Homework Assignment 4
Total Points: 100
(Due Thursday, October 12, 2023 at 5PM)

Please email your answer (compiled pdf file from R markdown) and R code to Cenxiao (CENXIAO@email.sc.edu)

1. Use the hdata.txt downloaded from the class website (command below), and

   (a) Find Mendelian inconsistency due to potential genotyping error in every genetic marker for every offspring using R assuming parents' genotypes are correct. [Hint: Use **intersect** and modify the R code in Lab8.R.] **(30 points)**

   Data Description:
   faid: Father's ID (0 indicates founder)
   moid: Mother's ID (0 indicates founder)
   iid: Individual ID
   sex: M: male, F: Female
   weight: weight
   m11: the first allele in m1 marker
   m12: the second allele in m1 marker

```
>hdata.url<-
"http://people.stat.sc.edu/hoyen/STAT588/Data/hdata.txt"
>hdata<-read.table(file=hdata.url, header=T, sep="\t")
```

   (b) Remove offsprings with Mendelian inconsistency, and create genotype data using R. For example, take

Original data:	m11	m12	m21	m22
	M1	M2	M3	M2
	M2	M1	M2	M3

   And create this:

Genotype data:	m1	m2
	M1_M2	M2_M3
	M1_M2	M2_M3

   Notice that both ("M1, "M2") and ("M2", "M1") should be the same "M1_M2" genotype.
   [Hint: use **sort** and **paste**] **(15 points)**

   (c) Perform gene association test using the genotype data generated from offspring. Identify any genetic markers that are associated

with weight adjusting for **faid** and **sex**. Interpret your findings: which allele in what marker increases weight by how much on average? **(15 points)**.

2. Use the FAMuSS data, determine the minor allele and its frequency for the actn3_1671064 SNP. Report these frequencies overall and stratified by the variable labeled Race. Interpret your findings. **(15 points)**

3. Based on the FAMuSS data, determine whether any of the four SNPs within the akt2 gene are associated with percentage change in non-dominant arm muscle strength as measured by NDRM.CH. **(15 points)**

4. Based on the FAMuSS data, determine whether there is a deviation from Hardy-Weinberg equilibrium (HWE) for the akt1_t10726c_t12868c SNP. Check whether stratifying by the variable Race alters your findings. **(10 points)**

Get FAMuSS data:

```
>fmsURL<-
"http://people.stat.sc.edu/hoyen/STAT588/Data/FMS_d
ata.txt"

>fms<-read.delim(file=fmsURL, header=TRUE,
sep="\t")
```