Chapter 4 Project: Fertility Rates in Developing Countries

Name _____   Name _____

Background: The negative impacts of a world population surging past 6 billion people have fallen most heavily on nations whose economies are less industrialized. These so-called "developing" countries, with population growth rates often exceeding 3% annually, face an uphill task meeting essential human needs such as housing, clean water, nutrition and basic health care.

In recent years, population growth rates have fallen remarkably in many developing countries because of declining fertility rates. Slower population growth helps protect the environment by easing demands on natural resources such as water, wood, soil and fish, and by lessening migration into overcrowded cities. In this project, you will explore several factors that influence fertility rates.

Data:
The data that you will examine in this project are from a Johns Hopkins School of Public Health Report titled *Why Family Planning Matters*, published in 1999. This report has data on numerous population variables for 139 developing countries. For this project, you will examine 6 of these variables: family planning through the use of contraception, economic wealth, women's age at marriage, infant mortality, female education, and total fertility rate. In the report, only 81 of the 139 countries have data for all 6 of these categories. The data for those 81 countries are provided on the attached 2-page data sheet.

More specific definitions of the variables are as follows:

CONTR	Percentage of couples using contraception as a family planning option. Source: United Nations.
GDP	Gross Domestic Product is the total output of goods and services per capita in 1997 U.S. dollars. GDP is a commonly-used indicator of economic wealth.
MEDAG	The median age of women at first marriage. The median is the number that lies in the middle of a data set when the data are arranged in order. Sources: U.S. Agency for International Development's Demographic and Health Surveys and United Nations.
MORT	Number of infant deaths (0 to exactly 5 years of age) per 1,000 live births. Source: UNICEF
SCHO	The percentage of females enrolled in secondary school. Source: UNICEF
TFR	Total Fertility Rate is the average number of children a woman would have during childbearing years (ages 15 to 49) at current birth rates. Source: Population Reference Bureau.

1. Examining Total Fertility Rate (TFR) data

a) The Total Fertility Rate for a country is the average number of children a woman will have during her reproductive lifetime. Scan the 81 TFR data values on your data sheet to find the minimum and maximum TFRs. Which countries have those rates?

b) The **replacement rate** is the TFR necessary to keep the long-term population constant. In theory, the replacement rate is 2.0—just enough children are born to grow up and replace the parents. In reality, the worldwide replacement rate is about 2.1 because some children do not live to reach breeding age. Of the 81 developing countries listed, how many have TFRs at or below the worldwide replacement rate? What percentage of the countries is that number?

c) Suppose a particular country has a replacement rate of 2.5. What can you conclude about the population characteristics of that country?

The birth rate is defined as the number of live births each year for every 1,000 individuals in the population. TFRs and birth rates are highly correlated: countries with high birth rates tend to have high TFRs, and those with low birth rates have low TFRs. Similarly, the death rate (number of deaths per 1,000 individuals) is highly correlated with childhood mortality.

d) Using TFRs as a proxy (substitute) for birth rates, and childhood mortality rates (MORT) as a proxy for death rates, find 3 developing countries, from the list of 81, that will fit each stage of the **demographic transition**. (See *Science in Depth* article in

Chapter 4.) If you think that there are no countries that fit a particular stage, write "none."

stage	countries *(list 3 for each stage)*
1	
2	
3	
4	

e) Briefly explain how you chose countries for each stage above:

2. TFR vs. Contraception – Straightedge Method

a) Examined the scatterplot of the (CONTR, TFR) data which is provided with this project. Is there an overall trend in the data? Explain.

b) Use the Straightedge Method to find the line that best approximates the scatterplot's data. Write the equation of the line using appropriate variables, and using 3 decimal places for the slope.

c) Give an interpretation of what the slope of the line means. In your interpretation, comment specifically about the numerical value of the slope.

d) Interpret the meaning of the TFR (vertical) intercept:

e) Find and interpret the meaning of the CONTR (horizontal) intercept.

f) List one country that falls very close to, or exactly on, your hand-drawn line. What are the (CONTR, TFR) coordinates for that country? Check that these coordinates fit the equation produced through the straightedge method.

g) The percentage of Brazil's couples that use contraception as a family planning method is 77%. Use this percentage and the linear equation to find Brazil's *predicted* TFR.

h) The **TFR deviation** is the actual value minus the predicted value. Determine the deviation for Brazil's TFR. Find and label the point on the scatterplot corresponding to Brazil. Explain what the TFR deviation represent on the graph.

i) Which country has the *largest negative* TFR deviation? Mark that country on the scatterplot. What is the value of the TFR deviation?

Before you begin the next set of questions, you will need to load the 2-page data sheet into StatCrunch. Go to the textbook's website (enviromath.com) and locate the Chapter 4 project materials. Click on the StatCrunch link to load the data. You will be prompted to log into StatCrunch. Note that you will need to register if this is your first time using StatCrunch.

Once you log into StatCrunch, 7 columns of information will load automatically. Scan the columns to ensure that there are 81 rows of data corresponding to the 81 developing countries.

The data will automatically be saved into your StatCrunch "account" under the filename Ch04ProjectData.txt. Click <u>MyData</u> to view and manage all of your saved data sets.

3. TFR vs. Contraception – Least Squares Regression

To ensure that the data is loaded correctly, create a scatterplot of the (CONTR, TFR) data on StatCrunch by selecting **Graphics > Scatter Plot**. Under X Variable, select CONTR and under Y Variable select TFR. Then go directly to the scatterplot by selecting **Create Graph!** Compare the StatCrunch scatterplot with the printed scatterplot included with the project – are they similar?

Now find the best-fitting regression line through the (CONTR, TFR) data using StatCrunch. To do this, select **Stat > Regression > Simple Linear**. Under X Variable, select CONTR and under Y Variable select TFR. Then press **Calculate**.

a) What is the best fitting regression equation?

b) What is the correlation coefficient? Is a country's percent contraceptive use a good predictor of its TFR? Explain

c) Summarize your results by listing the slope and *y*-intercept for each equation. Use 3 decimal places of precision.

	slope	y-intercept
Straightedge		
Least Squares Regression		

4. Exploratory Data Analysis

You've found that, among developing countries, percent contraceptive use is a very good predictor of TFR. Are there other variables that explain TFR equally as well? To find out, use StatCrunch to plot y = TFR against each of the other x variables (GDP, MEDAG, MORT, and SCHO). *See the previous section for instructions on how to create a scatterplot on StatCrunch.*

a) Make a "thumbnail" sketch of each scatterplot in the table below to show the general shape of the data. Label the axes, but don't be concerned with scales. Finally, compute and record the regression equations and correlation coefficients. *Again, revisit the previous section for directions on how to calculate regression lines.*

equation:

corr. coef:

equation:

corr. coef:

equation:

corr. coef:

equation:

corr. coef:

b) Compare the correlation coefficients. Which of the four independent variables does the best job in predicting TFR? How does that variable compare with percent contraception use in explaining TFR?

c) Which of these 4 variables does the poorest job in predicting TFR? Explain.

5. Removing the effect of outliers

a) In the previous section you should have found that the scatterplot of the (GDP, TFR) data was highly skewed by several **outliers**. Outliers are data points that do not fit the general trend. Go to the scatterplot and determine the **4 most significant outliers.** Which countries are those outliers?

On your data sheet, circle the GDP and TFR values of these 4 outlier countries. You would like to remove these outliers to better examine whether there is a pattern amongst the remaining (GDP, TFR) data. But you might need the outlier values at a later time. The solution is to store a copy of the GDP and CONTR data into new columns.

To copy the GDP data into a new column, select **Data > Transform data**. A new window will appear, with a box labeled **transformation**. In the box, enter **1*GDP**. See below.
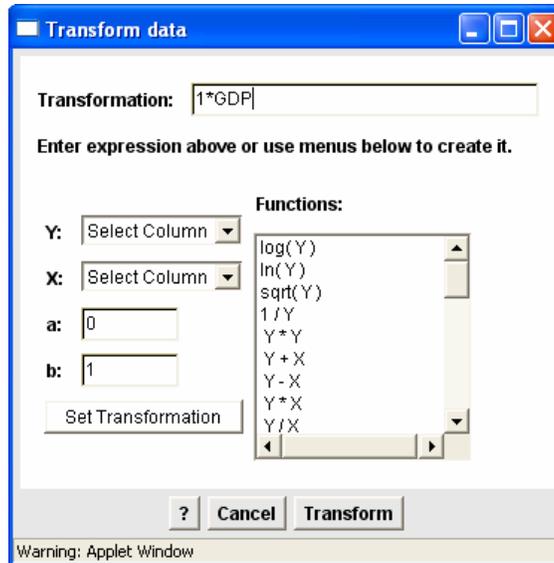
Now press the **Transform** button on the bottom of the window. A new window will appear indicating that you have created a new column. Close this window, and then inspect the new column (labeled 1*GDP) to ensure that it contains the GDP data.

Repeat the steps above to copy the TFR data into a new column (use a 1*TFR transformation).

Now scroll through the columns labeled 1*GDP and 1*TFR. When you come to an outlier value, press **Delete** on your keyboard.

b) On StatCrunch, make a scatterplot of the revised (GDP, TFR) data that excludes the 4 outliers. Then find the regression line and correlation coefficient. Sketch and record below. Remember to label axes.
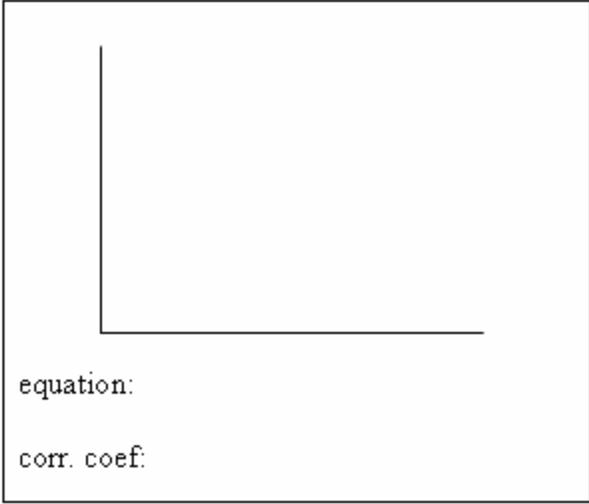

equation:

corr. coef:

c) How well did removing the outliers improve the linear nature of the scatterplot?

6. Cause and Effect

a) Examine the results again from question 4 on Exploratory Data Analysis. How well does SCHO, the percentage of females in secondary school, predict the TFR for a country?

b) Let's swap the order of these variables, so that total fertility rate is plotted on the horizontal axis, and percent schooling is on the vertical axis. Make a scatterplot of the (TFR, SCHO) data on StatCrunch. Also find the regression equation and correlation coefficient. Record below, labeling appropriately.

equation:

corr. coef:

c) How well do the TFR values predict the percentage of females that are enrolled in secondary school?

d) If a country reports that its total fertility rate is 4.8, what percentage of females would you predict to be enrolled in secondary school.

e) With regards to developing countries, some claim that an increase in female education causes a lower TFR; others claim that a lower TFR causes an increase in education. Expound on these cause and effect ideas, giving a few reasons why each might be true.

More schooling lowers TFRs:

Lower TFRs result in more schooling: