# STA10003 Supplement B – Reporting Information about Single Variables

***Classifying a variable according to its level of measurement***:

When reporting information about single variables, we need to take into account the level of measurement before making decisions about which procedure is required, and the statistics that are the most appropriate to report.

Data can be classified under the two broad headings: Categorical or Metric. Categorical data has a nominal or ordinal level of measurement, whereas metric data has an interval or ratio measure.

| **Levels of Measurement:** | |
|---|---|
| ***Categorical: Nominal & Ordinal*** | |
| ***Nominal:*** | used to **label** the group<br>eg 1 = Male 2 = Female |
| ***Ordinal:*** | used to **label** *AND* **order**<br>eg 1st, 2nd, 3rd, 4th in a race, Chapter 1, Chapter 2, Chapter 3 in Text |
| ***Metric: Interval & Ratio*** | |
| ***Interval:*** | numbers are used to **label** and **order *AND*** the **intervals** between the numbers **are equal**<br>eg Temperature in $^0$C or $^0$F |
| ***Ratio:*** | numbers are used to **label** and **order** and the **intervals** between the numbers **are equal *AND ZERO* means a complete absence** of something<br>eg number of correct answers in a test |

In an existing SPSS data file, categorical data will usually have named categories, as numbers are used to either represent different named categories only [nominal data] or to represent different named categories which also have some form of ranking [ordinal data]. By contrast, when a variable recorded in an SPSS file is metric, it occurs naturally as numbers which represent a reasonably accurate quantitative measurement, so these numbers will not have name labels assigned to them.

***Producing the analysis and writing the report***:

The report is a ***description*** of the information / findings from the sample.

**FOR CATEGORICAL VARIABLES:**

**SPSS FREQUENCIES procedure:**

Frequencies Table [default]

Percentage Bar Chart [requested]

# STA10003 Supplement B – Reporting Information about Single Variables

**REPORT SHOULD CONTAIN**:

An introductory sentence about the variable of interest, a description of the sample [size and who / what] and reference to the Figure representing graphical information about the distribution. Information relating to the 'most typical' [mode] response category – including percentages.

Any other information that will provide an 'overall' picture of the data. For example, was there one response category that was quite different from the others?

## FOR METRIC VARIABLES:

**SPSS EXPLORE procedure:**

Descriptives Table [default]

Histogram [requested] and Boxplot [default]

Percentiles Table [requested]

**REPORT SHOULD CONTAIN**:

An introductory sentence about the variable of interest, a description of the sample [size and who / what] and refer to the Figure representing graphical information of the distribution.

The shape of the distribution [symmetrical or skewed – if skewed the direction of the skew] and interpretation of the 'measure of centre' [mean / sd for symmetrical or median for skewed data]

An interpretation of the 'middle 50%' [25th to 75th percentiles] – which provides an indication of how spread out the data is

Any other information that will provide an 'overall' picture of the data – eg outliers.

# STA10003 Supplement B – Reporting Information about Single Variables

## FOR CATEGORICAL VARIABLES:

**Worked example**: [Dogs&Cats.sav] **Variable**: Pet Preference

Using the Frequencies procedure, the following output [Frequency Table and Percentage Bar Chart] was produced. The colour coding shows the linking of the information.

**Do you prefer Dogs or Cats?**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | Other Animals | 5 | 2.3 | 2.3 | 2.3 |
| | **Cats** | 96 | 43.6 | **44.4** | 46.8 |
| | **Dogs** | 115 | 52.3 | **53.2** | 100.0 |
| | Total | **216** | 98.2 | 100.0 | |
| Missing | 9 | 4 | 1.8 | | |
| Total | | 220 | 100.0 | | |

***Report***:

The distribution of pet preference for a sample of **216** pet owners is displayed in *Figure 1*. The most typical * reported preference was for **dogs** (**53.2**%), although **cats** were also popular (**44.4**%). Few pet owners reported a preference for other types of animals.

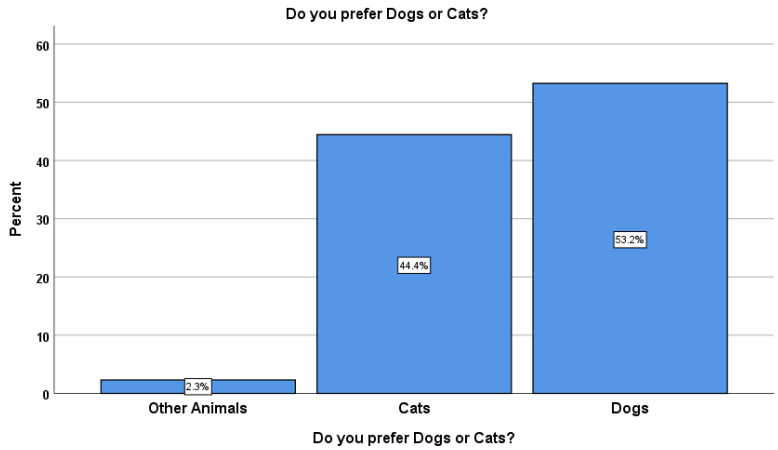



*Figure 1*: Distribution of pet preference

***Notes***:

***** In this instance, the word majority can be used rather than 'most typical', as the percentage is greater than 50%.

# STA10003 Supplement B – Reporting Information about Single Variables

## FOR METRIC VARIABLES:

**Worked example**: [Dogs&Cats.sav] **Variable**: Number of Pets

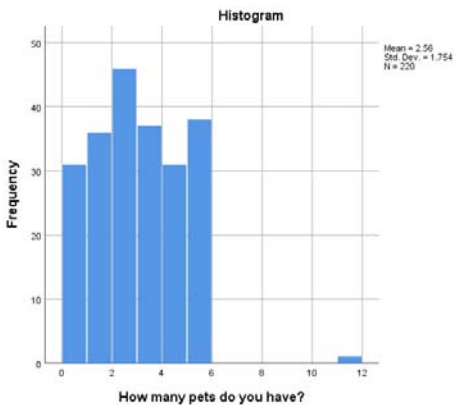
Using the Explore procedure, the following output [Descriptives Table / Histogram / Boxplot and Percentiles Table] was produced. The colour coding shows the linking of the information to the report.

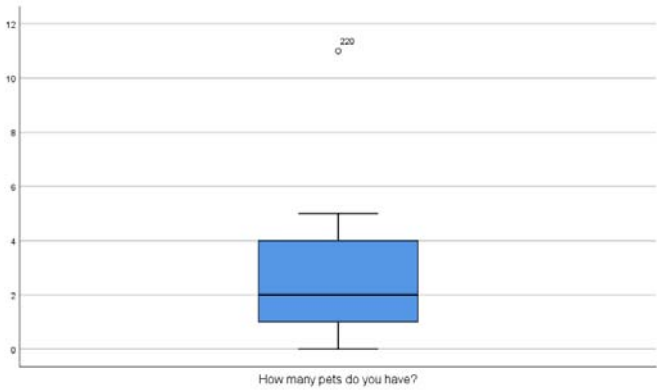
**Descriptives**

| | | | Statistic | Std. Error |
|---|---|---|---|---|
| How many pets do you have? | Mean | | 2.56 | .118 |
| | 95% Confidence Interval for Mean | Lower Bound | 2.33 | |
| | | Upper Bound | 2.80 | |
| | 5% Trimmed Mean | | 2.54 | |
| | **Median** | | **2.00** | |
| | Variance | | 3.078 | |
| | Std. Deviation | | 1.754 | |
| | Minimum | | 0 | |
| | **Maximum** | | **11** | |
| | Range | | 11 | |
| | Interquartile Range | | 3 | |
| | Skewness | | .492 | .164 |
| | Kurtosis | | .969 | .327 |

**Percentiles**

| | | Percentiles | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 5 | 10 | **25** | 50 | **75** | 90 | 95 |
| Weighted Average(Definition 1) | How many pets do you have? | .00 | .00 | **1.00** | 2.00 | **4.00** | 5.00 | 5.00 |
| Tukey's Hinges | How many pets do you have? | | | 1.00 | 2.00 | 4.00 | | |



# STA10003 Supplement B – Reporting Information about Single Variables

***Report***: [skewed distribution]

The distribution of the number of pets for a sample of **220 *** pet owners is displayed in *Figure 2*. This distribution is positively skewed, with 50% of people in the sample having **2 or more** pets [or could say … 2 or less]. Typically, the number of pets were between **1 and 4**, with half of 'the number of pets' values falling within this range. One respondent reported having an exceptionally high number of pets (11).

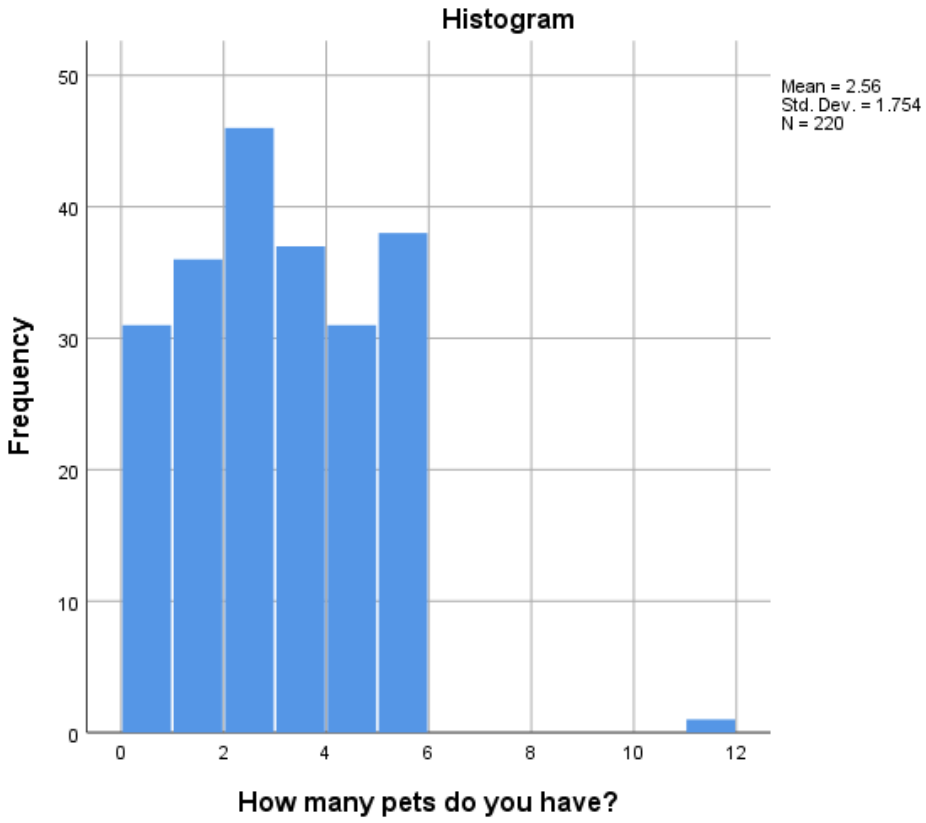



*Figure 2*: Distribution of the number of pets

***Notes***: * The size of the sample is shown in the legend of the histogram.

# STA10003 Supplement B – Reporting Information about Single Variables

**Worked example**: [Travel_Time_Work.sav] **Variable**: Travel

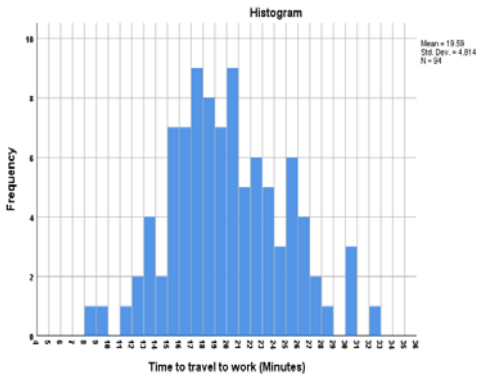
Using the Explore procedure, the following output [Descriptives Table / Histogram / Boxplot and Percentiles Table] was produced. The colour coding shows the linking of the information to the report.

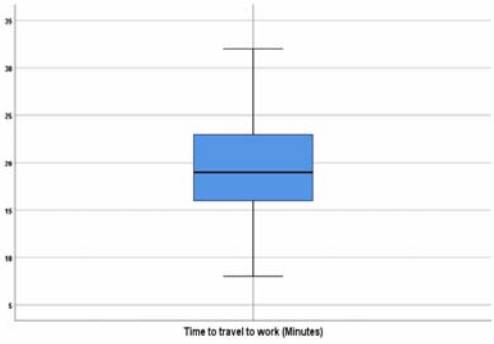
**Descriptives**

| | | | Statistic | Std. Error |
|---|---|---|---|---|
| Time to travel to work (Minutes) | **Mean** | | **19.59** | .496 |
| | 95% Confidence Interval for Mean | Lower Bound | 18.60 | |
| | | Upper Bound | 20.57 | |
| | 5% Trimmed Mean | | 19.52 | |
| | Median | | 19.00 | |
| | Variance | | 23.170 | |
| | **Std. Deviation** | | **4.814** | |
| | Minimum | | 8 | |
| | Maximum | | 32 | |
| | Range | | 24 | |
| | Interquartile Range | | 7 | |
| | Skewness | | .232 | .249 |
| | Kurtosis | | -.089 | .493 |

**Percentiles**

| | | Percentiles | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 5 | 10 | **25** | 50 | **75** | 90 | 95 |
| Weighted Average (Definition 1) | Time to travel to work (Minutes) | 12.00 | 13.50 | **16.00** | 19.00 | **23.00** | 26.00 | 28.50 |
| Tukey's Hinges | Time to travel to work (Minutes) | | | 16.00 | 19.00 | 23.00 | | |



# STA10003 Supplement B – Reporting Information about Single Variables

***Report***: [approximately symmetric distribution]

The distribution of 'travel time to work' for a random sample of **94 *** Australian adults is displayed in *Figure 3*. This distribution is approximately symmetric, with the average travel time to work being 19.59 minutes (s = 4.81 minutes). Typically, the average travel time to work was between 16.00 and 23.00 minutes, with half of the travel times to work falling within this range.

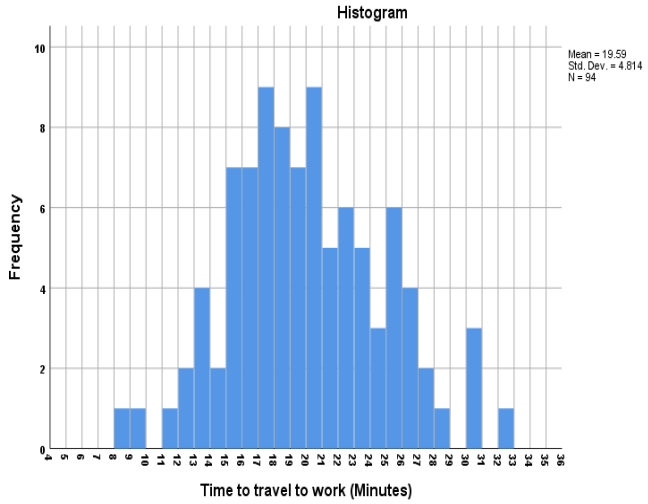



*Figure 3*: Distribution of travel time to work

***Notes***:

***** The size of the sample is shown in the legend of the histogram. There was some missing data in the data set.

The histogram was modified using the 'Binning' function.

Figures reported to two decimal places for consistency and to correspond with the output (although these figures could also have been rounded off to whole numbers throughout this report).