# Data Merging: Tolterodine Protocol A6121123

Russell V. Lenth
Department of Statistics
The University of Iowa

December 4, 2006

**Abstract**

This document details the importation of raw data, and the matching/merging of the various data sets into one dataset. This document constitutes a "literate program": it contains computer code that is run in the process of creating this document, and the merged data set is created simultaneously. Thus, it is a reliable, reproducible description of the data set used in further analysis.

## 1 Data tables

In this section we detail the importing of the raw data files, and provide brief structure summaries.

**randomization** This dataset gives information about the assignment of subjects to treatments. The randomization data were provided in a PDF file. I copied and pasted the information into a text file. We subtract a constant from each subject ID to remove the first 4 digits.

```
R> randomization = read.table("A6121123 A7.2-2.txt", header = TRUE)
R> randomization$Subj = randomization$Subj - 10010000
```

```
R> str(randomization)
```

```
`data.frame':       24 obs. of  6 variables:
 $ Subj        : num  1101 1102 1104 1106 1107 ...
 $ seq         : Factor w/ 6 levels "ABC","ACB","BAC",..: 2 1 5 1 2 5 1 3 1 2 ...
 $ ActSeq      : Factor w/ 6 levels "ABC","ACB","BAC",..: 2 1 5 1 2 5 1 3 1 2 ...
 $ FirstTrtDate: Factor w/ 6 levels "02FEB2005","14MAR2005",..: 1 1 6 5 4 4 2 5 5 6 ...
 $ FirstTrtTime: Factor w/ 4 levels "6:59","7:00",..: 3 4 2 1 2 2 4 2 4 2 ...
 $ AgeGrp      : Factor w/ 2 levels "46-64","65+": 1 2 2 1 2 2 1 1 1 1 ...
```

**driving** The driving data were sent as an Excel file that was opened and re-saved as a comma-delimited file.

```
R> driving = read.csv("Pfizer All Data Sent Final.csv", na = ".")
R> driving$Group = factor(driving$Group)
R> driving$Week = factor(driving$Week)
R> driving$Day = factor(driving$Day)
R> driving$Time = factor(driving$Time)
```

```
R> str(driving)
```

```
`data.frame':       72 obs. of  32 variables:
 $ Center.Num       : int  1001 1001 1001 1001 1001 1001 1001 1001 1001 1001 ...
 $ Screen.Num       : Factor w/ 24 levels "1001-1101","1001-1102",..: 5 6 12 13 4 8 9 14 3 10 ...
 $ Rand             : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Group            : Factor w/ 6 levels "1","2","3","4",..: 1 1 1 1 2 2 2 2 3 3 ...
 $ Week             : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
 $ Order            : Factor w/ 2 levels "NS","SN": 2 1 2 1 2 1 2 1 2 1 ...
 $ Day              : Factor w/ 2 levels "4","5": 1 2 1 2 1 2 1 2 1 2 ...
```

```
 $ Time              : Factor w/ 2 levels "1","2": 1 1 2 2 1 1 2 2 1 1 ...
 $ P1.Phase          : num  -0.710 -0.916 -1.102 -1.008 -0.888 ...
 $ P1.RMS.Error      : num  2.47 2.55 2.75 2.80 2.29 ...
 $ P1.Steer.Instab   : num  2.537 1.271 2.472 0.683 3.397 ...
 $ P1.Lane.Exc       : int  0 0 0 0 12 0 0 0 0 2 ...
 $ P1.Velo.Instab    : num  8.43 6.69 7.60 7.84 4.88 ...
 $ P1.SD.of.Lat.Pos  : num  0.478 0.367 0.294 0.311 0.243 ...
 $ P1.Fix.Rate       : num  2.752 0.239 0.363 0.275 0.842 ...
 $ P1.Sacc.Rate      : num  2.750 0.237 0.361 0.273 0.840 ...
 $ P1.Avg.Fix.Length : num  0.266 4.101 2.677 3.562 1.089 ...
 $ P1.Avg.Sacc.Length: num  0.0975 0.0824 0.0770 0.0745 0.0992 ...
 $ P2.Steer.Instab   : num  37.7 37.0 37.8   NA 37.6 ...
 $ P2.Lane.Exc       : int  5 14 9 NA 32 43 10 6 8 9 ...
 $ P2.Velo.Instab    : num  14.4 16.5 14.2   NA 19.5 ...
 $ P2.SD.of.Lat.Pos  : num  1.11 1.86 1.45   NA 1.20 ...
 $ P2.Fix.Rate       : num  2.726 0.424 0.877    NA 0.791 ...
 $ P2.Sacc.Rate      : num  2.726 0.423 0.876    NA 0.791 ...
 $ P2.Avg.Fix.Length : num  0.273 2.251 1.059    NA 1.173 ...
 $ P2.Avg.Sacc.Length: num  0.0942 0.1082 0.0815     NA 0.0917 ...
 $ FE.Accel.Rel      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FE.Min.TTC        : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FE.Fix.Rate       : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FE.Sacc.Rate      : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FE.Avg.Fix.Length : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FE.Avg.Sacc.Length: num  NA NA NA NA NA NA NA NA NA NA ...
```

**npdata** The neuropsychological data were provided as an SPSS `SAV` file. The dataset contains two missing-value codes: $-3$ denotes a value that is missing because the protocol does not call for data to be collected there; $-9$ is the code for an actual nonresponse. When we merge the datasets, the $-3$s will not be an issue at all. However, it is expedient to just set all the negative values to missing.

```
R> library(foreign)
R> npdata = read.spss("np-Data.sav", to.data.frame = TRUE)
R> for (i in 1:length(npdata)) npdata[[i]][npdata[[i]] < 0] = NA
R> rm(i)
```

```
R> str(npdata)
```

```
`data.frame':        192 obs. of  67 variables:
 $ GROUP   : num  4 3 3 4 6 5 4 1 1 5 ...
 $ ID      : num  1128 1104 1115 1101 1139 ...
 $ SESSION : num  1 1 1 1 1 1 1 1 1 1 ...
 $ AGE     : num  60 68 64 56 57 61 66 60 72 58 ...
 $ EDUYRS  : num  16 12 18 17 18 14 13.5 19 16 12 ...
 $ HAND    : Factor w/ 2 levels "right","left": 1 1 2 1 1 1 1 1 1 1 ...
 $ MFVPT   : num  13 16 21 18 21 20 18 18 17 20 ...
 $ REYCOPY : num  34 30 30 34 34 34 36 31 29 35 ...
 $ COPYTIME: num  530 133 133 434 291 245 342 120 84 208 ...
 $ COPYPERC: atomic  17 14 8 17 17 17 17 14 17 17 ...
  ..- attr(*, "value.labels")= Named num 17
  .. ..- attr(*, "names")= chr ">16%ile"
 $ REYRECAL: num  9.5 16 17 20.5 23.5 22 25 20 19 17.5 ...
 $ RECLTSCR: num  35 53 51 55 62 61 72 57 62 49 ...
 $ RECLPERC: num  7 62 54 69 88 86 99 76 88 46 ...
 $ BVMT1   : num  3 3 5 6 8 9 5 3 6 8 ...
 $ BVMT1TSC: num  38 40 49 52 61 68 49 38 56 62 ...
 $ BVMT1PER: num  12 16 46 58 86 96 46 12 73 88 ...
 $ BVMT2   : num  4 8 6 9 10 11 7 6 8 10 ...
 $ BVMT2TSC: num  31 52 41 54 58 65 47 40 53 59 ...
 $ BVMT2PER: num  3 58 18 66 79 93 38 16 62 82 ...
 $ BVMT3   : num  4 8 7 9 10 11 8 10 9 10 ...
 $ BVMT3TSC: atomic  24 47 41 48 53 60 46 55 52 54 ...
  ..- attr(*, "value.labels")= Named num 19
  .. ..- attr(*, "names")= chr "<20tscr"
 $ BVMT3PER: atomic  1 38 18 42 62 84 34 69 58 66 ...
  ..- attr(*, "value.labels")= Named num 0
  .. ..- attr(*, "names")= chr "<1%ile"
```

```
 $ BVMT123 : num  11 19 18 24 28 31 20 19 23 28 ...
 $ TSCR123 : num  28 46 43 51 58 66 47 43 54 59 ...
 $ PERC123 : num  1 34 24 54 79 95 38 24 66 82 ...
 $ BVMTLERN: num  1 5 2 3 2 2 3 7 3 2 ...
 $ LERNTSCR: num  35 57 41 46 41 41 46 68 46 41 ...
 $ LERNPERC: num  7 76 18 34 18 18 34 96 34 18 ...
 $ BVMTDELY: num  6 10 8 10 11 11 8 10 9 10 ...
 $ DELYTSCR: num  38 58 48 56 61 62 49 57 55 56 ...
 $ DELYPERC: num  12 79 42 73 86 88 46 76 69 73 ...
 $ BVMTRET : num  100 100 100 100 100 100 100 100 100 100 ...
 $ RETPERC : atomic  17 17 17 17 17 17 17 17 17 17 ...
  ..- attr(*, "value.labels")= Named num 17
  .. ..- attr(*, "names")= chr ">16%ile"
 $ TMTASCR : num  69 28 33 27 25 29 30 25 24 29 ...
 $ TMTAERR : num  0 0 0 0 0 0 0 0 0 0 ...
 $ TMTBSCR : num  96 47 83 45 62 69 61 40 54 38 ...
 $ TMTBERR : num  0 0 0 0 0 0 0 0 0 0 ...
 $ DIGITSYM: num  57 73 60 60 78 82 62 77 76 88 ...
 $ LNSSCORE: num  6 11 8 10 14 11 9 13 7 10 ...
 $ LNSSPAN : num  4 5 4 5 8 6 5 6 4 5 ...
 $ RAVLT1  : num  6 5 4 7 7 11 9 7 6 4 ...
 $ RAVLT2  : num  6 6 7 9 10 13 10 8 7 8 ...
 $ RAVLT3  : num  6 8 9 11 14 15 13 9 8 8 ...
 $ RAVLT4  : num  8 10 9 13 12 15 14 10 11 12 ...
 $ RAVLT5  : num  7 12 11 12 12 15 15 11 10 14 ...
 $ RAVLTSUM: num  33 41 40 52 55 69 61 45 42 46 ...
 $ RAVLTB  : num  4 3 5 6 6 8 7 5 3 4 ...
 $ RAVLT6  : num  5 9 6 10 15 15 13 8 5 12 ...
 $ RAVLTDLY: num  4 9 7 11 14 15 13 5 6 12 ...
 $ VOCTSCR : num  NA NA NA NA NA NA NA NA NA NA ...
 $ BLCKTSCR: num  NA NA NA NA NA NA NA NA NA NA ...
 $ SIMLTSCR: num  NA NA NA NA NA NA NA NA NA NA ...
 $ MATXTSCR: num  NA NA NA NA NA NA NA NA NA NA ...
 $ VERBTSCR: num  NA NA NA NA NA NA NA NA NA NA ...
 $ VERBALIQ: num  NA NA NA NA NA NA NA NA NA NA ...
 $ PERFTSCR: num  NA NA NA NA NA NA NA NA NA NA ...
 $ PERFIQ  : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FULL4TSC: num  NA NA NA NA NA NA NA NA NA NA ...
 $ FULL4IQ : num  NA NA NA NA NA NA NA NA NA NA ...
 $ FULL2TSC: num  NA NA NA NA NA NA NA NA NA NA ...
 $ FULL2IQ : num  NA NA NA NA NA NA NA NA NA NA ...
 $ MMSE    : num  NA NA NA NA NA NA NA NA NA NA ...
 $ TIMEPERC: atomic  1 17 17 4 17 17 17 17 17 17 ...
  ..- attr(*, "value.labels")= Named num 17
  .. ..- attr(*, "names")= chr ">16&ile"
 $ BUTTONS : num  353 394 292 388 254 ...
 $ CHOOSER : num  925 721 782 712 517 ...
 $ NBACK   : num  64 89 73 82 91 90 68 85 86 76 ...
 $ TAPPER  : num  46.2 51.2 54 44.6 46.6 44.4 42.8 49 38 47.2 ...
 - attr(*, "variable.labels")= Named chr  "participant study group" "participant study number" "participant visit number" "parti
  ..- attr(*, "names")= chr  "GROUP" "ID" "SESSION" "AGE" ...
```

**other** Some other data collected before and after each driving run were sent in comma-delimited format.

```
R> other = read.csv("Other.csv", na = ".")
```

```
R> str(other)
```

```
`data.frame':        144 obs. of  12 variables:
 $ Center    : int  1001 1001 1001 1001 1001 1001 1001 1001 1001 1001 ...
 $ ScreenNo  : Factor w/ 24 levels "1001-1101","1001-1102",..: 5 5 5 5 5 5 6 6 6 6 ...
 $ RAN       : int  1 1 1 1 1 1 2 2 2 2 ...
 $ Group     : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Week      : int  1 1 2 2 3 3 1 1 2 2 ...
 $ Order     : Factor w/ 2 levels "NS","SN": 2 2 2 2 2 2 1 1 1 1 ...
 $ Day       : int  4 5 4 5 4 5 4 5 4 5 ...
 $ Time      : int  1 1 1 1 1 1 1 1 1 1 ...
 $ PreDrowsy : int  8 5 9 9 13 9 1 1 4 1 ...
```

```
 $ PreImp    : int  8 7 8 11 13 9 1 1 4 2 ...
 $ PostDrowsy: int  13 NA 7 NA 12 NA NA 51 NA 33 ...
 $ PostImp   : int  10 NA 8 NA 13 NA NA 48 NA 36 ...
```

# 2 Merging the data

The data are matched by first sorting to ensure that the subject IDs line up (and also confirming that the subject IDs match among the tables)

```
R> ord = order(randomization$Subj)
R> randomization = randomization[ord, ]
R> print(randomization$Subj)

 [1] 1101 1102 1104 1106 1107 1108 1111 1113 1114 1115 1116 1117 1118 1119 1120
[16] 1125 1127 1128 1131 1132 1136 1139 1140 1143

R> ord = order(npdata$SESSION, npdata$ID)
R> npdata = npdata[ord, ]
R> print(npdata$ID[1:24])

 [1] 1101 1102 1104 1106 1107 1108 1111 1113 1114 1115 1116 1117 1118 1119 1120
[16] 1125 1127 1128 1131 1132 1136 1139 1140 1143

R> ord = order(driving$Week, driving$Screen.Num)
R> driving = driving[ord, ]
R> print(driving$Screen.Num[1:24])

 [1] 1001-1101 1001-1102 1001-1104 1001-1106 1001-1107 1001-1108 1001-1111
 [8] 1001-1113 1001-1114 1001-1115 1001-1116 1001-1117 1001-1118 1001-1119
[15] 1001-1120 1001-1125 1001-1127 1001-1128 1001-1131 1001-1132 1001-1136
[22] 1001-1139 1001-1140 1001-1143
24 Levels: 1001-1101 1001-1102 1001-1104 1001-1106 1001-1107 ... 1001-1143

R> ord = order(other$Day, other$Week, other$ScreenNo)
R> other = other[ord, ]
R> print(other$ScreenNo[1:24])

 [1] 1001-1101 1001-1102 1001-1104 1001-1106 1001-1107 1001-1108 1001-1111
 [8] 1001-1113 1001-1114 1001-1115 1001-1116 1001-1117 1001-1118 1001-1119
[15] 1001-1120 1001-1125 1001-1127 1001-1128 1001-1131 1001-1132 1001-1136
[22] 1001-1139 1001-1140 1001-1143
24 Levels: 1001-1101 1001-1102 1001-1104 1001-1106 1001-1107 ... 1001-1143

R> rm(ord)
```

We now need to select appropriate entries from each variable and bring them into a merged data set. Note that the data tables have different numbers of observations For analysis purposes, we want the merged data set to have $3 \times 24 = 72$ rows, consisting of an observation for each subject for each week of the experiment. Thus, the `driving` data set serves as the basic structure.

## 2.1 Driving variables

```
R> merged = driving
```

## 2.2 Treatment information

Variables from the `randomization` table need to come in three times:

```
R> attach(randomization)
R> merged$Seq = rep(seq, 3)
R> s = as.character(seq)
R> trt = c(substr(s, 1, 1), substr(s, 2, 2), substr(s, 3, 3))
R> merged$Trt = factor(trt, labels = c("Tol", "Oxy", "Pbo"))
R> levels(merged$Seq) = c("TOP", "TPO", "OTP", "OPT", "PTO", "POT")
R> rm(s, trt)
R> detach()
```

## 2.3 Neuropsychological variables

To bring-in variables from the `npdata` table, the following functions are useful. `forsess` returns the indices if the rows in `npdata` that correspond to the given session(s).

```
R> forsess = function(s) {
R+     which(!is.na(match(npdata$SESSION, s)))
R+ }
```

`addnp` adds the given variable `var` from `npdata` to `merged`, using the specified sessions `sess`. When it is only one session, the 24 values are replicated 3 times and the given `prefix` is prepended to the variable name. Otherwise it is assumed (without checking!) that the defaut sessions are `1:7`, and three variables are added: three copies of the variable for session 1 (prepended with `Base` for the baseline values), one for sessions 2, 4, and 6 (prepended with `Pre` because these are at the beginnings of the experimental weeks); and one for sessions 3, 5, and 7 (prepended with `Post` because these are at the ends of the experimental weeks).

```
R> addnp = function(var, sess = 1:7, prefix = "") {
R+     var = as.character(substitute(var))
R+     x = npdata[[var]]
R+     if (length(sess) == 1) {
R+         x = x[forsess(sess)]
R+         nm1 = paste(prefix, var, sep = "")
R+         merged[[nm1]] <<- rep(x, 3)
R+         cat(paste("Variable", nm1, "added\n"))
R+     }
R+     else {
R+         base = x[forsess(1)]
R+         x1 = x[forsess(c(2, 4, 6))]
R+         x2 = x[forsess(c(3, 5, 7))]
R+         nmbase = paste("Base", var, sep = "")
R+         nm1 = paste("Pre", var, sep = "")
R+         nm2 = paste("Post", var, sep = "")
R+         merged[[nmbase]] <<- rep(base, 3)
R+         merged[[nm1]] <<- x1
R+         merged[[nm2]] <<- x2
R+         cat(paste("Variables ", nmbase, ", ", nm1, ", and ",
R+             nm2, " added\n", sep = ""))
R+     }
R+ }
```

So here we go with the needed neuropsychological variables:

**Variables only measured in one session (0 or 1)**

```
R> addnp(AGE, 0)

Variable AGE added

R> addnp(EDUYRS, 0)

Variable EDUYRS added

R> addnp(VOCTSCR, 0)

Variable VOCTSCR added

R> addnp(BLCKTSCR, 0)

Variable BLCKTSCR added

R> addnp(SIMLTSCR, 0)

Variable SIMLTSCR added

R> addnp(MATXTSR, 0)

Variable MATXTSR added

R> addnp(FULL4IQ, 0)

Variable FULL4IQ added

R> addnp(MMSE, 0)

Variable MMSE added

R> addnp(MFVPT, 1)

Variable MFVPT added

R> addnp(REYCOPY, 1)

Variable REYCOPY added

R> addnp(REYRECAL, 1)

Variable REYRECAL added

R> addnp(BVMT123, 1)

Variable BVMT123 added

R> addnp(BVMTDELY, 1)

Variable BVMTDELY added

R> addnp(NBACK, 1)

Variable NBACK added
```

**Variables measured in sessions 1–7**

```
R> addnp(TMTASCR)

Variables BaseTMTASCR, PreTMTASCR, and PostTMTASCR added

R> addnp(TMTBSCR)

Variables BaseTMTBSCR, PreTMTBSCR, and PostTMTBSCR added

R> addnp(DIGITSYM)

Variables BaseDIGITSYM, PreDIGITSYM, and PostDIGITSYM added

R> addnp(LNSSCORE)

Variables BaseLNSSCORE, PreLNSSCORE, and PostLNSSCORE added

R> addnp(RAVLT1)

Variables BaseRAVLT1, PreRAVLT1, and PostRAVLT1 added

R> addnp(RAVLTSUM)

Variables BaseRAVLTSUM, PreRAVLTSUM, and PostRAVLTSUM added

R> addnp(RAVLTDLY)

Variables BaseRAVLTDLY, PreRAVLTDLY, and PostRAVLTDLY added

R> addnp(BUTTONS)

Variables BaseBUTTONS, PreBUTTONS, and PostBUTTONS added

R> addnp(CHOOSER)

Variables BaseCHOOSER, PreCHOOSER, and PostCHOOSER added

R> addnp(TAPPER)

Variables BaseTAPPER, PreTAPPER, and PostTAPPER added
```

## 2.4 Neuropsychological composite variables

The Pfizer protocol specifies two composite variables that we will label `Memory` and `Speed`, each an average of three standardized variables. For convenience, we define vectors with their names:

```
R> memvars = c("RAVLT1", "RAVLTSUM", "RAVLTDLY")
R> spdvars = c("TMTASCR", "TMTBSCR", "CHOOSER")
```

Here is a utility function to create these variables. It standardizes with respect to the values in sessions `wrt.sess`, and outputs the values for the sessions in `out.sess`. If `gain=TRUE`, the values from session 1 are subtracted first, and we make sure that session 1 is excluded from `wrt.sess`.

```
R> comp = function(vars, wrt.sess = 1:7, out.sess = c(3, 5, 7),
R+     gain = FALSE) {
R+     if (gain)
R+         wrt.sess = wrt.sess[wrt.sess != 1]
R+     i.in = forsess(wrt.sess)
R+     i.out = forsess(out.sess)
```

```
R+     std1 = function(v) {
R+         x = npdata[, v]
R+         if (gain)
R+             x = x - rep(x[25:48], 8)
R+         (x[i.out] - mean(x[i.in]))/sd(x[i.in])
R+     }
R+     Z = sapply(vars, std1)
R+     apply(Z, 1, mean)
R+ }
R> merged$BaseMemory = rep(comp(memvars, out = 1), 3)
R> merged$PreMemory = comp(memvars, out = c(2, 4, 6))
R> merged$PostMemory = comp(memvars)
R> merged$BaseSpeed = rep(comp(spdvars, out = 1), 3)
R> merged$PreSpeed = comp(spdvars, out = c(2, 4, 6))
R> merged$PostSpeed = comp(spdvars)
```

### 2.5 Other variables

The `other` table has 144 observations; after sorting, the first 72 are data collected on day 4 or each week, and the last 72 are data collected on day 5. The code below reflects this organization in extracting separate variables for days 4 and 5.

```
R> merged$PreDrowsy4 = other$PreDrowsy[1:72]
R> merged$PreDrowsy5 = other$PreDrowsy[73:144]
R> merged$PreImp4 = other$PreImp[1:72]
R> merged$PreImp5 = other$PreImp[73:144]
R> merged$PostDrowsy4 = other$PostDrowsy[1:72]
R> merged$PostDrowsy5 = other$PostDrowsy[73:144]
R> merged$PostImp4 = other$PostImp[1:72]
R> merged$PostImp5 = other$PostImp[73:144]
```

In these variables, "pre" and "post" refer to data collected before and after the test on that day (neuropsychological on one day and driving on the other, depending on `Order`).

## 3 Summary

Here is a structure listing of the merged dataset.

```
R> str(merged)
```

```
`data.frame':        72 obs. of  91 variables:
 $ Center.Num        : int  1001 1001 1001 1001 1001 1001 1001 1001 1001 1001 ...
 $ Screen.Num        : Factor w/ 24 levels "1001-1101","1001-1102",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ Rand              : int  13 15 9 5 1 2 24 6 7 10 ...
 $ Group             : Factor w/ 6 levels "1","2","3","4",..: 4 4 3 2 1 1 6 2 2 3 ...
 $ Week              : Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1 1 1 1 ...
 $ Order             : Factor w/ 2 levels "NS","SN": 2 2 2 2 2 1 1 1 2 1 ...
 $ Day               : Factor w/ 2 levels "4","5": 1 1 1 1 1 2 2 2 1 2 ...
 $ Time              : Factor w/ 2 levels "1","2": 1 2 1 1 1 1 2 1 2 1 ...
 $ P1.Phase          : num  -0.993 -1.080 -1.353 -0.888 -0.710 ...
 $ P1.RMS.Error      : num  2.39 2.91 2.58 2.29 2.47 ...
 $ P1.Steer.Instab   : num  1.14 2.01 1.13 3.40 2.54 ...
 $ P1.Lane.Exc       : int  0 0 0 12 0 0 0 0 0 2 ...
 $ P1.Velo.Instab    : num  5.88 7.27 3.63 4.88 8.43 ...
 $ P1.SD.of.Lat.Pos  : num  0.221 0.271 0.307 0.243 0.478 ...
```

```
$ P1.Fix.Rate       : num  0.802 0.812 0.590 0.842 2.752 ...
$ P1.Sacc.Rate      : num  0.800 0.810 0.588 0.840 2.750 ...
$ P1.Avg.Fix.Length : num  1.149 1.115 1.607 1.089 0.266 ...
$ P1.Avg.Sacc.Length: num  0.0974 0.1168 0.0881 0.0992 0.0975 ...
$ P2.Steer.Instab   : num  37.3 40.1 38.2 37.6 37.7 ...
$ P2.Lane.Exc       : int  5 21 8 32 5 14 1 43 10 9 ...
$ P2.Velo.Instab    : num  15.7 20.0 10.6 19.5 14.4 ...
$ P2.SD.of.Lat.Pos  : num  1.21 2.16 1.11 1.20 1.11 ...
$ P2.Fix.Rate       : num  1.342 0.873 0.806 0.791 2.726 ...
$ P2.Sacc.Rate      : num  1.341 0.872 0.805 0.791 2.726 ...
$ P2.Avg.Fix.Length : num  0.650 1.044 1.158 1.173 0.273 ...
$ P2.Avg.Sacc.Length: num  0.0954 0.1024 0.0836 0.0917 0.0942 ...
$ FE.Accel.Rel      : num  NA NA NA NA NA NA NA NA NA NA ...
$ FE.Min.TTC        : num  NA NA NA NA NA NA NA NA NA NA ...
$ FE.Fix.Rate       : num  NA NA NA NA NA NA NA NA NA NA ...
$ FE.Sacc.Rate      : num  NA NA NA NA NA NA NA NA NA NA ...
$ FE.Avg.Fix.Length : num  NA NA NA NA NA NA NA NA NA NA ...
$ FE.Avg.Sacc.Length: num  NA NA NA NA NA NA NA NA NA NA ...
$ Seq               : Factor w/ 6 levels "TOP","TPO","OTP",..: 2 1 5 1 2 5 1 3 1 2 ...
$ Trt               : Factor w/ 3 levels "Tol","Oxy","Pbo": 1 1 3 1 1 3 1 2 1 1 ...
$ AGE               : num  56 66 68 59 72 66 60 58 58 64 ...
$ EDUYRS            : num  17 13.5 12 15 16 16 14 20 16.5 18 ...
$ VOCTSCR           : num  60 58 53 65 64 64 52 60 57 64 ...
$ BLCKTSCR          : num  64 44 63 49 60 53 48 55 42 54 ...
$ SIMLTSCR          : num  63 47 50 64 63 62 56 59 53 61 ...
$ FULL4IQ           : num  121 101 112 118 123 119 106 112 102 118 ...
$ MMSE              : num  30 28 29 29 29 30 25 27 26 26 ...
$ MFVPT             : num  18 18 16 17 17 19 22 15 14 21 ...
$ REYCOPY           : num  34 36 30 30 29 34 35 30.5 25.5 30 ...
$ REYRECAL          : num  20.5 25 16 15 19 10 16.5 13 7.5 17 ...
$ BVMT123           : num  24 20 19 22 23 29 19 25 9 18 ...
$ BVMTDELY          : num  10 8 10 9 9 10 7 8 3 8 ...
$ NBACK             : num  82 68 89 75 86 80 61 81 74 73 ...
$ BaseTMTASCR       : num  27 30 28 27.5 24 40 53 32 48 33 ...
$ PreTMTASCR        : num  26 26 38 35 27 49 58 57 34 35 ...
$ PostTMTASCR       : num  20 20 18 19 22 38 37 32 32 37 ...
$ BaseTMTBSCR       : num  45 61 47 60 54 89 243 88 93 83 ...
$ PreTMTBSCR        : num  53 78 46 37 52 59 130 75 81 88 ...
$ PostTMTBSCR       : num  64 80 81 44 40 75 109 99 72 79 ...
$ BaseDIGITSYM      : num  60 62 73 81 76 62 37 59 62 60 ...
$ PreDIGITSYM       : num  66 65 81 82 76 63 40 68 62 72 ...
$ PostDIGITSYM      : num  64 76 84 79 80 72 56 69 68 60 ...
$ BaseLNSSCORE      : num  10 9 11 12 7 11 6 7 12 8 ...
$ PreLNSSCORE       : num  12 9 12 13 11 18 4 9 10 12 ...
$ PostLNSSCORE      : num  9 11 11 17 9 8 6 10 12 12 ...
$ BaseRAVLT1        : num  7 9 5 9 6 9 5 7 5 4 ...
$ PreRAVLT1         : num  8 8 6 9 7 6 4 6 5 7 ...
$ PostRAVLT1        : num  6 9 4 8 6 6 4 4 6 5 ...
$ BaseRAVLTSUM      : num  52 61 41 52 42 52 33 53 50 40 ...
$ PreRAVLTSUM       : num  53 62 36 64 44 49 29 44 50 44 ...
$ PostRAVLTSUM      : num  49 57 35 62 38 45 31 42 45 38 ...
$ BaseRAVLTDLY      : num  11 13 9 8 6 10 5 10 12 7 ...
$ PreRAVLTDLY       : num  9 15 7 9 6 6 5 10 8 6 ...
$ PostRAVLTDLY      : num  8 13 9 12 5 7 3 7 7 6 ...
```

```
$ BaseBUTTONS       : num  388 266 394 356 346 ...
$ PreBUTTONS        : num  256 410 428 313 441 ...
$ PostBUTTONS       : num  258 251 420 356 343 ...
$ BaseCHOOSER       : num   712  804  721  759 1088 ...
$ PreCHOOSER        : num  582 722 735 827 900 ...
$ PostCHOOSER       : num  591 648 720 787 988 ...
$ BaseTAPPER        : num  44.6 42.8 51.2 39.4 38 34 54 51.4 49 54 ...
$ PreTAPPER         : num  48.6 49.6 47.2 42.2 NA NA 55 50.6 57.8 53 ...
$ PostTAPPER        : num  46.2 49 48.2 40 41 39.6 60 49.6 47.4 47 ...
$ BaseMemory        : num   0.648  1.464 -0.231  0.712 -0.307 ...
$ PreMemory         : num   0.665  1.510 -0.404  1.185 -0.072 ...
$ PostMemory        : num   0.101  1.337 -0.593  1.230 -0.527 ...
$ BaseSpeed         : num  -0.1960  0.2812 -0.1252  0.0744  0.8182 ...
$ PreSpeed          : num  -0.5190  0.0758  0.2134  0.2989  0.3768 ...
$ PostSpeed         : num  -0.587 -0.296 -0.153 -0.247  0.365 ...
$ PreDrowsy4        : int  6 25 0 27 8 1 16 14 35 45 ...
$ PreDrowsy5        : int  13 26 0 21 5 1 12 4 7 1 ...
$ PreImp4           : int  11 27 1 31 8 1 25 27 13 25 ...
$ PreImp5           : int  20 21 0 25 7 1 15 6 6 1 ...
$ PostDrowsy4       : int  22 32 2 27 13 NA NA NA 46 NA ...
$ PostDrowsy5       : int  NA NA NA NA NA 51 21 53 NA 7 ...
$ PostImp4          : int  27 31 1 29 10 NA NA NA 27 NA ...
$ PostImp5          : int  NA NA NA NA NA 48 22 42 NA 8 ...
```

Here are statistical summaries of all the variables

```
R> summary(merged)

  Center.Num       Screen.Num      Rand       Group  Week   Order   Day
 Min.   :1001   1001-1101: 3   Min.   : 1.00   1:12   1:24   NS:36   4:36
 1st Qu.:1001   1001-1102: 3   1st Qu.: 6.75   2:12   2:24   SN:36   5:36
 Median :1001   1001-1104: 3   Median :12.50   3:12   3:24
 Mean   :1001   1001-1106: 3   Mean   :12.50   4:12
 3rd Qu.:1001   1001-1107: 3   3rd Qu.:18.25   5:12
 Max.   :1001   1001-1108: 3   Max.   :24.00   6:12
                (Other)  :54
 Time      P1.Phase        P1.RMS.Error   P1.Steer.Instab  P1.Lane.Exc
 1:36   Min.   :-1.3530   Min.   :1.602   Min.   :0.683   Min.   : 0.0000
 2:36   1st Qu.:-0.9573   1st Qu.:2.154   1st Qu.:1.014   1st Qu.: 0.0000
        Median :-0.8475   Median :2.414   Median :1.289   Median : 0.0000
        Mean   :-0.8499   Mean   :2.411   Mean   :1.619   Mean   : 0.5417
        3rd Qu.:-0.7138   3rd Qu.:2.656   3rd Qu.:2.101   3rd Qu.: 0.0000
        Max.   :-0.4180   Max.   :3.794   Max.   :3.909   Max.   :12.0000

 P1.Velo.Instab   P1.SD.of.Lat.Pos  P1.Fix.Rate      P1.Sacc.Rate
 Min.   : 2.742   Min.   :0.0711   Min.   :0.0564   Min.   :0.0544
 1st Qu.: 5.614   1st Qu.:0.1862   1st Qu.:0.4228   1st Qu.:0.4208
 Median : 7.016   Median :0.2582   Median :0.6669   Median :0.6650
 Mean   : 6.872   Mean   :0.2648   Mean   :0.8622   Mean   :0.8604
 3rd Qu.: 8.016   3rd Qu.:0.3182   3rd Qu.:1.1365   3rd Qu.:1.1345
 Max.   :12.796   Max.   :0.5866   Max.   :2.7524   Max.   :2.7504

 P1.Avg.Fix.Length P1.Avg.Sacc.Length P2.Steer.Instab  P2.Lane.Exc
 Min.   : 0.2659   Min.   :0.05860    Min.   :31.38   Min.   : 0.000
 1st Qu.: 0.8137   1st Qu.:0.06888    1st Qu.:36.09   1st Qu.: 2.000
```

```
Median : 1.4275   Median :0.07850    Median :37.11   Median : 6.000
Mean   : 2.0482   Mean   :0.08476    Mean   :37.24   Mean   : 9.521
3rd Qu.: 2.2656   3rd Qu.:0.09750    3rd Qu.:38.01   3rd Qu.:13.500
Max.   :17.6718   Max.   :0.16350    Max.   :43.59   Max.   :52.000
                                     NA's   : 1.00   NA's   : 1.000
P2.Velo.Instab   P2.SD.of.Lat.Pos  P2.Fix.Rate      P2.Sacc.Rate
Min.   : 4.954   Min.   :0.3625   Min.   :0.2682   Min.   :0.2677
1st Qu.:10.935   1st Qu.:0.9216   1st Qu.:0.7043   1st Qu.:0.7040
Median :13.292   Median :1.1259   Median :0.9197   Median :0.9192
Mean   :15.151   Mean   :1.2818   Mean   :1.1132   Mean   :1.1128
3rd Qu.:18.145   3rd Qu.:1.6287   3rd Qu.:1.4553   3rd Qu.:1.4548
Max.   :52.098   Max.   :2.6596   Max.   :2.7263   Max.   :2.7258
NA's   : 1.000   NA's   :1.0000   NA's   :1.0000   NA's   :1.0000
P2.Avg.Fix.Length P2.Avg.Sacc.Length  FE.Accel.Rel      FE.Min.TTC
Min.   :0.2727    Min.   :0.06080    Min.   : 0.150   Min.   : 0.538
1st Qu.:0.5941    1st Qu.:0.07065    1st Qu.: 0.780   1st Qu.: 1.893
Median :1.0040    Median :0.08150    Median : 1.250   Median : 2.119
Mean   :1.0826    Mean   :0.08652    Mean   : 1.177   Mean   : 2.067
3rd Qu.:1.3291    3rd Qu.:0.09480    3rd Qu.: 1.630   3rd Qu.: 2.427
Max.   :3.6399    Max.   :0.15220    Max.   : 2.080   Max.   : 3.235
NA's   :1.0000    NA's   :1.00000    NA's   :55.000   NA's   :48.000
 FE.Fix.Rate       FE.Sacc.Rate     FE.Avg.Fix.Length FE.Avg.Sacc.Length
Min.   : 0.1678   Min.   : 0.1398   Min.   : 0.2724   Min.   : 0.06460
1st Qu.: 0.7057   1st Qu.: 0.6725   1st Qu.: 0.5966   1st Qu.: 0.07460
Median : 0.9952   Median : 0.9537   Median : 0.9354   Median : 0.08430
Mean   : 1.2085   Mean   : 1.1821   Mean   : 1.1468   Mean   : 0.09068
3rd Qu.: 1.4950   3rd Qu.: 1.4522   3rd Qu.: 1.2932   3rd Qu.: 0.09645
Max.   : 2.7292   Max.   : 2.7292   Max.   : 5.8556   Max.   : 0.13810
NA's   :49.0000   NA's   :49.0000   NA's   :49.0000   NA's   :49.00000
 Seq       Trt           AGE            EDUYRS          VOCTSCR
TOP:12   Tol:24   Min.   :56.00   Min.   :11.00   Min.   :43.00
TPO:12   Oxy:24   1st Qu.:58.00   1st Qu.:13.38   1st Qu.:56.50
OTP:12   Pbo:24   Median :61.00   Median :16.00   Median :60.00
OPT:12            Mean   :63.25   Mean   :15.79   Mean   :58.96
PTO:12            3rd Qu.:66.50   3rd Qu.:18.00   3rd Qu.:64.00
POT:12            Max.   :74.00   Max.   :22.00   Max.   :65.00

   BLCKTSCR         SIMLTSCR         FULL4IQ           MMSE
Min.   :40.00   Min.   :45.00   Min.   : 94.0   Min.   :25.00
1st Qu.:48.75   1st Qu.:52.75   1st Qu.:106.8   1st Qu.:26.75
Median :54.50   Median :58.00   Median :118.0   Median :29.00
Mean   :54.96   Mean   :56.79   Mean   :113.2   Mean   :28.25
3rd Qu.:61.00   3rd Qu.:61.25   3rd Qu.:121.0   3rd Qu.:30.00
Max.   :74.00   Max.   :67.00   Max.   :123.0   Max.   :30.00

    MFVPT           REYCOPY          REYRECAL        BVMT123
Min.   :13.00   Min.   :25.50   Min.   : 7.50   Min.   : 9.00
1st Qu.:17.00   1st Qu.:30.00   1st Qu.:12.63   1st Qu.:18.75
Median :18.00   Median :32.00   Median :17.25   Median :23.50
Mean   :18.21   Mean   :32.02   Mean   :16.77   Mean   :21.67
3rd Qu.:20.00   3rd Qu.:34.00   3rd Qu.:20.13   3rd Qu.:26.50
Max.   :22.00   Max.   :36.00   Max.   :25.00   Max.   :31.00

   BVMTDELY          NBACK        BaseTMTASCR      PreTMTASCR
```

```
Min.   : 3.000   Min.   :61.00   Min.   :19.00   Min.   :18.00
1st Qu.: 7.750   1st Qu.:73.00   1st Qu.:27.38   1st Qu.:26.00
Median : 9.000   Median :77.00   Median :32.50   Median :32.00
Mean   : 8.208   Mean   :77.63   Mean   :36.02   Mean   :33.60
3rd Qu.:10.000   3rd Qu.:82.50   3rd Qu.:42.00   3rd Qu.:36.25
Max.   :11.000   Max.   :91.00   Max.   :69.00   Max.   :79.00

 PostTMTASCR     BaseTMTBSCR       PreTMTBSCR      PostTMTBSCR
Min.   :13.00   Min.   : 38.00   Min.   : 29.00   Min.   : 29.00
1st Qu.:21.00   1st Qu.: 52.69   1st Qu.: 46.75   1st Qu.: 48.00
Median :24.00   Median : 68.50   Median : 59.00   Median : 63.50
Mean   :26.26   Mean   : 81.28   Mean   : 71.10   Mean   : 73.83
3rd Qu.:32.00   3rd Qu.: 90.00   3rd Qu.: 86.50   3rd Qu.: 82.50
Max.   :50.00   Max.   :243.00   Max.   :249.00   Max.   :299.00

 BaseDIGITSYM    PreDIGITSYM      PostDIGITSYM     BaseLNSSCORE
Min.   :32.00   Min.   : 30.00   Min.   : 42.00   Min.   : 6.000
1st Qu.:59.00   1st Qu.: 66.00   1st Qu.: 68.75   1st Qu.: 8.000
Median :62.00   Median : 71.00   Median : 73.00   Median :10.000
Mean   :65.38   Mean   : 73.69   Mean   : 76.60   Mean   : 9.875
3rd Qu.:77.25   3rd Qu.: 81.25   3rd Qu.: 84.50   3rd Qu.:11.000
Max.   :89.00   Max.   :108.00   Max.   :106.00   Max.   :14.000

 PreLNSSCORE     PostLNSSCORE    BaseRAVLT1       PreRAVLT1
Min.   : 4.00   Min.   : 6.0   Min.   : 4.000   Min.   : 2.000
1st Qu.: 9.00   1st Qu.: 9.0   1st Qu.: 5.000   1st Qu.: 5.000
Median :11.00   Median :10.0   Median : 6.000   Median : 6.000
Mean   :10.85   Mean   :10.5   Mean   : 6.417   Mean   : 6.208
3rd Qu.:12.00   3rd Qu.:12.0   3rd Qu.: 7.000   3rd Qu.: 7.000
Max.   :18.00   Max.   :17.0   Max.   :11.000   Max.   :14.000

  PostRAVLT1      BaseRAVLTSUM    PreRAVLTSUM     PostRAVLTSUM
Min.   : 3.000   Min.   :30.00   Min.   :28.00   Min.   :22.00
1st Qu.: 5.000   1st Qu.:40.75   1st Qu.:36.75   1st Qu.:35.00
Median : 6.000   Median :48.50   Median :44.00   Median :43.50
Mean   : 6.111   Mean   :47.13   Mean   :45.47   Mean   :42.76
3rd Qu.: 7.250   3rd Qu.:53.25   3rd Qu.:55.00   3rd Qu.:51.00
Max.   :12.000   Max.   :69.00   Max.   :69.00   Max.   :64.00

 BaseRAVLTDLY    PreRAVLTDLY      PostRAVLTDLY     BaseBUTTONS
Min.   : 3.00   Min.   : 2.000   Min.   : 0.000   Min.   :253.6
1st Qu.: 7.75   1st Qu.: 6.000   1st Qu.: 4.750   1st Qu.:295.5
Median :10.00   Median : 8.000   Median : 7.000   Median :342.0
Mean   : 9.50   Mean   : 8.347   Mean   : 7.306   Mean   :366.7
3rd Qu.:11.25   3rd Qu.:11.000   3rd Qu.:10.000   3rd Qu.:388.1
Max.   :15.00   Max.   :15.000   Max.   :15.000   Max.   :795.4
                                                  NA's   :  6.0
  PreBUTTONS      PostBUTTONS     BaseCHOOSER      PreCHOOSER
Min.   :214.6   Min.   :214.9   Min.   : 516.8   Min.   :460.2
1st Qu.:277.0   1st Qu.:269.6   1st Qu.: 631.8   1st Qu.:558.9
Median :299.7   Median :292.7   Median : 716.7   Median :641.1
Mean   :323.5   Mean   :312.1   Mean   : 714.8   Mean   :654.9
3rd Qu.:358.7   3rd Qu.:333.1   3rd Qu.: 784.5   3rd Qu.:728.1
Max.   :621.7   Max.   :629.8   Max.   :1087.8   Max.   :960.5
```

```
NA's   :  1.0
 PostCHOOSER      BaseTAPPER      PreTAPPER       PostTAPPER
Min.   :409.6   Min.   :34.00   Min.   :36.40   Min.   :37.00
1st Qu.:544.1   1st Qu.:43.55   1st Qu.:44.35   1st Qu.:44.15
Median :619.0   Median :46.40   Median :47.30   Median :47.00
Mean   :631.5   Mean   :46.52   Mean   :47.21   Mean   :47.07
3rd Qu.:709.2   3rd Qu.:49.55   3rd Qu.:49.75   3rd Qu.:49.65
Max.   :988.1   Max.   :56.40   Max.   :58.40   Max.   :60.80
                                NA's   : 4.00
  BaseMemory        PreMemory          PostMemory        BaseSpeed
Min.   :-1.1392   Min.   :-1.46795   Min.   :-1.7511   Min.   :-0.7733
1st Qu.:-0.2428   1st Qu.:-0.62157   1st Qu.:-0.7174   1st Qu.:-0.2124
Median : 0.2399   Median :-0.07198   Median :-0.2989   Median : 0.3303
Mean   : 0.2534   Mean   : 0.05761   Mean   :-0.1421   Mean   : 0.3932
3rd Qu.: 0.6641   3rd Qu.: 0.72262   3rd Qu.: 0.5777   3rd Qu.: 0.7352
Max.   : 2.2478   Max.   : 2.57727   Max.   : 1.7455   Max.   : 2.1129

   PreSpeed           PostSpeed          PreDrowsy4      PreDrowsy5
Min.   :-1.025151   Min.   :-1.23514   Min.   : 0.00   Min.   : 0.00
1st Qu.:-0.505140   1st Qu.:-0.77800   1st Qu.: 2.00   1st Qu.: 1.50
Median : 0.003962   Median :-0.25787   Median :12.00   Median : 9.00
Mean   : 0.068492   Mean   :-0.19957   Mean   :14.46   Mean   :11.35
3rd Qu.: 0.425330   3rd Qu.: 0.07826   3rd Qu.:22.00   3rd Qu.:16.00
Max.   : 2.581937   Max.   : 2.14517   Max.   :53.00   Max.   :66.00
                                       NA's   : 1.00   NA's   : 1.00
   PreImp4         PreImp5       PostDrowsy4     PostDrowsy5
Min.   : 0.00   Min.   : 0.00   Min.   : 0.00   Min.   : 0.00
1st Qu.: 4.50   1st Qu.: 2.00   1st Qu.: 6.50   1st Qu.: 5.50
Median :12.00   Median :11.00   Median :18.50   Median : 9.50
Mean   :15.56   Mean   :12.56   Mean   :22.33   Mean   :15.22
3rd Qu.:24.50   3rd Qu.:16.00   3rd Qu.:32.50   3rd Qu.:19.25
Max.   :55.00   Max.   :54.00   Max.   :64.00   Max.   :71.00
NA's   : 1.00   NA's   : 1.00   NA's   :36.00   NA's   :36.00
   PostImp4        PostImp5
Min.   : 0.00   Min.   : 0.00
1st Qu.: 9.75   1st Qu.: 6.75
Median :19.00   Median :13.50
Mean   :19.86   Mean   :15.61
3rd Qu.:28.00   3rd Qu.:18.25
Max.   :67.00   Max.   :65.00
NA's   :36.00   NA's   :36.00
```

# 4 Save the results

We now save the data as both an R worksheet and a CSV file that can be imported into practically any statistical program. In the `CSV` file, missing values will be empty cells.

```
R> save.image("merged.RData")
R> write.csv(merged, file = "merged.csv", row.names = FALSE, na = "")
```