**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. Describe some ways in which data are used in different business functions.**

**Answer:**

Data is used in many different business functions:

a. Finance and Accounting – the data is the basic element from which a balance sheet is created, and the determination of costs and profits at a company or within a business unit.

b. Marketing – data is used to determine advertising impact, how, when, and where coupons and sales promotions are used by customers, in market research to determine customer satisfaction and where new product interests might lie.

c. Human Resources – data is used to determine employee turnover, attendance, success of orientation programs and the effectiveness of training programs.

d. Strategic planning – data is used to determine which countries a company may want to enter in a market and where to build manufacturing and warehouse facilities.

**4. Explain how a company might use internal sources of data, special studies, and external data bases.**

**Answer:**

Internal sources of data is the information that an organization already has within its own company data bases and is routinely collected by the accounting, marketing, and operations functions. Examples include: production output, material costs, sales, accounts receivables, and customer demographics.

Other data must be generated through special efforts.

External databases are often used for comparative purposes, marketing projects, and economic analyses. These might include population trends, interest rates, industry performance, consumer spending, and international trade data. Such data can be found in annual reports, Standard & Poor’s Compustat data sets, industry trade associations, or government databases.

01-01

**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 statistical thinking? Why is it an important managerial skill?**

01-02

**Answer:**

Statistical thinking is a philosophy of learning and the action for improvement based on three principles:

a. All work occurs in a system of interconnected processes.

b. Variation exists in all processes.

c. Understanding and reducing variation are keys to success.

Statistical thinking is an important management skill because managers need to be able to understand the difference between common and special cause of variation in the business processes that they are responsible for. This type of mindset allow managers to making decisions the help to reduce variation and to deliver more consistent performance over a long term time horizon.

**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.

**12. What types of chart would be best for displaying the data in each of the following data sets on the Companion Website? If several charts are appropriate, state this, but justify your best choice.**

01-03

**a. *Mortgage Rates***

**b. *Census Education Data***

**c. *Consumer Transportation Survey***

**d. *MBA Student Survey***

**e. *Vacation Survey***

**f. *Washington, DC, Weather***

**Answer:**

The types of charts best for the following would be:

1. Mortgage rates – stock chart if the high, low, and close rates are of interest, line chart to show trends over time, and area chart for rates over time.
2. Census Education Data – scatter diagram to show relationships between ad, gender, race, or marital status to type of degree held, pie charts for proportions of degrees, age, and marital status as part of the population, radar chart for multiple dimensions of demographics and a surface chart to plot 3 various dimensions of demographics, and doughnut charts to include more than one set of data.
3. Consumer transportation survey – pie charts to show proportions of demographics and types of vehicles among consumers, and scatter diagrams to show cause and effect between hours/week in car and miles driven.
4. MBA student survey – scatter diagram to plot relationship between age and nights out/week or undergraduate concentration and nights out/week. Pie charts for proportion of international students. A bubble chart for major vs. nights out/week vs. study hours/week.
5. Vacation survey – scatter chart on vacations per year vs. marital status, bar or column charts on number of vacations/year, pie chart on proportions for gender, age, or marital status.
6. Washington DC average temperatures – line chart for temperatures over time, area chart for temperatures over time. Scatter chart for month vs. temperature.

**Problems and Applications**

1. For the Excel file *Credit Approval Decisions*, identify each of the variables as categorical, ordinal, interval, and ratio.

**Answer:**

**Credit Approval Decisions:**

|  |  |
| --- | --- |
| Categorical | **Homeowner** |
| Ratio  | **Credit Score** |
| Ratio | **Years of Credit History** |
| Ratio | **Revolving Balance** |
| Ratio | **Revolving Utilization** |
| Categorical | **Decision** |

01-04

2. A survey handed out to individuals at a major shopping mall in a small Florida city in July asked the following:

• Gender

• Age

• Ethnicity

• Length of residency

• Overall satisfaction with city services (using a scale of 1–5 going from Poor to Excellent)

• Quality of schools (using a scale of 1–5 going from Poor to Excellent)

a. What is the population that the city would want to survey?

b. Would this sample be representative of the population?

c. What types of data would each of the survey items represent?

**Answer:**

1. The population that the city would want to survey would be those people who used city services such as public transportation and people with children in the K -12 school system who were residents of the city.
2. This subset would not represent the entire city because it would not include people without children, visitors who do not tend to use city services, and people with grown children.
3. Types of data:

Gender – categorical

Age – ratio

Ethnicity – categorical

Length of residency – ratio

Overall satisfaction – ordinal

Quality of schools – ordinal

3. Construct a column chart for the data in the Excel file *State Unemployment Rates* to allow comparison of the June rate with the historical highs and lows. Would any other charts be better to visually convey this information? Why or why not?

01-05

**Answer:**

**Column Chart:**

A line graph demonstrates all 3 scenarios

**Line Graph:**

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

01-13

a. Construct an area chart showing the fossil fuel production as a proportion of total energy production.

b. Construct line charts for each of the variables.

c. Construct a line chart showing both the total energy production and consumption during these years.

d. Construct a scatter diagram for total energy exports and total energy production.

e. Discuss what information the charts convey.

**Answer:**

1. **Area Chart:**

**B. Line Charts** for each of the variables:

01-14



01-15

01-15

01-16

**C. Line Chart** for Total Energy Production and Consumption during the years:

 **D. Scatter Diagram:**

01-17

1. The area chart show fossil fuel production over time compared with total energy production. The line charts show trends for each variable. The scatter diagram shows the relationship between energy exports and total energy production.

6. The Excel file *Internet Usage* provides data about users of the Internet.

a. Construct appropriate charts that will allow you to compare any differences due to age or educational attainment.

b. What conclusions can you draw from these charts?

**Answer:**

01-18

1. The possible conclusions would be that more than half of those 45 and under have Internet accounts and those with a Bachelor's degree or higher have the highest proportion of Internet users.

01-19

7. The Excel file *Freshman College Data* provides data from different colleges and branch campuses within one university over four years.

a. Construct appropriate charts that allow you to contrast the differences among the colleges and branch campuses.

b. Write a report to the academic vice president explaining the information.

**Answer:**

a.

01-20

b.

01-21

Dear President Smith,

Attached you will find summaries that compare the branch campus scores of students with that of the main campus students.

The ACT and SAT averages are lower for the branch campus students.

The average ACT scores are 3.75 points less, or 16% lower.

The SAT scores average about 150 points less, or 14% lower.

High school GPA's are comparable, but the percentage of students in the top 10% and top 20% is dramatically lower for the branch campuses.

The 1st year retention rate is also much lower for the branch campus students, 63% compared to 75% for the main campus, approximately.

Sincerely,

Mary Jones, University statistician.

8. Construct whatever charts you deem appropriate to convey comparative information on the two categories of televisions in the Excel file Hi‐Definition Televisions. What conclusions can you draw from these?

**Answer:**

|  |  |
| --- | --- |
| Comparison of Big Screen/Projection with LCD/Plasma TV's | **Price** |
| Bigscreen/Projection |  $ 2,124.00  |
| LCD/Plasma |  $ 2,472.00  |

01-22

|  |  |
| --- | --- |
| Comparison of Big Screen/Projection with LCD/Plasma TV's | **Screen Size** |
| Bigscreen/Projection | 52.76 |
| LCD/Plasma | 29.08 |



|  |  |
| --- | --- |
| Comparison of Big Screen/Projection with LCD/Plasma TV's | **Sound Quality** |
| Bigscreen/Projection | 4.71 |
| LCD/Plasma | 5 |



|  |  |
| --- | --- |
| Comparison of Big Screen/Projection with LCD/Plasma TV's | **Ease of Use** |
| Bigscreen/Projection | 3.4 |
| LCD/Plasma | 3 |

01-23

|  |  |
| --- | --- |
| Comparison of Big Screen/Projection with LCD/Plasma TV's | **Overall Score** |
| Bigscreen/Projection | 50.76 |
| LCD/Plasma | 52 |

**Conclusions:**

The average price for an LCD/Plasma television is nearly $350 more than a big screen/projection TV. Big screen/projection televisions have about 40% larger screens and are viewed as easier to use. Sound quality and overall score are higher for LCD/plasma televisions.

9. Construct whatever charts you deem appropriate to convey comparative information on deaths by major causes in the Excel file *Death Cause Statistics*. What conclusions can you draw from these?

**Answer:**

a. Causes of Death 1960 -2007

|  |  |
| --- | --- |
| **Types of Death** | **Deaths (thousands)** |
| Natural Causes | 786.441667 |
| Accident | 46.4625 |
| Suicide | 12.24 |
| Homicide | 7.9667 |
| Total | 853.110867 |



b. Average Number of deaths by heart disease and cancer vs. all other natural causes 1960 – 2007.

01-24

|  |  |
| --- | --- |
| **Types of Death** | **Deaths (thousands)** |
| Heart Disease | 385.3 |
| Cancer | 201.88 |
| All Other | 199.26 |

c. Heart Disease Deaths from 1960 – 2007

**Conclusions:** The largest proportion of deaths is by natural causes, 92% of all deaths.

Heart disease and cancer account for 75% of all natural deaths.

Deaths by heart disease have declined from over 500,000 per year to about 200,000 per year.

10. Construct an appropriate chart to show the proportion of funds in each investment category in the Excel file *Retirement Portfolio*.

**Answer:**

01-25

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Mutual Fund** | **Type** | **Net Asset Value** | **Shares** | **Total Value** |
| Fidelity Capital Appreciation | Large Cap |  $25.14  | 1225 |  $30,796.50  |
| Fidelity Contrafund | Large Cap |  $55.32  | 1500 |  $82,980.00  |
| Fidelity Equity Income | Large Cap |  $51.00  | 1400 |  $71,400.00  |
| Fidelity Export & Multinational | Large Cap |  $19.11  | 600 |  $11,466.00  |
| Fidelity Strategic Large Cap Value | Large Cap |  $12.68  | 1800 |  $22,824.00  |
| Total for category |  |  |  |  **$219,466.50**  |
| Fidelity Mid Cap Stock | Mid Cap |  $22.38  | 600 |  $13,428.00  |
| Fidelity Value | Mid Cap |  $69.92  | 850 |  $59,432.00  |
| Total for category |  |  |  |  **$72,860.00**  |
| Fidelity Small Cap Independence | Small Cap |  $19.02  | 1000 |  $19,020.00  |
| Fidelity Low Priced Stock | Small Cap |  $39.54  | 400 |  $15,816.00  |
| Total for category |  |  |  |  **$34,836.00**  |
| Fidelity Puritan | Blended |  $19.87  | 1500 |  $29,805.00  |
| Fidelity Fidelity Fund | Blended |  $30.44  | 675 |  $20,547.00  |
| Total for category |  |  |  |  **$50,352.00**  |
| Fidelity Mortgage Securities | Bond |  $11.23  | 700 |  $7,861.00  |
| Fidelity Strategic Income | Bond |  $10.59  | 400 |  $4,236.00  |
| Total for category |  |  |  |  **$12,097.00**  |

|  |  |
| --- | --- |
| **Asset class** | **Total value** |
| Large cap |  $219,466.50  |
| Mid cap |  $72,860.00  |
| Small cap |  $34,836.00  |
| Blended |  $50,352.00  |
| Bond |  $12,097.00  |
|  | $389,611.50  |

01-26

11. Modify the Excel file *Major League Baseball* to identify teams that have either a winning or losing record. Use Excel functions to find the minimum and maximum values for each type of data and count the number of teams with winning and losing records.

**Answer:**

|  |  |  |  |
| --- | --- | --- | --- |
| **TEAM** | **Won** | **Lost** | Win % |
| Philadelphia | 97 | 65 | 60% |
| Tampa Bay | 96 | 66 | 59% |
| NY Yankees | 95 | 67 | 59% |
| Minnesota | 94 | 68 | 58% |
| San Francisco | 92 | 70 | 57% |
| Atlanta | 91 | 71 | 56% |
| Cincinnati | 91 | 71 | 56% |
| San Diego | 90 | 72 | 56% |
| Texas | 90 | 72 | 56% |
| Boston | 89 | 73 | 55% |
| Chicago Sox | 88 | 74 | 54% |
| St. Louis | 86 | 76 | 53% |
| Toronto | 85 | 77 | 52% |
| Colorado | 83 | 79 | 51% |
| Detroit | 81 | 81 | 50% |
| Oakland | 81 | 81 | 50% |
| Florida | 80 | 82 | 49% |
| LA Angels | 80 | 82 | 49% |
| LA Dodgers | 80 | 82 | 49% |
| NY Mets | 79 | 83 | 49% |
| Milwaukee | 77 | 85 | 48% |
| Houston | 76 | 86 | 47% |
| Chicago Cubs | 75 | 87 | 46% |
| Cleveland | 69 | 93 | 43% |
| Washington | 69 | 93 | 43% |
| Kansas City | 67 | 95 | 41% |
| Baltimore | 66 | 96 | 41% |
| Arizona | 65 | 97 | 40% |
| Seattle | 61 | 101 | 38% |
| Pittsburgh | 57 | 105 | 35% |

The teams with winning records are highlighted in blue and the teams with losing records are highlighted in green.

01-27

|  |  |  |
| --- | --- | --- |
|  | **Max value** | **Min Value** |
| **Won** | 97 | 57 |
| **Lost** | 105 | 65 |
| Win % | 0.6 | 0.35 |
| **Runs** | 859 | 513 |
| **Hits** | 1556 | 1274 |
| **Doubles** | 358 | 227 |
| **Triples** | 54 | 16 |
| **Home Runs** | 257 | 101 |
| **Runs Batted In** | 823 | 485 |
| **Earned Run Average** | 5 | 3.36 |
| **Strike Outs** | 1529 | 905 |
| **Walks** | 672 | 415 |

|  |  |
| --- | --- |
| The count of winning teams is: | 14 |
| The count of teams with equal or losing records: | 16 |
| The count of teams with losing records: | 14 |

**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