

Excel Instructions

Life Expectancy of Canadians Linear Activity



Background Information

Life expectancy refers to the average number of years an individual is expected to live, as determined by statistics regarding mortality by age group. For instance, if the life expectancy in 1955 was 65, that would mean that people born in 1955 would be expected to live until they were 65 years old, on average.

The life expectancy of Canadians is among the highest in the world. Why do you think this is?

Part 1: Overall Life Expectancy


In this section, you will retrieve data on the life expectancy of Canadians from 1979 to 1999 from E-STAT. Then, you will import your data into Excel and perform analysis, using your knowledge of linear equations, to see if a linear model provides a good approximation of this data.

Excel Analysis

Retrieving, copying, and pasting the data

- 1) On the **E-STAT Output specification** page, under **Screen output – table:**, select **HTML, table: Time as rows**.
- 2) Click on the **Retrieve now** button.
- 3) Highlight all the data and the attribute names in the table, but not the title at the top of the page or the source line at the bottom of the page.
- 4) **Right-click** and select **Copy**.
- 5) Switch to **Excel**. If Excel isn't already running, you will need to launch it.
- 6) **Right-click** on the first cell (**A1**) and select **Paste**.
- 7) Rename cell **A1** as **Year** and cell **B1** as **Life Expectancy**. You may wish to adjust the column width in order to be able to better read the attribute names.
- 8) Highlight column B. From the **Insert** menu, select **Columns**.
- 9) Type **Years Since 1979** in cell **B1**.
- 10) In cell **B2**, type **=A2 – 1979** and press **Enter**. Your cell should now read 0, indicating there are 0 years from 1979 to 1979.
- 11) **Right-click** on cell **B2** and select **Copy**.
- 12) **Highlight** cells **B3** to **B22**. **Right-click** and select **Paste** to copy the formula from B2. The cells in column B should now read 0, 1, 2, 3, etc. indicating the number of years since 1979. For example, B12 should read 10, indicating there are 10 years from 1979 to 1989.
- 13) Save your Excel file as **Life Expectancy Total**.

Graphing and modelling the data

- 1) **Highlight columns B** (Years Since 1979) and **C** (Life Expectancy) by clicking on them and holding down the **Control** key.
- 2) Click on the **Chart Wizard** icon. 
- 3) On the first page (Chart Wizard – Step 1 of 4 – Chart Type), select **XY (Scatter)** and click on **Next**.
- 4) On the second page (Chart Wizard – Step 2 of 4 – Chart Source Data), click on **Next**.
- 5) On the third page (Chart Wizard – Step 3 of 4 – Chart Options):

- Click on the **Titles tab** and rename the **Chart title** as **Life Expectancy of Canadians**, rename the **Value (X) Axis** as **Years Since 1979**, and rename the **Value (Y) Axis** as **Life Expectancy**.
 - Click on the **Gridlines tab** and check the box **Major gridlines** under **Value (X) Axis**. (Major gridlines should already be checked under Value (Y) Axis by default.) Click on **Next**.
- 6) On the fourth page (Chart Wizard – Step 4 of 4 – Chart Location), click on **Finish**. Your graph should appear on top of your table.
 - 7) **Right-click** on your graph and select **Copy**.
 - 8) Open a word processing document and paste your graph by **right-clicking** and selecting **Paste**. You may wish to make your graph larger before you print, so make any necessary adjustments.
 - 9) Print your graph by selecting **File** and then **Print**.
 - 10) **By hand**, draw in your estimated **line of best fit** for this dataset.

***** Go to Worksheet #1 and answer questions #1 to 7. *****

- 11) Return to **Excel** and **click on any point** on your graph (the points will become blue and yellow).
- 12) Right-click on your graph and select **Add Trendline**.
- 13) **Linear** should be selected as the default. Ensure this is the case and then click on **OK**.
- 14) **Click on the line of best fit** that has appeared on your graph to select it. **Right-click** and select **Format Trendline**.
- 15) Click on the **Options tab**. **Check the box** beside **Display equation on chart** and then click on **OK**. You may wish to re-size your chart in order to better be able to read the equation.
- 16) Open a second word processing document and paste your graph by **right-clicking** and selecting **Paste**. You may wish to make your graph larger before you print, so re-size it if necessary.
- 17) Print your graph by selecting **File** and then **Print**.

***** Go to Worksheet #1 and answer questions #8 to 14. *****

Part 2: Life Expectancy by Sex


In this section, you will retrieve the life expectancy data for males and females separately. You will compare the equations of the lines of best fit of the two sexes. You will see where the two lines of best fit intersect and perform data analysis based on this graph.

Excel Analysis

Retrieving, copying, and pasting the data

- 1) On the **Output specification** page, under **Screen output – table:**, select **HTML, table: Time as rows**.
- 2) Click on the **Retrieve now** button.
- 3) Highlight all the data and the attribute names in the table, but not the title at the top of the page or the source line at the bottom of the page.
- 4) **Right-click** and select **Copy**.
- 5) Switch to **Excel**. If Excel isn't already running, you will need to launch it.
- 6) **Right-click** on the first cell (**A1**) and select **Paste**.
- 7) Rename cell **A1** as **Year**, cell **B1** as **Life Expectancy – Males**, and cell **C1** as **Life Expectancy – Females**. You may wish to adjust the column width in order to be able to better read the attribute names.
- 8) Highlight column B. From the **Insert** menu, select **Columns**.
- 9) Type **Years Since 1979** in cell **B1**.
- 10) In cell **B2**, type **=A2 – 1979** and press **Enter**. Your cell should now read 0, indicating there are 0 years from 1979 to 1979.
- 11) **Right-click** on cell **B2** and select **Copy**.
- 12) **Highlight** cells **B3** to **B22**. **Right-click** and select **Paste** to copy the formula from B2. The cells in column B should now read 0, 1, 2, 3, etc. indicating the number of years since 1979. For example, B12 should read 10, indicating there are 10 years from 1979 to 1989.
- 13) Save your Excel file as **Life Expectancy by Sex**.

Graphing and modelling the data

- 1) **Highlight columns B** (Years Since 1979), **C** (Life Expectancy – Males), and **D** (Life Expectancy – Females) by clicking on them and holding down the **Control** key.
- 2) Click on the **Chart Wizard** icon. 
- 3) On the first page (Chart Wizard – Step 1 of 4 – Chart Type), select **XY (Scatter)** and click on **Next**.
- 4) On the second page (Chart Wizard – Step 2 of 4 – Chart Source Data), click on **Next**.
- 5) On the third page (Chart Wizard – Step 3 of 4 – Chart Options):
 - Click on the **Titles** tab and rename the **Chart title** as **Life Expectancy of Canadians by Sex**, rename the **Value (X) Axis** as **Years Since 1979**, and rename the **Value (Y) Axis** as **Life Expectancy**.
 - Click on the **Gridlines** tab and check the box **Major gridlines** under **Value (X) Axis**. (Major gridlines should already be checked under Value (Y) Axis by default.) Click on **Next**.
- 6) On the fourth page (Chart Wizard – Step 4 of 4 – Chart Location), click on **Finish**. Your graph should appear on top of your table.

***** Go to Worksheet #2 and answer questions #1 to 3. *****

- 7) Return to **Excel** and **click on any point** on the line of best fit for your graph of **Life Expectancy – Females** (the points will become pink and green for the line of best fit for female life expectancy and blue and yellow for the line of best fit for male life expectancy).
- 8) Right-click on your graph and select **Add Trendline**.

- 9) **Linear** should be selected as the default. Ensure this is the case and then click on **OK**.
- 10) **Click on the line of best fit** that has appeared on your graph to select it. **Right-click** and select **Format Trendline**.
- 11) Click on the **Options tab**. **Check the box** beside **Display equation on chart** and then click on **OK**. You may wish to re-size your chart in order to better be able to read the equation.
- 12) Repeat Steps #7 to 11 for the line of best fit for your graph of **Life Expectancy – Males**.

******* Go to Worksheet #2 and answer questions #4 to 6. *******

- 13) Now, extend the lines of best fit into the future to see when the point of intersection of the two lines occurs, by **right-clicking** on the line of best fit for **Life Expectancy – Females** and selecting **Format Trendline**. Under the **Options tab**, under **Forecast Forward**, type in **100 units** and click **OK**. *This projects the line of best fit 100 years into the future beyond the last data point (1999).*
- 14) Repeat Step #13 for the line of best fit for **Life Expectancy – Males**.
- 15) Change the scale of the y-axis so you can better estimate the point of intersection. **Double-click** on the **y-axis values**. Click on the **Scale tab** and change the **Minimum** to **70** and the **Maximum** to **100**. Click on **OK**. *This changes the range of life expectancy values to 70 to 100 years.*
- 16) Change the step size on the x-axis so you can better estimate the point of intersection. **Double-click** on the **x-axis values**. Click on the **Scale tab** and change the **Major Unit** to **10**. Click on **OK**. *This changes the step size of the x-axis values from 50 years to 10 years.*
- 17) Open a third word processing document and paste your graph by **right-clicking** and selecting **Paste**. You may wish to make your graph larger before you print, so re-size it if necessary.
- 18) Print your graph by selecting **File** and then **Print**.

******* Go to Worksheet #2 and answer questions #7 to 9. *******

Student Worksheet #1: Overall Life Expectancy

Life Expectancy of Canadians Linear Activity

Excel Version



- 1) What variable represents x ?
- 2) What variable represents y ?
- 3) What is the y -intercept of your approximated line of best fit (to one decimal place)?
What does the y -intercept represent?
- 4) Calculate the slope of your approximated line of best fit, using the formula $m = \frac{y_2 - y_1}{x_2 - x_1}$.
Be sure to indicate which points you used in your calculations. What does the slope represent?
Note: Use whole numbers for Year and numbers to one decimal place for Life Expectancy.
- 5) Write your equation in two forms:
 - **Life Expectancy = $m(\text{Years Since 1979}) + b$**
 - **$y = mx + b$**
- 6) Rewrite the equation of your approximated line of best fit ($y = mx + b$ form) in standard form ($Ax + By + C = 0$).
- 7) Staple your printout with your approximated line of best fit to this worksheet.

******* RETURN TO YOUR INSTRUCTION SHEET AND COMPLETE STEPS #11 TO 17 BEFORE ANSWERING QUESTIONS #8 TO 14. *******

- 8) Write the equation of the software-generated line of best fit (in $y = mx + b$ form as displayed on the graph).
- 9) How does the software-generated line of best fit compare to your approximated line of best fit? How specifically do the differences in slope and y -intercept relate?
- 10) Why do you think the graph of life expectancy follows the trend it does? Provide at least three possible reasons for this trend.
- 11) Using the equation of the software-generated line of best fit, what does the model predict your life expectancy will be? Be sure to indicate your year of birth and show all calculations. Remember that the x value in your equation is the **number of years since 1979**, not the actual year of your birth.
- 12) Will this equation still be a valid model for someone who is born in 2107? Explain why or why not. Use calculations to support your claim. Remember that the x value in your equation is the **number of years since 1979**, not the actual year of birth.
- 13) If the model is assumed to be correct, in what birth year will Canadians have a life expectancy of 90 years on average? Round your answer to the nearest year.

Remember that the x value in your equation is the **number of years since 1979**, not the actual year of birth.

- 14) Staple a printout of your graph with the software-generated line of best fit to this worksheet.

Student Worksheet #2: Life Expectancy by Sex

Life Expectancy of Canadians Linear Activity

Excel Version



- 1) Who has a higher life expectancy, males or females? Approximately how many years of difference are there between the life expectancies of males and females (over the period from 1979 to 1999)?
- 2) By inspection, which line (life expectancy of males or females) appears to have a higher slope?
- 3) Select two points on each line (click on the points to obtain exact co-ordinates) and estimate the slope using the equation $m = \frac{y_2 - y_1}{x_2 - x_1}$. Show all calculations and be sure to note which points you chose for each line, as an ordered pair. Describe what each ordered pair means. Which line has a higher slope, according to your calculations? Was your guess by inspection correct?

******* RETURN TO YOUR INSTRUCTION SHEET AND COMPLETE STEPS #7 TO 12 BEFORE ANSWERING QUESTIONS #4 TO 6. *******

- 4) For the graph of the life expectancy of males, write the equation of the line of best fit in two ways:
 - Slope-intercept form ($y = mx + b$)
 - Standard form ($Ax + By + C = 0$)
- 5) For the graph of the life expectancy of females, write the equation of the line of best fit in two ways:
 - Slope-intercept form ($y = mx + b$)
 - Standard form ($Ax + By + C = 0$)
- 6) Which line of best fit has a higher slope? What does this higher slope represent?

******* RETURN TO YOUR INSTRUCTION SHEET AND COMPLETE STEPS #13 TO 18 BEFORE ANSWERING QUESTIONS #7 TO 9. *******

- 7) By inspection, approximately where do the lines intersect (as an ordered pair)? What does this represent?
- 8) Substitute the number of years since 1979 value (round to the nearest year) into your lines of best fit equations to confirm the life expectancy at the point of intersection.
- 9) Staple a printout of your graph with the lines of best fit and point of intersection to this worksheet.

Student Worksheet #1

Life Expectancy of Canadians Linear Activity

TEACHER VERSION



- 1) What variable represents x ?

Years Since 1979

- 2) What variable represents y ?

Life Expectancy

- 3) What is the y -intercept of your approximated line of best fit (to one decimal place)?
What does the y -intercept represent?

Answers vary. A y -intercept of 75.2 years was used in this example, which represents a life expectancy of 75.2 years at year zero (1979).

- 4) Calculate the slope of your approximated line of best fit, using the formula $m = \frac{y_2 - y_1}{x_2 - x_1}$.

Be sure to indicate which points you used in your calculations. What does the slope represent?

Note: Use whole numbers for Year and numbers to one decimal place for Life Expectancy.

Answers vary. Note: It is important to select points that are far apart from each other, but close to the line of best fit in order to accurately estimate slope. Points used in this example were (5, 76.2) and (15, 78.0).

$$\begin{aligned} m &= \frac{78.0 - 76.2}{15 - 5} \\ &= \frac{1.8}{10} \\ &= 0.18 \end{aligned}$$

A slope of 0.18 means that for every one year increase in birth year, the life expectancy on average increases by 0.18 years.

- 5) Write your equation in two forms:
- **Life Expectancy = $m(\text{Years Since 1979}) + b$**
 - **$y = mx + b$**

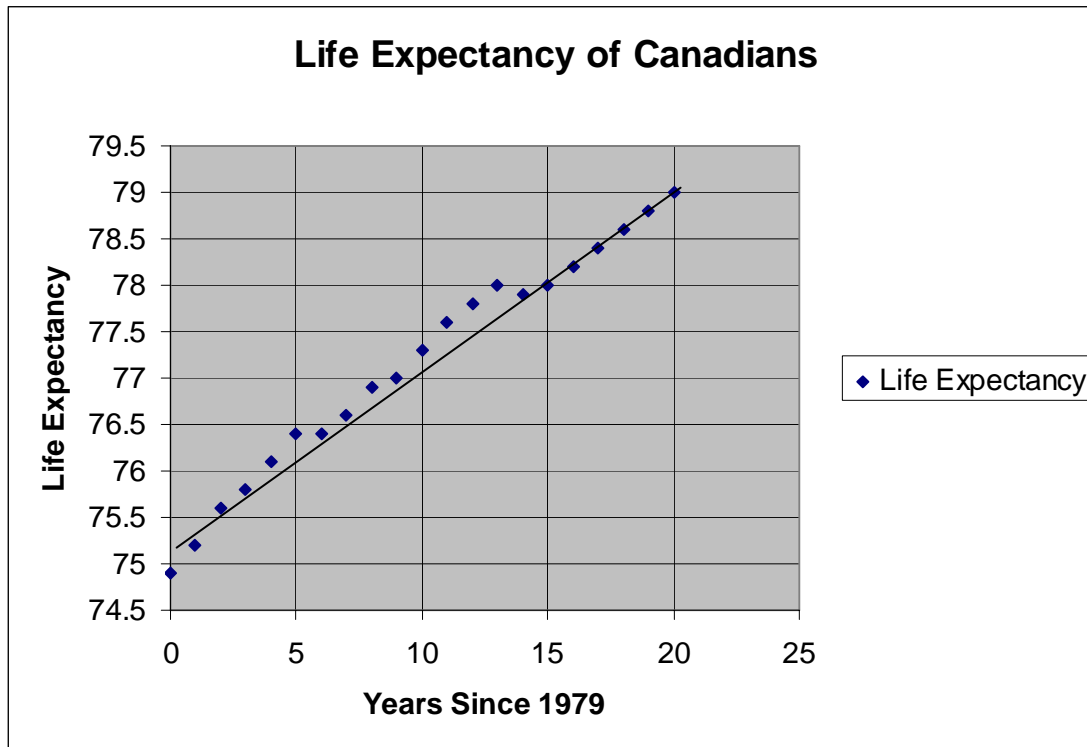
$$\begin{aligned} \text{Life Expectancy} &= 0.18(\text{Years Since 1979}) + 75.2 \\ y &= 0.18x + 75.2 \end{aligned}$$

- 6) Rewrite the equation of your approximated line of best fit ($y = mx + b$ form) in standard form ($Ax + By + C = 0$).

$$\begin{aligned} y &= 0.18x + 75.2 \\ 0 &= 0.18x + 75.2 - y \\ 0.18x - y + 75.2 &= 0 \end{aligned}$$

- 7) Staple your printout with your approximated line of best fit to this worksheet.

Graphs vary. Below is a sample of the Excel graph with a computer-generated line of best fit (using AutoShapes in MS Word), similar to the hand-drawn one used in the above calculations.



- 8) Write the equation of the software-generated line of best fit (in $y = mx + b$ form as displayed on the graph).

$$y = 0.1932x + 75.234$$

- 9) How does the software-generated line of best fit compare to your approximated line of best fit? How specifically do the differences in slope and y-intercept relate?

My line of best fit had a slightly lower slope, which means that I approximated a lesser increase in life expectancy for each year of birth. My y-intercept was slightly lower, which means I approximated a lower life expectancy for the base year (1979). Overall, my approximated line of best fit was a very close approximation.

- 10) Why do you think the graph of life expectancy follows the trend it does? Provide at least three possible reasons for this trend.

Answers vary. Possible answers include better health care (particularly in regard to treatments for heart problems), increased exercise, technological advances, medical discoveries, better diet and lifestyle choices, and decreased infant mortality.

See 'Life Expectancy' article (p. 45 of pdf file) in Health Reports (Vol. 17, No. 1, 2002) at <http://www.statcan.ca/english/freepub/82-003-XIE/0010582-003-XIE.pdf>.

- 11) Using the equation of the software-generated line of best fit, what does the model predict your life expectancy will be? Be sure to indicate your year of birth and show all calculations. Remember that the x value in your equation is the **number of years since 1979**, not the actual year of your birth.

Answers vary. For instance, for a student born in 1992:

$$\begin{aligned}\text{Years Since 1979} &= 1992 - 1979 \\ &= 13\end{aligned}$$

$$\begin{aligned}y &= 0.1932x + 75.234 \\ y &= 0.1932(13) + 75.234 \\ y &= 2.5116 + 75.234 \\ y &= 77.7456\end{aligned}$$

The life expectancy of a student born in 1992 is predicted to be approximately 77.7 years, according to this model.

- 12) Will this equation still be a valid model for someone who is born in 2107? Explain why or why not. Use calculations to support your claim. Remember that the x value in your equation is the **number of years since 1979**, not the actual year of birth.

$$\begin{aligned}\text{Years Since 1979} &= 2107 - 1979 \\ &= 128\end{aligned}$$

$$\begin{aligned}y &= 0.1932x + 75.234 \\ y &= 0.1932(128) + 75.234 \\ y &= 24.7296 + 75.234 \\ y &= 99.9636\end{aligned}$$

The life expectancy for someone who is born in 2107 would be approximately 100 years. This is likely not a valid model because, regardless of improvements in health care, cures for diseases, and other medical improvements that may occur in the next 100 years, it is highly unlikely that the average life expectancy would be 100 years. Human bodies simply would not survive for that long on average, plus there will always be accidents and other accidental causes of death that will lower the life expectancy, regardless of medical improvements.

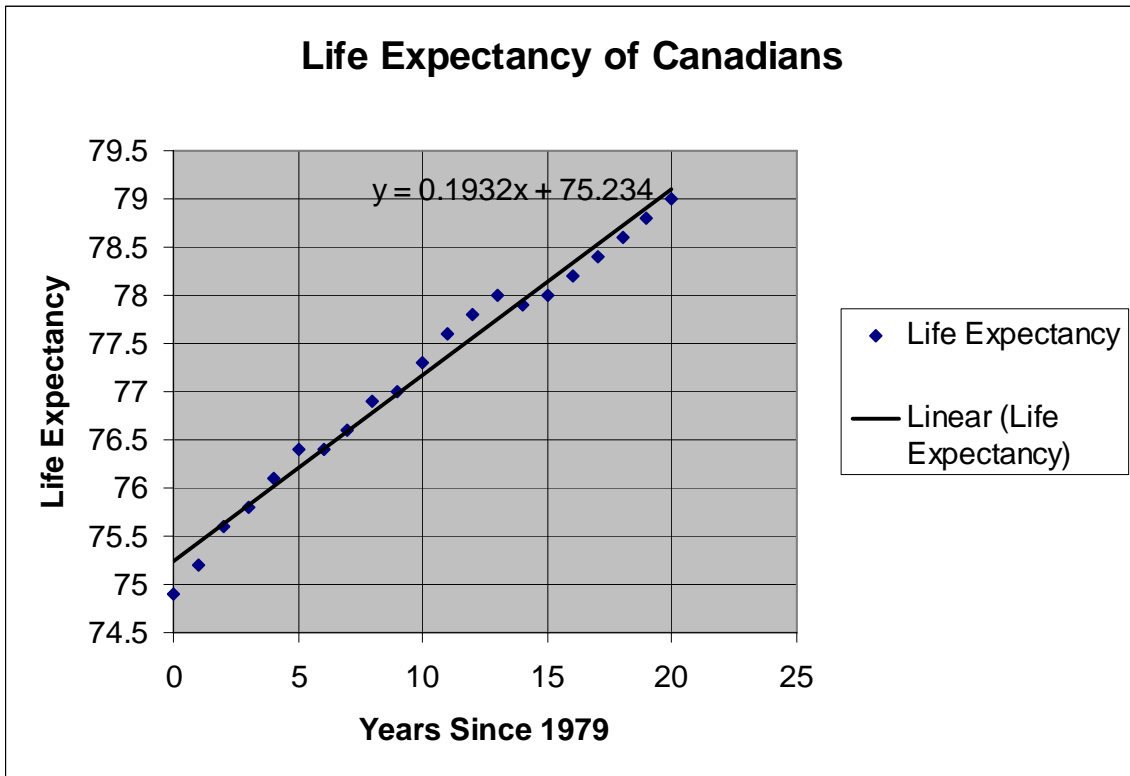
- 13) If the model is assumed to be correct, in what birth year will Canadians have a life expectancy of 90 years on average? Round your answer to the nearest year. Remember that the x value in your equation is the **number of years since 1979**, not the actual year of birth.

$$\begin{aligned}90 &= 0.1932x + 75.234 \\ 14.766 &= 0.1932x \\ x &= 14.766/0.1932 \\ x &\approx 76\end{aligned}$$

$$\begin{aligned}\text{Year} &= 1979 + 76 \\ &= 2055\end{aligned}$$

Canadians born in 2055 are predicted to have a life expectancy of 90 years, according to this model.

- 14) Staple a printout of your graph with the software-generated line of best fit to this worksheet.



Student Worksheet #2: Life Expectancy by Sex

Life Expectancy of Canadians Linear Activity

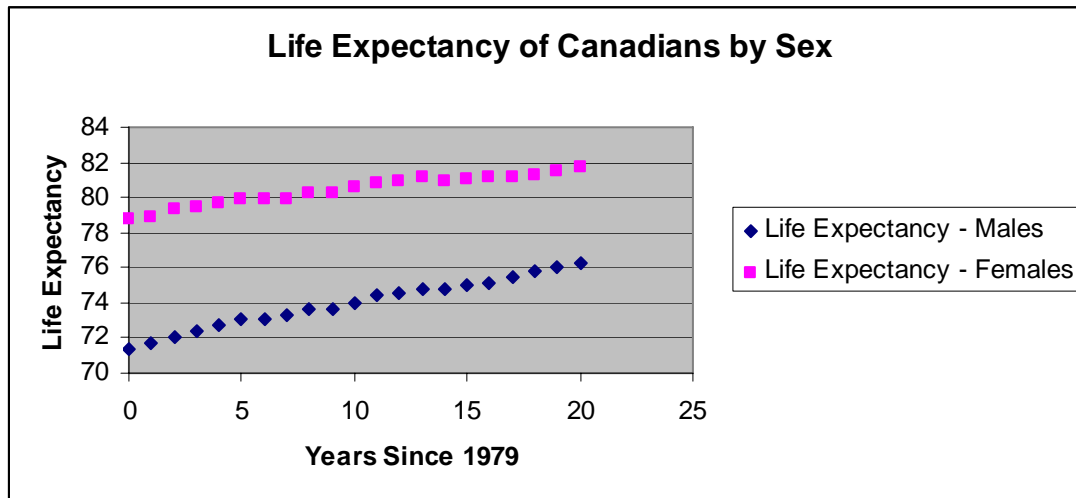
Excel Version



TEACHER VERSION

- 1) Who has a higher life expectancy, males or females? Approximately how many years of difference are there between the life expectancies of males and females (over the period from 1979 to 1999)?

Females have a higher life expectancy by approximately six years. See graph below.



- 2) By inspection, which line (life expectancy of males or females) appears to have a higher slope?

The line of male life expectancy appears to have a higher slope.

- 3) Select two points on each line (click on the points to obtain exact co-ordinates) and estimate the slope using the equation $m = \frac{y_2 - y_1}{x_2 - x_1}$. Show all calculations and be sure to note which points you chose for each line, as an ordered pair. Describe what each ordered pair means. Which line has a higher slope, according to your calculations? Was your guess by inspection correct?

Note: It is important to select points that are far apart from each other, but close to the line of best fit in order to accurately estimate slope.

Males: Points selected were (2, 72.1) and (19, 76). This means that for boys born two years after 1979 (1981), their life expectancy is 72.1 years whereas the life expectancy for boys born 19 years after 1979 (1998) is 76 years.

$$\begin{aligned}
 m &= \frac{76 - 72.1}{19 - 2} \\
 &= \frac{3.9}{17} \\
 &\approx 0.2294
 \end{aligned}$$

Females: Points selected were (3, 79.4) and (18, 81.3). This means that for girls born three years after 1979 (1982), their life expectancy is 79.4 years whereas the life expectancy for girls born 18 years after 1979 (1997) is 81.3 years.

$$\begin{aligned}m &= \frac{81.3 - 79.4}{18 - 3} \\ &= \frac{1.9}{15} \\ &= \sim 0.1267\end{aligned}$$

The slope of the line for males is 0.2294, which is higher than the slope of the line for females (0.1267). My guess by inspection was correct.

- 4) For the graph of the life expectancy of males, write the equation of the line of best fit in two ways:

- Slope-intercept form ($y = mx + b$)
- Standard form ($Ax + By + C = 0$)

$$\begin{aligned}y &= 0.2299x + 71.668 \\ 0.2299x - y + 71.668 &= 0\end{aligned}$$

- 5) For the graph of the life expectancy of females, write the equation of the line of best fit in two ways:

- Slope-intercept form ($y = mx + b$)
- Standard form ($Ax + By + C = 0$)

$$\begin{aligned}y &= 0.1357x + 79.052 \\ 0.1357x - y + 79.052 &= 0\end{aligned}$$

- 6) Which line of best fit has a higher slope? What does this higher slope represent?

The line of best fit for the life expectancy of males has a higher slope. This represents that for each year later that a child is born, the life expectancy will have a greater increase (0.2299 years for males versus 0.1357 years for females).

- 7) By inspection, approximately where do the lines intersect (as an ordered pair)? What does this represent?

The lines appear to intersect at the point (80, 90). This means that for babies born 80 years after 1979 (in the year 2059), life expectancies of both males and females are predicted to be 90 years, according to the linear model.

- 8) Substitute the number of years since 1979 value into your lines of best fit equations to confirm the life expectancy at the point of intersection.

$$\text{Years Since 1979} = 80$$

Males:

$$y = 0.2299x + 71.668$$

$$y = 0.2299(80) + 71.668$$

$$y = 18.392 + 71.668$$

$$y = 90.06$$

Females:

$$y = 0.1357x + 79.052$$

$$y = 0.1357(80) + 79.052$$

$$y = 10.856 + 79.052$$

$$y = 89.908$$

- 9) Staple a printout of your graph with the lines of best fit and point of intersection to this worksheet.

