

# MIS770 Foundation Skills in Business Analysis

DEPARTMENT OF INFORMATION SYSTEMS AND BUSINESS ANALYTICS  
DEAKIN BUSINESS SCHOOL  
FACULTY OF BUSINESS AND LAW



## Tutorial Topic 2 Organising and Visualising Data

### Introduction

Our underlying goal here is to faithfully represent our raw Sample Data visually, such that the **visual** story we're telling is accurate and not misleading.

So, specifically we will see how to construct various types of tables, histograms, frequency distributions, frequency polygons and Ogives. Then we will further illustrate our data using various charts.

Therefore, the aims of this tutorial are to:

- describe the distribution of a single categorical variable using tables and charts
- describe the distribution of a single numerical variable using tables and graphs
- describe the relationship between two categorical variables using contingency tables
- describe the relationship between two numerical variables using scatter diagrams and time-series plots
- correctly present data in graphs

### Textbook Questions

- 2.15 The values for a data set vary from 11.6 to 97.8
- a. If these values are grouped into nine classes, indicate appropriate class boundaries
  - b. What class width did you choose?
  - c. What are the corresponding class mid-points?
- 2.17 The following data represent the electricity cost (in dollars) during the month of July for a random sample of 50 two-bedroom apartments in a New Zealand city. [Dataset: ELECTRICITY.XLS]

Electricity charge (\$)									
96	171	202	178	147	102	153	197	127	82
157	185	90	116	172	111	148	213	130	165
141	149	206	175	123	128	144	168	109	167
95	163	150	154	130	143	187	166	139	149
108	119	183	151	114	135	191	137	129	158

- a. Construct a frequency distribution and a percentage distribution with upper class boundaries of <\$100, <\$120, and so on. [Do manually or use Excel – Instructions in Textbook pg. 80 - Summarising Numerical Data]

Note: For Excel, you need to manually setup Bins. Resultant Excel output - needs refining

- b. Plot the corresponding histogram and percentage polygon.
  - c. Construct the corresponding cumulative percentage distribution and plot the corresponding ogive (cumulative percentage polygon).
  - d. Around what amount does the monthly electricity cost seem to be concentrated?
- 2.19 The ordered arrays in the table below give the life (in hours of usage) of samples of forty 15-watt CFL (Compact Fluorescent Lamp) energy-saving light bulbs produced by two manufacturers, A and B. [Dataset: BULBS.XLS]

Manufacturer A									
5544	5814	6190	6307	6342	6423	6429	6485	6612	6667
6832	6868	6879	6930	6941	7007	7037	7043	7059	7136
7497	7645	7654	7773	7816	7838	7924	7999	8038	8067
8091	8119	8392	8416	8416	8514	8532	8542	8544	8731

Manufacturer B									
6701	6837	6961	7118	7133	7142	7156	7344	7493	7569
7607	7612	7651	7721	7754	7767	7806	7839	7888	7983
8298	8344	8535	8666	8792	8800	8856	8861	8993	9001
9036	9096	9262	9385	9460	9471	9521	9540	9693	9744

- Construct a frequency distribution and percentage distribution for each manufacturer.
- Plot the corresponding frequency histograms on separate graphs and the percentage polygons on the same graph.
- Form the cumulative percentage distributions and plot the ogives on one graph.
- Which manufacturer has bulbs with a longer life – Manufacturer A or Manufacturer B? Explain.

2.23 The following table classifies road fatalities in Australia from 2012 to 2016 (inclusive) by age and gender.  
[Dataset: ROAD\_FATALITIES\_2012\_2016.XLS]

Age	Gender			Total
	Male	Female	Unknown	
< 10	89	74	3	166
10 to < 20	402	182	0	584
20 to < 30	990	310	0	1300
30 to < 40	686	185	0	871
40 to < 50	693	182	0	875
50 to < 60	534	179	0	713
60 to < 70	412	212	0	624
70 to < 80	319	162	0	481
80 to < 90	243	178	0	421
90 or more	56	47	0	103
Unknown	3	1	0	4
<b>Total</b>	<b>4427</b>	<b>1712</b>	<b>3</b>	<b>6142</b>

Data obtained from the Australian Road Deaths Database

<[www.bitre.gov.au/statistics/safety/fatal\\_road\\_crash\\_database.aspx](http://www.bitre.gov.au/statistics/safety/fatal_road_crash_database.aspx)> accessed 18 March 2017

Ignore the unknown categories.

- Investigate the relationship between age and gender by constructing a side-by-side bar chart to highlight the pattern of male and female road fatalities.
- Discuss the pattern of male and female road fatalities for 2012 to 2016.

#### TEXTBOOK REFERENCE:

Basic Business Statistics: Concepts and Applications. *Berenson, M.L. Levine, D.M. Szabat, K.A. O'Brien, M. Jayne, N. Watson, J.* 5th edition. 2019. Pearson Australia Group Pty Ltd. ISBN 9781488617249. Chapter 2, sections 2 to 2.7