

2.2.1: Frequency Polygons and Time Series Graphs

Frequency Polygons

Frequency polygons are analogous to line graphs, and just as line graphs make continuous data visually easy to interpret, so too do frequency polygons. To construct a frequency polygon, first examine the data and decide on the number of intervals, or class intervals, to use on the x -axis and y -axis. After choosing the appropriate ranges, begin plotting the data points. After all the points are plotted, draw line segments to connect them.

You can create a Frequency Polygon from a grouped frequency distribution by using the class midpoints for the x -axis values. Frequency polygons are "anchored" on the x -axis on both ends, as you can see below.

Example 2.2.1.4

A frequency polygon was constructed from the frequency table below.

Frequency Distribution for Calculus Final Test Scores

Lower Bound	Upper Bound	Frequency	Cumulative Frequency
49.5	59.5	5	5
59.5	69.5	10	15
69.5	79.5	30	45
79.5	89.5	40	85
89.5	99.5	15	100

 A frequency polygon was constructed from the frequency table below.

Figure 2.2.1.4.

The first label on the x -axis is 44.5. This represents an interval extending from 39.5 to 49.5. Since the lowest test score is 54.5, this interval is used only to allow the graph to touch the x -axis. The point labeled 54.5 represents the next interval, or the first "real" interval from the table, and contains five scores. This reasoning is followed for each of the remaining intervals with the point 104.5 representing the interval from 99.5 to 109.5. Again, this interval contains no data and is only used so that the graph will touch the x -axis. Looking at the graph, we say that this distribution is skewed because one side of the graph does not mirror the other side.

Exercise 2.2.1.4

Construct a frequency polygon of U.S. Presidents' ages at inauguration shown in [Table](#).

Age at Inauguration	Frequency
41.5–46.5	4
46.5–51.5	11
51.5–56.5	14
56.5–61.5	9
61.5–66.5	4
66.5–71.5	2

Answer

The first label on the x -axis is 39. This represents an interval extending from 36.5 to 41.5. Since there are no ages less than 41.5, this interval is used only to allow the graph to touch the x -axis. The point labeled 44 represents the next interval, or the first “real” interval from the table, and contains four scores. This reasoning is followed for each of the remaining intervals with the point 74 representing the interval from 71.5 to 76.5. Again, this interval contains no data and is only used so that the graph will touch the x -axis. Looking at the graph, we say that this distribution is skewed because one side of the graph does not mirror the other side.

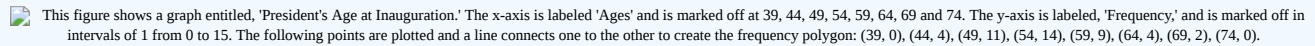
 This figure shows a graph entitled, 'President's Age at Inauguration.' The x -axis is labeled 'Ages' and is marked off at 39, 44, 49, 54, 59, 64, 69 and 74. The y -axis is labeled, 'Frequency,' and is marked off in intervals of 1 from 0 to 15. The following points are plotted and a line connects one to the other to create the frequency polygon: (39, 0), (44, 4), (49, 11), (54, 14), (59, 9), (64, 4), (69, 2), (74, 0).

Figure 2.2.1.5.

Frequency polygons are useful for comparing distributions. This is achieved by overlaying the frequency polygons drawn for different data sets.

Example 2.2.1.5

We will construct an overlay frequency polygon comparing the scores from [Example](#) with the students' final numeric grade.

Frequency Distribution for Calculus Final Test Scores

Lower Bound	Upper Bound	Frequency	Cumulative Frequency
49.5	59.5	5	5
59.5	69.5	10	15
69.5	79.5	30	45
79.5	89.5	40	85
89.5	99.5	15	100

Frequency Distribution for Calculus Final Grades

Lower Bound	Upper Bound	Frequency	Cumulative Frequency
49.5	59.5	10	10
59.5	69.5	10	20
69.5	79.5	30	50
79.5	89.5	45	95
89.5	99.5	5	100

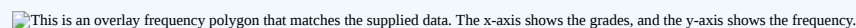
 This is an overlay frequency polygon that matches the supplied data. The x -axis shows the grades, and the y -axis shows the frequency.

Figure 2.2.1.6.

Suppose that we want to study the temperature range of a region for an entire month. Every day at noon we note the temperature and write this down in a log. A variety of statistical studies could be done with this data. We could find the mean or the median temperature for the month. We could construct a histogram displaying the number of days that temperatures reach a certain range of values. However, all of these methods ignore a portion of the data that we have collected.

One feature of the data that we may want to consider is that of time. Since each date is paired with the temperature reading for the day, we don't have to think of the data as being random. We can instead use the times given to impose a chronological order on the data. A graph that recognizes this ordering and displays the changing temperature as the month progresses is called a time series graph.

Constructing a Time Series Graph

To construct a time series graph, we must look at both pieces of our **paired data set**. We start with a standard Cartesian coordinate system. The horizontal axis is used to plot the date or time increments, and the vertical axis is used to plot the values of the variable that we are measuring. By doing this, we make each point on the graph correspond to a date and a measured quantity. The points on the graph are typically connected by straight lines in the order in which they occur.

Example 2.2.1.6

The following data shows the Annual Consumer Price Index, each month, for ten years. Construct a time series graph for the Annual Consumer Price Index data only.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul
2003	181.7	183.1	184.2	183.8	183.5	183.7	183.9
2004	185.2	186.2	187.4	188.0	189.1	189.7	189.4
2005	190.7	191.8	193.3	194.6	194.4	194.5	195.4
2006	198.3	198.7	199.8	201.5	202.5	202.9	203.5
2007	202.416	203.499	205.352	206.686	207.949	208.352	208.299
2008	211.080	211.693	213.528	214.823	216.632	218.815	219.964
2009	211.143	212.193	212.709	213.240	213.856	215.693	215.351
2010	216.687	216.741	217.631	218.009	218.178	217.965	218.011
2011	220.223	221.309	223.467	224.906	225.964	225.722	225.922
2012	226.665	227.663	229.392	230.085	229.815	229.478	229.104

Year	Aug	Sep	Oct	Nov	Dec	Annual
2003	184.6	185.2	185.0	184.5	184.3	184.0
2004	189.5	189.9	190.9	191.0	190.3	188.9
2005	196.4	198.8	199.2	197.6	196.8	195.3
2006	203.9	202.9	201.8	201.5	201.8	201.6
2007	207.917	208.490	208.936	210.177	210.036	207.342
2008	219.086	218.783	216.573	212.425	210.228	215.303
2009	215.834	215.969	216.177	216.330	215.949	214.537
2010	218.312	218.439	218.711	218.803	219.179	218.056
2011	226.545	226.889	226.421	226.230	225.672	224.939
2012	230.379	231.407	231.317	230.221	229.601	229.594

Answer


 This is a times series graph that matches the supplied data. The x-axis shows years from 2003 to 2012, and the y-axis shows the annual CPI.

Figure 2.2.1.7.

Exercise 2.2.1.5

The following table is a portion of a data set from www.worldbank.org. Use the table to construct a time series graph for CO₂ emissions for the United States.

CO ₂ Emissions			
	Ukraine	United Kingdom	United States
2003	352,259	540,640	5,681,664
2004	343,121	540,409	5,790,761
2005	339,029	541,990	5,826,394
2006	327,797	542,045	5,737,615
2007	328,357	528,631	5,828,697
2008	323,657	522,247	5,656,839
2009	272,176	474,579	5,299,563

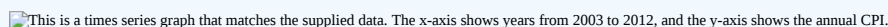
 This is a times series graph that matches the supplied data. The x-axis shows years from 2003 to 2012, and the y-axis shows the annual CO₂ emissions.

Figure 2.2.1.8.

Uses of a Time Series Graph

Time series graphs are important tools in various applications of statistics. When recording values of the same variable over an extended period of time, sometimes it is difficult to discern any trend or pattern. However, once the same data points are displayed graphically, some features jump out. Time series graphs make trends easy to spot.

Review

A **histogram** is a graphic version of a frequency distribution. The graph consists of bars of equal width drawn adjacent to each other. The horizontal scale represents classes of quantitative data values and the vertical scale represents frequencies. The heights of the bars correspond to frequency values. Histograms are typically used for large, continuous, quantitative data sets. A **frequency polygon** can be used instead of a histogram when graphing large data sets with data points that repeat. The data usually goes on *x*-axis with the frequency being graphed on the *y*-axis. **Time series graphs** can be helpful when looking at large amounts of data for one variable over a period of time.

References

1. Data on annual homicides in Detroit, 1961–73, from Gunst & Mason's book 'Regression Analysis and its Application', Marcel Dekker
2. "Timeline: Guide to the U.S. Presidents: Information on every president's birthplace, political party, term of office, and more." Scholastic, 2013. Available online at www.scholastic.com/teachers/a...-us-presidents (accessed April 3, 2013).
3. "Presidents." Fact Monster. Pearson Education, 2007. Available online at <http://www.factmonster.com/ipka/A0194030.html> (accessed April 3, 2013).
4. "Food Security Statistics." Food and Agriculture Organization of the United Nations. Available online at <http://www.fao.org/economic/ess/ess-fs/en/> (accessed April 3, 2013).
5. "Consumer Price Index." United States Department of Labor: Bureau of Labor Statistics. Available online at <http://data.bls.gov/pdq/SurveyOutputServlet> (accessed April 3, 2013).
6. "CO₂ emissions (kt)." The World Bank, 2013. Available online at <http://databank.worldbank.org/data/home.aspx> (accessed April 3, 2013).
7. "Births Time Series Data." General Register Office For Scotland, 2013. Available online at www.gro-scotland.gov.uk/statistics/me-series.html (accessed April 3, 2013).
8. "Demographics: Children under the age of 5 years underweight." Indexmundi. Available online at <http://www.indexmundi.com/g/r.aspx?t=50&v=2224&aml=en> (accessed April 3, 2013).
9. Gunst, Richard, Robert Mason. *Regression Analysis and Its Application: A Data-Oriented Approach*. CRC Press: 1980.

10. “Overweight and Obesity: Adult Obesity Facts.” Centers for Disease Control and Prevention. Available online at <http://www.cdc.gov/obesity/data/adult.html> (accessed September 13, 2013).

Frequency

the number of times a value of the data occurs

Histogram

a graphical representation in $x - y$ form of the distribution of data in a data set; x represents the data and y represents the frequency, or relative frequency. The graph consists of contiguous rectangles.

Relative Frequency

the ratio of the number of times a value of the data occurs in the set of all outcomes to the number of all outcomes

Contributors and Attributions

- Barbara Illowsky and Susan Dean (De Anza College) with many other contributing authors. Content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at <http://cnx.org/contents/30189442-699...b91b9de@18.114>.

This page titled [2.2.1: Frequency Polygons and Time Series Graphs](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- [2.3: Histograms, Frequency Polygons, and Time Series Graphs](#) by [OpenStax](#) is licensed [CC BY 4.0](#). Original source: <https://openstax.org/details/books/introductory-statistics>.