

model.rmd

Yanran Li, Yijin Wang, Shubo Zhang

2024-03-20

```
library(mice)
```

```
##
```

```
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
## filter
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## cbind, rbind
```

```
library(tidyr)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(visdat)
```

```
library(naniar)
```

```
library(ggplot2)
```

```
data <- read.csv("data.csv") %>%
```

```
janitor::clean_names()
```

```
# pre-processing: expand design matrix and fill in NAs
```

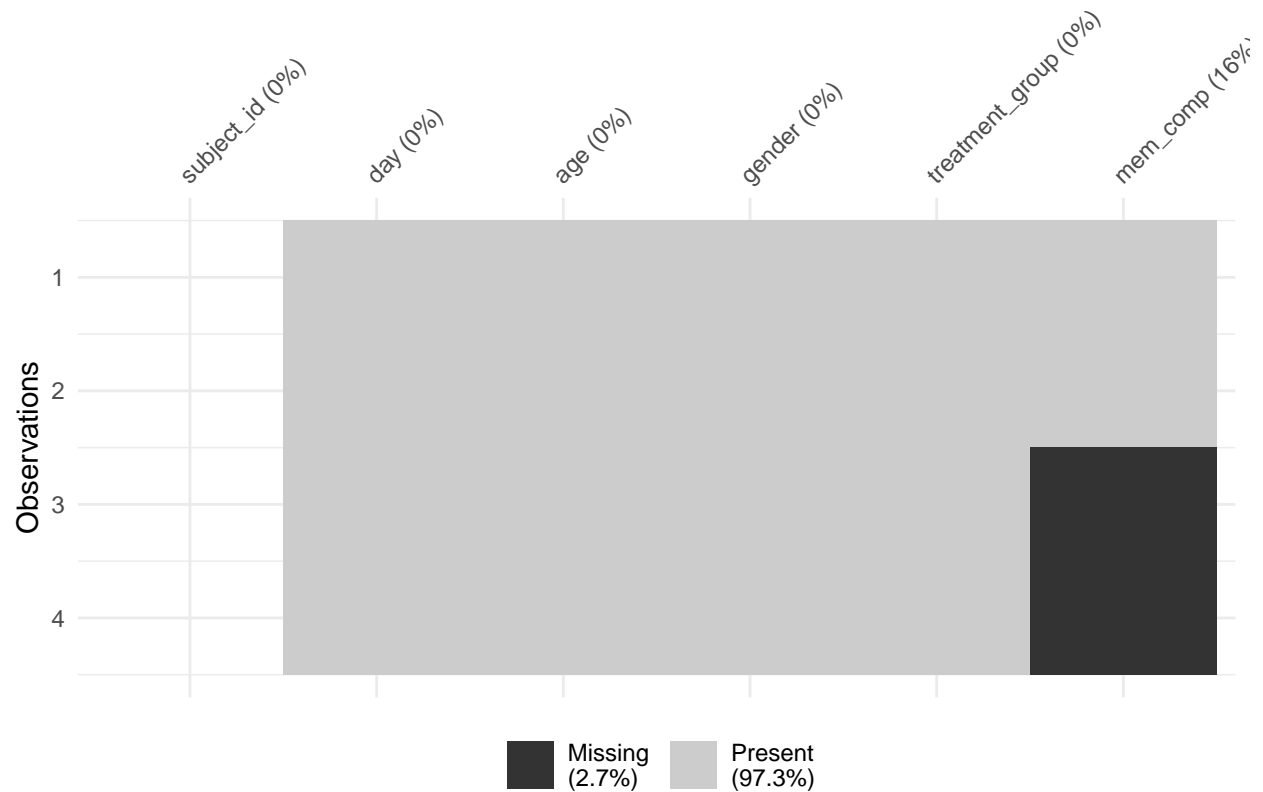
```
full_data <- data %>%
```

```
complete(day,
```

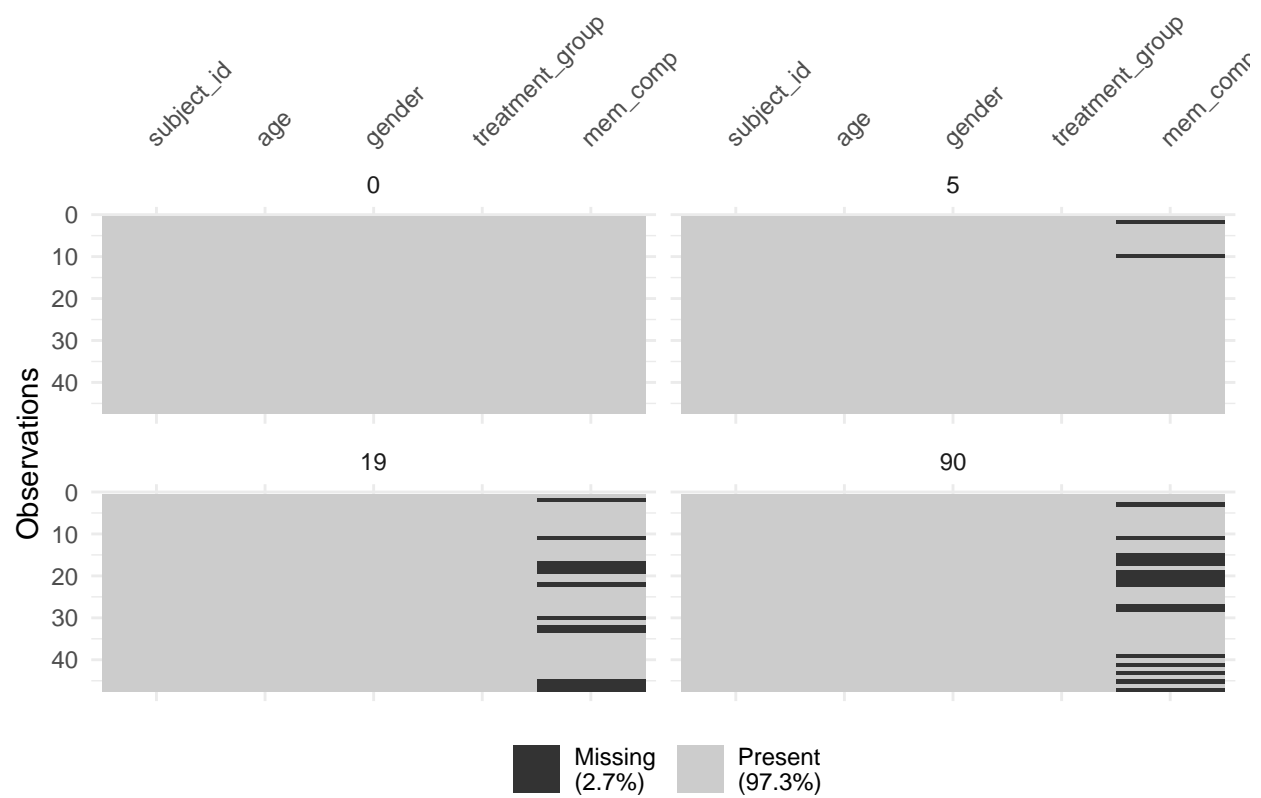
```
nesting(subject_id, treatment_group)) %>%
```

```
relocate(day, .after = subject_id) %>%
relocate(treatment_group, .after = gender) %>%
group_by(subject_id) %>%
fill(age, gender, .direction = "downup") %>%
arrange(subject_id, day)
```

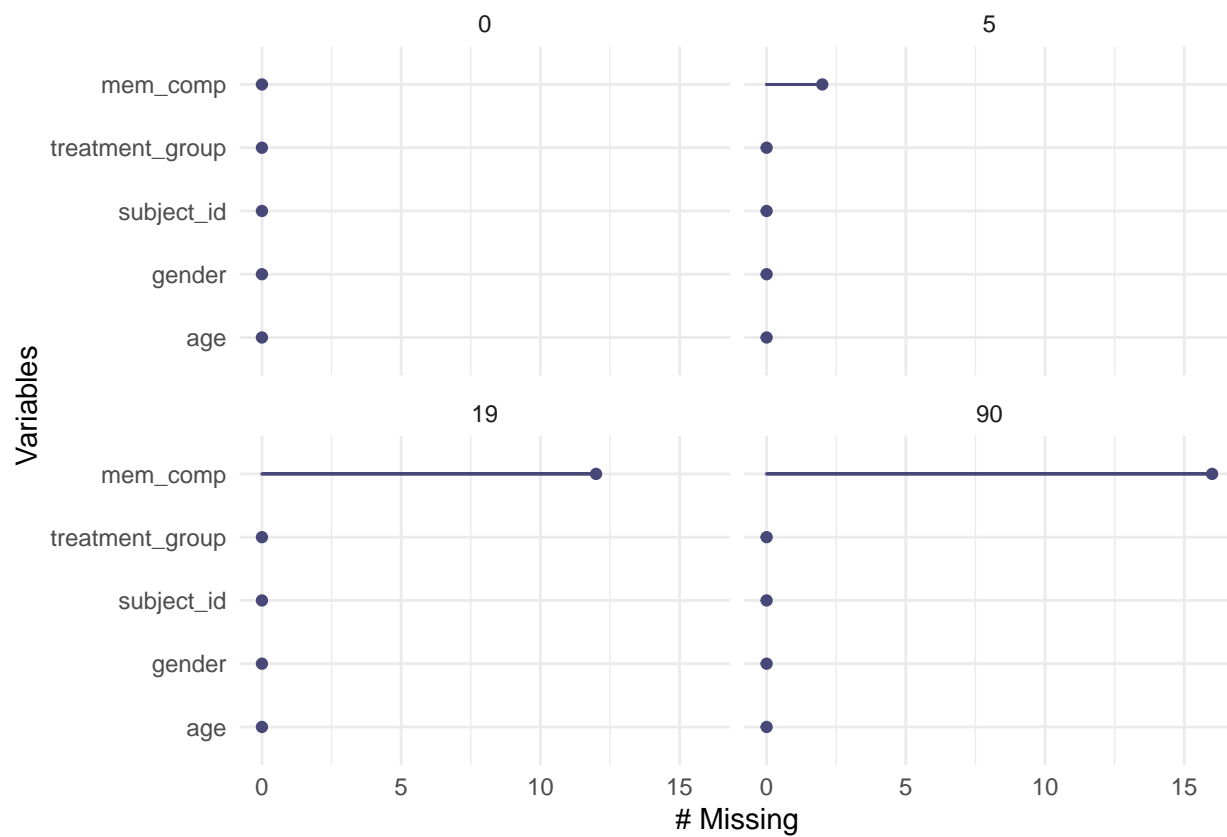
```
vis_miss(full_data)
```



```
vis_miss(full_data, facet = day)
```

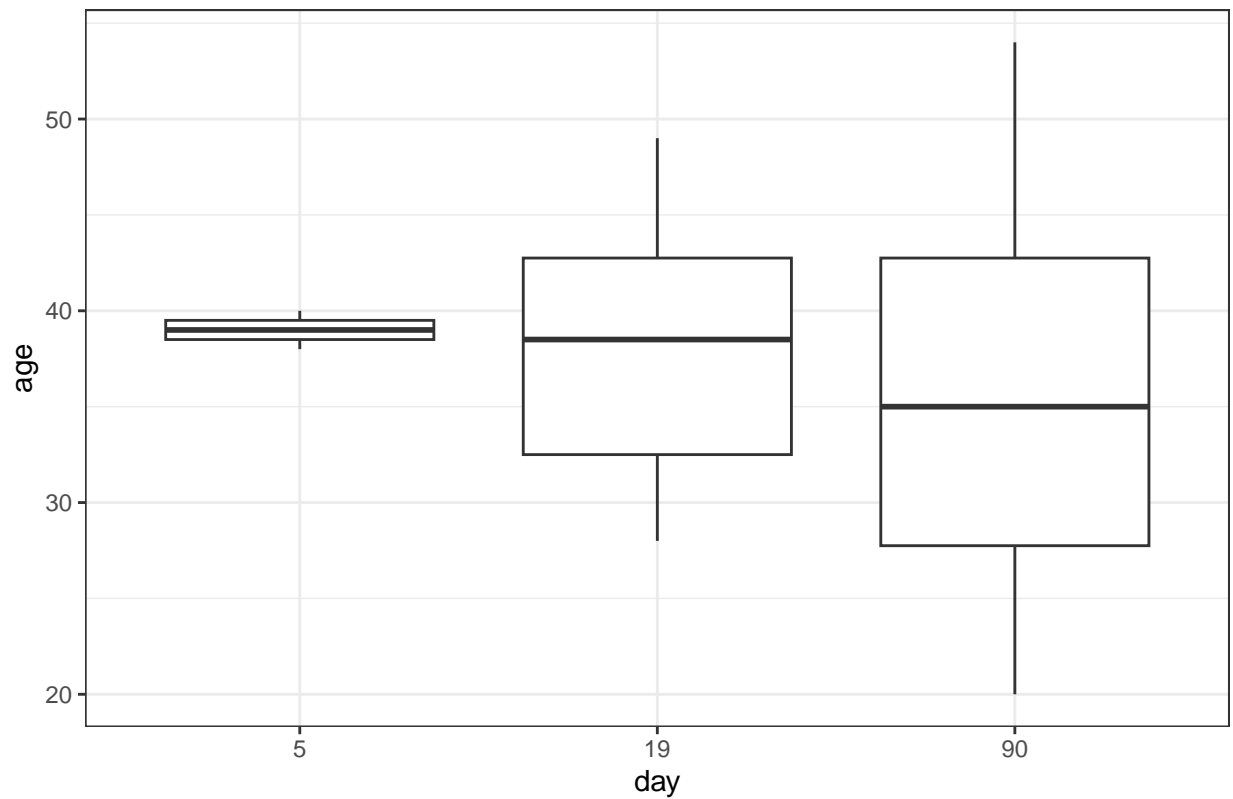


```
gg_miss_var(full_data,
  facet = day)
```



```
full_data %>% filter(is.na(mem_comp)) %>%
  ggplot(aes(x=as.factor(day)))+
  geom_boxplot(aes(y=age))+
  theme_bw()+
  xlab("day")+
  ggtitle("Relationship of age and missing values")
```

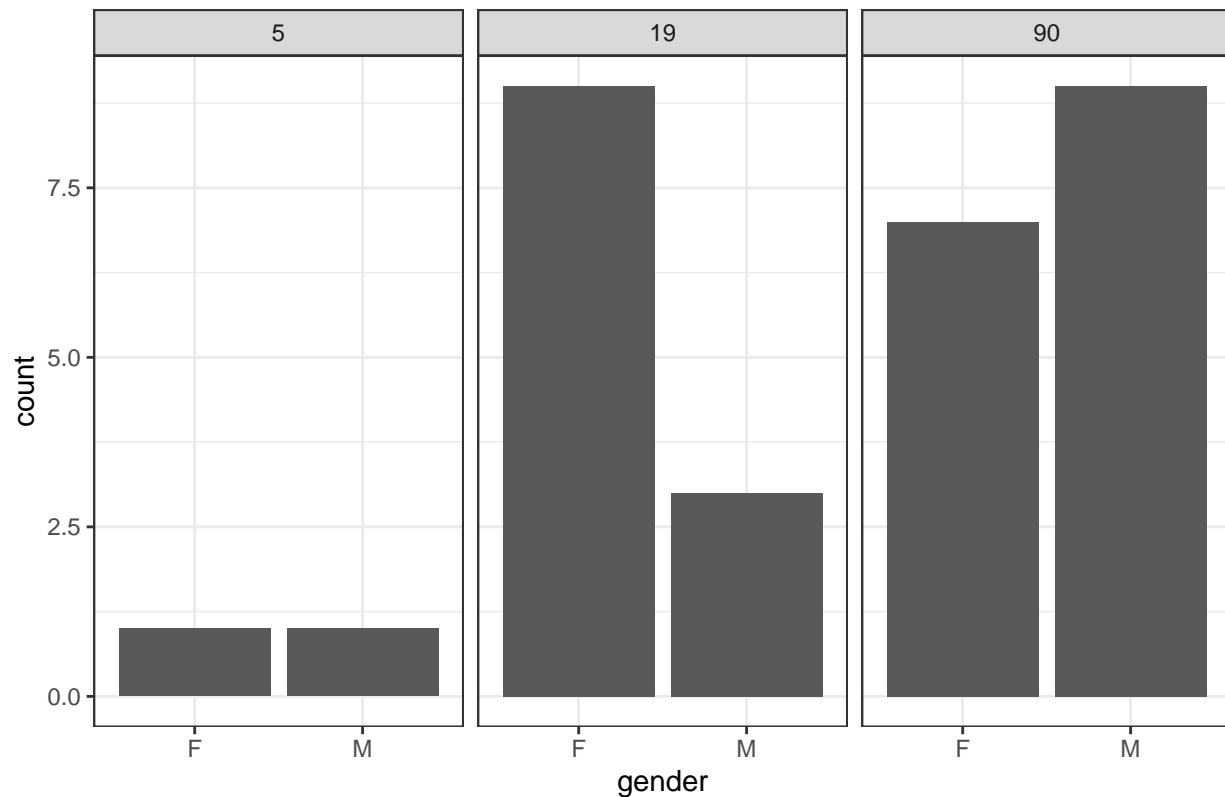
Relationship of age and missing values



```
full_data %>% filter(is.na(mem_comp)) %>%  
  ggplot()+  
  geom_bar(aes(x=gender, stat = "count"))+  
  facet_wrap(day~.)+  
  theme_bw()+  
  xlab("gender")+  
  ggtitle("Relationship of gender and missing values")
```

```
## Warning in geom_bar(aes(x = gender, stat = "count")): Ignoring unknown  
## aesthetics: stat
```

Relationship of gender and missing values



```
# assume MAR
# multiple imputations
imputed_data <- mice::mice(data = full_data, m = 6, seed = 2024, print = FALSE)
```

```
## Warning: Number of logged events: 2
```

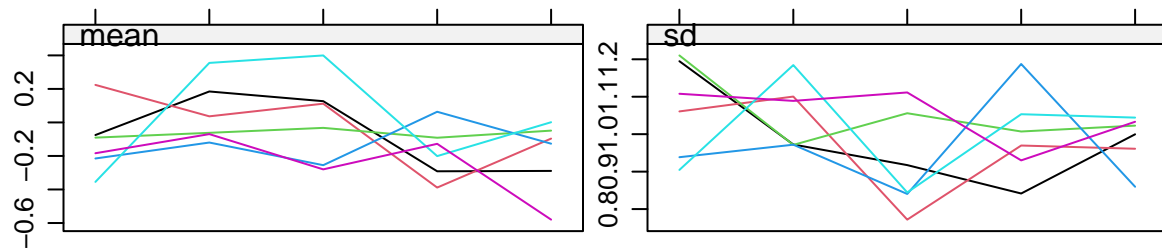
```
imputed_data
```

```
## Class: mids
## Number of multiple imputations: 6
## Imputation methods:
##   subject_id      day      age      gender treatment_group
##           ""      ""      ""      ""      ""
##   mem_comp
##   "pmm"
## PredictorMatrix:
##           subject_id day age gender treatment_group mem_comp
## subject_id      0  1  1  0      0      1
## day             1  0  1  0      0      1
## age             1  1  0  0      0      1
## gender          1  1  1  0      0      1
## treatment_group 1  1  1  0      0      1
## mem_comp        1  1  1  0      0      0
## Number of logged events: 2
##   it im dep      meth      out
```

```
## 1 0 0    constant    gender
## 2 0 0    constant treatment_group
```

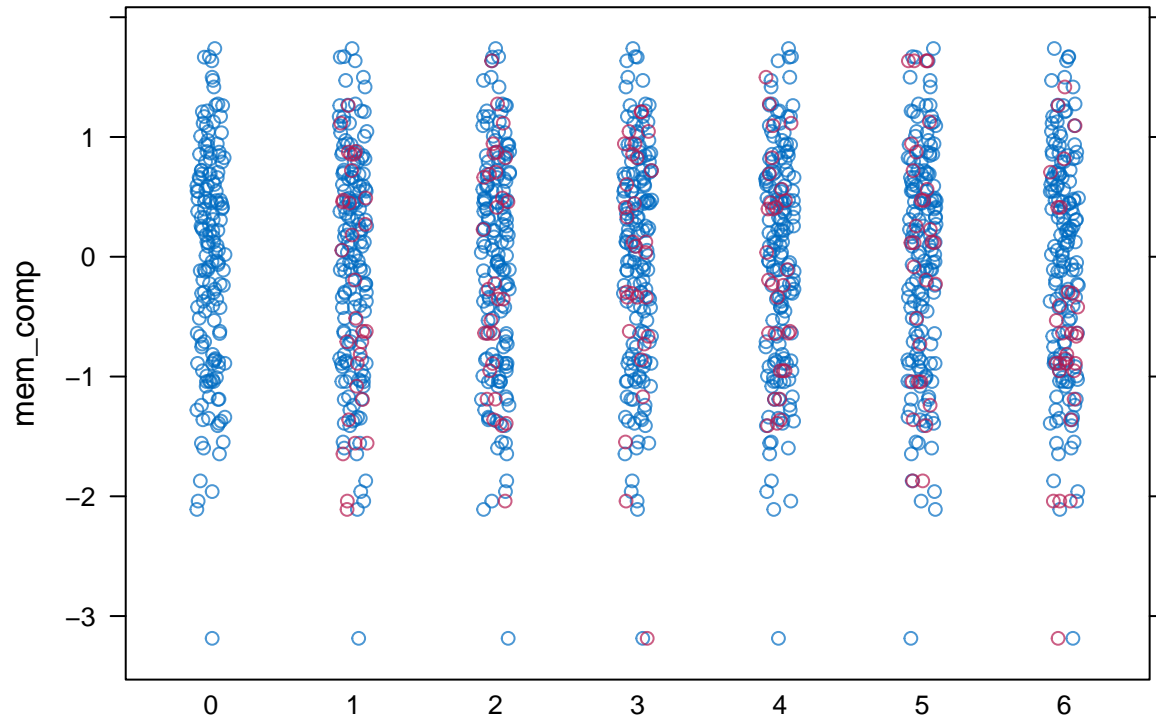
```
# used pmm method
```

```
# check convergence
plot(imputed_data)
```



Iteration

```
# check if imputed data makes sense (ie. in a reasonable bound)
# red is imputed, blue is observed
mice::stripplot(imputed_data, mem_comp)
```



```
# fit model to each imputed datasets
```

```
# pool coef estimates and standard errors to create 95% CIs
```

```
# sensitivity analysis for each data missing assumptions  
# mcar - fit data with completers data
```

```
# mmnr - fill na's with different delta --> fit model --> CI's
```

```
# Create a delta vector that represent the following adjustment values for mmHg: 0 for MAR, and -0.7, -  
delta <- c(0, -0.7, -0.2, 0.2, 0.7 )
```

```
imp.all <- vector("list", length(delta))  
post <- imputed_data$post  
for (i in 1:length(delta)){  
  d <- delta[i]  
  cmd <- paste("imp[[j]][,i] <- imp[[j]][,i] +", d)  
  post["mem_comp"] <- cmd  
  imp <- mice::mice(full_data, post = post, seed = i, print = FALSE)  
  imp.all[[i]] <- imp  
}
```

```
## Warning: Number of logged events: 2
```

```
## Warning: Number of logged events: 2
```

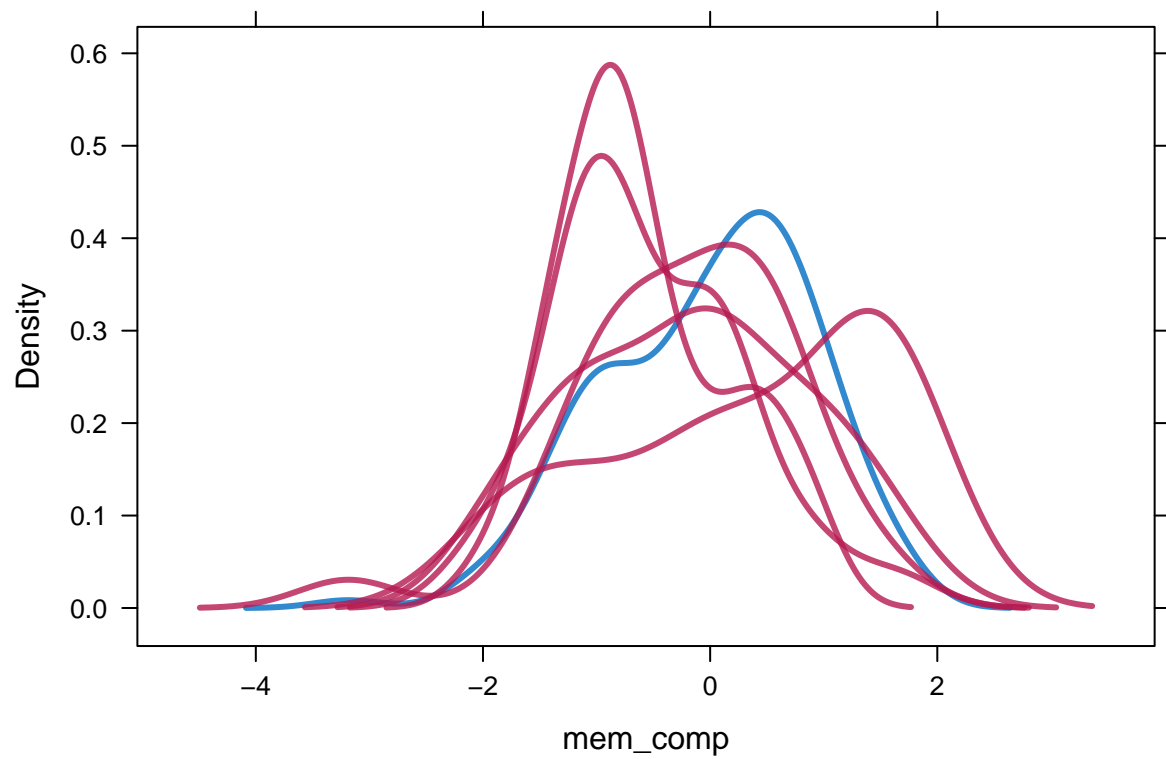


```
## Warning: Number of logged events: 2
```

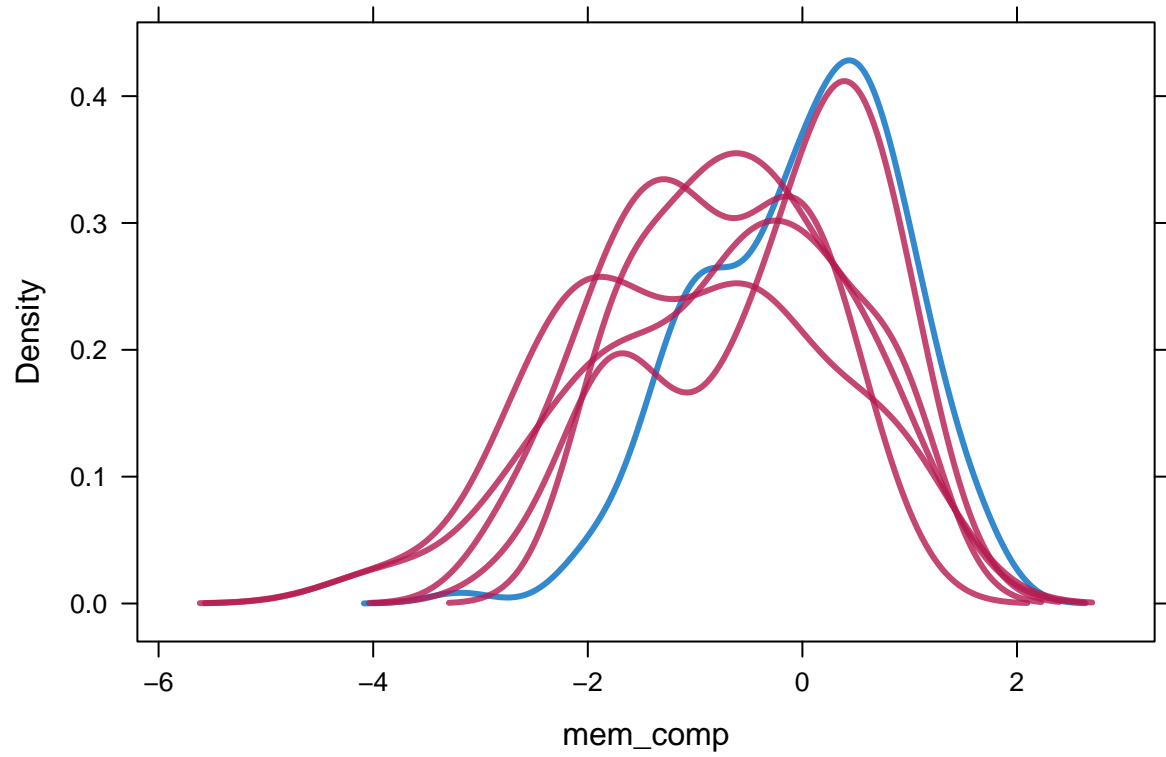
```
## Warning: Number of logged events: 2
```

```
## Warning: Number of logged events: 2
```

```
# imputation with no adjustment (delta = 0)  
densityplot(imp.all[[1]], lwd = 3)
```



```
# imputation with adjustment (delta = -0.7)  
densityplot(imp.all[[2]], lwd = 3)
```



```
# imputation with adjustment (delta = 0.7)  
densityplot(imp.all[[5]], lwd = 3)
```

