

Do you like to sleep in? Maybe you should move to Resolute Bay, Nunavut where the sun doesn't rise for three months. But you will have to leave by April, when the town begins a three-month stretch of daylight!



Inquire

Analysing Scatter Plots Using Software

Materials

- Microsoft Excel
- daylighthours.xls
- snowrain.xls
- access to the Internet and E-STAT

Work with a partner.



Part A: Using Microsoft Excel

- Open the file *daylighthours.xls*.

1. The spreadsheet shows the latitudes of different locations and the number of daylight hours on August 15. Describe the relationship between latitude and hours of daylight.

	A	B
1	Daylight Hours on August 15	
	Latitude	Length of daylight time
2	(°N)	(h)
3	0	12.1
4	10	12.5
5	20	12.8
6	30	13.3
7	40	13.8
8	50	14.5
9	60	15.8
10	70	18.5
11	80	24.0
12	90	24.0

- Highlight cells A3 to B12.

Click the **Chart Wizard** icon.

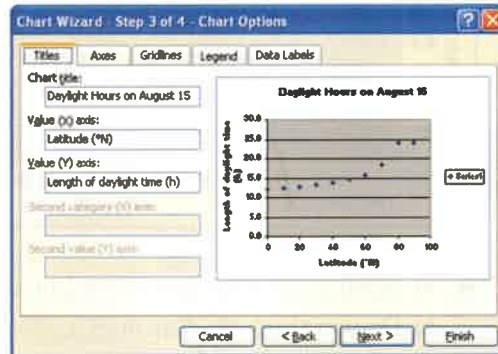


Under *Chart type*: select **XY (Scatter)**.

Click **Next**. Accept the data range by clicking **Next** again.

The title screen should appear. If it does not, click on the **Titles** tab.

Enter the data headings as the chart titles.

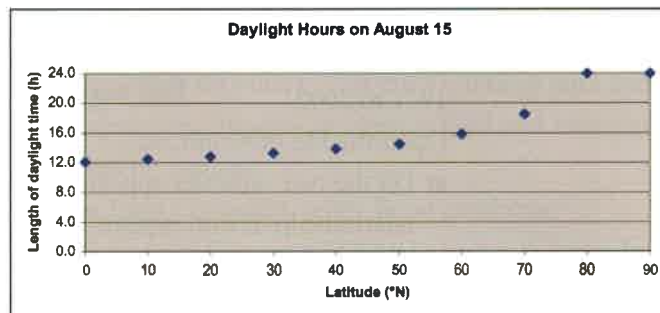


You may find it easier to type "degrees North" than "°N".

The latitude of the equator is 0°N.
The latitude of the North Pole is 90°N.

There are 24 h in 1 day.

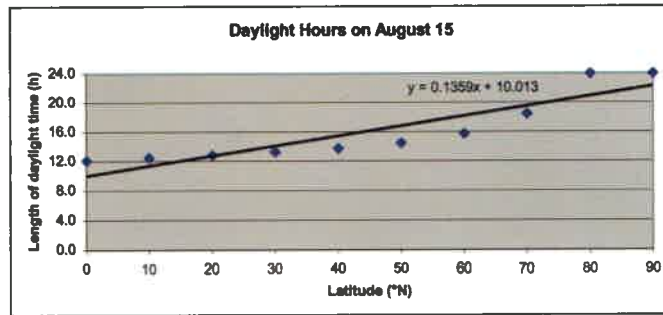
- Click on the **Legend** tab. Deselect **Show Legend**.
Click **Finish** to embed the graph into your spreadsheet.
- Right click the horizontal axis of the graph and select **Format Axis...**
Click on the **Scale** tab.
Set **Minimum** to 0 and **Maximum** to 90.
- Right click the vertical axis and select **Format Axis...**
Click on the **Scale** tab.
Set **Minimum** to 0 and **Maximum** to 24.



2. Describe the correlation between the variables. Compare it with your description of the relationship from question 1.
3. Do you think a linear model would represent the data well?
Explain your thinking.

The **Chart** menu will not appear unless a graph is selected.

- From the **Chart** menu, select **Add Trendline**.
Select **Linear** from the list of Trend/regression types.
Click on the **Options** tab, then select **Display equation on chart**.
Click **OK**. Your graph should look similar to the one shown here.



4. a) What is the equation of the line of best fit?
b) Do you think the line does a good job of representing these data? Justify your answer.
c) Would the linear model provide reliable estimates of daylight hours? Justify your answer.



5. Environment Canada calculates “weather normals” that represent typical weather data for different locations. The current normals are based on data collected from 1971 to 2000.

	A	B	C
1	City or town	Average annual snowfall (cm)	Average annual rainfall (mm)
2	Belleville	155.7	735.9
3	Cameron Falls	237.5	576.5
4	Chalk River	195.4	669.2
5	Cobourg	106.0	765.5
6	Dresden	84.6	759.5
7	Hamilton	161.8	764.5
8	Kapuskasing	313.0	544.5
9	Renfrew	195.5	616.0
10	Samia	125.0	732.5
11	Sault Ste. Marie	302.9	634.5
12	Timmins	313.4	558.5
13	Toronto	133.1	709.5

Open the file *snowrain.xls*.

- a) Do the two variables appear to be related? If so, describe the relationship. If not, explain why not.
- b) Create a scatter plot for the data. Describe any correlation you see. Does the graph support your answer to part a)?
- c) Add a line of best fit to the graph. How well do you think it represents the data? Justify your answer.
- d) Petawawa receives an average of 228.5 cm of snow each year. Based on the line of best fit, what would you expect the average annual rainfall to be in Petawawa? How close was the prediction to the actual average of 615.9 mm?

Part B: Using E-STAT

- Go to the Statistics Canada Web site.
Click **English**.
Select **Learning resources** from the menu on the left.
Click on **E-STAT** in the yellow box on the right.
Click on **Accept and enter**.
If you are working from home, you will need to enter the user name and password assigned to your school.

CANSIM provides data taken over time. *Census databases* offer information about entire populations taken once every 5 years.

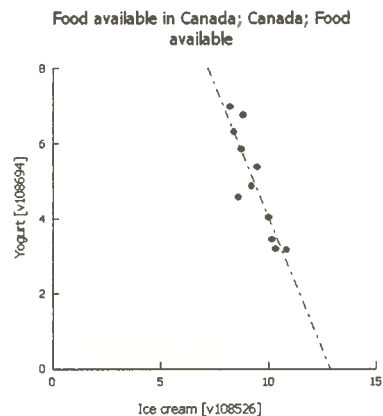
The screenshot shows the Statistics Canada website interface. At the top, there are navigation links for 'Français', 'Contact us', 'Help', 'Search', and 'Canada site'. Below this is a blue header with 'STATISTICS CANADA' and 'CANADA'S NATIONAL STATISTICAL AGENCY'. The main content area is titled 'E-STAT: Table of contents' and lists various categories of data. A left sidebar contains links like 'HOME', 'E-STAT', 'About E-STAT', 'What's new in E-STAT', 'Table of contents', 'User guides', 'Search CANSIM', 'Search Censuses', 'Search sep 2001', 'Help frequently asked questions', 'Contact E-STAT', and 'Learning resources'. The 'Table of contents' section is highlighted in yellow.

E-STAT: Table of contents	
Economy	
Business performance and ownership	Manufacturing
Business, consumer and property services	Prices and price indexes
Construction	Retail and wholesale
Economic accounts	Science and technology
Information and communications technology	Transportation
International trade	
Land and Resources	
Agriculture	Environment
Energy	
People	
Aboriginal peoples	Income, pensions, spending and wealth
Children and youth	Labour
Culture and leisure	Languages
Education, training and learning	Population and demography
Ethnic diversity and immigration	Seniors
Families, households and housing	Society and community
Health	Travel and tourism
Nation	
Crime and justice	Government
Historical Censuses of Canada	
1665-1971	
Elections Canada	
2000: Provinces and Territories	2000: Federal electoral districts
1997: Provinces and Territories	1997: Federal electoral districts

Per capita means per person.

- The E-STAT table of contents will be displayed.
In the *Land and Resources* section, click on **Agriculture**.
From the list of CANSIM data, click on **Food and nutrition**.
Click on table **002-0011**. This table contains data about the per capita consumption of various food and beverage items throughout Canada.
- Under *Food Categories*, select **Food available**.
Under *Commodity* you will be selecting two items, but there are too many choices to see all at once. Click on **View checklist and footnotes** to display all the items more conveniently.
Scroll down and select **Ice cream (litres per year)** and **Yogurt (litres per year)**.
Scroll up to the top of the page and click **Return to picklist**.
Set the Reference period from **1996** to **2006**.
Click on **Retrieve as individual Time Series**.

- Create a scatter plot. Select **Scatter graph with line of best fit (linear regression)** and click **Retrieve Now**. Click on **Modify Graphic**. After **Origin**: select **Start axis at 0**. Click **Replot**.



6. a) What two variables are being compared in this graph?
 b) What data are not included in the graph?
 c) Describe the correlation displayed by the scatter plot.
 d) What does the direction of the line of best fit tell us about the relationship between ice cream consumption and yogurt consumption? Why do you think this relationship occurs?

- Click on the back button three times to return to the *Output specification screen*. In the *Screen output – table* box, under *HTML, Table*, select **Time as rows** and click **Retrieve now**. Click and drag to highlight entire table, including the headings. With the table highlighted and the cursor on the highlighted table, right click and **Copy**.
- Open a new *Microsoft Excel* spreadsheet. Right click on cell A1 and select **Paste**. The data from E-STAT will appear.

Widen the columns to reveal the table headings.

	A	B	C
1	Annual	v108526 - Canada; Food available; Ice cream (litres per year)	v108694 - Canada; Food available; Yogurt (litres per year)
2	1996	10.87	3.17
3	1997	10.35	3.19
4	1998	10.18	3.46
5	1999	10.02	4.05
6	2000	8.62	4.59
7	2001	9.22	4.88
8	2002	9.49	5.39
9	2003	8.76	5.85
10	2004	8.4	6.31
11	2005	8.84	6.76
12	2006	8.22	6.98

- Highlight the columns containing the ice cream and yogurt data. Click on the **Chart Wizard** icon in the toolbar. Follow the steps you learned in the first part of this *Inquire* to create a scatter plot. Enter appropriate titles and embed the graph into your spreadsheet. Add a trend line using the **Chart** menu.
7. Compare this scatter plot to the one you created in E-STAT. Which one do you prefer? Justify your answer.



8. Return to E-STAT table 002-0011 on food consumption in Canada.
- a) Retrieve data about the consumption, in litres, of standard (3.25%) milk and partly skimmed (1%) milk between 1996 and 2006.
 - b) Would you expect the consumption of these two products to be related? Justify your answer.
 - c) Graph the data as a scatter plot with a line of best fit. Describe the shape and direction of the correlation. What does the scatter plot tell you about the consumption of these two kinds of milk?

Reflect

- How does creating a scatter plot help you identify and describe trends in data that might not be obvious from looking at a table?
- Choose a data set from Lesson 3.2 or 3.3. Describe how to use a spreadsheet to graph the data set. What would be an advantage of using spreadsheet software to construct the scatter plot and draw the line of best fit?