PROBLEM OVERVIEW:
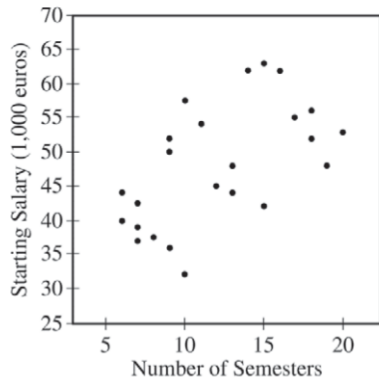
A newspaper reported that the more semesters needed to complete an academic program at this university, the higher the starting salary would be based on a random sample of 24 people who had recently completed an academic program.

Part a:

Students were asked to   determine if the scatterplot supported the newspaper's claim and include justification.


Source: collegeboard.com

Part b:

The following table was provided above part b.

Predictor	Coef	SECoef	T	P
Constant	34.018	4.455	7.65	0.000
Semesters	1.1594	0.3482	3.33	0.003
S= 7.37702	R-Sq = 33.5%		R-Sq(adj) = 30.5%	

Students were asked to identify the slope and interpret the slope in context.

Additional information provided prior to part c.

An independent researched conducted a new analysis by separating the graduates into 3 groups by major as indicated below.
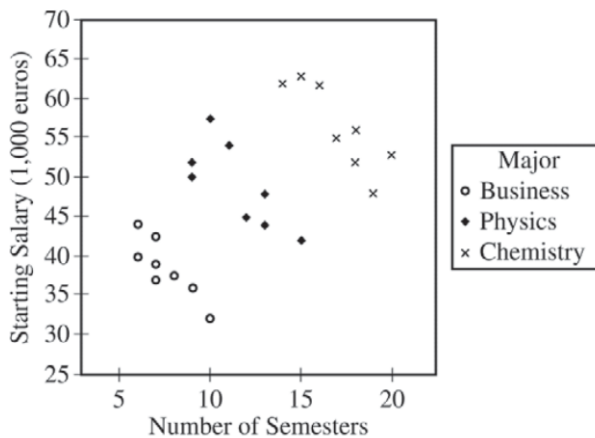
Part c:

Students were asked to describe the association between semesters and starting salary for business majors from the new plot.

Part d:
Students were asked to compare the median starting salaries of all three majors.

Part e:
Students had to tell how the newspaper should modify their article based on the independent researcher's report to give a better description of the relationship between the number of semesters and the starting salary.

Intent of the Questions:

This question assessed a student's ability to (1) use a scatterplot and comment on the relationship between two variables, (2) identify slope and interpret it in context (3) describe the relationship between 2 variables in a scatterplot when a categorical variable is added based on the categories and (4) describe how the associations between 2 variables change when a categorical variable is added.

Solution:

Part a

Students had to indicate a positive association (component 1) and must use the positive association to justify the newspaper's claim(component 2).

Reader Notes:
1. Students did not need to talk about strength and form, so these were ignored.
2. Words like "increasing relationship", "starting salaries are higher if semesters are greater" or "positive correlation" were acceptable.
3. Students who answered the question with words like "maybe", "no" or "kind of" etc. did not receive credit for component 2. You can't sit on the fence!
4. If they talk about the three clusters of points individually in part a, they did not receive credit for either component (ex: stating that each of the 3 clusters had a negative relationship).

Part b:

The slope is 1.1594. For each additional semester that a student is in an academic program, the predicted starting salary increases by 1159.40 euros. This was graded with 3 components: (1) identifying the slope, (2)

Reader notes:
1. Students had to identify the numerical value of the slope.
2. An equation written in context with the variables still required the slope being identified.
3. A wrong numerical value for the slope could still earn the last 2 components.

4. A 1-unit increase in semesters or an equivalent was required in the description of slope. Unspecified number of semesters does not get credit for component 4.
5. Non-deterministic language must be used: predicted starting salary, estimated starting salary, our model predicts, starting salary on average, etc. Words such as about, approximately or according to our model do not satisfy the 5th component.
6. No units are required when giving the predicted salary. Thus if a student said the salary increased in dollars, there was no penalization.
7. "Salary" could be used instead of "Starting Salary".

Part c:

Students had to satisfy 3 components: (1) Identify a negative relationship, (b) state the relationship is strong or linear or both and (3) refer to semesters and salary in context.

For business majors, one can see a strong, negative, linear relationship. Business majors who stayed in the academic program for more semesters tended to have lower starting salaries.

Reader notes:
1. Justification that said "negative association", "negative correlation", "decreasing relationship", "inversely related", or describe a negative association in context of the variables satisfies component 1. Comparison between only two points does not satisfy the component.
2. Students could say "moderate" but did not receive credit for component 2 if they indicated the relationship was weak or nonlinear.
3. Drawing a line through the points and having no statement that the relationship is linear did not receive credit.

Part d:

Students had to correctly describe the relationship between the three majors (component 1) and use estimates of median salaries to compare the majors(component 2).

Reader notes:
1. All three majors had to be considered in the comparison, but if a student indicated that chemistry had the highest median starting salary and business had the lowest median starting salary; it is implied that physics is in the middle.
2. If a student indicated one major's median starting salary was lower than another one, and did not include the third major, credit for component 1 could not be given.
3. Values for the median salaries of the majors are reasonable if the estimate for the median is between the highest and the lowest salary for the major.
4. If no values are provided for median starting salary, the words "median starting salary" must be included when making the comparison.
5. If an interval is provided for median starting salaries, all of the values of that interval must be between the minimum and the maximum for that major.

Part e:

Students had 2 components: (1) Recognizing the individual negative associations and (2)the overall was positive.

Modify the article to be more specific. When specific majors are taken into account, majors that take longer to complete tend to have higher salaries with chemistry the highest, then physic and business has the lowest. However, within each major, students who take a greater number of semesters tend to have lower salaries.

Reader notes:
1. Students could earn credit for component 1 by saying the original newspaper report was correct or by stating majors that require more semesters to complete, there is a negative
2. A well written response without the 3 majors named could receive credit if they refer to the three majors or indicates for an academic major, there is a negative association. (Academic major--- not just "an academic program".
3. Students who indicated the original newspaper report was wrong, incorrect, should be retracted, etc. did not receive credit for the overall positive association. It was acceptable to indicate the first report could be misleading.

Observations:
1. There were a lot of 1's on this question due to many technical requirements. A "P" got a 1.
2. Some of the students gave an analysis, but it was not based on statistics.
3. Many students listed values, but never actually compared the majors.
4. When interpreting slope, non-deterministic language is essential.
5. Many students did not know how to read the table and find slope and the y-intercept.
6. Be specific when comparing and include context. It's amazing how many comparisons lack sufficient context to receive credit.
7. "Median starting salary" or "median salary" were necessary for context. Median alone does not give context.
8. Readers cannot "read-in" what a student has not written and assume they know what a student mean. It has to be clear.
9. Many students did not answer the question for "a". Yes or supports was necessary as well as why.
10. Some students said yes and no for part a, so they did not receive credit.

Recommendations for Teachers:

1. Encourage students to use correct language to describe strength.
2. Work with students on non-deterministic language.
3. Discourage the use of "about", and "approximately" to students. These are not considered non-deterministic because most models have many decimals so any calculation that has any rounding is "approximately" or "about" the value you gave just by rounding. It's not the same as predicted value or salary on average.
4. Give students more practice at describing slope, and analyzing in context. Practice the use of non-deterministic language throughout the year and talk about why it is used.