Exercise Sheet 3

1. In the context of experimental design, what is:
   * a factor?
   * a level?
   * a treatment?
2. A mechanical engineer is studying the thrust force developed by a drill press. He suspects that the drilling speed and the feed rate of the material affect the thrust force. He selects four feed rates and uses a high and low drill speed chosen to represent the extreme operating conditions.
   * How many factors does this experiment have?
   * How many levels does each of the factors have?
   * How many treatments does this experiment have?
3. Give an example of an experiment that uses a randomized block design.
4. Why was a randomized block design appropriate?
5. If we test five null hypotheses (which are all true) using 0.01 as the critical significance level what is the probability that we will make at least one type I error?
6. A statistician carries out an ANOVA on a data set. Name two assumptions that should be satisfied before carrying out the test.
7. The Kenton Food Company wished to test four different package designs for a new breakfast cereal. The different designs were sold in twenty shops, with approximately equal sales figures. Each shop was randomly assigned one of the package designs so that each design was sold in five different shops. The number of cases sold were represented by the variable ‘Cases’ and the type of packaging was represented by the variable ‘Package’ which had four different levels, identified as 1, 2, 3 and 4. After the preliminary data analysis, a one way ANOVA was carried out on the data set, followed by Tukey’s HSD post hoc test. The R output is shown below:

**Figure 1**

Fligner-Killeen test of homogeneity of variances

data: kenton$Cases by kenton$Package

Fligner-Killeen:med chi-squared = 1.1321, df = 3, p-value = 0.7693

**Figure 2**

Df Sum Sq Mean Sq F value Pr(>F)

Package 3 586.8 195.6 19.56 1.33e-05 \*\*\*

Residuals 16 160.0 10.0

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

**Figure 3**

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 14.600 1.414 10.32 1.76e-08 \*\*\*

kenton$Package2 -1.200 2.000 -0.60 0.5569

kenton$Package3 4.600 2.000 2.30 0.0352 \*

kenton$Package4 12.600 2.000 6.30 1.06e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.162 on 16 degrees of freedom

Multiple R-squared: 0.7858, Adjusted R-squared: 0.7456

F-statistic: 19.56 on 3 and 16 DF, p-value: 1.333e-05

**Figure 4**

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = kenton$Cases ~ Package)

$Package

diff lwr upr p adj

2-1 -1.2 -6.92203965 4.52204 0.9305700

3-1 4.6 -1.12203965 10.32204 0.1395792

4-1 12.6 6.87796035 18.32204 0.0000569

3-2 5.8 0.07796035 11.52204 0.0463912

4-2 13.8 8.07796035 19.52204 0.0000194

4-3 8.0 2.27796035 13.72204 0.0051230

**Figure 5**



**Figure 6**

# A tibble: 4 x 3

Package mean\_Cases sd\_Cases

<fct> <dbl> <dbl>

1 14.6 2.30

2 13.4 3.65

3 19.2 2.39

4 27.2 3.96

* + What does the result of the Fligner-Killeen test (see Figure 1) tell you?
  + What are the null and alternative hypotheses associated with this analysis?
  + Why does the variable ‘Package’ have 3 degrees of freedom?
  + Report the results of the ANOVA including the pairwise comparisons, a table indicating which (if any) treatments are significantly different at the 5% level and the effect size η².
  + Write down an ANOVA model for this data set, indicating clearly what each term in the model represents.

1. A soft drink distributer knows that end-aisle displays are an effective way to increase sales of the product. However there are several ways to design these displays. The marketing group has designed three new end aisle displays and wants to test their effectiveness. They have identified 15 stores of similar size and type to participate in each study. Each store will test one of the displays for a period of one month. The displays are assigned at random around the stores with each display tested in 5 stores. The response variable is the percentage increase in sales activity over typical sales in that store when the display is not is use. The data from the experiment is available on Blackboard in the soft\_drink.txt file. Can you analyse the data to determine whether the end-aisle displays have a significant impact on the soft drink sales? Write up a brief report including the following outputs and any conclusions that you draw.
   * Exploratory data analysis
   * ANOVA table
   * Pairwise comparisons including a table summarizing which (if any) treatments are significantly different.
   * The effect size η²
   * An ANOVA model indicating clearly what each term in the model represents.
   * Model diagnostics
2. An experiment was run to determine whether four specific firing temperatures affect the density of a certain type of brick. A completely randomized experiment was run and the results are available on Blackboard in the brick.txt file. Write up a brief report including the following outputs and any conclusions that you draw.
   * Exploratory data analysis
   * ANOVA table
   * Pairwise comparisons including a table summarizing which (if any) treatments are significantly different.
   * The effect size η²
   * An ANOVA model indicating clearly what each term in the model represents.
   * Model diagnostics