Chapter 1 – Stats Starts Here

1. **Voters.** The response is a categorical variable.

2. **Job hunting.** The answer is a categorical variable.

3. **Medicine.** The company is studying a quantitative variable.

4. **Stress.** The researcher is studying a quantitative variable.

5. **The News.** Answers will vary.

6. **The Internet.** Answers will vary.

7. **Bicycle Safety.** *Who* – 2,500 cars. *What* – Distance from the bicycle to the passing car (in inches). *Population of interest* – All cars passing bicyclists.

8. **Investments.** *Who* – 30 similar companies. *What* – 401(k) employee participation rates (in percent). *Population of interest* – All similar companies.

9. **Honesty.** *Who* – Workers who buy coffee in an office. *What* – amount of money contributed to the collection tray. *Population of interest* – All people in honor system payment situations.

10. **Gaydar.** *Who* – 40 undergraduate women. *What* – Whether or not the women could identify the sexual orientation of men based on a picture. *Population of interest* – All women.

11. **Blindness.** *Who* – 24 patients. *What* – Whether the patient had Stargardt's disease or dry age-related macular degeneration, and whether or not the stem cell therapy was effective in treating the condition. *Population of interest* – All people with these eye conditions.

12. **Molten iron.** *Who* – 10 crankshafts at Cleveland Casting. *What* – The pouring temperature (in degrees Fahrenheit) of molten iron. *Population of interest* – All crankshafts at Cleveland Casting.

13. **Weighing bears.** *Who* – 54 bears. *What* – Weight, neck size, length (no specified units), and sex. *When* – Not specified. *Where* – Not specified. *Why* - Since bears are difficult to weigh, the researchers hope to use the relationships between weight, neck size, length, and sex of bears to estimate the weight of bears, given the other, more observable features of the bear.
How – Researchers collected data on 54 bears they were able to catch. *Variables* – There are 4 variables; weight, neck size, and length are quantitative variables, and

sex is a categorical variable. No units are specified for the quantitative variables. *Concerns* – The researchers are (obviously!) only able to collect data from bears they were able to catch. This method is a good one, as long as the researchers believe the bears caught are representative of all bears, in regard to the relationships between weight, neck size, length, and sex.

14. **Schools.** *Who* – Students. *What* – Age (probably in years, though perhaps in years and months), race or ethnicity, number of absences, grade level, reading score, math score, and disabilities/special needs. *When* – This information must be kept current. *Where* – Not specified. *Why* – Keeping this information is a state requirement. *How* – The information is collected and stored as part of school records. *Variables* – There are seven variables. Race or ethnicity, grade level, and disabilities/special needs are categorical variables. Number of absences, age, reading test score, and math test score are quantitative variables. *Concerns* – What tests are used to measure reading and math ability, and what are the units of measure for the tests?

15. **Arby's menu.** *Who* – Arby's sandwiches. *What* – type of meat, number of calories (in calories), and serving size (in ounces). *When* – Not specified. *Where* – Arby's restaurants. *Why* – These data might be used to assess the nutritional value of the different sandwiches. *How* – Information was gathered from each of the sandwiches on the menu at Arby's, resulting in a census. *Variables* – There are three variables. Number of calories and serving size are quantitative variables, and type of meat is a categorical variable.

16. **Age and party.** *Who* – 20,392 Americans. *What* – Gender, age (in years), race, party affiliation, whether or not the person was of Hispanic ethnicity, education, region, the number of adults in the household, and phone status. *When* – 2011. *Where* – United States. *Why* – The information was gathered for presentation in a Gallup public opinion poll. *How* – Poll. *Variables* – There are nine variables. Gender, race, party affiliation, whether or not the person is of Hispanic ancestry, education, region, and phone status are categorical variables, and age and number of adults are quantitative variables.

17. **Babies.** *Who* – 882 births. *What* – Mother's age (in years), length of pregnancy (in weeks), type of birth (caesarean, induced, or natural), level of prenatal care (none, minimal, or adequate), birth weight of baby (unit of measurement not specified, but probably pounds and ounces), gender of baby (male or female), and baby's health problems (none, minor, major).
When – 1998-2000. *Where* – Large city hospital. *Why* – Researchers were investigating the impact of prenatal care on newborn health. *How* – It appears that they kept track of all births in the form of hospital records, although it is not specifically stated. *Variables* – There are three quantitative variables: mother's age, length of pregnancy, and birth weight of baby. There are four categorical variables: type of birth, level of prenatal care, gender of baby, and baby's health problems.

18. **Flowers.** *Who* – 385 species of flowers. *What* – Date of first flowering (in days). *When* – Not specified. *Where* – Southern England. *Why* – The researchers believe that this indicates a warming of the overall climate. *How* – Not specified. *Variables* – Date of first flowering is a quantitative variable. *Concerns* - Hopefully, date of first flowering was measured in days from January 1, or some other convention, to avoid problems with leap years.

19. **Herbal medicine.** *Who* – experiment volunteers. *What* – herbal cold remedy or sugar solution, and cold severity. *When* – Not specified. *Where* – Major pharmaceutical firm. *Why* – Scientists were testing the efficacy of an herbal compound on the severity of the common cold.
How – The scientists set up a controlled experiment. *Variables* – There are two variables. Type of treatment (herbal or sugar solution) is categorical, and severity rating is quantitative. *Concerns* – The severity of a cold seems subjective and difficult to quantify. Also, the scientists may feel pressure to report negative findings about the herbal product.

20. **Vineyards.** *Who* – Vineyards. *What* – Size of vineyard (in acres), number of years in existence, state, varieties of grapes grown, average case price (in dollars), gross sales (probably in dollars), and percent profit. *When* – Not specified. *Where* – United States. *Why* – Business analysts hoped to provide information that would be helpful to producers of American wines. *How* – Not specified. *Variables* – There are five quantitative variables and two categorical variables. Size of vineyard, number of years in existence, average case price, gross sales, and percent profit are quantitative variables. State and variety of grapes grown are categorical variables.

21. **Streams.** *Who* – Streams. *What* – Name of stream, substrate of the stream (limestone, shale, or mixed), acidity of the water (measured in pH), temperature (in degrees Celsius), and BCI (unknown units). *When* – Not specified. *Where* – Upstate New York. *Why* – Research is conducted for an Ecology class. *How* – Not specified. *Variables* – There are five variables. Name and substrate of the stream are categorical variables, and acidity, temperature, and BCI are quantitative variables.

22. **Fuel economy.** *Who* – Every model of automobile in the United States. *What* – Vehicle manufacturer, vehicle type, weight (probably in pounds), horsepower (in horsepower), and gas mileage (in miles per gallon) for city and highway driving. *When* – This information is collected currently. *Where* – United States. *Why* – The Environmental Protection Agency uses the information to track fuel economy of vehicles. *How* – The data is collected from the manufacturer of each model. *Variables* – There are six variables. City mileage, highway mileage, weight, and horsepower are quantitative variables. Manufacturer and type of car are categorical variables.

23. Refrigerators. *Who* – 102 refrigerators. *What* – Brand, cost (probably in dollars), size (in cu. ft.), temperature performance, noise, ease of use, energy efficiency, estimated annual energy cost (probably in dollars), overall rating, exterior height, exterior length, and exterior width (all dimensions probably measured in inches). *When* – 2012. *Where* – United States. *Why* – The information was compiled to provide information to the readers of *Consumer Reports. How* – Not specified. *Variables* – There are 11 variables. Brand, temperature performance, overall rating, noise, ease of use, and energy efficiency are categorical variables. Cost, size, estimated energy cost, exterior height, exterior length, and exterior width are quantitative variables.

24. Walking in circles. *Who* – 32 volunteers. *What* – Sex, height, handedness, the number of yards walked before going out of bounds, and the side of the field on which the person walked out of bounds. *When* – Not specified. *Where* – Not specified. *Why* – The researcher was interested in whether people walk in circles when lost. *How* – Data were collected by observing the people on the field, as well as by measuring and asking the participants. *Variables* – There are 5 variables. Sex, handedness, and side of the field are categorical variables. Height and number of yards walked are quantitative variables.

25. Kentucky Derby 2012. *Who* – Kentucky Derby races. *What* – Year, winner, jockey, trainer, owner, and time (in minutes, seconds, and hundredths of a second. *When* – 1875 – 2012. *Where* – Churchill Downs, Louisville, Kentucky. *Why* – It is interesting to examine the trends in the Kentucky Derby. *How* – Official statistics are kept for the race each year. *Variables* – There are 6 variables. Winner, jockey, trainer and owner are categorical variables. Date and duration are quantitative variables.

26. Indy 2012. *Who* – Indy 500 races. *What* – Year, driver, time (in minutes, seconds, and hundredths of a second), and speed (in miles per hour). *When* – 1911 – 2012. *Where* – Indianapolis, Indiana. *Why* – It is interesting to examine the trends in Indy 500 races.
How – Official statistics are kept for the race every year. *Variables* – There are 4 variables. Driver is a categorical variable. Year, time, and speed are quantitative variables.