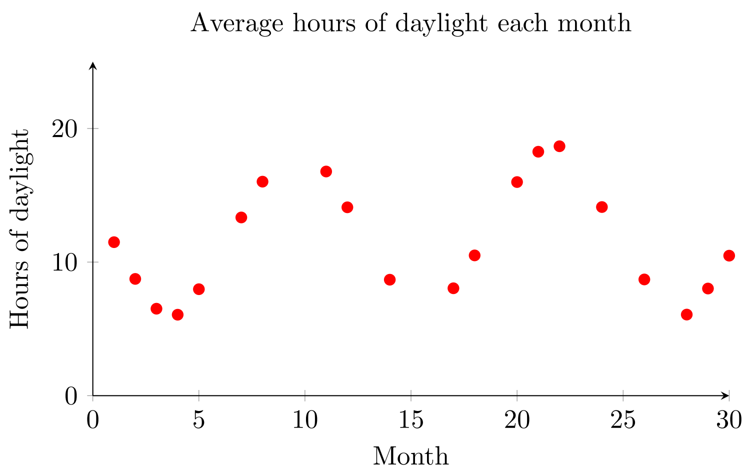
Modeling data

Nathan

May 23th 2023

Part I

1 Make a scatter plot of the data using appropriate technology



2 Give reasons why a quadratic model would not be appropriate for this data

The data appears to oscillate at a constant rate, meaning a quadratic would not be suitable for this kind of data. A quadratic would either only increase or decrease which would not model this data well. This data would require some sort of oscillatory function to model it instead of an exponential one like a quadratic.

3 Using your plot, can you recommend an appropriate function type to model this data? Give reasons why you would choose this type of function.

I believe that either a sin or cos function would work best due to the oscillatory nature of the data, as well as the data seeming to go from a high point (a max) to a low point (min). This is similar to how trigonometric functions work, making trigonometric functions suitable for modeling this data. $\cos$ would most likely be the better option as it is easier to match the max with the data.

4 Create an appropriate function to model this data. Explain how you found any relevant transformation coefficients.

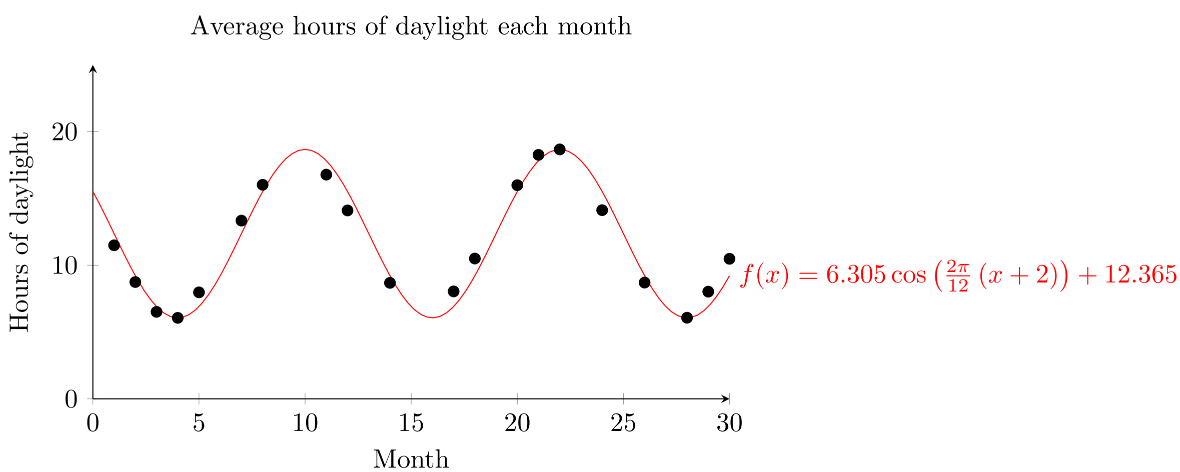
$$f(x) = 6.305 \cos\left(\frac{2\pi}{12}(x+2)\right) + 12.365 \tag{1}$$

I got this function starting off with the equation:

$$f(x) = \alpha \sin(b(x+c)) + d$$

where α is the amplitude, b the period, c is the phase shift, and d is the central axis. To get the amplitude, I got the maximum value and subtracted the lowest value from it and got 12.61 and then divided it by 2 to get 6.305, the amplitude. To get b, I knew that it would repeat every 12 months as hours of daylight repeats every 12 months, giving me $\frac{2\pi}{12}$ (2π is there to convert from radians). The phase shift, c, I got by looking at the highest point which was during month 22, and knowing that 24 is the closest multiple of 12 (the period), I knew that 2 would be the necessary phase shift to line up with the data. To get the central axis, I got the minimum point, 6.06, and added the amplitude, 6.305, which gave me 12.365.

5 Plot your function together with the data points



6 Explain/justify whether your solution makes sense in this context, detailing any limitations of your model

I believe that my solution makes sense as it pretty accurately models the data and accurately matches the maximums and the minimums. It also accurately repeats at the same rate as the data. However, the data points between the maxes and mins do not perfectly line up with the function and seem to change slightly ahead of my mathematical function.

7 Using your model, predict the daylight hours (missing data points) for June and July 2020. Justify the degree of accuracy of your response

For June (9th month):

$$\begin{aligned}
f(x) &= 6.305\cos\left(\frac{2\pi}{12}(x+2)\right) + 12.365 \\
f(9) &= 6.305\cos\left(\frac{2\pi}{12}(9+2)\right) + 12.365 \\
f(9) &\approx 17.8
\end{aligned}$$

For July (10th month):

$$\begin{aligned}
f(x) &= 6.305\cos\left(\frac{2\pi}{12}(x+2)\right) + 12.365 \\
f(10) &= 6.305\cos\left(\frac{2\pi}{12}(10+2)\right) + 12.365 \\
f(10) &\approx 18.7
\end{aligned}$$

I believe that 1 d.p. is a reasonable degree of accuracy based on the accuracy of my mathematical model and that no decimal points wouldn't be accurate enough. Anymore decimal points would be too accurate and unnecessary for number of hours as well as my mathematical model not being accurate enough to provide correct data at that level of accuracy.

Part II

The Summer Solstice is an annual astronomical phenomenon that brings the longest day of the year. It is related to the Earth's axial tilt with respect to the plane of its orbit around the Sun, also known as the ecliptic. Though the exact moment of solstice depends on different factors, you may assume for the purposes of this report, that it occurs on 21 June in the Northern Hemisphere. Likewise, the Winter solstice can be assumed to occur on 21 December, and results in the shortest day of the year in the Northern Hemisphere.

The actual daylight data for Oslo reveal that the longest day of 2021 occurred on the 21st of June with 18.83 hours of daylight. The shortest day in 2021 was 5.90 hours, occurring on the 21st of December.

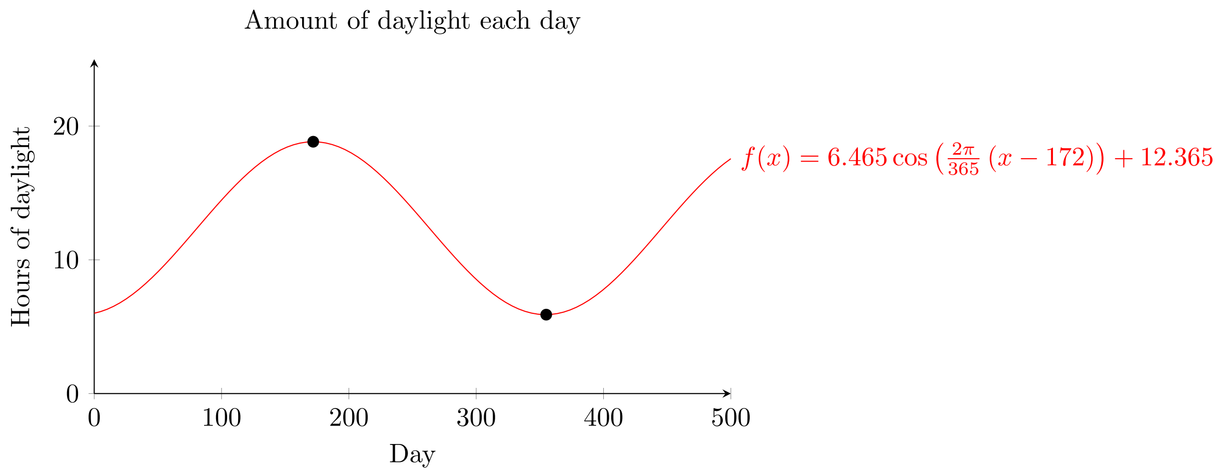
8 Using the information about Oslo given above, create a mathematical model to estimate the number of daylight hours on any day in 2021, where time is measured in days, instead of in months. You may ignore Leap years and assume that February contains 28 days. Show all your working.

Based on the data we have, we know that the data will repeat every year, or 365 days, and that the maximum point is at June 21st, the 172nd day of the year, with 18.83 hours of daylight, and the minimum point is at

December 21st, the 355th day of the year, with 5.90 hours of daylight.

From this, we know that the period is 365 days, max is 18.83, min is 5.90, amplitude is $\frac{18.83-5.90}{2} = 6.465$, the central axis is the amplitude plus the minimum, $6.465 + 5.90 = 12.365$, and the phase shift, using cosine, will be the amount of days after the start of the year until the maximum point, 172. Putting this all into an equation, we get:

$$f(x) = 6.465 \cos\left(\frac{2\pi}{365}(x-172)\right) + 12.365 \tag{2}$$



Part III

A list of locations throughout the world is given on the next page (in Part 4). The table includes the daylight hours for 21 June 2021 and 21 December 2021. Latitudes are also given; negative values for latitude indicate locations that are in the Southern Hemisphere. Length of daylight data is missing for Seoul and Bangkok. Let time in hours, minutes and seconds be expressed as hh:mm:ss.

9

9.1 Daylight for Seoul on 21 June was 14:45:37. Convert this to decimal form and verify that this is 14.76 when rounded to 2 d.p. Include this data point with the other data.

We start off by converting the seconds to decimal form. We will ignore hours for now. We have 45 minutes and 37 seconds. There are 60 seconds in a minute so we can just divide 37 by 60 and we get the decimal part of our answer: 0.6167, meaning that 45 minutes and 37 seconds is equal to 45.6167 minutes. Now to convert that to hours, we do the same process with the minutes, $\frac{45.6167}{60} = 0.760278$. We can now bring back the 14 hours from the original question to add to the 0.7603 to have an answer of 14.7603 which when rounded to 2 d.p. is equal to 14.76.

9.2 Daylight for Seoul on 21 December was 09:34:05. Verify that this is 9.57 when rounded to 2 d.p. Include this data point with the other data.

Starting by converting the seconds to minutes, we divide 5 by 60 which gives us 0.0833 minutes. We then convert the minutes, 34.0833, to hours by dividing it by 60, giving us 0.568056 hours, so in total its 9.56806 which when rounded to 2 d.p. gives us 9.57.

9.3 Daylight lengths for Bangkok on 21 June and 21 Dec of 2021 were 12:56:12 and 11:19:08 respectively. Convert these to decimal form and include with the other data.

$$\begin{aligned} \frac{12}{60} &= 0.2 \\ \frac{56.2}{60} &= 0.93667 \\ \therefore 12:56:12 &\approx 12.94 \end{aligned}$$

$$\begin{aligned} \frac{8}{60} &= 0.1333 \\ \frac{19.133}{60} &= 0.31889 \\ \therefore 11:19:08 &\approx 11.32 \end{aligned}$$

10 Formulate equations to model the number of daylight hours as a function of time in days during 2021 for Madrid, Veracruz, Singapore, Rio de Janeiro, and Christchurch in addition to your earlier result for Oslo.

Madrid:

$$2.885 \cos\left(\frac{2\pi}{365}(x-172)\right) + 12.175 \tag{3}$$

Veracruz:

$$1.155 \cos\left(\frac{2\pi}{365}(x-172)\right) + 12.135 \tag{4}$$

Singapore:

$$0.075 \cos\left(\frac{2\pi}{365}(x-172)\right) + 12.125 \tag{5}$$

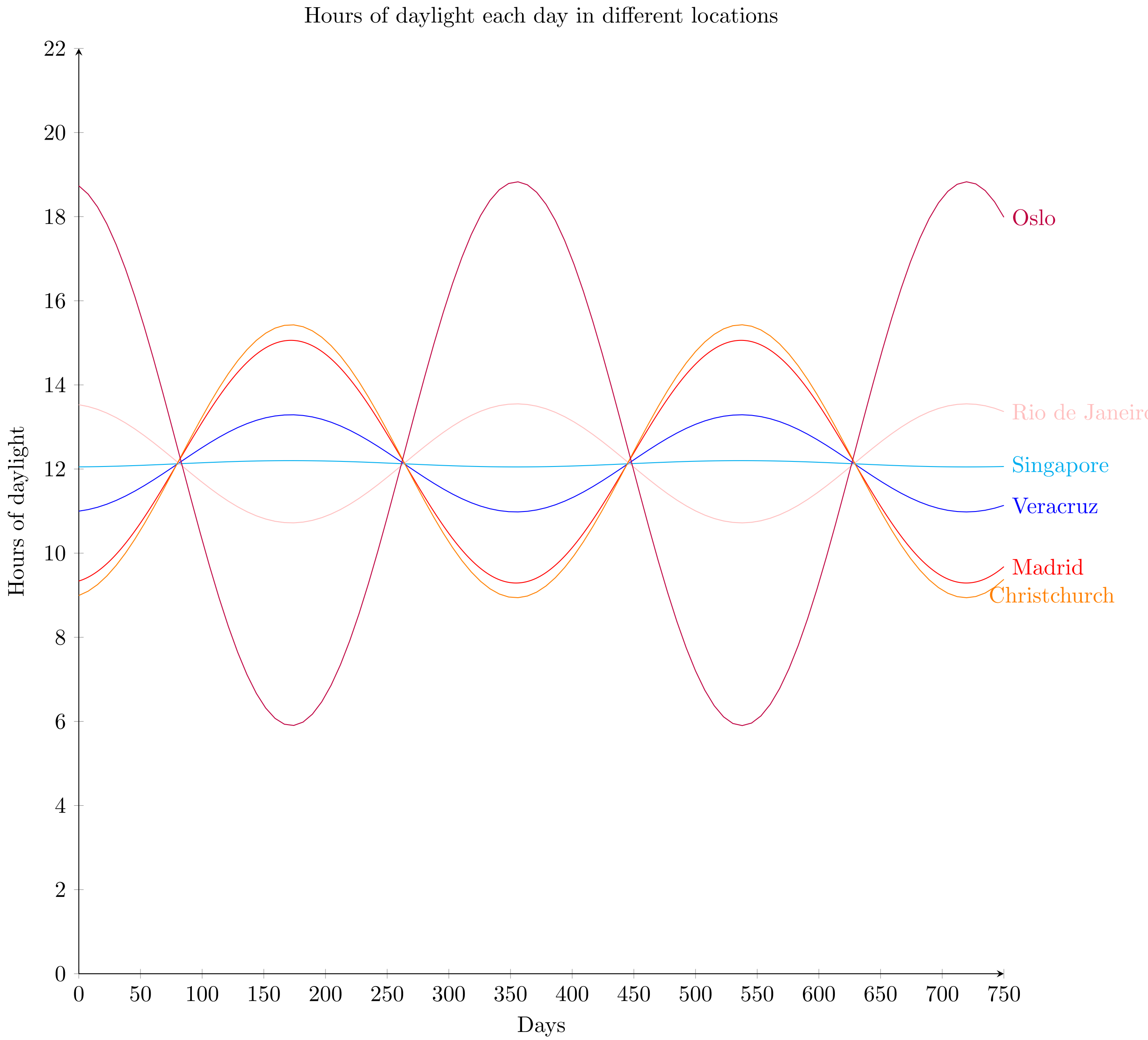
Rio de Janeiro:

$$1.415 \cos\left(\frac{2\pi}{365}(x-355)\right) + 12.135 \tag{6}$$

Christchurch:

$$3.245 \cos\left(\frac{2\pi}{365}(x-355)\right) + 12.185 \tag{7}$$

10.1 Plot these curves in an appropriate way so that you can discuss any similarities and differences there are between the graphs.



10.2 How do the components of the graphs relate to the real-world context?

As the latitude goes further away from 0, the amount of daylight each day will change more depending on the time of year. Locations found near the equator tend to have consistent amounts of daylight throughout the year. However, as the you go away from the equator, the amount of daylight varies more throughout the year. Additionally, negative latitudes mean that the max and min amounts of daylight is switched, meaning for negative latitudes, the max amount of daylight is December 21st and the min amount of daylight is June 21st. Singapore has a nearly consistent amount of daylight year round due to how close it is to the equator. Oslo has the greatest difference in most and least daylight due to its high latitude of 59.95°. Unlike the other graphs, Rio de Janeiro's and Christchurch's graph looks flipped due to it having a negative latitude.

10.3 Choose any of the cities you have modeled and use your equation to verify that the longest and shortest days of the year are correct (compare to the data table in Part 4)

The longest day in Madrid was 15.06 hours long and the shortest day was 9.29 hours long. Putting 172 (June 21st is the 172nd day) into my equation for Madrid ($2.885\cos(\frac{2\pi}{365}(x-172))+12.175$) gives 15.06 and putting in 355 (December 21st) gives 9.29 which are both correct.

11 Use your results from 10 to estimate the number of daylight hours there were for each location on 21 March 2021 and 21 September 2021. Comment on your results, and discuss whether they have a more general implication for other locations as well

	Hours of daylight			
Location	June 21 (172)	December 21 (355)	March 21 (80)	September 21 (264)
Christchurch	8.94	15.43	12.143	12.143
Madrid	15.06	9.29	12.138	12.138
Oslo	18.83	5.9	12.2815	12.2815
Rio de Janeiro	10.72	13.55	12.117	12.117
Singapore	12.2	12.05	12.124	12.124
Veracruz	13.29	10.98	12.12	12.12

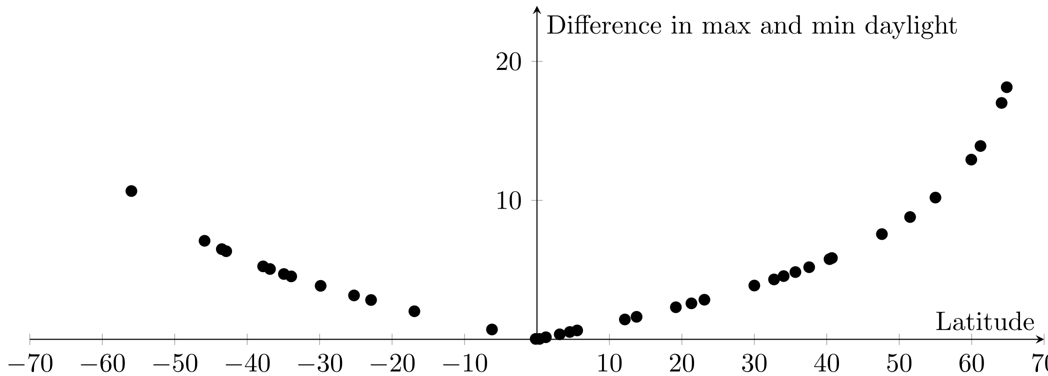
The days March 21 and September 21 have the same amount of daylight as each other. This is because March 21st (80th day) and September 21st (264th day) are both 92 days apart from 172 which is the max. This means that they'll have the same value as cosine is a symmetrical function at any max or min. This would also work at any location based on the mathematical model I have made and chosen.

Part IV

12 Using appropriate technology, utilize all 38 data points from the table to determine the nature of the relationship between latitude and the daylight hours profile for the various locations. You may attempt to formulate an equation, or use graphical methods.

Graphing the latitude and difference in most and least amount of daylight of each place gives us the following graph:

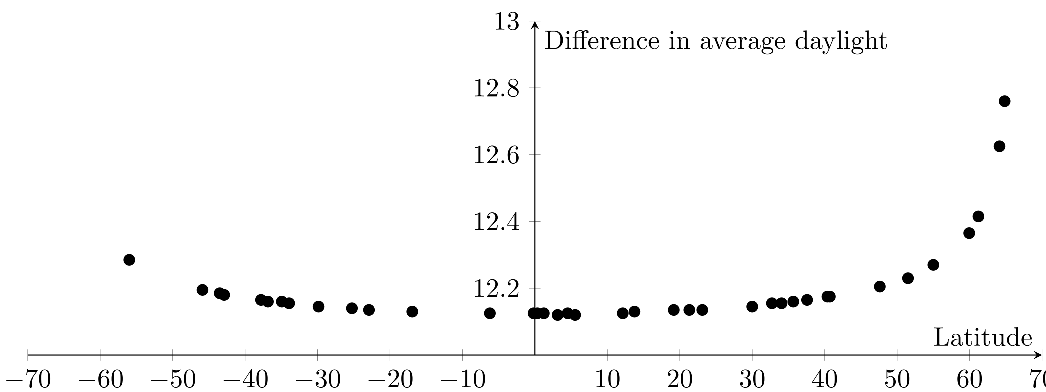
Difference in max and min daylight for different latitudes



This gives us a relationship similar to that of a quadratic. However, it seems to increase more suddenly at latitudes above 50, meaning that to model this relationship, more powers would be requires to accurately model this relationship with an equation. Putting these values into Desmos and using regression to a degree 6 polynomial gives us $0.396328 + 0.00522289x^2 - 0.00000174191x^4 + (3.546 \times 10^{-10})x^6$ as the equation that best models the relationship between the latitude and difference in max and min daylight.

Graphing the average amount of daylight (central axis) for different latitudes gives us the following graph.

Difference in max and min daylight for different latitudes



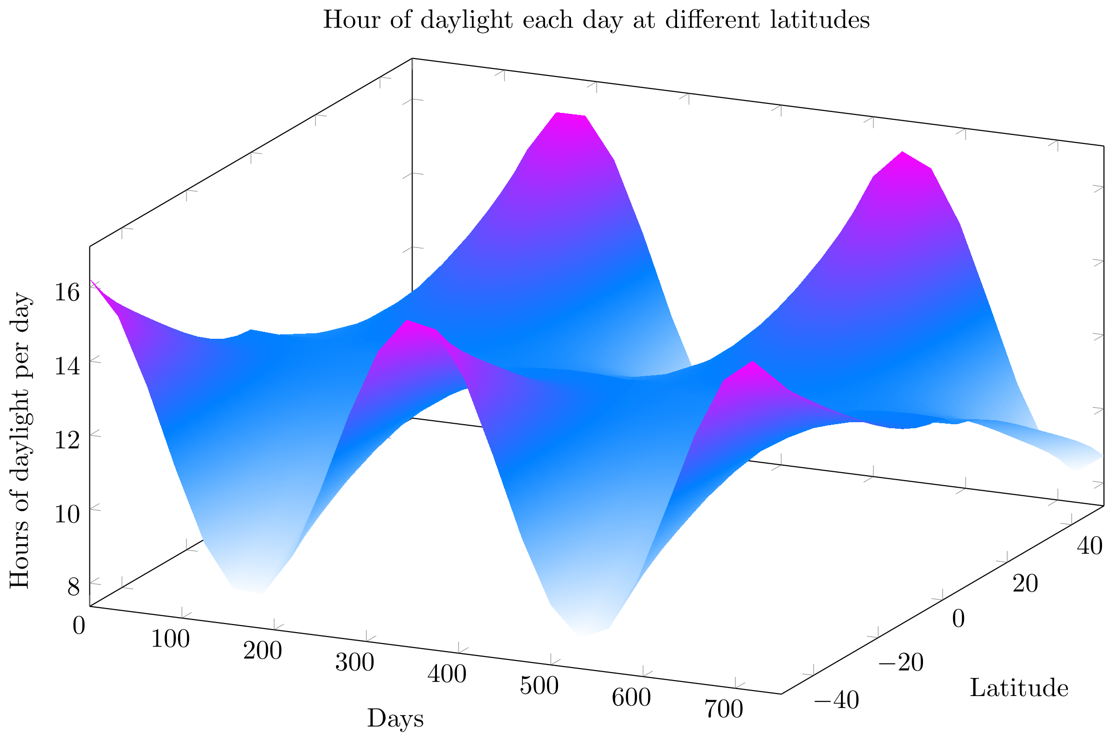
This gives us a fairly "flat" graph until around a latitude of 60 where the average amount of daylight suddenly increases. Once again putting this in Desmos and using a regression of the form $a + bx^2 + cx^6$ gives us the equation $12.125 + 0.000021898x^2 + 0.0000000000036647x^6$.

Additionally, when the latitude is negative, the day with the highest amount of daylight switches from June 21st to December 21st, and the day with the lowest amount of daylight switches from December 21st to June 21st. This can be expressed as an equation in the following form: $\frac{x}{|x|}$, where x is the latitude. This gives us $+1$ if x is positive and -1 if x is negative.

Putting this all into an equation based on both latitude and days after January 1st gives us the following:

$$\frac{0.396 + 0.0052y^2 - 1.74 \cdot 10^{-6}y^4 + 3.55 \cdot 10^{-10}y^6}{2}\cos\left(\frac{2\pi}{365}\left(x - \left(263.5 - 91.5\frac{y}{|y|}\right)\right)\right) + (12.125 + 2.19 \cdot 10^{-5}y^2 + 3.67 \cdot 10^{-12}y^6) \tag{8}$$

Where x is the number of days after January 1st and y is the latitude. Graphing this gives us the following:



13 Crater Lake in Oregon lies at a latitude of 42.94° North latitude.

13.1 Use your findings from 12 to predict the number of daylight hours there were at Crater Lake on 21 June 2021. Explain how you arrived at your answer. Explain/justify the accuracy of your answer and whether your answer makes sense in the context of this investigation.

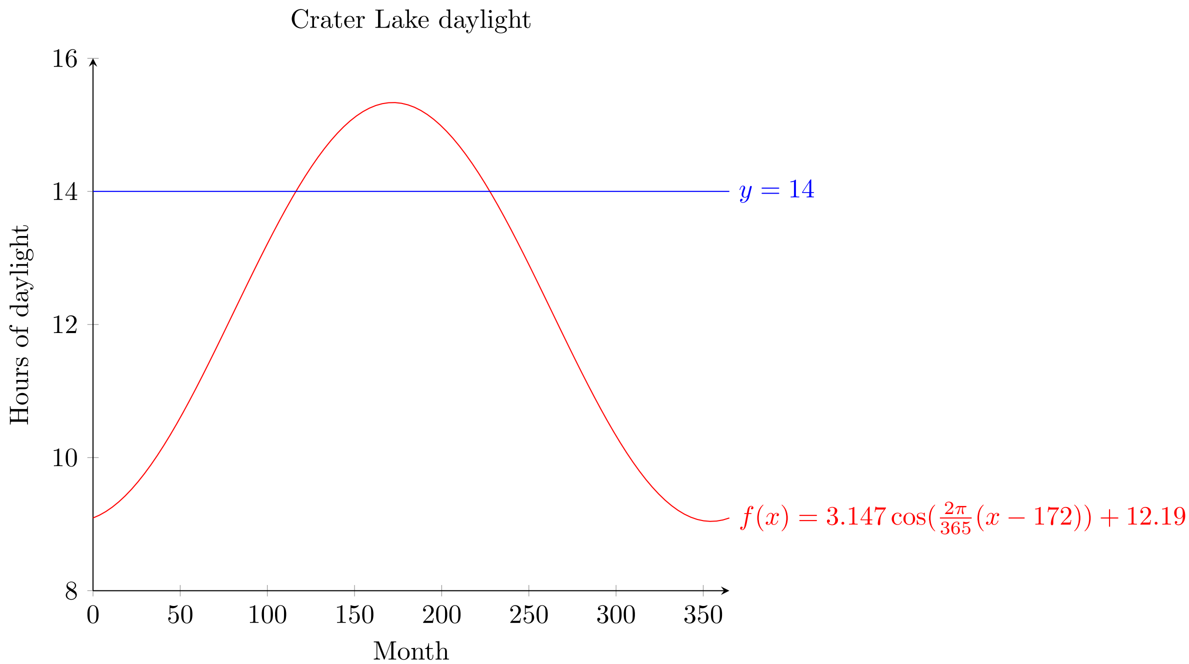
Using the equation I got from 12, I can put in 172 for x (June 21 is 172 days after January 1st) and 42.94 for y (the latitude). This gives us a result of approximately 15.352. I believe this value makes sense as Madrid which lies at a latitude of 40.38 had 15.06 hours of daylight on June 21st. Since Crater Lake lies at a slightly higher latitude, it should get slightly more daylight than Madrid, which is the case with my prediction of 15.352.

13.2 Estimate the number of days at Crater Lake that had more than 14 hours of daylight in 2021. Explain how you arrived at your answer.

Method 1: Graphical

Using the formula I found earlier and keeping y (the latitude) constant at 42.94, we get the following equation:

$$3.14688365\cos\left(\frac{2\pi}{365}(x-172)\right)+12.18838602 \tag{9}$$



Graphing this gives:
Looking at this, we can approximate the number of days with over 14 hours of daylight to be from 120 to approximately 230 which would mean 110 days had more than 14 hours of daylight. However this in highly inaccurate and would only be useful for getting a rough idea of how many days had more than 14 hours of daylight.

Method 2: Algebraic

Using the equation $3.14688365\cos(\frac{2\pi}{365}(x-172)) + 12.18838602$ that models the daylight at Crater Lake, we can make it equal 14, which would give us 2 points and we can find the different between those 2 points to see how many days had over 14 hours of daylight.

$$14 = 3.14688365\cos(\frac{2\pi}{365}(x-172)) + 12.18838602$$
$$1.81161398 = 3.14688365\cos(\frac{2\pi}{365}(x-172))$$
$$0.57568508 = \cos(\frac{2\pi}{365}(x-172))$$
$$\cos^{-1}(0.57568508) = \frac{2\pi}{365}(x-172)$$
$$\frac{365}{2\pi}\cos^{-1}(0.57568508) = x - 172$$
$$\frac{365}{2\pi}\cos^{-1}(0.57568508) + 172 = x$$
$$x = 116.386, 227.614$$

Taking the difference between these 2 points gives us 111, which is the amount of days with over 14 hours of daylight. This is quite close to my estimation using the graph of 110 days.

14 Compare your Crater Lake result with the data for Hobart in the data table. Comment on your observations. Can you make a generalization based on your observations?

Hobart has a similar latitude to Crater Lake except negative. This means that both graphs have similar maxes and mins except that they are out of phase. The day with the most daylight in Crater Lake is the 172nd day (June 21st) and the day with the least is the 355th day (December 21st). Hobart's day with the most daylight is the 355th day and lowest is the 172nd day which is the exact opposite of Crater Lake. Negative latitudes are exactly out of phase compared to their positive counterparts.

15 The northernmost town in the world is Longyearbyen, Norway. This town lies at 78.22° North latitude. Use your model to determine the number of hours of daylight there were on 21 June 2021. Explain/justify whether your answer makes sense in this context

If I input 78.22 as y and 172 as x into my equation earlier, I would get a result of 37.28 hours of daylight each day, which would be impossible. This result shows that my formula should always give answers between 0 and 24, and any values above or below this would represent always having daylight or no daylight at all. This result also shows that my formula can be unreliable at large latitudes such as the 78.22° of Longyearbyen. I believe that an answer of 24 hours of daylight on June 21st would make sense due to the extremely high latitude of Longyearbyen. The location with the highest latitude that we got in our data was Reykjavik with a latitude of 64.13° and having 21.13 hours of daylight, which is near to having 24 hours of daylight. Longyearbyen has a latitude 14.09° higher than Reykjavik which would imply more hours of daylight on June 21st. Therefore I think an answer of 24 hours of daylight makes sense.

16 Reflect on what you have learned in this investigation. What is required to develop a good mathematical model? How certain can we be that our derived model actually represents the situation accurately? What are some of the limitations of the models you have constructed?

This investigation has taught me how to use mathematical functions, specifically trigonometric functions, to model real world data. To develop good mathematical models, the model should be accurate and accurately model the data given. We can test out certain data points and compare them to our mathematical model such as the maximums or minimums to test how accurate our mathematical model is. We also should test and see if our result seems to make sense in the context of the question or situation. The models I have made are only based around the max and min points meaning for the other points, it could be slightly inaccurate. Also, trig functions do not allow for a lot of control so the data must increase and decrease at a similar rate to that of a trig function like sin or cos.

City	Latitude	21-Jun	21-Dec
Accra	5.55	12.44	11.8
Adelaide	-34.93	9.81	14.51
Anchorage	61.22	19.37	5.46
Asuncion	-25.25	10.56	13.72
Auckland	-36.85	9.63	14.69
Bangkok	13.75	12.94	11.32
Bogota	4.53	12.39	11.86
Cairns	-16.92	11.12	13.14
Calro	30	14.08	10.21
Cape Horn	-55.98	6.95	17.62
Cape Town	-33.92	9.89	14.42
Christchurch	-43.5	8.94	15.43
Dunedin	-45.88	8.65	15.74
Durban	-29.86	10.22	14.07
Fairbanks	64.84	21.83	3.69
Havana	23.11	13.56	10.71
Hobart	-42.89	9.01	15.35
Honolulu	21.32	13.43	10.84
Jakarta	-6.21	11.77	12.48
Kampala	0.35	12.14	12.11
Kuala Lumpur	3.14	12.3	11.94
London	51.5	16.63	7.83
Londonderry	55	17.37	7.17
Los Angeles	34.05	14.43	9.88
Madrid	40.38	15.06	9.29
Managua	12.13	12.84	11.41
Melbourne	-37.8	9.54	14.79
New York	40.71	15.1	9.25
Oslo	59.95	18.83	5.9
Quito	-0.18	12.11	12.14
Reykjavik	64.13	21.13	4.12
Rio de Janeiro	-22.9	10.72	13.55
San Diego	32.72	14.31	10
Seattle	47.61	15.99	8.42
Seoul	37.57	14.76	9.57
Singapore	1.23	12.2	12.05
Tokyo	35.67	14.58	9.74
Veracruz	19.17	13.29	10.98