FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF MATHEMATICS AND STATISTICS

LABORATORY EXAMINATION PAPER 2016

MODULE CODE: MA4605   SEMESTER: Autumn 2016

MODULE TITLE: Chemometrics   DURATION OF EXAM: 45 minutes

LECTURER: Mr. Kevin O'Brien   GRADING SCHEME: 20 marks

                 20% of module grade

EXTERNAL EXAMINER: Prof. J. King

INSTRUCTIONS TO CANDIDATES

- This exam will start at 13:05 / 14:05 , and will last 45 minutes.

- Each question will be worth either 5 Marks. There are 20 Marks worth of questions. All questions must be attempted (LENS students please see below).

- **IMPORTANT** Write all of your answers in the exam script. Write the script number on any other documents you submit.

- It is your responsibility to return the script to collection box. An audit of scripts will take place immediately after the exam. If your script is account for in that audit, you are deemed to be absent, and will receive no marks.

- **IMPORTANT for LENS Student:** Specifically approved LENS students have to answer any 3 of the 4 questions .

Question 1: Exploratory Data Analysis

(a) Analyse the data set `X148` and answer the following questions:

    (i) (1 Mark) What is the mean of the values of `X148`?

    (ii) (1 Mark) What is the 95% confidence intervalf for the mean of the values of `X148`?

(b) The Anderson-Darling test and the Shapiro-Wilk test are two inference procedures that provide similar tests (i.e. one cab be considered an alternative to the other). Perform the Shapiro-Wilk test on the data set `X148`.

    (i) (1 Mark) What is the null hypothesis for this test?

    (ii) (1 Mark) Interpret the p-value of this test?

(c) Run the following piece of R code on dataset `X148`, and answer the questions below.

```
qqnorm(X148)
qqline(X148)
```

    (i) (1 Mark) Interpret the outcome of this graphical procedure for data set `X148`. *(a Brief answer will suffice)*.

Question 2 : Regression Models

(a) Consider the Response variable Y (found in data set X208). We wish to use the independent variable X (found in data set X047) to fit an explanatory model for X.

    (i) (1 Mark) Write down the regression equation for this fitted model.

    (ii) (1 Mark) Write down the correlation coefficient for these two variable X and Y.

(b) We wish to determine if a quadratic model is more appropriate in explaining Y. Hence we will use X^2 (found in data set X048)

    (i) (1 Mark) Write down the regression equation for this fitted model.

    (ii) (1 Mark) Write down the Akaike Information Criterion values for both linear models (i.e. this model, and the model from Part A).

    (iii) (1 Mark) Which of the two models better fits the data? Explain your answer.

(SPARE)

Question 3 : One Way ANOVA for Experimental Design

Four analysts each made seven determinations of the paracetamol content of the same batch of tablets. The results are shown below. There are 28 determinations in total.

Analyst	Content						
A	84.32	84.61	84.64	84.62	84.51	84.63	84.51
B	84.24	84.13	84.00	84.02	84.25	84.41	84.30
C	84.29	84.28	84.40	84.63	84.40	84.68	84.36
D	84.14	84.48	84.27	84.22	84.22	84.02	84.33

- The samples for Analyst A,B,C and D are contained in data sets X050,X051,X052 and X053.

- The aggregate data is found in data set X054.

- The grouping data is found in data set X049.

(i) (3 Marks) Write the ANOVA table in your answer sheet.

(ii) (1 Mark) What hypothesis is being considered by this procedure?

(iii) (1 Mark) What is the conclusion following from the above analysis? State the null and alternative hypothesis clearly.

(SPARE)

Question 4 : Factorial Design

An experiment is run on an operating chemical process in which the aim is to reduce the amount of impurity produced. Three continuous variables are thought to affect impurity, these are concentration of NaOH, agitation speed and temperature. As an initial investigation two settings are selected for each variable these are

Factor:	low level	highlevel
Concentration of NaOH	40%	45%
Agitation speed (rpm)	15	25
Temperature (°F)	170	200

Readings were recorded of the impurity produced from the chemical process for each combination of the levels of these factors, and each combination was tested twice.

Conc NaOH C	Agitation A	Temperature T	Impurity y
-	-	-	39,34
+	-	-	40,47
-	+	-	23,34
+	+	-	25,36
-	-	+	75,89
+	-	+	61,75
-	+	+	59,43
+	+	+	21,20

For the sake of brevity and clarity - type the following before carrying out your analysis

```
C = X061
A = X062
T = X063
y = X064
```

(i) (2 Marks) Complete the ANOVA table - For the sake of brevity, you may omit the SS and MS columns.

(ii) (1 Mark) State which of the main effects are significant

(ii) (1 Mark) State which of the interaction effects are significant

(iv) (1 Mark) Sketch the Interaction Plot between Agitation and Temperature. Comment on this plot.

(SPARE)