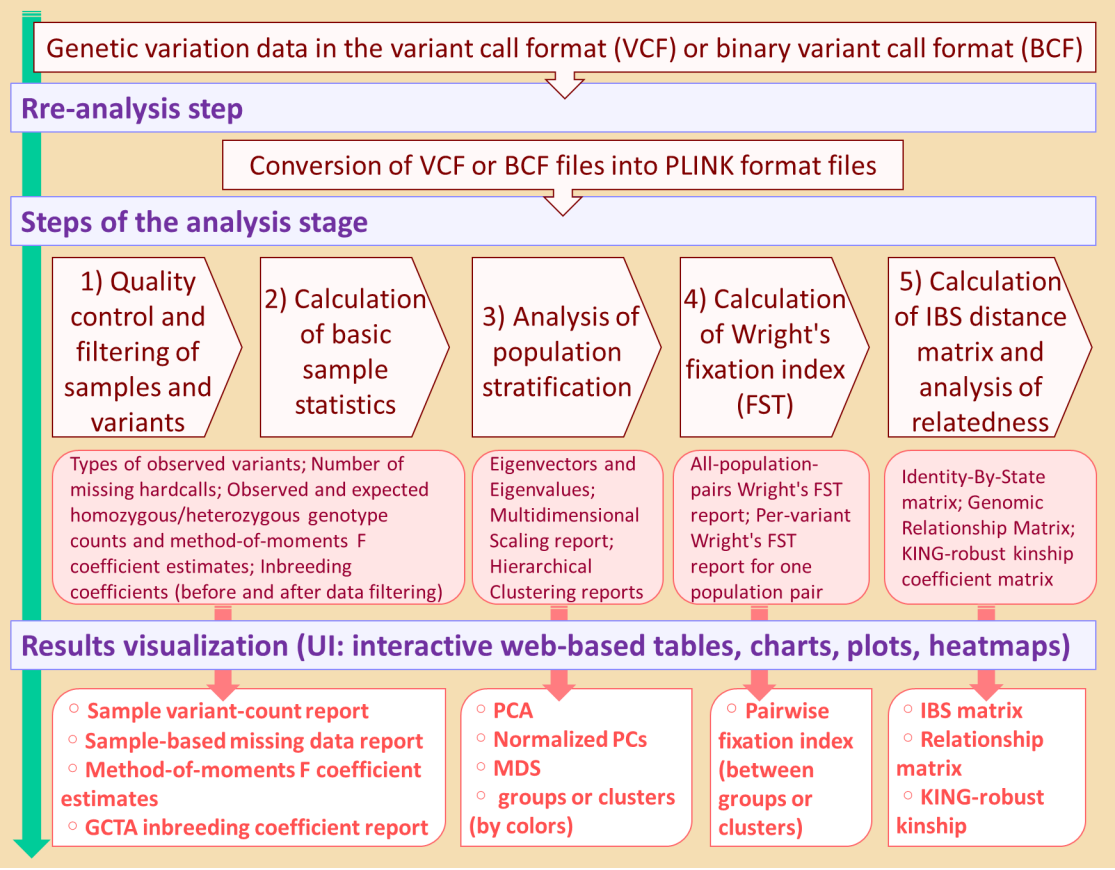
Genetic variation data in the variant call format (VCF) or binary variant call format (BCF)
Rre-analysis step
Conversion of VCF or BCF files into PLINK format files
Steps of the analysis stage
1) Quality control and filtering of samples and variants
2) Calculation of basic sample statistics
3) Analysis of population stratification
4) Calculation of Wright's fixation index (FST)
5) Calculation of IBS distance matrix and analysis of relatedness
Types of observed variants; Number of missing hardcalls; Observed and expected homozygous/heterozygous genotype counts and method-of-moments F coefficient estimates; Inbreeding coefficients (before and after data filtering)
Eigenvectors and Eigenvalues; Multidimensional Scaling report; Hierarchical Clustering reports
All-population-pairs Wright's FST report; Per-variant Wright's FST report for one population pair
Identity-By-State matrix; Genomic Relationship Matrix; KING-robust kinship coefficient matrix
Results visualization (UI: interactive web-based tables, charts, plots, heatmaps)
◦ Sample variant-count report
◦ Sample-based missing data report
◦ Method-of-moments F coefficient estimates
◦ GCTA inbreeding coefficient report
◦ PCA
◦ Normalized PCs
◦ MDS
◦ groups or clusters (by colors)
◦ Pairwise fixation index (between groups or clusters)
◦ IBS matrix
◦ Relationship matrix
◦ KING-robust kinship