

A Passion for Compassion

In the 19th century, statistics was not widely seen as an applicable skill. That is, until Florence Nightingale came onto the scene. When she arrived at the front line of the Crimean War, she was appalled by the situation. The mortality rate was too high, and the hospitals were in complete disarray. She immediately set about organizing what little records were kept and started to gather a lot of new data. Upon analyzing this new data, she discovered that the majority of deaths that were occurring in British military hospitals were due to preventable diseases. Using this new information, Nightingale was able to present a case to Parliament for improving the sanitary practices in British hospitals. She utilized data analysis and visualization to literally save thousands of lives, and in the process, her “rose diagram”, also known as a “coxcomb chart”, became an iconic data visualization.

A pie chart showing where government revenues originate is given in Figure 3.2.11. In order to determine from the pie chart how much was received in a particular category, multiply the total amount by the percentage given for that category in the pie chart. For example, to find the amount of government receipts contributed by individual income taxes, multiply the total amount of government receipts (\$4.738 trillion) by the percentage contributed by individual income taxes (49.3%). This means that $4.738 \cdot 0.493 = \$2.3358$ trillion of all government receipts come from individual income taxes.

Although pie charts are useful displays of categorical data, they have their limitations. Once the number of categories rises above ten, the information conveyed by a pie chart is less meaningful. As Figure 3.2.12 and Figure 3.2.13 demonstrate, when the number of employees increases from 5 to 40, using a pie chart to show the percentage of sales for each employee is not as informative.

Percentage of Sales
(5 Employees)

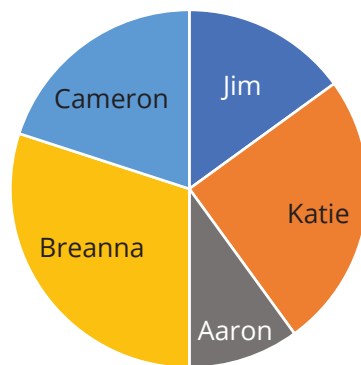


Figure 3.2.12

Percentage of Sales
(40 Employees)

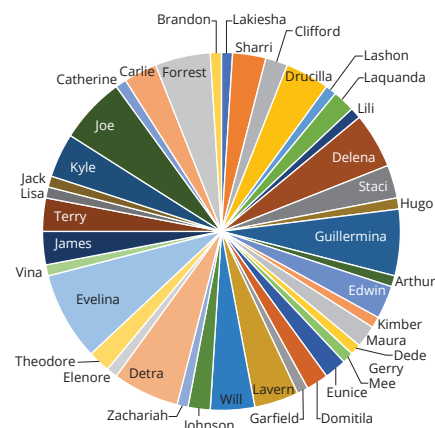


Figure 3.2.13

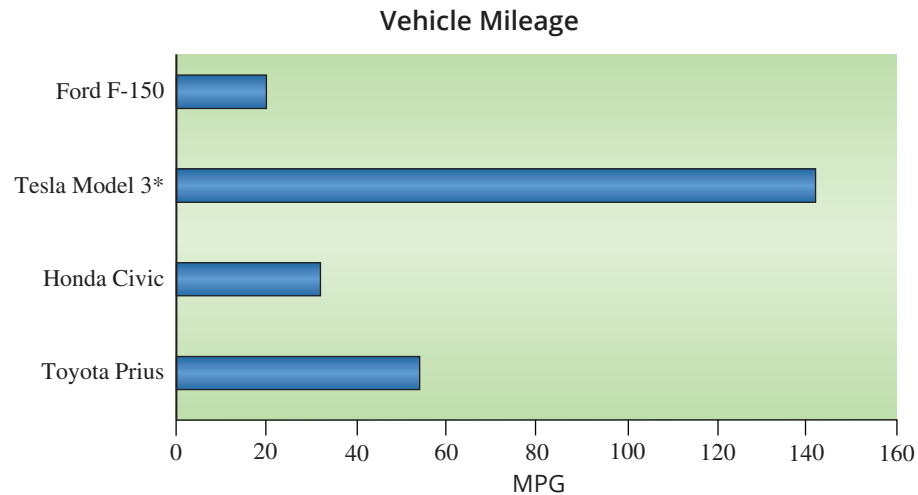
3.2 Exercises

Basic Concepts

1. What are some benefits of graphing?
2. What is the major disadvantage of graphing?
3. Describe the types of data that a bar chart would be useful in displaying.
4. Where should miscellaneous categories be displayed in a bar chart?
5. Explain how axis scales on bar charts can be misleading.
6. What is a stacked bar chart?
7. In what circumstances would a stacked bar chart be preferred over a standard bar chart?
8. What is a pie chart?
9. In Figure 3.2.7, the displayed data is not normalized. How could you normalize this data?

Exercises

10. A consumer magazine uses bar charts to compare four popular brands of automobiles. This particular bar chart represents a comparison of the miles per gallon (mpg) for the four brands.



*MPGe, or miles per gallon equivalent, for electric vehicles

- What is wrong with this picture?
 - Evaluate the bar chart using the guidelines suggested in the section on the aesthetics of bar chart construction.
11. The following frequency distribution is for a portion of the Employee Satisfaction data set.

Frequency Distribution of Employee Salaries			
Department	High Salary	Medium Salary	Low Salary
Accounting	74	358	335
Management	225	180	225
Marketing	80	402	376
Sales	269	2099	1772
Technical	201	1372	1147
Total	849	4411	3855

- Construct a bar chart for the number of employees with high salaries by department.
- Construct a bar chart for the number of employees with medium salaries by department.
- Construct a bar chart for the number of employees with low salaries by department.
- Construct a stacked bar chart for the number of employees with low, medium, and high salaries by department.
- Construct a pie chart for the number of employees with high salaries.
- Construct a pie chart for the number of employees with medium salaries.
- Construct a pie chart for the number of employees with low salaries.

Data

stat.hawkeslearning.com under **Discovering Statistics and Data, Fourth Edition > Data > Employee Satisfaction**

12. The Centers for Disease Control and Prevention (CDC) reports information about the leading cause of death for the United States each year. The frequency table displays the data from 2021.⁶

Leading Causes of Death in the U.S., 2021	
Cause of Death	Frequency
Heart disease	695,547
Cancer	605,213
COVID-19	416,893
Accidents (unintentional injuries)	224,935
Stroke (cerebrovascular diseases)	162,890
Chronic lower respiratory diseases	142,342
Alzheimer's disease	119,399
Diabetes	103,294
Chronic liver disease and cirrhosis	56,585
Nephritis, nephrotic syndrome, and nephrosis	54,358

- Construct a bar chart for the leading causes of death in the United States in 2021.
 - Did you create a Pareto chart? Explain your response.
 - What did you learn from the chart?
13. Consider the following data regarding professions with the highest projected percent of change in employment for the years 2021 through 2031.⁷

Occupation Growth Rates	
Occupation	Projected Increase 2021–2031
Nurse practitioners	46%
Wind turbine service technicians	44%
Ushers, lobby attendants, ticket takers	41%
Motion picture projectionist	40%
Cooks, restaurant	37%
Data scientist	36%
Athletes and sports competitors	36%
Information security analysts	35%

- Construct a bar chart for the projected growth rates of the various occupations.
- What did you learn from the chart?

14. Consider the following data regarding the payment methods which consumers reported using during October 2020. BANP: bank account number payment. OBBP: Online banking bill pay. Other includes Paypal, account-to-account transfers, mobile payments, and deductions from income.

Payment Methods	
Method of Payment	Relative Frequency
Cash	19%
Check	7%
Credit Card	27%
Debit Card	28%
Pre-paid Card	2%
BANP	7%
OBBP	5%
Other	5%

- Construct a bar chart for the relative frequencies of the various methods of payment.
- Construct a pie chart for the relative frequencies of the various methods of payment.
- Comment on any information about the relative frequencies of the various methods which you were able to ascertain by examining the charts.

3.3 Displaying Quantitative Data Graphically

Quantitative graphs are powerful tools used in data visualization to represent and analyze numerical data. They come in various forms, including bar graphs, line graphs, histograms, scatter plots, dot plots, pie charts, and choropleth maps, each with its unique utility. These graphs help to summarize large data sets, display patterns, trends, and correlations, and make complex information more comprehensible. Examining raw data as compared to viewing a well-designed graph of that data is akin to hearing a symphony described in words versus actually listening to the music; the latter offers an immediate, rich understanding that simple description can never capture.

Histograms

A **histogram** is a common graphical method that reveals the distribution of a set of data. Histograms are often constructed based on frequency distributions of quantitative data. They look similar to bar graphs but are used to analyze quantitative data rather than qualitative data.

Histograms literally show the shape of the data by graphically displaying a frequency distribution.