A QUANTITATIVE STUDY OF COUNTRIES USING POPULATION DATA AND PYRAMID GRAPHS

Henry Kranendonk
Milwaukee Public Schools, United States
hkranendonk@earthlink.net

While an equation can be solved manually or by the click of a button in multiple ways, the information this equation communicates should also be a reason why students study it. Although the "how to's" are important, they must be balanced with "why?" and "for what purpose?" The latter questions are essentially more meaningful, more interesting, and more engaging for many students. It is critical that students learn how to communicate their thinking to others. The project outlined in this paper was specifically designed for middle and high school students to improve their skills in posing quantitative questions and in articulating "why?" and "for what purpose?" these questions are asked.

INTRODUCTION

Students in the 8th and 9th grades from Milwaukee, Wisconsin were involved in a project entitled "*People Count!*" The first part of this project began with unpacking data displayed in population pyramid graphs that, at first glance, were simply viewed as "odd graphs" by students. Although pyramid graphs are used in biology, geography, and economics to study various conditions of countries, this project connected the shapes of pyramid graphs to basic numeric summaries that generated an understanding of population data.

The first lesson asked students to examine data displayed in a simplified population pyramid graph (see Figure 1). This type of graph displays a side-by-side histogram of the count of people by specific age categories and gender. Figure 1 summarizes a group of people attending an event. Students' conjectures identifying the event became more focused as they read or calculated the number and percent of females and males, the number and percent of various age categories, the number and percent of teenagers, and the mean and median ages of females and males.

Throughout this lesson, students were connecting the population data to several calculated number summaries and the "shape" of the pyramid graph. Initially students developed verbal explanations of the connections they observed. Their explanations were later incorporated as written summaries for questions presented in the project activities.

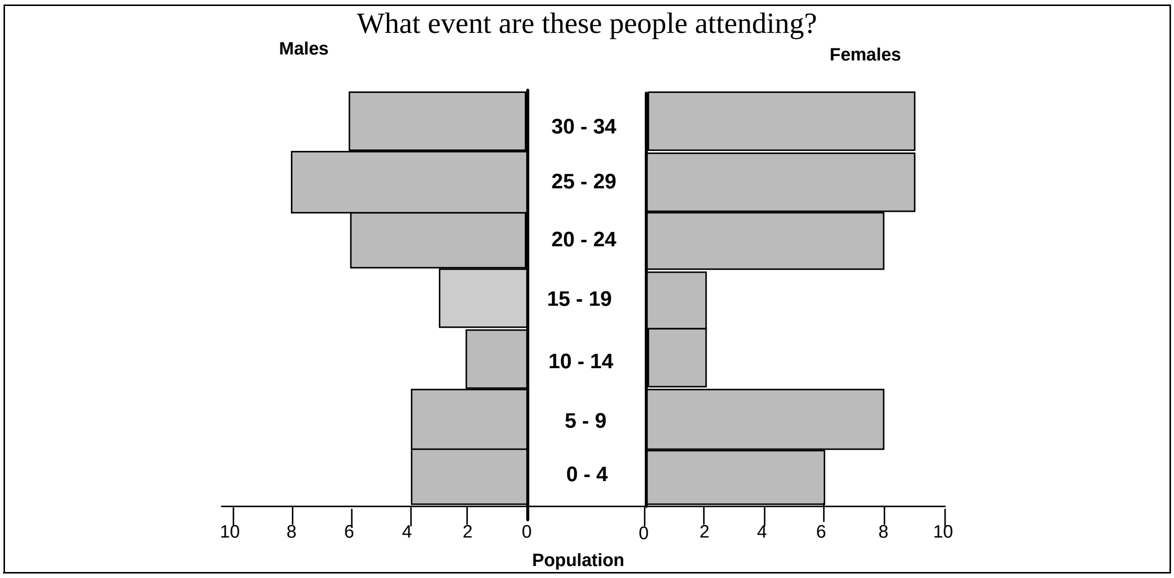


Figure 1: Introducing a Population Pyramid Graph

The mean age of males, females, and the entire population were estimated using two methods. The first method involved assigning the "centered" age, or the median age, of each interval to each person represented in an age category. Students were then able to estimate a mean age for males, females, and the combined population. The mean ages were also derived by a second method using a visual representation of balance. After turning the pyramid graph on its side, students placed a pencil at approximately the middle of the age axis of the males. Each student moved the pencil along the axis to a position that appeared to balance the distribution of the bars. The balance point identified by each student represented the mean age of the males and was compared to the value calculated in the first method.

Early in the process, a few students made conjectures about the event described in the graph, while others made their conjectures as more summaries of the data were derived. This graph was prepared from a summary of people attending a parent-teacher conference at a local day care center. (And yes, some students correctly identified the event!)

CLASSIFICATION OF PYRAMID SHAPES

After working with the simplified population pyramid graph, students were presented two sets of more complex pyramid graphs; one set represented the population in 2000 and a second set represented the projected population in 2050 of 16 countries including the United States. The second lesson developed two classification systems reflecting the general shapes of the pyramid graphs. Students were initially challenged to develop their own classification systems. Most students organized the countries around shapes the graphs resembled. For example, Japan (Figure 2c below) was classified as a ship or boat (along with a few other countries). Other descriptions included "blobs" or a layered cake (see the 2000 graph of Kenya, Figure 2a). After this discussion, a classification of the different countries was introduced that was based on a "layered" pattern of the pyramid graphs. This classification system would provide a way to later connect calculations of the population distributions to the shapes of the population pyramid graphs.

A classification by "layers" was introduced in class. A country in which the *longest* combined layer (or bar) of males and females was located in the first six layers (or ages 0 to 29) of its pyramid graph was classified as a "bottom-layered" country (see Kenya in Figure 2a). A country in which the longest layer of the pyramid graph was in the next six layers (or ages 30 to 59) was classified as a "middle-layered" country (see the United States and Japan in Figures 2b and 2c). A country in which the longest layer of its pyramid graph was located in the 13th or higher layer would be classified as a "top-layered" country (see Japan's 2050 graph, Figure 2c). Students did not classify countries as "top-layered" when the very top layer was the longest as they recognized the age scale for this layer was not displaying the ages of a specific 5-year category as the other layers were displaying (see the United States' 2050 graph, Figure 2b).

(All graphs were obtained from the United States Census Bureau Web site. See the reference: www.census.gov/ipc/www/idbpyr.html or navigate using the links to the International Data Base at www.census.gov).

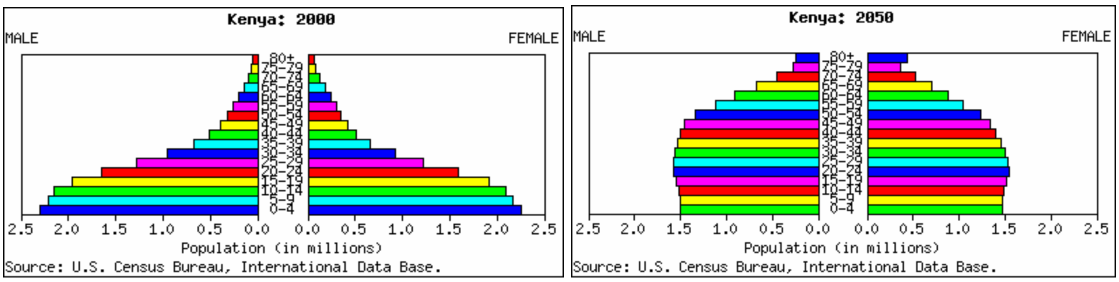


Figure 2a: Kenya

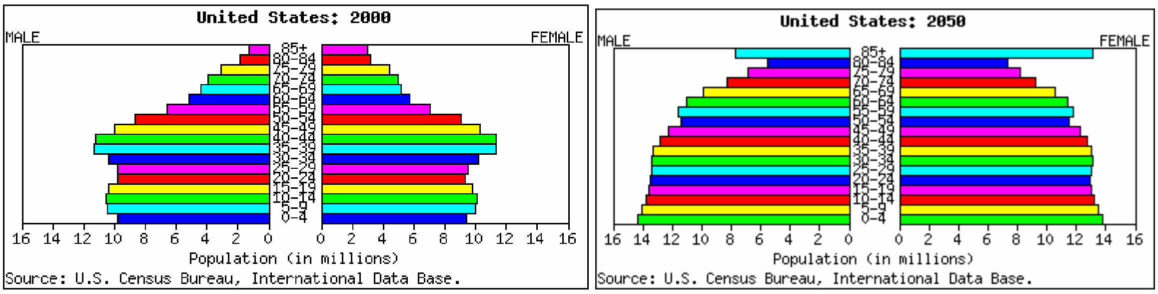


Figure 2b: The United States

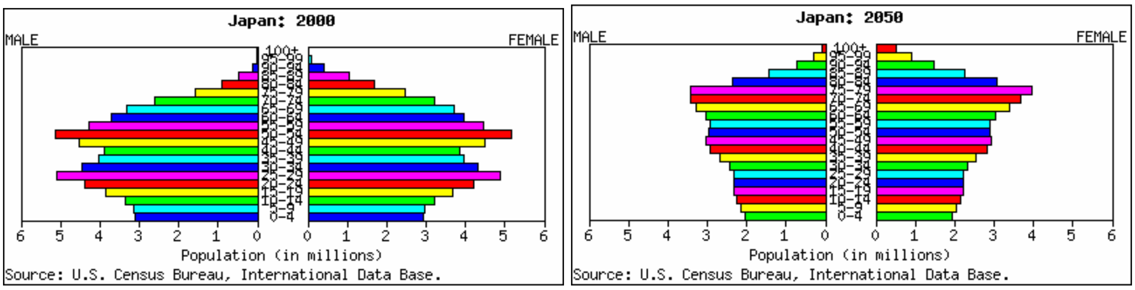


Figure 2c: Japan

A "GENERATIONS" CLASSIFICATION SYSTEM

The "layered" classification system was a relatively easy way to organize the various shapes of the graphs. A second classification system was also introduced that provided students with a language to create a more visual summary of a country's population. This second classification system required, however, that students estimate a value for each of the male and female layers of a country and to calculate the percents of four age ranges. As this process was more tedious, each student was only assigned one or two countries to classify. The process involved the following: Students divided the age categories into four "generations": the *"Child"* (ages 0-24), the *"Parent"* (ages 25-49), the *"Grandparent"* (ages 50-74), and the *"Great-Grandparent"* (75-99). A country was classified by one of these labels based on which generation had or will have the most number of people. For example, Kenya was a "Child" country in 2000 and is projected to be a "Child" country in 2050; the United States was a "Parent" country in 2000 and is projected to be a "Child" country in 2050; and Japan was a "Parent" country in 2000 and is projected to be a "Grandparent" country in 2050. Relating this classification to the pyramid shapes provided students a way to discuss the differences or similarities of the populations of several countries and the challenges faced by countries dominated by a particular generation.

After the generalizations about the shapes, students were directed in the third lesson of this series to derive several number summaries for the population distributions. Students worked in pairs to study one country other than the United States. Using the pyramid graphs, the students calculated at least 10 "almanac summaries," or quantitative descriptions of a country generally highlighted in an almanac. Included in these descriptions were the mean ages of the males and females, an estimate of the mean age of the total population, the percent of females and the percent of males, the percent of children under the age of 10, the percent of teenagers, etc. Mean ages were estimated by finding a balance point as well as through an estimation involving the "centered age" of the age intervals.

One specific summary students derived was used to discuss the spread of the distribution, specifically using a few top and bottom layers of the pyramid graphs. This summary was defined as the "old-to-young" indicator and was the decimal number formed by reducing the ratio of the estimated number of "old" people (or people aged 65 and older) to "young" people (or people aged under 10). As each country's "old-to-young indicator" was derived, the layered classification labels became more meaningful. For example, Kenya's bottom-layered classification explains its very small "old-to-young" indicator (approximately 0.12). Bottom-layered countries produced "old-to-young" indicators within the range of 0 to 0.7. The United States has an "old-to-young" indicator of approximately 1.0. Most middle-layered countries produced indicators in the ranges of 0.7 to 2.0. The countries with the greatest "old-to-young" indicators (Greece, Japan, Italy) are also the countries projected to be top-layered in 2050. Students connected the indicators to the balance points representing the mean ages. The value of the "old-to-young" indicator could be visualized as the pull the older age categories (for an indicator greater than 1) or the younger age categories (for an indicator less than 1) exert on the population distribution. A summary of the countries studied is provided in Table 1.

Table 1: Shape, Mean age, and "Old-to-Young" Indicators

	USA	Argen.	Brazil	Aust.	China	Cuba	France	Ghana
Shape Description	Middle-layered	Bottom-layered	Bottom	Middle	Middle	Middle	Middle	Bottom
Mean age	35.9	32.0	28.3	35.8	31.0	34.1	39.2	23.5
"Old-to-Young"	0.93	0.61	0.28	0.90	0.44	0.73	1.47	0.10

	Greece	Japan	Italy	Kenya	Mexico	Russia	Nigeria	Finland
Shape Description	Middle	Middle	Middle	Bottom	Bottom	Middle	Bottom	Middle
Mean age estimate	40.0	41.3	41.1	21.9	26.3	37.0	22.9	38.6
"Old-to-Young"	1.79	1.96	1.95	0.12	0.23	1.39	0.12	1.07

ASSESSMENT

The fourth lesson of this series represented a summative assessment of the project. Students, working in pairs with chart paper, markers, and colored construction paper, were directed to develop a poster that compared the country they studied with the United States. They were instructed that at least five items had to be connected to summaries obtained from the pyramid graphs of the two countries. The students were allowed to use their answers from the discussion questions as well as additional information they were able to research.

Questions comparing the 2000 to 2050 pyramid graphs for each country were also posed to the students. Students were asked how they thought the Census Bureau developed a projection of the count of people in a country for the next 50 years. The topics the students addressed were organized around various models of a changing population. Whatever model was used to estimate future counts, four dynamic factors were used to explain how a country's population would change, namely immigration (people moving into the country), emigration (people moving out of the country), deaths, and births. Each factor contributes to the "shape" the pyramid graph of a country's population would exhibit.

Students are often introduced to population applications as examples of linear or exponential models in their 8^{th} or 9^{th} grade algebra (or equivalent) course. In most cases, deriving an equation that will be used to estimate the projected count of people is the focus. Population models are derived to determine predictions that are based on specific assumptions of what will or will not change. Students rarely analyze the underlying questions about what might cause a country's population to increase (or decrease).

The organization of a country's population from a pyramid graph can be used to develop a prediction of a country's population by looking at how smaller "chunks" of the population change over time. The United States has a particularly ragged bulge in the age intervals of 45 to 65 year old adults of the 2000 pyramid graph that is known as the country's "Baby Boomers." How might the underlying ragged chunks of the population change the way a country's future population will look over time? How do the changes in these chunks affect the total population?

A spreadsheet application was designed that managed the tedious calculations needed to analyze the smaller chunks over 5-year intervals. By creating a graph similar to a pyramid graph, students were provided a glimpse of the changing shape of a country's population. The population data displayed in this graph combined the male and female count of each age category. See *Figure 3* for the format developed by the spreadsheet application.

Table 2 was used to begin the discussion of a different type of application to estimate future counts managed by the spreadsheet. This table introduced students to a specific connection that analyzes age intervals of five years. An expanded table of these connections was used to design this application. The arrow in Table 2 identifies a *connecting layer*. Connecting layers track the count of a specific group (or "chunk") of people over five years. For example, the count of people 5 to 9 years old in 2000 is determined by subtracting the number of people who died and emigrated plus adding the number of people who immigrated from the count of people 0 to 4 years old in 1995. The arrow in Table 2 identifies this specific connecting layer. It connects the "base" layer (1995 population) with its corresponding "next" layer (2000 population). This connecting layer is particularly interesting as the count increased, therefore, indicating the net effects from immigration.

Table 2: Example of a "connecting layer" from the population estimates of the USA in 1995 and 2000. This is particularly interesting as the count increased. The total refers to the total population of the country.

"*Base*" population

Ages	1995 population estimates (in thousands)
5 - 9	19,096
0 - 4	19,532
Total	262,087

"*Next*" population

Ages	2000 population estimates (in thousands)
5 - 9	19,805
0 - 4	18,946
Total	274,645

Students were directed to examine the connecting layers and to create the ratio of the "next to the base layer" count of people. The ratio was reduced to a number that was called a "connecting layer factor" as it was multiplied in subsequent iterations to estimate the count of people in the new "next" layers of this application. Although this factor does not in itself indicate the extent of immigration or emigration or death during the 5-year period, it does indicate whether or not incoming counts (immigration) were outweighed by outgoing counts (emigration or death). The connecting layer factor for the connecting layer illustrated in Table 2 is approximately 1.014, or the decimal approximation of 19805 / 19532. The connecting factors derived from the 1995 and 2000 population totals were used to derive estimates from 2000 to 2050. Table 3 summarizes several of the connecting factors students used in this application.

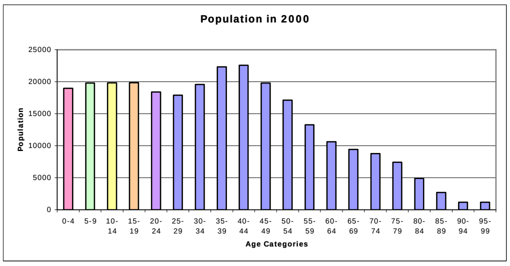
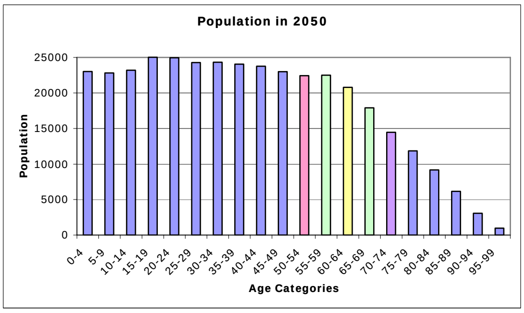
Table 3: Connecting Factors

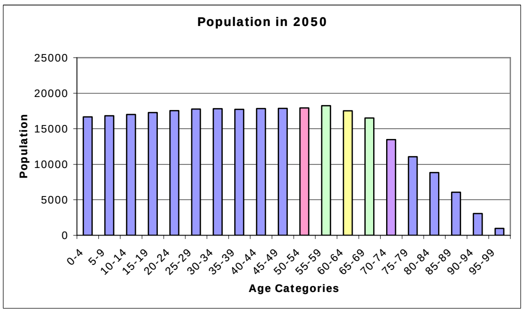
Base Layer Ages	*Base Population for Year:* 1995 (in thousands)	*Next Layer Ages*	*Next Population for Year:* 2000 (in thousands)	*Connecting Layer Ratio or Next / Base*	*Connecting Layer Factor*
50 - 54	13,642	55 – 59	13,262	13262/13642 =	.972
45 - 49	17,458	50 – 54	17,113	17113/17458 =	.980
40 - 44	20,259	45 – 49	19,810	19810/20259 =	.978
35 - 39	22,296	40 – 44	22,559	22559/22296 =	1.012
30 - 34	21,825	35 – 39	22,313	22313/21825 =	1.022
25 - 29	18,905	30 – 34	19,549	19549/18905 =	1.034
20 - 24	17,982	25 – 29	17,893	17893/17982 =	0.995
15 - 19	18,203	20 - 24	18,384	18384/18203 =	1.010
10 - 14	18,185	15 - 19	19,856	19856/18185 =	1.092
5 - 9	19,096	10 - 14	19,829	19829/19096 =	1.038
0 - 4	19,532	5 - 9	19,805	19805/19532 =	1.014

Population values were obtained from: http://www.census.gov/popest/archives/1990s/nat-agesex.txt

To estimate the number of 10 to 14 years old people in 2005, the number of people in the 5 to 9 years old in 2000 is multiplied by the connecting factor 1.038. Each 5-year age interval is estimated using the corresponding connecting factor. A cascading effect of projecting connecting layers into the future represented a recursive process with this application. Estimates for future counts in 5-year intervals are based on the assumption that the past changes in the 5-year age intervals remain proportionally constant.

The formulas used in the spreadsheet application allowed, however, students to also consider what if the assumptions of change are not held constant. The spreadsheet provided students control over several situations that might affect the future count of people. *"What if…?"* scenarios were investigated by simply changing the population of a specific age interval, or by changing the ratios of connecting layers, or by changing the ratio of the 0 to 4 years old count of people to the "parent population" of 15 to 44 years old highlighted in Table 3. (This last ratio essentially was used to estimate the 0 to 4 years old count for each 5-year application.) Similar to the algebraic equations of a closed linear or exponential model, the spreadsheet application provided a generalization of the changes in counts from one 5-year period to the next.

<table>
<tr>
<td>Shape of the population distribution of the United States in 2000 as summarized by the census data.

</td>
<td>Shape of the population distribution of the United States in 2050 as summarized by the census data.

</td>
</tr>
<tr>
<td>What if immigration was severely limited? Determine a maximum estimate of the population of the USA if a law limiting immigration was enforced from 2000 to 2050.

Graph to the right represents the resulting 2050 population in which the connecting layer factors greater than 1 (designating an obvious addition of the population from immigration) were reduced to a value slightly less than 1.</td>
<td>Shape of the population distribution of the United States in 2050.

</td>
</tr>
<tr>
<td>What if a disease wiped out approximately half of the population aged 65 or older beginning shortly after the count of people in 2000?

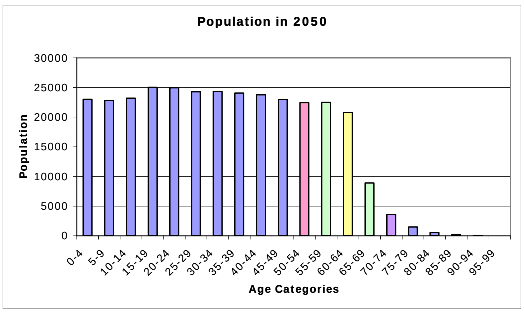
Graph to the right represents the resulting 2050 population in which the connecting layer factors were cut in half for the ages 65 and older.</td>
<td>Shape of the population distribution of the United States in 2050.

</td>
</tr>
</table>

Figure 3: Applying the *"What if…?"* questions to the spreadsheet application

The project concluded with students using the spreadsheet application with a data set formed by extracting population values from another country's pyramid graph. In each example, students addressed "why?" and "for what purpose?" this type of analysis would be important.