Phenotypic Data

Exploratory Data Analysis

1. **Distribution of the phenotypic data**
2. **Distribution of the residuals (error term)**

Genotypic Data

Quality Control

1. **Missing marker rate**
2. **Minor Allele Frequency (MAF)**
3. **Other filtering criteria**

Genotypic Data

Imputation

Preliminary step for a wide range of genetic analyses. Most models and software for GWAS do not handle missing data by default.

SNP Marker tested for association with trait
(fixed effect)

Genomic relationship matrix to correct for family structure
(random effect)

GWAS

$$y = X\beta + S\alpha + Qv + Zu + e$$

Fixed effects
(other than SNP under testing and population structure)

Population structure
Subpopulation effect
(fixed effect)