# MAT 300 Midterm Exam

Student Name:

Date:

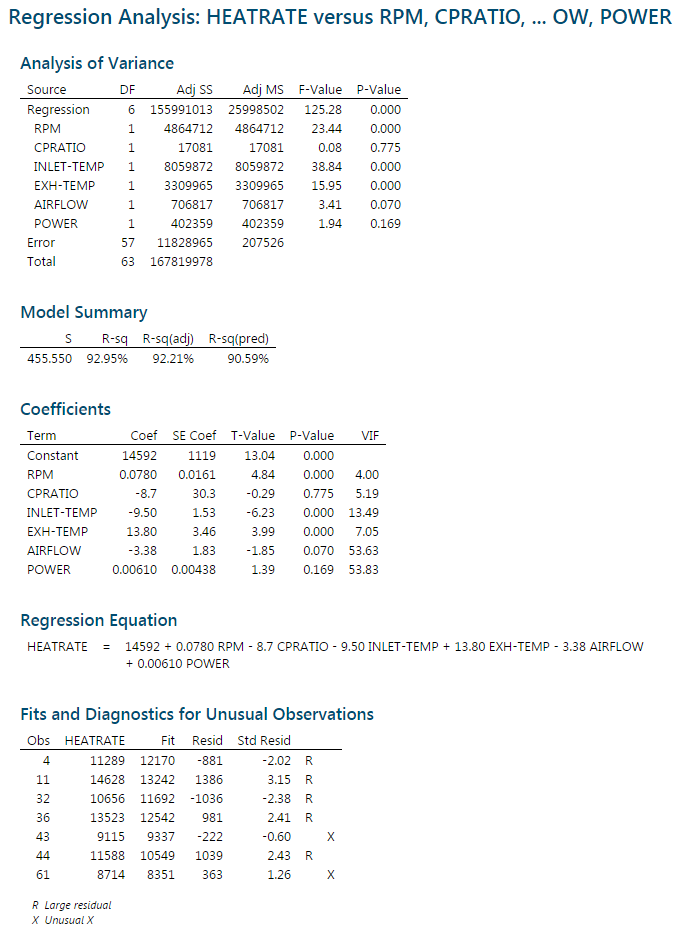
Reminder from the Syllabus: Your answers on this exam must be prepared individually in your own words. All assignments will be reviewed electronically for indications of cheating, plagiarism, misrepresentation, or unauthorized collaboration. Violations of this policy may result in serious consequences.

All calculations and relevant Minitab output must be included to receive full credit. Be sure to word-process your solutions and copy and paste the appropriate outputs from Minitab. Show all steps used in arriving at the final answers. Incomplete solutions will receive partial credit. This exam covers content from Modules One through Three.

Unless otherwise instructed, all inference tests should be performed at the α = 0.05 level.

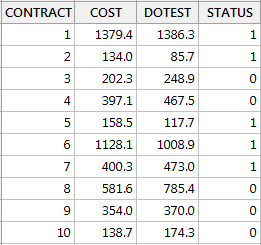
Problem 1. Consider the GASTURBINE data set and corresponding output from Minitab. Use the complete data set in your analysis. The first 10 observations are given for illustrative purposes. Complete parts a) through g) below.

| ENGINE | SHAFTS | RPM | CPRATIO | INLET-TEMP | EXH-TEMP | AIRFLOW | POWER | HEATRATE |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Traditional | 1 | 27245 | 9.2 | 1134 | 602 | 7 | 1630 | 14622 |
| Traditional | 1 | 14000 | 12.2 | 950 | 446 | 15 | 2726 | 13196 |
| Traditional | 1 | 17384 | 14.8 | 1149 | 537 | 20 | 5247 | 11948 |
| Traditional | 1 | 11085 | 11.8 | 1024 | 478 | 27 | 6726 | 11289 |
| Traditional | 1 | 14045 | 13.2 | 1149 | 553 | 29 | 7726 | 11964 |
| Traditional | 1 | 6211 | 15.7 | 1172 | 517 | 176 | 52600 | 10526 |
| Traditional | 1 | 6210 | 17.4 | 1177 | 510 | 193 | 57500 | 10387 |
| Traditional | 1 | 3600 | 13.5 | 1146 | 503 | 315 | 89600 | 10592 |
| Traditional | 1 | 3000 | 15.1 | 1146 | 524 | 375 | 113700 | 10460 |
| Traditional | 1 | 3000 | 15 | 1171 | 525 | 514 | 164300 | 10086 |



1. Write a first-order model **in general form** for the model that includes RPM, CPRATIO, INLET-TEMP, EXH-TEMP, AIRFLOW, and POWER to predict HEATRATE.
2. Use Minitab to compose regression equation for HEATRATE based on RPM, CPRATIO, INLET-TEMP, EXH-TEMP, AIRFLOW, and POWER. Paste the output here:
3. Write out the least squares prediction equation for the model that was fit above.
4. Calculate and give an interpretation of the coefficients based on a one-unit change in each xi. Calculate and give an interpretation of the effect on HEATRATE based on a 1-unit change in AIRFLOW together with a 210-unit change in POWER.
5. Interpret the overall model F-test. State the appropriate hypothesis test and associated numerator and denominator degrees of freedom used for this test as well as the critical value that the test statistic is compared to. State the conclusion you would make regarding the null hypothesis. Specifically, would you reject or fail to reject the null hypothesis, and what does this conclusion means about the model parameters? Does this tell us anything about the significance of the individual predictors? Why or why not?
6. Report and interpret the model R2.
7. Which predictors are significant in the model? Report the appropriate hypothesis test and formal conclusion you would make regarding RPM and CPRATIO. In your conclusion, state their p-values and test statistics. Would you suggest removing all non-significant predictors at once and refitting the model? Why or why not?

Problem 2. Consider the FLAG data set. The first 10 observations are given for informational purposes.



1. Use Minitab to fit a model that predicts COST based on DOTEST. Paste your output here:
2. Calculate a confidence and prediction interval for DOTEST = 110.
3. Interpret the confidence and prediction intervals given in the output. Do you see any problems with the interpretation of the prediction interval in terms of what we are trying to predict?
4. Why are confidence intervals always more narrow than prediction intervals?

Problem 3. Consider the EXPRESS data set. The first 10 observations are given for illustrative purposes.

| Weight | Distance | Cost |
| --- | --- | --- |
| 5.9 | 47 | 2.6 |
| 3.2 | 145 | 3.9 |
| 4.4 | 202 | 8 |
| 6.6 | 160 | 9.2 |
| 0.75 | 280 | 4.4 |
| 0.7 | 80 | 1.5 |
| 6.5 | 240 | 14.5 |
| 4.5 | 53 | 1.9 |
| 0.6 | 100 | 1 |
| 7.5 | 190 | 14 |

1. Draw a scatterplot of COST vs. each of the predictors. Do you see any evidence of a quadratic relationship?
2. Write a **general** second-order model (not including interaction terms) for Cost(y).
3. Give the null and alternative hypothesis for determining whether both of the second-order terms are statistically significant (nested model hypothesis).
4. Identify which of the two general nested models is the complete model and which is the reduced model.
5. Using Minitab, produce an output and write the least squares regression equation for the second-order model AND the reduced model that was fit in Minitab.
6. Compute the test statistic and perform the appropriate F-test. Be sure to state the degrees of freedom and the correct F critical value that you are comparing your test statistic to. Formally state your conclusion. Hint: To compute the test statistic, you need to separately fit both the complete and reduced models. To fit the complete model, you need to add the appropriate variables to your data set.

Problem 4. Once again you will be using the EXPRESS data set. The first 10 observations are given for illustrative purposes.

| Weight | Distance | Cost |
| --- | --- | --- |
| 5.9 | 47 | 2.6 |
| 3.2 | 145 | 3.9 |
| 4.4 | 202 | 8 |
| 6.6 | 160 | 9.2 |
| 0.75 | 280 | 4.4 |
| 0.7 | 80 | 1.5 |
| 6.5 | 240 | 14.5 |
| 4.5 | 53 | 1.9 |
| 0.6 | 100 | 1 |
| 7.5 | 190 | 14 |

1. Write out a complete **general** first-order model including an interaction term for Cost as the outcome.
2. Using Minitab, fit a model including the interaction term. Paste your output here.
3. Write the least squares regression equation with the interaction term.
4. State the null and alternative hypothesis to test if there is a significant interaction effect between weight and distance. Test the hypothesis at the α= 0.01 level. Write your formal conclusion in terms of the variable names.
5. If there is a significant interaction effect, but the individual predictors (main effects) that make up that interaction are not significant in the model, would you suggest removing the main effects? Why or why not?