The Population of Japan and Swaziland

Type 2 Portfolio
Rachel Timmons
Lee’s Summit West High School
IB Mathematics SL
5/18/2009

Population models are formulas that one can use to calculate the future population of a country based on past growth. These growths can sometimes be shown exponentially. In this portfolio, I will be finding population models for the countries of Swaziland and Japan.

I will begin with Swaziland. Using the data in the following table, I will find an exponential function algebraically to describe the population based on the year. A possible format for this function is $y = ab^{x}$, and this is the one I'll be basing my model on. The following data was taken from www.library.uu.nl/wesp/populstat/afica/swazilac.htm . The populations shown are estimates.

Year	Population (thousand)	Year	Population (thousand)
1911	100.0	1960	330.0
1921	112.8	1970	422.0
1927	122.0	1980	565.0
1936	156.7	1990	751.0
1944	171.3	2000	1083.3
1950	264.0	2005	1317.0

In order to make the data easier to work with, I'm going simplify the years according to 1911 being year 1.Therefore, year 1921 will be represented as year 11, 1927 as 17, 1936 as 26, etc. These new expressive values will be used as my x values and the population as my y values. I placed these values in the table below.

X	*Y*	*X*	*Y*
1	100.0	50	330.0
11	112.8	60	422.0
17	122.0	70	565.0
26	156.7	80	751.0
34	171.3	90	1083.3
40	264.0	95	1317.0

My process will be to first find the rate, b. I am going to obtain this value by finding the difference in population divided by the change in years. Then, I'll take that quotient and add it to the smaller of the two populations; that sum will then be divided by that same smaller population. I will do multiple trials of this process using only data adjacent to each other. The results of these trials will then be averaged. I'll demonstrate this process using the first two consecutive values in the data table with the use of my TI-84 Plus Silver Edition Graphics Display Calculator (GDC).

To begin with, I will use the two population values of 100 and 112.8. Since the values are taken from the year 1911 and 1921 the change will be 10.

```
(112.8-100)/10
                1.28
Ans+100
              101.28
Ans/100
              1.0128
```

1.01 is the approximate rate of change in population between the years 1911 and 1921.

I will keep repeating this process using each group of two successive data values in the table. However, I will only illustrate the next two computations which are subsequently displayed using my GDC.

Using years 1921 and 1927:

```
(122.0-112.8)/6
         1.533333333
Ans+112.8
         114.3333333
Ans/112.8
         1.013593381
```

I conclude, 1.01 is the approximate rate of change.

Furthermore, from 1927 to 1936:

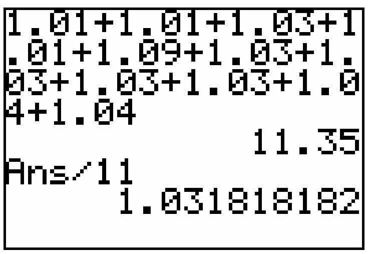
```
(156.7-122.0)/9
         3.855555556
Ans+122.0
         125.8555556
Ans/122.0
         1.031602914
```

1.03 is the approximate rate of change.

Continuing this same process for the remaining data, I obtained the following values shown in the table below. I rounded each approximate rate of change to the nearest tenth place

Range of Years	Approximate Rate of Change in Population	Range of Years	Approximate Rate of Change in Population
1911-1921	1.01	1960-1970	1.03
1921-1927	1.01	1970-1980	1.03
1927-1936	1.03	1980-1990	1.03
1936-1944	1.01	1990-2000	1.04
1936-1950	1.09	2000-2005	1.04
1950-1960	1.03		

Now I'll find the average approximated rate of change using my GDC.



I will round this value to the nearest hundredth place which will enable me to state that my rate, *b*, is equal to 1.03. In other words, my population model thus far is $y = a(1.03)^x$.

To find my *a* value I will plug in each set of corresponding values from the original table into the *x* and *y* variables in my model, $y = a(1.03)^x$, and solve algebraically for *a*. After finding an *a* value for each set of data I will again execute the simple process of averaging to find a value that will take the place of *a* in the population model. As previously stated in the beginning of this portfolio I will be using the population measured in thousands for my *y* values and their expressive year values as my *x*. I will exemplify a few of these procedures below.

For my first demonstration I will be using the data from year 1 which has the population of 100.0 (thousand).

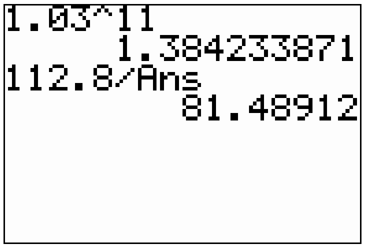
$100.0 = a(1.03)^1$

$100.0 = a(1.03)$

$\frac{100.0}{1.03} = a$

$a \approx 97.087$ (This value will be utilized in finding the averaged *a* value.)

I will complete another demonstration now using the data from year 11 which has the population of 112.8 (thousand) with the use of my GDC. The problem will be set up as follows: $112.8 = a(1.03)^{11}$.

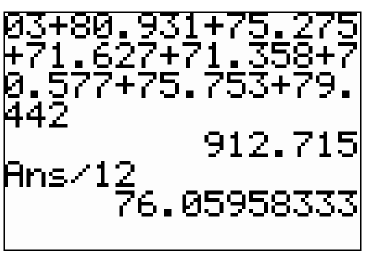


$a \approx 81.489$

When I continued this method of calculation with my GDC for the rest of the data I found the following answers in the data table below. I rounded each *a* value to the nearest thousandth place.

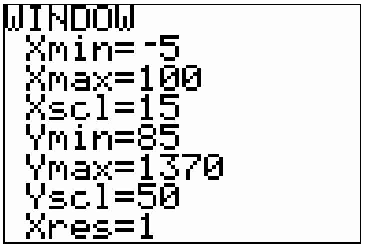
X	Y	a	X	Y	a
1	100.0	97.087	50	330.0	75.275
11	112.8	81.489	60	422.0	71.627
17	122.0	73.812	70	565.0	71.358
26	156.7	72.661	80	751.0	70.577
34	171.3	62.703	90	1083.3	75.753
40	264.0	80.931	95	1317.0	79.442

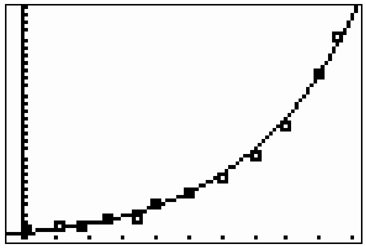
Finally, I'll find the average of these *a* values using my GDC.



Rounding these values to the nearest hundredth place results in $a \approx 76.06$.

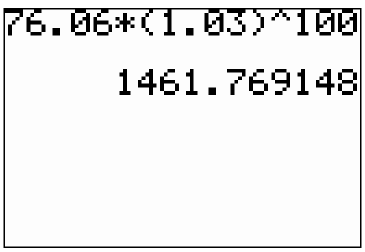
Therefore, the population model of Swaziland is $y = 76.06(1.03)^x$, where *x* is equal to the years starting with 1911 as the first year and *y* equals the population of that year in thousands. In comparison to the data, I believe this model fits fairly well. When I created a graphical model of the actual population data and compared and contrasted it with the graph of my model, I observed a pretty close relation in the trend and pattern. I made this graph with my GDC shown below.



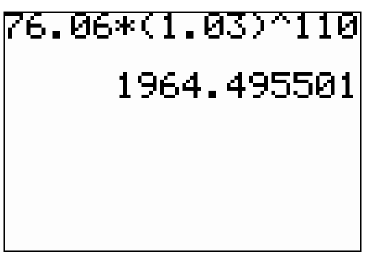


(The square dots represent a scatter plot of the population data, and the curved line represents the graph of my model)

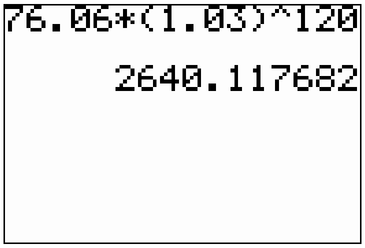
I am now going to use this model to predict future populations of this country. I will then compare these predictions to ones I will find from other sources and observe how reliable my model is. The years I will be using are the following: 2010, 2020, and 2030.The way I simplified the other year values will be applied to these, thus, 2010 will now be represented by 100, 2020 as 110, and 2030 as 120. In order to find these predictions I will plug the expressive year values in for x in my population model, y=76.06(1.03)^x which will then give me the possible population for that year in thousands. I will implement these calculations using my GDC.



If x=100 then y≈1461.769 (thousand). Thus, the projected population of Swaziland in 2010 is 1,461,769 people.



If x=110 then y≈1964.496 (thousand). Therefore, the possible population of Swaziland in 2020 is 1,964,496 people.



If x=120 then y≈2640.118 (thousand). Thus, the estimated population of Swaziland in 2030 is 2,640,118 people.

According to the CIA Factbook https://www.cia.gov/library/publications/the-world-factbook/geos/wz.html#People, the estimated 2009 population is expected to be 1,123,913 people. Compared to my prediction of the year 2010 having the population 1,461,769 people and having the record of 2005 having the population approximately 1,317,000 people, this estimation can elicit a conclusion that the population of Swaziland is actually shrinking instead of growing as my population model portends. This shrinking could be a result of disease such as AIDS, or other environmental conditions.

I am now going to proceed in finding a population model for Japan using the same methods and ideas I did for Swaziland. I will again use the exponential function, $y = db^x$, and find each variable using algebraic processes. The following table shows approximate populations measured in millions of given years.

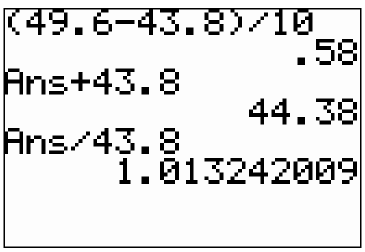
Year	Population (million)	Year	Population (million)
1900	43.8	1960	93.4
1910	49.6	1970	103.7
1920	56.0	1980	117.1
1930	64.5	1990	123.5
1940	71.9	2000	123.9
1950	83.2	2005	127.7

I reused the same idea of simplifying the years that I did when finding the model for Swaziland. I started with 1900 as the first year, 1910 as the eleventh, 1920 as the twenty-first, and so on. These simplified values will express my x value and the years (in millions) will express my y values. I placed these values in the table below.

X	***Y***	***X***	***Y***
1	43.8	61	93.4
11	49.6	71	103.7
21	56.0	81	117.1
31	64.5	91	123.5
41	71.9	101	123.9
51	83.2	106	127.7

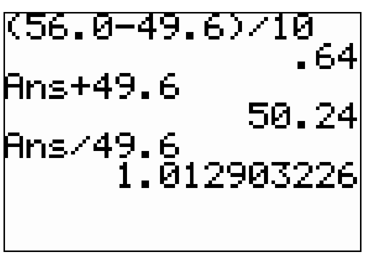
I am going to repeat the same procedures I used to find the model for Swaziland in finding the model for Japan. Therefore, I will first find the rate, *b*. To reiterate what I did in finding this value for Swaziland, I took the change in population divided by the change in year. I then added that quotient to the smaller of the two populations and divided that sum by the same smaller population. I will repeat this process using every group of two consecutive populations and their respective year. I will then average these answers to find the value that will replace the *a* in the model.

I will proceed using my GDC. My first demonstration will use the population 43.8 and 49.6, and the change in years is 10.



The approximate rate of change between year 1900 and 1910 is 1.01.

I will only exhibit one more example of this process, using the data from years 1920 and 1930.

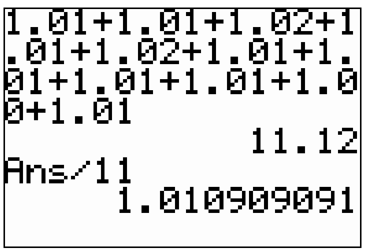


Therefore, the approximate rate of change between these two years is 1.01.

By completing this process for the rest of the data, I obtained the following values that I placed in the table below. I rounded each approximate rate of change to the nearest hundredth place.

Range of Years	Approximate Rate of Change in Population	Range of Years	Approximate Rate of Change in Population
1900-1910	1.01	1960-1970	1.01
1910-1920	1.01	1970-1980	1.01
1920-1930	1.02	1980-1990	1.01
1930-1940	1.01	1990-2000	1.00
1940-1950	1.02	2000-2005	1.01
1950-1960	1.01		

Now I'll find the average approximate rate of change using my GDC.



By rounding this averaged value to the nearest hundredth place, I can say that my population model of Japan to this point is $y = a(1.01)^x$.

In the same way I found my *a* value for the model of Swaziland I will use for my model of Japan. I will plug in each set of values from the original table into the *x* and *y* places of my model thus far, $y = a(1.01)^x$, and algebraically solve for *a*. After finding each individual *a* value for each set of data I will again carry out the process of averaging to find a value that will replace *a* in the population model. As already stated I will be using the population measured in millions for my *y* values and their simplified year values as my *x*. I will display an example of this algebraic process below.

For the example I will be using the data from year 1, which has the population of 43.8 (million).

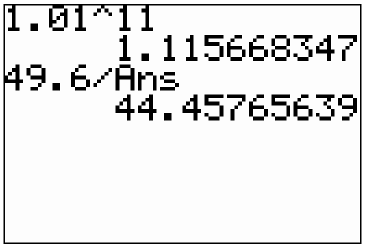
$$43.8 = a(1.01)^1$$

$$43.8 = a(1.01)$$

$$\frac{43.8}{1.01} = a$$

$a \approx 43.366$ (This will be used in finding the averaged *a* value.)

To find the remaining individual *a* values I will be using my GDC. I will only show one example of how this will be done below. For this example I will be using the data from the year 11. The set up for this example is $49.6=a(1.01)^{11}$.

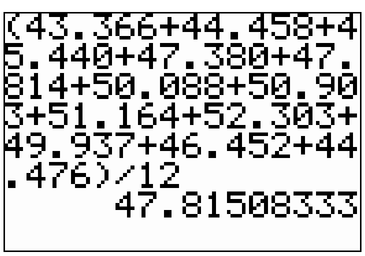


$a \approx 44.458$

After continuing this process of calculating individual *a* values with my GDC for the rest of the data, I found the following answers I placed in the data table below. I rounded each *a* value to the nearest thousandth place.

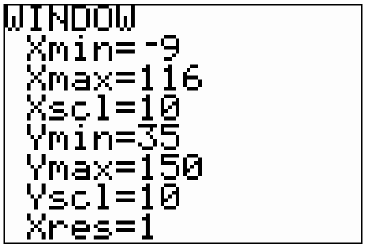
X	***Y***	***a***	***X***	***Y***	***a***
1	43.8	43.366	61	93.4	50.903
11	49.6	44.458	71	103.7	51.164
21	56.0	45.440	81	117.1	52.303
31	64.5	47.380	91	123.5	49.937
41	71.9	47.814	101	126.9	46.452
51	83.2	50.088	106	127.7	44.476

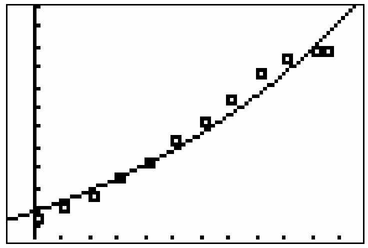
I will now average all the *a* values using my GDC.



I will round this value to the nearest hundredth place which it will then be substituted for the *a* value in my population model of Japan. Therefore, the population model is $y = 47.82(1.01)^x$.

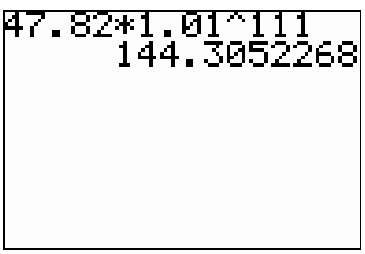
This model fits reasonably well with the data; however, the data seems to have reached a point of leveling off as the model continues on. I have come to this conclusion by studying the graph I made on my GDC.



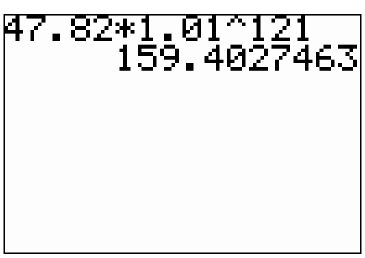


The square dots represent the data from the table and the curved line represents the graph of the model I made. As you can see the pattern of the scatter plot appears to be leveling off while the curve continues.

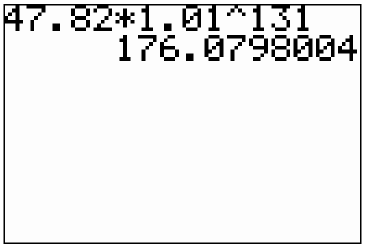
Even though I do not think my model would be the best choice to predict future populations, I am going to utilize it in finding estimated populations of the years 2010, 2020, and 2030. I will go ahead and simplify these year values like I did with the other data, consequently, 2010 will be represented by 111, 2020 by 121, and 2030 by 131. I will then compare these estimations with predictions from other sources to examine the reliability of my model. In order to predict these estimations I will plug the simplified year values in for x in my population model, $y=47.82(1.01)^x$ which will then give me the possible population for that year in millions. I will accomplish this using my GDC.



If $x=111$ then $y \approx 144.305$. I can, therefore, say that the estimated population of year 2010 in Japan will be 144,305,000 people.



If $x=121$ then $y \approx 159.403$. Hence, the estimated population of Japan in 2020 is 159,403,000 people.



If x=131 then y≈176.080. Thus, the population of Japan in 2030 is estimated to be 176,080,000 people.

According to the CIA Factbook, https://www.cia.gov/library/publications/the-world-factbook/geos/ja.html, the estimated population of Japan in July 2009 is 127,080,000. In comparison to the estimated population in 2005, my conjecture of the population leveling off seems to be true. The CIA Factbook states that the estimated population growth in 2009 is in fact a decrease of -.191%. This decrease is consistent with my observation of the leveling off in the graph. This percent decrease in population could be due to the idea that Japan's population has reached its capacity, or disease and other economical issues could be the cause.

The Verhulst-Pearl Model might be a better function to depict the relationship between years and population of Japan than the exponential one I created. This model is a logistic function that could represent my conclusion of the data leveling off, because this model contains a capacity, or limit. This means that the graph will reach a kind of peak and level off. The Verhulst-Pearl Model is the following: $P(t)=\dfrac{MP_0}{P_0+(M-P_0)e^{-kt}}$, where t is time, measured in years after 1900, and $P(t)$ is the population, measured in millions, at time t. k is approximately constant. P_0 is the population in 1900 and M is the limiting population (in millions) which may be considered to be the maximum population capacity of the country could support under given social and environmental conditions. I am going to assume that M=150.

Due to what I already stated about P_0 being the population in 1900, since I am using 1900 as my initial starting year, I can conclude 1910 as being the tenth year, 1920 as the twentieth, etc. This also enables me to declare that P_0 = 43.8. I can now plug everything I know into the Verhulst-Pearl Model. Since I know M=150, and P_0 = 43.8, I can rewrite the model, as the following:

$$P(t)=\frac{(150)(43.8)}{(43.8)+(150-43.8)e^{-kt}}$$

This is capable of being simplified even further to $P(t)=\dfrac{(6570)}{(43.8)+(106.2)e^{-kt}}$.

If I take known values from the data table and place them into their appropriate places in this function, the only unknown variable will be k, the constant. This value can be solved for algebraically in order for this function to be utilized in finding either population or year of its corresponding time or population. In this portfolio I will be using this model to find a more valid prediction of the population of Japan in 2010, 2020, and 2030.

I am going to find this constant by means of inserting the years from the data table in for the *t* variable and its respective population for the *P(t)* variable. I will do this for every set of data with the exclusion of the initial year, and then I'll average the individual *k* values. The exclusion is due to the fact that *k* would be impossible to calculate, since the variable *t* would be equal to zero. To start with, I'll display a revised data table that will show how I simplified the year values for this model.

Years (since 1900)	**Population (million)**	**Year (since 1990)**	**Population (million)**
0	43.8	60	93.4
10	49.6	70	103.7
20	56.0	80	117.1
30	64.5	90	123.5
40	71.9	100	126.9
50	83.2	105	127.7

I am now going to demonstrate how I am going to algebraically step by step solve for *k*. I arbitrarily chose to use the data from year 50. Therefore, *P(t)*=83.2 and *t*=50. Therefore, by replacing these variables, the equation can be shown as

$$83.2 = \frac{6570}{43.8 + (106.2)e^{-k(50)}}.$$

First, I will cross multiply to eliminate the fraction. Then I will use the distributive property and other algebraic operations to simplify. I also will be utilizing logarithmic properties.

$$83.2 \cdot (43.8 + (106.2)e^{-k(50)}) = 6570 \cdot 1$$

$$43.8 + (106.2)e^{-k(50)} = \frac{6570}{83.2}$$

$$(106.2)e^{-k(50)} = 78.96634615 - 43.8$$

$$e^{-k(50)} = \frac{35.16634615}{106.2}$$

$$e^{-k(50)} = 0.331133203$$

$$(\ln e)^{-k(50)} = \ln(0.331133203)$$

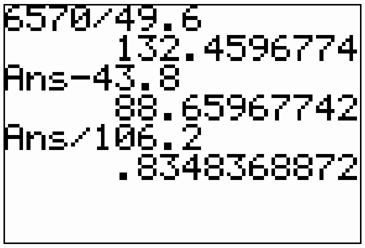
$$-k(50) = -1.105234559$$

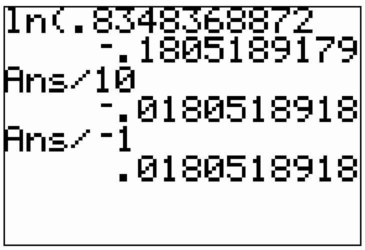
$$-k = \frac{-1.105234559}{50}$$

$-k$= -0.0221046912

k≈0.022

I will now repeat this process for the rest of the sets of data in a more efficient way with the use of my GDC. I will demonstrate one of these calculations below using the data from year 10.



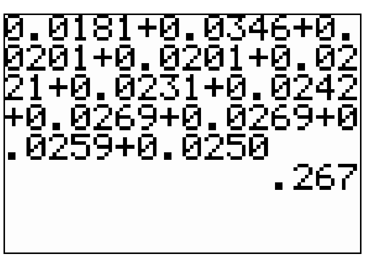


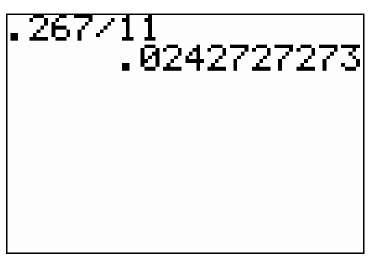
Therefore, for this set of data, $k \approx 0.004$.

By proceeding to calculate these individual k values for every set of data except for the initial year, I found the following values placed in the data table below. This table is to delineate what t and $P(t)$ I used in the function obtain each value. I rounded each value to its hundred thousandth place

P(t)	***t***	***k***	***P(t)***	***t***	***k***
49.6	10	0.0181	103.7	70	0.0242
56.0	20	0.0346	117.1	80	0.0269
64.5	30	0.0201	123.5	90	0.0269
71.9	40	0.0201	126.9	100	0.0259
83.2	50	0.0221	127.7	105	0.0250
93.4	60	0.0231			

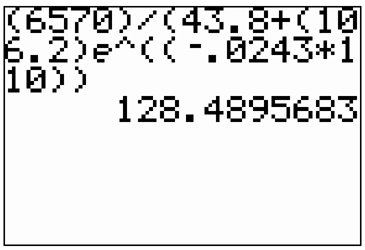
I will now average all of these individual k values using my GDC.



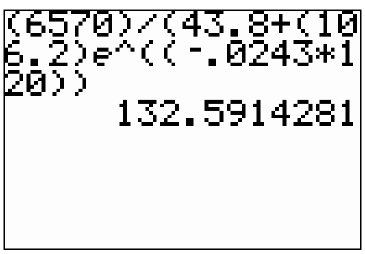


Therefore $k \approx 0.024$.

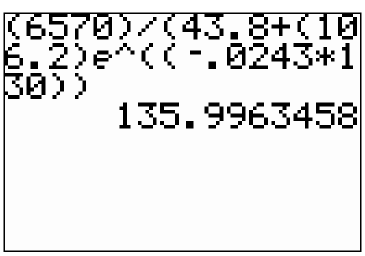
Thus, the Verhulst-Pearl Model for Japan according to my calculations is $P(t) = \frac{6570}{43.8 + (106.2)e^{-0.0243t}}$. I am now able to use this function to find a probable population for the years 2010, 2020, and 2030. I will do this by first simplifying the years in proportional values consistent with the simplification I applied to the years in my original data table for the population of Japan. Therefore, 2010 can be now represented by 110, 2020 by 120, and 2030 by 130. Then I will perform a calculation for each by simply replacing the *t* value with one of these year values and solving with my GDC. These calculations are shown below.



If *t*=110 then *P(t)*≈128.490. In other words, the estimated population of Japan in 2010 is about 128,490,000.

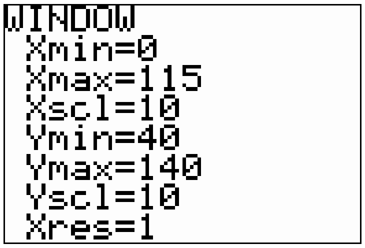


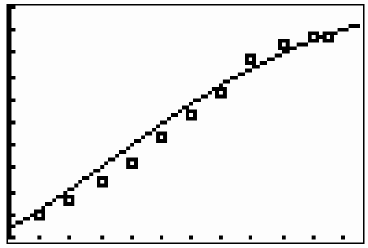
If *t*=120 then *P(t)*≈132.591Therefore, the population of Japan in 2020 is predicted to be approximately 132,591,000.



If *t*=130 then *P(t)*≈135.996. Thus, in 2030 the population of Japan is expected to be near 135,996,000.

I think these predictions are more probable. They fit with the trend of the data more closely, as could be observed by the graph I made with my GDC.

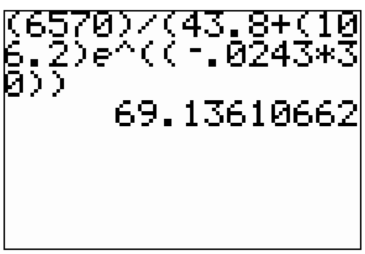




The last three points on this graph are data predicted from my model. As can be observed, the graph of the model follows the pattern of the seeming “leveling off”. This results in the predictions being more feasible.

These estimates fit better with the projections from CIA Factbook, which stated that the estimate 2009 population would be 127,078,679. Compared to the 2010 projections from my model, it is fairly close. I believe this form of model is more preferable because of its having a capacity, which I had mentioned earlier as a possible part of the reason Japan’s population appears to be leveling off.

Another way I can check for accuracy of predictability in this model would be to use it to calculate population estimates for past years and compare them to the actual recorded populations. I will demonstrate this below using my GDC with year 30.



As listed in earlier pages, the actual population in year 30 was about 64.5 million. According to my model, the population should be approximately 69.1 million. As you can see, the population values are similar but not equivalent. This is a representation of how accurate is my model’s ability to predict. Although it is not a perfect model, it is able to calculate reasonable results that can assist in population estimates. Overall, I think the logistic model is a better fit for determining estimations of future populations.

Works Cited

"Japan." CIA - The World Factbook. 23 Apr. 2009. 13 May 2009 <https://www.cia.gov/library/publications/the-world-factbook/geos/ja.html>.

"Swaziland." CIA - The World Factbook. 23 Apr. 2009. 13 May 2009 <https://www.cia.gov/library/publications/the-world-factbook/geos/wz.html>.