Problem 1. [8 marks]

a) **0.344**
[1]
b) **0.318**
[1]
c)
[1] **H0: the defect issue is independent on the parts vendor**
HA: the defect issue is dependent on the parts vendor
d)
[2] Expected Frequency = **58.48** Chi-square term = **0.34**
e)
[1] **4**
f)
[1] P-value > 0.10
g)
[1] There is insufficient evidence of an association between inspection results and vendor.

Problem 2. [15 marks]

a) i)
[1] Matched pairs t-test, since the two observations per car are paired (explanation necessary for mark)
ii)
[4] **H0: $\mu d = 0$ Ha: $\mu d \neq 0$**
t = $\bar{d}/(s_d/\sqrt{n})$ = 2.56
Rejection region is | t | > 2.262, based on 9 d.f.
Reject Ho, There is evidence at the 5% significance level that the appraisers differ in their assessments.
iii)
[1] The probability of deciding the appraisers differ in their assessments when in fact there is no difference.
iv)
[1] The boxplot of the differences is symmetric and there are no outliers. We can assume the differences are normally distributed. The t-test is appropriate.
0.5 for identifying the t-test, 0.5 for the reason.

(b)
i)
[1] **2-sample t-test** Matched pairs t-test
ii)
[1] **There is no difference between assessments (since P-value = 0.664 > 0.05)**
1 mark for conclusion, reason not necessary
iii)
[1] No, because the variation between cars is much greater than the variation between appraisers.
Iv)
[3] (838 – 766) ± 2.101×361×SQRT(1/10 + 1/10) = 72 ± 339=(-267, 411)
-1 for difference of 76 and proper use of sp
-1 for critical value of 2.101
-1 for calculation of standard error

v)

[1] The two boxplots for the two appraisers are positively skewed. Since we cannot assume the data are normally distributed, the Mann-Whitney would be a better test (0.5 for identifying test and 0.5 for reason)

Problem 3. [7 marks]

(a)

[5]

$$H_0: p_1 - p_2 = 0, \quad H_a: p_1 - p_2 \neq 0.$$

$$\hat{p} = \frac{97 + 122}{1443 + 1558} = .073$$

$$z_{stat} = \frac{\frac{97}{1443} - \frac{122}{1558}}{\sqrt{.073 \times .927\left(\frac{1}{1443} + \frac{1}{1558}\right)}} = -1.17$$

Leads to a p-value of 0.242 or a critical value of 1.645. Clearly not significant. There is insufficient evidence of a difference.

-1 for hypotheses
-1 for pooled proportion
-1 for z-statistic
-1 for p-value or critical value
-1 for decision and conclusion

(b)

[2]

$$n = \frac{2.575^2 \times .0672 \times .9328}{0.01^2} = 4158$$

-0.5 for critical value, 0.5 for p and 1-p, 0.5 for denominator,
0.5 for formula and final answer

Problem 4. [18 marks]

(a)
[1] Designed. Subjects are randomly assigned to treatments.
(b)
[1] 2*3=6
[1] There are no blocks in this design. The number of replications is 10.
(c)
[2]

- The boxplots suggest that within each treatment group, the data are symmetric and not particularly skewed. The data for each group should be normal, with equal variances, and the box plots are consistent with this.

- The normal plot of residuals hugs the line, which suggests normality of the errors.
- The histogram of residuals is symmetric – albeit with a couple of "large" negative values.
- Plot of residuals vs fitted values suggests homoscedasticity – the residuals seem equally spread out within each treatment group.
- Plot of residuals vs fitted values suggests that the residuals follow the empirical rule, suggesting that the model errors are normally distributed.
- The sequential plot of residuals reveals no outliers, nor any sequential correlations in the data. Observations should be independent. We see no evidence against that.

I would suggest that the two main assumptions to mention are the normality of the errors and constant error variance. (1 mark for each)

Source	DF	Sum of Squares	Mean Square	F
Protein	1	3168.3	3168.3	14.766
Food	2	266.5	133.3	0.6211
Interaction	2	1178.1	589.1	2.7455
Error	54	11586.0	214.6	NA
Total	59	16198.93	NA	NA

-total 3 marks: deduct 0.5 for each error, provided that error is not caused by another error

(d)
[4]

- Ho: no interaction; Ha: some interaction between protein level and food source
- The test statistic is 2.7455 (or whatever number was calculated above). The rejection region consists of all values greater than the critical value of an F with 2 and 54 degrees of freedom, F > 3.18 (2, 50 df is closest)
- Do not reject the null hypothesis, conclude there is insufficient evidence of interaction
- No interaction means that the difference in weight gain due to a high or low protein diet does not change with the source of that protein. (equivalently, the difference in weight gain between for different food sources is the same whether the diet is high or low protein).

(e)
[2]

- There is no apparent correspondence. With the F-test showing insufficient evidence of interaction, we might expect all three lines of the protein chart to be parallel (or both lines of the food chart).
- There may be interaction but the experimental error is too high to detect the non-zero interaction. A larger sample could result in a significant result. Alternatively, the apparent non-parallelism could result from sampling error.

(f)
[4]

- We need a t-critical value with 54 degrees of freedom (the DF's of the MSE in the ANOVA table). There are 15 possible pairwise comparisons amongst the 6 treatments (6*5/2). The new "alpha" is .05/15. For a central confidence interval, take the t-value corresponding to an upper tail probability of 0.05/30 = 0.00167. (1 mark for explanation)
- CI is: $$6.7 \pm 3.07\sqrt{214.56\left(\frac{1}{10}+\frac{1}{10}\right)} = [-13.4, 26.8]$$ or [-26.8, 13.4]

3 marks for CI: 1 for ± 6.7, 1 for standard error calculation, 1 for final interval.

Problem 5. [23 marks]

(a) [2]

Cannot assume the errors are normally distributed since the residuals are positively skewed;

Cannot assume the errors have constant variance since the residuals have different spreads.

(b) [2]

For startups receiving VC from a Canadian fund:
AmountFinanced = -11.34+ 1.3055*Age

For startups receiving VC from a US fund:
AmountFinanced = (-11.34 + 25.847) + 1.3055*Age

(c) [1]

Assume the amount financed increases linearly with age at the same rate for startups receiving VC from a US fund and for those receiving VC from a Canadian fund, but that startups receive more from US funds than from Canadian funds, if they are the same age. Or that the difference between VC received from the two sources is the same for all ages.

(d) [2]

For startups receiving VC from a Canadian fund:
AmountFinanced = -7.40 + 0.596*Age

For startups receiving VC from a US fund:
AmountFinanced = -7.40 + (0.596 + 3.34)*Age

(e) [1]

Assume the amount financed changes linearly with age but at different rates for the two groups. Or that the difference in the amounts financed between the two groups changes with age.

(f) [2]

Since both models have the same number of predictor variables, we can compare the R-squares (or, of course, the adjusted R-squares); based on this comparison, Model 2 is a better fit. Moreover, the standard error is lower with Model 2. (Look for at least two of these three measures; could allow also the F-statistic.)

(g) [2]

The problem of non-constant variance is resolved, but we cannot assume the errors are normally distributed since there are too many extremely negative residuals (well beyond three standard errors). -1 mark for mentioning each assumption with comment.

(h) [2]

Model 3 is slightly better than Model 4 (higher R-square or adj R-sq, and lower standard error) but only marginally. Could also mention the F-statistic is higher for Model 3.(Hard to pick one over the other, but Model 4 clearly not better) --1 mark for mentioning each summary statistic for a maximum of 2.

(i) [3]

Ho: beta1=0, beta2=0, Ha: one beta nonzero
F = 17.62 is in the rejection region F > 3.03 (based on 2, 300 df instead of 2, 288)
Reject Ho, conclude the model is useful.

(j) [3]

Ho: beta(US)=0; Ha: beta(US) nonzero
T = 5.57 is in the rejection region |t| > 1.96
Reject Ho, conclude there is a difference between the amounts financed by US versus Canadian funds, assuming the companies are the same age.

(k) [3]

95% PI for lnAmount is 1.616 +/- 1.96 sqrt(0.256^2 + 6.43) = 1.616 +/- 1.96 * 2.55

= 1.616 +/- 5.0 = (-3.384 , 6.616)

The interval for the actual amount financed is (0.0339, 747) millions of dollars

-1 for standard error calculation, 1 for CI for lnAmount, 1 for CI in dollars