Unit 1 - Working with Data

MAP 4C

Foundations for College Mathematics

BIG PICTURE

Students will:
Personalize the course, and capitalize on their interests, post-secondary and career pathways
Collect, analyze, and summarize one-variable data using a variety of tools and strategies, and interpret and draw conclusions from the data
Distinguish situations requiring one-variable and two-variable data analysis
Analyze the use and misuse of data in the media

Day	Lesson Title	Math Learning Goals	Expectations
1	A Survey of Surveys *Lesson Included*	• Analyze a variety of surveys/questionnaires (e.g. Teen Magazine, Match Making Valentine Questionnaire, Census at Schools, etc.) in order to describe the characteristics of an effective survey/questionnaire	DM1.2
2	Designing a Questionnaire *Lesson Included* ***New Jan/08****	• Design and critique questionnaires to collect data about the class (e.g. college destination, career interests, personal interests, mathematics background, etc.) • Create a class questionnaire in order to conduct a survey about the class (consider incorporating questions from the Census at School questionnaire for later comparisons in Day 6) • Assessment of class interests	DM1.2
3		• Use examples from the media that include common statistical terms (e.g. percentile, quartile, standard deviation) and expressions in order to review and interpret them. • Analyze the class data using the statistical terms and expressions for use by the media	DM2.1
4-5	Statistics in the Media *Lessons Included*	• Interpret statistics presented in the media. • Explain how the media misuses statistics. • Create a media advertisement from the class data that would promote a certain point of view in order to lobby for a school interest • Assess the validity of the conclusions presented by the class media advertisements • Assess the validity of the conclusions presented in the media	DM2.3, 2.4
6-7	Are more or Less People Smoking *Lesson Included* ***New Jan/08****	• Analyze data from a secondary source (e.g. Census at School) with technology (e.g. Fathom, spreadsheet, graphing calculator) • Validate class analysis of common attributes using the secondary source (e.g. sample size, demographic bias) • Look for mathematical relationships in the data • Distinguish situations requiring one-variable and two-variable data analysis	DM2.1, 2.3, 2.4. 1.1, 1.3
8	How Popular is your Program *Lesson included*	• Summative Assessment (e.g. collection of case studies with individual report, data project with report)	

Unit 1: Day 2: Designing a Questionnaire

Minds On: 10 Action: 50 Consolidate:15 Total=75 min	**Math Learning Goals:** • Design and critique questionnaires to collect data about the class (e.g. college destination, career interests, personal interests, mathematics background, etc.) • Create a class questionnaire in order to conduct a survey about the class (consider incorporating questions from the Census at School questionnaire for later comparisons in Day 6) • Assessment of class interests	**Materials** • Chart paper and markers • Post-its

				Assessment Opportunities
	Minds On…	**Whole Class → Brainstorm** Lead students in a brainstorming session to determine what survey topics they may be interested in pursuing (e.g. post-secondary destinations, extra-curricular involvement, health/smoking habits, part time jobs, etc.). Write each survey topic at the top of a sheet of chart paper.		You may want refer back to the surveys from last day to peak student interest &/or encourage ideas before/during the brainstorming
	Action!	**Small Groups → Brainstorm** Divide the class into groups according to how many survey topics the class came up with in the brainstorming session. Distribute one piece of chart paper to each group. Give students time to write down as many survey questions (either open ended or not) as they can think of that would fit that category. Once the group has exhausted possibilities (or a time limit has been reached) instruct students to rotate chart papers to the next group. The groups can then read what has been written so far and add to it. As the chart papers continue to rotate, students will need less time to add to the existing questions. **Teamwork/Observation/Mental Notes:** Circulate during the brainstorming and observe student work and encourage students to stay on task. **Whole Class → Gallery Walk** Post completed chart papers around the room and allow students to walk around and read what has been written on each. Have students stand in front of the chart paper that they wish to work on. You may need to ask some students to make a second choice if one chart paper is 'over populated.' **Small Groups → Discussion** Have new groups formed above take chart paper back to a desk and select the best 8 – 10 questions from their chart paper to create a survey with. **Reflecting/Observation/Anecdotal:** Observe students' reasoning as they make their selections for the survey and give them feedback. **Mathematical Process Focus: Reflecting –** students will **reflect** on questions brainstormed and choose the best options for their survey.		If you have access to computers in the classroom you may wish to require that students type their survey. These surveys or 'the best one' can be used as part of the lesson next day.
	Consolidate Debrief	**Small Groups → Peer Assessment** Once they have put together a rough draft of their survey have groups swap surveys. Groups should then read through the survey, mentally answering the questions as they go and write feedback on post-its. When they are finished, the survey & feedback should be returned to the group it came from for groups to create their good copy for submission. **Expectations/Questionnaire/Anecdotal:** Assess the student's questionnaires for quality, accuracy and clarity.		
	Application	**Home Activity or Further Classroom Consolidation** Complete and hand-in the survey if you have not finished it already. Find in the media (online/newspaper/magazine/...) an article that references statistics and be prepared to present it to the class next day.		

Unit 1: Day 6 & 7: Are More or Less People Smoking?

Minds On: 30 Action: 90 Consolidate:30 Total=150 min	**Learning Goal**: • Retrieve and analyze data from the internet with technology (e.g. Fathom, spreadsheet, graphing calculator) • Determine which statistical measures are meaningful for the data and determine any forms of bias • Look for mathematical relationships in the data • Distinguish situations requiring one-variable and two-variable analysis	**Materials** • Computer Lab • graphing calc. (optional) • BLM 1.6.1 • BLM 1.6.2 • BLM 1.6.3 or BLM 1.6.4 or BLM 1.6.5

			Assessment Opportunities
Minds On…	**Small Groups → Think, Pair, Share** Have students work individually to answer the questions below. Then discuss their answers in pairs. *Which measure of central tendency (mean, median, or mode) is best to use in each situation? Why?* ***1.*** *Marks on a test: 42 68 72 73 73 75 77 81 82 82 83 84* ***2.*** *Sizes of a particular shoe sold at Payless Shoes Store in one day:* *5 5 5 6 6 6 6 6 6 7 7 7 8 8 9 9 9 10 11* **Whole Class → Discussion** Point out to students that mean, median and mode are all examples of **one variable statistics** (along with standard deviation and range) as they only ever refer to one variable. Introduce the idea of **two variable statistics** (relationships between two different variables) using the following game, where one variable is given and students must think of the second variable that would fit the situation. **Groups of 4 → Game** Write the following six sentences on the board, leaving the blanks in each. Divide students into groups of four. • *As population of a city increases, _____ increases.* • *As population of a city increases, _____ decreases.* • *As temperature increases, _____ increases.* • *As temperature increases, _____ decreases.* • *As a person's age increases, _____ increases.* • *As a person's age increases, _____ decreases.* For each sentence, a group must think of **at least** one word that could be used in the blank. Students may list more than one possibility for each sentence. Each group should share their answers with the class, scoring a point for each original suggestion not given by any other group. **Expectations/Presentation/Anecdotal Feedback:** Give groups feedback on the answers that they give and discuss reasons for answers to ensure understanding.		Literacy Strategy: Think, Pair, Share
Action!	**Pairs → Activity** Distribute and review BLM 1.6.1 and one of BLM 1.6.3 or BLM 1.6.4 or BLM 1.6.5 (depending on technology). Students should complete the activity in pairs. **Mathematical Process Focus: Reasoning and Proving** – Students will make conclusions based on the data collected using different statistical methods.		On BLM 1.6.1, you need to fill in what type of technology the students should use to create their graphs (part b) and do the calculations (part c).
Consolidate Debrief	**Whole Class→ Discussion** Discuss the difference between one variable statistics and two variable statistics. Focus on what questions required the use of one variable statistics and what questions required the use of two variable statistics.		
Application	**Home Activity or Further Classroom Consolidation** Complete the questions on BLM 1.6.2.		

1.6.1: Activity: Are More or Less People Smoking?

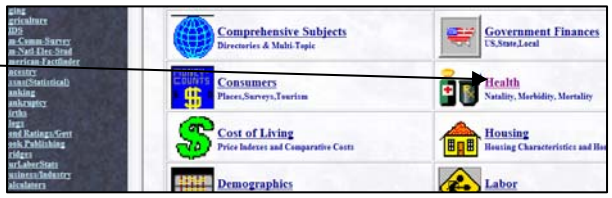
In this activity you will:

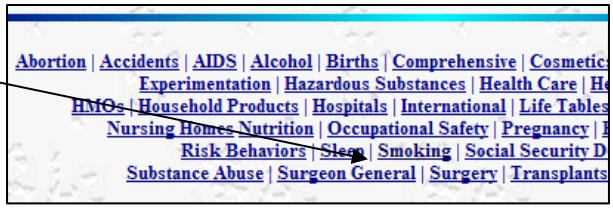
- Analyze data from the internet on the number of smokers over a span of years in the U.S. using technology (e.g. Fathom, spreadsheet, graphing calculator)
- Determine which statistical measures are meaningful for the data
- Look for mathematical relationships in the data and determine any forms of bias
- Distinguish situations requiring one-variable and two-variable analysis

Part 1: Investigating Data on Smoking

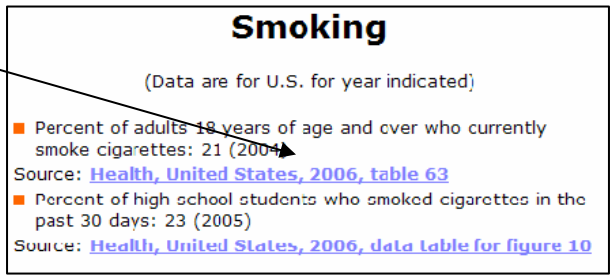
Part A: Retrieving Data on Smoking

- Go to http://www.lib.umich.edu/govdocs/stats.html .
- Click on **Health**
- Click on **Smoking**
- Click on **Cigarette Smoking**





- Click on **Health, United States, 2006, table 63**



Part B: Two Variable Analysis - Creating A Time Series Graph of Data

Using <Fill in Excel/Fathom/Graphing Calculator>:

- Create a scatter plot showing the percentage of high school seniors (male and female) who smoke (with the year as your independent variable). Use the data for all years listed. This data can be found on **page 290** of the website given in part A.

- Create a scatter plot showing the percentage of tenth graders (male and female) who smoke (with the year as your independent variable). Use the data for all years listed. This data can be found on **page 290** of the website given in part A.

1.6.1: Activity: Are More or Less People Smoking? (Continued)

- Create a scatter plot showing the percentage of all persons age 18 and older who smoke (with the year as your independent variable). Use the data for all years listed. This data can be found on **page 283** of the website given in part A.

Answer the following 3 questions for each of the age groups given below using the scatter plots you created above.

i) Has the number of smokers been increasing or decreasing since 1980?
ii) What are some reasons for this trend?
iii) Were you surprised by the trend you saw? Why?/Why not?

a) Percentage of high school seniors (male and female) who smoke

b) Percentage of 10th graders (male and female) who smoke

c) Percentage of persons age 18 and older (male and female) who smoke

Part C: One Variable Analysis - Calculating Measures of Central Tendency and Dispersion

Use <Fill in Excel/Fathom/Graphing Calculator> to calculate the measures below.

a) Percentage of high school seniors (male and female) who smoke

Mean = _____ Median = _____ Mode = _____ Standard Deviation = _____

b) Percentage of 10th graders (male and female) who smoke

Mean = _____ Median = _____ Mode = _____ Standard Deviation = _____

c) Percentage of persons age 18 and older (male and female) who smoke

Mean = _____ Median = _____ Mode = _____ Standard Deviation = _____

1.6.1: Activity: Are More or Less People Smoking? (Continued)

Answer the following questions based on your calculations.

1. Which measure of central tendency above is most meaningful (mean, median, or mode)? Why?

2. What does the measure you chose in part (a) above tell you about the data?

3. What does the standard deviation tell you about the data?

4. Use the measures of central tendency you calculated to determine which age group (grade 10's, high school seniors, or persons age 18 and over) has the highest percentage of smokers.

5. Use your calculations of standard deviation to determine which age group (grade 10's, high school seniors, or persons age 18 and over) fluctuated the most with the percentage of smokers.

1.6.1: Activity: Are More or Less People Smoking? Part 2 (Continued)

Part 2: Investigating Data on Cancer

Part A: Retrieving Data on Cancer

- Go to the website for The National Centre for Health Statistics at http://www.cdc.gov/nchs/howto/w2w/w2welcom.htm.
- Click on **FactStats A to Z** (found in menu on left side of page)
- Scroll down and click on **Cancer**
- Scroll down and under **"More Data"** click on **Cancer incidence rates by sex, race, and Hispanic origin, 1990-2003 Health, United States, 2006, table 53.**

Part B: Two Variable Analysis - Creating a Time Series Graph of Cancer Data

Using <Fill in Excel/Fathom/Graphing Calculator>:

- Create a scatter plot showing the number of new cases of **all types of cancer** per 10, 000 population (both male and female) (with the year as your independent variable). This data can be found on **page 261** of the website given in part A.
- Create a scatter plot showing the number of new cases of **lung and bronchus cancer** per 10, 000 population (include both male and female). **Note:** You will need to add the male and female data to get a **combined total** (with the year as your independent variable). This data can be found on **page 261** of the website given in part A.

Answer the following questions based on the scatterplots you created above.

1. Examine the scatter plot that shows the number of new cases of **all types of cancer** per 10,000 population (both male and female) in the United States. Has the number of new cases of cancer been increasing or decreasing since 1990? What are some reasons for this trend? Were you surprised by the trend you saw?

1.6.1: Activity: Are More or Less People Smoking? Part 2
(Continued)

2. Examine the scatter plot that shows the number of new cases of **lung and bronchus cancer** per 10,000 population (both male and female) in the United States. Has the number of new cases of cancer been increasing or decreasing since 1990? What are some reasons for this trend? Were you surprised by the trend you saw?

Part 3: Comparing Smoking and Cancer Data

1. Does there seem to be a relationship between the number of smokers over the years and the number of new cancer cases over the years? Describe the relationship you see. What are some things that should be investigated before you can be certain about this relationship?

2. Do you think similar results would have occurred if the surveys for the number of smokers and cancer cases were completed in Canada? What do you base your opinion on?

3. What other variable(s) could you collect that may help to explain the trend(s) in your data?

1.6.1: Activity: Are More or Less People Smoking?
(Continued)

Part 4: Assessing Bias

List any possible sources of bias that may exist in:

a) Sample Selection _____

_____

_____

b) Data Collection _____

_____

_____

c) Data Analysis _____

_____

_____

d) Reporting of Data _____

_____

_____

1.6.2: Analyzing Data – Follow-Up Questions

1. For each of the data sets below, decide whether <u>one variable or two variable</u> analysis should be used.

a) A class set of test marks.

b) The percentage of Canadian teenagers who smoke and the number of programs available to help people quit smoking.

c) Number of people in a household that have jobs.

d) The winning times for the women's 100m dash in the Olympics from 1975 to 2000.

e) Number of canned drinks sold and the number of cans recycled.

f) The ages of individuals involved in car accidents.

g) The salaries of players on a professional sports team.

h) Students' marks and the number of hours spent watching television.

i) Model year of cars on the road.

j) The percentage of Canadians who are overweight between years 1980 and 2005.

k) Team scores in a national mathematics competition.

2. For each of the data sets above involving one variable analysis, state whether the mean, median or mode would be the most appropriate measure.

3. Using one of the surveys from earlier in the unit, select two questions where the answers would be used for single variable analysis (eg. mean, median, mode, standard deviation, range). For each selected question, explain which measure would be most appropriate to analyse the data collected.

4. Using one of the surveys from earlier in the unit, select two questions where the answers would be used for two variable analyses. Predict the relationship that exists between the variables. (For example, each person taking a particular survey must give their age and their height. These two variables are most likely related: the older a person is, the taller they are likely to be.)

1.6.3: Statistical Analysis Using a Graphing Calculator

Creating a Scatter Plot

Step 1

- Press ***Stat, 1*** to access the list editor. Enter your data into L1 and L2. L1 refers to x-values and L2 refers to y-values.

Step 2

- Turn the stat plot on by pressing ***2nd, y=, enter, enter***. Make sure your screen looks like Figure 1.

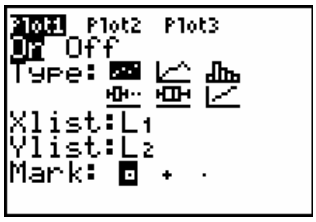


Figure 1

Step 3

- Press ***graph***.
- Press ***zoom, 9***. This will adjust the window settings, providing a graph of the data.

Calculating Measures of Central Tendency and Standard Deviation

- Clear the memory by pressing ***2nd, +, 7, 1, 2***.
- Press ***Stat, 1*** to access the list editor.
- Enter your data into L1.

To Find the Median and Standard Deviation:

- Press ***1*** (1-Var Stats) and ***enter***.
- $\bar{x}$ is the mean and ϑx is the standard deviation

To Find the Median:

- Press ***2nd, Stat*** and move your cursor to **MATH** (use right arrow key)
- Press 2nd, 1, enter to find the median of the data in L1.
- Press ***Stat*** and then move your cursor to **CALC** (use right arrow key).

To Find the Mode:

- Press Stat and then 2, 2nd, 1, enter (this will sort your data in L1 in ascending order).
- Now you can go back to L1 by pressing Stat, 1 and check to see what the mode is.

1.6.4: Statistical Analysis Using Excel

Creating a Scatter Plot

Step 1

- Enter your data into the spreadsheet by putting x-values in column A and y-values in column B. See Figure 1.

Step 2

- Highlight your data. Click on the insert pull-down menu at the top of the screen and then select chart. See figure 2.

Step 3

- Highlight ***XY (Scatter)*** and click on ***next***.

Step 4

- Click on ***next***.
- Under ***Chart Title***, enter a title for your graph.
- Under ***Value (X) axis***, enter a title for the x-axis.
- Under ***Value (Y) axis***, enter a title for the y-axis.
- Click on ***Finish***.
- Your graph will be displayed on the screen.

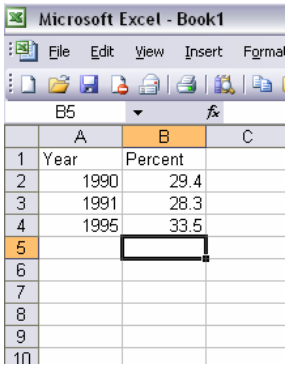


Figure 1

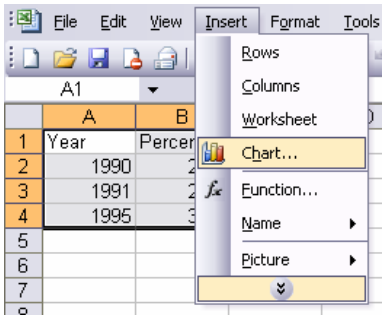


Figure 2

Calculating Measures of Central Tendency and Standard Deviation

- Enter data values in cells A1 to A10 (if you have 10 values)
- In any other cell, type in the expressions below to find what you need
- =average(A1:A10) (mean)
- =median(A1:A10) (median)
- =mode(A1:A10) (mode)
- =stdev(A1:A10) (standard deviation)

1.6.5: Statistical Analysis Using Fathom

Creating a Scatter Plot using Fathom

Step 1

- Create a **case table** by clicking and dragging the table icon from the menu bar onto the main screen. See Figure 1.



Figure 1

Step 2

- Click on **<new>** and type in a name for the x-axis (e.g. Year). To put a space between each word, use the underscore symbol.
- Enter your x-values into the column.

.

Step 3

- Click on **<new>** (2nd column) and type in a name for the y-axis (e.g. Percentage of High School Seniors that Smoke).
- Enter your y-values into the column.
- Your table should like similar to Figure 2 (you will have more data entered)

.

Collection 1

	Year	Percent...
1	1990	29.4
2	1991	28.3
3	1995	33.5

Figure 2

Step 4

- Create a graph by clicking and dragging the graph icon from the menu bar onto the main screen. See Figure 3.



Figure 3

Step 5

- In your table, left click on the heading in your first column (i.e. Year) and drag it onto the horizontal axis of your graph.
- In your table, left click on the heading in your second column and drag it onto the vertical axis of your graph.
- Left click in the top right corner of your graph and select scatter plot. You may be able to skip this step if scatter plot is already selected for you.

Step 6

- Add a title to your graph by double-clicking on "Collection 1" which can be found in the top left corner of your graph.
- Type in your title (remember to use underscores for spaces between words.

1.6.5: Statistical Analysis Using Fathom (Continued)

Calculating Measures of Central Tendency and Standard Deviation

- Open a new case table and enter in the data values.
- Label the column of data as "**values.**" See Figure 1.
- Under the ***Edit*** menu choose ***inspect collection***.
- Click on **Measures**.
- Click on new and type "Mean".
- Double click in the box for the formula for the mean and a formula box will appear.

Figure 1

- Type **mean(values)** and then press enter. See Figure 2.

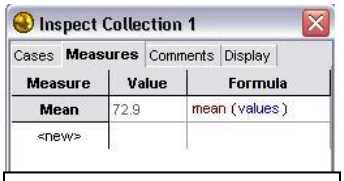


Figure 2

- repeat the above 3 steps to find the median, and standard deviation
- **Note:** To get the standard deviation type **sampleStdDev(values)**