Prerequisite Review for ECO220Y1Y, 2024/25

Written by Jennifer Murdock; Version: August 27, 2024

Contents

1 Overview of this Required Supplement

This supplement emphasizes prerequisite skills that we use throughout ECO220Y1Y. Because the applications are in contexts relevant to our course, working with this supplement helps you start learning some of the curriculum. In addition, you can obtain a deeper understanding of content you have encountered before and the skill to apply that understanding to real data and authentic contexts. For most people in our course, here is the recommended approach to this supplement:

1. Study Section 3 starting on page 11. Solve the exercises and check your answers against Section 9. Often people struggle in our course when we talk about percents, the magnitude of changes, and conditional statements. Partly this is because of a shaky foundation in prerequisite skills.

2. Study Section 4 starting on page 15, paying particular attention to all subsections having to do with logarithms in Section 4.4 starting on page 18. Our use of logs is challenging for most students. Solve the exercises and check your answers against Section 9. Also, be advised that we have a separate supplement devoted to the use of logarithms in regression analysis, which you study about mid-way through the first term. (Regression analysis allows us to fit a line to data.) This prerequisite review also *previews* some things we will study further and logarithms as used in regression analysis are an excellent example of that.

3. Work through the quiz in Section 2 starting on page 2. You will practice applying your skills in contexts directly relevant to our course. Section 10, beginning on page 36, allows you to grade yourself. It also directs you to specific sections in this supplement (where applicable).

4. Finally, review the remaining sections, and study anything that is not easy for you.

2 Quiz: Review, Refresh, and Update Your Prerequisite Skills

1. If an interest rate is 10 percent and rises by 5 *percentage points* then what is the new rate?

   (a) 10.05
   (b) 10.50
   (c) 10.75
   (d) 12.50
   (e) 15.00

2. If an interest rate is 10 percent and rises by 5 *percent* then what is the new rate?

   (a) 10.05
   (b) 10.50
   (c) 10.75
   (d) 12.50
   (e) 15.00

3. Give a realistic example in an authentic context to illustrate each of these situations.

   (a) A 1 percentage point increase is a large increase.
   (b) A 1 percentage point increase is a small increase.
   (c) A 200 percent increase is a small increase.
   (d) A 2 percent increase is a large increase.

4. A receipt including taxes is \$22.00. If the tax rate is 10%, what is the amount due before taxes?
   (a) 19.60
   (b) 19.80
   (c) 20.00
   (d) 20.20
   (e) 20.40

5. A receipt including taxes is \$X. If the tax rate is r%, what is amount due before taxes?
   (a) X/r
   (b) $X \times (100 - r)$
   (c) $X - (r/100) \times X$
   (d) $\frac{X}{1+r/100}$
   (e) $\frac{X}{1-r/100}$

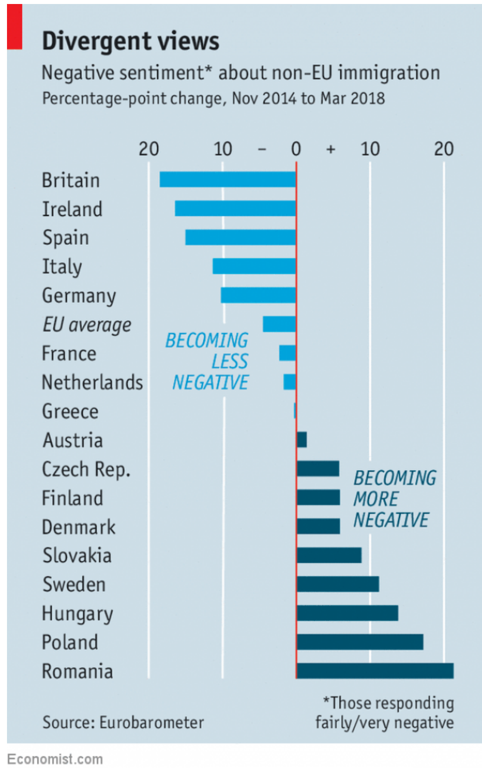
6. On February 21, 2019 the Auditor General released: "Review of Toronto Transit Commission's Revenue Operations: Phase One – Fare Evasion and Fare Inspection." It revealed a shocking amount of fare evasion – people not paying for their rides – on the TTC. The 2018 audit involved a considerable data collection effort: observing TTC use for a random sample of rides and recording whether the rider properly paid. It summarizes the results of a detailed data analysis with both a technical report and a video. Let's consider a part of the report about streetcars. On page 21, Table 5 summarizes the data separately for newer streetcars (image on left below) versus older legacy streetcars (image on right below).


Table 5: Fare Evasion by Type of Vehicles on the Streetcar Routes

	Newer streetcar	Older legacy streetcar
Invalid payments	609	174
Total observations	3,272	2,299
Fare evasion rate	18.61%	7.57%

   (a) How does the evasion rate compare for newer versus older streetcars? Answer with 1 - 2 sentences and make the comparison both ways (percentage point and percent).
   (b) What is the overall fare evasion rate on streetcars (regardless of whether it is an older or newer vehicle type)?
   (c) What is the value 18.61% conditional on?
   (d) What other conditional values of the fare evasion rate would you expect to see in the full report?

7. See the figure from “Xenophobia’s ups and downs: Europeans, on the whole, are becoming more positive about foreigners, but those from the north and the east tend to be less tolerant” in *The Economist*, July 21, 2018 (link).



(a) If, in November 2014, 59% of Germans had negative sentiments about non-EU immigration, then *approximately* what percent had negative sentiments in March 2018?

(b) If, in November 2014, 54% of Poles (i.e. people in Poland) had negative sentiments about non-EU immigration, then *roughly* what percent had negative sentiments in March 2018?

(c) The article states: “In November 2014 Eurobarometer began asking citizens of EU countries about their sentiments towards immigrants. Since then, the overall share of people who have negative feelings about arrivals from outside the bloc has fallen from 57% to 52%.” What is the percent change? What is the percentage point change?

(d) Consider adding major cities to the figure (e.g. Paris). Suppose, in a particular city, 46% had negative sentiments about non-EU immigration in November 2014 but only 33% did in March 2018. This city would be between which two countries in the figure?

8. Consider the statement: “as few as 15 murders.” If x is the number of murders, which is a correct translation into a mathematical statement?

(a) $x > 15$

(b) $x < 15$

(c) $x <= 15$

(d) $x >= 15$

(e) $x >= 16$

9. Consider the statement: "at least half the votes." If x is the number of "yes" votes in a random sample of 10 votes, which is a correct translation into a mathematical statement?

   (a) $x > 1/2$
   (b) $x > 4$
   (c) $x > 5$
   (d) $x <= 5$
   (e) $x >= 6$

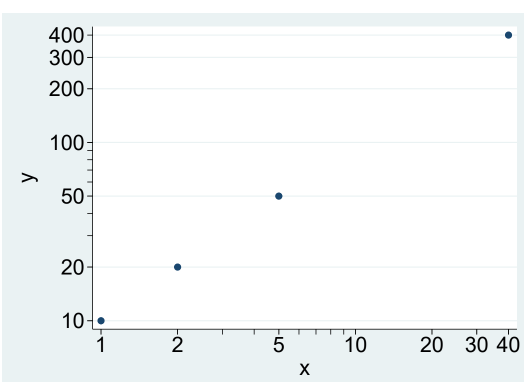
10. Consider the statement: "at most 160 pounds." If x is the allowable weight of a horse-rider in pounds, which is a correct translation into a mathematical statement?

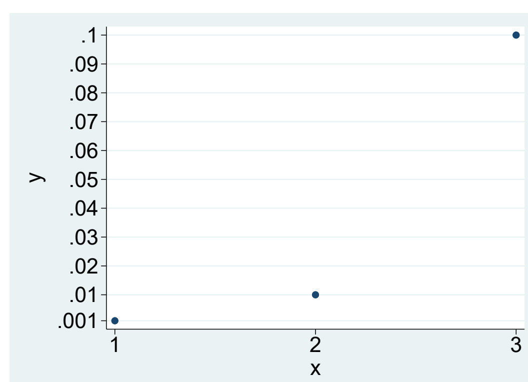
   (a) $x <= 159$
   (b) $x <= 160$
   (c) $x < 161$
   (d) $x >= 160$
   (e) $x >= 161$

11. Assume a, b, c, d are real, non-zero numbers. Assume that n, x, and y are also real numbers and that x and y are non-negative and n is positive.

   (a) Is it true or false that $\frac{1}{c/d} = \frac{d}{c}$?
   (b) Is it true or false that $\frac{c}{a+b} = \frac{c}{a} + \frac{c}{b}$?
   (c) Is it true or false that $\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}$?
   (d) Is it true or false that $\sqrt{x+y} = \sqrt{x} + \sqrt{y}$?
   (e) Is it true or false that $\sqrt{x \times y} = \sqrt{x} \times \sqrt{y}$?
   (f) Is it true or false that $\sqrt{x/n^2 + y/n^2} = \frac{\sqrt{x+y}}{n}$?
   (g) Is it true or false that $a^0 = 1$?

12. The figure on the left uses a log scale for both axes. The figure on the right shows entirely different data and does not use a log scale on either axis.





   (a) Redraw the graph on the left without the log scales on either axis.
   (b) Redraw the graph on the right using a log scale for the y axis.

13. In a 2012 report "The Demand for Disaggregated Food-Away-From-Home and Food-at-Home Products in the United States," the USDA estimates the own-price elasticity of demand for many composite goods.

    (a) The own-price elasticity of dairy is -0.05 (Table 5, p. 18). Interpret -0.05.

    (b) The own-price elasticity of cereals and bakery is -0.58 (Table 5, p. 18). Interpret -0.58.

    (c) The own-price elasticity of white bread is -1.54 (Table 6, p. 20). Interpret -1.54.

14. Consider a 2016 *NBER Working Paper* "Does 'Ban the Box' Help or Hurt Low-Skilled Workers? Statistical Discrimination and Employment Outcomes When Criminal Histories are Hidden."

    ABSTRACT: Jurisdictions across the United States have adopted "ban the box" (BTB) policies preventing employers from conducting criminal background checks until late in the job application process. Their goal is to improve employment outcomes for those with criminal records, with a secondary goal of reducing racial disparities in employment. However, removing information about job applicants' criminal histories could lead employers who don't want to hire ex-offenders to try to guess who the ex-offenders are, and avoid interviewing them. In particular, employers might avoid interviewing young, low-skilled, black and Hispanic men when criminal records are not observable. This would worsen employment outcomes for these already-disadvantaged groups. In this paper, we use variation in the details and timing of state and local BTB policies to test BTB's effects on employment for various demographic groups. We find that BTB policies decrease the probability of being employed by 3.4 percentage points (5.1%) for young, low-skilled black men, and by 2.3 percentage points (2.9%) for young, low-skilled Hispanic men. These findings support the hypothesis that when an applicant's criminal history is unavailable, employers statistically discriminate against demographic groups that are likely to have a criminal record.

    What is the employment rate for young, low-skilled black men?

    (a) 66.7%

    (b) 72.4%

    (c) 77.8%

    (d) 78.1%

    (e) 86.5%

15. Continuing, what is the employment rate for young, low-skilled Hispanic men?

16. When interpreting lines, pick easy-to-understand units of measurement. For example, if y_i is a quiz mark (%) and x_i is seconds spent studying and $y_i = 25.3 + 0.002x_i$, we say each extra *hour* of study is associated with a quiz score that is 7.2 percentage points higher ($= 0.002 \times 60 \times 60$). For each case below, interpret the *slope* by *choosing* the units for the x_i and y_i variables. For now, do not worry about how these lines are obtained and how that affects the interpretation.

    (a) For Canadian workers, $y_i = 43.0283 + 6.1884x_i$ where y_i is annual salary in 1,000s of dollars (e.g. 76.022) and x_i is years of education (e.g. 13). How to interpret 6.1884?

    (b) For people in a sleep study, $y_i = 6.4021 + 0.1562x_i$ where y_i is hours of sleep (e.g. 6.541) and x_i is milligrams of a drug taken at bedtime (e.g. 10). How to interpret 0.1562?

(c) For electric vehicles, $y_i = -34,580 + 330x_i$ where y_i is suggested retail price in dollars and x_i is battery range in miles. How to interpret 330? Assume your readers are most comfortable with kilometers and recall that 1 mile equals 1.60934 kilometers.

(d) For corporate boards of publicly traded companies in North America, $y_i = 0.2452 - 0.0052x_i$ where y_i is the fraction of a board that is female (e.g. 0.33) and x_i is the number of years since the corporation was founded (e.g. 10). How to interpret -0.0052?

(e) For the 500 most popular bottles of wine sold by the LCBO in 2016, $y_i = -64.0105 + 0.9681x_i$ where y_i is the price of a bottle of wine in dollars (e.g. \$16.99) and x_i is its rating (by a wine critic) on a scale of 0 to 100 (e.g. 89). How to interpret 0.9681?

(f) For 100 large retail firms operating in North America, $y_i = 0.2419 + 2.0585x_i$ where y_i is the fraction of employees that leave a company within one year of hire (e.g. 0.28) and x_i is the fraction of employees paid the minimum wage (e.g. 0.15). How to interpret 2.0585?

(g) For colonies of bacteria being exposed to varying intensities of heat in a lab, $y_i = 30,854,882,104 + 1,612,115,885x_i$ where y_i is the number of bacteria that die within five minutes and x_i is the temperature in Fahrenheit (e.g. 110). How to interpret 1,612,115,885? Assume your readers are most comfortable with Celsius and recall $T_{Celsius} = (T_{Fahrenheit} - 32) \times \frac{5}{9}$.

(h) For ten years of annual observations of household shopping trips in the European Union, $y_t = 0.8199 - 0.0104x_t$ where y_t is the share of shopping trips with a total purchase amount less than 10 euro (e.g. 0.76) and x_t is the year (e.g. 2010). How to interpret -0.0104?

(i) For NBA basketball players, $y_i = -65,966,401.2252 + 114,264.0054x_i$ where y_i is salary in dollars and x_i is height in inches. How to interpret 114,264.0054? Assume your readers are most comfortable with centimeters and recall that 1 inch equals 2.54 centimeters.

(j) For students, $y_i = 61.028894 + 0.295204x_i$ where y_i is a student's final exam mark as a percentage (e.g. 80 means 80%, an A-) and x_i measures the number of lectures attended. How to interpret 0.295204? Assume your readers are unfamiliar with the context – there are 48 lectures in the course – and explain how final exam marks relate to a change in the percent of lectures attended (e.g. 24 out of 48 lectures is 50 percent attended whereas 30 out of 48 lectures is 62.5 percent attended).

17. Suppose $\sum_{k=1}^{50} x_k^2 = 258251$ and $\sum_{k=1}^{50} x_k = 3565$. Also, $\sum_{k=1}^{50} y_k = 3469$

(a) Find $\sum_{k=1}^{50}(x_k - \bar{X})^2$.

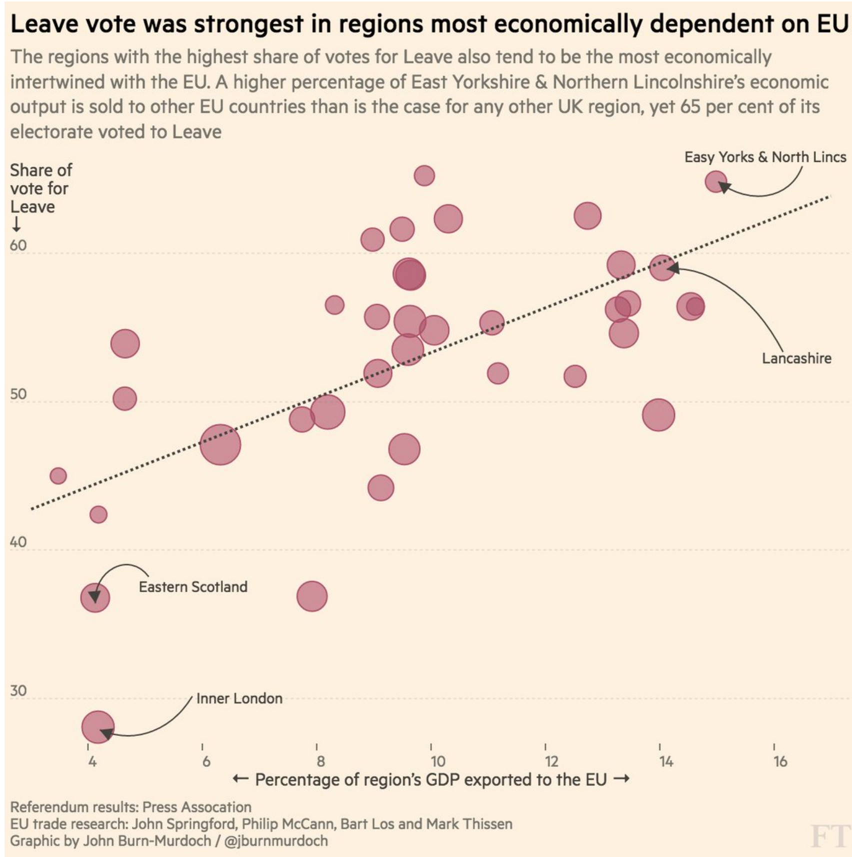
(b) Find $\sum_{k=1}^{50}(2 + 0.4x_k + 0.6y_k)$.

(c) Suppose x_k is in thousands of dollars (i.e. 70 means \$70,000). The variable z_k measures the same thing as x_k except that z_k is measured in dollars. Find $\sum_{k=1}^{50} z_k$ and $\sum_{k=1}^{50} z_k^2$.

18. Consider the figure on the next page posted on Twitter by *The Financial Times* (FT) on June 25, 2016 https://twitter.com/FT/status/746681886131486720.

(a) Approximately, what is the equation of the dashed line? (Define x_i and y_i with units.)

(b) What would the (approximate) equation of the line be if the horizontal axis were the fraction of the region's GDP exported to the EU (instead of the percentage)?



(c) What would the (approximate) equation of the line be if the vertical axis were the fraction voting for "Leave" (instead of the percentage) but the other variable continued to be measured as in the FT figure?

(d) What would the (approximate) equation of the line be if the horizontal axis were the fraction of the region's GDP exported to the EU (instead of the percentage) and the vertical axis were the fraction voting for "Leave" (instead of the percentage)?

19. Recall the figure in Question 18. Consider how the relationship may differ by sex: instead of a single dashed line, two lines (one for female voters and one for male voters). (Such an analysis would use exit polls and not actual voting data, which is anonymous and cannot be analyzed by sex or any other voter-specific information, only by geographic area.) Define a variable $male_i$ to be equal to 1 if the voter is male and zero otherwise. Suppose that male voters are 5 percentage points more likely than female voters to vote "Leave" regardless of export dependency on the EU, which means that the dashed lines for males and females are parallel. Consider this equation $y_i = 35.7 + 1.5x_i + b_2 male_i$ where y_i is the percent voting for "Leave", x_i is the percent of exports to the EU, $male_i$ is equal to 1 if the voter is male and zero otherwise, and b_2 is a constant.

(a) Given the presented facts, what is the implied value of the constant b_2?

(b) What is the meaning of 35.7?

(c) If you used a variable fem_i that is equal to 1 if the voter is female and zero otherwise, then what would be the values of the constants b_0, b_1, and b_2 in $y_i = b_0 + b_1 x_i + b_2 fem_i$?

20. Section 3.3 on page 14 shows research on the flu shot (influenza vaccination). The y_i variable is $shot\_pct\_th\_yr_i$: the percent of people getting a flu shot this year. The x_i variable is $ill\_la\_yr_i$: it equals 1 for those that suffered a flu illness last year and 0 otherwise (i.e. 0 means no flu illness last year). $shot\_la\_yr_i$ equals 1 for those that got a flu shot last year and 0 otherwise (i.e. 0 means no flu shot last year). The **key equation** is:

$$shot\_pct\_th\_yr_i = 13.56 + 13.66 ill\_la\_yr_i + 60.81 shot\_la\_yr_i - 17.99 shot\_la\_yr_i \times ill\_la\_yr_i$$

(a) Draw a Cartesian graph starting at (0,0) and with the y-axis and x-axis as specified next. Label the vertical axis "% getting flu shot this year" and tick values from 0 to 100 by tens. Label the horizontal axis "Suffered a flu illness last year" and tick only 0 and 1.

(b) Using the **key equation**, add a line to the graph for those that *DID NOT GET* a flu shot last year. Label the line. (Hint: For this group the variable $shot\_la\_yr_i$ is 0. Hence, two terms cancel in the key equation.)

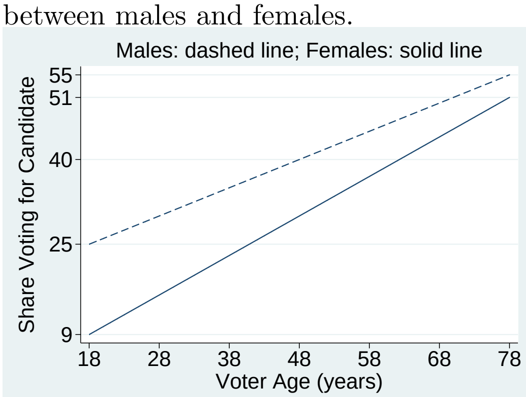
(c) Again, using the **key equation**, add another line to the graph for those that *DID GET* a flu shot last year. Label the line. (Hint: For this group the variable $shot\_la\_yr_i$ is 1. Hence, to find the intercept you need to add two values in the key equation and to find the slope you need to add two values in the key equation.)

(d) What is the *difference* in the intercepts between the two lines you graphed? How does that relate to the **key equation**?

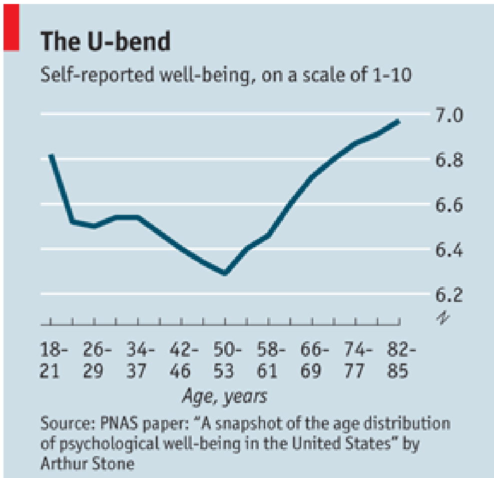
(e) What is the *difference* in the slopes between the two lines you graphed? How does that relate to the **key equation**?

(f) What would the key equation be if $shot\_last\_yr_i$ were replaced with $no\_shot\_la\_yr_i$? The new variable $no\_shot\_la\_yr_i$ is 1 for those that did not get a flu shot last year and 0 otherwise. In other words, what are the values of the coefficients b_0, b_1, b_2, and b_3 in $shot\_pct\_th\_yr_i = b_0 + b_1 ill\_la\_yr_i + b_2 no\_shot\_la\_yr_i + b_3 no\_shot\_la\_yr_i \times ill\_la\_yr_i$?

21. A political candidate's support varies by voters' age and sex: s/he has higher support among older voters and among males. Let the variable x_i be voter age and y_i be the percent intending to vote for the candidate. Define $male_i$ to be 1 for male voters and 0 otherwise. For the figure below, what are the values of the constants a, b, c, and d in $y_i = a + bx_i + cmale_i + d(x_i \times male_i)$? Hint: First, *separately* find the equation for males ($y_i = a_m + b_m x_i$) and females ($y_i = a_f + b_f x_i$). Next, think about how the intercepts *differ* between males and females and how the slopes *differ* between males and females.



22. See the figure from "Age and happiness: The U-bend of life, why, beyond middle age, people get happier as they get older" in *The Economist*, December 16, 2010 (link). A 2012 paper "The mystery of the U-shaped relationship between happiness and age" gives the numeric values below. Define y to be happiness on a ten-point scale (where 10 is happiest and 0 most miserable) and define x to be age in years.



(a) Consider $y = 8.4890 - 0.0417x + 0.0006x^2$ for ages 22 to 80. Draw a fully-labeled graph.

i. What is the slope when age is 25?

ii. What is the slope when age is 40?

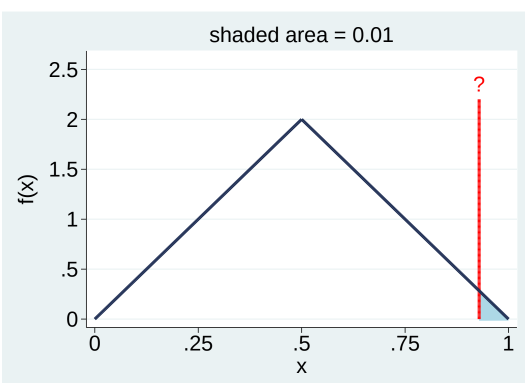
iii. What is the slope when age is 60?

iv. Why are your answers to the previous three parts different?

(b) How would your graph look different if the function were $y = 8.4890 - 0.0617x + 0.0006x^2$?

(c) How would the graph look different if the function were $y = 8.4890 - 0.0417x + 0.00075x^2$ instead of $y = 8.4890 - 0.0417x + 0.0006x^2$?

23. Given that the shaded area in the figure below is equal to 0.01, what is the value of "?"?



3 Percent Changes, Percentage Points, and Conditional Statements

3.1 Percent Changes versus Absolute Changes

To assess if a change is large or not, we often convert to percent changes. For example, suppose a price rose by $1. Is that large or small? It depends on the context. For a new phone a $1 increase is tiny, but for a cup of coffee it's huge. Converting to percent changes sometimes clarifies the magnitude. If a phone is $500, it's a 0.2 percent increase. If a cup of coffee is $2, it's a 50 percent increase.

Further, converting from an absolute change to a percent change yields a unit-free measure, which is often desirable and especially with relatively unfamiliar contexts and units. For example, if you learn that Toronto Hydro raises electricity rates by 1 cent per kilo-watt hour (kWh) is that a big rate increase or not? Most of us would not be familiar enough with those units and context to assess the magnitude of that rate increase. Once you learn that effective June 1, 2020 the electricity rate is 12.8 cents per kWh[1] we realize that a 1 cent increase is considerable: it would correspond to a 7.8 percent jump in households' electricity bills.

3.1.1 Misleading Headline: "Air Travel Surges by 123%!"

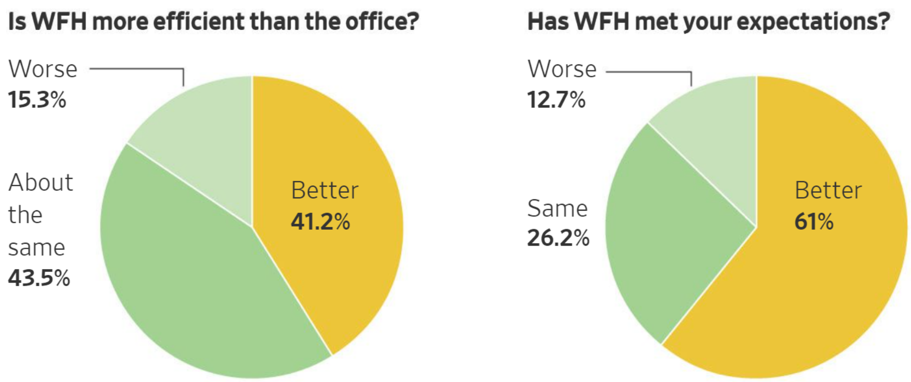
Converting to percent changes does *not* always clarify the size of change. A May 19, 2020 article in *The New York Times* titled "Air Travel Surges by 123%! (Beware of Misleading Data Like That)" (https://www.nytimes.com/2020/05/19/upshot/virus-economic-data-upended.html) gives a great example.

> **Excerpt:** In normal times, percentage change is a helpful guide to what's happening in the economy. But these are not normal times.
>
> Did you hear about the booming air travel industry? It's up 123 percent in just the last month!
>
> Technically, that's an accurate number. Over the seven days ended Sunday, an average of 212,580 people went through U.S. airport security checkpoints, up from 95,161 in the week ended April 17.
>
> But of course, that is all wrong if you know anything about the underlying reality of the air travel industry. This time a year ago, 2.4 million people a day went through those same checkpoints. By any reasonable measure, these remain disastrous times for air traffic. It's just that the shutdown in March and early April made even the slight recovery that has taken place seem like an enormous surge in percentage terms.
>
> Get ready for the same effect to apply to all sorts of numbers – most notably with economic data. These swings are artifacts of the arithmetic of percentage change. But if you aren't attuned to the yo-yo effect that we are likely to see in crucial data in the coming months, you could get a misleading impression of where the United States stands.
>
> In particular, as some of the business activity shut down by the pandemic begins to come back, economic data could create the impression of a soaring economy in the summer and fall – even though it is the equivalent of those air travel numbers.

[1] Retrieved from the Toronto Hydro page titled "Residential electricity rates: Understanding current rates and how they're set" on July 29, 2020, https://www.torontohydro.com/for-home/rates.

3.1.2 Rounding and Percents Not Summing to 100%

When percents are rounded off, the sum over an exhaustive and mutually exclusive set will not necessarily be 100 percent, even though it should be in theory. A January 2, 2021 article in *The Wall Street Journal* titled "Is a Home Office Actually More Productive? Some Workers Think So" (https://www.wsj.com/articles/is-a-home-office-actually-more-productive-some-workers-think-so-11609563632) gives a great example. The acronym "WFH" means work from home (i.e. remote work).



Note: Data might not equal 100% due to rounding.
Source: Becker Friedman Institute for Economics at the University of Chicago online survey of 10,000 employees in August-November 2020 (efficiency); 15,000 employees in May- October 2020 (expectations)

In the pie chart on the left, the sum of the percents, which are each rounded to the nearest first decimal place, is 100 ($(43.5 + 15.3 + 41.2) = 100$). However, in the pie chart on the right, the sum is 99.9 ($(26.2+12.7+61.0) = 99.9$). While you may think that maybe 0.1% of people skipped answering the question, that would be an incorrect inference. The note below the figure makes it clear that any discrepancy between the sum and 100% is due to rounding. They rounded to the nearest first decimal place. Further, they are reporting the percents *conditional* on answering the multiple-choice question, which means everyone gave one of those three answers: same, worse, or better. Those answers are exhaustive: survey repondents could choose among only those three options (no chance to say other). Those answers are also mutually exclusive: survey respondents could pick one and only one answer (no chance to select two or three of the choices). So, how can it be that they don't sum to 100%? The note below says the pie chart to the right is based on a sample of 15,000 employees. Suppose that 14,791 answered the question on expectations and that 9,028 said "better," 3,882 said "same," and 1,881 said "worse" ($9,028+3,882+1,881 = 14,791$). That means that 61.037117% said "better" ($= 100 * 9,028/14,791$), which is 61.0% to the nearest first decimal place with standard rounding. Further, 26.245690% said "same" ($100*3,882/14,791$), which rounds to 26.2%, and 12.717193% said "worse" ($100 * 1,881/14,791$), which rounds to 12.7%. In general, with standard rounding to the nearest first decimal place, it is not uncommon that the sum is off by 0.1 percentage points (i.e. sum could be 99.9 or 100.1) and it can be off by more. Similarly, with standard rounding to the nearest integer, it is not uncommon that the sum is off by plus or minus 1 percentage point (i.e. the sum could be 99 or 101) and it can be off by more.[2]

[2] To illustrate how the sum could be off by more than 1 percentage point with standard rounding to the nearest integer, suppose we had 10,000 survey responses to a multiple choice question with five options from strongly agree (5) to strongly disagree (1). If 4,549 answered (5), 1,649 answered (4), 1,749 answered (3), 1,849 answered (2), and 204 answered (1), that would round to 45%, 16%, 17%, 18%, and 2%, which sums to only 98%.

3.2 Percent Changes versus Percentage Point Changes: Watch Out!

Is there a difference between a *percent change* and a *percentage point change*? Yes! Percent changes make sense regardless of the units of measurement. For example, if City A had 26 smog-alert days in 2013 and 32 in 2014, then it is up by 6 days, which is a 23 percent increase ($= 100 \times (32 - 26)/26$). However, percentage point changes only make sense for percents: percentage points are the units of measurement for a variable measured as a percent. For example, if City A had an unemployment rate of 7.2 percent in 2013 and 8.1 percent in 2014, then unemployment is up by 0.9 percentage points.

Comparisons involve various levels of conditional statements. We ask about the smog-alert days and the unemployment rate given that the year is 2013 and the location is City A (i.e. for 2013 in City A). These can be compared with 2014. The phrases "given that" or "for" indicate the numbers are conditional on the year and the location.

Consider this sentence: "Lottery winners are 14.1 percentage points more likely to enroll in college the fall after their senior year, a 49.0 percent increase from the control mean of 28.8 percent."[3] What does this mean exactly? 28.8 percent of lottery losers (the control group) enroll in college while 42.9 percent of lottery winners (the treatment group) enroll. The key conditioning variable is whether or not the student is in the control group (lottery losers) or the treatment group (lottery winners). There are two different ways to describe the (large) difference. You may say that the lottery winners are 14.1 *percentage points* more likely to enroll ($= 42.9 - 28.8$). Alternatively, you may say that the lottery winners are 49.0 *percent* more likely to enroll ($= 100 \times (42.9 - 28.8)/28.8$). Obviously 14.1 and 49.0 differ: percent and percentage point mean *different* things.

There is no rule about the order you make comparisons. For example, rather than say that lottery winners are 14.1 percentage points *more* likely to enroll, you could say that lottery losers are 14.1 percentage points *less* likely to enroll ($= 28.8 - 42.9$). However, be careful with *percent* changes if you change the base of comparison. You may say that the lottery losers are 32.9 *percent* less likely to enroll ($= 100 \times (28.8 - 42.9)/42.9$). While the percentage point difference stays the same magnitude, the relative difference depends on what is the base of comparison: whether the enrolment rate of the lottery winners or losers is in the denominator. In other words, "lottery winners are *49.0 percent* more likely to enrol" compared to "lottery losers are *32.9 percent* less likely to enrol." As always, remember that you must *interpret* the sign of the difference. We say lottery losers are 32.9 percent *less* likely to enrol, *not* that the difference is -32.9 percent.

Returning to the original excerpt, suppose the sentence were shorter: "Lottery winners are 14.1 percentage points more likely to enroll in college the fall after their senior year, a 49.0 percent increase." Can you figure out what percent of the control group (lottery losers) enrolled in college? Yes. Define x to be the percent of the control group that enroll in college the fall after their senior year. We know that $\frac{x+14.1}{x} = (1 + 0.49)$. We can solve to obtain $x = 28.8$ percent. Hence, in the original sentence, the last phrase ("from the control mean of 28.8 percent") is redundant. However, given that the authors are interpreting the results, it is a great idea to include that extra phrase to

[3]This appears on page 16 of a 2013 *NBER Working Paper* "The Medium-Term Impacts of High-Achieving Charter Schools on Non-Test Score Outcomes" by Will Dobbie and Roland Fryer http://www.nber.org/papers/w19581.

make sure the reader understands.

3.3 Jin and Koch (2018) "Learning by Suffering? Patterns in Flu Shot Take-up"

Economists and other experts in data analysis often break up data into subgroups to examine whether the results are different or similar. For example, you may have data on whether or not individuals got their flu shots. You can break the data up into various groups (e.g. by income, by race, by sex, etc.) and examine how the rates vary across groups. However, often researchers cut data even finer: sub-subgroups. For example, consider the November 2018 *NBER Working Paper* "Learning by Suffering? Patterns in Flu Shot Take-up" (link) by Ginger Zhe Jin and Thomas G. Koch.

Consider ONLY the subset of people who did *not* get the flu shot last year. The flu shot rate this year is 27.22% among those that suffered a flu illness last year: these people may remember paying for not getting a flu shot last year and decide to get a flu shot this year. 27.22% is conditional on: (1) not getting the flu shot last year and (2) suffering a flu illness last year. In contrast, the flu shot rate this year is 13.56% among those that did not suffer a flu illness last year: these people remember getting away with no shot last year. 13.56% is conditional on: (1) not getting the flu shot last year and (2) not suffering a flu illness last year. (Notice that (1) is identical for both percents.)

Why analyze sub-subgroups? It shows evidence of learning (to get the flu shot) by suffering from a flu illness when you skipped the flu shot. What if we did not break up the subgroup of people who did not get the shot last year into two sub-subgroups (those that suffered a flu illness last year and those that did not suffer a flu illness last year)? Imagine there are 1,421 people who did not get the flu shot last year and suffered the flu illness last year and 3,002 people who did not get the flu shot last year and did not suffer a flu illness last year. What is the overall flu shot rate this year among those that did not get a flu shot last year? It is 17.95% $\left(= 100\frac{(0.2722\times 1{,}421+0.1356\times 3{,}002)}{(1{,}421+3{,}002)}\right)$. However, by combining the sub-subgroups we hide an interesting difference between them.

How do Jin and Koch (2018) concisely report 27.22% and 13.56%? The table below is cut-and-paste from page 27 of that paper. The subscript t means this year. The subscript $t-1$ means last year. The four numbers refer to the flu shot rate *this year* as per the title Flu Shot Rate$_t$. The table reports that rate under four possible conditions for *last year*, which are about whether or not you had the flu *illness* last year and whether or not you had the flu *shot* last year.

	Flu Shot Rate$_t$	
	Had Shot$_{t-1}$	Did Not$_{t-1}$
Had Flu$_{t-1}$	70.04	27.22
Did Not$_{t-1}$	74.37	13.56

Section 3 Exercises:

1. Suppose that the interest rate is currently 2.76%.
   (a) What would be the new interest rate if the bank increased it by 1 percentage point?
   (b) What would be the new interest rate if the bank increased it by 1 percent?
2. Suppose that among 502 males, 162 got the flu shot. Among 777 females, 289 got the flu shot. What is the overall flu shot rate (regardless of sex)?

3. Recall Jin and Koch (2018). Consider ONLY the subset of people who *got* the flu shot last year. The flu shot rate this year is 70.04% among those that suffered a flu illness last year: these people may remember the flu shot not working for them (but also may be in the habit of getting a flu shot). In contrast, the flu shot rate this year is 74.37% among those that did not suffer a flu illness last year: these people may remember the flu shot working for them.

   (a) What is the value 70.04% conditional on? What is the value 74.37% conditional on?

   (b) What if we did not break up the subgroup of people who got the shot last year into two sub-subgroups (those that suffered a flu illness last year and those that did not)? Imagine there are 4,555 people who got the flu shot last year and suffered the flu illness last year and 46,387 people who got the flu shot last year and did not suffer a flu illness last year. What is the overall flu shot rate this year among those that got a flu shot last year?

4 Functions

A function f assigns to each element x in a set A exactly one element, called $f(x)$, in a set B. For example, $f(x) = 2x + 1$ (or $y = 2x + 1$). This section reviews linear and nonlinear functions. While it starts out quite basic, keep working all the way through Section 4.4, which gives important insights about logs and explores case studies.

4.1 Linear Functions

A line can be written in terms of the y-intercept, a, and the slope, b, where a and b are constants.

$$y = a + bx \tag{1}$$

The intercept, a, is the value of y when x is equal to zero. The slope, b, can be found by using any two points, (x_1, y_1) and (x_2, y_2), that a line passes through.

$$b = \frac{y_2 - y_1}{x_2 - x_1} \tag{2}$$

One way to find the equation of a line is to use the *point-slope formula*, which requires that you know the slope of the line, b, and one point the line passes through (x_1, y_1):

$$y - y_1 = b(x - x_1) \tag{3}$$

Section 4.1 Exercises:

1. Consider a line that passes through the points (1,2) and (2,2).

   (a) What is the equation of the line? (Write it in the same format as Equation 1.)

   (b) According to the line, what happens to the value of y when x increases from 1 to 100?

2. Consider a line that passes through the points (1,2) and (2,0).

   (a) What is the equation of the line? (Write it in the same format as Equation 1.)

   (b) According to the line, what happens to the value of y when x increases from 1 to 2?

3. Consider the line $y = 10 - x$. All of the question subparts are relative to this "original" line.
   (a) Imagine a sketch of the line. What is the equation of line parallel to the original but two units higher everywhere?
   (b) What is the equation of a line parallel to the original but three units lower everywhere?
   (c) What is the equation of a line that has a slope twice as steep but the same y-intercept?
   (d) What is the equation of a line that has a slope twice as steep but the same x-intercept?

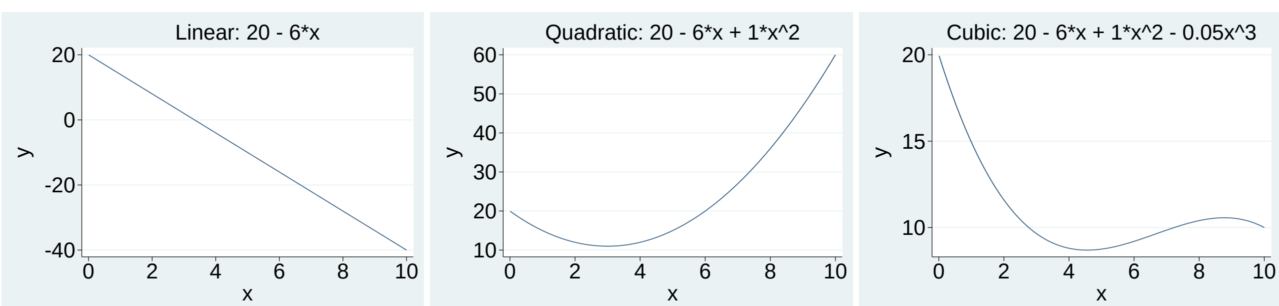
4.2 Power Functions and Polynomials

One type of nonlinear function is power functions: $f(x) = x^a$, where a is a constant. Recall the rules:

- $f(x) = x^a x^b = x^{a+b}$
- $f(x) = (x^a)^b = x^{ab}$
- $f(x) = \frac{x^a}{x^b} = x^{a-b}$
- $f(x) = x^{1/a} = \sqrt[a]{x}$ (which implies $f(x) = x^{0.5} = \sqrt{x}$)
- $f(x) = x^0 = 1$
- $f(x) = \frac{1}{x^a} = x^{-a}$ (which implies $f(x) = \frac{1}{x} = x^{-1}$)

A polynomial function is $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + ... + a_{n-1}x^{n-1} + a_nx^n$, where a_i for $i = 1, 2, 3, ..., n$ are $n+1$ constants. The numbers $a_0, a_1, ..., a_n$ are the coefficients and n is the degree. For example, if $n = 3$ and $a_i = 2$ for all $i = 0, 1, 2, 3$ then we'd have $f(x) = 2 + 2x + 2x^2 + 2x^3$, which is a third degree polynomial (cubic function). If $n = 2$ then it's a second degree polynomial (quadratic function, parabola). If $n = 1$ then it's a first degree polynomial (line).

Higher degree polynomials allow more twists and turns: they are more *flexible*. When researchers say *flexible functional form*, it may mean including terms in a model (such a cube) to allow it to fit the data. However, researchers also value *parsimonious* models: the most simple one possible.



For a quadratic function, recall the quadratic formula that allows you solve quadratic equations (even if you cannot factor): if $f(x) = ax^2 + bx + c$, then $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$.

Section 4.2 Exercises:

1. Consider $w = 23.1 + 2.4z + 5.9z^2 - 6.2z^3$. What is the degree of this polynomial? The coefficients?
2. Solve $5x^2 + 3x - 3 = 0$.

4.3 Exponential and Logarithmic Functions

An exponential function is $f(x) = a^x$, where a is a positive constant, such as $f(x) = e^x$, $e \approx 2.71828$.

- If $x = n$, n a positive integer, then $a^n = a \times a \times ... \times a$ and if $x = 0$ then $a^0 = 1$
- If $x = -n$, n a positive integer, then $a^{-n} = \frac{1}{a^n}$
- If x is rational, $x = m/n$, where m and n are integers and $n > 0$, then $a^x = a^{m/n} = \sqrt[n]{a^m}$

Laws of Exponents:

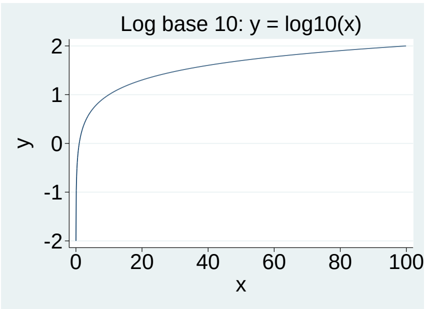
- $a^{x+y} = a^x a^y$ and $a^{x-y} = \frac{a^x}{a^y}$
- $(a^x)^y = a^{xy}$ and $(ab)^x = a^x b^x$

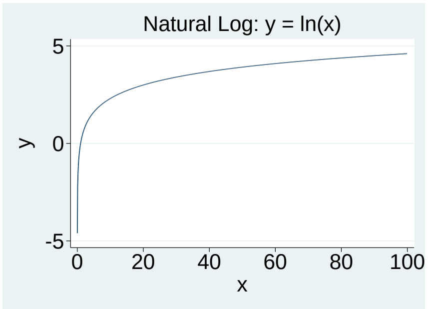
The inverse of the exponential function $f(x) = a^x$ is the logarithmic function to the base a. (The symbol $\Leftrightarrow$ is read "is equivalent to.")

$$log_a x = y \Leftrightarrow a^y = x$$

The logarithmic function, log_a, has a domain (possible inputs: x values) of $(0, \infty)$. The logarithm of 0 or a negative number does not exist. Its range (possible outputs: y values) is all real numbers.

A logarithm with base e is the natural logarithm, $log_e(x) = ln(x)$, where e is ≈ 2.71828. Among economists, "logarithm" means "natural log." It is common to write $log()$ even for a natural log. Popular software reflects this convention. In STATA both the function log and ln return the natural log: only log10 returns the base 10 logarithm. Excel is an exception: log means a base 10 logarithm.





Rules for Logarithms and Natural Logarithms:

- For all x: $log_a(a^x) = x$
- For $x > 0$: $a^{log_a x} = x$
- $log_a(xy) = log_a(x) + log_a(y)$
- $log_a\left(\frac{x}{y}\right) = log_a(x) - log_a(y)$
- $log_a(x^y) = y log_a(x)$
- $ln(e) = 1$ and $ln(1) = 0$
- For all x: $ln(e^x) = x$
- For $x > 0$: $e^{ln(x)} = x$

4.4 What's So Special About Logs?

Economists use logs extensively. Since COVID-19, logs have even become common in the popular press (e.g. *The New York Times*, *The Wall Street Journal*, *The Economist*, *The Toronto Star*, and so on). Logs can transform non-linear relationships into linear ones, rein in outliers, remove skew from histograms, and address compound growth. We start with a quick refresher of basics in Section 4.4.1 and an investment example in Section 4.4.2. Section 4.4.3 reviews the use of logs in figures. Section 4.4.4 reviews linear and logarithmic functions more thoroughly within a COVID-19 case study.

4.4.1 Basics of Logarithms: A Look at the Numbers

To recall the mechanics of logarithms, see this table and answer the questions below to explore it.

(1)	(2)	(3)		(4)	(5)	(6)
x	$log(x)$	$ln(x)$		x	$log(x)$	$ln(x)$
0.01 $\left(= 10^{-2}\right)$	-2	-4.605170186		0.135335283 $\left(= e^{-2}\right)$	-0.868588964	-2
0.1 $\left(= 10^{-1}\right)$	-1	-2.302585093		0.367879441 $\left(= e^{-1}\right)$	-0.434294482	-1
1 $\left(= 10^{0}\right)$	0	0		1 $\left(= e^{0}\right)$	0	0
10 $\left(= 10^{1}\right)$	1	2.302585093		2.718281828 $\left(= e^{1}\right)$	0.434294482	1
100 $\left(= 10^{2}\right)$	2	4.605170186		7.389056099 $\left(= e^{2}\right)$	0.868588964	2
1,000 $\left(= 10^{3}\right)$	3	6.907755279		20.08553692 $\left(= e^{3}\right)$	1.302883446	3
10,000 $\left(= 10^{4}\right)$	4	9.210340372		54.59815003 $\left(= e^{4}\right)$	1.737177928	4
100,000 $\left(= 10^{5}\right)$	5	11.51292546		148.4131591 $\left(= e^{5}\right)$	2.17147241	5

- Noticing the pattern above, and without using a calculator, the value of $log(23,521)$ would fall between which two integers? The value of $ln(47.5)$ would fall between which two integers?
- For Column (1), are the values increasing in a linear or exponential manner? Column (2)? Column (3)? (For (3), check if the slope – the change from one row to the next – is constant.)
- For Column (4), are the values increasing in a linear or exponential manner? Column (6)? Column (5)? (For (5), check if the slope – the change from one row to the next – is constant.)

4.4.2 Natural Logs and Growth Rates

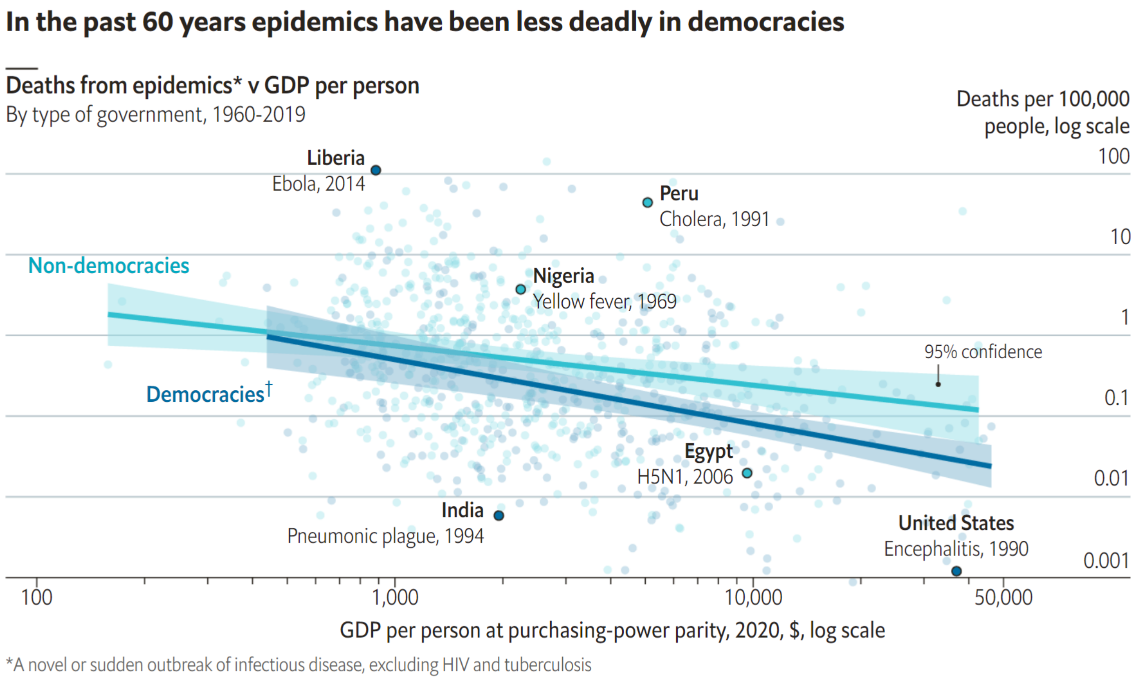
An initial investment of $2,000 with a constant annual return of 9 percent ($r = 0.09$) yields $4,734.73 at the end of 10 years ($4,734.73 = 2,000 * (1 + 0.09)^{10}$). For each year t, $V(t) = 2,000 * (1 + 0.09)^{t}$. Take the natural log of both sides: $ln(V(t)) = ln(2,000) + t * ln(1 + 0.09)$. $ln(1 + 0.09) = 0.086177696$.

Time $= t$	Amount $= V(t)$	Change	ln(Amount) $= ln(V(t))$	Change
0	$2,000.00		7.60090246	
1	$2,180.00	$180.00	7.687080156	0.086177696
2	$2,376.20	$196.20	7.773257852	0.086177696
3	$2,590.06	$213.86	7.859435548	0.086177696
4	$2,823.16	$233.11	7.945613245	0.086177696
5	$3,077.25	$254.08	8.031790941	0.086177696
6	$3,354.20	$276.95	8.117968637	0.086177696
7	$3,656.08	$301.88	8.204146333	0.086177696
8	$3,985.13	$329.05	8.290324029	0.086177696
9	$4,343.79	$358.66	8.376501726	0.086177696
10	$4,734.73	$390.94	8.462679422	0.086177696

Because of compound interest, the dollar change increases at an increasing rate. However, the natural log of the total amount increases at a constant rate: $ln(V(t)) = 7.60090246 + 0.086177696 * t$. In fact, that change is an excellent approximation of the growth rate (0.086177696 rounds to 0.09). Economists *love* that $ln(1 + r) \approx r$ for smallish values of r (like $r = 0.09$).

4.4.3 Log Base 10 in Figures

See the figure from "Democracies contain epidemics most effectively: People living under freely elected governments have been more responsive to lockdown measures" in *The Economist*, June 6, 2020 (https://www.economist.com/graphic-detail/2020/06/06/democracies-contain-epidemics-most-effectively). Is gives an ideal opportunity to review the basics of logs and preview other course concepts, including scatter plots, panel data, confidence intervals, and regression analysis.



While economists nearly always use natural logs in analysis, one exception is in visually displaying data where a log base 10 is used. Why? As you saw in Section 4.4.1, if $log_{10}(x)$ is 2 then x is 100 and if $log_{10}(x)$ is 3 then x is 1,000, and so on. Hence, it is intuitive to set the log scale of 2, 3, 4, ... (equally spaced) where 2 corresponds to 100, 3 corresponds to 1,000, and 4 corresponds to 10,000, like the x-axis above. It starts at $log_{10}(x) = 2$ – not at $log_{10}(1) = 0$ or $log_{10}(10) = 1$ – because no country had GDP per capita less than \$100. Figures focus on the region where the data lie, and only include $(0, 0)$ if that is within the range of the data.

The y-axis is also a log scale and shows deaths per 100,000 of population, not a count of deaths. Making it relative to population allows comparing big and small countries. However, why per 100,000? Why not per million? Or per 1,000? That choice is dictated by how rare something is: the more rare, the bigger the per number. For example, about 9 to 10 people die of a lightning strike in Canada each year[4] and in 2020 the Canadian population is about 38 million people. Hence, we would say

[4] Retrieved from the page "Lightning in Canada: frequently asked questions" on August 1, 2020 from https://www.canada.ca/en/environment-climate-change/services/lightning/frequently-asked-questions.html.

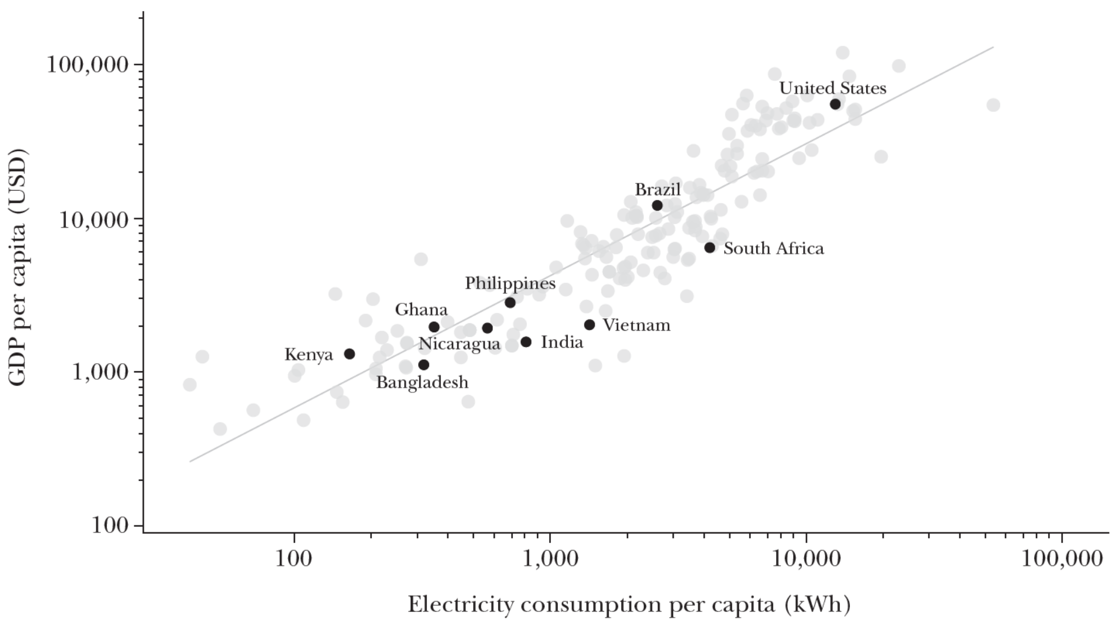
deaths by lightning strikes in Canada is about 2.5 deaths per 10 million people. Death by lightning strike is very unlikely so we make it per 10 million people. Unfortunately, death by epidemic is far more likely so it is per 100,000 in the figure. Again notice the log scale on the y-axis: the tics 0.001, 0.01, 0.1, 1, 10, and 100 are equidistant whereas they are NOT equidistant on a real number line.

But why does the death by epidemics figure use a log scale on both the x-axis and the y-axis? The figure would look terrible if it did not. A few countries have GDP per capita that is *one hundred times* bigger than other countries (e.g. \$500 versus \$50,000), which means there would be lots of poorer countries crowded together with a lot of white space between the richer ones. Similarly for the y-axis.

To illustrate how logs can make a figure more readable, consider Figure 1 from a 2020 article "Does Household Electrification Supercharge Economic Development?" in the *Journal of Economic Perspectives* (*JEP*) (https://doi.org/10.1257/jep.34.1.122).

Figure 1

The Positive Correlation between Electricity Consumption and GDP per Capita



Source: 2014 data obtained from the World Bank DataBank.
Note: Both variables are presented on a logarithmic scale. GDP per capita data are in current US dollars.

Section 4.4.3 Exercises:

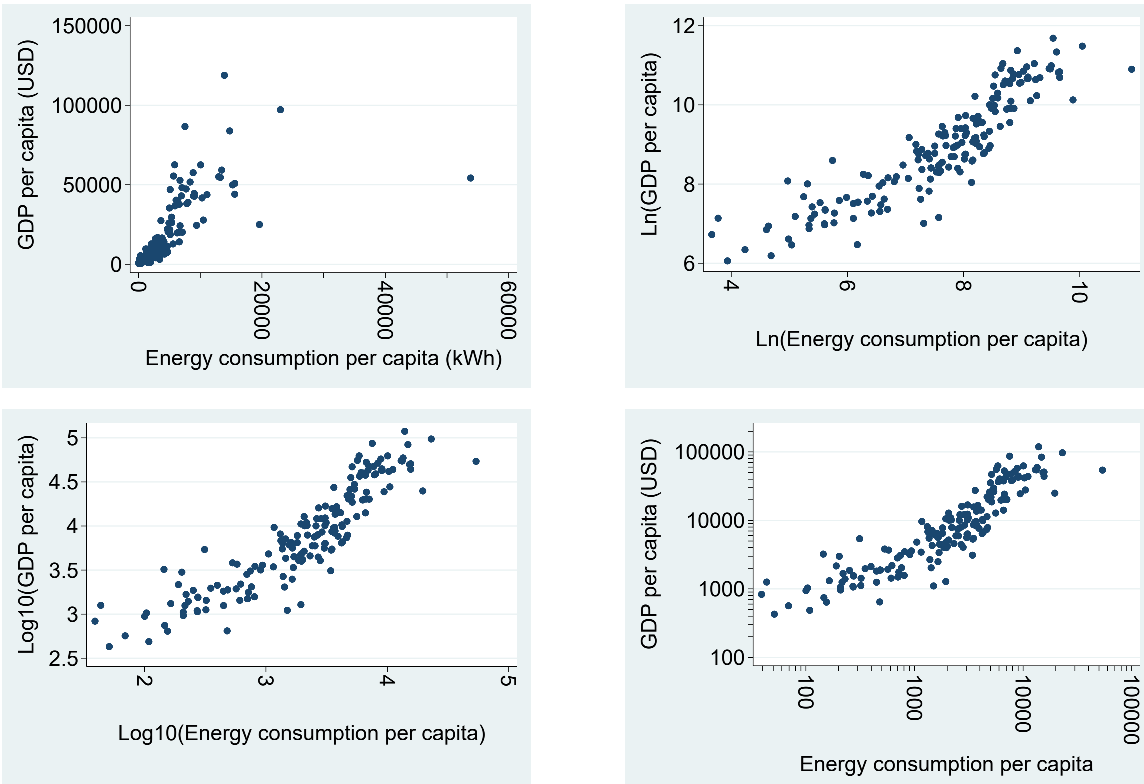
1. Given Figure 1, *approximately* what is the electricity consumption per capita and *approximately* what is the GDP per capita for:

    (a) Bangladesh?

    (b) Brazil?

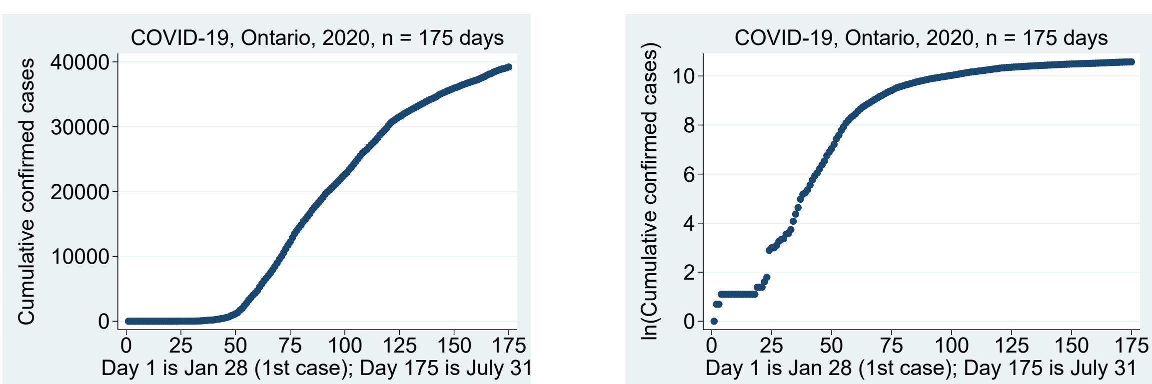
    (c) South Africa?

The *JEP* is one of the journals that publish replication data. Hence, we can redraw the graph without the logs to see why it's not very effective. The four graphs below all plot the same dots, which we call observations, as Figure 1 above. The only difference is how they use logarithms. The graph on the top left does not use logarithms at all, and it is hard to read with lots of dots piled on top of each other. The top right graph uses a natural log transformation for each variable and the bottom left graph uses a log base 10 transformation. As you should have expected from Section 4.4.1, it makes no difference to the visual look of the graph whether we use ln or log_{10}, but does affect the numeric scale of the x and y axes. Finally, the graph on the bottom right shows the logarithmic scale like the original Figure 1, which helps people remember that if $log_{10}(x) = 3$ then x is 1,000.

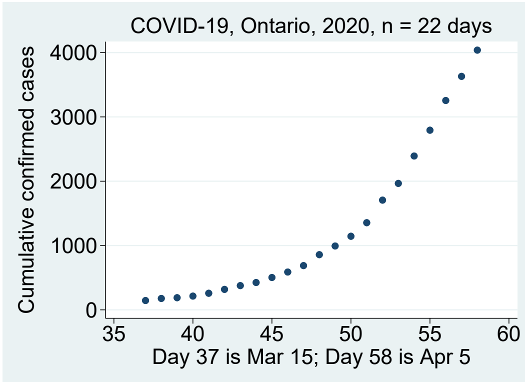


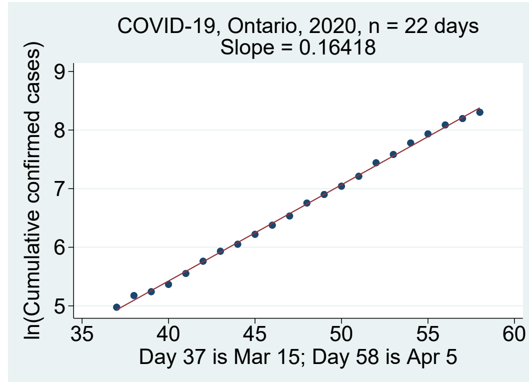
4.4.4 Review Linear and Logarithmic Functions: COVID-19 Case Study

Consider data from the "Status of COVID-19 cases in Ontario" from the Ontario Data Catalog (retrieved on July 31, 2020 https://data.ontario.ca/dataset/status-of-covid-19-cases-in-ontario). The left graph shows the cumulative number of confirmed cases daily from January 28, 2020 through July 31, 2020. The right graph shows the natural log of cumulative number of confirmed cases.

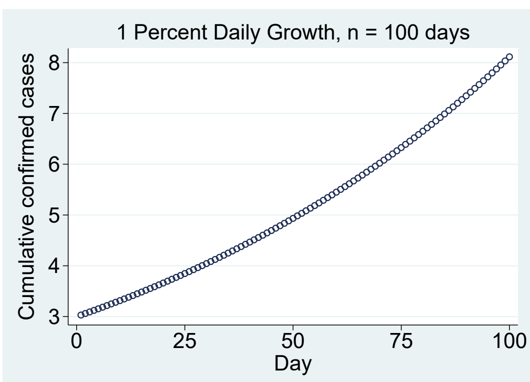


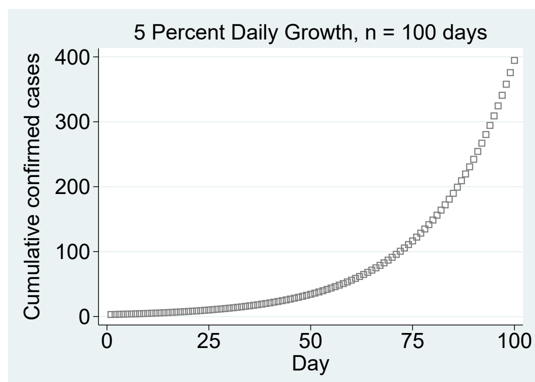
The next two graphs zoom in on the period between March 15, 2020 and April 5, 2020.

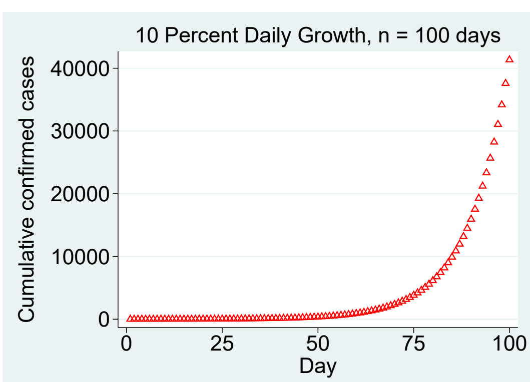


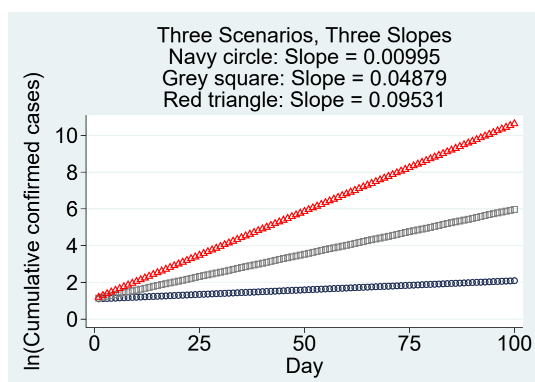


In real data, like for Ontario, there is random variation from day to day and evolving trends. To illustrate most clearly, we explore some clean fake data and then return to Ontario. Consider three situations where a city starts with 3 confirmed cases on Day 1 and then experiences either a 1 percent, 5 percent, or 10 percent daily growth rate, which are illustrated below. Like in Section 4.4.2, all three curves show the effect of compound growth: they are increasing at an increasing rate. The vertical axis of the fourth graph is the natural log of the number of cumulative confirmed cases: it shows three *linear* relationships (constant slope). This means that the percentage growth rate is constant. Like in Section 4.4.2, the slope after the natural log transformation of the y variable is an excellent approximation of the growth rate.

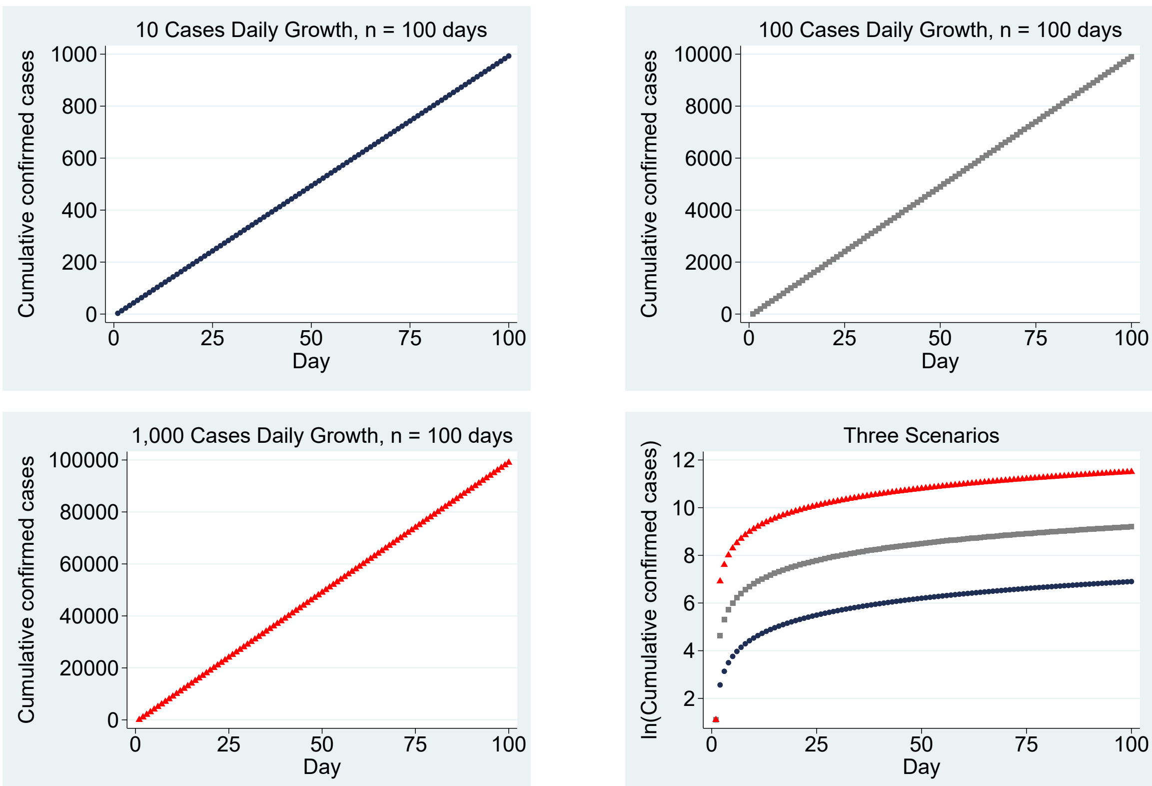








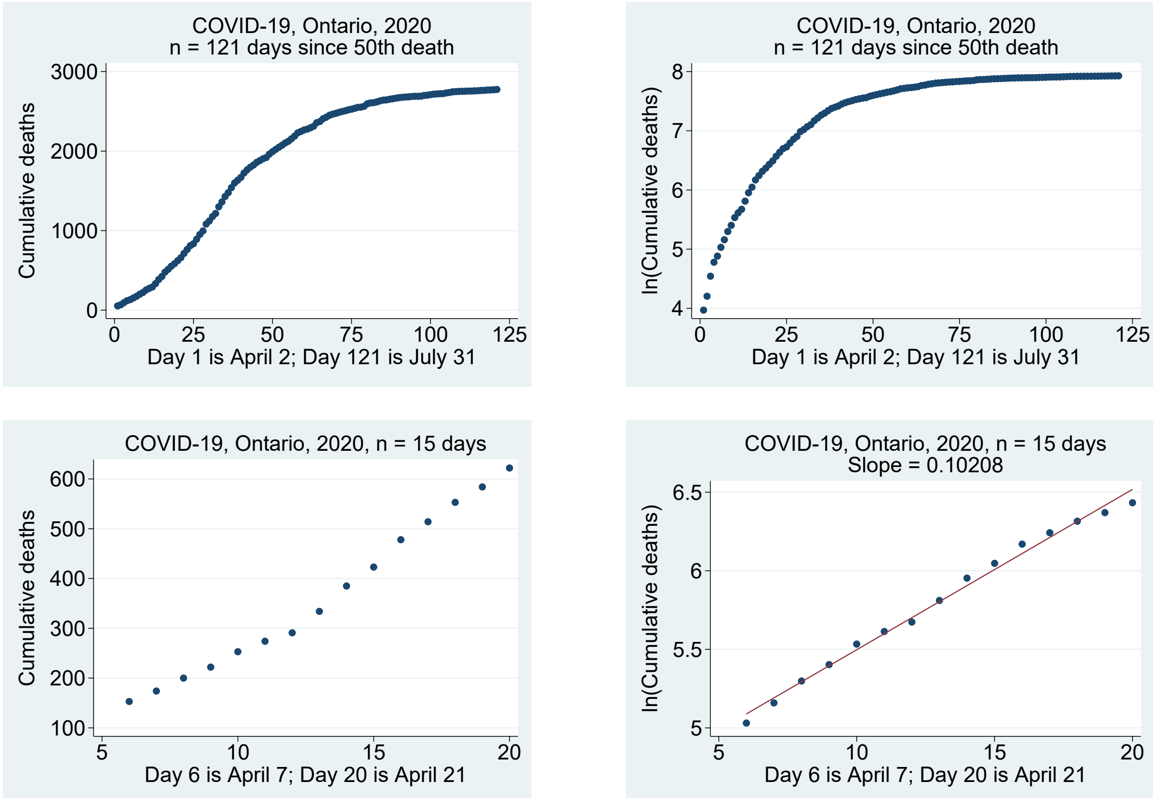
For a useful contrast, consider another set of hypothetical scenarios. In this case a city starts with 3 confirmed cases on Day 1 and then experiences either 10 new cases per day, 100 new cases per day, or 1,000 new cases per day. In contrast to percent growth, the cumulative cases are increasing at a constant rate, not an increasing rate. In other words, the relationship between cumulative confirmed cases and time is *linear*, which means a constant slope. If we use a natural log transformation on these, the fourth graph shows it is increasing at a decreasing rate. To understand, consider a city experiencing 100 new cases per day. If Day 1 is 3 cases then Day 2 is 103 cases which is a 3,333% increase! However, from Day 2 (103 cases) to Day 3 (203 cases) there is a 97% increase. From Day 3 (203) to Day 4 (303) there is 49% increase. The downward trend in the growth rate continues and by Day 100 (with 9,903 cases) there is only a 1% increase from Day 99 (with 9,803 cases). This is why the relationship between the natural log of cumulative confirmed cases and time is increasing at a decreasing rate.



Now, returning to the real Ontario data, we experienced many different things from January to July 2020. For example, in the 22 days from March 15 through April 5, the daily count of cases was increasing at an increasing rate, which corresponded to approximately a 16.4 percent daily growth rate in confirmed cases, which is alarming. However, after that, the daily percent jumps in cases started coming down as evidenced by the natural log curve becoming less steep. After roughly Day 90, even the absolute daily increase in cases started coming down: the cumulative case curve started becoming less steep. After the natural log transformation the curve became nearly flat as the daily percent increases became increasingly small.

Section 4.4.4 Exercises:

1. Continue with the Ontario data, but focus on deaths rather than cases. See the graphs below.



(a) What does the estimated slope (0.10208) in the bottom right graph mean?

(b) Why would we focus on April 7 to April 21? Why not the entire period of these data?

5 Derivatives

To find the rate of change, slope, at a particular point, take the derivative of a function. It is common to write the derivative of a function as: $f'(x)$, y' and $\frac{dy}{dx}$. Derivatives are useful in economics. For example, to obtain MC (marginal cost) from TC (total cost), take the derivative of TC. While we only need derivatives for part of ECO220Y, we work with functions – both linear and nonlinear – extensively, which adds much value to reviewing derivatives.

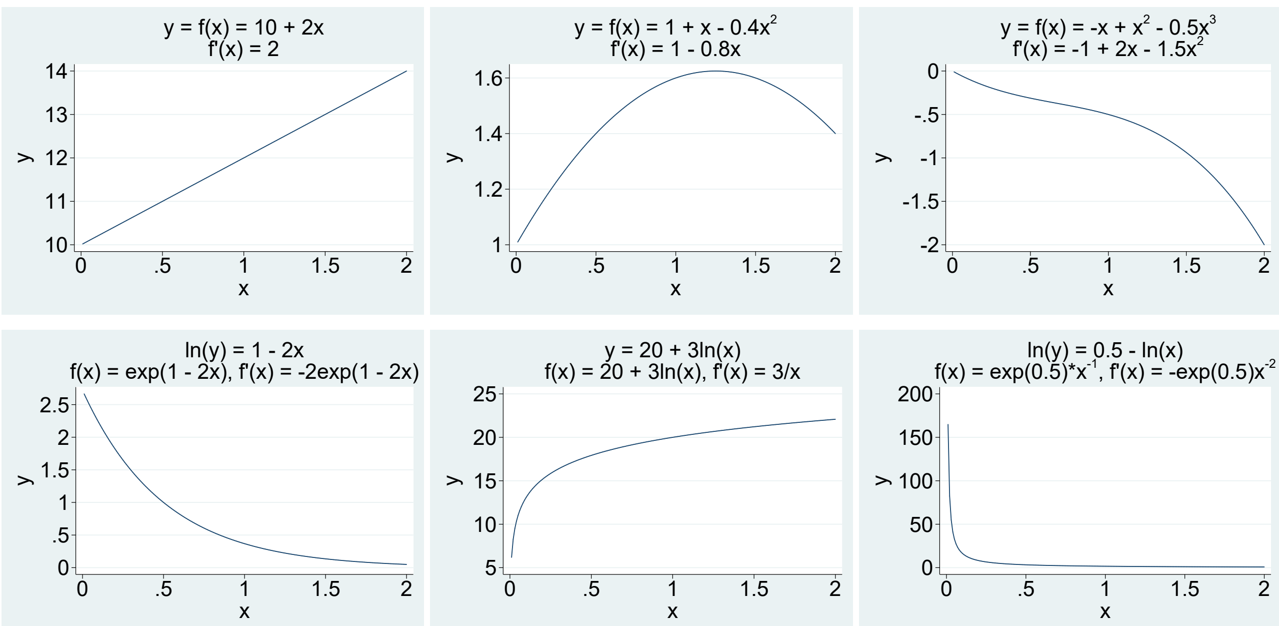
Rules of Differentiation: (Suppose c is a constant and $f'(x)$ and $g'(x)$ exist)

- If the function is constant, $f(x) = c$, then $f'(x) = 0$
- If n is any real number and $f(x) = x^n$, then $f'(x) = nx^{n-1}$ ("Power Rule")
- If the function is exponential, $f(x) = e^x$, then $f'(x) = e^x$
- If the function is exponential, $f(x) = a^x$, then $f'(x) = a^x ln(a)$
- If the function is a natural log, $f(x) = ln(x)$ then $f'(x) = \frac{1}{x}$
- If the function is a log base a, $f(x) = log_a(x)$ then $f'(x) = \frac{1}{xln(a)}$
- If $h(x) = cf(x)$ then $h'(x) = cf'(x)$

- If $h(x) = f(x) + g(x)$ then $h'(x) = f'(x) + g'(x)$
- If $h(x) = f(x) - g(x)$ then $h'(x) = f'(x) - g'(x)$
- If $h(x) = f(x)g(x)$ then $h'(x) = f'(x)g(x) + f(x)g'(x)$ ("Product Rule")
- If $h(x) = \frac{f(x)}{g(x)}$ then $h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$ ("Quotient Rule")
- If F is the composite function defined by $F(x) = f(g(x))$, then $F'(x) = f'(g(x))g'(x)$ ("Chain Rule": In words, take the derivative of the outside function, $f(x)$, keeping the inside function, $g(x)$, and then multiply by the derivative of the inside function, $g'(x)$.)

For example, what is the derivative of $ln(y) = a + bx + cx^2$ with respect to x? To simplify, rewrite as $y = e^{a+bx+cx^2}$. Use the Chain Rule to obtain $y' = e^{a+bx+cx^2}(b + 2cx)$.

Next, consider six graphical examples highlighting some functional forms relevant for ECO220Y1Y.



Notice that, except for the linear function with a constant slope, the slope depends on the value of x (i.e. where you are on the curve). In some cases the sign of the slope depends on the value of x.

- A *monotonic function* is one that is either always increasing or decreasing: in other words, the *sign* of the slope does not vary. Examples of monotonic functions: $y = a + bx$, $ln(y) = a + bx$, $y = a + bln(x)$ and $ln(y) = a + bln(x)$ where a and b are constants.
- A *non-monotonic function* is one that is sometimes increasing and sometimes decreasing: in other words, the *sign* of the slope varies. Examples: quadratic and higher degree polynomials.

5.1 Partial Derivatives

There is an important moment in ECO220Y requiring partial derivatives: multiple regression. Consider a function of two variables, x and w, such that $y = f(x, w)$. It is common to write the partial derivative with respect to x as $f_x(x, w)$, $\frac{\partial f}{\partial x}$, $\frac{\partial}{\partial x} f(x, w)$, and $\frac{\partial y}{\partial x}$ and likewise for w.

Rules of Partial Differentiation:

- To find the partial derivative with respect to x, treat w as a constant and differentiate $f(x, w)$ with respect to x
- To find the partial derivative with respect to w, treat x as a constant and differentiate $f(x, w)$ with respect to w

For example, consider $y = f(x, w) = a + bx + cw$, where a, b, and c are constants. The partial derivative with respect to x is $\frac{\partial y}{\partial x} = b$ and the partial derivative with respect to w is $\frac{\partial y}{\partial w} = c$. Remember that this means b is the change in y given a change in x while *holding w constant*. Similarly, c is the change in y given a change in w while *holding x constant*.

As another example, consider $y = f(x, w) = a + bx + cw + d(x \times w)$, where a, b, c, and d are constants. $\frac{\partial y}{\partial x} = b + dw$ and $\frac{\partial y}{\partial w} = c + dx$. In other words, the slope of the relationship between y and x depends on the level of w. These kinds of interactions are very common. A classic example in economics are the complementarities between labor and capital in producing output: the productivity of labor often depends on capital. For example, the value of a farmer spending an extra hour tilling his field greatly depends on whether or not he can use a tractor. In this example the farm output is the y variable and level of labor and capital are the x and w variables.

Section 5 Exercises:

1. What is the slope of $y = 22.1 + 0.8x - 0.05x^2$ when $x = 2$?
2. What is the slope of $ln(y) = 4.2 - 2.1ln(x)$ when $x = 10$?
3. What is the effect on y of increasing x by one unit if $y = 10 + 1x + 3w + 4(x \times w)$ and $w = 2$?

6 Inequalities, Intervals, and Absolute Values

Inequalities

Consider any two real numbers a and b:

$a < b$	a is less than b	a is *strictly* less than b
$a \leq b$	a is less than or equal to b	a is *weakly* less than b
$a > b$	a is greater than b	a is *strictly* greater than b
$a \geq b$	a is greater than or equal to b	a is *weakly* greater than b
$a \neq b$	a is not equal to b	

An important skill is translating between plain English statements and precise mathematical statements. Consider these illustrations, but remember that there are many ways to precisely state something in plain English (i.e. this is not a comprehensive list).

- "At least four people" means $x \geq 4$ where x is the integer number of people. It's a weak inequality because the statement implies that exactly four people is possible (four is at least four). It's also correct to write $x > 3$, which is equivalent to $x \geq 4$ when x is an integer.

- "More than one defective item" means $x > 1$ where x is the integer number of defects. It's a strict inequality because the statement implies that exactly one is excluded (one is not more than one). It's also correct to write $x \geq 2$, which is equivalent to $x > 1$ when x is an integer.

- "Fewer than 10 heat alert days" means $x < 10$ where x is the integer number of days. It's a strict inequality because the statement implies that exactly ten is excluded (ten is not fewer than ten). It's also correct to write $x \leq 9$, which is equivalent to $x < 10$ when x is an integer.

- "As few as 3 complaints" means $x \leq 3$ where x is the integer number of complaints. It's a weak inequality because the statement implies that exactly three complaints is a possibility. It's also correct to write $x < 4$, which is equivalent to $x \leq 3$ when x is an integer. (Note: If you are confused about the direction of the inequality, "as few as" means that few or fewer: "as few as" implies that the amount is surprisingly small.)

- "No more than 2 pieces of carry-on luggage" means $x \leq 2$ where x is the integer number of pieces. It's a weak inequality because the statement implies that exactly two pieces is a possibility (two is no more than two). It's also correct to write $x < 3$, which is equivalent to $x \leq 2$ when x is an integer.

- "As few as 10 percent of people exit during a fire alarm" means $x \leq 10$ where x is the percent that exit. x is *not* logically restricted to integer values (9.25 percent is possible).

- "Weigh less than 25 kg" means $x < 25$ where x is weight. It's a strict inequality because the statement implies that exactly 25 is excluded (25 is not less than 25).

- "Growth of 2.5 percent at a minimum" means $x \geq 2.5$ where x is the percentage growth rate. It's a weak inequality because the statement implies that exactly 2.5 is a possibility.

Rules for Inequalities:

- If $a < b$, then $a + c < b + c$
- If $a < b$ and $c < d$, then $a + c < b + d$
- If $a < b$ and $c > 0$, then $ac < bc$
- If $a < b$ and $c < 0$, then $ac > bc$
- If $0 < a < b$, then $\frac{1}{a} > \frac{1}{b}$

Intervals

Intervals are sets of real numbers. For real numbers a and b where $a < b$, possible intervals are:

(a, b)	$\{x \mid a < x < b\}$	Open interval of all values of x such that $x > a$ and $x < b$
$[a, b]$	$\{x \mid a \leq x \leq b\}$	Closed interval of all values of x such that $x \geq a$ and $x \leq b$
$[a, b)$	$\{x \mid a \leq x < b\}$	Interval of all values of x such that $x \geq a$ and $x < b$
$(a, b]$	$\{x \mid a < x \leq b\}$	Interval of all values of x such that $x > a$ and $x \leq b$
(a, ∞)	$\{x \mid x > a\}$	Infinite interval of all values of x such that $x > a$
$[a, \infty)$	$\{x \mid x \geq a\}$	Infinite interval of all values of x such that $x \geq a$
$(-\infty, b)$	$\{x \mid x < b\}$	Infinite interval of all values of x such that $x < b$
$(-\infty, b]$	$\{x \mid x \leq b\}$	Infinite interval of all values of x such that $x \leq b$
$(-\infty, \infty)$	All Real numbers	Infinite interval of all values of x

- "Two to four requests are expected" means $2 \le x \le 4$ where x is the integer number of requests. It's also correct to write $1 < x < 5$, which is equivalent to $2 \le x \le 4$ when x is an integer.
- "Concentrations must not exceed 6 ppm (parts per million) but must be at least 3 ppm" means $3 \le x \le 6$ where x is the concentration. x is *not* logically restricted to integer values.
- "Exposure should exceed 30 minutes but not surpass 60 minutes" means $30 < x \le 60$ where x is minutes. x is *not* logically restricted to integer values (40.75 minutes is possible).

Other useful symbols:

- $\approx$, $\simeq$ or $\cong$ is read as "approximately equal to." For example, $27.8889 \approx 28$.
- $\pm$ is read as "plus or minus." For example 4 ± 3 means [1, 7].

Absolute Values

The absolute value of a number a, denoted by $|a|$, is the distance from a to 0 on the real number line. Distances are always positive or 0, so $|a| \ge 0$ for every number a. $|a| = a$ if $a \ge 0$ and $|a| = -a$ if $a < 0$. (Recall: If a is negative then $-a$ is positive.)

7 Areas

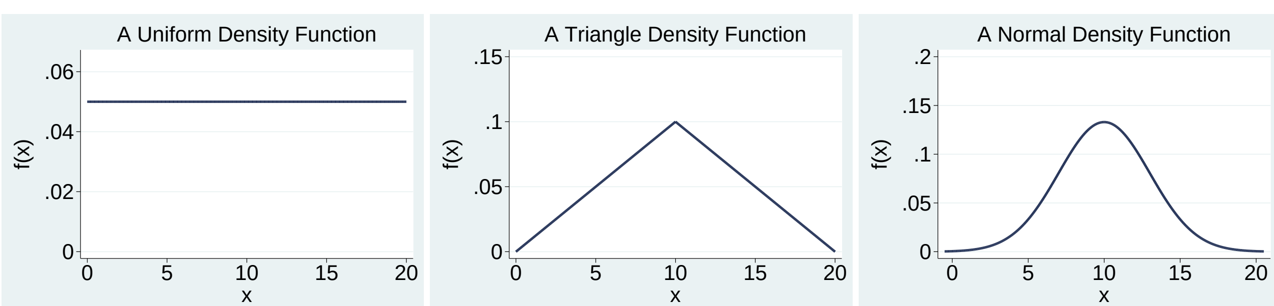
Finding areas is important in statistics. This may suggest that integration (calculus) is necessary. However, the types of functions we work with are either too simple or too complex to use integration. That means you will either be using simple formulas for finding areas or relying on statistical tables.

- Area of a rectangle $= b \times h$ where b is the length of the base and h is the height.
- Area of a triangle $= \frac{1}{2} b \times h$ where b is the length of the base and h is the height.

Whether using simple formulas or statistical tables, you will need to also use some basic logic. For example, if the total area is 5 and a sub-part has an area of 3 then the non-sub-part area is 2. The exercises below are specifically constructed for practice on commonly needed skills in ECO220Y. In particular, all of the figures show *density functions*. Density functions are always non-negative. Further, the total (positive) area under a density function is exactly 1. The *support* of the density function is the range of x-axis values that have a (non-zero) line above them. To summarize:

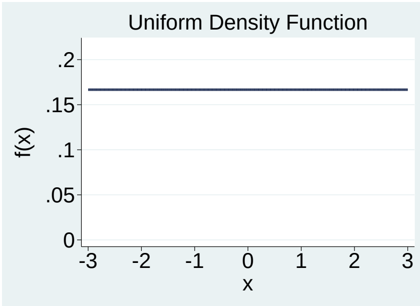
- A density function is non-negative: $f(x) >= 0$
- The total (positive) area under a density function is exactly 1

Here are some sample density functions. Finding areas for the first two is simple. Finding areas under the Normal density function is harder: a statistical table is required as there is no formula for the area.



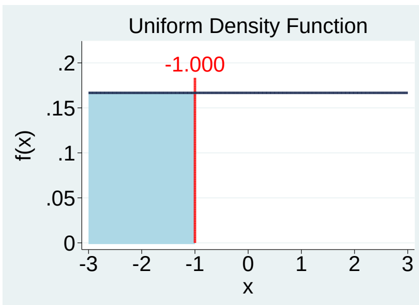
Section 7 Exercises:

1. Consider the Uniform density function given below. Its support is $[-3, 3]$.

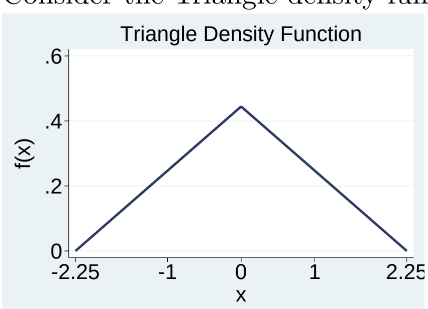


   (a) What is the *exact* height of the density function? In other words, what is the constant a in $f(x) = a$? (Hint: Remember that the total area under a density function is exactly 1.)

   (b) What is the shaded area below under the density function between -3 and -1?

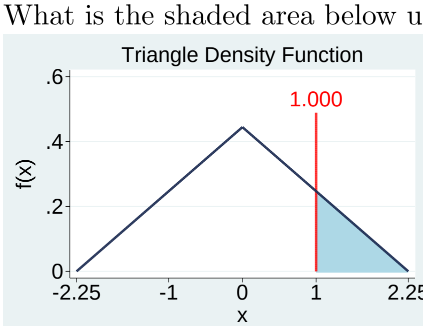


2. Consider the Triangle density function given below. Its support is $[-2.25, 2.25]$.

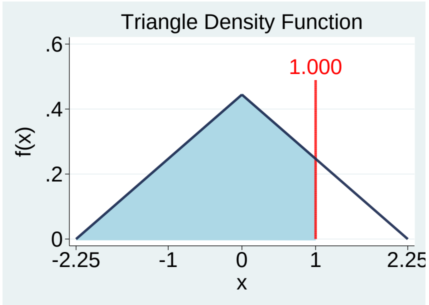


   (a) What is the *exact* maximum height of the density function? In other words, what is the value of $f(x)$ at the peak of the triangle? (Hint: Again, remember the total area is 1.)

   (b) What is the shaded area below under the density function between 1 and 2.25?



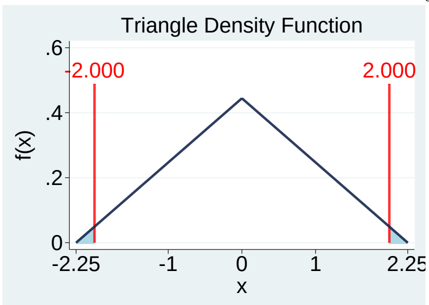
(c) What is the shaded area below under the density function between -2.25 and 1?



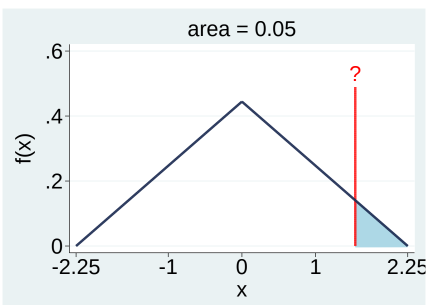
(d) What is the area under the density function between -2.25 and 0?

(e) What is the area under the density function between 0 and 1? (Draw a graph.)

(f) What is the area under the density function between -2.25 and -2 and between 2 and 2.25?



(g) What is the value of "?" in the diagram below if the shaded area shown is exactly 0.05?



8 Summation

A quick browse of the formulas in ECO220Y reveals many occurrences of summation in sigma notation: Σ. For example, to find the mean, add up all values and divide by the number of observations.

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + ... + x_n}{n} \tag{4}$$

Consider a general definition of summation. The *index* of summation is i. If $a_m, a_{m+1}, ..., a_n$ are real numbers and m and n are integers such that $m \leq n$, then

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + ... + a_{n-1} + a_n \tag{5}$$

For example, write $2 + 2 = 4$ in sigma notation as $\sum_{i=1}^{2} 2 = 4$. The index i takes two values (1 and 2) because the original sum has two terms and $a_1 = a_2 = 2$. Write $2 + 2 + 2 + 2 = 8$ in sigma notation as $\sum_{i=1}^{4} 2 = 8$. The index i takes four values (1, 2, 3 and 4) because the original sum has four terms and $a_1 = a_2 = a_3 = a_4 = 2$. These two examples are a special case where we add up a constant. In general, adding up a constant n times yields n times the constant:

$$\sum_{i=i}^{n} a = n \times a \tag{6}$$

Next, consider annual compensation for 10 randomly selected CEO's in Fortune 500 companies measured in millions of dollars. Name the variable x_i. Hence, x_1 records the first CEO's compensation (\$14.2 million) and x_{10} records the tenth CEO's compensation (\$22.4 million).

$$\sum_{i=1}^{10} x_i = 14.2 + 8.8 + 25.6 + 22.7 + 4.1 + 3.0 + 9.5 + 9.1 + 50.1 + 22.4 = 169.5 \tag{7}$$

The sample mean compensation is \$16.95 million: $\bar{X} = \frac{169.5}{10} = 16.95$.

Laws of Summation: where c is a constant and n is a positive integer

- $\sum_{i=m}^{n} ca_i = c \sum_{i=m}^{n} a_i$
- $\sum_{i=m}^{n} (a_i + b_i) = \sum_{i=m}^{n} a_i + \sum_{i=m}^{n} b_i$
- $\sum_{i=m}^{n} (a_i - b_i) = \sum_{i=m}^{n} a_i - \sum_{i=m}^{n} b_i$
- $\sum_{i=1}^{n} 1 = n$
- $\sum_{i=1}^{n} c = nc$
- $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$

Notice that an i subscript, as in a_i and b_i, means a *variable*: not all observations must be the same value. In contrast, no subscript, as in c, means a *constant*. More common notation in economic applications is x_i and y_i rather than a_i and b_i.

For example, suppose x_i records days to complete stage 1 of a production process for each of 25 items and $\sum_{i=1}^{25} x_i = 23.14$ days. (Notice that x_i is a variable: there is variation in completion time across items.) Similarly, y_i records hours to complete stage 2 and $\sum_{i=1}^{25} y_i = 114.81$ hours. For stage 3, every item needs exactly 105 minutes of drying time (no variation). What is the total time it takes to complete all three stages for all 25 items? To answer, choose common units, say hours. Then, find $\sum_{i=1}^{25}(x_i \times 24 + y_i + 105/60)$. Using the laws of summation, $\sum_{i=1}^{25}(24x_i + y_i + 105/60) = 24\sum_{i=1}^{25} x_i + \sum_{i=1}^{25} y_i + \sum_{i=1}^{25} 105/60 = 24 \times 23.14 + 114.81 + 25 \times 105/60 = 713.92$ hours.

While you can distribute a summation over ***linear transformations***, where a linear transformation of a variable is simply adding a *constant*, subtracting a *constant*, and/or multiplying by a *constant*, and you can distribute a summation over ***linear combinations***, where a linear combination of two or more variables is simply adding or subtracting variables and/or multiplying by *constants*, you *CANNOT* distribute a summation over ***non-linear transformations*** or ***non-linear combinations***. In other words, notice what is *NOT* in the Laws of Summation above. These are *INEQUALITIES*:[5]

[5]These are examples of Jensen's Inequality.

- $\left(\sum_{i=1}^{n} x_i\right)^c \neq \sum_{i=1}^{n} (x_i)^c$ (For example, $\left(\sum_{i=1}^{n} x_i\right)^2 \neq \sum_{i=1}^{n} (x_i)^2)$
- $\sum_{i=1}^{n} x_i \times y_i \neq \sum_{i=1}^{n} x_i \times \sum_{i=1}^{n} y_i$
- $\sum_{i=1}^{n} \frac{x_i}{y_i} \neq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}$
- $ln\left(\sum_{i=1}^{n} x_i\right) \neq \sum_{i=1}^{n} ln(x_i)$

For example, how to find $\sum_{i=1}^{n} \frac{(3x_i-4y_i)^2}{12}$? First, rewrite as $\sum_{i=1}^{n} \left(\frac{3}{4}x_i^2 - 2x_iy_i + \frac{4}{3}y_i^2\right)$. Next, use the Laws of Summation: $\left(\frac{3}{4}\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_iy_i + \frac{4}{3}\sum_{i=1}^{n} y_i^2\right)$. However, that is as far as we can go. We need to compute $\sum_{i=1}^{n} x_i^2$, $\sum_{i=1}^{n} x_iy_i$, and $\sum_{i=1}^{n} y_i^2$. It is *NOT* enough to know $\sum_{i=1}^{n} x_i$ and $\sum_{i=1}^{n} y_i$.

Section 8 Exercises:

1. Consider these data with two variables and 30 observations. *None* of the questions require a lot of computation. Make sure to use all given information and the laws of summation.

i	x	y
1	18.1	3.8
2	32.8	7.5
3	4.9	4.2
4	22.1	6.6
5	18.7	4.0
6	44.7	5.6
7	21.6	4.1
8	10.8	6.3
9	47.2	6.4
10	10.9	4.2
11	21.2	3.3
12	8.0	2.6
13	24.5	5.4
14	32.8	1.8
15	3.2	7.9
16	28.7	1.0
17	33.4	1.1
18	41.7	5.5
19	1.9	0.3
20	46.9	8.2
21	27.2	3.6
22	0.7	8.1
23	20.4	0.5
24	47.9	2.1
25	40.7	5.9
26	35.8	9.8
27	28.0	6.7
28	39.9	1.9
29	36.1	2.5
30	45.2	7.2
Total	796.0	138.1

(a) What is $\sum_{i=1}^{30} x_i$?

(b) What is $\sum_{i=1}^{30} y_i$?

(c) What is $\sum_{i=1}^{30}(x_i + y_i)$?

(d) What is $\sum_{i=1}^{30}(x_i - 3y_i)$?

(e) What is $\sum_{i=1}^{3} x_i y_i$? (Note it says 3, not 30.)

(f) What is $(\sum_{i=1}^{3} x_i)(\sum_{i=1}^{3} y_i)$? (Note it says 3, not 30.)

(g) Are your answers equal in (e) and (f)? Why or why not?

(h) What is $\sum_{i=29}^{30} x_i^2$? (Note it says 29, not 1.)

(i) What is $(\sum_{i=29}^{30} x_i)^2$? (Note it says 29, not 1.)

(j) Are your answers equal in (h) and (i)? Why or why not?

(k) What is $\sum_{i=1}^{30}(10 + 22x_i - 30y_i)$?

(l) What is $\sum_{i=1}^{30}((x_i - 30) + (y_i - 2))$?

(m) Suppose the observations represent a cross-section of small businesses. Further, suppose x is reported in the above table in 1000's of dollars and is a measure of profits for the entire year (12 months). This means that for Firm 1 (the first observation) x is 18.1, which corresponds to profits of \$18,100. Suppose we create a new variable z that is x measured in dollars (hence, z_1 is \$18,100). What is $\sum_{i=1}^{30} z_i$?

(n) Suppose y represents 1000's of dollars lost to theft over two years. Suppose we create a new variable w that is an annualized value of y (average loss per year) in 1000's of dollars. What is $\sum_{i=1}^{30} w_i$?

(o) Suppose that the government is considering charging a lump-sum profit tax of \$1,000 per year per small business. If x measures each firm's profits before this profit tax, what is $\sum_{i=1}^{30} q_i$ where q measures profits in 1000's of dollars after this profit tax?

(p) Suppose that the government is considering charging a profit tax of 10% of profits per year per small business. If x measures each firm's profits before this profit tax, what is $\sum_{i=1}^{30} q_i$ where q measures profits in 1000's of dollars after this profit tax ?

2. Suppose $\sum_{k=1}^{1000} x_k^2 = 427744.2$ and $\sum_{k=1}^{1000} x_k = 20051.8$.

(a) Find $\sum_{k=1}^{1000}(x_k - 10)^2$.

(b) Find $\sum_{k=1}^{1000}(x_k - \bar{X})^2$.

(c) Find $\sum_{k=1}^{1000}(x_k - 4)^2$.

(d) In which case is the sum the smallest? Intuition for that finding?

9 Answers to Exercises that Appear in Sections 3 - 8

Section 3 Exercises, Answers:

1. (a) 3.76

(b) 2.7876

2. The overall flu shot rate is 35.26% $\left(= 100\frac{(162+289)}{(502+777)}\right)$.

3. (a) 70.04% is conditional on (1) got the flu shot last year and (2) suffered a flu illness last year. 74.37% is conditional on (1) got the flu shot last year and (2) did not suffer a flu illness last year.

   (b) The overall flu shot rate this year among those that got a flu shot last year is 73.98% $\left(= 100\frac{(0.7004\times 4{,}555+0.7437\times 46{,}387)}{(4{,}555+46{,}387)}\right)$. However, 73.98% hides the evidence that people who got the shot and still got sick may become discouraged and skip the flu shot next time.

Section 4.1 Exercises, Answers:

1. (a) $y = 2$

   (b) No change: this is a horizontal line

2. (a) $y = 4 - 2x$

   (b) Decreases by 2 units

3. (a) $y = 12 - x$

   (b) $y = 7 - x$

   (c) $y = 10 - 2x$

   (d) $y = 20 - 2x$

Section 4.2 Exercises, Answers:

1. This is a third degree polynomial (cubic). $a_0 = 23.1, a_1 = 2.4, a_2 = 5.9, a_3 = -6.2$

2. x = 0.5307 and x = -1.1307

Section 4.4.3 Exercises, Answers:

1. (a) Bangladesh consumes a bit more than 300 kWh of electricity per capita and has GDP per capita that is a bit more than $1,000. If you download the replication data, you can obtain the exact values of 320.20416 kWh and $1,118.8537, respectively.

   (b) Brazil consumes a bit less than 3,000 kWh of electricity per capita and has GDP per capita that is roughly $12,000. If you download the replication data, you can obtain the exact values of 2619.9606 kWh and $12,112.59, respectively.

   (c) South Africa consumes a bit more than 4,000 kWh of electricity per capita and has GDP per capita a bit higher than $6,000. If you download the replication data, you can obtain the exact values of 4197.907 kWh and $6,428.2936, respectively.

Section 4.4.4 Exercises, Answers:

1. (a) In 2020, the daily growth rate in deaths from COVID-19 between April 7 and April 21 in Ontario is approximately 10 percent on average.

   (b) The graph on the top right clearly is not straight. Instead it is concave, which means that the daily percent growth rate is becoming smaller over time. Hence, the percent growth rate is changing dramatically over the overall period. The specific range of April 7 to April 21 has a quite constant daily percent growth rate, which means it makes sense to talk about the typical growth over that period.

Section 5 Exercises, Answers:

1. 0.6
2. -0.1112
3. 9

Section 7 Exercises, Answers:

1. (a) $\frac{1}{6}$
   (b) $\frac{1}{3}$
2. (a) $\frac{2}{4.5} = \frac{4}{9}$
   (b) The shaded area is $\frac{25}{162}$. To find the area, requires first finding the equation of the line that connects the triangle peak $(0, \frac{4}{9})$ to the right corner $(\frac{9}{4}, 0)$. That line is $f(x) = \frac{4}{9} - \frac{16}{81}x$.
   (c) The shaded area is $\frac{137}{162}$, which is most easily found as $1 - \frac{25}{162}$.
   (d) The shaded area is $\frac{1}{2}$, which is most easily found by noting that that is exactly half of the total area and the total area is 1.
   (e) The area is $\frac{56}{162}$, which is most easily found as $\frac{1}{2} - \frac{25}{162}$.
   (f) The shaded area is $\frac{2}{162}$. Note that you may find the area in one of the tails and multiple it by two: the two shaded areas are exactly the same.
   (g) $\frac{90-9\sqrt{10}}{40}$. To find this is not hard but requires some work. You will need to use the quadratic formula, reviewed in Section 4.2 to solve.

Section 8 Exercises, Answers:

1. (a) 796.0
   (b) 138.1
   (c) 796.0 + 138.1 = 934.1
   (d) 796.0 - 3*138.1 = 381.7
   (e) 18.1*3.8 + 32.8*7.5 + 4.9*4.2 = 335.36
   (f) (18.1 + 32.8 + 4.9)*(3.8 + 7.5 + 4.2) = 864.9
   (g) No. No rule that says they should be equal and in general they are not.
   (h) $36.1^2 + 45.2^2 = 3346.25$
   (i) $(36.1 + 45.2)^2 = 6609.69$
   (j) No. No rule that says they should be equal and in general they are not.
   (k) 30*10 + 22*796.0 - 30*138.1 = 13,669
   (l) 796.0 - 30*30 + 138.1 - 2*30 = -25.9
   (m) 796,000
   (n) 138.1 / 2 = 69.05
   (o) 796.0 - 30*1 = 766.0

(p) 796.0*(1 - 0.10) = 716.4

2. (a) 126708.2
   (b) 25669.5
   (c) 283329.8
   (d) The sum is smallest when computing the squared difference between each observation and the mean. The mean is a measure of the center of the data, so it makes sense that the data are "closer" to the center than some other arbitrary point. For example, if x_i records cGPA, we would expect $\sum_{k=1}^{100}(x_k - \bar{X})^2$ to be smaller than $\sum_{k=1}^{100}(x_k - 4.0)^2$: there is going to be more people close to average cGPA than a 4.0 cGPA.

10 Quiz Answers

1. (e) 15.00 (See Section 3)
2. (b) 10.50 (See Section 3)
3. See Section 3, including the subsections, for the concepts. Below are some sample answers and, of course, your examples should be different but illustrate the same concepts.
   (a) If the annual inflation rate in Canada were to go up by 1 percentage point from 2019 to 2020 that would be a huge increase. In 2019 it was about 1.95 percent as measured by the consumer price index so an increase to 2.95 would be gigantic headline news. The Bank of Canada tries to keep inflation near the midpoint of the range from 1 to 3 percent and 2.95 would definitely start to make policy makers uncomfortable.
   (b) If a poll in 2019 finds that 42 percent of Canadians believe that climate change is national emergency and it increases to 43 percent in 2020 that is a small increase. That is not much movement in public opinion and because it is a poll based on a random sample, it could simply be sampling error.
   (c) If someone tracking the daily cases of COVID-19 in a city sees it is down to 1 new case one day but up to 3 new cases the next day that is a 200 percent increase. However, if the daily new cases had been bouncing around between 0 and 9 for some time, that 200 percent increase would be a small increase.
   (d) If global sea levels on average were rising by 2 percent per year that would be a large increase and coastal cities would become inundated in the foreseeable future.
4. (c) 20.00 (Solve with algebra)
5. (d) $\frac{X}{1+r/100}$ (Solve with algebra)
6. (Another application of Section 3)
   (a) In the 2018 audit of TTC streetcars, the newer streetcars have a shockingly higher fare evasion rate compared to the older streetcars: 18.61% (newer streetcars) is 11.04 percentage points higher than 7.57% (older streetcars) and 145.8% higher (more than twice as big)! One potential cause of the increase is the longer length of the newer streetcars: maybe riders feel more tempted to steal if the driver is far away.

(b) The overall fare evasion rate for streetcars is 14.05% $\left(= 100\frac{(609+174)}{(3{,}272+2{,}299)}\right)$.

(c) The value 18.61% is conditional on two things: (1) TTC mode is a streetcar and (2) the rider was entering a newer streetcar model.

(d) We would certainly expect to see the fare evasion rate for other major modes of transport operated by the TTC: buses and the subway. (They also break some of those down further, like they did for streetcars. For example, comparing automatic subway entrances with staffed subway entrances.)

7. (Another application of Section 3)

   (a) The figure clearly indicates that it shows the "percentage-point change" for each country from 2014 to 2018. Further, it says that, in Germany, the percent with negative sentiments declined by *about* 10 percentage points. Hence, if it had been 59% in 2014, then about 49% of Germans have negative sentiments about non-EU immigration in 2018 (just three and a half years later).

   (b) The figure shows that it increased by *about* 17 percentage points. Hence, if it had been 54% in 2014, then about 71% of Poles have negative sentiments about non-EU immigration in 2018 (just three and a half years later).

   (c) The change from 57% to 52% is a 5 percentage point decline and a 8.77 percent decline.

   (d) That is a 13 percentage point decline, which would put this city between Italy and Spain in the figure.

8. (c) $x <= 15$ (See Section 6)

9. (b) $x > 4$. Note that $x >= 5$ would have also been correct had it been one of the choices. $x > 4$ is equivalent to $x >= 5$ because in this example x is an integer. (See Section 6)

10. (b) $x <= 160$ (See Section 6)

11. (a) True

    (b) False

    (c) True

    (d) False

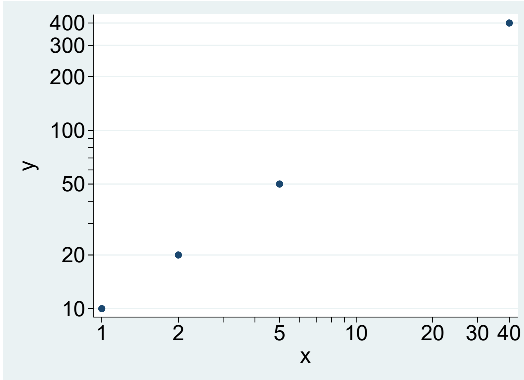
    (e) True

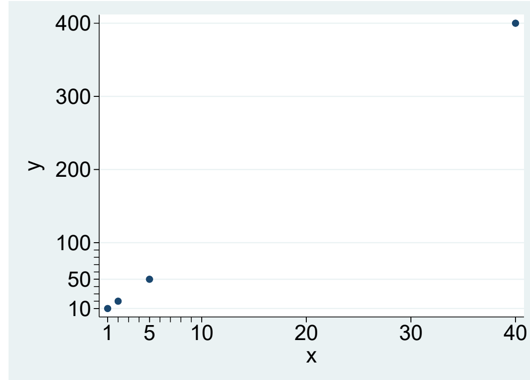
    (f) True

    (g) True (See Section 4.2)

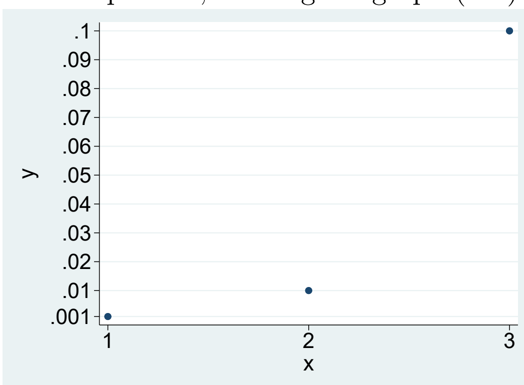
12. (An application of Section 4.4.3)

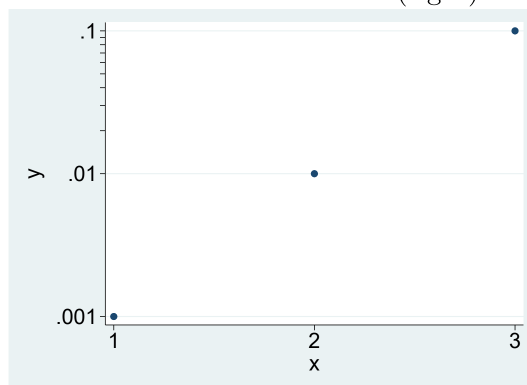
    (a) For comparison, the original graph (left) is provided in addition to the answer (right).





(b) For comparison, the original graph (left) is provided in addition to the answer (right).





13. Recall from your principles of economics course, that the elasticity of demand is $\frac{\%\Delta Q_D}{\%\Delta P}$.

   (a) When the price of dairy increases by 10%, the quantity demanded of dairy declines by 0.5%. The demand for dairy is extremely inelastic.

   (b) When the price of cereals and bakery products increases by 1%, the quantity demanded of cereals and bakery products decreases by 0.6%. The demand for cereals and bakery products is inelastic. Generally the more you aggregate up products the more inelastic demand becomes. Cereals and bakery products is a highly aggregate group that includes a range of things such as flour, breakfast cereals, rice, pasta, bread, muffins, cakes, etc. There are not a lot of good substitutes for this broad group, which explains why demand is inelastic.

   (c) When the price of white bread increases by 1 percent, the quantity demanded of white bread decreases by 1.55 percent. The demand for white bread is elastic.

14. (a) 66.7% (See Section 3)

15. 79.3% (See Section 3)

16. (a) For Canadian workers, we observe annual salaries increase by $6,188 Canadian with each additional year of education. (We tend to think about salaries in dollars. If we are talking about CEOs, celebrities, or athletes, we may think in millions of dollars.)

   (b) For people participating in a sleep study, an extra milligram of the drug yields an extra 9 minutes of sleep. (If you tell someone to meet you in 0.1562 hours, they will not understand you. For odd fractions of an hour, we tend to think in minutes.)

(c) For the electric vehicles in the study, each additional kilometer of battery range is associated with a suggested retail prices that is \$205 higher.

(d) For corporate boards of publicly traded companies in North America, for each additional year older the company is we see a slightly lower percentage of the board being female, 0.5 percentage points lower. An extra 10 years of age is associated with a decline of female representation on the board of 5 percentage points. (We tend to think in percentages, not fractions.)

(e) For the 500 most popular bottles of wine sold by the LCBO in 2016, each additional rating point (on a 100-point scale) is associated with the price of the bottle being nearly \$1 (Canadian) higher. (In this case the original units are good. We can just round 0.9681 to 1. Alternatively, it would have also been acceptable to say "being 97 cents (Canadian) higher.")

(f) For 100 large retail firms operating in North America, for every additional percentage point of the workforce making the minimum wage, we observe annual turnover that is about two percentage points higher.

(g) For colonies of bacteria being exposed to varying intensities of heat in a lab, increasing the temperature by one degree Celsius increase the number of bacteria killed by about 2.9 billion. (For huge numbers, we usually upgrade the units of measurement. For example, nobody talks about the Canadian GDP in dollars, but rather trillions of dollars. Also, if you answered 0.9 billion instead of 2.9 billion you got mixed up on the temperature conversion. An extra degree Celsius corresponds to an extra $\frac{9}{5}$ degrees Fahrenheit.)

(h) For ten years of annual observations of household shopping trips in the European Union, the share of shopping trips with a small purchase (less than 10 euro) has been declining by about one percentage point per year.

(i) For NBA basketball players, players that are 1 centimeter taller on average earn salaries that are about 45 thousand dollars higher. Remember that centimeters are smaller than inches: if each additional inch is associated with salaries that are \$114,264 higher, each additional cm will be associated with a smaller rise in salary. Also, when we are talking about such big salaries, we typically would not measure things in dollars (but in thousands or millions of dollars).

(j) For students, those whose lecture attendance is 1 percentage point higher earn final exam marks that are 0.14 percentage points higher on average. To get a more meaningful interpretation, try a less tiny increase in lecture attendance: those whose lecture attendance is 10 percentage points higher earn final exam marks that are 1.4 higher on average. Attending 1 percentage point more lectures would mean attending an extra 0.48 lecture ($= 0.01 \times 48$), which is about an extra half of a lecture (a tiny increase in attendance). From the original equation, we know that attending an extra 0.48 lectures corresponds to an 0.141698 ($= 0.295204 \times 0.48$) increase in the final exam mark and attending an extra 4.8 lectures (a 10 percentage point increase) corresponds to a 1.416979 ($= 0.295204 \times 4.8$) increase in the final exam mark. People are often confused by questions like this one and there are a lot of ways to understand it. Let's offer a couple of more ways to explain this.

    - According to the original equation, as the number of lectures attended increase by one lecture (e.g. attend 38 lectures instead of 37), the final exam mark is 0.295204 higher on average. One extra lecture out of 48 is a 2.083333 percentage point increase in attendance ($= 100 * 1/48$): for each extra 2.083333 percentage points of attendance, the final exam mark is higher by 0.295204 on average. Hence, for each extra 1 percentage point of attendance, the final exam mark is higher by 0.141698 on average ($= 0.295204/2.083333$).

- According to the original equation, a student that attends 24 lectures (50% attendance) has a predicted final exam mark of 68.1 ($68.11379 = 61.028894 + 0.295204 \times 24$). Similarly, a student that attends 30 lectures (62.5% attendance) has a predicted final exam mark of 69.9 ($69.885014 = 61.028894 + 0.295204 \times 30$). Hence, a 12.5 percentage point increase in attendance ($= 62.5 - 50$) is associated with a 1.771224 increase in the final exam mark ($= 69.885014 - 68.11379$). That means a 1 percentage point increase is associated with a 0.141698 increase ($= 1.771224 \div 12.5$).

17. (a) 4066.5 (See Section 8)

    (b) 3607.4 (See Section 8)

    (c) Note that $z_k = 1000x_k$. $\sum_{k=1}^{50} z_k = \sum_{k=1}^{50} 1000x_k = 1000 \sum_{k=1}^{50} x_k = 3565000$ and $\sum_{k=1}^{50} z_k^2 = \sum_{k=1}^{50} (1000x_k)^2 = 1000^2 \sum_{k=1}^{50} x_k^2 = 258251000000$ (See Section 8)

18. (a) Define y as the share (a percent) that voted for "Leave" and define x as the percentage of the region's GDP exported to the EU. Two points determine a line. We can very accurately approximate when the y value is either 50 or 60 as the figure clearly marks those: we just need to find (approximately) the corresponding x values. Hence, (x_1, y_1) is (about) $(7.85, 50)$ and (x_2, y_2) is (about) $(14.5, 60)$. Hence, after finding the slope and then using the point-slope formula, the approximate equation of the line is $y = 38.2 + 1.5x$. Thus, when x is 1 percentage point higher, y is 1.5 percentage points higher and the base value (intercept) is 38.2 percentage points, which is outside the range of the data and hence not shown in the figure (no region has zero EU exports). (See Section 4.1)

    (b) $y = 38.2 + 150x$. This still says that when x is 0.01 higher (which means one percentage point higher), y is 1.5 percentage points higher. (If you are confused, redraw the graph with the new units and/or re-do the point-slope formula calculations with the new units.) (See Section 4.1)

    (c) $y = 0.382 + 0.015x$. This still says that when x is 1 higher (which means one percentage point higher), y is 0.015 higher, which is 1.5 percentage points. (If you are confused redraw the graph with the new units and/or re-do the point-slope formula calculations with the new units.) (See Section 4.1)

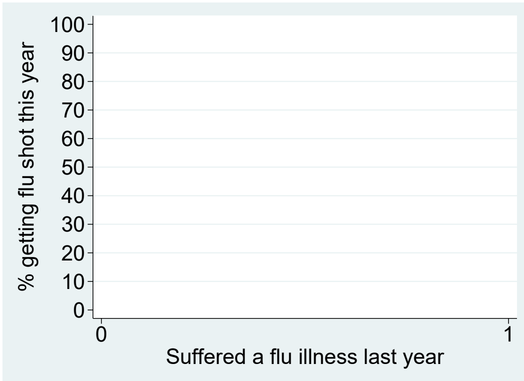
    (d) $y = 0.382 + 1.5x$. This still says that when x is 0.01 higher (which means one percentage point higher), y is 0.015 higher, which is 1.5 percentage points. (If you are confused redraw the graph with the new units and/or re-do the point-slope formula calculations with the new units.) (See Section 4.1)

19. (a) $b_2 = 5$ because we are told that males are 5 percentage points more likely to vote "Leave" regardless of exports. (See Section 4.1)

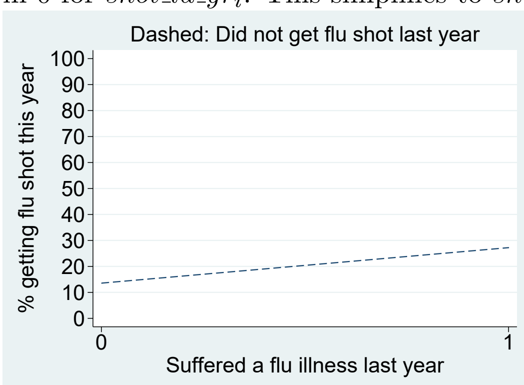
    (b) 35.7 is the intercept for the line for female voters (for whom $male_i = 0$), whereas the line for male voters has an intercept 5 units higher ($40.7 = 35.7 + 5$). (See Section 4.1)

(c) $y_i = 40.7 + 1.5x_i - 5fem_i$, which, just like in the previous part, implies that the male-voter line is $y_i = 40.7 + 1.5x_i$ (because for males $fem_i = 0$) and the parallel female-voter line is $y_i = 35.7 + 1.5x_i$ (because for females $fem_i = 1$). (See Section 4.1)

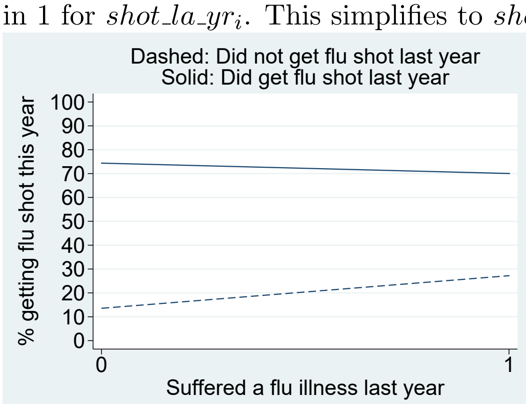
20. (a) Graphical answer:



(b) $shot\_pct\_th\_yr_i = 13.56 + 13.66ill\_la\_yr_i + 60.81 \times 0 - 17.99 \times 0 \times ill\_la\_yr_i$, which plugs in 0 for $shot\_la\_yr_i$. This simplifies to $shot\_pct\_th\_yr_i = 13.56 + 13.66ill\_la\_yr_i$.



(c) $shot\_pct\_th\_yr_i = 13.56 + 13.66ill\_la\_yr_i + 60.81 \times 1 - 17.99 \times 1 \times ill\_la\_yr_i$, which plugs in 1 for $shot\_la\_yr_i$. This simplifies to $shot\_pct\_th\_yr_i = 74.37 - 4.33ill\_la\_yr_i$.



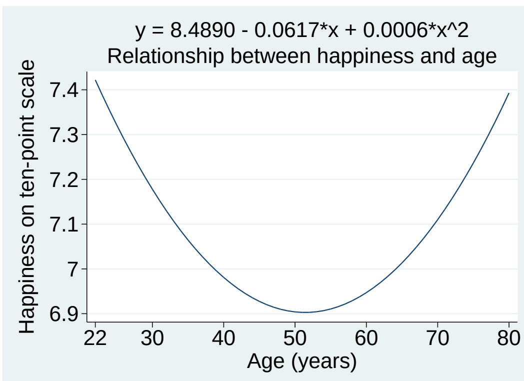
(d) The *difference* in the intercepts is $(74.37 - 13.56) = 60.81$. It is the coefficient on the $shot\_la\_yr$ term in the **key equation**.

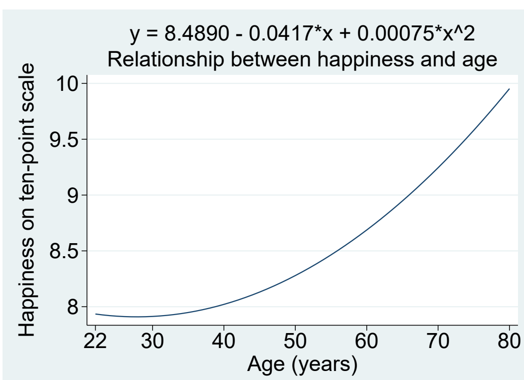
(e) The *difference* in the slopes is $(-4.33 - 13.66) = -17.99$. It is the coefficient on the $shot\_la\_yr \times ill\_la\_yr_i$ term in the **key equation**.

(f) To obtain the alternate key equation, start with the two lines obtained in the previous parts. Next, you must set the coefficients such that when $no\_shot\_la\_yr_i$ is 0 you obtain $shot\_pct\_th\_yr_i = 74.37 - 4.33 ill\_la\_yr$ and when $no\_shot\_la\_yr$ is 1 you obtain $shot\_pct\_th\_yr_i = 13.56 + 13.66 ill\_la\_yr_i$. This yields $shot\_pct\_th\_yr_i = 74.37 - 4.33 ill\_la\_yr_i - 60.81 no\_shot\_la\_yr_i + 17.99 no\_shot\_la\_yr \times ill\_la\_yr_i$.

21. The line for male voters is $y_i = 16 + 0.5x_i$ and the line for female voters is $y_i = -3.6 + 0.7x_i$. Hence, the intercept for male voters is 19.6 percentage points higher compared to female voters and the slope for male voters is 0.2 less steep: $y_i = -3.6 + 0.7x_i + 19.6male_i - 0.2(x_i \times male_i)$. (See Sections 4.1 and 5.1)

22. (a) (See Sections 4.2 and 5)
    i. -0.0117 (See Section 5)
    ii. 0.0063 (See Section 5)
    iii. 0.0303 (See Section 5)
    iv. Because a quadratic function does not have a constant slope. This function is non-monotonic: sometimes it is decreasing and sometimes increasing. (See Section 5)

(b) 

(c) 

23. 0.929. If you cannot solve this, make sure to work through Sections 4.1, 4.2 and 7.