Chapter 2: Frequency Distributions

A table reporting the number of observations falling into each category of the variable.

Death Penalty Statutes 1993

(data from which to create a frequency distribution)

State	Minimum Age	State	Minimum Age
Arkansas	14	Texas	17
Virginia	15	California	18
Alabama	16	Colorado	18
Delaware	16	Connecticut	18
Indiana	16	Illinois	18
Kentucky	16	Louisiana	18
Mississippi	16	Maryland	18
Missouri	16	Nebraska	18
Nevada	16	New Jersey	18
Oklahoma	16	New Mexico	18
Wyoming	16	Ohio	18
Georgia	17	Oregon	18
New Hampshire	17	Tennessee	18
North Carolina	17		

Source: Kathleen Maguire and Ann L. Pastore, eds., *Sourcebook of Criminal Justice Statistics. 1994.* U.S. Department of Justice, Bureau of Justice Statistics. Washington, D.C.: U.S. Government Printing Office, 1995, pp. 115-116.

Creating a Frequency Distribution

Minimum Age	Tally	Frequency
14	|	1
15	|	1
16	|||||||||	9
17	||||	4
18	||||||||||||	12
Total N		27

Proportions and Percentages

- **Proportion (P):** a relative frequency obtained by dividing the frequency in each category by the total number of cases.
- **Percentage (%):** a relative frequency obtained by dividing the frequency in each category by the total number of cases and multiplying by 100.
- **N:** total number of cases
- Proportions and percentages are *relative frequencies*

$$P = \frac{f}{N}$$

$$(\%) = P(100)$$

Proportions and Percentages

Minimum Age	Frequency	Proportion	Percentage
14	1	1/27=.037	3.7
15	1	.037	3.7
16	9	.333	33.3
17	4	.148	14.8
18	12	.444	44.4
Total N	**27**	**1.0**	**100.0**

Cumulative Frequency Distribution

Minimum Age	Freq. (f)	Percentage	Cumulative Frequency
14	1	3.7	01
15	1	3.7	02
16	9	33.3	11
17	4	14.8	15
18	12	44.4	27
Total (N)	**27**	**99.9***	

* Doesn't total to 100% due to rounding

Cumulative Percentage Distribution

Minimum Age	Frequency	Percentage	Cumulative Percentage
14	1	3.7	3.7
15	1	3.7	7.4
16	9	33.3	40.7
17	4	14.8	55.5
18	12	44.4	99.9*
Total N	27	99.9*	

* Doesn't total to 100% due to rounding

Rates

A number obtained by dividing the number of *actual* occurrences in a given time period by the number of *possible* occurrences.

$$\text{Marriage rate, 1990} = \frac{\text{Number of marriages in 1990}}{\text{Total population in 1990}}$$

$$\text{Marriage rate, 1990} = \frac{\text{2,448,000 marriages}}{\text{250,000,000 Americans}}$$

Marriage rate, 1990 = .0098
(9.8 marriages for every 1000 people)

Reading Statistical Tables

Basic principles for understanding what the researcher is trying to tell you (that is, questions you should ask yourself when reading a table):

- What is the source of this table?
- How many variables are presented? What are their names?
- What is represented by the numbers presented in the first column? In the second column?

Example of Table Format for Research Paper

Table 1: The Effect of Sex on Attitudes Toward the Death Penalty

In Favor of the Death Penalty
(actual number of respondents reported)

		Yes	No	Total
Gender	Male	36	19	55
	Female	33	18	51
	Total	69	37	106

(Source: non-random sample obtained by students in a college statistics class)

SPSS Output Looks Something Like This:

V13 * V6 Crosstabulation

			V6 Yes	V6 No	Total
V13	Male	Count	36	19	55
		% within V6	52.2%	51.4%	51.9%
		% within V13	65.5%	34.5%	100.0%
	Female	Count	33	18	51
		% within V6	47.8%	48.6%	48.1%
		% within V13	64.7%	35.3%	100.0%
Total		Count	69	37	106
		% within V6	100.0%	100.0%	100.0%
		% within V13	65.1%	34.9%	100.0%

Example of Table Format for Research Paper

Table 1: The Effect of Sex on Attitudes Toward the Death Penalty

Column %		Percent In Favor of the Death Penalty*		
Row%		Yes	No	Total
Gender	Male	52	51	
		65	35	100
		(36)	(19)	(55)
	Female	48	49	
		65	35	100
		(33)	(18)	(51)
	Total	100	100	100
		(69)	(37)	(106)

*Numbers in parentheses are actual numbers of respondents

Chapter 4:

Central Tendency

(mean, median, mode)

The Mode: An Example

- **Which of the three candidates represents the "mode" for these candidates**
- **Variable=Candidates**
  Candidate A – 11,769 votes
  Candidate B – 39,443 votes
  Candidate C – 78,331 votes

Level of measurement? =

The Mode?=

The Mode: An Example

- **Which of the three candidates represents the "mode" for these candidates**
- **Variable=Candidates**
  Candidate A – 11,769 votes
  Candidate B – 39,443 votes
  Candidate C – 78,331 votes

Level of measurement = nominal (why?)

The Mode= Candidate C (why?)

Finding Median Among Individual Cases

# of hate crimes by state

Cases	
NC	= 39
Penn	= 141
TX	= 287
Ohio	= 255
Fla	= 240

States ordered low to high

NC	=39
Penn	=141
Fla	=240
Ohio	=255
TX	=287

\# of cases = 5

Steps to Determine:

1. Order the cases from highest to lowest or vice versa
2. Add 1 to the total number of cases (5+1=6)
3. divide resulting number 2 (6/2 = 3)
4. Count down that many cases to identify the middle or median (Fla)

Finding Median for a frequency distribution

I am very satisfied with my job

Values	Freq	Cummulative Frequency
Agree Strongly	5	5
Agree	10	15
Undecided	3	18
Disagree	7	25
Dis. Strongly	3	28

Total Cases: 28

Steps to Determine:

1. divide total N by 2 to locate middle case: **28/2 = 14**
2. determine cumulative frequencies
3. locate the category that holds the middle case: "**agree**" contains the 14^{th} case

Formula for the Mean

$$\overline{Y} = \frac{\sum f\,Y}{N}$$

"Y bar" ($\overline{Y}$) equals the average or the sum of all the scores, Y, divided by the number of scores, N.

Calculating the mean with frequency distributions (grouped scores):

Satisfaction with Health

	Freq	Category x Freq
1 - Very High	5	5
2 – High	7	14
3 – Moderate	6	18
4 - Low	7	28
5 - Very Low	3	15

Totals: N=28; Cat.x Freq=80

$$\overline{Y} = \frac{\sum f Y}{N}$$

Steps to Determine:

1. multiply each category by its frequency (category x frequency)
2. sum all the "category x freq" scores to determine total (80)
3. divide total (80) by total number of cases (total N or 28) to get average score (2.86)

Considerations for Choosing a Measure of Central Tendency

- For a nominal variable, the mode is the only measure that can be used.
- For ordinal variables, the mode and the median may be used. The median provides more information (taking into account the ranking of categories).
- For interval-ratio variables, the mode, median, and mean may all be calculated. The mean provides the most information about the distribution, but the median is preferred if the distribution is skewed.

Chapter 5: The Importance of Measuring Variability

- Measures of Central Tendency - Numbers that describe what is typical or average (central) in a distribution (e.g., mean, mode, median).
- Measures of Variability - Numbers that describe diversity or variability in the distribution (e.g., range, interquartile range, variance, standard deviation).

The Range

- **Range – A measure of variation in <u>interval-ratio variables</u>.**
- **It is the difference between the highest (maximum) and the lowest (minimum) scores in the distribution.**

Range = highest score - lowest score

What is the range for these diversity scores? (higher number means more diversity)?

Steps to determine: subtract the lowest score ____from the highest _____to obtain the range of IQV scores_____.

State	IQV	State	IQV	State	IQV
California	0.80	Alabama	0.51	Indiana	0.27
New Mexico	0.76	North Carolina	0.51	Utah	0.26
Texas	0.74	Delaware	0.49	Nebraska	0.24
New York	0.66	Colorado	0.45	South Dakota	0.24
Hawaii	0.64	Oklahoma	0.44	Wisconsin	0.24
Maryland	0.62	Connecticut	0.42	Idaho	0.23
New Jersey	0.61	Arkansas	0.40	Wyoming	0.22
Louisiana	0.61	Michigan	0.40	Kentucky	0.20
Arizona	0.61	Tennessee	0.39	Minnesota	0.20
Florida	0.61	Washington	0.37	Montana	0.20
Mississippi	0.61	Massachusetts	0.34	North Dakota	0.17
Georgia	0.59	Missouri	0.31	Iowa	0.13
Nevada	0.57	Ohio	0.31	West Virginia	0.11
Illinois	0.57	Pennsylvania	0.31	New Hampshire	0.08
South Carolina	0.56	Kansas	0.30	Maine	0.06
Alaska	0.56	Rhode Island	0.30	Vermont	0.06
Virginia	0.53	Oregon	0.28		

What is the range for these diversity scores? (higher number means more diversity)?

Steps to determine: subtract the lowest score _.06___from the highest _____to obtain the range of IQV scores_____.

State	IQV	State	IQV	State	IQV
California	0.80	Alabama	0.51	Indiana	0.27
New Mexico	0.76	North Carolina	0.51	Utah	0.26
Texas	0.74	Delaware	0.49	Nebraska	0.24
New York	0.66	Colorado	0.45	South Dakota	0.24
Hawaii	0.64	Oklahoma	0.44	Wisconsin	0.24
Maryland	0.62	Connecticut	0.42	Idaho	0.23
New Jersey	0.61	Arkansas	0.40	Wyoming	0.22
Louisiana	0.61	Michigan	0.40	Kentucky	0.20
Arizona	0.61	Tennessee	0.39	Minnesota	0.20
Florida	0.61	Washington	0.37	Montana	0.20
Mississippi	0.61	Massachusetts	0.34	North Dakota	0.17
Georgia	0.59	Missouri	0.31	Iowa	0.13
Nevada	0.57	Ohio	0.31	West Virginia	0.11
Illinois	0.57	Pennsylvania	0.31	New Hampshire	0.08
South Carolina	0.56	Kansas	0.30	Maine	0.06
Alaska	0.56	Rhode Island	0.30	Vermont	0.06
Virginia	0.53	Oregon	0.28		

What is the range for these diversity scores?

(higher number means more diversity)?

Steps to determine: subtract the lowest score _.06___from the highest ___.80___to obtain the range of IQV scores_____.

State	IQV	State	IQV	State	IQV
California	0.80	Alabama	0.51	Indiana	0.27
New Mexico	0.76	North Carolina	0.51	Utah	0.26
Texas	0.74	Delaware	0.49	Nebraska	0.24
New York	0.66	Colorado	0.45	South Dakota	0.24
Hawaii	0.64	Oklahoma	0.44	Wisconsin	0.24
Maryland	0.62	Connecticut	0.42	Idaho	0.23
New Jersey	0.61	Arkansas	0.40	Wyoming	0.22
Louisiana	0.61	Michigan	0.40	Kentucky	0.20
Arizona	0.61	Tennessee	0.39	Minnesota	0.20
Florida	0.61	Washington	0.37	Montana	0.20
Mississippi	0.61	Massachusetts	0.34	North Dakota	0.17
Georgia	0.59	Missouri	0.31	Iowa	0.13
Nevada	0.57	Ohio	0.31	West Virginia	0.11
Illinois	0.57	Pennsylvania	0.31	New Hampshire	0.08
South Carolina	0.56	Kansas	0.30	Maine	0.06
Alaska	0.56	Rhode Island	0.30	Vermont	0.06
Virginia	0.53	Oregon	0.28		

What is the range for these diversity scores?

(higher number means more diversity)?

Steps to determine: subtract the lowest score _.06___from the highest __.80___to obtain the range of IQV scores__.74___.

State	IQV	State	IQV	State	IQV
California	0.80	Alabama	0.51	Indiana	0.27
New Mexico	0.76	North Carolina	0.51	Utah	0.26
Texas	0.74	Delaware	0.49	Nebraska	0.24
New York	0.66	Colorado	0.45	South Dakota	0.24
Hawaii	0.64	Oklahoma	0.44	Wisconsin	0.24
Maryland	0.62	Connecticut	0.42	Idaho	0.23
New Jersey	0.61	Arkansas	0.40	Wyoming	0.22
Louisiana	0.61	Michigan	0.40	Kentucky	0.20
Arizona	0.61	Tennessee	0.39	Minnesota	0.20
Florida	0.61	Washington	0.37	Montana	0.20
Mississippi	0.61	Massachusetts	0.34	North Dakota	0.17
Georgia	0.59	Missouri	0.31	Iowa	0.13
Nevada	0.57	Ohio	0.31	West Virginia	0.11
Illinois	0.57	Pennsylvania	0.31	New Hampshire	0.08
South Carolina	0.56	Kansas	0.30	Maine	0.06
Alaska	0.56	Rhode Island	0.30	Vermont	0.06
Virginia	0.53	Oregon	0.28		

Inter-quartile Range

- **Inter-quartile range (IQR) – The width of the middle 50 percent of the distribution.**
- **The IQR helps us to get a better picture of the variation in the data than the range.**
- **The shortcoming of the range is that an "outlying" case at the top or bottom can increase the range substantially.**

Inter-quartile Range

- **Inter-quartile range (IQR) – The width of the middle 50 percent of the distribution.**
- **It is defined as the difference between the lower and upper quartiles (Q1 and Q3.)**
- **IQR = q3 – q1**

What is the IQR for these Diversity Scores?

State	IQV	State	IQV	State	IQV
California	0.80	Alabama	0.51	Indiana	0.27
New Mexico	0.76	North Carolina	0.51	Utah	0.26
Texas	0.74	Delaware	0.49	Nebraska	0.24
New York	0.66	Colorado	0.45	South Dakota	0.24
Hawaii	0.64	Oklahoma	0.44	Wisconsin	0.24
Maryland	0.62	Connecticut	0.42	Idaho	0.23
New Jersey	0.61	Arkansas	0.40	Wyoming	0.22
Louisiana	0.61	Michigan	0.40	Kentucky	0.20
Arizona	0.61	Tennessee	0.39	Minnesota	0.20
Florida	0.61	Washington	0.37	Montana	0.20
Mississippi	0.61	Massachusetts	0.34	North Dakota	0.17
Georgia	0.59	Missouri	0.31	Iowa	0.13
Nevada	0.57	Ohio	0.31	West Virginia	0.11
Illinois	0.57	Pennsylvania	0.31	New Hampshire	0.08
South Carolina	0.56	Kansas	0.30	Maine	0.06
Alaska	0.56	Rhode Island	0.30	Vermont	0.06
Virginia	0.53	Oregon	0.28		

(Steps are provided on the next slides)

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

1. Order the categories from highest to lowest (or vice versa)
2. To obtain Q1, begin by dividing N (total number of categories or states) by 4 (or alternatively multiply N by .25). This equals______?
3. We now know that Q1 falls between the 12th and 13th category or, in this case, states.
4. To find the exact number for Q1, determine the midpoint between the 12th and 13th states or between .59 and .57)
5. Q1 = ____

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

1. Order the categories from highest to lowest (or vice versa)
2. To obtain Q1, begin by dividing N (total number of categories or states) by 4 (or alternatively multiply N by .25). This equals___12.5___?
3. We now know that Q1 falls between the 12th and 13th category or, in this case, states.
4. To find the exact number for Q1, determine the midpoint between the 12th and 13th states or between .59 and .57)
5. Q1 = ____

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

1. Order the categories from highest to lowest (or vice versa)
2. To obtain Q1, begin by dividing N (total number of categories or states) by 4 (or alternatively multiply N by .25). This equals___12.5___?
3. We now know that Q1 falls between the 12th and 13th category or, in this case, states.
4. To find the exact number for Q1, determine the midpoint between the 12th and 13th states or between .59 and .57)
5. Q1 = __.58__

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

6. To obtain Q3, begin by multiplying 12.5 by 3 (or alternatively multiply 12.5 by .75). This will give us____.
7. Based on this number, Q3 falls between the 37th and 38th states.
8. Determine the midpoint between these two states. This equals_____. This tells us that 50% of the cases fall between .58 and .24.
9. To obtain the IQR subtract Q3 from Q1 which equals_____or the middle of the middle 50% of the cases.

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

6. To obtain Q3, begin by multiplying 12.5 by 3 (or alternatively multiply 12.5 by .75. This will give us__37.5__.
7. Based on this number, Q3 falls between the 37th and 38th states.
8. Determine the midpoint between these two states. This equals_____. This tells us that 50% of the cases fall between .58 and .24.
9. To obtain the IQR subtract Q3 from Q1 which equals_____or the middle of the middle 50% of the cases.

What is the IQR for the Diversity Scores?

Steps to determine the IQR (Q3 – Q1):

6. To obtain Q3, begin by multiplying 12.5 by 3 (or alternatively multiply N (50) by .75). This will give us__37.5__.
7. Based on this number, Q3 falls between the 37th and 38th states.
8. Determine the midpoint between these two states. This equals___.24__. This tells us that 50% of the cases fall between .58 and .24.
9. To obtain the IQR subtract Q3 from Q1 which equals_.34____ or the middle of the middle 50% of the cases.

Measures of Variability:
Shortcomings of the Range and IQR

- **The range is based on only two categories (the highest and lowest)**
- **Likewise, only two categories are used to calculate the inter-quartile range.**
- **Neither allows us to know how much variation there is among all the categories.**

Measures of Variability: the Variance

- The variance allows us to account for the total amount of variation that includes the variation of all the categories.
- The amount of variation in each category is considered when calculating the variance.
- The variance is an important statistic that is used in most other sophisticated statistics. Therefore, it is important for you to give it particular attention.

Determining Variation in the "Percentage Increase" in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage
Pacific	15.7
West North Central	16.2
New England	17.6
East North Central	23.2
West South Central	24.3
Middle Atlantic	28.5
East South Central	38.0
Mountain	47.9
South Atlantic	71.7

How might we take into account the variation that exists for each category?

Determining Variation in the "Percentage Increase" in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage
Pacific	15.7
West North Central	16.2
New England	17.6
East North Central	23.2
West South Central	24.3
Middle Atlantic	28.5
East South Central	38.0
Mountain	47.9
South Atlantic	71.7

How might we take into account the variation that exists for each category?

Hint: the "mean" is the mid-point so you could measure how far each category is from the mid-point and then add up all these distances. The larger this sum the more variation.

Variation in the "Percentage Increase" in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage
Pacific	15.7
West North Central	16.2
New England	17.6
East North Central	23.2
West South Central	24.3
Middle Atlantic	28.5
East South Central	38.0
Mountain	47.9
South Atlantic	71.7

Average "% increase"

$\sum fY = 283.1$ mean = $\bar{Y} = \frac{\sum fY}{N}$ = 31.45

Variation in the "Percentage Increase" in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage
Pacific	15.7
West North Central	16.2
New England	17.6
East North Central	23.2
West South Central	24.3
Middle Atlantic	28.5
East South Central	38.0
Mountain	47.9
South Atlantic	71.7

Average "% increase"

$\sum fY = 283.1$ mean = $\bar{Y} = \frac{\sum fY}{N}$ = 31.45

Percentage Change in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage	$Y - \bar{Y}$
Pacific	15.7	15.7 - 31.5 = -15.8
West North Central	16.2	16.2 - 31.5 = -15.3
New England	17.6	17.6 - 31.5 = -13.9
East North Central	23.2	23.2 - 31.5 = - 8.3
West South Central	24.3	24.3 - 31.5 = - 7.2
Middle Atlantic	28.5	28.5 - 31.5 = - 3.0
East South Central	38.0	38.0 - 31.5 = 6.5
Mountain	47.9	47.9 - 31.5 = 16.4
South Atlantic	71.7	71.7 - 31.5 = 40.2
(mean = 31.5)	$\sum Y$ = 283.1	$\sum (Y - \bar{Y}) = 0$

How might we take into account the variation that exists for each category?

Problem: when you add up the distances you end up with zero rather than the total variation from all the categories.

Percentage Change in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage	$Y - \bar{Y}$
Pacific	15.7	15.7 - 31.5 = -15.8
West North Central	16.2	16.2 - 31.5 = -15.3
New England	17.6	17.6 - 31.5 = -13.9
East North Central	23.2	23.2 - 31.5 = - 8.3
West South Central	24.3	24.3 - 31.5 = - 7.2
Middle Atlantic	28.5	28.5 - 31.5 = - 3.0
East South Central	38.0	38.0 - 31.5 = 6.5
Mountain	47.9	47.9 - 31.5 = 16.4
South Atlantic	71.7	71.7 - 31.5 = 40.2
(mean = 31.5)	$\sum Y =$ 283.1	$\sum (Y - \bar{Y}) = 0$

- One solution would be to take the absolute value for each number (ignore the minus signs). Unfortunately, absolute values are very difficult to work with mathematically.
- Fortunately, there is another alternative.

Percentage Change in the Nursing Home Population, 1980-1990

Nine Regions of U.S.	Percentage	$Y - \bar{Y}$	$(Y - \bar{Y})^2$ (squared deviations)
Pacific	15.7	15.7 - 31.5 = -15.8	249.64
West North Central	16.2	16.2 - 31.5 = -15.3	234.09
New England	17.6	17.6 - 31.5 = -13.9	193.21
East North Central	23.2	23.2 - 31.5 = - 8.3	68.89
West South Central	24.3	24.3 - 31.5 = - 7.2	51.84
Middle Atlantic	28.5	28.5 - 31.5 = - 3.0	9.00
East South Central	38.0	38.0 - 31.5 = 6.5	42.25
Mountain	47.9	47.9 - 31.5 = 16.4	268.96
South Atlantic	71.7	71.7 - 31.5 = 40.2	1616.04
(mean = 31.5)	$\sum Y =$ 283.1		$\sum (Y - \bar{Y})^2 = 2733.92$

- The best solution is to square the differences before adding them up (when two negative numbers are multiplied the resulting product is a positive number).

Measures of Variability: the Variance

The Variance is the average of the squared deviations from the mean.

$$s_Y^2 = \frac{\sum (Y - \bar{Y})^2}{N - 1}$$

In our example we would take the sum of the squared deviations (2733.92) and divide this number by the total number of cases minus one (9 – 1 = 8). This would give us _341.74___ or the variance for the Percent Increase in the Nursing Home population by region.

Measures of Variability: The Variance

To Sum Up:

The Variance is the average of the squared deviations from the mean.

The Variance is a measure of variability for interval-ratio variables.

$$s_Y^2 = \frac{\sum (Y - \bar{Y})^2}{N - 1}$$

Measures of Variability: Standard Deviation vs Variance

- One problem with the variance is that the final number obtained is in a squared form

  (that is, we squared all the deviations from the mean and so the final number is still "inflated" in this way making it difficult to interpret)

- One solution is to take the square root of the variance so that the number is no longer in a squared form (or "inflated") and it is back to its original form. The square root of the variance is called the Standard Deviation.

Measures of Variability: Standard Deviation

- To obtain the square root of the variance simply enter the number (variance) into your calculator and then push the square root button.

- If the variance is 341.74 the standard deviation would be ____18.49______. This tells us that the percent of change in the nursing home population for the nine regions is widely dispersed around the mean (mean = 31.45).

- Thus, the standard deviation is a measure of the average amount of variation (or deviation) around the mean.

Measures of Variability: Standard Deviation

In Sum

- Standard Deviation – A measure of variation for interval-ratio variables; it is equal to the square root of the variance.

$$s = \sqrt{s_Y^2} = \sqrt{\frac{\sum (Y - \overline{Y})^2}{N - 1}}$$

Measures of Variability: Standard Deviation

(a look at what's to come in future chapters)

We will see later that when the data are "normally distributed" around the mean (produce a normal curve), 34% of the scores will be one standard deviation above the mean and 34% will be one standard deviation below the mean.

Scores are often "normally distributed" around the mean when a random sample has been used to obtain the scores or there are a large number of cases.

Considerations for Choosing a Measure of Variability

- For ordinal variables, you can calculate the IQR (inter-quartile range.)
- For interval-ratio variables, you can use IQR, or the variance/standard deviation. The standard deviation (also variance) provides the most information, since it uses all of the values in the distribution in its calculation.

$$\sqrt{\frac{\sum (Y - \overline{Y})^2}{N - 1}}$$

Σ

$\overline{Y} = \frac{\sum fY}{N}$ $\overline{Y} = \frac{\sum fY}{N}$ $\overline{Y} = \frac{\sum fY}{N}$