BST5220 multilevel HW1

Due Wednesday by 12:00 pm, 2/11

Miao Cai *miao.cai@slu.edu*

2/4/2019

You will first need to run the SAS code (homework1_data.sas) to create the SAS data set. The purpose of this study is to assess how job-related stress is related to hospital size (0=small, 1=medium, 2=large) and nurse's job experience (years). Data are from nurses working in 25 hospitals. In each hospital, a sample of about 40 nurses is selected and given a test that measures job-related stress (coded on a scale of 0-7).

1. Identify the type of the data structure (cross-sectional clustered, longitudinal, or clustered longitudinal).
2. Identify the variables at each level.
3. Graphically examine the association between job stress and nurse's experience within each hospital. Refer to slides 8-11, lecture 2.
4. Use the model building strategies discussed in lecture 3 to select the best model for the data

```
dat = rio::import("data/a1.sas7bdat")
head(dat)
```

```
##   hospital hospital_size nurse stress experience
## 1        1             2     1      7         11
## 2        1             2     2      7         20
## 3        1             2     3      7          7
## 4        1             2     4      6         25
## 5        1             2     5      6         22
## 6        1             2     6      6         22
```

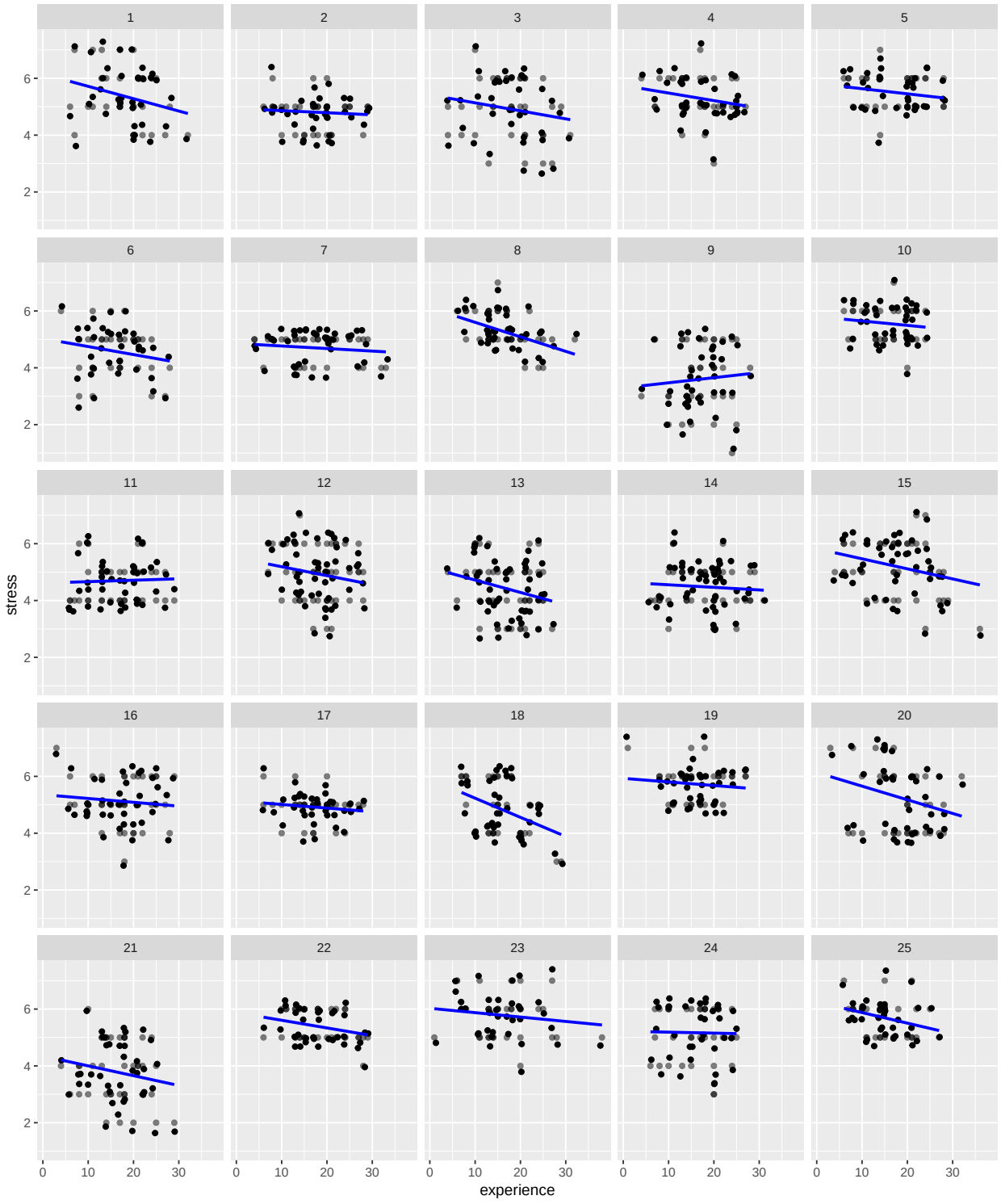
1. this is a cross-sectional data struture
2. Variables at each level:

- Hospital level: hospital, hospital_size
- nurse, stress, experience

3. Graphically examine the association between job stress and nurse's experience within each hospital

```
pacman::p_load(tidyverse)

dat %>%
  ggplot(aes(experience, stress,)) + geom_point(alpha = 0.5) + geom_jitter() +
  geom_smooth(method = "lm", se=FALSE, color="blue") +
  facet_wrap(.~hospital, ncol = 5)
```




```
pacman::p_load(rstanarm)

fit0 = stan_lmer(stress ~ (1|hospital), data = dat,
                 chains = 1, iter = 1000, warmup = 500)
```

```
fit0
```

```
## stan_lmer
##  family:       gaussian [identity]
##  formula:      stress ~ (1 | hospital)
##  observations: 1000
## ------
##             Median MAD_SD
## (Intercept) 5.0    0.1
##
## Auxiliary parameter(s):
##       Median MAD_SD
## sigma 0.8    0.0
##
## Error terms:
##  Groups   Name        Std.Dev.
##  hospital (Intercept) 0.58
##  Residual             0.83
## Num. levels: hospital 25
##
## Sample avg. posterior predictive distribution of y:
##          Median MAD_SD
## mean_PPD 5.0    0.0
##
## ------
## * For help interpreting the printed output see ?print.stanreg
## * For info on the priors used see ?prior_summary.stanreg
```