Multilevel Mixed Models on Music Perfromance Anxiety Analysis

Introduction

Case Design

The study explored the emotional state of musicians before performances and factors which may affect their emotional state. The subjects here were 37 undergraduate music majors from a competitive undergraduate music program fill out diaries prior to performances over the course of an academic year.

Factors which were examined for their potential relationships with performance anxiety included: performance type (solo, large ensemble, or small ensemble); audience (instructor, public, students, or juried); if the piece was played from memory; age; gender; instrument (voice, orchestral, or keyboard); and, years studying the instrument. In addition, the personalities of study participants were assessed at baseline through the Multidimensional Personality Questionnaire (MPQ). The MPQ provided scores for one lower-order factor (absorption) and three higher-order factors: positive emotionality (PEM—a composite of well-being, social potency, achievement, and social closeness); negative emotionality (NEM—a composite of stress reaction, alienation, and aggression); and, constraint (a composite of control, harm avoidance, and traditionalism).

Variables

There interested factor information were collected and described in variables as follows:

- `id` = unique musician identification number
- `diary` = cumulative total of diaries filled out by musician
- `previous` = number of previous performances with a diary entry
- `perf_type` = type of performance (Solo, Large Ensemble, or Small Ensemble)
- `audience` = who attended (Instructor, Public, Students, or Juried)
- `memory` = performed from Memory, using Score, or Unspecified
- `na` = negative affect score from PANAS
- `gender` = musician gender
- `instrument` = Voice, Orchestral, or Piano
- `years_study` = number of previous years spent studying the instrument
- `mpqab` = absorption subscale from MPQ
- `mpqsr` = stress reaction subscale from MPQ
- `mpqpem` = positive emotionality (PEM) composite scale from MPQ
- `mpqnem` = negative emotionality (NEM) composite scale from MPQ
- `mpqcon` = constraint scale from MPQ

Hypotheses

Primary scientific hypotheses of the researchers included:

- Lower music performance anxiety will be associated with lower levels of a subject's negative emotionality.
- Lower music performance anxiety will be associated with lower levels of a subject's stress reaction.
- Lower music performance anxiety will be associated with greater number of years of study.

Methodologies

Multilevel Model

The models we used for this study was the multilevel model. With multilevel models, exploratory analyses must eventually account for the level at which each variable is measured. In a two-level study such as this one, *Level One* will refer to variables measured at the most frequently occurring observational unit, while *Level Two* will refer to variables measured on larger observational units. For example in this study, level two factors can be variables measure characteristics of participants that remain constant over all performances for a particular musician

Specifically in this study, the level one and level two factors can include as the follows:

- Level One:
  - positive affect (our response variable)
  - performance characteristics (type, audience, if music was performed from memory)
  - number of previous performances with a diary entry
- Level Two:
  - demographics (age and gender of musician)
  - instrument used and number of previous years spent studying that instrument
  - baseline personality assessment (MPQ measures of positive emotionality, negative emotionality, constraint, stress reaction, and absorption)

Analytics Plans

- Exploratory data analysis (question a)
  - Explore the distributions of interested response, level one and level two variables, as well as their correlations with respect to response variable
- Multilevel models (question b-g)
  - Fit predicitive models with level one or level two factors, make interpreations on the models
- Model Comparison (question h)
  - Explore differences between predictive models, make comparisons on their performances

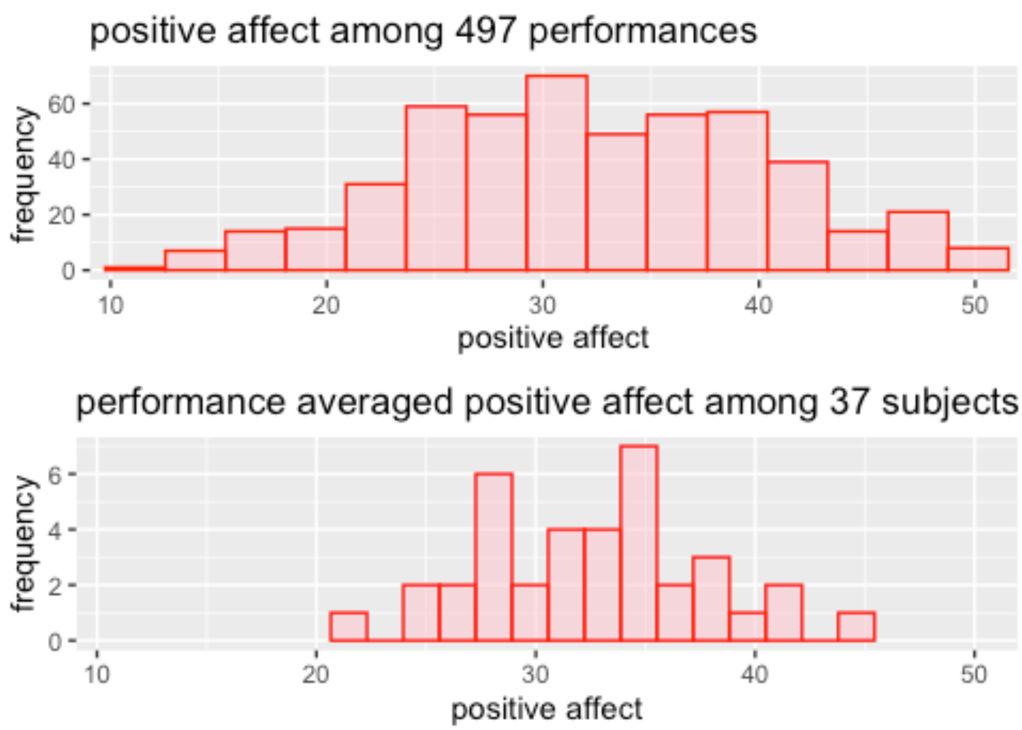
Results

Exploratory Data Analysis (a)

-- Response variable: positive affect

Firstly, it's about the exploratory analysis in response variable `pa`. Here in figure1, the distributions of `pa` among 497 performances as well as their averaged value among 37 subjects were shown as follows:

figure1



From figure1, we can see that he positive affect was nearly symmetrically distributed.

-- Level one categorical variables: performance type, audience, if music was performed from memory

`perform_type`, `audience` and `memory` are level one categorical variables we were interested in this study. To get an overview about their distributions, I summarised their frequency tables as well as visualized their bar charts, which were shown in table1-3 and figure2.

table1

`perform_type`	Large ensemble	Small Ensemble	Solo
frequency	136	82	279

table2

`audience`	Instructor	Juried Recital	Public Performance	Student(s)
frequency	149	44	204	100

table3

`memory`	Memory	Score	Unspecified
frequency	149	274	74

figure2

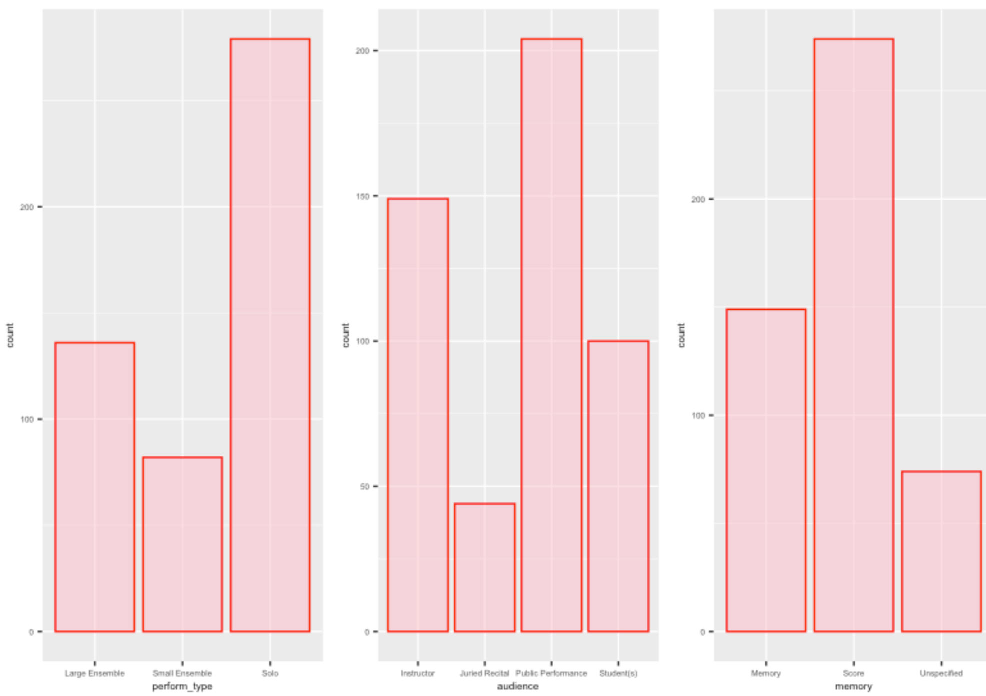
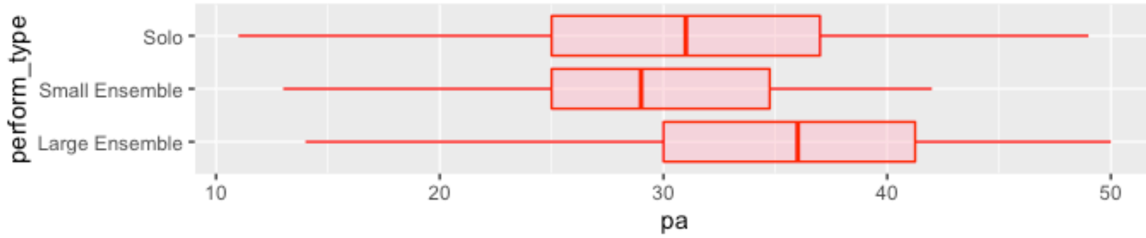
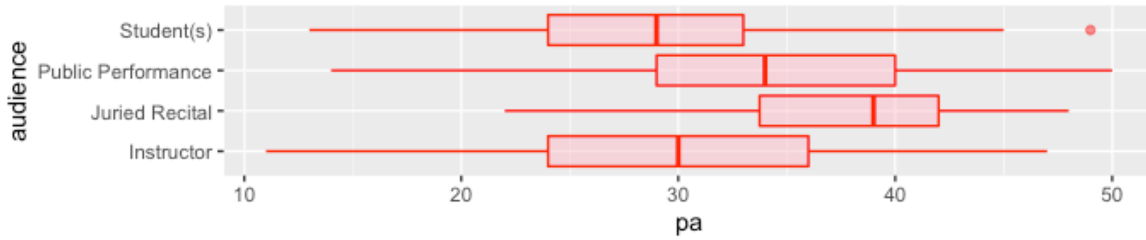
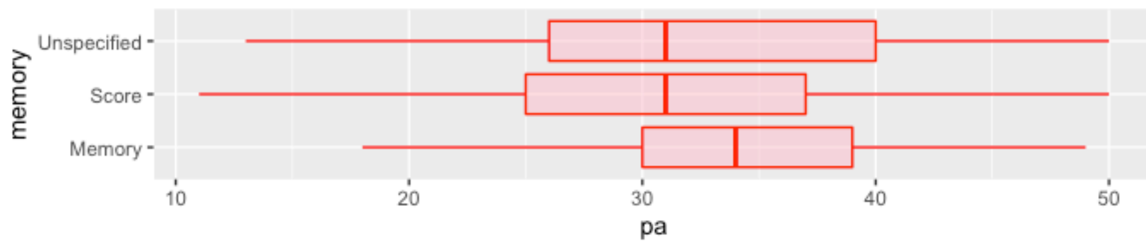


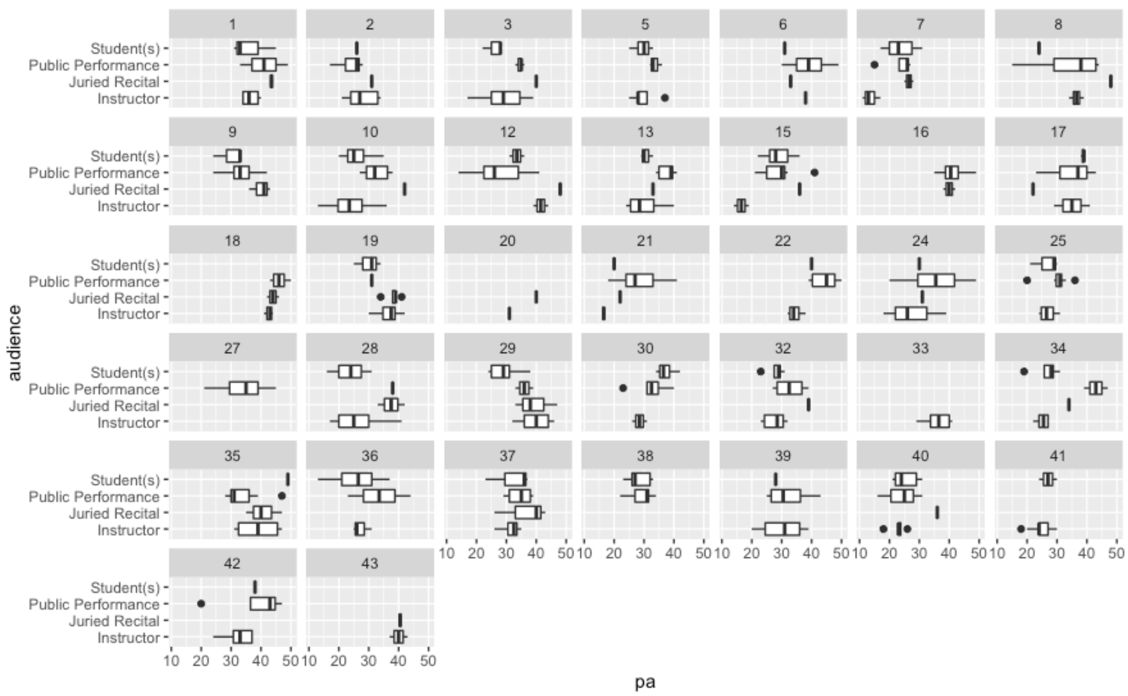
Table1-3 and figure2 showed the overall distributions of these level one categorical variables in the study samples. To roughly explore their corelations with our response variable `pa`. Boxplots were plotted and shown in figue3.

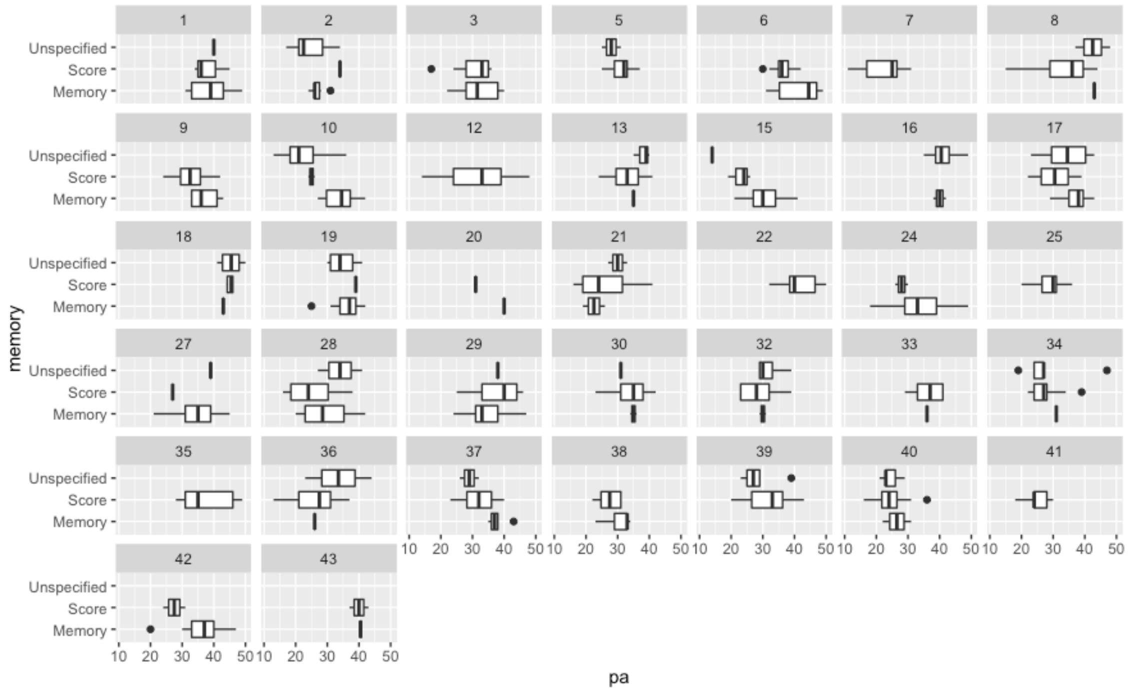
figure3

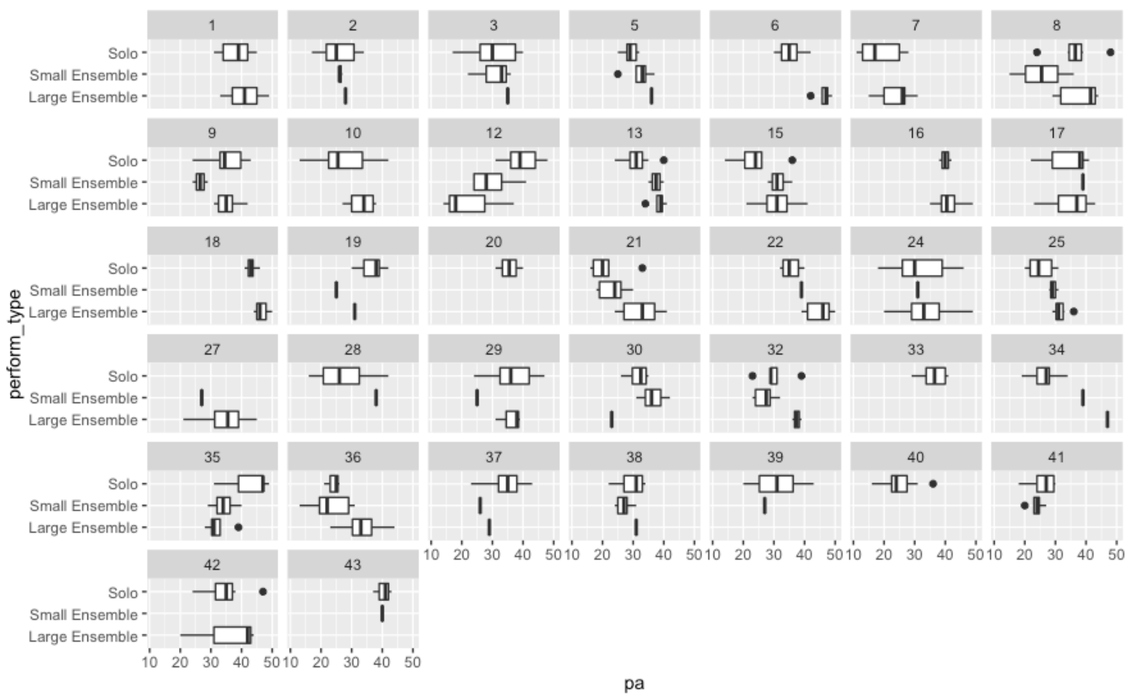

perform_type
Solo
Small Ensemble
Large Ensemble
10
20
30
40
50
pa


audience
Student(s)
Public Performance
Juried Recital
Instructor
10
20
30
40
50
pa


memory
Unspecified
Score
Memory
10
20
30
40
50
pa


1
2
3
5
6
7
8
9
10
12
13
15
16
17
18
19
20
21
22
24
25
27
28
29
30
32
33
34
35
36
37
38
39
40
41
42
43
audience
Student(s)
Public Performance
Juried Recital
Instructor
10 20 30 40 50
pa





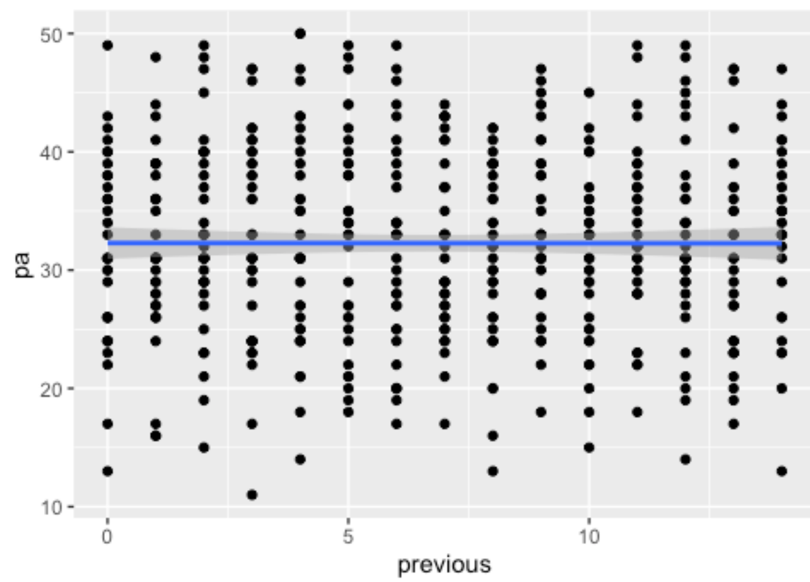
From figure3, we can see that positive affects from different category types can vary strongly. For example, the *large ensemble* performance type may have higher positive affect; the *student or instructor* audience may be related to lower positive affect; while memorized performance can have higher positive affect. While in terms of individual, the effects from perform_type, audience and memory may vary, which is also quite reasonable.

-- Level one numeric variable: number of previous performances with a diary entry number of previous performances with a diary entry `previous` was also an interested factor explored here. The 5 number statistical summaries were shown in table4. Also, scatterplot was plotted here to see its correlation with `pa`, which was displayed in figure4.

table4

Min	1st Qu	Median	3rd Qu	Max
0	3	7	10	14

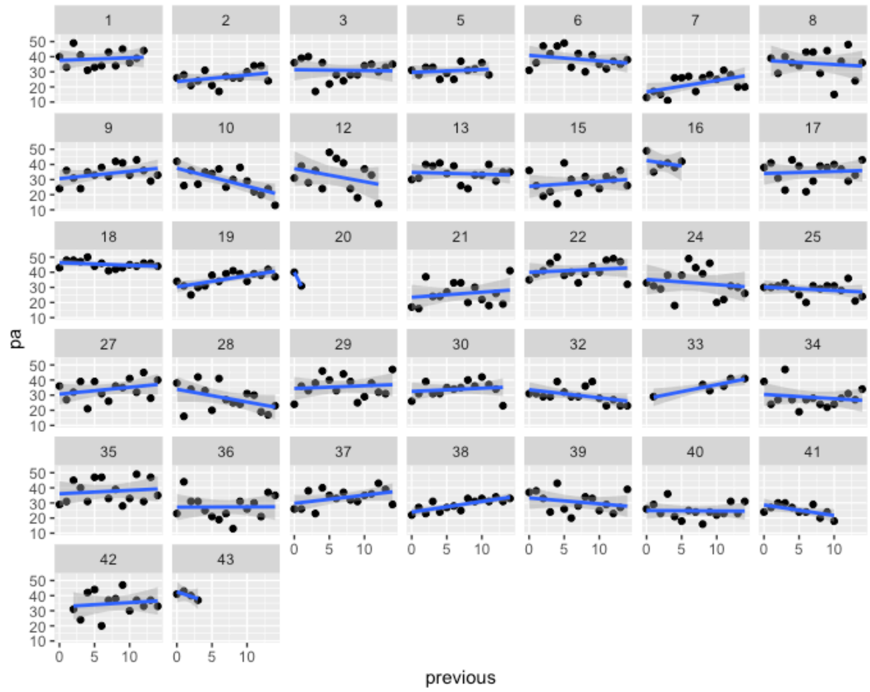
figure4



From figure4, it seemed that the number of previous performance on a diary may not so connected with positive affect in terms of performance.

To explore whether the `previous` correlated with `pa` in terms of subject. A lattice plot was produced and shown in figure5.

figure5



From figure5, we can see that the correlation between previous number of performance and positive affect may vary for different people.

-- Level two categorical variables: `gender` , instrument used (`instrument`)

Frequency tables and bar charts were generated to show the overall distributions of these 2 variables, we can find them in table5-6, and figure6.

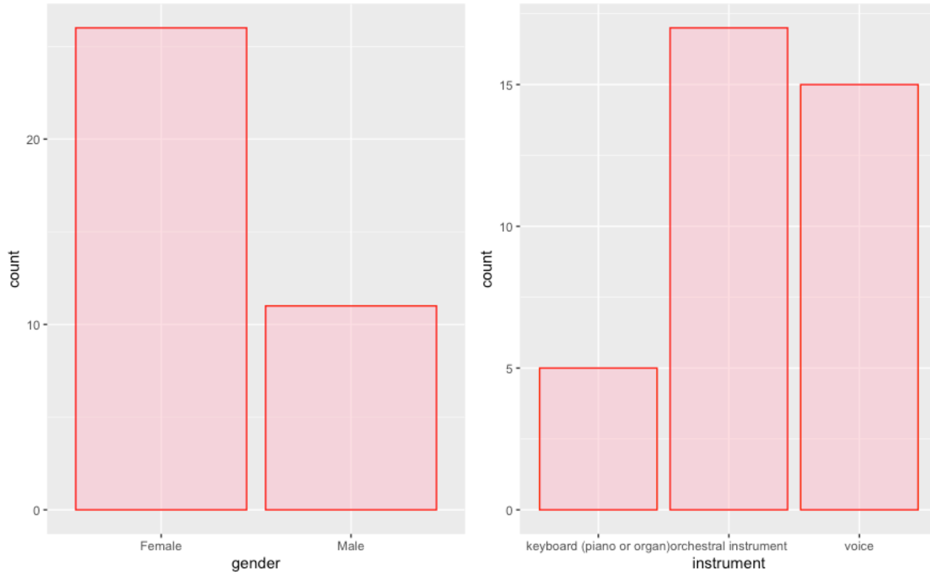
table5

`gender`	**Female**	**Male**
Frequency	26	11

table6

`instrument`	**keyboard**	**Orchestral**	**vocie**
frequency	5	17	15

figure6



Also, their correlations with response variable were shown in figure7-8.

figure7

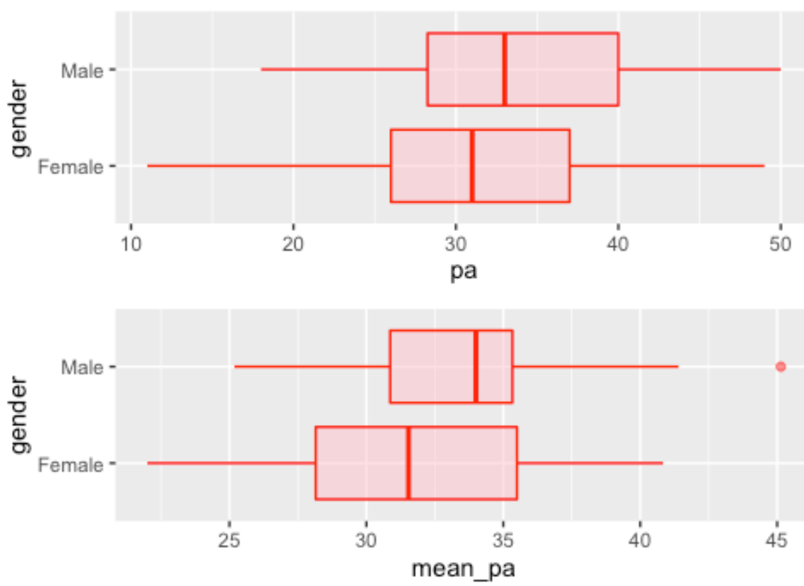
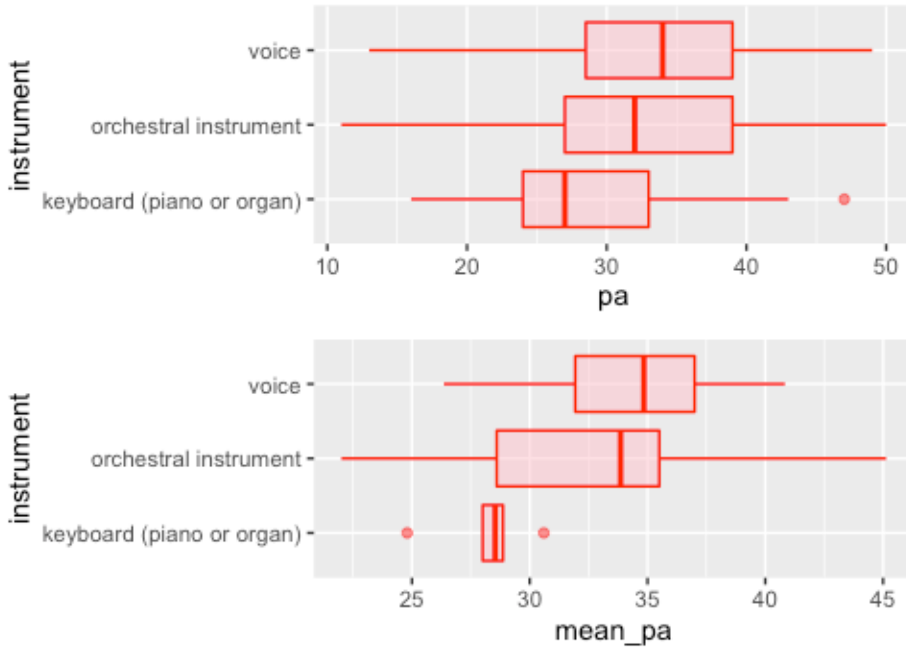


figure8

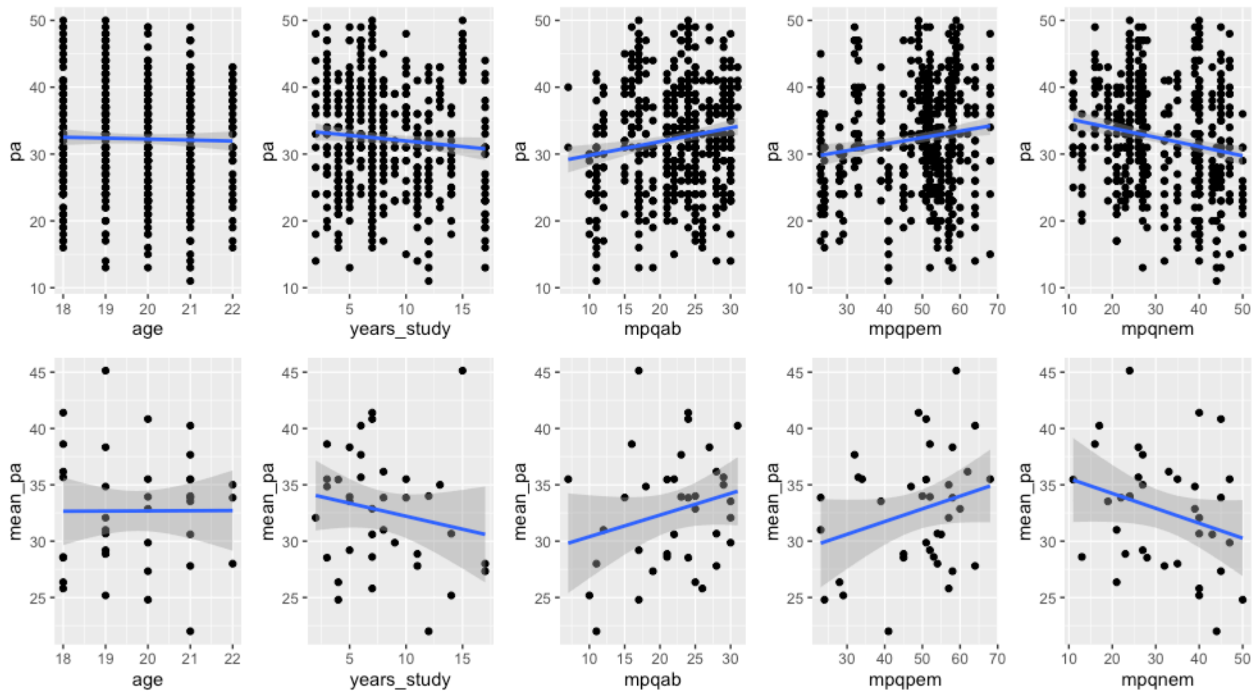


From figure6-7, we can see that the male musicians seemed to have higher positive affect either in performance level or in subject level. While in terms of instrument, the keyboard performers seem to have lower positive affect either in performance level or subject level.

-- Level two numeric variables: `age`, previous years spent studying that instrument(`years_study`), MPQ compsite scales

To see the correlations between numeric level two variables and response variable, scatterplots with linear regression smooth line were produced for a better visualization. Here in figure9, the top row showed scatter points in terms of each performance, while the bottom row showed points in terms of each subject, where the response variable changed accordingly into mean of positive affect for each subject.

figure9



From figure8, we can see that `age` seemed not to be so associated with `pa`. While other factors like `years_study`, `mpqab`, `mpqpem` and `mpqnem` may potentially closely correlated with positive affect.

Statistical Analysis

-- (**b**) Model A: unconditional means model

- Level One:

$$Y_{ij} = a_i + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma^2)$$

- Level Two:

$$a_i = \alpha_0 + u_i, u_i \sim N(0, \sigma_u^2)$$

-

```
Random effects:
 Groups   Name        Variance Std.Dev.
 id       (Intercept) 23.72    4.871
 Residual             41.70    6.457
Number of obs: 497, groups:  id, 37

Fixed effects:
            Estimate Std. Error t value
(Intercept)  32.5622     0.8584   37.93
```

$\hat{\alpha}_0$ = 32.5622 = the estimated mean positive affect across all performances and all subjects.

$\hat{\sigma}^2$ = 41.70 = the estimated variance in within-person deviations.

$\hat{\sigma}_u^2$ = 23.72 = the estimated variance in between-person deviations.

intraclass correlation coefficient:

$$\hat{\rho} = \frac{\text{Between-person variability}}{\text{Total variability}} = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}^2} = \frac{23.72}{23.72 + 41.70} = 0.3626$$

Thus, 36.26% the total variability in positive affect are attributable to differences among subjects. Here ρ approaches 0, which means responses from an individual are essentially independent and accounting for the multilevel structure of the data becomes less crucial, also, the effective sample size approaches the total number of observations.

-- (**c**) Model B: indicator for instructor audience type and indicator for student audience type at Level One; no Level Two predictors

- Level One:

$$Y_{ij} = a_i + b_i InstrAud_{ij} + c_i StudAud_{ij} + \epsilon_{ij}$$

- Level Two:

$$\begin{aligned} a_i &= \alpha_0 + u_i \\ b_i &= \beta_0 + v_i \\ c_i &= \gamma_0 + w_i \end{aligned}$$

Where $\epsilon_{ij} \sim N(0, \sigma^2)$ and

$$\begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & & \\ \rho\sigma_u\sigma_v & \sigma_v^2 & \\ \rho\sigma_u\sigma_w & \rho\sigma_v\sigma_w & \sigma_w^2 \end{bmatrix} \right)$$

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 20.34    4.509
          ia          11.61    3.407     0.08
          sa          12.08    3.475    -0.63  0.72
 Residual             36.39    6.033
Number of obs: 497, groups:  id, 37

Fixed effects:
            Estimate Std. Error t value
(Intercept)  34.7296     0.8604  40.366
ia           -4.1901     0.9271  -4.520
sa           -4.4540     0.9735  -4.575
```

$\hat{\alpha}_0 = 34.7296, \hat{\beta}_0 = -4.1901, \hat{\gamma}_0 = -4.4540$
$\hat{\sigma}_u^2 = 20.34, \hat{\sigma}_v^2 = 11.61, \hat{\sigma}_w^2 = 12.08, \sigma^2 = 36.39$

Where $\hat{\alpha}_0 = 34.7296$ means expected positive affect across all performance, $\hat{\beta}_0 = -4.1901$ means the positive affect decreases by 4.1901 unit if the audience is instructor among all performance. $\hat{\sigma}_u^2 = 20.34$ means estimated variance in between person deviation

$$\text{Pseudo } R^2_{L1} = \frac{\hat{\sigma}^2(ModelA) - \hat{\sigma}^2(ModelB)}{\hat{\sigma}^2(ModelA)} = \frac{41.7 - 36.39}{41.7} = 0.1273$$

Pseudo R^2_{L2} shows an improvment in Model B over Model A with addition of instructor audience indicator and student audience indicator, which has decreased the between-person variability in mean positive affect by 12.73%.

-- (**d-e**) Model C: indicator for instructor audience type and indicator for student audience type at Level One; centered MPQ absorption subscale as Level Two predictor for intercept and all slope terms

- Level One:

$$Y_{ij} = a_i + b_i InstrAud_{ij} + c_i StudAud_{ij} + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma^2)$$

- Level Two:

$$\begin{aligned} a_i &= \alpha_0 + \alpha_1 MPQab_i + u_i \\ b_i &= \beta_0 + \beta_1 MPQab_i + v_i \\ c_i &= \gamma_0 + \gamma_1 MPQab_i + w_i \\ u_i \sim N(0, \sigma_u^2);\ v_i &\sim N(0, \sigma_v^2);\ w_i \sim N(0, \sigma_w^2) \end{aligned}$$

- As a composite model:

$$\begin{aligned} Y_{ij} = [&\alpha_0 + \alpha_1 MPQab_i \\ &+ \beta_0 InstrAud_{ij} + \beta_1 MPQab_i InstrAud_{ij} \\ &+ \gamma_0 StudAud_{ij} + \gamma_1 MPQab_i StudAud_{ij}] \\ &+ [u_i + v_i InstrAud_{ij} + w_i StudAud_{ij} + \epsilon_{ij}] \end{aligned}$$

There're 10 parameters (13 parameters if also include variances between subjects) to estimate, the estimated parameters were shown below:

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 20.321   4.508
          ia           8.062   2.839     0.09
          sa          10.725   3.275    -0.73  0.60
 Residual             36.487   6.040
Number of obs: 497, groups:  id, 37

Fixed effects:
             Estimate Std. Error t value
(Intercept)  35.31895    3.38607  10.431
ia          -12.36795    3.38355  -3.655
```

```
sa           -11.07164    3.74141  -2.959
mpqab         -0.02279    0.14711  -0.155
ia:mpqab       0.36914    0.14793   2.495
sa:mpqab       0.29220    0.16314   1.791

Correlation of Fixed Effects:
         (Intr) ia     sa     mpqab  i:mpqb
ia       -0.340
sa       -0.632  0.536
mpqab    -0.967  0.326  0.610
ia:mpqab  0.325 -0.966 -0.513 -0.325
sa:mpqab  0.608 -0.514 -0.967 -0.624  0.519
```

$\hat{\alpha}_0 = 35.3190$, the estimated mean positive affect for public performance or juried recital audiences (`InstrAud=0, StudAud=0`) is 35.3190 with an average level of absorption at baseline (`mpqab=21.9859`)

$\hat{\alpha}_1 = -0.0228$, for performances with public performance or juried recital audiences, as `mpqab` score 1 unit higher than average level of absorption, the estimated mean positive affect lowers by 0.0228 point.

$\hat{\gamma}_0 = -11.0716$, under an average level of baseline absorption (`mpqab`=21.9859), performances with student audience have an estimated mean positive affect 11.0716 points lower than that with public performance or juried recital audience.

$\hat{\beta}_1 = 0.3691$, with instructor audience, as MPQ absorption level 1 unit higher than average, the estimated mean positive affect lower by 0.3691 point.

$\hat{\sigma}_u = 4.508$, with audience being public performance or juried recital, the estimated standard deviation between person is in deviation of 4.508 on positive affect

$\hat{\sigma}_v = 2.839$, the estimated standard deviation between person on postive affect for instructor audience is in deviation of 2.389 higher than that of public performance or juried recital audience

$\hat{\rho}_{uv} = -0.340$, there's negative correlation between a subject's positive affect and whether or not the audience are instructors, typically, the instructor audience would decrease the positive affect by 0.34.

-- (**f-g**) Model D: indicator for instructor audience type and indicator for student audience type at Level One; centered MPQ absorption subscale and a male indicator as Level Two predictors for intercept and all slope terms

- Level One:

$$Y_{ij} = a_i + b_i InstrAud_{ij} + c_i StudAud_{ij} + \epsilon_{ij}, \epsilon_{ij} \sim N(0, \sigma^2)$$

- Level Two:

$$
\begin{aligned}
a_i &= \alpha_0 + \alpha_1 MPQab_i + \alpha_2 male_i + u_i \\
b_i &= \beta_0 + \beta_1 MPQab_i + \beta_2 male_i + v_i \\
c_i &= \gamma_0 + \gamma_1 MPQab_i + \gamma_2 male_i + w_i \\
u_i \sim N(0, \sigma_u^2);\ v_i &\sim N(0, \sigma_v^2);\ w_i \sim N(0, \sigma_w^2)
\end{aligned}
$$

- As a composite model:

$$
\begin{aligned}
Y_{ij} = [&\alpha_0 + \alpha_1 MPQab_i + \alpha_2 male_i \\
&+ \beta_0 InstrAud_{ij} + \beta_1 MPQab_i InstrAud_{ij} + \beta_2 male_i InstrAud_{ij} \\
&+ \gamma_0 StudAud_{ij} + \gamma_1 MPQab_i StudAud_{ij}] + \gamma_2 male_i StudAud_{ij} \\
&+ [u_i + v_i InstrAud_{ij} + w_i StudAud_{ij} + \epsilon_{ij}]
\end{aligned}
$$

There're 13 parameters (16 parameters if also include the variances between subjects) to estimate, the estimated parameters were shown below:

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 id       (Intercept) 18.53    4.304
          ia           8.19    2.862     0.28
          sa          10.72    3.274    -0.75  0.43
 Residual             36.41    6.034
Number of obs: 497, groups:  id, 37

Fixed effects:
             Estimate Std. Error t value
(Intercept)  35.16859    3.28182  10.716
ia          -12.49970    3.39174  -3.685
sa          -11.17194    3.72231  -3.001
mpqab        -0.05677    0.14278  -0.398
male          3.04969    1.80837   1.686
ia:mpqab      0.39935    0.14835   2.692
ia:male      -2.20752    1.94755  -1.133
sa:mpqab      0.29307    0.16477   1.779
sa:male      -0.21167    2.05087  -0.103

Correlation of Fixed Effects:
         (Intr) ia     sa     mpqab  male   i:mpqb ia:mal s:mpqb
ia       -0.256
sa       -0.641  0.479
mpqab    -0.953  0.251  0.613
male     -0.075 -0.043  0.031 -0.095
ia:mpqab  0.251 -0.953 -0.457 -0.259  0.071
ia:male  -0.047 -0.082 -0.007  0.074 -0.153 -0.082
sa:mpqab  0.602 -0.451 -0.951 -0.629  0.071  0.466 -0.060
sa:male   0.035 -0.005 -0.014  0.069 -0.601 -0.065  0.362 -0.167
```

$\hat{\alpha}_0 = 35.1686$, the estimated mean positive affect for public performance or juried recital audiences (`InstrAud=0, StudAud=0`) is 35.1686 with an average level of absorption and male at baseline (`mpqab=21.9859`, `male=0.3179`).

$\hat{\alpha}_1 = -0.0568$, for performances with public performance or juried recital audience, as `mpqab` score 1 unit higher than average level of absorption, the estimated mean positive affect lowers by 0.0568 point.

$\hat{\gamma}_0 = -11.1719$, under an average level of baseline absorption and male indicator (`mpqab=21.9859`, `male=0.3179`), performances with student audience have an estimated mean positive affect 11.1719 points lower than that of public performance or juried recital audience.

$\hat{\beta}_1 = 0.3994$, for performances with instructor audience, as MPQ absorption level 1 unit higher than average, the estimated mean positive affect lower by 0.3994 point.

$\hat{\sigma}_u = 4.304$, for performances with public performance or juried recital audience, the estimated standard deviation between person is in deviation of 4.304 on positive affect

$\hat{\sigma}_v = 2.862$, the estimated between-person standard deviation on positive affect for instructor auidence is in deviation of 2.862 higher than public performance or juried recital audience

$\hat{\rho}_{uv} = -0.256$, there's negative correlation between a subject's positive affect and whether or not the audience is instructor, typically, the instructor audience would decrease the positive affect by 0.256.

$\hat{\alpha}_2 = 3.0497$, with audience being public performance or juried recital, the estimated mean positive affect for male is 3.0497 points higher than female.

$\hat{\beta}_2 = -2.2075$, with instructor audiences, the estimated mean positive affect for male is 2.2075 higher than female.

-- (**h**) Model Comparisons

LR-test:

```
        npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
model.c   13 3310.2 3364.9 -1642.1   3284.2
model.d   16 3310.4 3377.7 -1639.2   3278.4 5.7914  3     0.1222
```

$LR = -2(logLik_c - logLik_d) = -2(-1642.1 + 1639.2) = 5.8$

$\text{p-value} = 0.1222$

From likelihood test, we can see that the test statistic is around 5.8, which follows chi-square distribution with df=3. p-value is 0.1222, which was greater than 0.05, thus the additional factor of male indicator didn't improve the model so much, and model C can be relatively better.

$\text{Pseudo } R^2_{L2} = \frac{\hat{\sigma}^2(ModelC) - \hat{\sigma}^2(ModelD)}{\hat{\sigma}^2(ModelC)} = \frac{36.487 - 36.41}{36.487} = 0.0021$

	AIC	BIC
ModelC	3310.2	3364.9
ModelD	3310.4	3377.7

From Pseudo R^2 values, we can see that the additional male indicator contributed very little on the variety between subjects. For AICs, there were not obvious difference between model C and D. While for BICs, the BIC for Model C can be smaller than that of Model D, which means Model C was better.

Conclusions

The positive affect can be associated with a lot of factors, some are level one which may vary among different performance, others are level two which may vary in terms of different subjects. Based on the fitted models considering several level one or level two factors, it is likely that the model c made a relative better performance, where factors like instructor audience, student audience and MPQ aborption level were considered to have pretty strong association with positive affect.

References

[1] Roback, P., & Legler, J. (2021). *Beyond Multiple Linear Regression: Applied Generalized Linear Models And Multilevel Models in R*. CRC Press. https://bookdown.org/roback/bookdown-BeyondMLR/ch-multilevelintro.html.

[2] Sadler, M. E., & Miller, C. J. (2010). Performance anxiety: A longitudinal study of the roles of personality and experience in musicians. Social Psychological and Personality Science, 1(3), 280-287.

Appendix

```
df <- read.csv('musicdata.csv')
head(df)

library(MASS)
library(gridExtra)
library(mnormt)
library(lme4)
library(knitr)
library(kableExtra)
library(tidyverse)
library(ggplot2)
library(ggpubr)

####  --------------------- EDA ----------------------- ####
df1 <- df%>%group_by(id)%>%summarise(mean_pa=mean(pa))
h1 <- ggplot(df, aes(pa))+
  geom_histogram(bins = 15, col='red', fill='pink', alpha=0.5)+
  coord_cartesian(xlim = c(min(c(df$pa, df1$mean_pa)), max(c(df$pa,df1$mean_pa))))
```

```
h2 <- ggplot(df1, aes(mean_pa))+
  geom_histogram(bins = 15, col='red', fill='pink', alpha=0.5)+
  coord_cartesian(xlim = c(min(c(df$pa, df1$mean_pa)), max(c(df$pa,df1$mean_pa))))
ggpubr::ggarrange(h1, h2, ncol=1)

ggplot(df, aes(audience, pa))+
  geom_boxplot(col='red', fill='pink', alpha=0.5)+
  coord_flip()

ggplot(df, aes(pa, ia))+
  geom_point()+
  geom_smooth(method = 'lm', col='pink')
ggplot(df, aes(pa, sa))+
  geom_point()+
  geom_smooth(method = 'lm')

ggplot(df, aes(ia, pa))+
  geom_point()+
  geom_smooth(method = 'lm')+
  facet_wrap(~id)

ggplot(df, aes(sa, pa))+
  geom_point()+
  geom_smooth(method = 'lm')+
  facet_wrap(~id)

pairs(df[,c('pa', 'ia', 'sa', 'mpqab', 'male')])

## Level 1 - pa; audience, memory, perform_type; diary

## histogram for positive affect distributions
p.pa <- ggplot(df, aes(x=pa))+
  geom_histogram(bins = 15, col='red', fill='pink', alpha=0.5)+
  coord_cartesian(xlim = c(min(df$pa),max(df$pa)))+
  labs(title = 'positive affect among 497 performances',
       x = 'positive affect', y = 'frequency')
p.pa_mean <- df%>%group_by(id)%>%summarise(mean_pa=mean(pa))%>%
  ggplot(aes(x=mean_pa))+
  geom_histogram(bins=15,col='red', fill='pink', alpha=0.5)+
  coord_cartesian(xlim = c(min(df$pa),max(df$pa)))+
  labs(title = 'performance averaged positive affect among 37 subjects',
       x = 'positive affect', y = 'frequency')
ggarrange(p.pa, p.pa_mean, ncol = 1)

## categorical factors: audience, memory, perform_type
table(df$perform_type)
```

```
table(df$audience)
table(df$memory)
p.type <- ggplot(df, aes(x=perform_type))+
  geom_bar(col='red', fill='pink', alpha=0.5)+
  theme(text = element_text(size = 6))
p.ad <- ggplot(df, aes(x=audience))+
  geom_bar(col='red', fill='pink', alpha=0.5)+
  theme(text = element_text(size = 6))
p.memory <- ggplot(df, aes(x=memory))+
  geom_bar(col='red', fill='pink', alpha=0.5)+
  theme(text = element_text(size = 6))
ggarrange(p.type, p.ad, p.memory, ncol=3)

p.type2 <- ggplot(df, aes(x=perform_type, y=pa))+geom_boxplot(col='red',
fill='pink', alpha=0.5)+coord_flip()
p.ad2 <- ggplot(df, aes(x=audience, y=pa))+geom_boxplot(col='red', fill='pink',
alpha=0.5)+coord_flip()
p.memory2 <- ggplot(df, aes(x=memory, y=pa))+geom_boxplot(col='red', fill='pink',
alpha=0.5)+coord_flip()
ggarrange(p.type2, p.ad2, p.memory2, ncol=1)

ggplot(df, aes(x=perform_type, y=pa))+
  geom_boxplot()+coord_flip()+
  facet_wrap(~id)

ggplot(df, aes(x=audience, y=pa))+
  geom_boxplot()+coord_flip()+
  facet_wrap(~id)

ggplot(df, aes(x=memory, y=pa))+
  geom_boxplot()+coord_flip()+
  facet_wrap(~id)

## numerical factors: previous (# previous performance with a diary entry)
summary(df$previous)
ggplot(df, aes(x=previous, y=pa))+geom_point()+geom_smooth(method = 'lm')
ggplot(df, aes(x=previous, y=pa))+geom_point()+geom_smooth(method = 'lm')+
  facet_wrap(~id)


## Level 2 - demographics (age, gender); instrument, years_study; MPQ (ab, sr,
pem, nem, con)
df1 <- df%>%group_by(id)%>%summarise(age, gender, instrument, years_study,
                                     mpqab, mpqsr, mpqpem, mpqnem, mpqcon,
                                     mean_pa=mean(pa))
df1 <- unique(df1)
```

```r
## l2 categorical: genderm instrument
table(df1$gender)
table(df1$instrument)

p.g <- ggplot(df1, aes(x=gender))+
  geom_bar(col='red', fill='pink', alpha=0.5)
p.i <- ggplot(df1, aes(x=instrument))+
  geom_bar(col='red', fill='pink', alpha=0.5)
ggarrange(p.g, p.i, ncol = 2)


## instrument & years_study
table(df1$instrument)
ggplot(df1, aes(x=years_study))+
  geom_bar(col='red', fill='pink', alpha=0.5)

## MPQ
summary(df1$mpqab)
ggplot(df1, aes(x=mpqab))+
  geom_bar()
ggplot(df1, aes(x=mpqsr))+
  geom_bar()
ggplot(df1, aes(x=mpqpem))+
  geom_bar()
ggplot(df1, aes(x=mpqnem))+
  geom_histogram()
ggplot(df1, aes(x=mpqcon))+
  geom_histogram()

## factors vs pa
## L2 categorical factors: gender, instrument
p.g1 <- ggplot(df, aes(x=gender, y=pa))+geom_boxplot(col='red', fill='pink',
alpha=0.5)+coord_flip()
p.g2 <- ggplot(df1, aes(x=gender, y=mean_pa))+geom_boxplot(col='red', fill='pink',
alpha=0.5)+coord_flip()
ggarrange(p.g1, p.g2, ncol = 1)
p.i1 <- ggplot(df, aes(x=instrument, y=pa))+geom_boxplot(col='red', fill='pink',
alpha=0.5)+coord_flip()
p.i2 <- ggplot(df1, aes(x=instrument, y=mean_pa))+geom_boxplot(col='red',
fill='pink', alpha=0.5)+coord_flip()
ggarrange(p.i1, p.i2, ncol = 1)
## L2 numerical: age, years_study, mpq
p.a1 <- ggplot(df, aes(x=age, y=pa))+geom_point()+geom_smooth(method = 'lm')
p.a2 <- ggplot(df1, aes(x=age, y=mean_pa))+geom_point()+geom_smooth(method = 'lm')
ggarrange(p.a1, p.a2, ncol=1)
```

```
p.ys1 <- ggplot(df, aes(x=years_study, y=pa))+geom_point()+geom_smooth(method =
'lm')
p.ys2 <- ggplot(df1, aes(x=years_study,
y=mean_pa))+geom_point()+geom_smooth(method = 'lm')
ggarrange(p.ys1, p.ys2, ncol=1)

p.ab1 <- ggplot(df, aes(x=mpqab, y=pa))+geom_point()+geom_smooth(method = 'lm')
p.ab2 <- ggplot(df1, aes(x=mpqab, y=mean_pa))+geom_point()+geom_smooth(method =
'lm')
ggarrange(p.ab1, p.ab2, ncol=1)

p.pem1 <- ggplot(df, aes(x=mpqpem, y=pa))+geom_point()+geom_smooth(method = 'lm')
p.pem2 <- ggplot(df1, aes(x=mpqpem, y=mean_pa))+geom_point()+geom_smooth(method =
'lm')
ggarrange(p.pem1, p.pem2, ncol=1)

p.nem1 <- ggplot(df, aes(x=mpqnem, y=pa))+geom_point()+geom_smooth(method = 'lm')
p.nem2 <- ggplot(df1, aes(x=mpqnem, y=mean_pa))+geom_point()+geom_smooth(method =
'lm')
ggarrange(p.nem1, p.nem2, ncol=1)

ggarrange(p.a1, p.ys1, p.ab1, p.pem1, p.nem1,
          p.a2, p.ys2, p.ab2, p.pem2, p.nem2, ncol = 5, nrow = 2)

#### ------------------- Statistical Analysis ---------------------- ####

df$ia <- ifelse(df$audience=='Instructor', 1, 0)
df$sa <- ifelse(df$audience=='Student(s)', 1, 0)
df$male <- ifelse(df$gender=='Male', 1, 0)
model.a <- lmer(pa ~ 1+(1|id),REML=T, data = df)
summary(model.a)

model.b <- lmer(pa ~ ia+sa+(ia+sa|id),data=df)
summary(model.b)

model.c <- lmer(pa ~ ia+sa+mpqab+mpqab:ia+mpqab:sa+(ia+sa|id), REML=T, data=df )
summary(model.c)
mean(df$mpqab)

model.d <- lmer(pa ~ ia+sa+mpqab+male+mpqab:ia+male:ia+mpqab:sa+male:sa+
(ia+sa|id), REML = T, data=df )
summary(model.d)
mean(df$male)

## model comparison C vs D
```

```r
drop_in_dev <- anova(model.d, model.c, test = "Chisq")
drop_in_dev
```