```
#!/bin/bash

##### GWAS_pipeline_TWO_COHORTs

#### Author : Dina Gamaleldin Mansour Aly

#### Date : 4th October 2017

#### Program Usage: ./gwas_pipeline_2.sh Study_name Cohort_1_Sample_file
####Cohort_2_Sample_file Exclusion file covariates_names phenotype_name phenotype_data_type
#### method

SCRIPT_DIR=${0%/*}

##### Tools path:

SNPTEST= “USER_SUPPLIED”

RSCRIPT= “USER_SUPPLIED”

#### Genotyping data files:

COHORT_1_CHROMOSOME_RAW_DATA= “USER_SUPPLIED”

COHORT_2_CHROMOSOME_RAW_DATA= “USER_SUPPLIED”

#### Input files and information; Command_line arguments:

STUDY_NAME=$1

COHORT_1_SAMPLE=$2

COHORT_2_SAMPLE=$3

EXCLUSION_FILE=$4

COV_NAMES=$5

PHENO_NAME=$6

PHENO_VARIABLE_TYPE=$7

METHOD=$8

#### Association_Analysis using SNPTEST:

for chr in $(seq 1 22);do

  $SNPTEST  -data $COHORT_1_CHROMOSOME_RAW_DATA/chr"$chr".dose.vcf.gz
$COHORT_1_SAMPLE $COHORT_2_CHROMOSOME_RAW_DATA/chr"$chr".dose.mdc.vcf.gz
$COHORT_2_SAMPLE -genotype_field GP -exclude_samples $EXCLUSION_FILE -cov_names
$COV_NAMES -o GWAS_$STUDY_NAME_CHR"$chr".out -log
log_GWAS_$STUDY_NAME_CHR"$chr".txt  -frequentist 1 -method $METHOD -hwe -pheno
$PHENO_NAME;done

  cat   *.out |  grep -v "#" | head -1 > header.txt

  cat   *.out |  grep -v "#" | grep -v rsid > DATA.txt

  cat    header.txt DATA.txt > GWAS_unfiltered.txt
```

Filteration of SNPTEST output and PLOTS for GWAS_analysis:

```
$RSCRIPT $SCRIPT_DIR/"GWAS_plots.r"

    cat SORTED_GWAS_snps.csv | tr -d "\"" | tr "," "\t" | head -100 > TOP_100_GWAS_snps.txt

    cat SORTED_GWAS_snps.csv | tr -d "\"" | tr "," "\t" | cut -f 1,2,3,10 | sed "s/rsid/SNP/" | sed "s/alternate_ids/CHR/" | sed "s/position/BP/" | sed "s/frequentist_add_pvalue/P/" > GWAS_plot.txt
```