

Figure 5.7. DATA-BASED INVESTIGATING: Error – Its Categories and Sources
Section (and subsection) headings to provide a content overview

1. Error in Data-based Investigating (page 5.19)
2. Six Categories of Error in Data-based Investigating (page 5.19)
3. Study Error, Sample Error and Measurement Error (pages 5.19 to 5.22)
4. Plan Components to Manage Study Error, Sample Error and Measurement Error (pages 5.22 to 5.24)
5. Non-Response Error (pages 5.24 to 5.26)
6. Plan Components to Manage Non-Response Error (pages 5.26 and 5.27)
7. Statistical Modelling and Model Error (pages 5.27 and 5.28)
8. Plan Components to Manage Model Error (page 5.28)
9. Investigating Statistical Relationships: Changing and Comparing (pages 5.28 and 5.29)
10. Comparison Error – Lurking Variates and Confounding (pages 5.39 and 5.40)
11. Association – Statistical Issues (pages 5.30 and 5.31)
12. Causation – Statistical Issues (pages 5.32 and 5.33)
13. Association Among Variates and Causation (pages 5.34 and 5.35)
14. Investigating Statistical Relationships (pages 5.35 and 5.36)
15. Terminology for Comparative Plans – The Protocol for Choosing Groups (pages 5.36 to 5.38)
16. Plan Components to Manage Comparison Error (page 5.38)
17. Experimental Plans – Sample selecting and Blocking (pages 5.38 and 5.39)
18. Observational Plans – Sample selecting, Matching and Subdividing (pages 5.39 and 5.40)
19. Comparative Plans and Causal Structures (pages 5.40 to 5.43)
20. Comparative Plans – The Protocol for Setting Levels and Interaction (pages 5.43 to 5.45)
21. Experimental Plans – Quantifying a Treatment Effect Under EPA (pages 5.45 to 5.49)
22. Observational Plans – The Confounding Effect (page 5.49)
23. Comparison Error in Experimental and Observational Plans (pages 5.50 to 5.52)
24. Summary of Error Management Strategies (pages 5.52 to 5.55)
25. Appendix 1: Populations and Processes (page 5.55)
26. Appendix 2: Equiprobable (Simple random) Selecting – The Protocol for Selecting Units (pages 5.56 and 5.57)
27. Appendix 3: Fishbone Diagrams for Comparative Plans (pages 5.57 to 5.59)
28. Appendix 4: Sample Size and Sample Error under EPS (page 5.59)
29. Appendix 5: Measuring Processes (pages 5.59 to 5.62)
30. Appendix 6: Bias and Rms Error (page 5.63)
31. Appendix 7: Illustrative Newspaper Article (page 5.64)
32. Appendix 8: Lurking Variates – Scatter Diagrams (page 5.65)
33. Appendix 9: Lurking Variates – a Broader Perspective (pages 5.65 to 5.70)
 1. Illustrations of Simpson's Paradox (pages 5.66 and 5.67)
 2. 'Simpson's Paradox' with a quantitative response variate (page 5.67)
 3. Reasons for Simpson's Paradox – properties of proportions (pages 5.67 and 5.78)
 4. Reasons for Simpson's Paradox – population subgroups and weighted averages (page 5.68)
 5. Reasons for Simpson's Paradox – probability distributions (page 5.68)
 6. A Plan for an investigation to answer the Question of sex discrimination (pages 5.68 to 5.70)
34. Appendix 10: Confounding – Usage in Statistics (pages 5.70 to 5.73)
35. Appendix 11: Reality and Illusion in Statistical Association and Causation (pages 5.73 to 5.76)
36. Appendix 12: Connections Among Three Variates (pages 5.76 and 5.77)
37. Appendix 13: Simpson's Paradox and Interaction (pages 5.77 to 5.79)
38. Appendix 14: Question Aspect and Method of Sample Selecting (pages 5.79 to 5.82)
39. Appendix 15: Limitations on Answers from Observational Plans (pages 5.82 to 5.84)
40. Appendix 16: Error Categories, Samples and Populations (page 5.84)

(continued overleaf)

41. Appendix 17: Error Categorization in the Social Sciences (pages 5.84 and 5.85)
 42. Appendix 18: Selecting Protocols and Unit Inclusion Probabilities (pages 5.85 and 5.86)
 43. Appendix 19: Themes, Symbols and Acronyms for this Figure 5.7 (page 5.86)

Fifty-three tables (plus three repeats)

One hundred and twenty-two diagrams (plus fifty-five repeats)

Eighteen equations (plus three repeats)

Tables		Equations		NOTES	
No.	Page(s)	No.	Page(s)	No.	Page
5.7.1	5.21	5.7.1	5.25	1	5.19
5.7.2	5.22	5.7.2	5.26	2, 3	5.20
5.7.3	5.25, 5.54	5.7.3	5.28	4, 5	5.21
5.7.4	5.26	5.7.4	5.33	6, 7	5.22
5.7.5	5.28	5.7.5	5.49, 5.51	8-11	5.23
5.7.6	5.28	5.7.6	5.50, 5.54	11-14	5.24
5.7.7	5.33	5.7.7	5.51, 5.54	15, 16	5.26
5.7.8	5.35	5.7.8	5.63	16	5.27
5.7.9	5.37	5.7.9	5.63	17-20	5.28
5.7.10	5.38	5.7.10	5.63	21, 22	5.29
5.7.11	5.39, 5.61	5.7.11	5.63	23	5.31
5.7.12	5.39	5.7.12	5.63	24	5.32
5.7.13	5.42	5.7.13	5.63	24-29	5.33
5.7.14	5.45	5.7.14	5.63	30, 31	5.35
5.7.15	5.45	5.7.15	5.63	32	5.36
5.7.16	5.46	5.7.16	5.83	33-36	5.37
5.7.17	5.47	5.7.17	5.83	36, 37	5.38
5.7.18	5.52 and 5.53	5.7.18	5.83	38	5.39
5.7.19	5.55			39	5.40
5.7.20	5.59			40, 41	5.42
5.7.21	5.66			41, 42	5.43
5.7.22	5.66			43-46	5.44
5.7.23	5.66			46-48	5.45
5.7.24	5.66			49, 50	5.46
5.7.25	5.67			50-52	5.47
5.7.26	5.67			52, 53	5.48
5.7.27	5.67			53, 54	5.49
5.7.28	5.68			55	5.51
5.7.29	5.68			55, 56	5.52
5.7.30	5.68			57-59	5.55
5.7.31	5.68			60, 61	5.56
5.7.32	5.68			61-63	5.57
5.7.33	5.68			64	5.58
5.7.34	5.69			64	5.59
5.7.35	5.69			65	5.61
5.7.36	5.69			65-68	5.62
5.7.37	5.69			69-71	5.63
5.7.38	5.69			72-74	5.65
5.7.39	5.71			75, 76	5.69
5.7.40	5.75			76	5.70
5.7.41	5.75			77-79	5.72
5.7.42	5.76			79	5.73
5.7.43	5.78			80-82	5.75
5.7.44	5.78			82	5.76
5.7.45	5.78			83-85	5.78
5.7.46	5.78			85-88	5.79
5.7.47	5.83			89-93	5.82
5.7.48	5.84			94-96	5.84
5.7.49	5.84			97	5.85
5.7.50	5.85			98-102	5.86

Figure 5.7, in part because of its length, involves a number of sequences of ideas; these can only be developed one at a time. Clarity of presentation and dependencies among ideas in different sequences often restricts us to *partial* development of a particular sequence in separate parts because of need for an idea not yet reached in another sequence. The four hundred and seventy or so cross-references in the text are given, in part, to assist the reader in recognizing both the continuation of a previous sequence and the relationship(s) of its ideas to those in other sequences. [Recall also the comment near the top of page 5.55 about a strictly sequential presentation of the ideas of introductory statistics.]