A Population Model and Temperature Model for XXXXXX, XX.

XXXXXXXX XXXX
Math 251
Month, Day, Year.

Introduction

This project concerns modeling both population and temperature data we were able to find at `wikipedia.org` about the author's hometown of XXXXXX, XX.

First, however, we shall tell the reader a little bit about XXXXXX. While it is not a large town, there is a lot history there. It started out as a mining camp by migrants seeking their fortune. They named the town XXXXXX since the area's terrain reminded them of their homeland [3].

It is a thoroughfare for many adventure seekers as it is a gateway to many sought-after destinations including XXXXXXXX National Park, XXXX Lake, and The XXXXXXX Wilderness to name a few. In the 20th century, timber took the place of mining as a principal economic driver of the region. However, since the 1990s, timber has been on the decline. Today, tourism fuels a large part of the economy of the region. Moreover, urban flight has caused the population to increase at historically unprecedented rates This was especially true in the 1980s, and the population continues to grow.

The above image is a recent photograph of XXXXXX, XX [3].

Now that the reader is more familiar with the author's hometown, we shall describe the main task of this project. We were asked to estimate the following:

1. The size of the population of XXXXXX, XX in 2020.
2. The rate at which the population is growing (or shrinking) in 2020.
3. The rate at which the average high temperature changes in May.
4. The rate at which the average high temperature changes in October.
5. The hottest time of the year as predicted by the temperature model.

Our first order of business was to log on to `https://wikipedia.org` and see what population and temperature data we could find. Both temperature and population data were available, and we created models for both which were able to answer all five of the above questions.

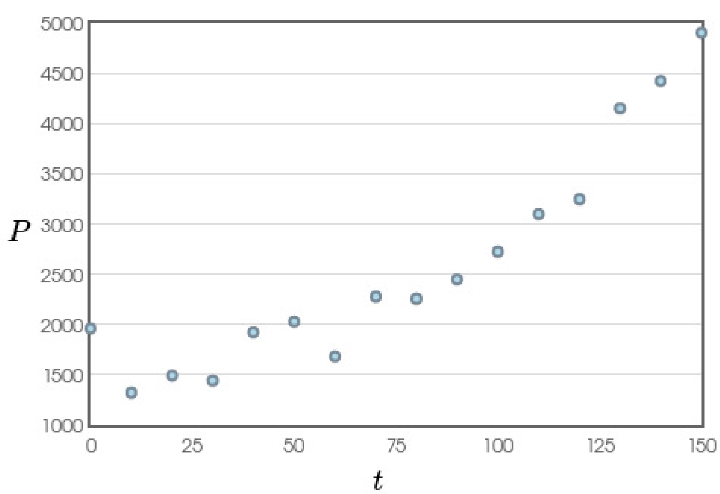
Population

We shall first analyze the population data we found in order to predict the size and rate of population growth of XXXXXX, XX. We used demographic data provided by Wikipedia [3], which is given below.

Historical population		
Census	Pop.	%±
1860	1,960	—
1870	1,322	−32.6%
1880	1,492	12.9%
1890	1,441	−3.4%
1900	1,922	33.4%
1910	2,029	5.6%
1920	1,684	−17.0%
1930	2,278	35.3%
1940	2,257	−0.9%
1950	2,448	8.5%
1960	2,725	11.3%
1970	3,100	13.8%
1980	3,247	4.7%
1990	4,153	27.9%
2000	4,423	6.5%
2010	4,903	10.9%
Est. 2016	4,823 [4]	−1.6%
U.S. Decennial Census[10]		

We note that this same data is also available at `https://factfinder.census.gov`.

To analyze the above data, we used the website `http://holt.blue` [2], to generate our plots and to calculate the statistical models we will use to make our prediction for the population of XXXXXX, XX in 2020.

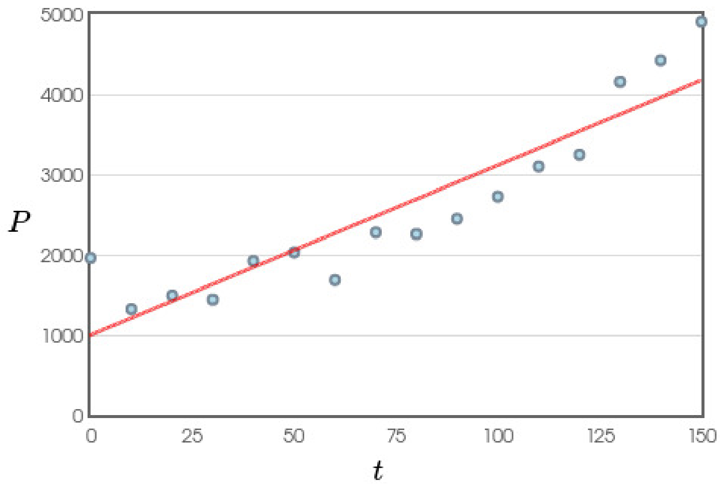


In the above figure, P denotes the number of people and t denotes number of years since 1860.

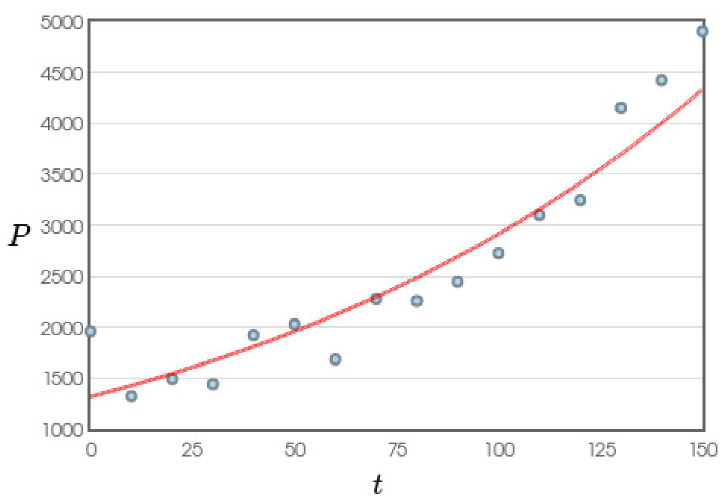
Since the data at first glance look like they could be linear, we first fitted a linear model,

$$P = 21.1735 \cdot t + 998.4853,$$

which is plotted as the red line along with the population data.



Noticing that the model consistently underpredicts at the ends, and overpredicts in the middle, we decided to try an exponential model instead. Running a least-squares regression on the data we obtained the following figure which contains both the data and the exponential curve of best fit in red.



The equation for the exponential curve of best fit to four decimal places is

$$P = 1318.3462 \cdot (1.008)^t.$$

Comparing both models visually, the reader can see that the exponential model is a much better fit to the data than the linear model.

We note that we also tried a quadratic fit, but the exponential fit was superior on all accounts. Namely, exponential models are a natural choice for populations growing at a fixed percentage per time period [1].

At this point, we may now answer the questions about the overall rate of growth of XXXXXX,XX and its predicted population for the year 2020:

Overall, XXXXXX, XX is growing at a rate of 0.8% per year.

In the year 2020, that is $t = 160$ for our model, we predict that XXXXXX, XX will have 4718 residents. We obtained this number by evaluating the model at $t = 160$: $P(160) = 1318.3462(1.008)^{160} \approx 4717.53$ and then rounding to the nearest whole number.

We now estimate the rate at which the population of XXXXXX, XX will be growing in 2020. To find the rate of change of population with respect to time, we find the derivative of P with respect to t of our exponential model:

$$\frac{dP}{dt} = \frac{d}{dt}(1318.3462 \cdot (1.008)^t)$$

$$= 1318.3462 \cdot \frac{d}{dt}(1.008)^t$$

$$= 1318.3462 \cdot \ln(1.008) \cdot (1.008)^t$$

$$= 10.5048 \cdot (1.008)^t.$$

The last is an approximation to four decimal places. Evaluating the derivative function $P'(t) = 10.5048 \cdot (1.008)^t$ at $t = 160$, we obtain $P'(160) \approx 37.5901$ people per year.

We interpret the above to mean that in the year 2020, XXXXXX, XX will be growing at a rate of about 38 people per year.

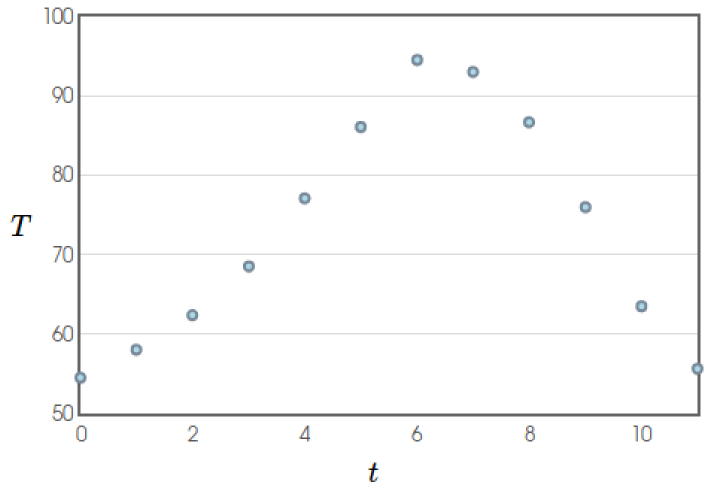
Temperature

We now move on to the average-high temperature of XXXXXX, XX. The data we used was again obtained from `wikipedia.org` [3], who in turn used data from The Weather Channel and The Desert Research Institute [4]. The data are displayed below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Record high °F (°C)	75 (24)	78 (26)	84 (29)	92 (33)	103 (39)	113 (45)	113 (45)	110 (43)	108 (42)	100 (38)	89 (32)	81 (27)	113 (45)
Average high °F (°C)	54.5 (12.5)	58.0 (14.4)	62.4 (16.9)	68.5 (20.3)	77.1 (25.1)	86.1 (30.1)	94.5 (34.7)	93.0 (33.9)	86.7 (30.4)	76.0 (24.4)	63.5 (17.5)	55.6 (13.1)	73.0 (22.8)
Average low °F (°C)	33.4 (0.8)	35.5 (1.9)	38.0 (3.3)	41.7 (5.4)	46.7 (8.2)	52.7 (11.5)	58.7 (14.8)	57.4 (14.1)	52.7 (11.5)	45.2 (7.3)	38.1 (3.4)	33.8 (1.0)	44.5 (6.9)
Record low °F (°C)	13 (−11)	15 (−9)	20 (−7)	24 (−4)	24 (−4)	33 (1)	39 (4)	40 (4)	33 (1)	25 (−4)	21 (−6)	8 (−13)	8 (−13)
Average precipitation inches (mm)	6.13 (156)	5.55 (141)	5.10 (130)	2.77 (70)	1.27 (32)	0.33 (8.4)	0.04 (1.0)	0.08 (2.0)	0.39 (9.9)	1.62 (41)	3.46 (88)	5.39 (137)	32.14 (816)
Average snowfall inches (cm)	2.3 (5.8)	1.0 (2.5)	0.5 (1.3)	0.2 (0.51)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0.7 (1.8)	4.7 (12)

We will be modeling the average high temperature displayed in the second row of the above table.

Again, we analyzed our data using software available at `http://holt.blue` [2]. The scatter plot of the above data is given below.



We note that temperature (in degrees Fahrenheit) is denoted T and time (in months) is denoted t.

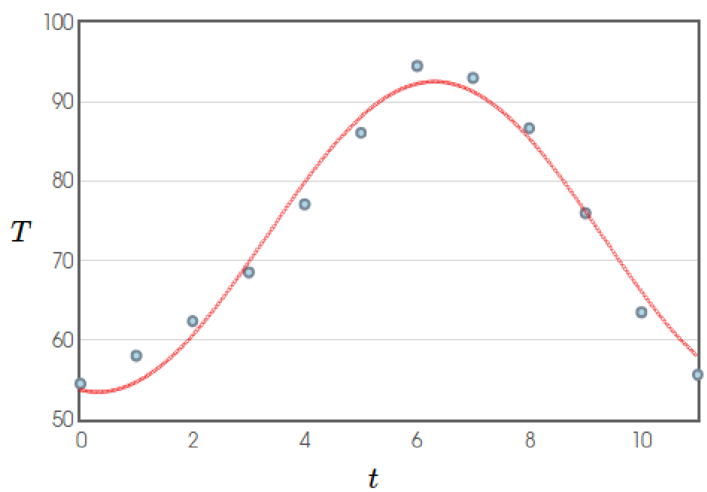
Due to the naturally periodic nature of the data, we decided to model this data set using periodic functions. However, experimenting other models, we obtained a very good fit as well using a polynomial model. For this project we shall present both since they both seem appropriate.

The Sinusoidal Temperature Model

Performing a sinusoidal regression on the above data, we obtain the following curve of best fit

$$T = 19.5379 \cdot \sin(0.5236t - 1.7337) + 72.9917$$

which is plotted below with the data.



With our formula for average temperature as a function of time, we now determine the rate at which the average high temperature of XXXXXX, XX changes during the beginning of the months of May and October.

We first compute the derivative:

$$\frac{dT}{dt} = \frac{d}{dt}(19.5379 \cdot \sin(0.5236t - 1.7337) + 72.9917)$$

$$= 19.5379 \cdot \frac{d}{dt}(\sin(0.5236t - 1.7337) + 72.9917))$$

$$= 19.5379 \cdot \cos(0.5236t - 1.7337)\tfrac{d}{dt}(0.5236t - 1.7337)$$

$$= 19.5379 \cdot \cos(0.5236t - 1.7337) \cdot 0.5236$$

$$= 10.23 \cdot \cos(0.5236t - 1.7337).$$

We now use the derivative function to compute the rate of change of the average high temperature. Since our year starts at time $t = 0$, we consider the beginning of May to coincide with $t = 4$ months. Similarly, the beginning of October is corresponds to $t = 9$ months.

In terms of the above, we now evaluate and interpret $T'(4)$ and $T'(9)$.

$T'(4) \approx 9.57$. We interpret this to mean that in XXXXXX,XX, the average high temperature is increasing at a rate of 9.57 degrees Fahrenheit per month at the beginning of May.

$T'(9) \approx -10.09$. This means that the average high temperature is decreasing by about 10.09 degrees Fahrenheit per month at the beginning of October.

We shall now find the time at which the simusoidal model predicts maximum average high temperature by setting the derivative equal to zero. From the above computations we have

$$\frac{dT}{dt} = 0$$

$$\Rightarrow \quad 10.23 \cdot \cos(0.5236t - 1.7337) = 0$$

$$\Rightarrow \quad \cos(0.5236t - 1.7337) = 0$$

$$\Rightarrow \quad 0.5236t - 1.7337 = \tfrac{n\pi}{2}$$

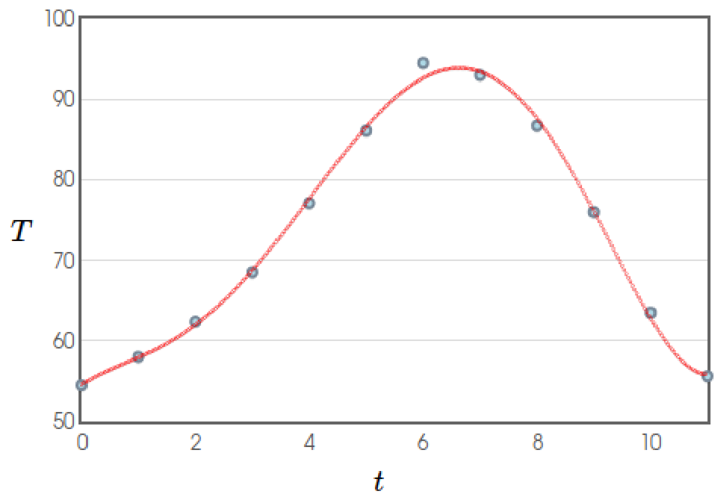
where n is an odd number. Then $t = \dfrac{n\pi/2 + 1.7337}{0.5236}$, where n is an odd number, represents the local extrema of our function T. Since we know that $0 \le t < 12$, we consider only values of n for which the previous inequality holds: $n = -1$, and $n = 1$. Then the extreme points are $t = 0.31$ months, and $t = 6.31$ months. We see that the former corresponds to the minimum value and the latter corresponds to the maximum value of T on $[0, 12)$. Thus, the maximum average high temperature of XXXXXX, XX happens at about one third of the way into July. This is no surprise as most locations in the northern hemisphere exibit similar weather patterns. The maximum average temperature at this is then approximately $T(6.31) \approx 92.5$ degrees Fahrenheit.

The Polynomial Temperature Model

We now do a similar analysis using our polynomial model. Performing a polynomial regression on our data, we obtained the following fifth-degree best-fit polynomial

$$T = 0.0078t^5 - 0.1843t^4 + 1.2416t^3 - 2.202t^2 + 4.4965t + 54.5822$$

which is graphed below in red along with the data.



Performing similar computations as above, we see that

$$\begin{aligned}
\frac{dT}{dt} &= \frac{d}{dt}(0.0078t^5 - 0.1843t^4 + 1.2416t^3 - 2.2020t^2 + 4.4965t + 54.5822) \\
&= 0.0078 \cdot 5t^4 - 0.1843 \cdot 4t^3 + 1.2416 \cdot 3t^2 - 2.2020 \cdot 2t + 4.4965 \\
&= 0.039t^4 - 0.7372t^3 + 3.7248t^2 - 4.404t + 4.4965.
\end{aligned}$$

Now, $T'(4) \approx 9.28$. Thus, our polynomial model tells us that the average high temperature of XXXXXX increases by 9.28 degrees per month at the beginning of May. We note that this is a similar estimate to

Also, $T'(9) \approx -14.97$. Thus, our polynomial model estimates that the average high temperature is decreasing by 14.97 degrees per month at the beginning of October. We note that this is a much steeper decline than the sinusoidal model predicts. However, since the polynomial curve seems to be a better fit, we take more stock in the polynomial estimate.

Finally, we determine the maximum value of our polynomial model on $[0, 12)$. Again, we set the derivative equal to zero:

$$\begin{aligned}
& \frac{dT}{dt} = 0 \\
\Rightarrow \quad & 0.039t^4 - 0.7372t^3 + 3.7248t^2 - 4.404t + 4.4965 = 0
\end{aligned}$$

We are not going to solve this last equation by algebraic methods, so we will use Newton's method to find the root near $t = 6$ (as seen from our plots).

To find where the local extrema of our polynomial model $T(t)$ are located, we will iterate the recursive rule given by Newton's Method [1] to two decimal places

$$\begin{aligned} t_{n+1} &= t_n - \frac{T'(t_n)}{T''(t_n)} \\ &= t_n - \frac{0.039t_n^4 - 0.7372t_n^3 + 3.7248t_n^2 - 4.404t_n + 4.4965}{0.156t_n^3 - 2.2116t_n^2 + 7.4496t_n - 4.404} \end{aligned}$$

with $t_0 = 6$ as our initial guess. The above quickly gives us our two decimal places of accuracy: $t \approx 6.56$ is our maximum value on $[0, 12)$. In terms of the problem situation, this means that the maximum average high temperature occurs a little over half way through July with a maximum temperature of 93.3 degrees.

Comparing $t = 6.56$ to our other estimate, $t = 6.31$, we see that it occurs a little bit later. As for which estimate is better, it is difficult to say since the rather large data value at $t = 6$ months complicates the analysis. But since both models give a reasonable approximation, we would suggest keeping both in mind as reasonable estimates. We also not that the maximum average high temperature predicted by both models are slight under-predictions.

Summary and Conclusions

We have modeled both the population and average high temperature of XXXXXX, XX. We explored multiple models for each in order to find the model which best represents the data. We concluded that an exponential model was the best and most natural model for describing population growth. We saw that by 2020, XXXXXX, XX will have approximately 4718 residents and will be growing at a rate of about 38 people per year.

As for temperature, we constructed two models: a sinusoidal model and a polynomial model. The reason for the two model is that both made sense in their own way: the sinusoidal model seems the most natural for cyclic, periodic data like temperature. However, a fifth-degree polynomial, although more ad hoc and less natural, modeled the data much more closely.

For the beginning of May, the sinusoidal model predicts a rate of increase in average high temperature of 9.57 degrees Fahrenheit per month; the polynomial model predicts 9.28 degrees per month. So both models agree on this point.

On the other hand, for the beginning of October, the sinusoidal model predicts a rate of decrease in average high temperature of 10.09 degrees Fahrenheit per month, whereas the polynomial model predicts a rate of decrease of 14.97 degrees per month. The disagreement between the two models is close to 5 degrees, which is fairly substantial. However, since the polynomial fit is much better fit to the data, we trust the polynomial estimate more.

The sinusoidal model predicted the hottest time of year to be at $t = 6.31$ months, and the polynomial model predicted $t = 6.56$ months. We take both to be reasonable estimates (somewhere in mid-July).

References

[1] Gilbert Strang, *Calculus: Volume 1*, Open Stax: `https://openstax.org/details/books/calculus-volume-1`, 2018.

[2] Holt.Blue Statistical Software Suite, available at `http://holt.blue`

[3] Wikpedia Data for XXXXXX, XX: https://en.wikipedia.org/wiki/XXXXXX,_XX

[4] Desert Research Institute Data, available at https://wrcc.dri.edu/cgi-bin/cliMAIN.pl?ca8353