**Chapter 1 Data and Business Decisions**

**Basic Concepts Review Questions**

**1. Explain the importance of statistics in business.**

**Answer:**

Statistics is the science of collecting, organizing, analyzing, interpreting, and presenting data. In business, statistics is quite important because it allows managers to make fact-based decisions instead of “gut feel” type decisions. In addition, if various claims are made about a product or service, the use of statistics can prove or disprove claims which can prevent legal issues and can allow for true and ethical decisions about a hypothesis.

**2. Explain the difference between data and information.**

**Answer:**

Data typically refers to the raw data, which is an important ingredient in producing useful information. Information is what managers can use to make appropriate decisions.

**3. What do the terms “Descriptive Statistics” and “Statistical Inference” refer to?**

**Answer:**

The decision making process of any scientific discipline requires the appropriate use of statistical data. The process of collection, organization and description of data is commonly called descriptive statistics. Statistical inference, on the other hand, refers to the process of drawing appropriate conclusions about unknown characteristics of a population based on sample data.

**4.** **What is a population? Why do we generally use sample information to make conclusions about the population.**

**Answer:**

In any decision making process our interest is in the characteristics of a group of individauls having some common feature. The totality of all such individuals is called the population of interest. In most practical situations the population is too large to do any meaningful analysis by collecting the relevant information from all the individuals in the population. Costraint on resources like time and money make the enumeration of the entire population unfeasible. In such situations the analysts actions have to be based on a suitably chosen subset of the population, which we refer to as the sample.

**5. What is a metric, and how does it differ from a measure?**

**Answer:**

A metric is a unit of measurement that provides a method for objectively quantifying performance. A measurement is the act of obtaining data. Measurement creates measures which are numerical values associated with a metric.

**6. Explain the difference between a discrete and a continuous metric.**

**Answer:**

A discrete metric is countable and finite number of distinct values and is expressed as counts or proportions. Continuous metrics are results of measurements, such as length, time or weight, and assume an infinite (continuous) range of possibilities.

**7. Explain the differences between categorical, ordinal, interval, and ratio data.**

**Answer:**

Categorical data or nominal data is data that is sorted into categories according to specified characteristics, without any natural order, such as male/female by geographic regions.

Ordinal data are ordered or ranked according to some relationship to one another. Rating a service as poor, average, good, very good, or excellent is an example of ordinal data.

Interval data are ordered, have a specified measure of the distance between observations but have no natural zero. Common examples are time and temperature.

Ratio data is interval data which have a natural zero. Most business and economic data fall into this category, and statistical methods are the most widely applicable to them.

**8. Explain the difference between cross‐sectional and time‐series data.**

**Answer:**

Cross sectional data is the data that are collected over a single period of time, such as responses to market questionnaires. Time series data is the data collected over a period of time, such as NASDAQ’s daily returns.

**9. What is the Six Sigma initiative?**

01-02

**Answer:**

Six sigma is a business process improvement approach which seeks to find and eliminate causes of defects and errors, reduce cycle times and cost of operations, improve productivity, better meet customer expectations and achieve higher asset use and returns on investment in manufacturing and service processes. Six Sigma studies because they address the Measure, Analyze and Control aspects Six Sigma problem solving paradigm: DMAIC (Define, Measure, Analyze, Improve and Control).

**10. What is the difference between a population and a sample?**

**Answer:**

A population consists of all items of interest for a particular decision or investigation, such as all the residents of a county or all the students at a university. A sample is a subset of a population, such as the residents in a neighborhood or the students in a business statistics class.

**11. List the different types of charts available in Excel, and explain characteristics of data sets that make each chart most appropriate to use.**

**Answer:**

There are many different types of charts that Excel can generate:

1. Column and bar charts can be used to compare types of data against each other or against a standard. Column charts are vertical and bar charts are horizontal.
2. Line charts provide a useful means for displaying data over time.
3. Pie charts show the relative proportion of each data source to the total.
4. Area charts combines the features of a pie chart with those of line charts. Area charts present more information than pie or line charts alone, but may clutter the observer’s mind with too many details if too many data sets are used.
5. Scatter diagrams show the relationship between two variables.
6. Stock charts allow a manager to plot stock prices, including the high, low, and close.
7. Doughnut charts are similar to pie charts, but can include more than one set of data.
8. Surface charts show 3 dimensional data.
9. A bubble chart is a type of scatter chart, but the size of the data marker corresponds to the value of a 3rd variable.
10. A radar chart allows for the plotting of multiple dimensions of several data series.
11. **12. What kind of graphical displays are appropriate for representing the frequencies of several mutually exclusive categories? What are the options for representing such data over the levels of another variable, such as year, gender or race?**

01-03

**Answer:**

The possible options for graphically representing the frequencies of several mutually exclusive categories are the column chart, the stacked column chart, and the pie chart. All of them give good visual representations for this kind of data. When such data have to be represented over the levels of another variable one can use several column charts or stacked column charts over the same axes.

**Problems and Applications**

1. For the Excel file *Surface Finish,* identify each of the variables as either categorical, ordinal, interval, and ratio.

**Answer:**

**Datasheet: Surface Finish**

Surface Finish: Interval

RPM: Ratio

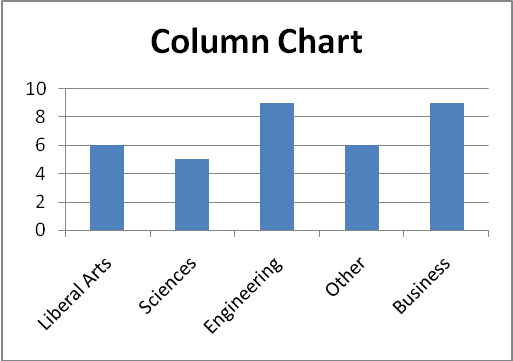
Cutting Tool: Categorical

2. Consider the data given in the file MBA student survey, and look at the variable “Undergraduate Concentration”.

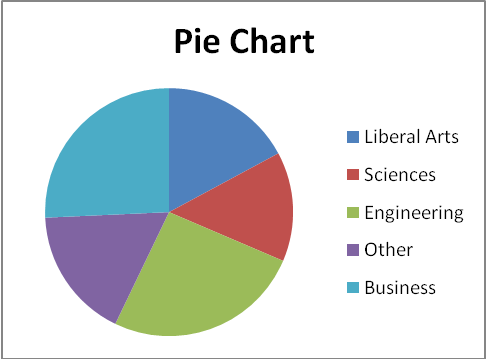
1. Construct a column a chart that visually represents the frequencies of each of the undergraduate concentrations.
2. Consider a pie chart showing the proportion of individuals in each concentration.

**Answer:**

01-05

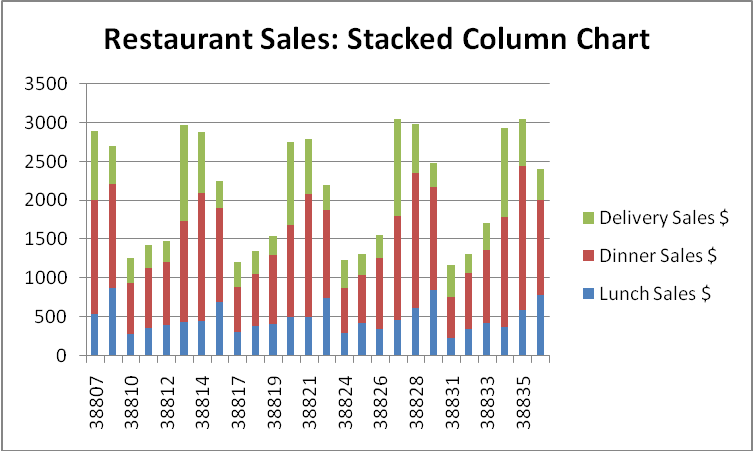
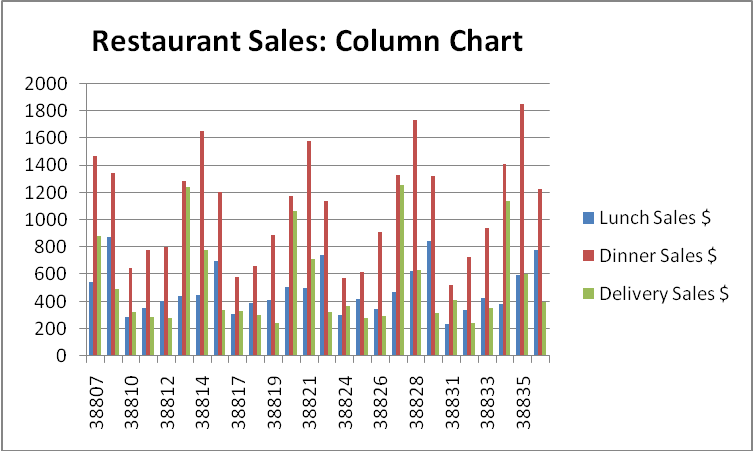


3. Construct a column chart for the data in the Excel file *Restaurant Sales* to allow a comparison of the “Delivery Sales” with the “Lunch Sales” and “Dinner Sales”. What other charts may be used to display this information?



**Answer:**

**Column Chart:**



A bar chart would be another representation, but is too complicated.

01-06

4. Data from the 2000 U.S. Census show the following distribution of ages for residents of Ohio:

|  |  |
| --- | --- |
| Total Households | 4,445,773 |
| Family households (families) | 2,993,023 |
| With own children under 18 years | 1,409,912 |
| Married‐couple family | 2,285,798 |
| With own children under 18 years | 996,042 |
| Female householder, no husband present | 536,878 |
| With own children under 18 years | 323,095 |
| Nonfamily households | 1,452,750 |
| Householder living alone | 1,215,614 |
| Householder 65 years and over | 446,396 |

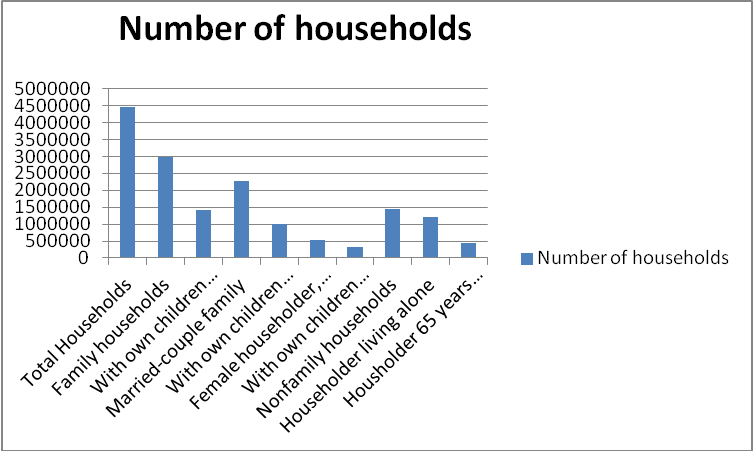
a. Construct a column chart to visually represent these data.

b. Construct a stacked bar chart to display the sub categories where relevant. (Note that you will have to compute additional subcategories, for instance, under Family households, the number of families without children under 18, so that the total of the subcategories equals the major category total. The sum of all categories does not equal the total.)

c. Construct a pie chart showing the proportion of households in each category.

**Answer:**

1. **Column Chart:**

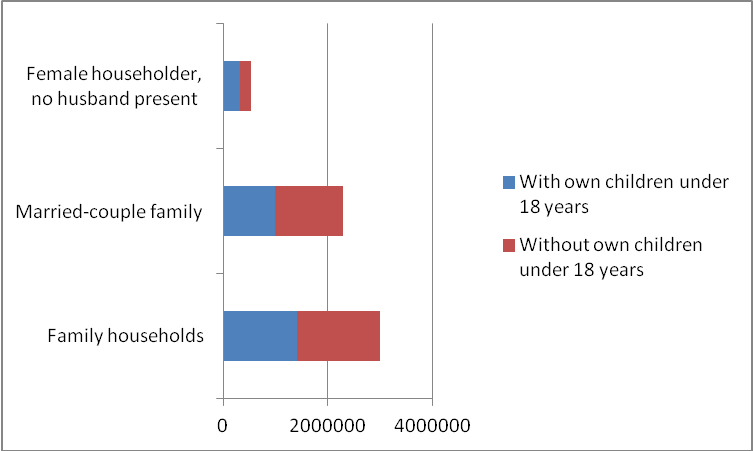


1. **Stacked Bar Chart:**

01-07

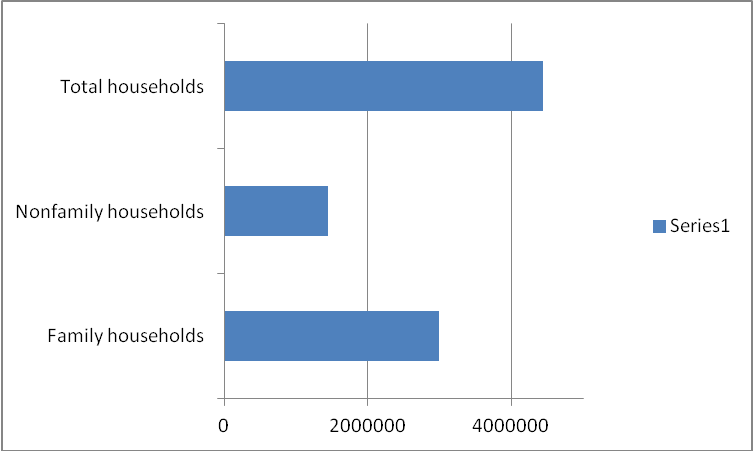
|  |  |
| --- | --- |
|  | Number of households |
| Total Households | 4,445,773 |
| Family households | 2,993,023 |
| Married-couple family | 2,285,798 |
| Female householder, no husband present | 536,878 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | Family households | Married-couple family | Female householder, no husband present |
| With own children under 18 years | 1,409,912 | 996,042 | 323,095 |
| Without own children under 18 years | 1,583,111 | 1,289,756 | 213,783 |



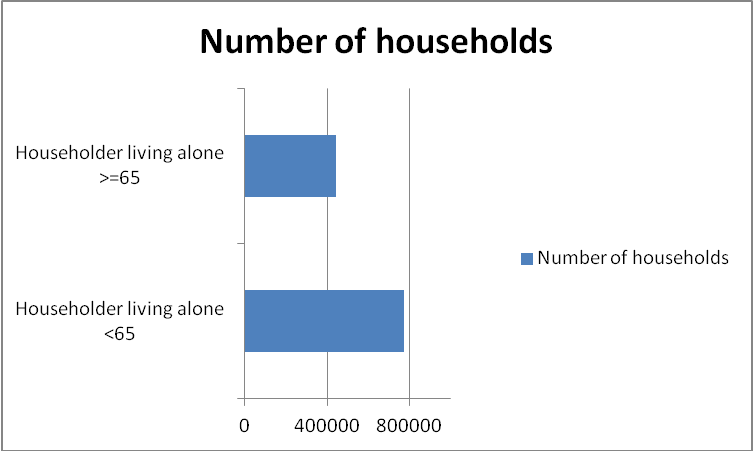
01-08

|  |  |  |
| --- | --- | --- |
| subcategories | Category total | Overall total |
| Family households | 2993023 | 4445773 |
| Nonfamily households | 1452750 | 4445773 |
| Total households | 4445773 |  |



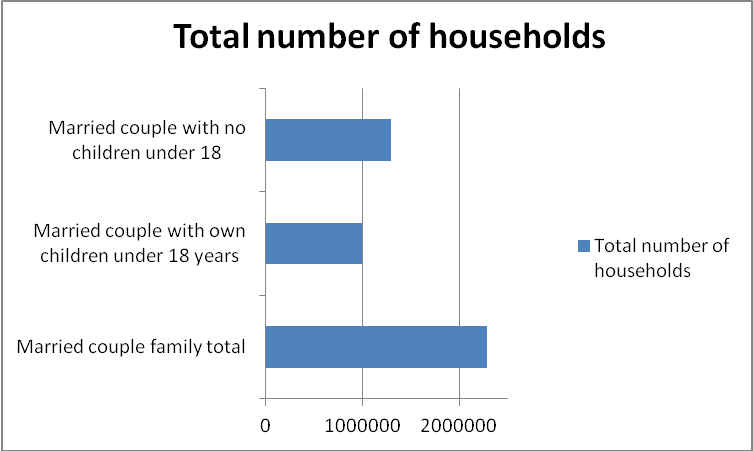
|  |  |  |
| --- | --- | --- |
| subcategories |  |  |
|  | Number of households | Total number of households living alone |
| Householder living alone <65 | 771218 | 1215614 |
| Householder living alone >=65 | 444396 | 1215614 |

|  |  |  |
| --- | --- | --- |
| subcategories | Householders living alone <65 | Householders alone >=65 |
| Both households | 771218 | 444396 |
| Total households | 1215614 | 1215614 |



01-09

|  |  |
| --- | --- |
| subcategories | Total number of households |
| Married couple family total | 2285798 |
| Married couple with own children under 18 years | 996042 |
| Married couple with no children under 18 | 1289756 |

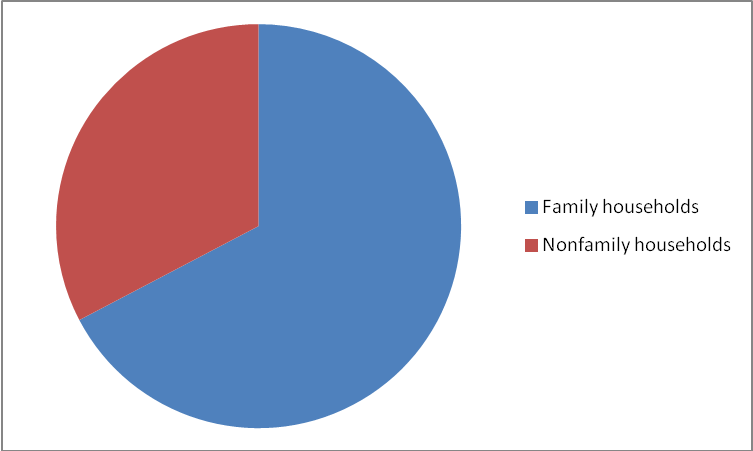


|  |  |
| --- | --- |
| Subcategories | Total number of households |
| Female householder, no husband present | 536878 |
| Female householder, no husband present, with children under 18 | 323095 |
| Female householder, no husband present, with children under 18 | 213783 |

01-10

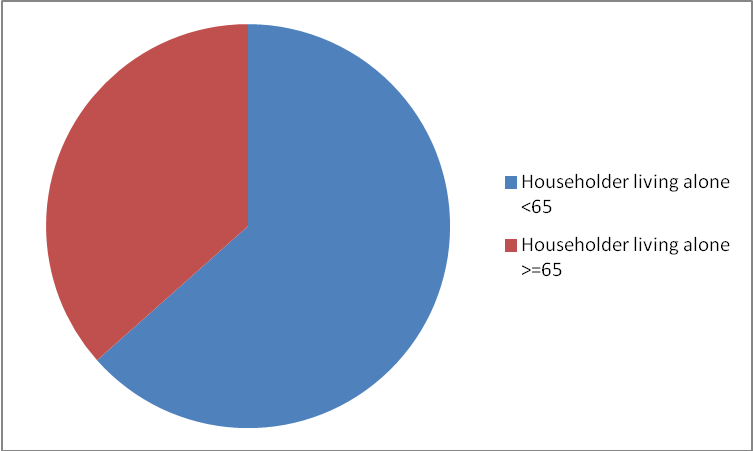
1. **Proportion Charts:**

|  |  |  |
| --- | --- | --- |
| Subcategories | Category total | Overall total |
| Family households | 2,993,023 | 4,445,773 |
| Nonfamily households | 1,452,750 | 4,445,773 |
| Total households | 4,445,773 |  |



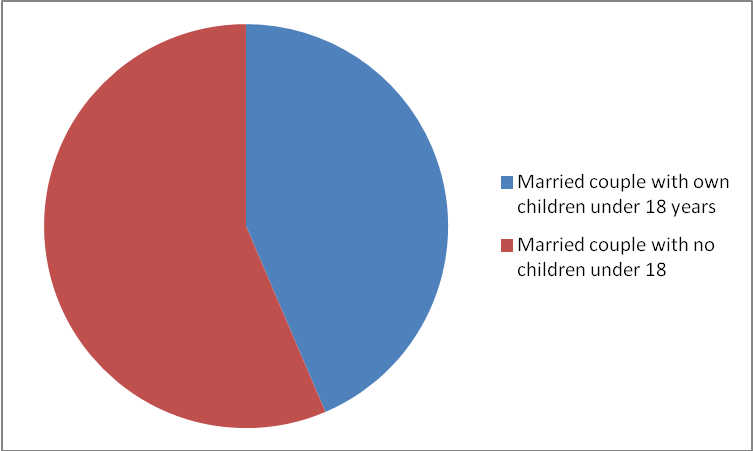
|  |  |  |
| --- | --- | --- |
|  | Number of households | Total number of households living alone |
| Householder living alone <65 | 771,218 | 1,215,614 |
| Householder living alone >=65 | 444,396 | 1,215,614 |

|  |  |  |
| --- | --- | --- |
| Subcategories | Householders living alone <65 | Householders alone >=65 |
| Both households | 771,218 | 444,396 |
| Total households | 1,215,614 | 1,215,614 |

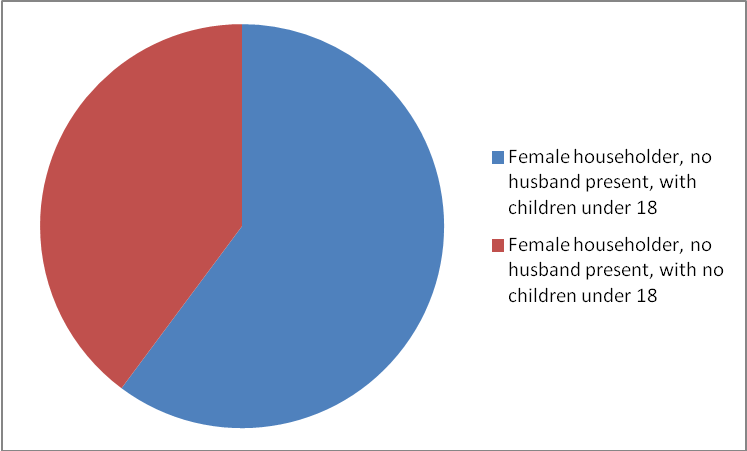


01-11

|  |  |
| --- | --- |
| Subcategories | Total number of households |
| Married couple family total | 2,285,798 |
| Married couple with own children under 18 years | 996,042 |
| Married couple with no children under 18 | 1,289,756 |

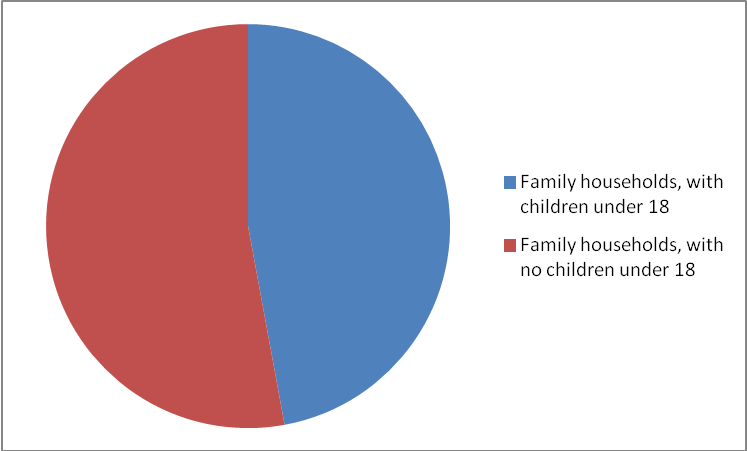


|  |  |
| --- | --- |
| Subcategories | Total number of households |
| Female householder, no husband present | 536,878 |
| Female householder, no husband present, with children under 18 | 323,095 |
| Female householder, no husband present, with no children under 18 | 213,783 |



01-12

|  |  |
| --- | --- |
| Subcategories | Total number of households |
| Family households | 2,993,023 |
| Family households, with children under 18 | 1,409,912 |
| Family households, with no children under 18 | 1,583,111 |

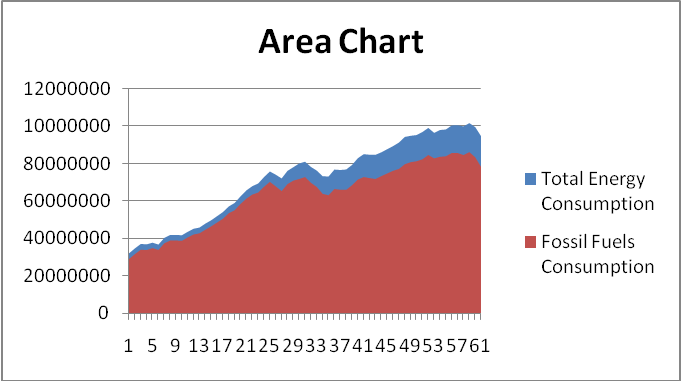


5. A The Excel file *Energy Production & Consumption* provides various energy data since 1949.

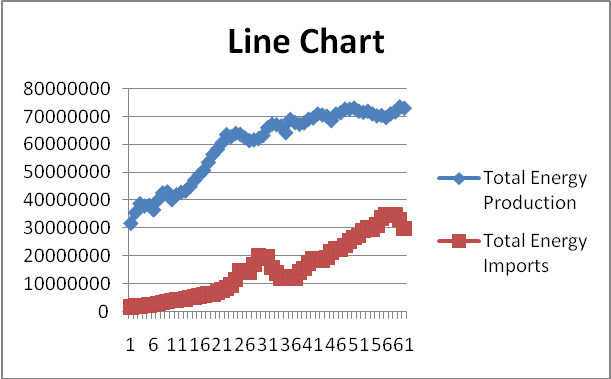
01-13

1. Construct an area chart showing the fossil fuel consumption as a proportion of the total energy consumption.
2. Construct line charts for total energy production and total energy imports.
3. Construct a scatter diagram for total energy consumption and total energy production.

**Answer:**



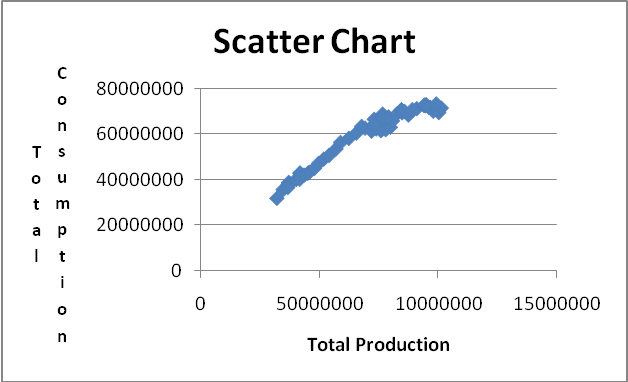
01-14



01-15

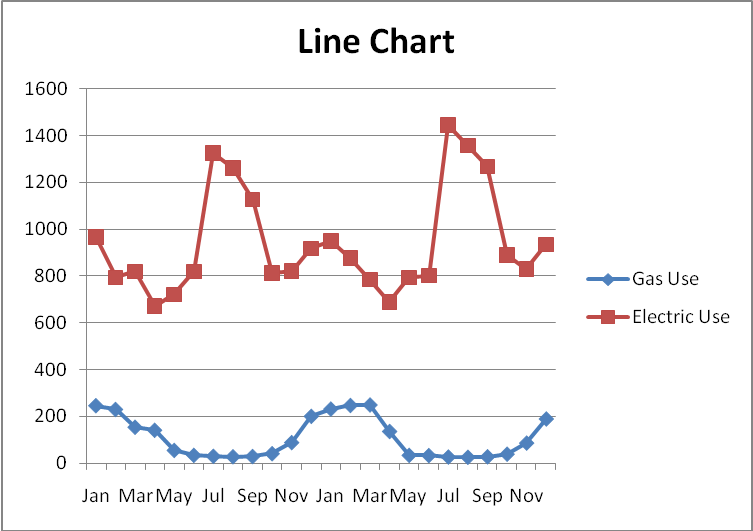
01-15

01-16



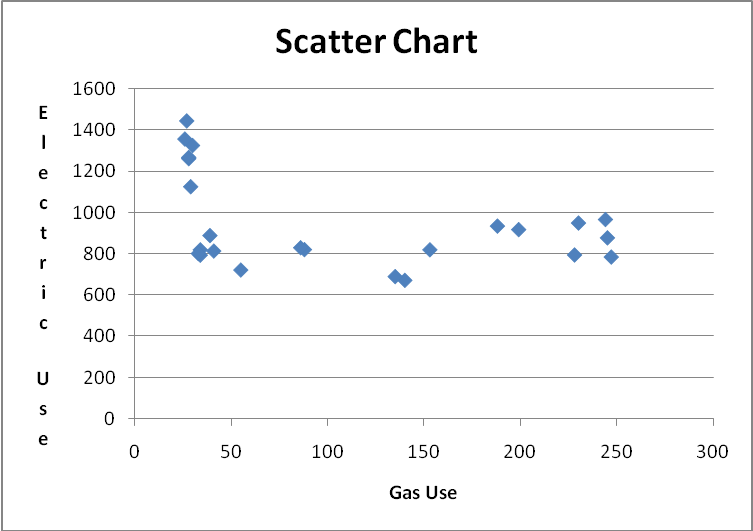
6. Construct whatever charts you deem appropriate to visually display the information contained in the Excel file *Gas and Electric.* What conclusions can you draw from these?

**Answer:**



In general there is a negative relationship between the variables, but the relationship is clearly nonlinear.

01-18

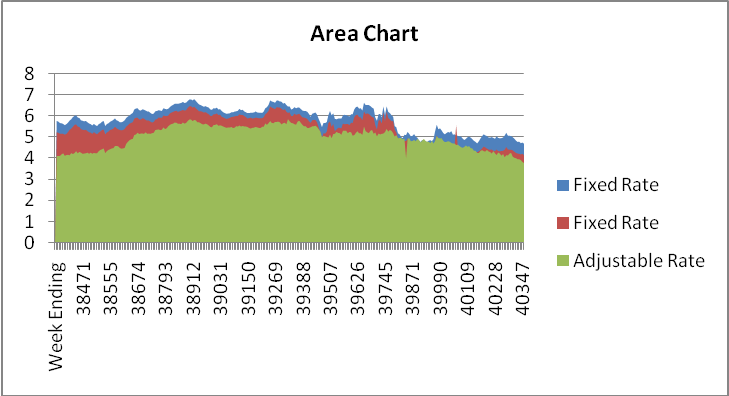
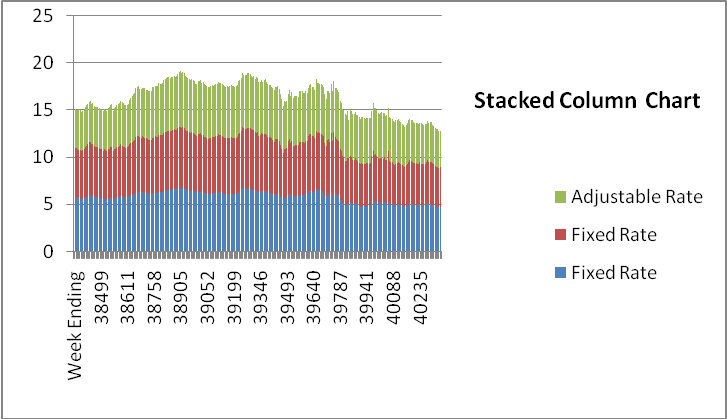
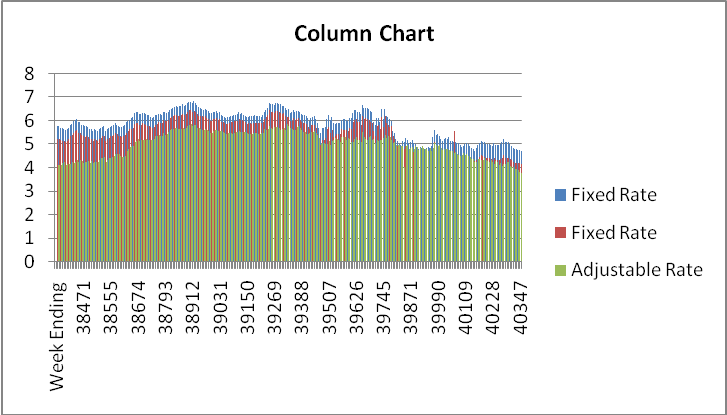


.

01-19

7. Construct whatever charts you deem appropriate to visually display the information contained in the Excel file *Mortgatge Rates.* What conclusions can you draw from these?

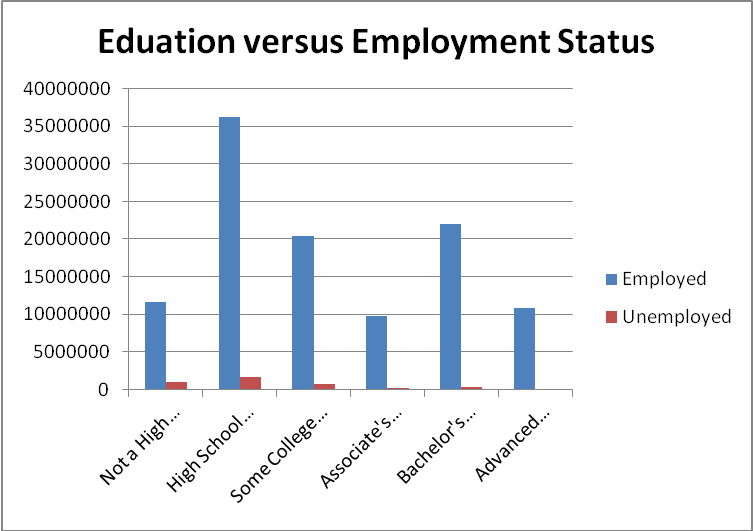
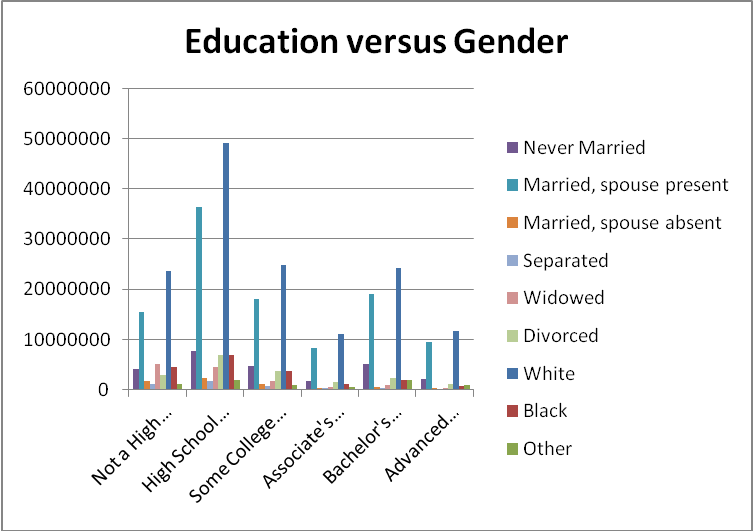
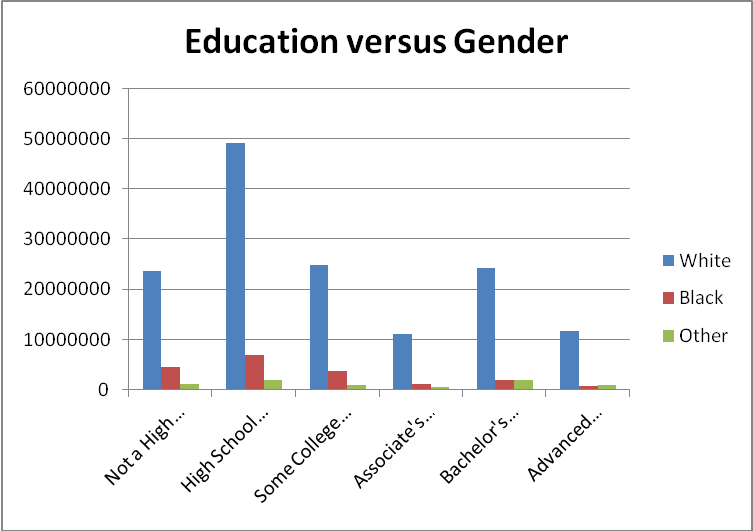
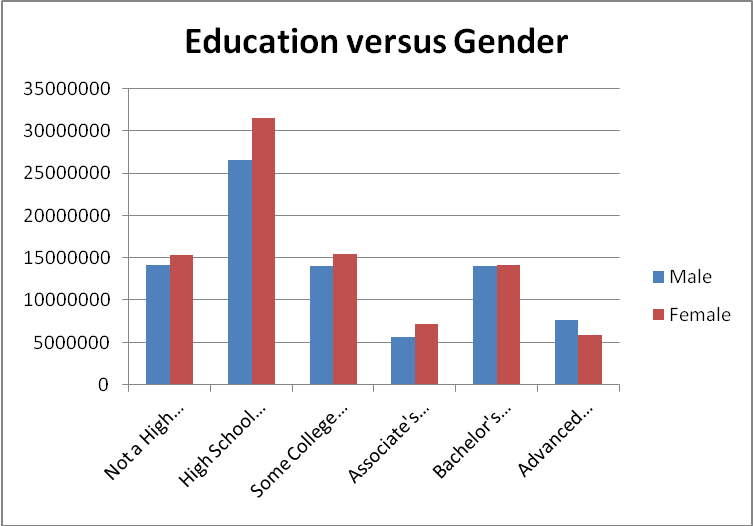
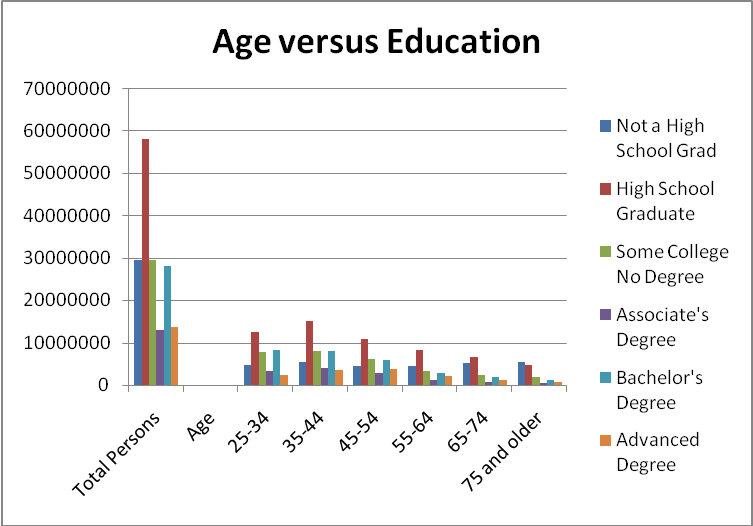
**Answer:**



All the charts provide similar information. The rates are initially rising, reach a peak, and then fall. During the initial part, the 30 year fixed rates and 15 year fixed rates are closer. During the later part, the 15 year fixed rates are closer to the adjustable rates.

8. Construct appropriate charts to visually display the information contained in the Excel file *Census Education Data.*

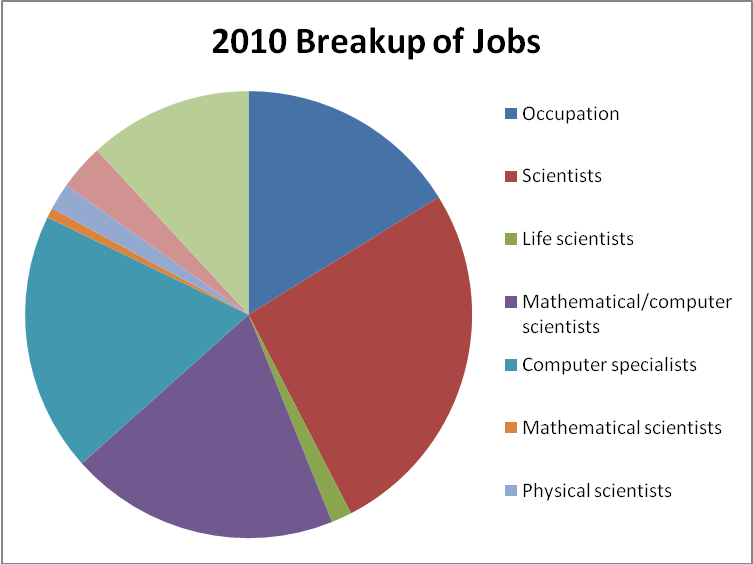
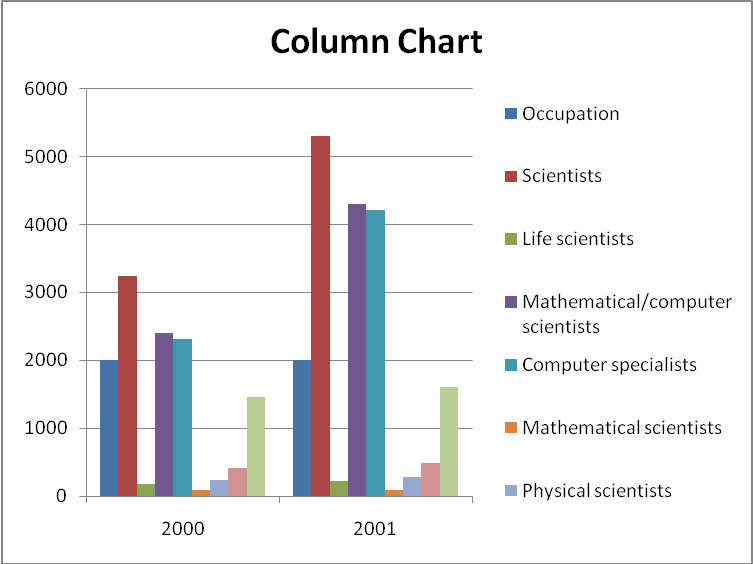
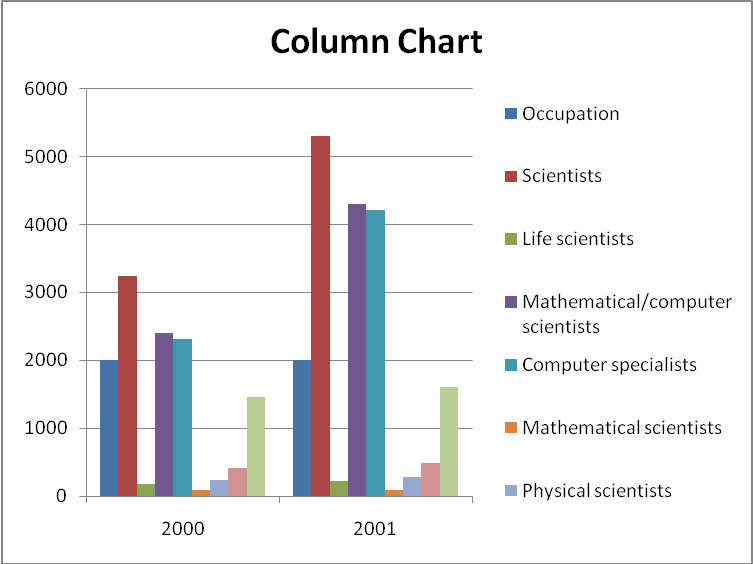
**Answer:**



|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

9. Construct appropriate charts to visually display the information contained in the Excel file *Science and Engineering Jobs.*

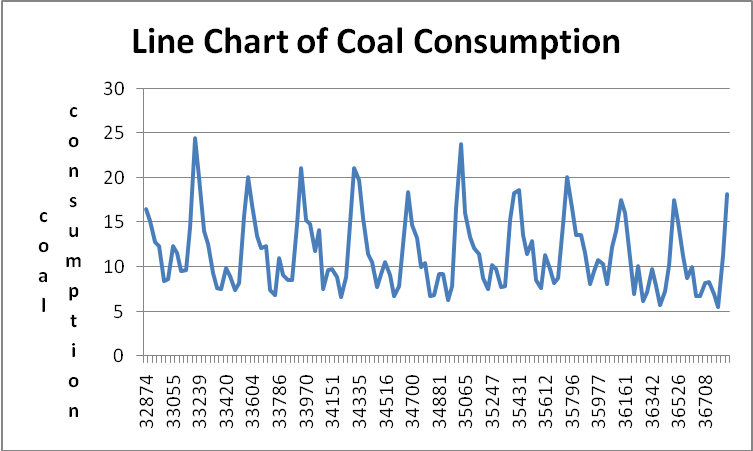
**Answer:**



|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

10. Construct appropriate charts to visually display the information contained in the Excel file *Coal Consumption.*

**Answer:**



The line charts clearly shows the cyclical variations over the years.

11. Modify the Excel file *Hi-Definition Television* to identify televisions that belong to the Bigscreen/Projection and LCD/Plasma category. Use Excel functions to find the minimum and maximum values for the overall scores for each type of television. Count the number of televisions in the Bigscreen/Projection category and the LCD/Plasma category.

**Answer: See Excel Folder.**

**Case**

**A Data Collection and Analysis Project:**

Develop a simple questionnaire to gather data that include a set of both categorical variables and ratio variables. In developing the questionnaire, think about some meaningful questions that you would like to address using the data. The questionnaire should pertain to any subject of interest to you, for example, customer satisfaction with products or school‐related issues, investments, hobbies, leisure activities, and so on—be creative! (Several Web sites provide examples of questionnaires that may help you. You might want to check out www.samplequestionnaire.com or www.examplequestionnaire.com for some ideas.) Aim for a total of 6–10 variables. Obtain a sample of at least 20 responses from fellow students or coworkers. Record the data on an Excel worksheet and construct appropriate charts that visually convey the information you gathered, and draw any conclusions from your data. Then, as you learn new material in Chapters 2–7, apply the statistical tools as appropriate to analyze your data and write a comprehensive report that describes how you drew statistical insights and conclusions, including any relevant Excel output to support your conclusions. (Hint: a good way to embed portions of an Excel worksheet into a Word document is to copy it and then use the Paste Special feature in Word to paste it as a picture. This allows you to size the picture by dragging a corner.)

01-28

**Answer:**

|  |
| --- |
| **Questionnaire** - covers different types of variables, the results should be amenable to PivotTables slicing and dicing, as well as subsequent analysis in Chapters 2 through 7 (although it does not include any time series). |

|  |  |  |
| --- | --- | --- |
|  | | |
| **Example Types of Questions** | **Name of the variable** | **Values** |
| What is your gender? (0 = female; 1 = male) | Gender | 0, 1 |
| What is your height? (5 foot 10 = 5\*12 + 10 = 70 inches) | Height | inches |
| What is your approximate weight? | Weight | pounds |
| What state are you from? | State | OH, Other |
| How many siblings do you have? | Siblings | 0, 1, 2, 3 |
| Do you play any college sports? (0 = no; 1 = yes) | Sports | 0, 1 |
| How many hours per week do you exercise? | Exercise | hours/week |
| How many hours per day do you sleep? | Sleep | hours/day |
| Do you regularly go home on weekends? (0 = no; 1 = yes) | Home | 0, 1 |
| You like the college because of its social and extracurricular activities. (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree) | Sec | 1, 2, 3, 4, 5 |
| How many hours per week do you spend on Facebook? | Online | hours/week |
| How many times per day do you login to Facebook? | Logins | times/day |
| How many friends do you have on Facebook? | Friends | 1, 2, ... |
| How many pictures do you have posted on Facebook? | Pics | 0, 1, 2, .... |
| How many times per week do you go out to socialize (parties, bars, etc.)? | Out | times/week |
| You like the college because of its scholarly focus and academic challenges. (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree) | Sfac | 1, 2, 3, 4, 5 |
| What was your high school GPA? | HS\_GPA | Range 1 - 5 |
| What was your overall SAT score? | SAT |  |
| What is your college GPA? | Coll\_GPA | Range 1 - 5 |
| How much money (including allowance) do you make per week? | Income | dollars/week |
| How many hours per week do you study? | Study | hours/week |
| What is your business major? Acct = accounting, Fin = finance, Mgmt = Management, Mkt = Marketing, BusAdm = Bus Admin, Other | Major | Acct, Fin, Mgmt, Mkt, MIS, Other |
| What starting salary do you expect? | Salary | dollars/year |
|  |  |  |
| Other examples might deal with social activities such as recreation, sports, music, theater interests, etc. | | |

01-29