Question 1. For each of the claims that follow, choose which of the following tests would be most appropriate.

(a) One sample t test
(b) Two sample t test
(c) Matched pair test
(d) One sample proportion test
(e) Two sample proportion test
(f) Regression slope test
(g) Chi-square goodness-of-fit test
(h) Analysis of variance

1. More than half of all babies born are boys.
2. The probability a student will pass Math 3 (pre-calculus) is the same as the probability a student will pass AMS 3 (also pre-calculus).
3. A person’s blood pressure is higher when taking an exam compared to reading a novel.

4. Smokers have a shorter life expectancy than non-smokers.

5. Increased spending on schools corresponds to better student performance.

6. Gas prices in Santa Cruz are the same as in San Jose.

7. The average GPA is the same for students in each of the ten colleges.

8. A coin is fair.

9. The average rent for a one-bedroom apartment in Santa Cruz is $1000.

10. Coho salmon caught off the coast of Santa Cruz are equally likely to have been hatched in any of the four nearby rivers.

† Question 2. In Orange County, 51% of the adults are males, and 49% are females. One adult is randomly selected for a survey involving credit card usage.

1. Find the probability that the selected person is a male.

2. It is later learned that the selected survey subject was smoking a cigar. Also, 9.5% of males smoke cigars, whereas 1.7% of females smoke cigars (based on data from the Substance Abuse and Mental Health Services Administration). Use this additional information to find the probability that the selected subject is a male.
† Question 3. The president of a particular college says “99% of the alumni support my firing of coach Boggs.” However in the simple random sample of 200 living alumni only 76 of them support firing coach Boggs. Use the 6-step hypothesis method to test the claim that at least 99% support firing coach Bogg (explicitly state the critical region for your test).

† Question 4. (a) Scenario: 161 people who visited one hospital’s emergency room in a 6-month study period with injuries from in-line skating were interviewed. The interviewer found that 53 people were wearing wrist guards and 6 of them had wrist injuries. Of the 108 who did not wear wrist guards, 45 had wrist injuries. We are interested in the difference between the proportions of wrist injuries in the population wearing wrist guard and the population without - perform a 6-step hypothesis test to test the claim that the proportions are different. Then, (b) provide a 95% CI for the population difference between the proportions.
Question 5. Scenario: these data describe the miles per gallon for U.S. cars (sample 1) and for Japanese cars (sample 2); the summary statistics for each sample are shown below.

<table>
<thead>
<tr>
<th></th>
<th>Sample 1</th>
<th>Sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td># of observations</td>
<td>249</td>
<td>79</td>
</tr>
<tr>
<td>mean</td>
<td>20.1</td>
<td>30.5</td>
</tr>
<tr>
<td>sd</td>
<td>6.4</td>
<td>6.1</td>
</tr>
</tbody>
</table>

(a) Perform a 6-step hypothesis test (report the p-value) on the claim that Japanese cars get better mpg than american cars.
(b) What is the 99% CI for the population mean difference?

Question 6. Scenario: these data come from a study of the effects of a drug called Captopril on both the systolic and diastolic blood pressures of patients one half hour after taking the drug. It is believed the drug will decrease both readings. To conduct the study 15 patients’ blood pressures were recorded and then they were given the drug. After 30 minutes their blood pressures were again taken. Test the claim (using the 6-step hypothesis) at a 0.01 level of significance that the drug decreased the diastolic blood pressure.

<table>
<thead>
<tr>
<th>Diastolic Reading</th>
<th>Pre</th>
<th>Post</th>
<th>Post - Pre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>112.3</td>
<td>103.1</td>
<td>-9.3</td>
</tr>
<tr>
<td>SD</td>
<td>10.4</td>
<td>12.6</td>
<td>8.6</td>
</tr>
</tbody>
</table>
Question 7. A byproduct of coal burning is flyash. Flyash is the ash particles that rise with flue gases and contribute to airborne pollution. Much of the flyash produced can be captured by scrubbers in the flue stack. In fact, it is mandated that flyash be captured, because along with the primary constituents of flyash, silicon dioxide (SiO2) and calcium oxide (CaO), other elements present in coal may be present, such as arsenic, mercury, thallium, dioxons, etc., which classify as hazardous waste. A beneficial use of flyash is to add it to cement where the cement acts as a binder holding together the various components and the flyash acts a micro-level aggregate to strengthen concrete. However, too much flyash (as with any high level of aggregate) can weaken the strength of concrete. Flyash % as a strength (sten) factor in concrete compression test (PSI) for 28 day cured concrete. Levine et al. 2008, pg. 615.
(a) According to the JMP output, is the linear model significant at the 0.05 level? Why or why not?

(b) Interpret the meaning of $r^2$ in the context of the problem.

(c) Do the residuals appear healthy? Why or why not? What assumptions (if any) of the model are being violated?

† Question 8. Consider the dataset consisting of the Pineo-Porter prestige score for occupation (response), from a social survey conducted in the mid-1960s. Also available to us are the following explanatory variables: Average education of occupational incumbents, years (education); Average income of incumbents ($) (income); Percentage of incumbents who are women (women).
(a) Which independent variable is most strongly correlated with the dependent variable? Provide this sample correlation.

(b) Provide the 95% CI for the correlation between prestige and women. Is the correlation statistically significant? Why or why not?
(c) Perform a 6-step hypothesis test for the linear model illustrated in the JMP output on slide 13. Recall that the test statistic is given by $F = \frac{MS(Regression)}{MS(Error)}$.

(d) Do the residuals appear to be healthy? Explain.

(e) Provide the linear equation for predicting prestige from education, income, and women. Compute the expected prestige when education is 13, income is 8000, and women is 20.
(f) Interpret the meaning of the slope parameter associated with income in the context of the problem.

(g) Perform a 6-step hypothesis test for the slope of income. Recall that the test stat is $t = \frac{b_{\text{income}}}{s_{b_{\text{income}}}}$ with $df = 98$.

† Question 9. In the dataset “Popular Kids”, students in grades 4-6 were asked whether good grades, athletic ability, or popularity was most important to them. A two-way table separating the students by grade and by choice of most important factor is shown below:

<table>
<thead>
<tr>
<th>Goals/Grade</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades</td>
<td>49</td>
<td>50</td>
<td>69</td>
<td>168</td>
</tr>
<tr>
<td>Popular</td>
<td>24</td>
<td>36</td>
<td>38</td>
<td>98</td>
</tr>
<tr>
<td>Sports</td>
<td>19</td>
<td>22</td>
<td>28</td>
<td>69</td>
</tr>
<tr>
<td>Total</td>
<td>92</td>
<td>108</td>
<td>135</td>
<td>335</td>
</tr>
</tbody>
</table>

Using the 6-step method, test the claim that Goals are different depending on the Grade of the child.
Question 10. Consider the following JMP ANOVA results for a dataset describing a measure of alertness at three different dosage levels of a particular drug.

Using the 6-step method, test the claim that the mean alertness is not the same under all three dosages. Recall that the test statistic is \( F = \frac{MS_{\text{treatment}}}{MS_{\text{error}}} \).