Class Activity: Introduction to Missing Data

STAT 340: Applied Regression

For this activity, consider this data set on chemotherapy for Stage B/C colon cancer, available as `colon_s` from the `finalfit` package in R. The help page for this data set gives the following description:
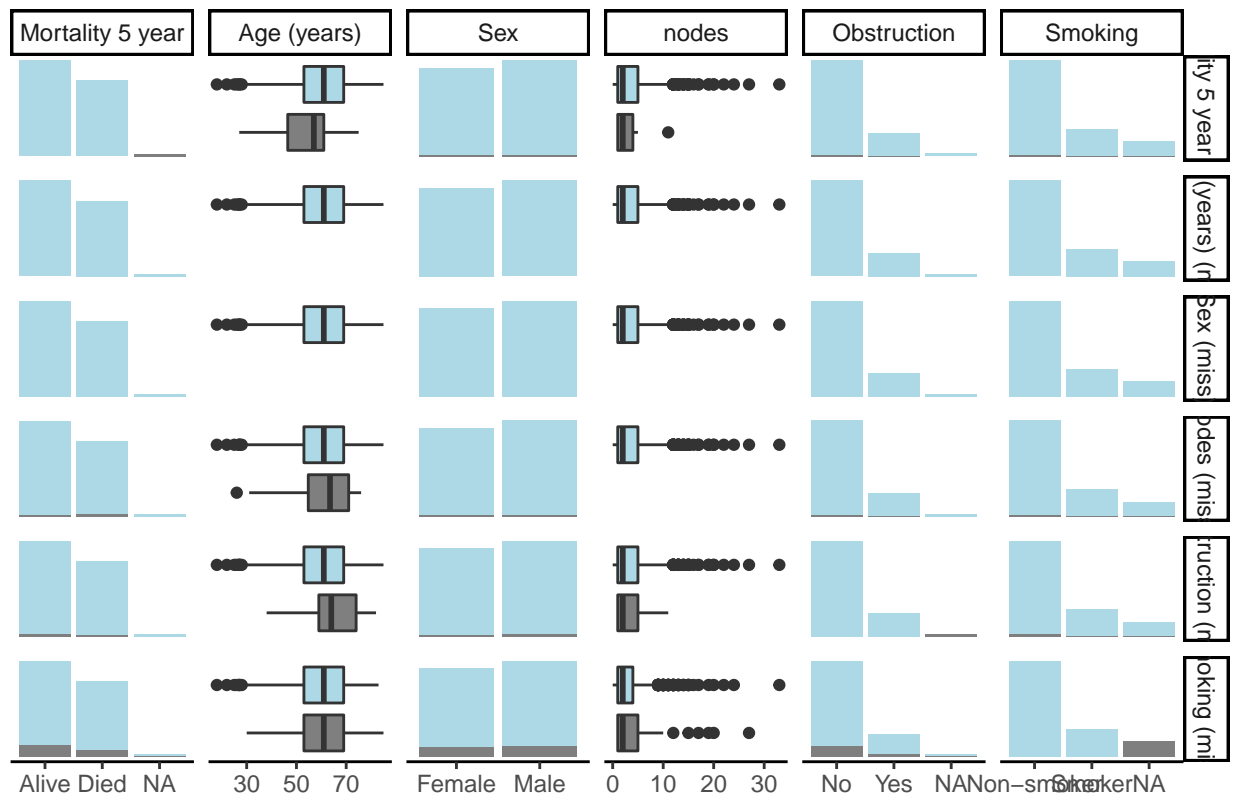
"These are data from one of the first successful trials of adjuvant chemotherapy for colon cancer. Levamisole is a low-toxicity compound previously used to treat worm infestations in animals; 5-FU is a moderately toxic (as these things go) chemotherapy agent. There are two records per person, one for recurrence and one for death."

Scenario 1

Consider the following pairs plots. Decide whether the data appear to be MCAR, MAR, or MNAR.
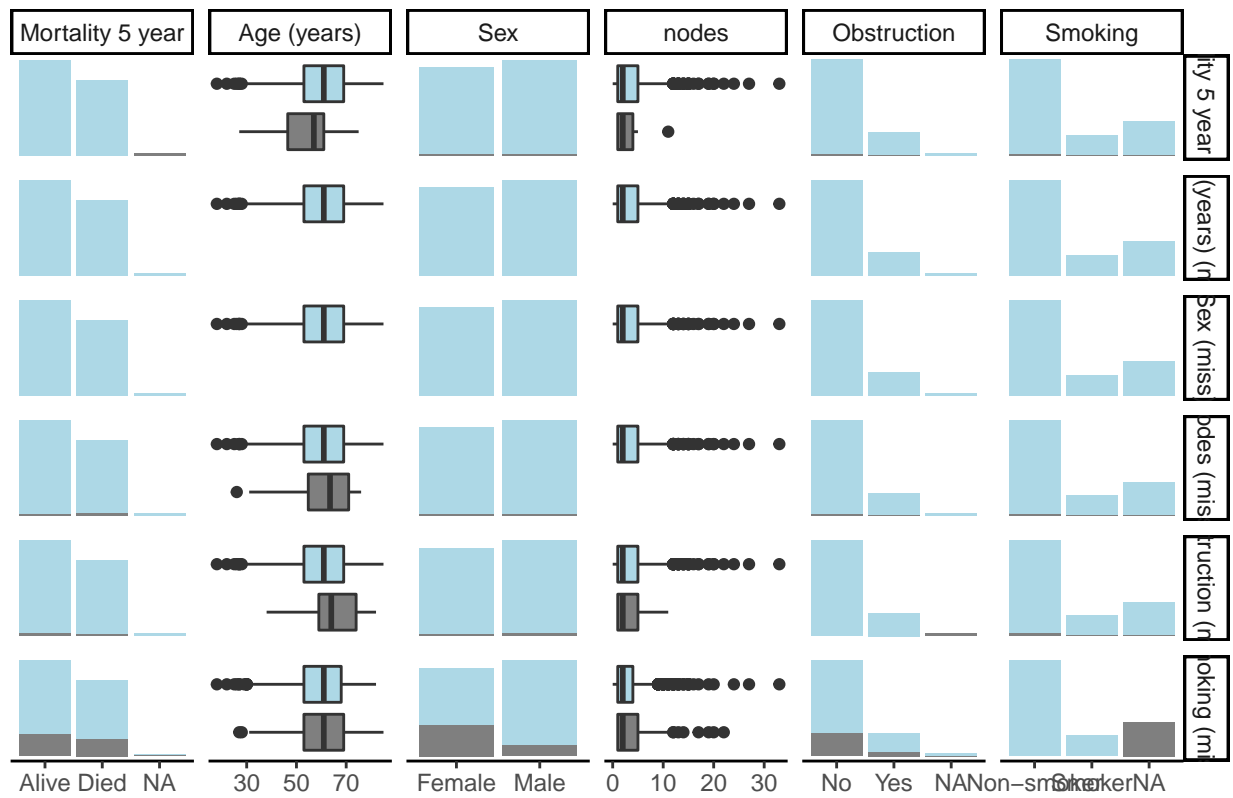
```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```


Missing data matrix

Scenario 2

Missing data matrix

Consider the following pairs plots. Decide whether the data appear to be MCAR, MAR, or MNAR. Support your conclusions with observations from the pairs plot.