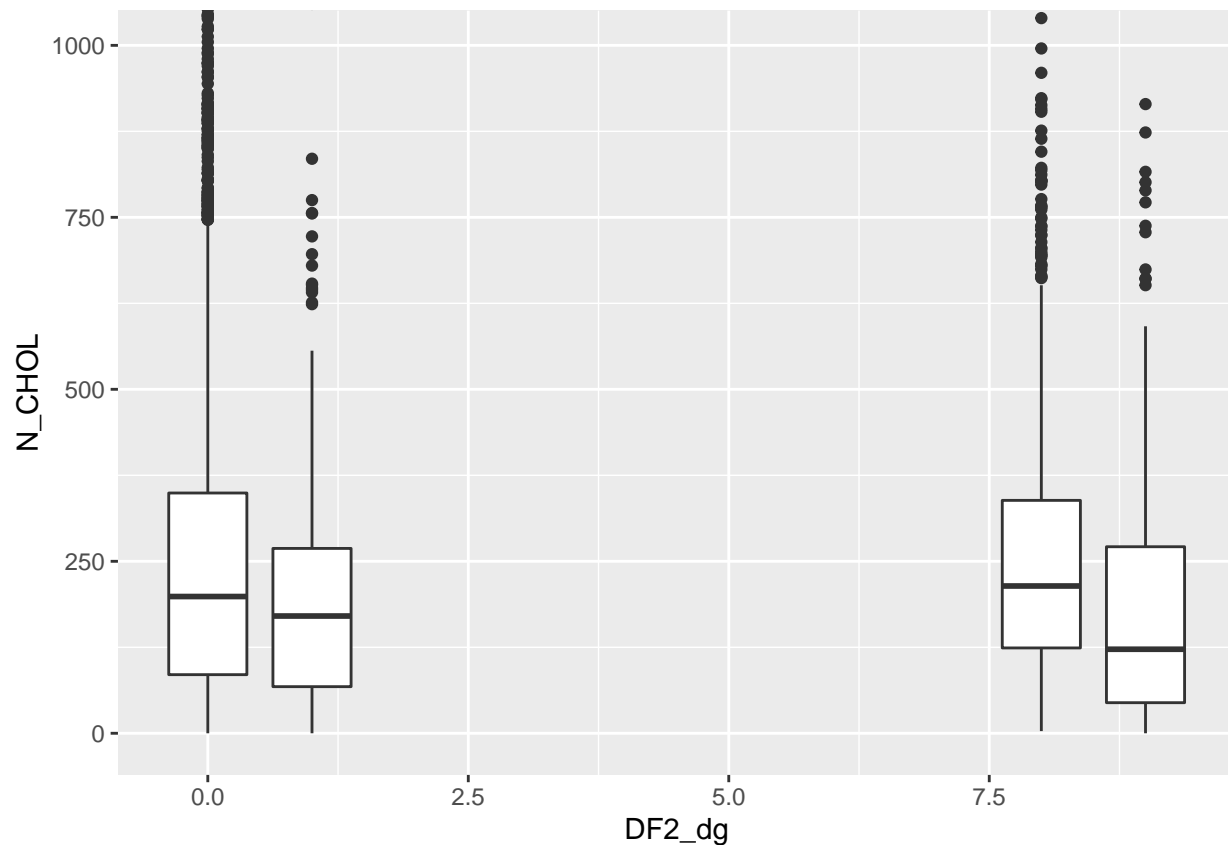
(a)

```
hn19b <- hn19 %>%
  select(ID, DF2_dg, N_CHOL) %>%
  filter(!is.na(DF2_dg), !is.na(N_CHOL))

max(hn19b$N_CHOL)
```

```
## [1] 2636.978
```

```
ggplot(data = hn19b) +
  geom_boxplot(aes(x = DF2_dg, y = N_CHOL, group = DF2_dg)) +
  coord_cartesian(ylim = c(-10,1000))
```



colnames(hn19)

```
## [1] "mod_d"      "ID"          "ID_fam"      "year"
## [5] "region"     "town_t"      "apt_t"        "psu"
## [9] "sex"        "age"         "age_month"    "incm"
## [13] "ho_incm"    "incm5"       "ho_incm5"     "edu"
## [17] "occip"      "wt_hs"       "wt_itvex"     "wt_oe"
## [21] "wt_pft"     "wt_ex1"      "wt_ige"       "wt_ntr"
## [25] "wt_tot"     "wt_oent"     "wt_pfnt"      "wt_ex1nt"
## [29] "wt_igent"   "kstrata"     "cfam"         "genertn"
## [33] "allownc"    "house"       "live_t"       "ainc_unit1"
## [37] "ainc_1"     "ainc"        "marri_1"      "marri_2"
## [41] "fam_rela"   "tins"        "npins"        "id_f"
## [45] "id_m"       "D_1_1"       "D_2_1"        "D_2_wk"
## [49] "DI1_dg"     "DI1_ag"      "DI1_pr"       "DI1_pt"
## [53] "DI1_2"      "DI2_dg"      "DI2_ag"       "DI2_pr"
## [57] "DI2_pt"     "DI2_2"       "DI3_dg"       "DI3_ag"
## [61] "DI3_pr"     "DI3_pt"      "DI3_2"        "DI4_dg"
## [65] "DI4_pr"     "DI4_pt"      "DI5_dg"       "DI5_ag"
## [69] "DI5_pr"     "DI5_pt"      "DI6_dg"       "DI6_ag"
## [73] "DI6_pr"     "DI6_pt"      "DM1_dg"       "DM1_pr"
## [77] "DM1_pt"     "DM2_dg"      "DM2_ag"       "DM2_pr"
## [81] "DM2_pt"     "DM3_dg"      "DM3_ag"       "DM3_pr"
## [85] "DM3_pt"     "DM4_dg"      "DM4_ag"       "DM4_pr"
## [89] "DM4_pt"     "DJ2_dg"      "DJ2_ag"       "DJ2_pr"
## [93] "DJ2_pt"     "DJ4_dg"      "DJ4_ag"       "DJ4_pr"
```

## [97]	"DJ4_pt"	"DJ4_3"	"DE2_dg"	"DE2_ag"
## [101]	"DE2_pr"	"DE2_pt"	"DE1_dg"	"DE1_ag"
## [105]	"DE1_pr"	"DE1_pt"	"DE1_3"	"DE1_31"
## [109]	"DE1_32"	"DE1_33"	"DE1_34"	"DE1_35"
## [113]	"DE1_4"	"DC1_dg"	"DC1_ag"	"DC1_pr"
## [117]	"DC1_pt"	"DC2_dg"	"DC2_ag"	"DC2_pr"
## [121]	"DC2_pt"	"DC3_dg"	"DC3_ag"	"DC3_pr"
## [125]	"DC3_pt"	"DC4_dg"	"DC4_ag"	"DC4_pr"
## [129]	"DC4_pt"	"DC5_dg"	"DC5_ag"	"DC5_pr"
## [133]	"DC5_pt"	"DC6_dg"	"DC6_ag"	"DC6_pr"
## [137]	"DC6_pt"	"DC7_dg"	"DC7_ag"	"DC7_pr"
## [141]	"DC7_pt"	"DC11_tp"	"DC11_dg"	"DC11_ag"
## [145]	"DC11_pr"	"DC11_pt"	"DF2_dg"	"DF2_ag"
## [149]	"DF2_pr"	"DF2_pt"	"DL1_dg"	"DL1_ag"
## [153]	"DL1_pr"	"DL1_pt"	"DJ8_dg"	"DJ8_ag"
## [157]	"DJ8_pr"	"DJ8_pt"	"DJ6_dg"	"DJ6_ag"
## [161]	"DJ6_pr"	"DJ6_pt"	"DH4_dg"	"DH4_ag"
## [165]	"DH4_pr"	"DH4_pt"	"DN1_dg"	"DN1_ag"
## [169]	"DN1_pr"	"DN1_pt"	"DK8_dg"	"DK8_ag"
## [173]	"DK8_pr"	"DK8_pt"	"DK9_dg"	"DK9_ag"
## [177]	"DK9_pr"	"DK9_pt"	"DK4_dg"	"DK4_ag"
## [181]	"DK4_pr"	"DK4_pt"	"DM8_dg"	"DM8_ag"
## [185]	"DM8_pr"	"DM8_pt"	"BP17_dg"	"DF1_yd"
## [189]	"DF1_ya"	"DJ9_yd"	"DJ9_ya"	"M_2_yr"
## [193]	"M_2_rs"	"M_2_et"	"mt_nontrt"	"BH9_11"
## [197]	"BH1"	"BH1_1"	"BH1_2"	"BH1_3"
## [201]	"BH1_8"	"BH1_6"	"BH1_7"	"BH2_61"
## [205]	"BH2_62"	"BH2_63"	"BH2_66"	"BH2_67"
## [209]	"BH2_64"	"BH2_65"	"LQ4_00"	"LQ4_01"
## [213]	"LQ4_02"	"LQ4_03"	"LQ4_04"	"LQ4_05"
## [217]	"LQ4_06"	"LQ4_07"	"LQ4_08"	"LQ4_09"
## [221]	"LQ4_10"	"LQ4_11"	"LQ4_12"	"LQ4_13"
## [225]	"LQ4_14"	"LQ4_15"	"LQ4_16"	"LQ4_21"
## [229]	"LQ4_22"	"LQ4_25"	"LQ4_26"	"LQ4_27"
## [233]	"LQ4_28"	"LQ4_29"	"LQ4_23"	"LQ4_24"
## [237]	"LQ4_17"	"LQ4_18"	"LQ4_19"	"LQ4_20"
## [241]	"LQ1_sb"	"LQ1_mn"	"LQ2_ab"	"LQ2_mn"
## [245]	"LQ_1EQL"	"LQ_2EQL"	"LQ_3EQL"	"LQ_4EQL"
## [249]	"LQ_5EQL"	"EQ5D"	"LQ_1HT"	"LQ_2HT"
## [253]	"LQ_3HT"	"LQ_4HT"	"LQ_5HT"	"LQ_6HT"
## [257]	"LQ_7HT"	"LQ_8HT"	"AC1_yr"	"AC3"
## [261]	"AC3_1_01"	"AC3_2_01"	"AC8_1_01"	"AC8_1e_01"
## [265]	"AC3_1_02"	"AC3_2_02"	"AC8_1_02"	"AC8_1e_02"
## [269]	"AC3_1_03"	"AC3_2_03"	"AC8_1_03"	"AC8_1e_03"
## [273]	"AC3_1_04"	"AC3_2_04"	"AC8_1_04"	"AC8_1e_04"
## [277]	"AC3_1_05"	"AC3_2_05"	"AC8_1_05"	"AC8_1e_05"
## [281]	"ij_expr"	"MH1_yr"	"MH1_1"	"M01_wk"
## [285]	"M01_1"	"educ"	"graduat"	"CH2_1"
## [289]	"CH2_2"	"EC1_1"	"EC_occp"	"EC_stt_1"
## [293]	"EC_stt_2"	"EC_wht_0"	"EC_wht_23"	"EC_wht_5"
## [297]	"EC_wht_6"	"EC_lgw_2"	"EC_lgw_4"	"EC_lgw_5"
## [301]	"EC_pedu_1"	"EC_pedu_2"	"B01"	"B01_1"
## [305]	"B01_2"	"B01_3"	"B02_1"	"B03_01"
## [309]	"B03_02"	"B03_03"	"B03_14"	"B03_05"

## [313]	"B03_04"	"B03_12"	"B03_07"	"B03_09"
## [317]	"B03_10"	"B03_11"	"BD1"	"BD2"
## [321]	"BD1_11"	"BD2_1"	"BD2_14"	"BD2_31"
## [325]	"BD2_32"	"BD7_4"	"BD7_5"	"BD7_61"
## [329]	"BD7_62"	"BD7_63"	"BD7_64"	"BD7_65"
## [333]	"BD7_66"	"BD7_67"	"BD9_1"	"BD9_2"
## [337]	"dr_month"	"BA2_12"	"BA2_13"	"BA2_14"
## [341]	"BA2_2_1"	"BA2_2_2"	"BA2_2_3"	"BA2_2_4"
## [345]	"BA2_2_5"	"BA2_2_6"	"BA2_22"	"BA1_1"
## [349]	"BA1_3"	"BA1_2"	"sc_seatblt"	"sc_seatblt2"
## [353]	"sc_seatblt3"	"BP16_1"	"BP16_2"	"BP17_2"
## [357]	"BP17_3"	"BP17_4"	"BP1"	"BP5"
## [361]	"BP6_10"	"BP6_2"	"BP6_31"	"BP7"
## [365]	"mh_stress"	"BS1_1"	"BS2_1"	"BS3_1"
## [369]	"BS3_2"	"BS3_3"	"BS6_2"	"BS6_2_1"
## [373]	"BS6_2_2"	"BS6_3"	"BS6_4"	"BS6_4_1"
## [377]	"BS6_4_2"	"BS12_37"	"BS12_47"	"BS12_47_1"
## [381]	"BS12_47_2"	"BS12_1"	"BS12_2"	"BS12_31"
## [385]	"BS12_32"	"BS12_33"	"BS12_34"	"BS12_35"
## [389]	"BS12_36"	"BS12_41"	"BS12_42"	"BS12_43"
## [393]	"BS12_44"	"BS12_45"	"BS12_46"	"BS5_4"
## [397]	"BS5"	"BS5_1"	"BS5_21"	"BS5_28"
## [401]	"BS5_26"	"BS5_33"	"BS5_34"	"BS5_32"
## [405]	"BS5_29"	"BS5_30"	"BS5_31"	"BS8_2"
## [409]	"BS9_2"	"BS13"	"BS10_1"	"BS2"
## [413]	"BS10_2"	"BS10_3"	"sm_presnt"	"BE3_71"
## [417]	"BE3_72"	"BE3_73"	"BE3_74"	"BE3_81"
## [421]	"BE3_82"	"BE3_83"	"BE3_84"	"BE3_91"
## [425]	"BE3_92"	"BE3_93"	"BE3_94"	"BE3_75"
## [429]	"BE3_76"	"BE3_77"	"BE3_78"	"BE3_85"
## [433]	"BE3_86"	"BE3_87"	"BE3_88"	"BE8_1"
## [437]	"BE8_2"	"BE3_31"	"BE3_32"	"BE3_33"
## [441]	"BE5_1"	"BE9"	"pa_aerobic"	"LW_ms"
## [445]	"LW_mp_e"	"LW_mp_a"	"LW_ms_a"	"LW_pr"
## [449]	"LW_pr_1"	"LW_mt"	"LW_mt_a1"	"LW_mt_a2"
## [453]	"LW_br"	"LW_br_ch"	"LW_br_dur"	"LW_br_yy"
## [457]	"LW_br_mm"	"LW_oc"	"HE_fst"	"HE_HPdg"
## [461]	"HE_HPdr"	"HE_DMdg"	"HE_DMdr"	"HE_mens"
## [465]	"HE_prg"	"HE_dprg"	"HE_fh"	"HE_HPfh1"
## [469]	"HE_HPfh2"	"HE_HPfh3"	"HE_HLfh1"	"HE_HLfh2"
## [473]	"HE_HLfh3"	"HE_IHDfh1"	"HE_IHDfh2"	"HE_IHDfh3"
## [477]	"HE_STRfh1"	"HE_STRfh2"	"HE_STRfh3"	"HE_DMfh1"
## [481]	"HE_DMfh2"	"HE_DMfh3"	"HE_THfh1"	"HE_THfh2"
## [485]	"HE_THfh3"	"HE_HBfh1"	"HE_HBfh2"	"HE_HBfh3"
## [489]	"HE_rPLS"	"HE_PLS"	"HE_mPLS"	"HE_nARM"
## [493]	"HE_sbp1"	"HE_dbp1"	"HE_sbp2"	"HE_dbp2"
## [497]	"HE_sbp3"	"HE_dbp3"	"HE_sbp"	"HE_dbp"
## [501]	"HE_HP"	"HE_ht"	"HE_wt"	"HE_wc"
## [505]	"HE_nc"	"HE_BMI"	"HE_obe"	"HE_wt_pct"
## [509]	"HE_BMI_pct"	"HE_glu"	"HE_HbA1c"	"HE_insulin"
## [513]	"HE_insulin_etc"	"HE_DM_HbA1c"	"HE_chol"	"HE_HDL_st2"
## [517]	"HE_TG"	"HE_LDL_drct"	"HE_HCHOL"	"HE_HTG"
## [521]	"HE_ast"	"HE_alt"	"HE_alt_etc"	"HE_hepaB"
## [525]	"HE_hepaC"	"HE_HB"	"HE_HCT"	"HE_anem"

## [529]	"HE_BUN"	"HE_crea"	"HE_WBC"	"HE_RBC"
## [533]	"HE_Bplt"	"HE_Uacid"	"HE_Uacid_etc"	"HE_df"
## [537]	"HE_dog"	"HE_cat"	"HE_birch"	"HE_oak"
## [541]	"HE_Jhop"	"HE_ragweed"	"HE_Uph"	"HE_Unitr"
## [545]	"HE_Usg"	"HE_Upro"	"HE_Uglu"	"HE_Uket"
## [549]	"HE_Ubil"	"HE_Ubld"	"HE_Uro"	"HE_Ucrea"
## [553]	"HE_Una"	"HE_Ualb"	"HE_Ualb_etc"	"HE_Ukal"
## [557]	"HE_Ukal_etc"	"HE_Ucot"	"HE_Ucot_etc"	"HE_fvc"
## [561]	"HE_fvcp"	"HE_fev1"	"HE_fev1p"	"HE_fev1fvc"
## [565]	"HE_fev6"	"HE_fef25_75"	"HE_pef"	"HE_COPD"
## [569]	"HE_cough1"	"HE_cough2"	"HE_sput1"	"HE_sput2"
## [573]	"HE_PFTdr"	"HE_PFTag"	"HE_PFTtr"	"HE_PFThs"
## [577]	"O_DTD"	"O_DTP"	"O_DID"	"O_DIP"
## [581]	"O_DFTD"	"O_DMFTP"	"O_DFID"	"O_DMFIP"
## [585]	"OR1"	"O_pain"	"O_ortho"	"BM1_0"
## [589]	"BM1_1"	"BM1_2"	"BM1_3"	"BM1_4"
## [593]	"BM1_5"	"BM1_6"	"BM1_7"	"BM1_8"
## [597]	"BM2_1"	"BM2_3"	"BM2_2"	"BM2_4"
## [601]	"BM2_5"	"BM13"	"BM13_1"	"BM13_2"
## [605]	"BM13_3"	"BM13_4"	"BM13_5"	"BM13_6"
## [609]	"BM7"	"O_chew_d"	"BM8"	"OR1_2"
## [613]	"M04_00"	"M04_4"	"M04_9"	"M04_7"
## [617]	"M04_8"	"M04_17"	"M04_11"	"M04_12"
## [621]	"M04_18"	"M04_15"	"BM14"	"BM14_1"
## [625]	"BM14_2"	"GS_use"	"GS_mea_r_1"	"GS_mea_r_2"
## [629]	"GS_mea_r_3"	"GS_mea_l_1"	"GS_mea_l_2"	"GS_mea_l_3"
## [633]	"T_ex"	"T_Q_HR"	"T_Q_HR_1"	"T_Q_HR_2"
## [637]	"T_NQ_OCP"	"T_Q_VN"	"T_Q_VN_1"	"T_Q_VN_2"
## [641]	"T_Q_DZ"	"T_Q_DZ2"	"T_Q_DZ3"	"T_Q_DZ4"
## [645]	"T_Q_DZ5"	"T_Q_DZ6"	"T_Q_DZ7"	"T_Q_CR"
## [649]	"T_Q_CR1"	"T_Q_CR2"	"T_Q_CR3"	"T_Q_CR4"
## [653]	"T_Q_ALLE"	"T_Q_SNST1"	"T_Q_SNST2"	"T_Q_SNST3"
## [657]	"T_Q_SNST4"	"T_Q_SINU"	"T_Q_SINU_1"	"T_Q_VC"
## [661]	"T_Q_VC1"	"T_HR_EX_R"	"T_HR_EX_R_1"	"T_HR_EX_R_2"
## [665]	"T_HR_EX_L"	"T_HR_EX_L_1"	"T_HR_EX_L_2"	"T_HR500_pass"
## [669]	"T_HR1000_pass"	"T_HR2000_pass"	"T_HR4000_pass"	"T_HR500_RT"
## [673]	"T_HR1000_RT"	"T_HR2000_RT"	"T_HR4000_RT"	"T_HR8000_RT"
## [677]	"T_HR500_LT"	"T_HR1000_LT"	"T_HR2000_LT"	"T_HR4000_LT"
## [681]	"T_HR8000_LT"	"T_HL_mod"	"T_HL_mild"	"T_IM_EX_R"
## [685]	"T_IM_EX_R_1"	"T_IM_EX_R_2"	"T_IM_EX_L"	"T_IM_EX_L_1"
## [689]	"T_IM_EX_L_2"	"T_IM_A_R"	"T_IM_A_L"	"T_IM_B_R"
## [693]	"T_IM_B_L"	"T_IM_C_R"	"T_IM_C_L"	"T_IM_FL_R"
## [697]	"T_IM_FL_L"	"T_Tymp_Dis"	"L_BR"	"L_LN"
## [701]	"L_DN"	"L_BR_FQ"	"L_LN_FQ"	"L_DN_FQ"
## [705]	"L_BR_TO"	"L_BR_WHO"	"L_LN_TO"	"L_LN_WHO"
## [709]	"L_DN_TO"	"L_DN_WHO"	"L_OUT_FQ"	"LS_1YR"
## [713]	"LK_EDU"	"LK_LB_CO"	"LK_LB_US"	"LK_LB_IT"
## [717]	"LK_LB_EF"	"DQ_CWATER"	"DA_CWATER"	"DQ_SODA"
## [721]	"DA_SODA"	"DQ_SPORTS"	"DA_SPORTS"	"DQ_CAFFEINE"
## [725]	"DA_CAFFEINE"	"DQ_FVJUICE"	"DA_FVJUICE"	"DQ_SFVJUICE"
## [729]	"DA_SFVJUICE"	"DQ_COFFEE"	"DA_COFFEE"	"DQ_SCOFFEE"
## [733]	"DA_SCOFFEE"	"DQ_TEA"	"DA_TEA"	"DQ_STEA"
## [737]	"DA_STEA"	"DQ_MILK"	"DA_MILK"	"DQ_SMILK"
## [741]	"DA_SMILK"	"DQ_YOGURT"	"DA_YOGURT"	"Y_BTH_WT"

(b)

’ ’ ’ ’ ’