Unit 2: Day 10: Summative - Analysing Two-Variable Data

Minds On: 5 Action: 65 Consolidate: 5 Total=75 min	**Learning Goal**: • Analyse two-variable data using techniques learned in Unit 2 • Interpret graphs • Draw conclusions based on their analysis	**Materials** • Computer Lab with Fathom • BLM 2.S.1, 2.S.2

				Assessment Opportunities
	Minds On…	**Whole Class → Discussion** Ask the students to reflect on their prior learning in Unit 2 and how it could be used to help choose a college destination. Point out that colleges often publish statistics that they use to sell their programs to students. **Whole Group → Guided Instruction** Inform students that they will be analyzing fictional data on a college program. Distribute BLM 2.S.1 to each student and review the expectations of the Statistical Analysis and Report assignment. **Mathematical Process Focus: Representing** - Students think of different ways to present and analyze data: tables, graphs, and statistics.		
	Action!	**Individual → Completion of Statistical Analysis and Report** Take students to a computer lab and have them work on completing the Analysis of Two-Variable Data BLM 2.S.1. While students are working, circulate around the room to assist students and ensure students are on-task. **Mathematical Process Focus: Connecting** - Students apply the mathematics they learned in Unit 2 to understand and analyze information related to choosing a college program. **Connecting/Oral Questions/Anecdotal Comments:** Pose questions as you circulate about how these techniques could be used to help students decide on which college to attend.	↷	When using Fathom, remind students that they should not have spaces between digits in data (i.e. 4206 not 4 206)
↷	**Consolidate Debrief**	**Whole Group → Teacher Instruction** Inform students that they are to complete the Statistical Analysis and Report at home if it was not completed in class. Collect the report on the decided due date and evaluate it using the rubric provided on BLM 2.S.2. **Expectations/Summative Assignment/Rubric and Scoring Guide:** Evaluate submitted work using the scoring guide given and the communication rubric.		
	Application	**Home Activity or Further Classroom Consolidation** Complete the Analysis of Two-Variable Data and Report at home, if you have not completed it during class time.		

2.S.1 Analysing Two-Variable Data

KU	APPS	Thinking	COM
___/24	___/17	___/11	Level

Part 1

The Municipal Applied Practice College (or MAP College, for short) would like to construct a mathematical model that uses the grade 12 marks of incoming students to predict these same students' marks in their first-year of college. A comparison of these marks for a random sample of first year students is shown in the following table:

Grade 12 Average (%)	78	85	75	76	81	79	82	74	80	83
First-Year Average (%)	74	83	68	70	75	72	73	74	76	78

1. Use Fathom to create a scatter plot for these data. Which variable should be placed on the x-axis? Why? ***(3 marks APP, 1 mark KU)***
2. Describe the correlation for this data, based on the scatter plot. Is there no correlation, positive correlation or negative correlation? If there is a correlation, is it strong or weak? ***(2 marks KU)***
3. Add a line-of-best-fit to the scatter plot and write the equation. ***(2 marks APP)***
4. How well does the line-of-best-fit fit the data in the scatter plot? Justify your answer. ***(2 marks KU)***
5. Identify outliers, if any, and suggest possible reasons for an outlier. ***(1 mark Thinking)***
6. Use the equation of the line-of-best-fit to predict the first-year average of a student with a grade 12 average of 81. ***(3 marks KU)***

Part 2

Suzy is a student in the second year of her program at MAP College. She is starting to think about a job after graduation, and she is wondering about the job market as well as the starting salary she can expect to earn. Suzy does some research and finds employment statistics for graduates of her program and industry surveys of entry-level salaries.
300 students are expected to graduate from Suzy's program in 2011.

Year	Total Number of graduates	Number hired upon graduation	Mean starting salary ($ 000)
1997	218	185	28
1998	205	180	30.5
1999	192	168	30
2000	220	188	32
2001	218	208	33
2002	225	184	32.5
2003	232	185	34
2004	238	193	37
2005	249	198	38.5
2006	266	209	40

1. Model the relationship between the total number of graduates and the number hired by creating a scatter plot and line-of-best-fit. Write the equation. ***(4 marks APP)***
2. How well does the line-of-best-fit fit the data in the scatter plot? Justify your answer. ***(2 marks KU)***
3. Use the equation of this line-of-best-fit to predict how many graduates will be hired in 2011 if there are 300 graduates. ***(3 marks KU)***
4. Identify any outliers in this scatter plot. List some reasons why the data could include an outlier. Do any of these reasons justify deleting the outlier from the scatter plot? ***(3 marks Thinking)***
5. Remove the outlier from the data. Create the new line-of-best-fit. Write the equation of this new line-of-best-fit. ***(2 marks APP)***
6. Use this new equation to recalculate the number of graduates that will be hired in 2011 if there are 300 graduates. ***(3 marks KU)***
7. Compare the results using the original line-of-best-fit and the new line-of-best-fit. Which result seems more reliable? Why? ***(2 marks Thinking)***
8. Which of the two lines gives a more optimistic prediction for the number hired in 2011? ***(1 mark APP)***
9. Are your calculations of graduates hired in 2011 an interpolation or extrapolation? How do you know? ***(2 marks KU)***

Part 3

Use the table of information in part 2.

- List all possible combinations of two variables in the table that have not already been graphed. ***(5 marks Thinking)***
- Use Fathom to create scatter plots for 3 of the combinations listed above. Then, create a line-of-best-fit for each of the scatter plots. Print out all of your graphs. ***(5 marks APP)***
- Describe the correlation for each of the scatter plots. For example, how strong is the relationship between the two variables chosen? For each graph, how well does the line-of-best-fit fit the data? ***(6 marks KU)***

Expectations

You are expected to hand in a typed report. Part 1 and Part 2 should have numbered responses that include the graphs and important information from Fathom. Part 3 should include a report for each of the combinations that you consider, including the graph and discussions of correlation and fit for each graph. Your mark will be determined using the scoring guide and a communication rubric.

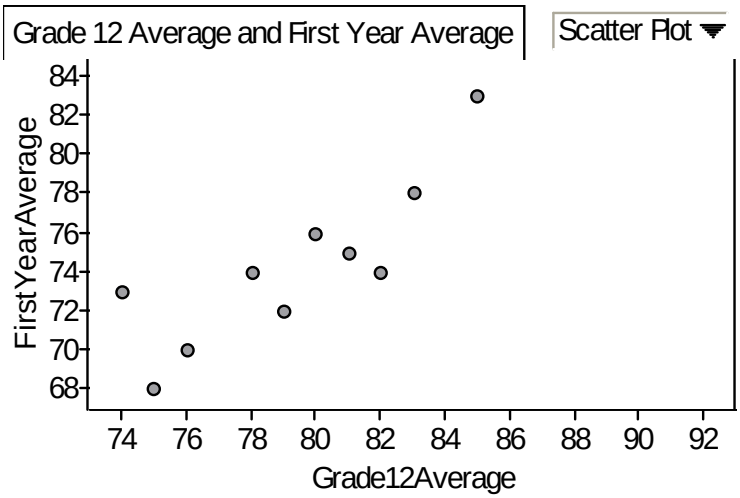
2.S.1 Analysing Two-variable Data (Teacher Notes)

Part 1

The Municipal Applied Practice College (or MAP College, for short) would like to construct a mathematical model that uses the grade 12 marks of incoming students to predict these same students' marks in their first-year of college. A comparison of these marks for a random sample of first year students is shown in the following table:

Grade 12 Average (%)	78	85	75	76	81	79	82	74	80	83
First-Year Average (%)	74	83	68	70	75	72	73	74	76	78

1. Use Fathom to create a scatter plot for these data. Which variable should be placed on the x-axis? Why?

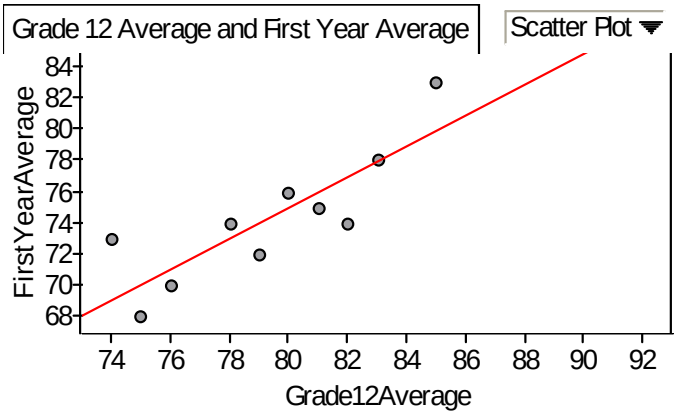


The Grade 12 average is the independent variable and therefore goes on the x-axis.

2. Describe the linear correlation for this data, based on the scatter plot. Is there no correlation, positive or negative correlation. If there is a correlation, is it strong or weak?

The correlation is positive and relatively strong

3. Add the line-of-best-fit to the scatter plot. Write the equation.



FirstYearAverage = 0.983Grade12Average - 3.6; r^2 = 0.71

4. How well does the line-of-best-fit fit the data in the scatter plot? Justify your answer.

   The line-of-best-fit fits the data well because the points do not deviate far from the line.

5. Identify outliers, if any, and suggest possible reasons for an outlier.

   There are no outliers.

6. Use the equation of the line-of-best-fit to predict the first-year average of a student with a grade 12 average of 81.

   First Year Average = 0.983* Grade 12 average – 3.6

   First Year Average = 0.983* (81) – 3.6

   First Year Average = 76.02

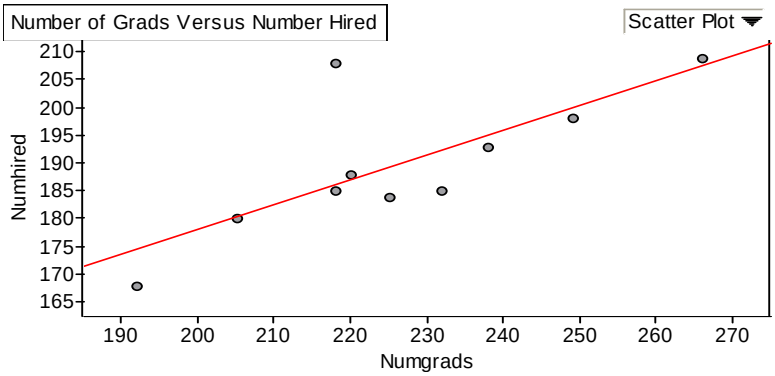
Part 2

Suzy is a student in the second year of her program at MAP College. She is starting to think about a job after graduation, and she is wondering about the job market as well as the starting salary she can expect to earn. Suzy does some research at the college's placement centre and finds employment statistics for graduates of her program and industry surveys of entry-level salaries.
300 students are expected to graduate from Suzy's program in 2011.

Year	Number of graduates	Number hired upon graduation	Mean starting salary ($ 000)
1997	218	185	28
1998	205	180	30.5
1999	192	168	30
2000	220	188	32
2001	218	208	33
2002	225	184	32.5
2003	232	185	34
2004	238	193	37
2005	249	198	38.5
2006	266	209	40

2.S.1 Analysing Two-variable Data (Teacher Notes)

1. Model the relationship between the total number of graduates and the number hired by creating a scatter plot and line-of-best-fit. Write the equation.



Numhired = 0.448Numgrads + 88; r^2 = 0.57

2. How well does the line-of-best-fit fit the data in the scatter plot? Justify your answer.

   The line-of-best-fit fits the data well because the points do not deviate far from the line.

3. Use the equation of this line-of-best-fit to predict how many graduates will be hired in 2011 if there are 300 graduates.

   Number Hired = 0.448* Number of Grads + 88

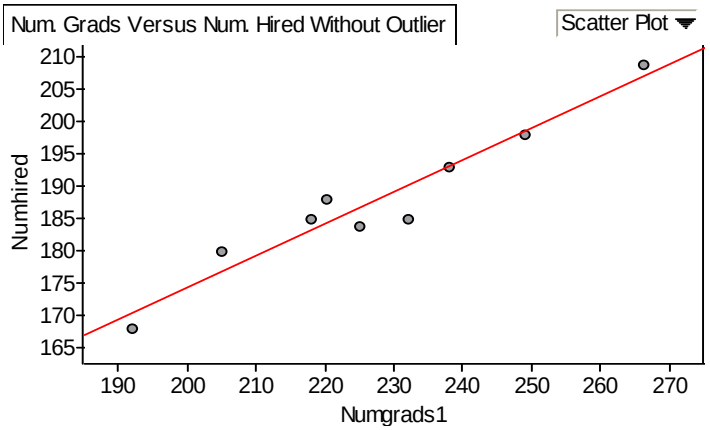
   Number Hired = 0.448* (300) + 88

   Number Hired = 222.4

4. Identify any outliers in this scatter plot. List some reasons why the data could include an outlier. Do any of these reasons justify deleting the outlier from the scatter plot?

   (218, 208) is an outlier. Answers may vary.

5. Remove the outlier from the data. Create the new line-of-best-fit. Write the equation of this new line-of-best-fit.



Numhired = 0.498Numgrads1 + 74.5; r^2 = 0.93

6. Use this new equation to recalculate the number of graduates that will be hired in 2011 if there are 300 graduates.

   Number Hired = 0.498* Number of Grads + 74.5

   Number Hired = 0.498* (300) + 74.5

   Number Hired = 223.9

7. Compare the results using the original line-of-best-fit and the new line-of-best-fit. Which result seems more reliable? Why?

   The new line-of-best-fit seems more reliable because there is no outlier causing undue influence.

8. Which of the two lines gives a more optimistic prediction for the number hired in 2011?

   The new line-of-best-fit gives a slightly more optimistic prediction for the number hired in 2011.

9. Are your calculations of graduates hired in 2011 an interpolation or extrapolation? How do you know?

   The prediction for 2011 is an extrapolation because the point is outside the range of data provided.

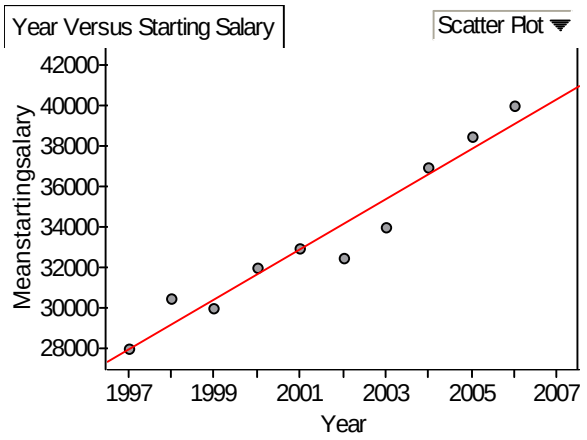
Part 3

Use the table of information in part 2.

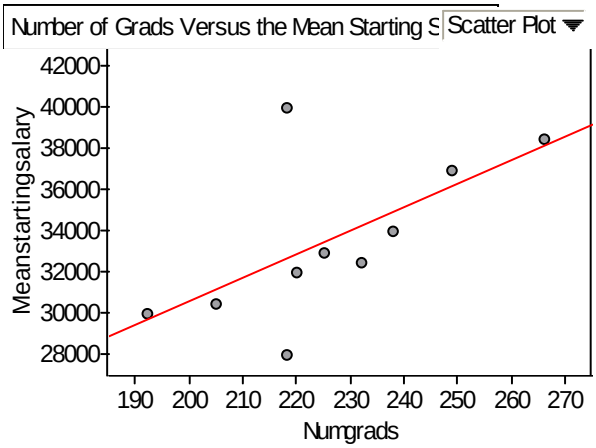
- List all possible combinations of two variables in the table that have not already been graphed.
- Use Fathom to create scatter plots for all of the combinations. Then, create a line-of-best-fit for each of the scatter plots. Print out all of your graphs.
- Describe the correlation for each of the scatter plots. For example, how strong is the relationship between the year and the mean starting salary? For each graph, how well does the line-of-best-fit fit the data?

The scatter plots of all possible relationships are below.

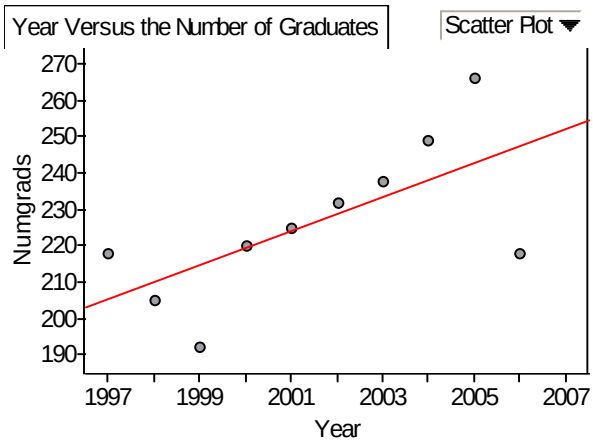
2.S.1 Analysing Two-variable Data (Teacher Notes)



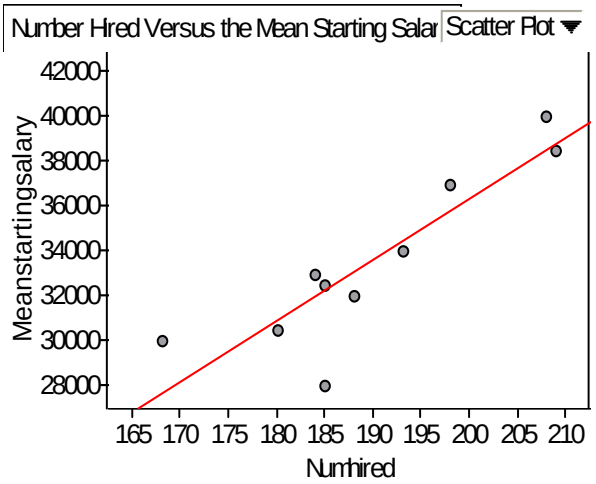
Meanstartingsalary = 1239.4Year - 2447100; r^2 = 0.94



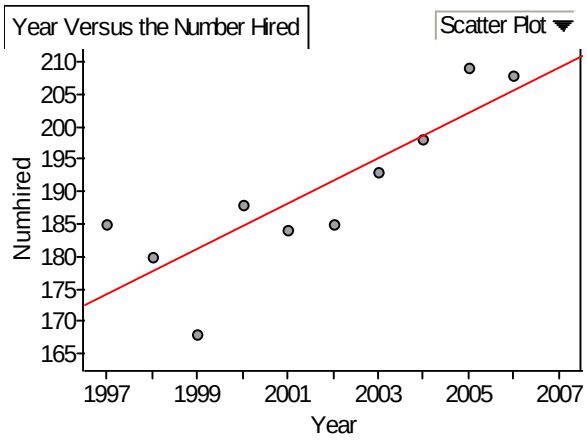
Meanstartingsalary = 114Numgrads + 7800; r^2 = 0.39



Numgrads = 4.68Year - 9200; r^2 = 0.44



Meanstartingsalary = 270.4Numhired - 17800; r^2 = 0.78



Numhired = 3.491Year - 6797; r^2 = 0.70

2.S.2: Two-Variable Data - Evaluation Rubric

	Category	Level 1	Level 2	Level 3	Level 4
Communication	Correct use of mathematical symbols, labels, units and conventions across a range of media. (Including proper labelling of graphs)	Sometimes uses mathematical symbols, labels and conventions correctly.	Usually uses mathematical symbols, labels and conventions correctly	Consistently uses mathematical symbols, labels and conventions correctly.	Consistently and meticulously uses mathematical symbols, labels and conventions, recognizing novel opportunities for their use.
	Appropriate use of mathematical vocabulary.	Sometimes uses mathematical vocabulary correctly	Usually uses mathematical vocabulary correctly when expected.	Consistently uses mathematical vocabulary correctly when expected.	Consistently uses mathematical vocabulary correctly, recognizing novel opportunities for its use.
	Degree of clarity in explanations and justifications in reporting.	Explanations and justifications are partially understandable.	Explanations and justifications are understandable by me, but would likely be unclear to others.	Explanations and justifications are clear for a range of audiences.	Explanations and justifications are particularly clear and detailed.