Random.useful.stuff

Rebekah Wrigley

26/01/2015

Dataframe organisation/naming

```
# Changing a single cell in a dataframe
index <- which(data$maj==99, arr.ind=TRUE) # finds the row number of the cell with a particular value (
data$maj[index]<- NA  # changes the the value in this row to NA


# Remove rows which have the same values in a particular column- good to check for duplicated IDs after
NPH_2<- NPH[duplicated(NPH$minor)!=TRUE,]


# Reordering columns in a dataframe- might want to do this if going to run a loop on adjacent columns.
Pol_NPH_nm_4<- Pol_NPH_nm_3[, c(1:49, 65, 50:64)]  # put column numbers in the order you want.


# Merging contents of mutually exclusive columns (where either one or the other column has data for a g
Cau_new$RS2231142_g<- ifelse(is.na(Cau_new$RS2231142_g.x)!=TRUE, Cau_new$RS2231142_g.x, Cau_new$RS223114

# Shortening long column names
name<- sapply(strsplit(x=colnames(Cau_imp_RG), split="_", fixed=TRUE), "[[", 1)  # Create vector of sho

colnames(Cau_imp_RG)<- name  # Make columnnames the same as those in the vector.

write.table(Cau_imp_RG, file="Cau_imp_short.txt",row.names=FALSE,quote=FALSE,sep="\t",na="NA")
```

Getting small p values from regression

```
## linear regression
2*pt(-abs(t), df)  # substitute t value and degrees of freedom from regression output.

## logistic regression
2*pnorm(-abs(z))  # substitute Z score from regression output.
```