

HW 2

Due 2/12 at 9:30 am

Use your own dataset to perform the following:

1. For each numerical variable which has missing values perform mean imputation.
2. For each categorical variable which has missing values perform mode imputation.

For mode imputation you can use the following R functions:

```
mode = function(x)
{
  ta = table(x)
  tam = max(ta)
  if (all(ta == tam))
    mod = NA
  else
    mod = names(ta)[ta == tam]
  return(mod)
}

mode.imp <- function (a)
{
  missing <- is.na(a)
  a.obs <- a[!missing]
  imputed <- a
  imputed[missing] <- mode(a.obs)
  # Output the imputed vector
  return (imputed)
}
```

3. Perform the analysis you planned to do on the imputed data and compare to the analysis done with the complete cases method (or complete data if you created the missingness yourself).