

# STATS 419 Survey of Multivariate Analysis

## Convert measure.xls data from panel form and save as TXT file

Michael Smith  
([cornell.smith@wsu.edu](mailto:cornell.smith@wsu.edu))  
[————]

Instructor: Monte J. Shaffer

15 October 2020

```
path.measure = "C:/Program Files/Git/WSU_STATS419_FALL2020/project/";
# change this to wherever measure.xls lives for you...
# YOU SHOULD USE MINE IN THE DROPBOX AS THE EXAMPLE...
file.measure = paste0(path.measure,"measure.xlsx");
# BELOW, we will import yours ...

library(readxl);

covariates = read_excel(file.measure, sheet="covariates");
panel = read_excel(file.measure, sheet="panel");

# these are currently in "tibble form"

covariates = as.data.frame(covariates);
panel = as.data.frame(panel);

# install_github("MonteShaffer/humanVerseWSU/humanVerseWSU");
#library(humanVerseWSU);

panel.first = panel$`person_id`[1]; # 1c2408654ef5a2fe1fc962088312266c

temp = panel[panel$`person_id` == panel.first , ];

mySide = (panel$side);
myKeys = (panel$mkey);

myJointKeys = unique( paste0(myKeys,"-",mySide) );

## we know these are an exhaustive list of keys ...
# covariates have 13 variables ...
# side + jointkeys (23) + values
tabularized.df = data.frame( matrix(NA, nrow=0, ncol=37) );
tabularized.cols = c( colnames(covariates)[1:2], "side", myJointKeys, colnames(covariates)[3:13]);
```

```
colnames(tabularized.df) = tabularized.cols;
```

Using the sample file provided, we have all the known keys. Now, we will use those keys and “flatten” the covariates and panel into one dataframe that has lots of columns (37).

```
# HERE, we will import yours ...

path.your.measure = "C:/Program Files/Git/WSU_STATS419_FALL2020/WEEK-05"; # trailing slash or not?
file.your.measure = paste0(path.your.measure, "/", "measure_MSmith.xlsx"); # if no trailing slash, notice

covariates = read_excel(file.your.measure, sheet="covariates");
panel = read_excel(file.your.measure, sheet="panel");

# these are currently in "tibble form"

covariates = as.data.frame(covariates);
panel = as.data.frame(panel);

n.cov = dim(covariates)[1];
for(i in 1:n.cov)
{
  covariates.row = covariates[i,];

  final.row = as.data.frame( t(rep("NA",37)) );
  colnames(final.row) = tabularized.cols;

  final.row$data_collector = collector = covariates.row$data_collector;
  final.row$person_id = person = covariates.row$person_id;
  final.row[27:37] = covariates.row[3:13];

  subpanel = panel[panel$data_collector == collector & panel$person_id == person, ];
  n.sub = dim(subpanel)[1];
  for(j in 1:n.sub)
  {
    subpanel.row = subpanel[j,];
    s = subpanel.row$side;
    k = subpanel.row$mkey;
    jk = paste0(k, "-", s);
    v = subpanel.row$mvalue;
    final.row$side = s;
    final.row[jk] = v;
  }

  tabularized.df = rbind(tabularized.df, final.row);

}

tabularized.df;
```

```
##                data_collector                person_id  side
```

```

## 1 2061a45a083b486247280c306b37feb4 fb386c060cabd086675716cd4de9409f right
## 2 fb386c060cabd086675716cd4de9409f 2061a45a083b486247280c306b37feb4 right
## 3 6f8b7e70c1381ce6c51419d1292b3701 a5cb2bc1b66c498981834e5488d4c762 right
## 4 128dbfc653511a24d93997418cc816a4 c18e1211f0fc8ba938e878bf786960e5 right
## 5 45bc06f2f391861e8e66aa70cec5e0d8 fb9952be581f009f0bf2206ca677f07d right
## 6 6f8b7e70c1381ce6c51419d1292b3701 341d68b5f84e711c2c974a202f11ba11 right
## 7 6f8b7e70c1381ce6c51419d1292b3701 7d4eeb7c6a177f908724d8d3f4bbb535 right
## 8 341d68b5f84e711c2c974a202f11ba11 6f8b7e70c1381ce6c51419d1292b3701 right
## 9 c18e1211f0fc8ba938e878bf786960e5 128dbfc653511a24d93997418cc816a4 right
## 10 fb9952be581f009f0bf2206ca677f07d 45bc06f2f391861e8e66aa70cec5e0d8 right
## height-NA head.height-NA head.circumference-NA hand.length-left
## 1 61.00 8.50 21.25 7.00
## 2 60.00 9.00 23.50 6.50
## 3 54.00 8.50 20.50 5.75
## 4 69.50 7.50 22.75 7.50
## 5 68.50 10.00 21.50 6.50
## 6 68.50 10.00 21.50 6.50
## 7 63.00 54.75 22.50 7.25
## 8 71.00 62.00 23.00 8.00
## 9 61.00 9.50 23.00 6.50
## 10 71.75 11.00 23.00 8.00
## hand.length-right hand.width-left hand.width-right hand.elbow-left
## 1 7.00 7.25 7.50 16.00
## 2 6.50 8.00 8.25 15.50
## 3 6.00 7.00 6.50 14.00
## 4 7.50 8.00 8.25 19.00
## 5 6.25 7.50 7.25 15.50
## 6 6.25 7.50 7.25 15.50
## 7 7.25 8.50 8.50 16.00
## 8 8.00 9.00 9.00 19.00
## 9 7.00 6.75 7.25 15.75
## 10 8.00 8.50 8.50 20.00
## hand.elbow-right elbow.armpit-left elbow.armpit-right arm.reach-left
## 1 16.0 9.5 9.50 77.50
## 2 15.0 9.0 9.50 76.00
## 3 14.0 9.5 9.00 67.00
## 4 18.5 15.0 15.00 87.25
## 5 15.0 12.5 12.25 78.50
## 6 15.0 12.5 12.25 78.50
## 7 16.0 12.0 12.00 80.25
## 8 19.0 55.5 55.50 89.50
## 9 16.0 8.5 9.00 77.00
## 10 20.0 13.0 13.00 92.00
## arm.reach-right arm.span-NA foot.length-left foot.length-right
## 1 78.50 62.0 9.5 9.50
## 2 75.00 62.0 8.5 8.75
## 3 66.00 52.5 9.0 8.50
## 4 86.75 69.0 10.5 10.00
## 5 78.00 60.0 8.5 8.00
## 6 78.00 60.0 8.5 8.00
## 7 80.25 3.5 9.5 9.50
## 8 89.50 3.5 11.0 11.00
## 9 76.25 56.0 9.5 9.50
## 10 92.00 73.5 11.0 11.25

```

```

##      floor.kneepit-left floor.kneepit-right floor.hip-left floor.hip-right
## 1          18.00          18.00          36.00          37.00
## 2          16.00          15.50          35.00          34.50
## 3          15.00          14.00          31.00          30.50
## 4          18.00          17.75          41.00          40.75
## 5          19.00          19.00          39.00          39.25
## 6          19.00          19.00          39.00          39.25
## 7          17.00          17.00          30.25          30.25
## 8          19.00          19.00          40.00          40.00
## 9          15.75          16.00          34.00          33.50
## 10         21.00          20.50          44.50          44.00
##      floor.navel-NA floor.earmpit-left floor.earmpit-right units writing eye
## 1          37.00          51.00          51.00    in    right right
## 2          35.50          47.00          47.00    in    right right
## 3          31.25          40.50          40.50    in    right right
## 4          40.00          56.00          56.50    in    right right
## 5          38.00          50.00          50.00    in    right right
## 6          38.00          50.00          50.00    in    right left
## 7          37.25          51.00          51.00    in    right left
## 8          42.00          54.75          54.75    in    right left
## 9          34.00          47.50          48.00    in    right right
## 10         44.50          55.00          55.00    in    right right
##      eye_color swinging age gender quality minutes ethnicity
## 1      brown    right  12  male     10     20  indian, caucasian
## 2      blue    right  21 female     10     21    caucasian
## 3      green    right  11  male     9     18  caucasian, hispanic
## 4      hazel    right  72  male     9     19    caucasian
## 5      brown    right  66 female     9     21    caucasian
## 6      brown    right  48 female    10     21  caucasian, hispanic
## 7      hazel    right  14 female     9     18  caucasian, hispanic
## 8      blue    right  43  male     8     17  caucasian, hispanic
## 9      brown    right  69 female     9     19  caucasian, hispanic
## 10     hazel    right  69  male    10     22    caucasian
##
##                                     notes
## 1 Possible ancestry may include: Indian, Polish, French, Israely, Caucasian, Croatian
## 2                                     <NA>
## 3                                     <NA>
## 4                                     <NA>
## 5                                     <NA>
## 6                                     <NA>
## 7                                     <NA>
## 8                                     <NA>
## 9                                     <NA>
## 10                                    <NA>

```

```
str(tabularized.df);
```

```

## 'data.frame':  10 obs. of  37 variables:
## $ data_collector      : chr  "2061a45a083b486247280c306b37feb4" "fb386c060cabd086675716cd4de9409f"
## $ person_id           : chr  "fb386c060cabd086675716cd4de9409f" "2061a45a083b486247280c306b37feb4"
## $ side                 : chr  "right" "right" "right" "right" ...
## $ height-NA            : num  61 60 54 69.5 68.5 ...
## $ head.height-NA       : num  8.5 9 8.5 7.5 10 ...
## $ head.circumference-NA: num  21.2 23.5 20.5 22.8 21.5 ...

```

```
## $ hand.length-left      : num  7 6.5 5.75 7.5 6.5 6.5 7.25 8 6.5 8
## $ hand.length-right     : num  7 6.5 6 7.5 6.25 6.25 7.25 8 7 8
## $ hand.width-left       : num  7.25 8 7 8 7.5 7.5 8.5 9 6.75 8.5
## $ hand.width-right      : num  7.5 8.25 6.5 8.25 7.25 7.25 8.5 9 7.25 8.5
## $ hand.elbow-left       : num  16 15.5 14 19 15.5 ...
## $ hand.elbow-right      : num  16 15 14 18.5 15 15 16 19 16 20
## $ elbow.armpit-left     : num  9.5 9 9.5 15 12.5 12.5 12 55.5 8.5 13
## $ elbow.armpit-right    : num  9.5 9.5 9 15 12.2 ...
## $ arm.reach-left        : num  77.5 76 67 87.2 78.5 ...
## $ arm.reach-right       : num  78.5 75 66 86.8 78 ...
## $ arm.span-NA           : num  62 62 52.5 69 60 60 3.5 3.5 56 73.5
## $ foot.length-left      : num  9.5 8.5 9 10.5 8.5 8.5 9.5 11 9.5 11
## $ foot.length-right     : num  9.5 8.75 8.5 10 8 ...
## $ floor.kneepit-left    : num  18 16 15 18 19 ...
## $ floor.kneepit-right   : num  18 15.5 14 17.8 19 ...
## $ floor.hip-left        : num  36 35 31 41 39 ...
## $ floor.hip-right       : num  37 34.5 30.5 40.8 39.2 ...
## $ floor.navel-NA        : num  37 35.5 31.2 40 38 ...
## $ floor.armpit-left     : num  51 47 40.5 56 50 ...
## $ floor.armpit-right    : num  51 47 40.5 56.5 50 ...
## $ units                 : chr  "in" "in" "in" "in" ...
## $ writing                : chr  "right" "right" "right" "right" ...
## $ eye                   : chr  "right" "right" "right" "right" ...
## $ eye_color             : chr  "brown" "blue" "green" "hazel" ...
## $ swinging              : chr  "right" "right" "right" "right" ...
## $ age                   : num  12 21 11 72 66 48 14 43 69 69
## $ gender                : chr  "male" "female" "male" "male" ...
## $ quality               : num  10 10 9 9 9 10 9 8 9 10
## $ minutes               : num  20 21 18 19 21 21 18 17 19 22
## $ ethnicity             : chr  "indian, caucasian" "caucasian" "caucasian, hispanic" "caucasian" ...
## $ notes                 : chr  "Possible ancestry may include: Indian, Polish, French, Israely, Cauc"
```

```
## verify your data ... make certain it seems to make sense
```

```
### Data Confirmed
```

```
## make certain notes or other fields don't have the "/" pipe character
```

```
### Pipe char not present
```

```
# save as a TXT file, pipe-delimited using your data_collector id
```

```
collector = "8e77a4a615183225a08c495fe0e88d62";
```

```
outfile = paste0(path.measure, "/", "measure", "-", collector, ".txt");
```

```
utils::write.table( tabularized.df , file=outfile, quote=FALSE, col.names=TRUE, row.names=FALSE, sep="|"
```

```
# SUBMIT THIS FILE as measure-txt homework...
```