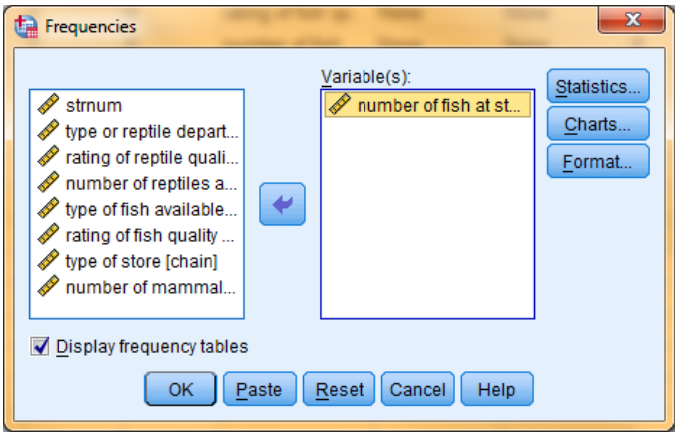
SPSS: Frequency Distributions & Univariate Statistics

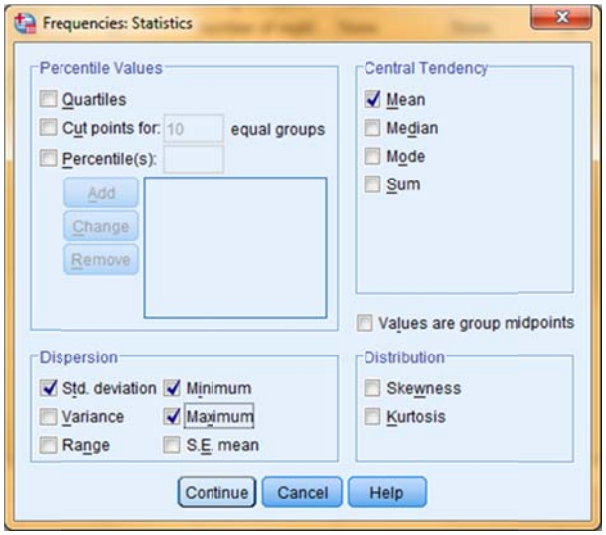
Application: To obtain a frequency table, a graph, and univariate statistics

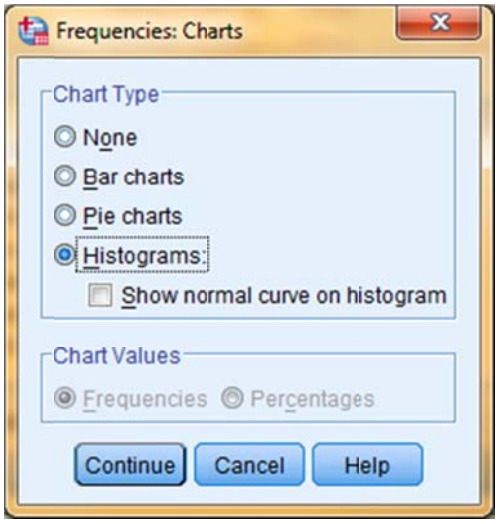
Analyze ➔ Descriptive Statistics ➔ Frequencies

- Highlight each desired variable and click the arrow button
- Click "Statistics" — indicate the univariate statistics you want computed (get mean, std dev, S.E. mean)
- Click "Charts" — click "Histogram" for quantitative variables and "Bar Chart" for qualitative variables



There are several other SPSS procedures that can be used to get univariate stats and various depictions of the data. "Frequencies" is an easy starting place, because it gets some of each – graphs, frequency tables & univariate stats.





SPSS Syntax

```
FREQUENCIES VARIABLES= reptnum reptdept chain      ← list variables
  /STATISTICS= MEAN STDDEV SEMEAN MINIMUM MAXIMUM  ← list statistics (more below)
  /NTILES=2                                        ← get cutpoints for #groups (e.g., 2)
  /FORMAT=NOTABLE                                  ← include if you don't want f tables
  /HISTOGRAM.                                      ← include if you want a histogram of the data
```

Available univariate statistics

Center ➔ MEAN MEDIAN MODE SUM
Variability ➔ STDDEV VARIANCE SEMEAN RANGE MINIMUM MAXIMUM
Shape ➔ SKEWNESS SESKEW KURTOSIS SEKURT

Statistics

number of reptiles at store

N	Valid	12
	Missing	0
Mean		9.25
Std. Error of Mean		1.232
Std. Deviation		4.267
Minimum		4
Maximum		15

number of reptiles at store

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4	3	25.0	25.0	25.0
	5	1	8.3	8.3	33.3
	7	1	8.3	8.3	41.7
	10	2	16.7	16.7	58.3
	12	2	16.7	16.7	75.0
	14	2	16.7	16.7	91.7
	15	1	8.3	8.3	100.0
	Total	12	100.0	100.0	

Univariate Stats for Quantitative Data

Mean and standard deviation are meaningful.

Combining the mean and std tells us...

- about 68% of stores in this population display between 4.983 & 13.517 reptiles
- about 96% of the stores display between .716 & 17.784 reptiles

Frequency Listing

Tells the number of occurrences of each value for this variable (e.g., there were 3 stores with 4 reptiles and 2 stores with 10

If there are missing values, then you should use the "Valid Percent" column to find the % of cases in each value/category.

Chart

Gives a quick picture of the data.

Can be used to look for noncontinuities (gaps) and skewness (asymmetry) in the distribution.

Thinking about the Standard Deviation and the Standard Error of the Mean

The **standard deviation** (Std) tells about the expected variation of individual scores in the distribution around the mean of that distribution. The **standard error of the mean** (SEM) tells about the expected variation in estimates of the mean of the population across multiple samplings of a specific size.

Huh? We're used to drawing a single sample and treating the mean of that sample as our estimate of the population mean. We know that estimate isn't exactly accurate, but by how much is likely to be off? If we took multiple samples of the same size from the same population and computed the mean for each, we could compute the standard deviation of those means, as an index of the expected variability in sample estimates of the population mean. The standard error of the mean gives us this information, even if we have only one sample.

How does this work? Assuming the data are normally distributed, we can estimate the variability of multiple sample means around the true population mean from the standard deviation of a single sample (the more variability there is among the individuals, the more variability there is likely to be among the sample means - get it !!).

Statistics

type or reptile department

N	Valid	12
	Missing	0
Mean		1.50
Std. Error of Mean		.151
Std. Deviation		.522

type or reptile department

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	not separate	6	50.0	50.0	50.0
	separate dept	6	50.0	50.0	100.0
	Total	12	100.0	100.0	

Univariate Statistics for Binary variables

The decimal part of the mean will tell you the proportion of the scores in the distribution with the higher coded value.

For these data...
Delete the whole number part of the mean, leaving .5, which tell us that .5 (1/2, 50%) of the sample had separate reptile departments

The mean gives useful information for a binary variable because there is only one combination of condition frequencies that will yield this particular mean! Only a 50%-50% combination of these two types of departments will give a mean of 1.5). Any other combination would give a different unique mean (e.g., 25% not separate and 75% separate would give a mean = 1.75, compared to a combination of 75% not separate and 25% separate would give a mean = 1.25)

Statistics

type of store

N	Valid	12
	Missing	0
Mean		1.92
Std. Error of Mean		.260
Std. Deviation		.900

type of store

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	chain store	5	41.7	41.7	41.7
	privately owned	3	25.0	25.0	66.7
	coop	4	33.3	33.3	100.0
	Total	12	100.0	100.0	

Univariate Statistics for multi-category Qualitative variables

The mean and standard deviation are **not meaningful! Why?** Because there are multiple combinations of frequencies of the three conditions that will yield this particular mean.

For example, a different mix of these frequencies, say 5 privately owned stores, 3 chain stores and 4 coops would also give a mean of 1.917. In fact, any combination that has the same sum ((5 coded 1 = 5) + (3 coded 2 = 6) + (4 coded 3 = 12) = 23) will have the same mean (23 / 12 = 1.917).

Example of a "Table 1"

- Used to show univariate statistics for a mix of quantitative and qalitative variables.
- Be sure to arrange variables with all the quantitative variables together, and all the qualitative variables together

Table 1
Summary of Story Type, Reptile Department Type and Number of Reptiles

Variable	Univariate Statistics		
Number of reptiles	M = 9.25	S = 4.27	N = 12
Type of Reptile Department	Separate Department		6 (50%)
	Not Separate Department		6 (50%)
Type of Store	Part of National Chair		5 (41.7%)
	Privately Owned		3 (25%)
	Member of Regional Coop		4 (33.3%)