*Unit 2- Descriptive Statistics

Content Area: **Mathematics**
Course(s):
Time Period: **November**
Length: **26 Blocks**
Status: **Published**

Enduring Understandings

Data can be organized and displayed in a variety of ways.

Understanding the distribution of data is important to determine how to analyze the data.

Describing the variation of data is as important as defining the center of a data set.

Correlation does not mean causation.

Essential Questions

What method of displaying data would best represent my purpose?

Why can technology support but not replace our mathematics skills and understanding?

How can we compare data that is measured differently?

What can we learn from correlated bivariate data?

Content

Topics Addressed:

Analyzing Categorical Data

Displaying Quantitative Data with Graphs

Describing Quantitative Data with Numbers

Describing Location in a Distribution

Density Curves and Normal Distribution

Scatterplots and Correlation

Least Square Regression

Skills

Analyzing Categorical Data:

Display categorical data with a bar graph.

Decide whether it would be appropriate to make a pie chart.

Identify what makes some graphs of categorical data deceptive.

Calculate and display the marginal distribution of a categorical variable from a two-way table.

Calculate and display the conditional distribution of a categorical variable for a particular value of the other categorical variable in a two-way table.

Describe the association between two categorical variables by comparing appropriate conditional distributions.

Displaying Quantitative Data with Graphs:

Make and interpret dot plots and stemplots of quantitative data.

Make and interpret dot plots and stemplots of quantitative data.

Describe the overall pattern (shape, center, and spread) of a distribution and identify any major departures from the pattern (outliers).

Identify the shape of a distribution from a graph a roughly symmetric or skewed.

Make a interpret histograms of quantitative data.

Compare distributions of quantitative data using dot plots, stemplots, or histograms.

Describing Quantitative Data with Numbers:

Calculate measures of center (mean and median).

Calculate and interpret measures of spread (range, IQR, and standard deviation)

Choose the most appropriate measure of center and spread in a given setting.

Identify outliers using the 1.5(IQR) rule.

Make and interpret boxplots of quantitative data.

Use appropriate graphs and numerical summaries to compare distributions of quantitative variables.

Decribing Location in a Distribution:

Find and interpret the percentile of an individual value within a distribution of data.

Estimate percentiles and individual values using a cumulative relative frequency graph.

Find and interpret the standardized score (z-score) of an individual value within a distribution of data.

Describe the effect of adding, subtracting, multiplying by, or dividing by a constant on the shape, center, and spread of a distribution of data.

Density Curves and the Normal Distribution:

Estimate the relative locations of the median and mean on a density curve.

Use the 68-95-99.7 rule to estimate areas (proportions of values) in a Normal distribution.

Use technology to find the proportion of z-values in a specified interval or a z-score from a percentile in the standard Normal distribution.

Use technology to find the proportion of values in a specified interval or the value that corresponds to a given percentile in any Normal distribution.

Determine whether a distribution of data is approximately Normal from graphical and numerical evidence.

Scatterplots and Correlation:

Identify explanatory and response variables in situations where one variable helps to explain or influences the other.

Make a scatterplot to display the relationship between two quantitative variables.

Describe the direction, form, and strength of a relationship displayed in a scatterplot and recognize outliers in a scatterplot.

Interpret the correlation.

Understand basic properties of correlation, including how the correlation is influenced by outliers.

Use technology to calculate correlation.

Explain why association does not imply causation.

Least Squares Regression:

Interpret the slope and y-intercept of a least squares regression line.

Use the least squares regression line to predict y for a given x.

Explain the dangers of extrapolation.

Calculate and interpret residuals.

Explain the concept of least squares.

Determine the equation of a least-squares regression line using technology or computer output.

Construct and interpret residual plots to assess whether a linear model is appropriate.

Interpret the standard deviation of the residuals and r^2 and use these values to assess how well the least squares regression line models the relationship between two variables.

Describe how the slope, y intercept, standard deviation of the residuals, and r^2 are influenced by outliers.

Find the slope and y intercept of of the least squares regression line from the means and standard deviations of x and y and their correlation.

Resources

Rossman-Chance Applet Collection

StatsMonkey

Rice Virtual Lab in Statistics

Khan Academy Mission: AP Statistics

Standards

NJSLS 2016

Statistics and Probability

INTERPRETING CATEGORICAL AND QUANTITATIVE DATA

A. Summarize, represent, and interpret data on a single count or measurement variable

1. Represent data with plots on the real number line (dot plots, histograms, and box plots).

2. Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

3. Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible

effects of extreme data points (outliers).

B. Summarize, represent, and interpret data on two categorical and quantitative variables

4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

5. Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.

INTERPRETING CATEGORICAL AND QUANTITATIVE DATA

A. Summarize, represent, and interpret data on a single count or measurement variable

4. Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.

B. Summarize, represent, and interpret data on two categorical and quantitative variables

6. Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

a. Fit a function to the data (including with the use of technology); use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear and exponential models.

b. Informally assess the fit of a function by plotting and analyzing residuals, including with the use of technology.

c. Fit a linear function for a scatter plot that suggests a linear association.

C. Interpret linear models

7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

8. Compute (using technology) and interpret the correlation coefficient of a linear fit.

9. Distinguish between correlation and causation.

Mathematics | Standards for Mathematical Practice

2 Reason abstractly and quantitatively.

Mathematically proficient students make sense of quantities and their relationships in problem situations. They bring two complementary abilities to bear on problems involving quantitative relationships: the ability to decontextualize—to abstract a given situation and represent it symbolically and manipulate the representing symbols as if they have a life of their own, without necessarily attending to their referents—and the ability to contextualize, to pause as needed during the manipulation process in order to probe into the referents for the symbols involved. Quantitative reasoning entails habits of creating a coherent representation of the problem at hand; considering the units involved; attending to the meaning of quantities, not just how to compute them; and knowing and flexibly using different properties of operations and objects.

4 Model with mathematics.

Mathematically proficient students can apply the mathematics they know to solve problems arising in everyday life, society, and the workplace. In early grades, this might be as simple as writing an addition equation to describe a situation. In middle grades, a student might apply proportional reasoning to plan a school event or analyze a problem in the community. By high school, a student might use geometry to solve a design problem or use a function to describe how one quantity of interest depends on another. Mathematically proficient students who can apply what they know are comfortable making assumptions and approximations to simplify a complicated situation, realizing that these may need revision later. They are able to identify important quantities in a practical situation and map their relationships using such tools as diagrams, two-way tables, graphs, flowcharts and formulas. They can analyze those relationships mathematically to draw conclusions. They routinely interpret their mathematical results in the context of the situation and reflect on whether the results make sense, possibly improving the model if it has not served its purpose.

5 Use appropriate tools strategically.

Mathematically proficient students consider the available tools when solving a mathematical problem. These tools might include pencil and paper, concrete models, a ruler, a protractor, a calculator, a spreadsheet, a computer algebra system, a statistical package, or dynamic geometry software. Proficient students are sufficiently familiar with tools appropriate for their grade or course to make sound decisions about when each of these tools might be helpful, recognizing both the insight to be gained and their limitations. For example, mathematically proficient high school students analyze graphs of functions and solutions generated using a graphing calculator. They detect possible errors by strategically using estimation and other mathematical knowledge. When making mathematical models, they know that technology can enable them to visualize the results of varying assumptions, explore consequences, and compare predictions with data. Mathematically proficient students at various grade levels are able to identify relevant external mathematical resources, such as digital content located on a website, and use them to pose or solve problems. They are able to use technological tools to explore and deepen their understanding of concepts.

MA.K-12.2	Reason abstractly and quantitatively.
MA.S-ID.A.1	Represent data with plots on the real number line (dot plots, histograms, and box plots).
MA.S-ID.A.2	Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation) of two or more different data sets.

MA.S-ID.A.3	Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers).
MA.S-ID.A.4	Use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages. Recognize that there are data sets for which such a procedure is not appropriate. Use calculators, spreadsheets, and tables to estimate areas under the normal curve.
MA.K-12.4	Model with mathematics.
MA.S-ID.B.5	Summarize categorical data for two categories in two-way frequency tables. Interpret relative frequencies in the context of the data (including joint, marginal, and conditional relative frequencies). Recognize possible associations and trends in the data.
MA.S-ID.B.6	Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.
MA.K-12.5	Use appropriate tools strategically.
MA.S-ID.B.6a	Fit a function to the data (including with the use of technology); use functions fitted to data to solve problems in the context of the data.
MA.S-ID.B.6b	Informally assess the fit of a function by plotting and analyzing residuals, including with the use of technology.
MA.S-ID.B.6c	Fit a linear function for a scatter plot that suggests a linear association.
MA.S-ID.C.7	Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.
MA.S-ID.C.8	Compute (using technology) and interpret the correlation coefficient of a linear fit.
MA.S-ID.C.9	Distinguish between correlation and causation.