1. In a study of beetle damage on oats, researchers measured the number of beetle larvae per stem in small plots of oats after randomly applying one of two treatments: no pesticide or Malathion at the rate of 0.25 pound per acre. The data gave the summary statistics below.

Treatment	n	$\bar{x}$	s
No pesticide	13	3.47	1.21
Malathion	14	1.36	0.52

(a) Assuming that the two samples are from Normally distributed populations with a common variance σ^2, obtain a pooled estimate of σ^2. [2 marks]

(b) Test whether the population mean larvae counts differ between the two treatments. Interpret your results. [8 marks]

(c) Calculate a 90% confidence interval for the ratio of the true population variances. Is it reasonable to assume a common variance? [10 marks]

2. A University lecturer is interested in whether businesses of different sizes respond more or less readily to questionnaires sent out by business schools. She sent questionnaires to 200 randomly selected businesses of each of three sizes. The data on the responses are given below.

	Size		
	Small	**Medium**	**Large**
Response	125	81	40
No response	75	119	160
Total	200	200	200

(a) State in words an appropriate null hypothesis, H_0, for this problem. What is the alternative hypothesis, H_1? [2 marks]

(b) Test your hypothesis. [14 marks]

(c) Give a full report of your conclusions. [4 marks]

3(a)   Given a random sample of n observations from a Normal distribution with mean μ and variance σ^2, derive the mean and variance of the sample mean $\bar{x}$. [6 marks]

(b)   In a blood transfusion unit, the blood pressures of twenty donors were measured. The results in millimetres of mercury, mmHg, were as follows:

105 115 117 120 120 122 125 128 130 131 132 136 137 140 145 147 151 155 163 168

Display the data in a stem-and-leaf plot and describe the distribution. [8 marks]

The unit wishes to estimate how many people will have blood pressure above 160mmHg, as it is unwise for them to donate blood. By making suitable assumptions about the distribution of blood pressure, give an estimate of the proportion of donors with blood pressure above 160mmHg. [6 marks]

4.	The following data give the red cell folate levels ($\mu g/L$) in three groups of cardiac bypass patients given different levels of nitrous oxide ventilation.

Group 1: 243, 251, 275, 291, 347, 354, 380, 392

Group 2: 206, 210, 226, 249, 255, 273, 285, 295, 309

Group 3: 241, 258, 270, 293, 328

On the same graph, display the data for each level of ventilation separately and comment

[6 marks]

Complete the ANOVA table below and test the null hypothesis that the level of nitrous oxide ventilation has no effect on red cell folate level. State your conclusions. [12 marks]

	SS	d.f.	MS	F statistic
Between groups				
Within groups				
Total	55231.87			

Would it be appropriate to test for differences between pairs of groups? Justify your answer. [2 marks]

5.   To investigate the effect of birth weight in ounces (oz) upon the increase in weight between the 70th and 100th day of life expressed as a percentage of the birth weight, data were collected on 32 babies. Analysis of the data using Minitab gave the following output.

```
Predictor          Coef          StDev          T              P
Constant           167.87        19.88          8.44           0.0001
Birth weight       -0.8643       0.1757         -4.92          0.0001
     S = 17.80     R-Sq = 90.6%     R-Sq(adj) = 90.4%


Analysis of Variance

Source             DF            SS             MS             F              P
Regression         1             7666.39        7666.39        24.2           0.0001
Error              30            9502.08        316.74
Total              31            17168.47
```

Write down the estimated relationship between birth weight and weight increase.

[2 marks]

Does birth weight appear to have a significant effect upon weight increase, and does the variation in weight increase seem to be mostly due to birth weight differences?

[3 marks]

Find 95% confidence intervals for the intercept and slope of the fitted line. Comment on your results. [7 marks]

Using x to denote birth weight given that $\bar{x} = 111.75$ and $S_{xx} = \sum (x_i - \bar{x})^2 = 10262$ find a 95% prediction interval for the weight increase of a baby weighing 100oz at birth.

[8 marks]

[Hint: a 95% prediction interval for a new response y_0 when $x = x_0$ is given by

$$\hat{\alpha} + \hat{\beta} x_0 \pm t_{n-2}(0.025) s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}} \quad \text{where} \quad s^2 = \frac{SSE}{n-2} \quad \text{and} \quad S_{XX} = \sum (x_i - \bar{x})^2 \,]$$

6.   Extensive monitoring in a supermarket has suggested that waiting time at checkouts at the weekend may be modelled as a Normal random variable with standard deviation 30 seconds. A new training scheme for checkout operators has recently been introduced. The store manager is planning a study to estimate the true average waiting time μ for the new environment.

(a)   Why is the sample size important in such a study?   [2 marks]

(b)   Assuming that waiting times are still normally distributed with $\sigma = 30$, what sample size is necessary to ensure that the resulting 95% confidence interval for μ has a length of at most 20?   [5 marks]

(c)   The store manager wants to test the null hypothesis that the mean waiting time is 90 seconds. He considers a mean waiting time of 2 minutes to be the alternative of interest. Assuming that waiting times are still normally distributed with $\sigma = 30$, what sample size is necessary to ensure that the power of the hypothesis test is 90% at a significance level of $\alpha = 0.05$?   [11 marks]

(d)   For what alternative mean waiting times would the calculated sample size in (c) give a test of $H_0 : \mu_0 = 90$ of power at least 90% for $\alpha = 0.05$?   [2 marks]

7.   The table below gives the pre-test and post-test scores on a listening test in Spanish for 20 high school Spanish teachers who attended an intensive summer course in Spanish.

Teacher	1	2	3	4	5	6	7	8	9	10
Pre-test	30	28	31	26	20	30	34	15	28	20
Post-test	29	30	32	30	16	25	31	18	33	25

Teacher	11	12	13	14	15	16	17	18	19	20
Pre-test	30	29	31	29	34	20	26	25	31	29
Post-test	32	28	34	32	32	27	28	29	32	32

(a)   Perform a test of the hypothesis that there is no effect of the summer course on listening scores. State any assumptions you are making. Interpret your results.   [12 marks]

(b)   Obtain a 95% confidence interval for the mean difference in scores. Interpret your results.

[5 marks]

(c)   To obtain continued funding for the course, the organisers need to demonstrate an improvement in listening score of at least 5 points. Would you say funding should be continued ? Justify your answer.   [3 marks]