```
/* Lab3 ANOVA analysis */

/* part1: tensile data. !!! you need to change the path of the data file to your own */

PROC IMPORT DATAFILE="/home/u63048916/STAT571B/Labs/Lab Data/tensile.csv"
    DBMS=CSV
    OUT=one
    REPLACE;
    GETNAMES=YES;
RUN;

/* take a look at the data */
title 'Plot of Strength vs Percent Blend';
proc sgplot data=one;
scatter x=percent y=strength;
run;

 /* run ANOVA analysis */
title 'ANOVA analysis';
 proc glm data=one;
 class percent;
 model strength=percent;
 output out=diag p=pred r=res;
run;



/* normality checking */
/* take a look at the residuals: residual plot */
title 'Residual plot';
proc sgplot data=diag;
scatter y=res x=pred;
refline 0;
run;

title 'Normality checking';
proc univariate data=diag normal;
var res;
qqplot res/normal(mu=est sigma=est);
run;

/* outlier checking */
title 'Outlier';
data outlier;
set diag;
stdres=res/sqrt(8.06);
run;

proc print data=outlier;
run;

/* independence checking */
title 'Plot of residuals vs time';
proc sgplot data=diag;
scatter y=res x=time;
refline 0;
run;


 /* constant variance checking */;
title 'constant variance checking';
proc glm data=one;
class percent;
model strength=percent;
means percent / hovtest=bartlett hovtest=levene;
output out=diag1 p=pred r=res;
run;


 /* if the constant variance assumption does not hold, use Welch ANOVA, see the code below */;
PROC anova data=one;
CLASS percent;
MODEL strength=percent;
MEANS percent / WELCH;
RUN;



/* part2: if non-constant variance, do transformation: boxcox */

PROC IMPORT DATAFILE="/home/anling0/folder1/boxcox.xlsx"
    DBMS=XLSX
    OUT=two
    REPLACE;
    GETNAMES=YES;
RUN;

proc glm data=two;
class trt;
model resp=trt;
output out=diag2 p=pred r=res;
run;

title 'Residual Plot';
proc sgplot data=diag2;
scatter y=res x=pred;
refline 0;
run;
```

```
/* find an appropriate transformation parameter */
proc transreg data=diag2;
model boxcox(resp/convenient lambda=-2.0 to 2.0 by 0.1)=class(trt);
run;

/* do transformation on the response */
data new;
set two;
newY=log(resp);
run;

/* run a new ANOVA analysis on the transformed data */
proc glm data=new;
class trt;
model newY=trt;
output out=newout p=pred r=res;
run;

/* check the constant variance for the new result */
title 'New: Residual Plot';
proc sgplot data=newout;
scatter y=res x=pred;
refline 0;
run;

title 'New: variance checking';
proc glm data=newout;
class trt;
model newY=trt;
means trt / hovtest=bartlett hovtest=levene;
output out=newdiag p=newpred r=newres;
run;

/* part 3: what if non-normal data? */

data three;
input strain nitrogen @@;
datalines;
1 0.80 1 1.04 1 0.41 1 0.73 1 2.18
2 0.60 2 1.14 2 4.14 2 0.16 2 1.70
3 2.05 3 1.17 3 5.18 3 0.46 3 1.87
4 8.20 4 5.39 4 6.22 4 5.77 4 4.24
5 4.14 5 8.20 5 5.12 5 4.59 5 4.27
6 1.16 6 6.26 6 1.47 6 2.46 6 1.26
;
run;

proc sgplot data=three;
scatter y=nitrogen x=strain;
run;

proc glm data=three;
class strain;
model nitrogen=strain;
output out=diag3 p=pred r=res;
run;

title 'normality checking';
proc univariate data=diag3 normal;
var res;
qqplot res/normal(mu=est sigma=est);
run;

title 'Non-parametric analysis';
proc npar1way data=three;
class strain;
var nitrogen;
run;


proc univariate data=diag2 normal;
var res;
qqplot res/normal(mu=est sigma=est);
run;
proc npar1way data=two;
class trt;
var resp;
run;


/* part 4: post-ANOVA contrast */

proc glm data=one;
class percent;
model strength=percent;
contrast 'C1' percent 0 0 0 1 -1;
contrast 'C2' percent 1 0 1 -1 -1;
contrast 'C3' percent 1 0 -1 0 0;
contrast 'C4' percent 1 -4 1 1 1;
run;

/* part 5: simultinous comparisons */

proc glm data=one;
class percent;
model strength=percent;
/* Construct CI for Treatment Means*/
means percent /alpha=.05 clm lsd;
means percent /alpha=.05 clm bon;
means percent /alpha=.05 clm scheffe;
/* Pairwise Comparison*/
means percent /alpha=.05 lines lsd;
means percent /alpha=.05 lines bon;
```

```
means percent /alpha=.05 lines tukey;
means percent /alpha=.05 dunnett ('15');
run;

/* part 6: sample size calculation */

proc power;
  onewayanova
  groupmeans = 575 | 600 | 650 | 675
  stddev = 80
  alpha = 0.05
  npergroup = 15
  power = .;
run;
```