

Resource Book

Market and Consumer Research Semester 1

-2020

Table of Contents

Part I - General Guidelines for Assessment Preparation	1
1. Self-directed Learning	2
(a) How Much Time Should One Allocate Towards this Course?	2
(b) MKTG7510 offering is an internal-mode course.....	2
2. Guidelines for Submitting Assignments.....	3
(a) Assignments	3
(b) Electronic Submission	3
(c) A Note on Plagiarism.....	4
3. Guidelines for Requesting an Extension.....	5
4. Explanation of Course Grading or What should I do to Achieve a 7?	6
(a) IF there is a Discrepancy in the marks awarded to you.....	8
(b) If you feel that the mark allocated on an assessment piece does not reflect the standard of your work:.....	8
5. Guidelines for Presentation, Style and Referencing.....	10
(a) Style	10
(b) Headings	10
(c) Numbers.....	10
(d) Structure	11
(e) Writing.....	11
(f) Spelling and punctuation.....	11
(g) Tables and figures.....	12
(h) Referencing	12
6. Guidelines for Preparing Assignment 1 (Market Research Proposal)	16
(a) The Task	17
(b) Some of the important questions your report needs to answer.	17
(c) Topics and Related Background Readings	18

(d) Appendices	25
(e) Length, Weight, Due Date and Venue	26
(f) Assignment 1 Scenarios	27
(g) Assignment 1 (30%) - Research Proposal (Individual Task) - Marking Sheet.....	34
(h). List of Journals Useful for Preparing the Assignments.....	37
7. Guidelines for Preparing Assignment 2 (Market Research Report) (30%).....	41
(a) Assignment 2 (Individual Task).....	42
(b) The Task	43
(c) General Guidelines.....	43
(d) Length, Weight, Due Date and Venue	51
(e) Assignment 2 (30%) - Research Report (Individual Task) - Marking Sheet	52
8. Guidelines for Final Examination (40%) Preparation –	56
(a) Examination Structure	57
(b) Examination Strategy.....	57
(c) Useful Tips	57
Part II - Practicals	60
Workshop Timetable	61
Workshop 1: Introduction to SPSS.....	62
Workshop 2: Descriptive Statistics	72
Additional Resources	80
Symbols Frequently Used in Statistics	81
Workshop 3: Correlation Analysis	83
Workshop 4: Cross Tabulations	94
Workshop 5: t-test	109
Workshop 6: Analysis of Variance	122
Workshop 7: Regression Analysis	132
Part III – Class Exercises.....	157

Exercise 1: Building Blocks of Research.....	158
Exercise 2: Problem Formulation.....	160
Exercise 3: Research Design	162
Exercise 4: Research Design- II-Ethics	164
Exercise 5: Survey Methods.....	167
Exercise 6: Scaling	170
Exercise 7: Questionnaire Design	173
Exercise 8: Sampling Design	176
Exercise 9: Data Analysis	178

Part I - General Guidelines for Assessment Preparation

1. Self-directed Learning

(a) How Much Time Should One Allocate Towards this Course?

MKTG7510 is a #2 unit course. UQ's expectation is that a #2 unit course requires a time allocation of, on average, 10 hours per week. Lectures and computer labs only account for 3 hours per week. Therefore, if you are attending all of these contact sessions, it is expected that you will be spending a further 7 hours per week in self-directed learning. This could include attending consultation sessions, textbook study, attempting practice questions at the end of each textbook chapter, etc.

150 Hours approximately*

- Attendance – 38 hours approximately
 - Class attendance – 26 hours (2 hours per week)
 - Lab attendance – 12 hours (1 hour per week)
- Weekly reading – 39 hours (3 hours per week)
- Blackboard – 13 hours approximately
- Assignment 1 – 25 hours? Assignment 2 – 25 hours?
 - Exam – 10 hours?

*Note: This should be used as a rough guide only.

Please give serious consideration to whether you are willing and able to make this time commitment before enrolling. Experience shows that students who do not put in this basic time allocation tend to struggle.

(b) MKTG7510 offering is an internal-mode course

This MKTG7510 offering is an internal-mode course. That is, although attendance is not mandatory, UQ offers lectures and computer labs and expects that you will be attending them. Results from previous years show the relationship between attendance and assessment performance is positive and statistically significant.

2. Guidelines for Submitting Assignments

(a) Assignments

(i) Assignment 1 (Market Research Proposal) - 30 % (Individual task)

- Word count: 1500-2000 (maximum 2000 words).
- Due: Latest by **2 pm, Friday, 10 April 2020 (Teaching Week 7)**
- The report involves the preparation of a market research proposal.

(ii) Assignment 2 (Market Research Project Report) – 30% (Individual task)

- Word count: 3000-5000 (maximum 5000 words).
- Due: Latest by **2 pm, Friday, 29 May 2020 (Teaching Week 13)**
- Research report for a client.

(b) Electronic Submission

- Students should submit an electronic copy of the assignment (as a Word document, not as pdf or in other form) through the course Blackboard, by the due date and time.
- Multiple files cannot be accepted. Seek help on using the word processing software if you need assistance with combining your tables and main body.
- If any parts of the assignment are missing (e.g. tables) in the electronic copy:-the assignment may not be considered for marking. Or, it may result in the loss of marks for the missing section(s).
- The electronic copy submitted through BlackBoard will be screened through the plagiarism detection software “turn-it-in”.
- Non-submission of the assignment will result in a zero mark being awarded to the concerned student.
- Late submissions incur penalty (5% marks per day or part delay).

(c) A Note on Plagiarism

- **See Section 6.1 of the course ECP**

The University has adopted the following definition of plagiarism:

Plagiarism is the act of misrepresenting as one's own original work the ideas, interpretations, words or creative works of another. These include published and unpublished documents, designs, music, sounds, images, photographs, computer codes and ideas gained. These ideas, interpretations, words or works may be found in print and/or electronic media.

Students are encouraged to read the UQ Student Integrity and Misconduct policy (<http://ppl.app.uq.edu.au/content/3.60.04-student-integrity-and-misconduct>) which makes a comprehensive statement about the University's approach to plagiarism, including the approved use of plagiarism detection software, the consequences of plagiarism and the principles associated with preventing plagiarism.

3. Guidelines for Requesting an Extension

- **See Sections 5.3 & 6.1 of the course Electronic Course Profile (ECP)**
- **We strongly encourage you to submit the assignments in time**
 - Most students who do well in this course submit their assignments within time.
 - Be proactive; take the time to plan the submission of assignments. The assignment submission due dates are made known to you **before the start of the semester** so that you can plan the submission of your assignments. Lack of planning cannot be a valid excuse for requesting extra time.
 - Extra time can only be granted to the needy: In case of extraordinary circumstances (e.g. sickness/bereavement). That is, for unplanned circumstances which are beyond one's control.
- **Please avoid seeking extensions on trivial grounds**
 - That is, “having other assignments due at the same time” or “having work commitments”. If you foresee any difficulty in submitting on Friday (e.g. because of work commitments), we encourage you to submit before (say on the Thursday before).
- **Requests for extensions should be made at least a day before the due date:**
 - Students must fill in an Application for Extension of Assessment Due Date and provide supporting documentation to **Assessment, Examinations & Misconducts Coordinator**
<https://my.uq.edu.au/information-and-services/manage-my-program/exams-and-assessment/applying-extension> at least 24 hours prior to the submission date.
 - The request must explicitly state how much time is being requested for. The request **must also contain documentation** satisfactorily supporting the request.
 - Students will be notified of the outcome of their request via the myRequests section of my.UQ (and an email sent to their student account).
 - Additional details associated with extension requests, including acceptable and unacceptable reasons, may be found at my.UQ.
 - If you do not receive a reply to your extension request, do not assume that an extension might be/has been granted.

4. Explanation of Course Grading or What should I do to Achieve a 7?

Grade 7, High Distinction:

- **Demonstrates substantial understanding of fundamental concepts of the field of study and ability to apply these concepts** in a variety of contexts;
- develops or adapts **convincing arguments** and provides coherent justification;
 - with consistent evidence of substantial originality and insight in identifying, generating and communicating competing arguments, perspectives or problem solving approaches;
- **critically evaluates problems**, their solutions and implications;
- **communicates information and ideas clearly and fluently** in terms of the **conventions of the discipline**; and
 - uses a level, style and means of communication appropriate to the discipline and the audience:

Note:

- Usually achieving Grade 7 requires some reading outside the prescribed text and lecture notes (e.g. from the additional learning sources) and the use of those references.
- The assignment submitted will be similar to a professional consulting report.

Grade 6, Distinction:

- Demonstrates substantial understanding of fundamental concepts of the field of study and ability to apply these concepts in a variety of contexts;
- develops or adapts convincing arguments and provides coherent justification;
 - with frequent evidence of originality in defining and analysing issues or problems and in creating solutions; and
- communicates information and ideas clearly and fluently in terms of the conventions of the discipline; uses a level, style and means of communication appropriate to the discipline and the audience.

Grade 5, Credit:

- Demonstrates substantial understanding of fundamental concepts of the field of study and ability to apply these concepts in a variety of contexts;
- develops or adapts convincing arguments and provides coherent justification; and
- communicates information and ideas clearly and fluently in terms of the conventions of the discipline.

Grade 4, Pass:

- Demonstrates adequate understanding and application of the fundamental concepts of the field of study;
- develops routine arguments or decisions and provides acceptable justification; and
- communicates information and ideas adequately in terms of the conventions of the discipline.

Grade 3, Fail:

- Demonstrates superficial or partial or faulty understanding of the fundamental concepts of the field of study and limited ability to apply these concepts;
- presents undeveloped or inappropriate or unsupported arguments; and
- communicates information or ideas with lack of clarity and inconsistent adherence to the conventions of the discipline.

Grade 2, Fail:

- Demonstrates clear deficiencies in understanding and applying fundamental concepts; and
- communicates information or ideas in ways that are frequently incomplete or confusing and give little attention to the conventions of the discipline.

Grade 1, Fail:

- Fails to demonstrate most or all of the basic requirements of the course.

Grade X:

- No assessable work received.

Grade Cut Off %

Grade (Description)		Range of Marks
1	(Serious Fail)	0 - 29
2	(Fail)	30 - 46
3	(Fail)	47 - 49
4	(Pass)	50 - 64
5	(Credit)	65 - 74
6	(Distinction)	75 – 84
7	(High Distinction)	85 – 100

Other Requirements:

- In addition, students must achieve minimum 50% from the two individual assessment tasks combined (Assignment 1 plus final examination), to achieve a pass grade from this course.

On Marking

(a) IF there is a Discrepancy in the marks awarded to you

- If you notice any discrepancies in the marks awarded to you (e.g. incorrect totalling/incorrect penalty levied):
 - you **MUST** contact the Course Coordinator, **in writing**, with the hard copy of the marked assignment, **within a week after the release of the marks**.

(b) If you feel that the mark allocated on an assessment piece does not reflect the standard of your work:

See Section 6.1 of the course ECP

- First, contact the Marker, with the hard copy of the assignment, and seek feedback, **only a week after the release of the marked assignments**.
- **Students will not be seen during the one week period, after the release of marks**, with the above requests.

If you are not satisfied with the feedback provided by the marker:

- contact the course coordinator (by email) within a week after the release of the marks, with evidence:
 - provide the hardcopy of the marked assignment and

- provide evidence in support of your claim that the assignment deserves a higher mark.

That is, provide a clear case (e.g., half a page) outlining why you believe you had received less marks than what you deserve (Against which marking criteria?)

- **No requests for remarking will be considered without this note**

- **Remarking is rare**

- Extensive care is taken in the marking.
 - Markers receive extensive training. Briefing sessions are conducted for each assignment for arriving at a consensus among the markers regarding the marking scheme and marking criteria.
 - After the marking has been completed, the Coordinator conducts moderation to check if marking is consistent across the markers. Special attention is given to any assignments that receive a fail mark.
 - All the markers are experienced and each would have marked hundreds of similar assignments.
- **If remarking is permitted, the revised mark (after the remarking) will be the final mark, irrespective of whether it is higher or lower than the originally awarded mark.**
- The chances of succeeding with remarking are slim.
- Overall, the onus is on you to develop a clear understanding of the assignment requirements.
 - Detailed marking templates and explicit guidelines on the preparation of the two assignments, briefings on the assignments (in the lectures and labs, and additional student consultations) are aimed at providing greater clarity on the requirements.

5. Guidelines for Presentation, Style and Referencing

Source: Adapted from the Journal of International Business Studies

(a) Style

- Indicate your assignment topic (e.g. “Brand extensions”) on the front page of the report.
- Provide a Table of Contents.
- Number all pages in your report, starting with the “Introduction” page.
- Use 2.5 cm margins at the top, bottom, and on both sides of the page.
- **Use 12-point** Times New Roman font for the text throughout the document.
- Use double-line-spacing for all body text.
- Use single line spacing for the text in the references, tables, appendices and the executive summary.
- Text should be left-aligned.

(b) Headings

- Use only a maximum of three levels of headings.
 - Use **main headings** to designate your major sections.
 - Center main headings, the words should be in bold, and only the first letter of major words should be capitalized (e.g. **Data Analysis**).
 - **Second-level headings** should be flush with the left margin, and only the first letter of major words should be bold and capitalized (e.g. **Data analysis**).
 - **Third-level headings** should use regular and italicized font; begin the first word with a capital (e.g. *Data analysis*).
 - **Do not use fourth level of headings.**

(c) Numbers

- Write out numbers below 10 (e.g. say ‘nine’ instead of 9).
- Insert leading zeros before decimal points in text and tables (e.g., '0.3' rather than '.3'). Report only two decimal places for statistics.

(d) Structure

- **Do not** start each section of the report on a new page.
- Each table starts on a new page.
- **Do not** use footnotes.
- **Use endnotes** sparingly, and keep them short. Include citations to the literature in the text, not in the endnotes.
- Place **references** and **appendices** at the end of your report.
 - Each component at the end of the report (e.g. endnotes, references and appendices) should begin on a new page.

(e) Writing

- Write in an interesting, readable manner.
 - Write **short paragraphs**.
 - Keep sentences short so the reader does not get lost before the end of a sentence.
- **Do not use code names or abbreviations** for variable names, unless the terms are well known to business managers (e.g. GDP and FDI).
 - Use the same name for a variable throughout your report.
 - **Do not italicize variable names.**

(f) Spelling and punctuation

- Check and correct spelling and punctuation before submitting your report.
 - Be consistent in your capitalization, spelling, hyphenation and formatting throughout the report.
 - **Avoid common usage errors** such as "it's" and "its", "affect" and "effect", and "that" and "which".
 - Singular (plural) nouns require singular (plural) pronouns and verbs.
 - "Who" should be used for people; "that" and "which" for organizations.
 - Use semicolons to help define long lists that include both groups and subgroups.

(g) Tables and figures

- **If a Table spans more than a single page: repeat the column headings and the title of the Table on subsequent pages.**
- Number tables and figures consecutively, using Arabic numerals, in order of appearance (one series for tables [Table 1, Table 2...], one for figures [Figure 1, Figure 1...]).
- Check that the sequence of any numbered elements (tables, figures etc.) is correct.
- **Each table or figure must have at least one sentence in your text that introduces it.**
- In-text references to tables should be in sequential order throughout the paper.
- A table should be understandable on its own. The text should highlight the main points in a table and summarize its message, but not duplicate the details.
- Titles of tables and figures should be short and descriptive.
 - They should not contain acronyms, abbreviations or symbols.
 - The number and title for each table or figure should be typed on separate lines. Cite sources directly below each table or figure.

(h) Referencing

- When citing a list of references in the text, put the **list in alphabetical order** and separate authors by semicolons.
 - For example, "Several studies (Pappu & Quester, 2006a, 2006b; Pappu, Quester & Cooksey, 2005) support this conclusion."
- If a paper being cited has two authors, give both names each time you cite it.
 - For example, Pappu and Quester (2006).
- If a paper being cited has three authors, give all 3 names the first time the work is cited and then use the first author's name and "et al." for all subsequent citations.
 - For example, say Pappu, Quester and Cooksey (2005) the first time, and Pappu et al. (2005) when the same source is cited a second time or thereafter.

Referencing (Continued...)

- For works with four or more authors, use the first author's name and "et al." for all in-text citations, including the first citation.
 - For example, say Brailsford et al. (2009) for the following publication).
 - Brailsford, T., Dash, S., Hutchinson, M., Pappu, R. & Richie, B. (2009) "Measuring Brand Equity of A Geographic Region," *Tourism Analysis*, 14(6), 765-779.
- **Quotations:** To cite a direct quotation, give page numbers after the year of publication, separated by a colon and a space.
 - For example: Pappu and Quester defined country equity as "the value endowed by the name of the country on to a product" (2001; p. 258).
- The reference list should follow the notes at the end of the report in alphabetical order by author name (use the "corporate author" or the journal name where no individual author's name is given).
 - This reference list must include all the works you have cited, and only the works you have cited.
- Make certain that the reference for each citation in the text is complete, and that the cited dates and the spellings of the authors' names in the text and references are in agreement.
 - Include full page range for all journal references and book chapters.
 - Include both the volume and issue (number [if applicable], season, month, or date) for journal/periodical references.

The following are examples of proper form:

- **Journal/periodical articles**

Pappu, R., Quester, P. & Cooksey, R.W. (2007), "Country Image and Consumer-Based Brand Equity: Relationships and Implications for International Marketing," *Journal of International Business Studies*, 38(5), 726-745.

Financial Times. (1996). Survey - Czech Republic: Message from the people. December 6: 3.

Referencing (Continued...)

- **Books**

Donahoe, J. D. (1989). *The privatization decision*. New York: Basic Books.

- **Papers**

Pappu, R. and Quester, P. (2005), Does Retailer Equity Vary Across Store Categories: Empirical Evidence from Australia, In C. De Moranville (Ed.), *Proceedings of World Marketing Congress*, Vol XII, p. 212, 7–8 July, Muenster, Germany: Northern Illinois University, USA.

- **Chapters in edited books**

Chien, P. M., Cornwell, T.B., & Pappu, R. (2008). A Theoretical Framework for Analysis of Multiple Sponsorship Effects on Consumer responses. In F. C. Pereira, J. Veríssimo & P. Neijens (Eds.), *New Trends in Advertising Research*:189-206. Lisbon, Portugal: Edições Sílabo.

- **Dissertations**

Pappu, R. (2003). *Evaluation of a theoretical framework for explaining the interrelationships between country-of-origin effects and consumer-based brand equity*. Unpublished PhD Dissertation, School of Marketing and Management, University of New England, Australia.

- **Online documents**

The Investment Company Institute. (2004). Worldwide mutual fund assets and flows, third quarter 2003. <http://www.ici.org>. Accessed 4 February 2004.

- **Online journal articles / advance online publication articles**

Hutzschenreuter, T., & Voll, J. C. (2007). Performance effects of "added cultural distance" in the path of international expansion: The case of German multinational enterprises. *Journal of International Business Studies*, advance online publication August 30. doi:10.1057/palgrave.jibs.8400312.

This page has been intentionally left blank

6. Guidelines for Preparing Assignment 1 (Market Research Proposal)

6. Guidelines for Preparing Assignment 1 (Market Research Proposal)

- **Start work on the assignment early in the semester if you want a decent grade**

(a) The Task

This assignment involves the preparation of a market research proposal/plan.

(b) Some of the important questions your report needs to answer.

- What is the management/decision problem?
- Is there a management/decision problem?
- What is/are the research problem(s)?
- Why is it important to conduct the above research?
- What is/are the hypothesis/hypotheses?
 - What does current literature say about the relationship/s hypothesized?
 - How do you define the variables conceptually?
- What type of research design is the most appropriate? Why?
 - What is the research method? Why?
 - What is the appropriate data collection method? Why?
 - What is the appropriate sampling design? Why?
 - What is the population of interest?
 - How do you plan to reduce sampling/non-sampling errors?
- How to measure the variables of interest? What are the operational definitions?
 - How do you plan to ensure validity and reliability?
- What is/are the appropriate data analysis technique(s)? Why?

Step 1: Select One Scenario from those Suggested (Provided separately)

- Although the scenario topics are fixed, their content may be slightly modified to suit your own purposes.

- That is, you may slightly change the scenarios to fit your particular interests as your thoughts develop. You are not allowed to make major changes to the scenario. That is, you can not change the scenario drastically or invent a new scenario.
- You may change/drop/include variables in line with your interests.
- **You can not duplicate an existing research study: that would constitute plagiarism.**

(a) Select a topic and complete the background reading (Download the reading from the course Blackboard or Library)

(c) Topics and Related Background Readings

1. Brand Innovativeness

Pappu, R. & Quester, P. (2016), "How does Brand Innovativeness Affect Brand Loyalty?" *European Journal of Marketing*, 50(1/2), 1-43.

2. Sponsorship

Pappu, R. & Cornwell, T.B. (2014), "Corporate Sponsorship as An Image Platform: Understanding the Roles of Relationship Fit and Sponsor-sponsee Similarity," *Journal of the Academy of Marketing Science*, 42(5), 490-510.

3. Celebrity endorsement

Spry, A., Pappu, R., & Cornwell, T. B. (2011). "Celebrity Endorsement, Brand Credibility and Brand Equity". *European Journal of Marketing*, 45(6), 882-909.

4. Retailer equity

Pappu, R., Quester, P. & Cooksey, R.W. (2005), "Consumer-based Brand Equity: Improving the Measurement - Empirical Evidence", *Journal of Product & Brand Management*, 14(3), 143-154.

5. Country-Brand Relationships

Pappu, R., Quester, P. & Cooksey, R.W. (2007), "Country Image and Consumer-Based Brand Equity: Relationships and Implications for International Marketing," *Journal of International Business Studies*, 38(5), 726-745.

6. Brand Identification

Bhattacharya, C.B. & Sen, S. (2003), "Consumer-Company Identification: A Framework for Understanding Consumers' Relationships with Companies," *Journal of Marketing*, 67(2), 76-88.

(b) Read the selected scenario (see Step 3 - Section f) thoroughly

Step 2: Identify management/decision problem and research problems

The decision maker described in the scenario is facing a **management problem**. Your task is to identify the **management/decision problem** and state this in a question form.

Note that, if you make any changes to the scenario, the management and research problems would change. Also, there may be more than one research problem.

Two students working on the same topic/scenario can come up with different management and research problems. Don't panic if your solution is not similar/identical to that of other students working on the same topic.

Step 3: Apply the relevant concepts and principles and provide a research design. Prepare the Report

(3)Structure

The report should include the following sections.

Table of Contents

1. Executive summary (1 page) (Not included in the word count)
2. Background and Problem Definition (3 – 5 pages)
 - Management/Decision problem
 - Importance and Benefits
 - 2.1. Exploratory research
 - 2.2. Research questions
 - 2.3. Hypotheses development
3. Research Design (3 – 5 pages)
 - 3.1 Type of research design
(e.g. exploratory/descriptive/causal/a combination)
 - 3.2 Sources of data
 - Secondary data sources
 - Primary data collection plan (survey, focus group, interview, experiment)

Step 3: Prepare the Report (Continued...)

3.3. Sampling design

- Population
- Sampling method
- Sampling frame
- Sample size

4. Data collection

- Questionnaire/interview guide

5. Data Analysis

References (Not included in the word count)

Appendix: (Not included in the word count)

- Table 1. Conceptual definitions
- Table 2. List of Hypotheses
- Table 3. List of Measures
- Table 4. Special facilities and resources (if required) and
- Table 5. Schedule

(ii) Content

Executive summary

The executive summary should provide an overview of the entire proposal. Briefly discuss the nature and scope of the proposal.

That is, indicate the type of study being proposed (i.e., the broad type of research design), how the study would be conducted, and tell the reader the most significant implications of the research.

(Remember, the proposal serves an important secondary purpose: it is a marketing document!).

This section should provide an overview to a busy executive the nature and content of the proposal. If interested, the reader would look inside the proposal.

Generally, this section will be of a full page (e.g., 400 words).

Step 3: Prepare the Report (Continued...)

1. Background and Problem Definition

- This introductory section should provide an overview of the research plan.
- Think about the issues that you will be discussing subsequently and in more detail (e.g., the problem statement, research objectives, literature, and importance of the study).
- Signal to the reader that you have a clearly defined problem, and one that is important and worth pursuing.
- Show that you are familiar with the literature, by integrating some key citations.
- Remember that a research proposal is a selling document. You need to keep readers interested and convinced of the importance of your planned study.

The **problem definition** section should provide the reader with some background information (e.g., about the firm and scenario you have selected for study), as well as clearly state the management problem and the research problem (you will need to specify these yourself, based on one of the scenarios).

- Try to convince the reader that the problem is interesting and important, and that the firm would benefit if it were solved.
- The problem definition section sets the broad framework for the research proposal; that is, the objectives (research hypotheses), literature review, and later, the research design, should flow on logically from the problem statement.
- Be sure to identify the independent and dependent variables for explanatory studies, and the nature of their relationship (e.g., positive or negative, linear or curvilinear).

Step 3: Prepare the Report (Continued...)

Importance and benefits of the study

This section would not be more than half a page in length.

In this section, provide a clear statement of the importance of the study and the benefits that would be gained if the study were conducted.

Think about both specific benefits and possible broader outcomes.

- A study could be important because it generates new insights into a problem or because a solution to the problem generates economic benefits or both!
 - For example, a business research problem could be connected to the goals of a firm or help to evaluate strategies for reaching those goals.
- More generally, solving a research problem might lead to superior effectiveness, quality and innovation in some process or processes.
- Lastly, stress the importance of conducting the study now!

Support for your hypotheses using relevant Literature

This section would not be more than a couple of pages long.

- Provide clear **conceptual definitions** of the variables of interest.
- Identify and briefly describe **theoretical perspectives** authors have used in the past (e.g., the theories to which authors refer).
- Use the summary of literature to build a **case for your research hypotheses**. Put forward an argument, but be objective and present a balanced view.
 - Stating other researchers' findings, including both supporting and contradictory evidence is important.
 - Contradictory evidence or areas of disagreement might be interesting discussion points and suggest a need for further research.
- Later, in the Research Design section, you will also be required to specify measures or “**operational**” **definitions**.
- Should give you insight into the research design/methods that could be used

Step 3: Prepare the Report (Continued...)

2. Research design (Focus on “Why”)

(Generally, this section is around 3-5 pages in length)

This section is **key** to the research proposal.

- **Type of research:**
 - Start this section by giving an indication of the **overall research design** (exploratory/descriptive/causal) and why it is appropriate. It is a big problem if this section is missing.
 - Then, you need to describe and justify the **research method(s)** (experiment/survey/observation). What is the specific experimental design used (if used)? What is the specific survey mode (if used)? What will be the specific mode of observation (if used)?
 - Specific design issues need to be addressed in this section, such as **sampling, data collection** and **measurement**.
- **Sampling design:**
 - Clearly state the sources of data for the study.
 - Define the population for the study and describe how a sampling frame would be developed. Justify the chosen probability/non-probability sample.
 - What is the sample size? Why?
- **Measurement:**
 - Indicate how you are going to measure the variables included in your hypotheses. [Hint: most good researchers use valid and reliable measures available in the literature.](#)
 - What are the sources (references) for the measures you plan to use?
 - What type of scale will be used?
 - How can you tell us that the measures are reliable and valid?
 - How many items are used for measuring each construct?

Step 3: Prepare the Report (Continued...)

- **Data collection:**

- What are the procedures that would be used to collect the data?
- How are you planning to collect the data (communication/observation)? Why?
- How will be the various types of error reduced?

Consider **ethical issues** here.

- The discussion on research ethics might include a reference to the researcher's responsibilities to: the client, research participants and researcher workers.
- Maintaining respondent anonymity and confidentiality is a crucial concern.
- Researchers also have responsibilities to their clients (e.g., professionalism, quality research).
- Think about the practices that might be help to maintain a researcher's responsibilities to clients and respondents (and research workers).

- **Data analysis**

In this section, you need to specify the methods of analysis that would be appropriate for testing your research hypotheses (or sorting qualitative data, if the design is exploratory).

Specific analytical methods should be identified – how will each hypothesis actually be tested? (Through correlation or regression analysis, for example.) The reason you are using a data analysis technique is because you need information from the data.

- How many pieces of information do you need? How many techniques do you need to generate the information required?
- Keep in mind that the analysis procedures should be consistent with the design, objectives and sampling methods employed.

Step 3: Prepare the Report (Continued...)

Also, give the reader an indication of how the outcomes of the research might look. Think about specific and more general outcomes. What do you hope to achieve through adapt analysis?

- Specific outcomes might include statistical conclusions.
- More general outcomes should also be identified (e.g., a strategy for increasing customer loyalty, a new compensation and evaluation system).
- A statement about how the results will be communicated should also be included here (e.g., an oral presentation, a research report).

References

- It is unlikely for a good proposal to have less than **15 references** to relevant sources (for the purpose of this course).
- Use author-date system. If you follow the referencing style used in any of the six lead articles is a good practice.
- It is unlikely for a good report not to have references to the prescribed text and/or something equivalent.

(d) Appendices

Table 1: Conceptual definitions

- Provide definitions of the key concepts used in a table including references to the original sources and page numbers.

Table 2: Hypotheses

- Provide the hypotheses developed in a table.

Table 3: List of Measures

- Provide a list of the measures and scales to be used, including references to the original sources and page numbers.

Table 4: Special facilities and resources required

- Special facilities and resources might be required to complete a project. Outline any special equipment, services, facilities or personnel capabilities that would be needed in this section (e.g., statistical software, a room for focus group interviews, travel for data collection purposes).

Table 5: Schedule

- At a minimum, a schedule should state the proposed time frame for the study.
- A more detailed schedule would include a listing of each step in the research process and the time required for each step – allowing the reader to identify critical paths!
- Think about how much time would be required to complete each stage. (Can you be specific and state the total number of hours?)

Hint: don't spend too... much time on Table 4 and Table 5.

(e) Length, Weight, Due Date and Venue

- Word count: 1500-2000 words; **Weightage: 30%**
 - Note that **you would lose marks if you exceed the suggested word count**. Quality of analysis, not the length, of the proposal is important.
 - References, headings, executive summary, table of contents, tables and appendices **are not included** in the word count.
- Due: Latest by 2 pm, [Friday, 10 April 2020, \(Teaching Week 7\)](#)
- Submit an electronic copy via the course Blackboard. Late submission will incur penalty.

(f) Assignment 1 Scenarios

General instructions

Your Assignment will need to be prepared using one of the following scenarios. There are many variables to choose from each scenario so that each research report will be a little bit different.

The most frequently asked questions are: “Which variables should I choose?” and “What is the minimum number of variables I should include in my study?” There are no perfect answers to these questions. You are supposed to answer these questions. Don’t ask for answers.

Let me say that you don’t get more marks by including more variables in the design. The objective behind the assignment is to test your ability to apply market research principles for designing research proposals to generate information for decision making. We would like to see how well your suggested design gets the manager the information they need.

You could design a very good study by examining the impact of say 3 or 4 variables on the focal variable. You would not want to design a study with 10 or 15 variables, making it too complicated. Overall, do not get bogged down by focusing too much on the number of variables. Note: When the proposal is marked the markers will have a model answer. So if you make radical changes to this scenario, there is the risk of not addressing the main management problem in this scenario.

The background reading suggested towards the end of this brief introduces you to the relevant literature. You are expected to search the library data bases to find articles and information relevant to the selected topic.

Importantly, your research plan should allow you in the second assignment to use the data analysis techniques you study in the computer labs during the semester for analyzing the data you collect for the study.

Scenario 1: Brand Innovativeness

The brand manager of a large firm (you can select one) suspects that high brand innovativeness perceptions from consumers may lead to several beneficial outcomes (e.g., satisfaction, loyalty, perceived quality among others) but is unsure.

You are asked to write a research plan that will help him determine how innovative their brand is considered by consumers.

The manager is also interested in predicting how different factors (e.g., consumer familiarity with the brand, consumer attitude towards the brand, perceived quality of the brand and others) influence people's perceptions of the brand's innovativeness.

He also wants to understand whether brand innovativeness perceptions vary by the types of customers (e.g. males/females; different countries-of-origin: Australian/non-Australian born; income levels: high or low).

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Pappu, R. & Quester, P. (2016), "How does brand innovativeness affect brand loyalty?" *European Journal of Marketing*, 50(1/2), 1- 43.

Scenario 2: Sponsorship

The CEO of a charity/nonprofit is seeking your input for developing a research plan. The charity/nonprofit is planning to take up a corporate social sponsorship. Several potential sponsors (e.g., Nike/KFC/Sony/Visa) have approached him.

The CEO speculates that not all of these sponsorships would be “helpful”. She believes that while some may positively influence people’s attitudes toward the charity (e.g., create value and enhanced brand credibility, increase consumer satisfaction and improve image for the charity organization), others would not.

She is unsure how different consumer- and sponsor-related factors (e.g., consumer familiarity with the sponsor, consumer attitude towards the sponsor, sponsor’s fit with the non-profit organization, negative information about the sponsor, sponsor credibility, sponsor motive) influence people’s attitudes towards and perceptions of the non-profit organization (e.g., attitude toward the charity, intention to donate to the charity) sponsored.

She also wants to understand whether attitudes toward commercial sponsorship of the charity vary by consumer’s gender (e.g. males/females) and country-of-origin (Australian/non-Australian born).

You have been asked to develop a research plan to address this problem. Select one of the potential sponsors (e.g. Nike/KFC/Sony/Visa) and develop a research plan to address some of the above issues.

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Pappu, R & Cornwell, T. B. (2014), “Corporate sponsorship as an image platform: understanding the roles of relationship fit and sponsor–sponsee similarity,” *Journal of the Academy of Marketing Science*, 42(5), 490-510.

Scenario 3: Celebrity Endorsement

The Research Director of a large non-profit organization wants to know how the use of celebrities (e.g., Nicole Kidman/Brett Lee/Ricky Ponting/Patrick Rafter) to endorse their organization would affect people's support (e.g. donations, volunteering time, WOM) as well as people's attitudes toward their organization. She would benefit by knowing how celebrity endorsement affects customer evaluations of the nonprofit organization.

The Director is unsure how different consumer- and celebrity-related factors (e.g., celebrity's fit with the non-profit organization, negative information about the celebrity, credibility of the celebrity, expertise of the celebrity) and consumer-related factors (e.g. consumer familiarity with the celebrity, consumer attitude towards the celebrity) influence people's attitudes towards and perceptions of the non-profit organization endorsed.

She is keen to be able to predict customer attitudes and non-profit brand evaluations based on the celebrity and consumer related factors mentioned above.

She also wants to understand whether people's attitudes towards the endorsed non-profit organization vary depending on different types of customers (e.g., males/females; different nationalities: Australian/non-Australian born).

Select a celebrity (e.g. Nicole Kidman/Brett Lee/Ricky Ponting/Patrick Rafter) and develop a research plan to explore *some* of these issues in further detail in case of the selected celebrity.

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Spry, A., Pappu, R., & Cornwell, T. B. (2011), "Celebrity endorsement, brand credibility and brand equity". *European Journal of Marketing*, 45(6), 882-909.

Scenario 4: Retailer Equity

The marketing manager of a large retail chain (e.g. Myer/David Jones/K-Mart) is reviewing the performance of their brand name. In particular, she is keen to improve consumer-based brand equity (i.e. the value consumers attach with a brand name) of their chain as a brand.

She wants to predict the consumer-based equity of their retail chain based on consumer-related factors. She suspects that several consumer-related factors (e.g. customer age, prior experience with the store, knowledge of the store, culture) may affect the consumer based brand equity of the retail chain.

She also wants to understand whether male and females consumers associate different levels of equity with the retail chain brand, and whether consumer brand equity of the retail chain varies according to their (consumer) location (rural vs urban).

Design a research study that will help her determine the level of the chain's consumer-based brand equity and how some of these consumer-related factors affect the consumer-based equity of the retail chain.

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Pappu, R. & Quester, P. (2006), "A consumer-based method for retailer equity measurement: Results of an empirical study", *Journal of Retailing and Consumer Services*, 13(5), 317-329.

Pappu, R., Quester, P. & Cooksey, R.W. (2005), "Consumer-based Brand Equity: Improving the Measurement - Empirical Evidence", *Journal of Product & Brand Management*, 14(3), 143-154.

Scenario 5: Country-Brand relationships

A large distributor imports products and sells them to department stores in the Australian market. The distributor is concerned about changing consumer perceptions of foreign products in the Australian market. She believes the Global Financial Crisis has negatively affected people's perception of products made in certain countries.

Through exploratory research, the distributor understands that country of origin of a product influences consumers' attitudes, perceptions, brand evaluations and even purchase intentions.

She wants to understand how consumers view brands from other countries (e.g. France) made available to consumers in Australia. Specifically, she is keen to understand how Australian consumers perceive the image of a country (e.g. China/India/Vietnam/Italy/France) and how consumer evaluations of a brand are affected by their image of the selected country (e.g. France's country image).

She also wants to understand whether consumers' attitudes, perceptions, brand evaluations and purchase intentions in relation to products made in the selected foreign country, differ by their gender, country-of-birth (e.g., Australian/non-Australian born) and according to peoples' income levels (high vs low).

Select one foreign country (e.g., China/India/Vietnam/Italy/France) and develop a research plan to examine (some of) these issues.

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Pappu, R., Quester, P. & Cooksey, R.W. (2007), "Country Image and Consumer-Based Brand Equity: Relationships and Implications for International Marketing," *Journal of International Business Studies*, 38(5), 726-745.

Scenario: 6. Brand Identification

The Director of Leukaemia Foundation is keen to understand how their organization is perceived in the market. She believes that people who identify with a non-profit or charity brand are more likely to support the organization through donations, time and word of mouth.

The Director wants to assess the level of people's identification with their nonprofit brand. She also wants to understand how consumer-related factors (e.g. word-of-mouth, trust, age, prior experience, knowledge, culture, location) may affect people's identification with their organization (brand identification). She is also very keen to be able to predict consumer brand identification levels based on various consumer-related factors mentioned previously.

She also wants to understand whether: (1) people in different states (e.g. NSW, SA) identify with their non-profit brand to different levels and (ii) brand identification varies according to gender).

You are asked to write a research plan to address this problem.

The following background reading introduces you to the relevant literature. Read this journal article first before starting work on your assignment. You are also expected to search the library databases to find other journal articles and information relevant to this topic.

Background Reading

Bhattacharya, C.B. & Sen, S. (2003), "Consumer-Company Identification: A Framework for Understanding Consumers' Relationships with Companies," *Journal of Marketing*, 67 (2), 76-88.

(g) Assignment 1 (30%) - Research Proposal (Individual Task) - Marking Sheet

Student Number:

Criterion (Weightage)						
1. Problem formulation (50%) - Task: The information required by the decision maker is correctly identified. Management and research questions and research objectives are identified. The research questions and hypotheses are developed based on a review of relevant literature. - Content: The material included in this section adequately covers relevant parts of the research plan (e.g. management question, research question/s, objectives, importance and background information, conceptual definitions and hypotheses). Conceptual definitions are included in an appendix. - Quality: The quality of analysis reflects the ability to apply relevant marketing research principles. References to texts and journal articles on relevant literature and research design/methods.						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
2. Quality of the suggested research design (30%) - Task: The research design (e.g. type of data, research method, data collection method, sampling design, measurement) suggested is appropriate for generating quality information and for achieving the research objectives. - Content: The material included in this section adequately covers relevant parts of the research design (e.g. broad design choice, type of data, research method, sampling design, data collection method, measures). - Quality: The quality of analysis reflects the ability to apply relevant marketing research principles. Links to management and research problems and the hypotheses. References to relevant literature on research methods including journal articles.						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
3. Presentation, Structure & Referencing (20%) - Task: The report conveys information clearly and follows the presentation, structure and referencing standards suggested. - Content: The report adequately covers the style, presentation, and referencing guidelines suggested (e.g. Provides an executive summary includes key details; Double line spacing in text; Executive summary one-and-half-line spacing; 2.5 cm margins; 12 size font; Table of contents; Page numbers; Appropriate referencing). Good quality tables and figures, where required. - Quality: style, presentation, and referencing are of <u>a professional standard</u>. The proposal is logically organised and <u>free of grammatical errors</u>. Clear introduction. Suitable sub-headings. Clear expression.						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
Overall grade for the Report The research proposal is of professional standard, convincing to the decision maker, includes required details to undertake research and if carried out allows the generation of quality information that can help the decision maker.						
1	3	3	4	5	6	7

Legend

1 - Fails to demonstrate most or all of the basic requirements

2 - Demonstrates clear deficiencies in understanding and applying relevant concepts; communicates information or ideas in ways that are frequently incomplete or confusing and give little attention to the conventions of the discipline

3 - Demonstrates superficial or partial or faulty understanding of relevant concepts and limited ability to apply these concepts; presents undeveloped or inappropriate or unsupported arguments; communicates information or ideas with lack of clarity and inconsistent adherence to the conventions of the discipline

4 - Demonstrates adequate understanding and application of the relevant concepts; develops routine arguments or decisions and provides acceptable justification; communicates information and ideas adequately in terms of the conventions of the discipline:

5 - Demonstrates substantial understanding of relevant concepts and ability to apply these concepts in a variety of contexts; develops or adapts convincing arguments and provides coherent justification; communicates information and ideas clearly and fluently in terms of the conventions of the discipline:

6 - As for 5, with frequent evidence of originality in defining and analysing issues or problems and in creating solutions; uses a level, style and means of communication appropriate to the discipline and the audience:

7 - As for 6, with consistent evidence of substantial originality and insight in identifying, generating and communicating competing arguments, perspectives or problem solving approaches; critically evaluates problems, their solutions and implications:

Part A: General comments (what was done well and not so well):

.....

.....

.....

Part B:

Suggestions for improvement (what could have been done to move to the next higher grade):

.....

.....

.....

.....

.....

.....

.....

Marker:

Date:

Note: All the comments included above are only aimed at providing constructive feedback

This page has been intentionally left blank

(h). List of Journals Useful for Preparing the Assignments

European Journal of Marketing	Health Marketing Quarterly
Industrial Marketing Management	International Journal of Nonprofit and Voluntary Sector Marketing
International Journal of Research in Marketing	International Journal of Sports Management and Marketing
Journal of Consumer Research	International Journal of Sports Marketing and Sponsorship
Journal of Marketing	Journal of Consumer Behaviour
Journal of Marketing Research	Journal of Marketing Communications
Journal of Retailing	Journal of Marketing Education
Journal of Service Research	Journal of Marketing Theory and Practice
Journal of the Academy of Marketing Science	Journal of Nonprofit and Public Sector Marketing
Marketing Science	Journal of Personal Selling and Sales Management
International Marketing Review	Journal of Product and Brand Management
Journal of Advertising	Journal of Social Marketing
Journal of Advertising Research	Qualitative Market Research: an international journal
Journal of Brand Management	Services Marketing Quarterly
Journal of Business Research	Sport Marketing Quarterly
Journal of Consumer Affairs	Academy of Management Journal
Journal of Consumer Psychology	Academy of Management Learning and Education
Journal of Interactive Marketing	Academy of Management Review
Journal of International Marketing	Administrative Science Quarterly
Journal of Macromarketing	Advances in Experimental Social Psychology
Journal of Marketing Management	Annual Review of Psychology
Journal of Public Policy and Marketing	Decision Sciences
Journal of Retailing and Consumer Services	Journal of Applied Psychology
Journal of Services Marketing	Journal of International Business Studies
Journal of Strategic Marketing	Journal of Management
Marketing Intelligence and Planning	Journal of Management Studies
Marketing Letters	Journal of Organizational Behavior
Marketing Theory	Journal of Personality and Social Psychology
Psychology and Marketing	Journal of Product Innovation Management
Quantitative Marketing and Economics	Management Science
Academy of Marketing Science Review	Organization Science
Advances in Consumer Research	Organization Studies
Asia Pacific Journal of Marketing and Logistics	Strategic Management Journal
Australasian Marketing Journal	The Leadership Quarterly
Organizational Research Methods	Personality and Social Psychology Bulletin
Psychological Bulletin	Academy of Management Annals
Personality and Social Psychology Review	Psychological Science

This page has been intentionally left blank

This page has been intentionally left blank

This page has been intentionally left blank

7. Guidelines for Preparing Assignment 2 (Market Research Report) (30%)

(a) Assignment 2 (Individual Task)

1. You MAY choose a topic from the following domains (indicative list).

Assignment Topics
1. Brand Innovativeness
2. Sponsorship
3. Celebrity Endorsement
4. Retailer equity
5. Country-Brand Relationships
6. Brand Identification

Assignment 2 (Market Research Project Report) Preparation Guidelines

- **Start work on the assignment early, in order to do well**

(b) The Task

This assignment involves the preparation of a market research project report for a client, based on one of the six scenarios suggested for Assignment 1. Unlike for Assignment 1, here in this assignment, you will need to **collect data** from other students in your class (assume they are your sample) and **analyse** the data converting it into **information**.

Mainly, the report should convince the client that you have used an appropriate research design and generated quality information useful for their decision making. The level of detail should be such that the client will be able to evaluate the findings included in the report.

(c) General Guidelines and Indicative Work Plan

Choose a topic towards Assignment 2 IDEALLY by the end of Teaching Week 7.

Note on Data collection

Note that students are not permitted to collect data from outside the class: either from students or general public. You will need permission from the UQBS Ethics Committee. You do not have the permission.

Therefore, **you can only collect data from your other class members enrolled in the course** and if you are compelled to do so from family members. Assume for the purpose of the assignment that the data have been collected from a representative sample of your target population.

You may use Qualtrics, if you prefer, for data collection.

Step 1: Select One Scenario from those Suggested (Provided separately)

- You should select the scenario you are most comfortable with or interested in. It is strongly advised that you continue with theme chosen in Assignment 1 in Assignment 2 as well.
 1. Brand Innovativeness
 2. Sponsorship
 3. Celebrity Endorsement
 4. Retailer equity

5. Country-Brand Relationships
6. Brand Identification

Step 2: Identify management and research problems, and research objectives

The decision maker described in the scenario is facing a **management problem**. Your task is to identify the management problem and state this in a question form. Ideally, the best from the different research proposals (from Assignment 1) will be taken for the purpose of this assignment.

Step 3: Apply the relevant concepts and principles and provide a research design, collect the data, analyse the data. Prepare the Report

(i) Structure

The report should include the following sections.

Table of Contents

Executive summary (1 page)

1. Background and Problem Definition

- Problem Statement & Research Objectives (What is the management/research problem?)
- Importance and Benefits
- Literature and Hypotheses

2. Research Design (What did you do?)

- Research method, sampling design, data collection, measurement

3. Data analysis and Results (What did you find?)

4. Conclusions, Recommendations and Limitations

(What is the information collected? What do the results mean? What is your recommendation?)

References

Appendix 1:

- Table 1: Conceptual definitions
- Table 2: List of measures

Appendix 2: (Not included in the word count)

- Measurement instrument (Questionnaire)

List of Tables (Not included in the word count)

(ii) Content

Executive summary

The executive summary should provide an overview of the entire report. Briefly discuss the nature and scope of the report. That is, indicate the type of study conducted (i.e., the broad type of research design), how the study was conducted, and tell the reader the most significant implications of the research.

- This section should provide an overview to a busy executive the nature and content of the report. If interested, the reader would look inside the report.
- Generally, this section will be of a full page.

1. Background and Problem Definition

This section would around 4 pages in length.

- This introductory section should provide an overview of the research report.
- Think about the issues that you will be discussing subsequently and in more detail (e.g., the problem statement, research objectives, literature, and importance of the study).
- Signal to the reader that you have a clearly defined problem, and one that is important and worth pursuing.
- Show that you are familiar with the literature, by integrating some key citations.

The **problem statement** section should provide the reader with some background information (e.g., about the firm and scenario you have selected for study), as well as clearly state the management problem

and the research problem (you will need to specify these yourself, based on one of the scenarios).

- Try to convince the reader that the problem is interesting and important, and that the firm would benefit if it were solved.
- The problem statement section sets the broad framework for the research proposal; that is, the objectives (research hypotheses), literature review, and later, the research design, should flow on logically from the problem statement.
- Be sure to identify the independent and dependent variables for explanatory studies, and the nature of their relationship (e.g., positive or negative, linear or curvilinear).

Importance and benefits of the study

This part would not be more than half a page in length.

Provide a clear statement of the importance of the study and the benefits expected. Think about both specific benefits and possible broader outcomes.

- A study could be important because it generates new insights into a problem or because a solution to the problem generates economic benefits or both!
 - For example, a market research problem could be connected to the marketing goals or help to evaluate alternative strategies for reaching those goals.
- More generally, solving a research problem might lead to superior effectiveness, customer satisfaction, market share, profit, brand loyalty and the like.

Support for your hypotheses using relevant Literature

This part would not be more than a couple of pages long.

- Provide clear **conceptual definitions** of the variables of interest.
- Identify and briefly describe **theoretical perspectives** authors have used in the past (e.g., the theories to which authors refer).

- Use the summary of literature to build a **case for your research hypotheses**. Put forward an argument, but be objective and present a balanced view.
 - Stating other researchers' findings, including both supporting and contradictory evidence is important.
 - Contradictory evidence or areas of disagreement might be interesting discussion points and suggest a need for further research.
- Later, in the Research Design section, you will also be required to specify measures or “**operational**” **definitions**.

2. Research Design

(Generally, this section may be around 6 pages in length)

This section one of the **key** sections of the research report. This section significantly affects your grade.

- Start this section by giving an indication of the **overall research design** (exploratory/descriptive/causal) and why it was appropriate. It is a big problem if this section is missing.
- Then, you need to describe and justify the **research method(s) used** (experiment/survey/observation). What was the specific experimental design used (if used)? What was the specific survey mode (if used)? What was the specific mode of observation (if used)?
- Specific design issues need to be addressed in this section, such as **sampling, data collection** and **measurement**.
- **Sampling:**
 - Clearly state the sources of data for the study.
 - Define the population for the study and describe how a sampling frame was developed. Justify the chosen probability/non-probability sample.
 - What was the sample size? Why?
- **Data collection:**
 - What were the procedures that would be used to collect the data?
 - How did you collect the data (communication/observation)? Why?

- How were the various types of error reduced?

Consider **ethical issues** here.

- The discussion on research ethics might include a reference to the researcher's responsibilities to: the client, research participants and researcher workers.
- Maintaining respondent anonymity and confidentiality is a crucial concern.
- Researchers also have responsibilities to their clients (e.g., professionalism, quality research).
- Think about the practices that might be help to maintain a researcher's responsibilities to clients and respondents (and research workers).
- **Measurement:**
 - Indicate how the variables included in your hypotheses were measured. [Hint: most good researchers use valid and reliable measures available in the literature.](#)
 - What were the sources (references) for the measures you used?
 - What type of scale were used?
 - How can you tell us that the measures used are reliable and valid?
 - How many items were used for measuring each construct?
- Examples of specific **operational definitions** of variables could be listed.

3. Data analysis and Results

This section would around 6 pages in length or more. This section significantly affects your grade.

In this crucial section, you need to specify the methods of analysis that were used for testing your research hypotheses.

Specific analytical methods should be described – how was each hypothesis tested? (Through correlation or regression analysis, for example) The reason you used a data analysis technique is because you need information from the data.

- How many pieces of information did you need? How many techniques did you need to generate the information required?
- Keep in mind that the analysis procedures should be consistent with the design, objectives and sampling methods employed.
- My expectation is that you will identify appropriate analytic techniques, briefly describe them and provide a rationale for why they were appropriate (Be sure to justify your choices!).
- In general, the scenarios are designed such that you would use all the data analysis techniques taught in the computer lab during the semester.

Then, give the reader specific and more general outcomes.

Specific outcomes might include statistical conclusions. You could provide statistical outcomes based on the analytic techniques you select for analysis. (e.g. regression model and results).

- More general outcomes should also be identified (e.g., a strategy for increasing customer loyalty, a new compensation and evaluation system).

4. Conclusions, Recommendations and Limitations

- Discuss the results. Provide opinions based on the results and suggestions for action.
- The conclusions and recommendations should be presented in this section in more detail than in the summary, and the text should include justification as needed.

References

- It is unlikely for a good proposal to have less than **10 references** to relevant sources (for the purpose of this course).
- Use author-date system. If you follow the referencing style used in any of the six lead articles is a good practice.
- It is unlikely for a good report not to have references to the prescribed text and/or something equivalent.

Appendix 1

Table 1: Conceptual definitions

- Provide definitions of the key concepts used in a Table including the original source and page numbers.

Table 2: List of Measures

- Provide a list of the measures and scales used, including the original source and page numbers.

Appendix 2: Measurement instrument

- Provide a copy of the measurement instrument (questionnaire) used for data collection

(d) Length, Weight, Due Date and Venue

- Word count: 3500-5000 words; **Weightage: 30%**
- Due: **Latest by 2 pm, Friday, 29 May 2020** (Teaching Week 13)
- Submit an electronic copy via the course Blackboard. Late submission will incur penalty.
- Note that you would lose marks if you exceed the suggested word count. Quality of analysis, not the length, of the proposal is important.
- References, headings, executive summary, table of contents, tables and appendices are not included in the word count.

(e) Assignment 2 (30%) - Research Report (Individual Task) - Marking Sheet

Criterion (Weightage)						
1. Presentation (10%) - Overall presentation, title page, table of contents - Clarity of expression, grammar, sentence structure, referencing style, word length						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
2. Problem definition (10%) - Purpose of study including research problem and objectives - Reporting of any background information obtained to refine the problem (i.e. secondary data, observation, depth interviews)						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
3. Research Design (20%) - Research Method - Questionnaire - Target population and sample						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
4. Data Analysis and Findings (40%) - Detailed presentation with supporting tables, figures, and graphs - Well structured, logical, and clear exposition of results						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
5. Conclusions, Recommendations, and Limitations (20%) - Quality of conclusions - Relevance of recommendations derived from the findings - Correspondence between research problem and objectives outlined						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7
Overall grade for the Report The research report is of professional standard, providing the decision maker with quality information, includes required details of the method undertaken so that the decision maker is able to evaluate the credibility of the findings.						
Poor 1	2	3	Satisfactory 4	Good 5	Very Good 6	Excellent 7

Legend

1 - Fails to demonstrate most or all of the basic requirements

2 - Demonstrates clear deficiencies in understanding and applying relevant concepts; communicates information or ideas in ways that are frequently incomplete or confusing and give little attention to the conventions of the discipline

3 - Demonstrates superficial or partial or faulty understanding of relevant concepts and limited ability to apply these concepts; presents undeveloped or inappropriate or unsupported arguments; communicates information or ideas with lack of clarity and inconsistent adherence to the conventions of the discipline

4 - Demonstrates adequate understanding and application of the relevant concepts; develops routine arguments or decisions and provides acceptable justification; communicates information and ideas adequately in terms of the conventions of the discipline:

5 - Demonstrates substantial understanding of relevant concepts and ability to apply these concepts in a variety of contexts; develops or adapts convincing arguments and provides coherent justification; communicates information and ideas clearly and fluently in terms of the conventions of the discipline:

6 - As for 5, with frequent evidence of originality in defining and analysing issues or problems and in creating solutions; uses a level, style and means of communication appropriate to the discipline and the audience:

7 - As for 6, with consistent evidence of substantial originality and insight in identifying, generating and communicating competing arguments, perspectives or problem solving approaches; critically evaluates problems, their solutions and implications:

Part A: General comments (what was done well and not so well):

.....

.....

.....

Part B:

Suggestions for improvement (what could have been done to move to the next higher grade):

.....

.....

.....

.....

.....

.....

.....

Marker:

Date:

Note: All the comments included above are only aimed at providing constructive feedback

This page has been intentionally left blank

This page has been intentionally left blank

8. Guidelines for Final Examination (40%) Preparation –

8. Final Examination (40%) - Preparation Guidelines

(a) Examination Structure

The final examination is worth 40 percent of your score in this course.

The examination is scored out of 100 marks:

- There will be four questions. Each question is worth 25 marks for a total of 100 marks. Answer ALL the questions.
- The four questions mostly test your “process” knowledge. These questions are based on research scenarios and the interpretation of statistical output, and have a *short essay* response format. There will also be some knowledge testing questions. These will require shorter answers.
- You may need a calculator in the exam.

(b) Examination Strategy

- The examination is two hours in length plus 10 minutes for perusal. A suggestion for the perusal time is to carefully read the scenario and interpretation questions. More generally, consider the following examination strategy:
- Spend approximately 25 minutes on each of the four questions (100 minutes in total). The remaining 20 minutes could be used for re-checking or editing your answers. Write about one to two pages for each answer (not more than 10 pages in total).
- Recall that the final examination is worth 40 percent of your score in this course.
 - This is your single best opportunity to score highly in this course!
 - Evidence indicates final examination performance is strongly correlated with the final grade in the course.

(c) Useful Tips

Before the exam:

- Focus on the topics highlighted in the lectures
- Practice computer lab questions
- Practice questions from past examination papers

During the exam:

- **Legibility is very important**
- Spend the time on each section judiciously: plan your answers -how much time on each question?
 - Do not spend too much time answering a single question.
- Quality of the analysis is important not how much you write.

This page has been intentionally left blank

Part II - Practicals

Workshop Timetable

Teaching Week	Topic
1	No workshop this week
2	1. Introduction to SPSS & Data Entry
3	2. Descriptive Statistics & Assignment 1 briefing
4	3. Correlations
5	4. Cross Tabulations
6	5. t-test (Assignment 1 due this week)
7	6. Recap & Assignment 2 briefing
8	7. Analysis of Variance
	Mid-semester break (No workshop this week)
9	8. Multiple Regression Analysis I
10	9. Multiple Regression Analysis II
11	10. Time for Student to work on their Research Project in the Computer Lab
12	11. Time for Student to work on their Research Project in the Computer Lab
13	12. Recap & Final exam briefing (Assignment 2 due this week)

Recommended

Field, A. (2017). Discovering Statistics Using IBM SPSS Statistics. 5th edn, Sage.

Useful Reading

Allen, Peter and Kellie, Bennett, and Heritage, Brody (2014), SPSS 22: A Practical Guide (3rd edition), Cengage Learning; Melbourne.

Workshop 1: Introduction to SPSS

This is a brief on how to use SPSS. The purpose of this chapter is to introduce you to SPSS. You are encouraged to actively access the interactive help of SPSS and ask your tutor questions if you get lost. **You are expected to spend time outside of class (i) familiarising yourself with SPSS and (ii) completing the homework exercises to complete Assignment 2 (Research Report) for assessment (30% of your MKTG7510 Grade).** SPSS is accessible in all computer labs in Building 39A and 35. Please check current times of access to the computer labs and plan your study time accordingly.

Learning Objectives

The following exercises introduce you to SPSS. Although some of you may have used SPSS previously, this workshop makes no such assumption. Those of you who are experienced SPSS users can treat this as a refresher.

Session Objectives

1. Learn how to create a new project
2. Understand the key components of SPSS project
3. Learn how to input data

1. How to create a new project?

- (1) Click on the 'Start' menu. Click on 'All Programs'.
- (2) You will find the folder 'Data Analysis and Stats Programs' - click it to open the folder.
- (3) Click on the 'IBM SPSS statistics' folder. Click on the 'IBM SPSS Statistics' icon.
- (4) Once SPSS is open, the screen (see Fig. 1) asks you what you would like to do: either New Dataset (to type in the data manually), or Open another file (to find an existing data file) (as circled below).

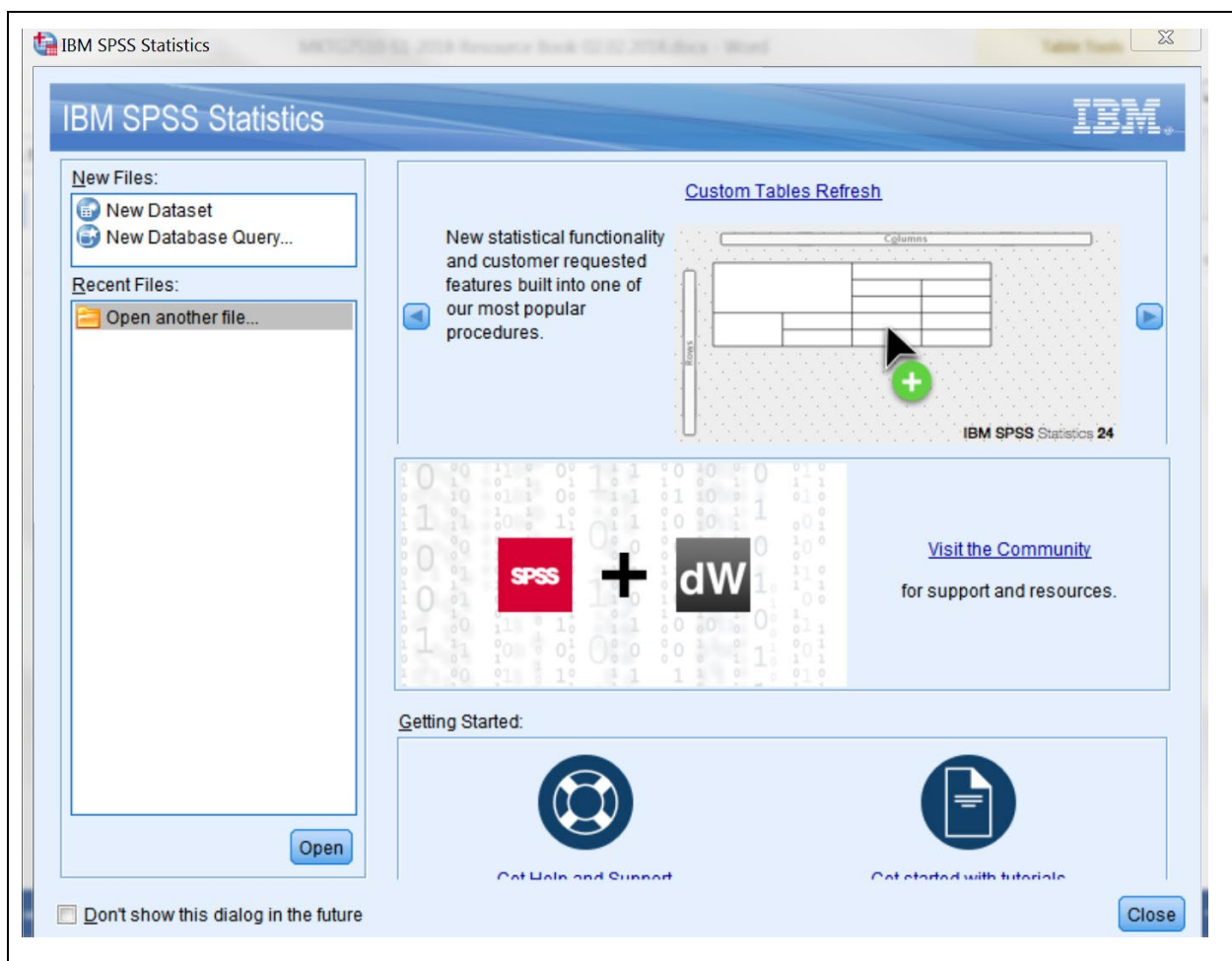
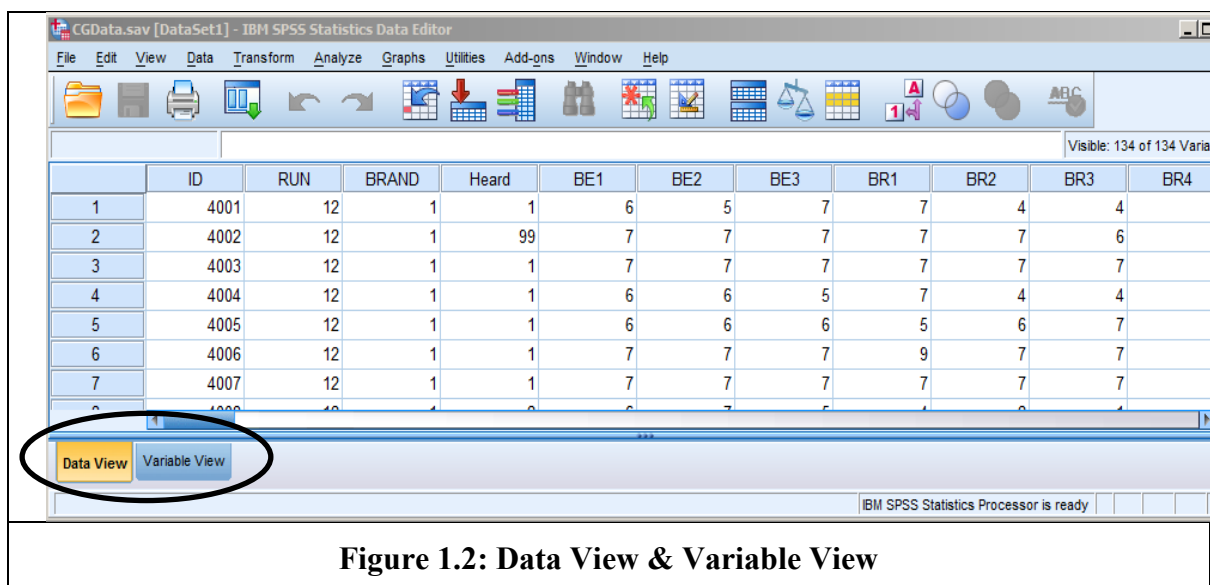


Figure 1.1: Create a New Project

2. What are the main components of an SPSS ‘project’?

When you open a new ‘project’, you will see the **Data Editor** window where the raw data are visible. However, there are multiple windows that allow you to interact with the data in different ways. The following is a list of the main components of a ‘project’.

- i. **Data Editor**: screen containing all the data and information for one study. You can have multiple worksheets open when working in SPSS. However, the **Data Editor** window only displays your current worksheet similar to a spread sheet format, with rows and columns that intersect to form individual cells. Although the **Data Editor** window has rows and columns, it is different from Microsoft Excel. In SPSS, you enter all data in a single worksheet.



The two screens in the **Data Editor** worksheet can be accessed by the tabs at the bottom left of the spread sheet (see Fig. 1.2): **Data View** and **Variable View**.

- ii. **Data View** is where raw data are entered into SPSS. Data are organised by Variable (in the columns) and Respondent (in the rows). Data are entered according to this structure so that SPSS can read variables and conduct data analysis correctly.
- iii. **Variable View** allows you to label each variable and provide information about the nature (e.g. nominal / ordinal etc.) of each of the variables. The information that can be specified for each of the variables is displayed in the columns. Each variable that takes up a column in the '**Data View**' sheet is now shown a row in the '**Variable View**' window.

Important!

Spend time setting up the data/variables correctly. Any errors affect the data analysis and lead to your analysis not working or being incorrect.

Results from all SPSS analysis commands are displayed in the **Output Window**, with the syntax commands for all SPSS actions you have taken. The syntax history is accessed in the left pane of the **Output Window** (Fig. 3), which logs actions chronologically. **Save your output** as a file for later use. You can create multiple output windows from each data file; however, SPSS only logs actions in the output window active at the time of the command.

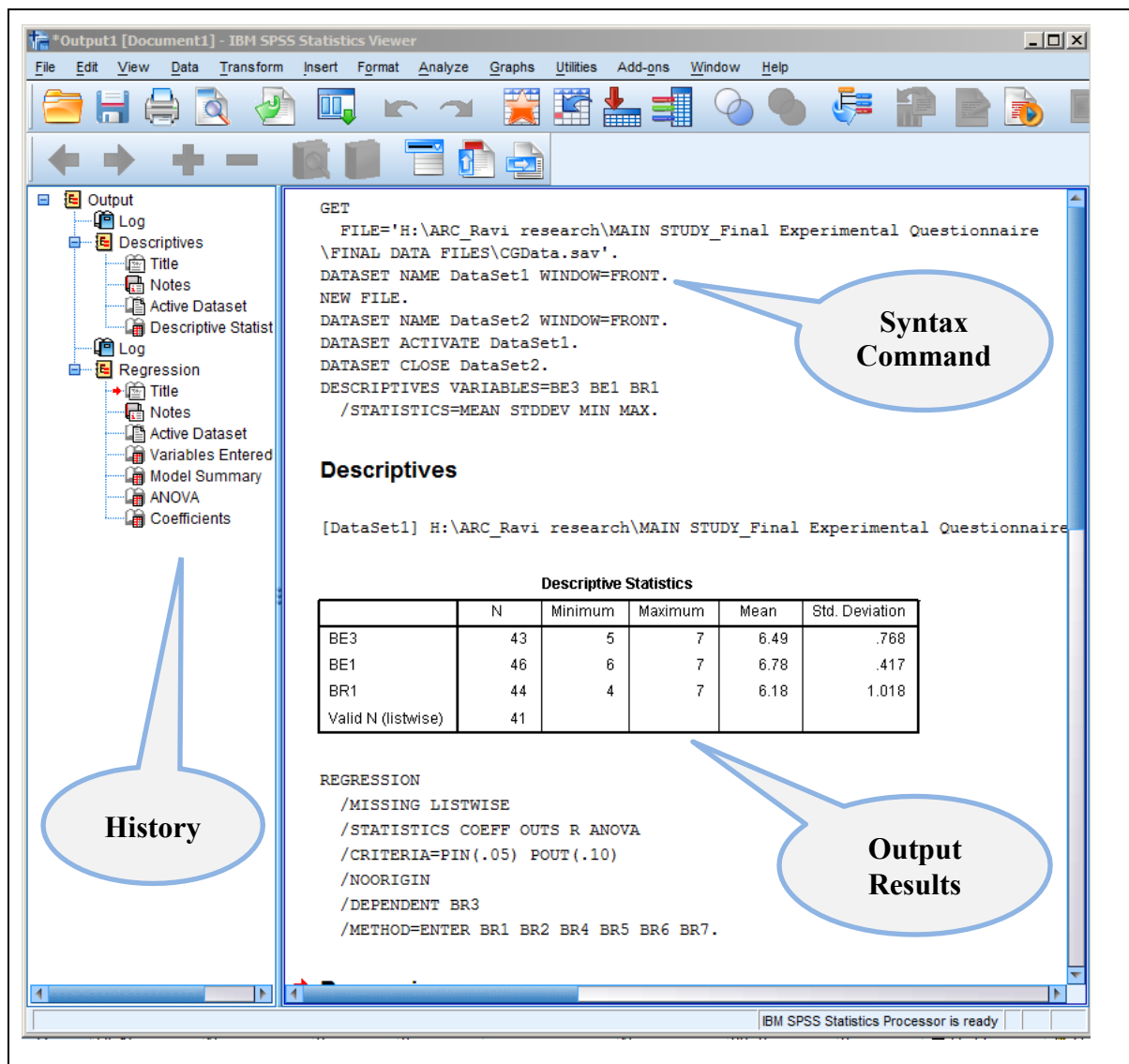


Figure 1.3: SPSS Statistics Output Viewer Window

3. How to open an already existing ‘project’?

Go to the “**File**” menu and click on the sub-menu “Open”, choose what type of file you would like to open (you will most often be opening a Data file or an Output file). A dialogue box appears, look in the folder where your project is located, select the project file and click on “Open”. Your project is opened for you! **Remember to select the correct file extension: .sav for a data file and .spv for output files.**

4. How to save a ‘project’?

- a) Go to the “File” menu, click on the sub-menu “Save as”. A dialog box appears. Give a name to the project, click on “Save” in the dialog box. The project is saved for you!

Remember the file name you have given to the project and the location where you saved it. Remember the data file extension will always be .sav.

- b) You can save an output window using that same procedure as above within the output window. Output files will have the file extension .spv.

5. How to print a ‘Window’?

Select the window that you want to print, go to the “File” Menu, select and click on the “Print” sub menu. You can print all of the text in a window or only a part of it.

- (1) Make sure the cursor is in the Session window.
- (2) Go to the **File** menu. Click on **Print**.
- (3) Choose **All pages** (or select a number of pages). Select **number of copies**. Click **OK**.

STEP 1: Create a new project

Before you create a new project file you need an actual project. Let's say you are interested in finding how many dates your friends have been on during the last year. You have conducted a small survey (see Fig. 1.4) to get an idea of frequency and quality of the dates to satisfy your interest. Ten of your friends participated in the survey.

Dating Survey Questionnaire									
1. How many dates have you had in the past year? _____									
2. How well did your last date go?									
Very bad Very good									
1	2	3	4	5	6	7	8	9	10
3. What is your annual income before tax? (\$) _____									
4. Are you male or female? (Circle one option) Male Female									

Figure 1.4: Dating Survey Questionnaire

First, you need to code each variable. You have 4 questions and therefore 4 variables. You also need a fifth variable to represent the respondents.

Provide a unique number to each of the respondents ($n = 10$). If you have 10 respondents it is a good idea to start numbering from 100. For example, use 101 to refer to the first respondent and 102 to refer to the second respondent and so on. Data received from your friends (respondents) is shown in Table 1.1.

Table 1.1: Data from Friend's Dating Survey

ID¹	How many dates have you had in the past year?	How well did your last go? 1 = very bad/ 10 = very good	What is your annual Income? (\$)	Are you male or female? 1 = Male 0 = Female
101	3	1	30000	0
102	3	10	32000	0
103	8	9	70000	0
104	1	8	25000	1
105	1	66	25000	1
106	1	2	25000	1
107 ²	3	.	35000	0
108	4	5	40000	0
109	2	7	35000	1
110	20	5	40000	1

Notes

¹It is important that you assign a unique ID number to each respondent

²You may use coding such as “99” or “.” for the missing values

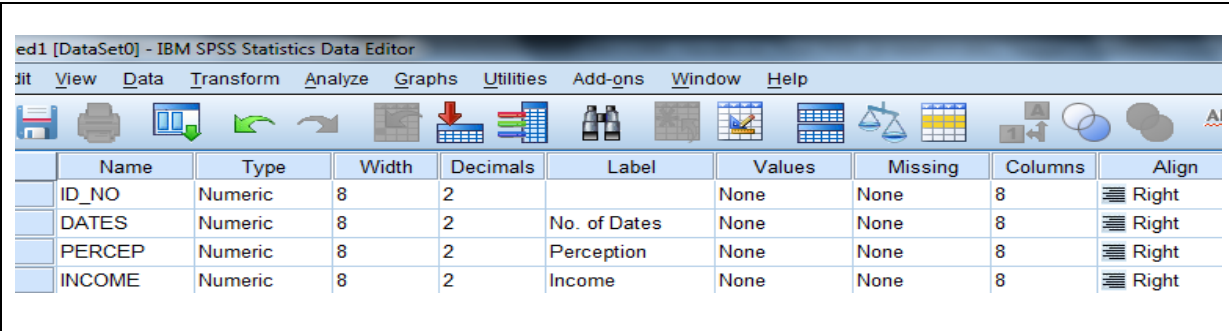
STEP 2: Structuring a data set and inputting data

Think about the structure you are going to require in the data. Click the **Variable View** (see Fig. 1.5) tab and create your variables. You will need (i) an ID variable to give each respondent a unique number and (ii) three question variables to mirror the four questions asked in the survey.

Generally, when naming variables you think about the key variable you are trying to measure and think of a code so you will remember what each variable is. Keep the variable name short. For example, we may shorten the questions using the following code:

- Question about the number of dates question shortened to **DATES**;
- Question about the perception of how well the date went to **PERCEP**;
- The income question to **INCOME**.
- Gender question to **GEN**

Figure 1.5 shows the **Variable View** window. Once you input a variable name, other details automatically fill themselves into the other columns in the Variable View window. You can change these details dependent on the characteristics of your variables. For this case we need to give each of the variables a **label** so that we can keep track of the questions asked. Give a label to each variable that describes what the variable is (i.e. the full text of the question in the survey).



The screenshot shows the IBM SPSS Statistics Variable View window. The window title is 'ed1 [DataSet0] - IBM SPSS Statistics Data Editor'. The menu bar includes 'Edit', 'View', 'Data', 'Transform', 'Analyze', 'Graphs', 'Utilities', 'Add-ons', 'Window', and 'Help'. The toolbar contains various icons for file operations, data manipulation, and analysis. The main area is a table with the following columns: Name, Type, Width, Decimals, Label, Values, Missing, Columns, and Align. The table contains four rows of variables: ID_NO (Numeric, Width 8, Decimals 2, Label 'ID_NO', Values 'None', Missing 'None', Columns 8, Align 'Right'), DATES (Numeric, Width 8, Decimals 2, Label 'No. of Dates', Values 'None', Missing 'None', Columns 8, Align 'Right'), PERCEP (Numeric, Width 8, Decimals 2, Label 'Perception', Values 'None', Missing 'None', Columns 8, Align 'Right'), and INCOME (Numeric, Width 8, Decimals 2, Label 'Income', Values 'None', Missing 'None', Columns 8, Align 'Right').

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
ID_NO	Numeric	8	2		None	None	8	Right
DATES	Numeric	8	2	No. of Dates	None	None	8	Right
PERCEP	Numeric	8	2	Perception	None	None	8	Right
INCOME	Numeric	8	2	Income	None	None	8	Right

Figure 1.5: SPSS Variable View

Variable View (see Fig. 1.6) provides more details about each of the variables. For example, let us say you want to describe the variable PERCEP. First select the variable PERCEP by selecting on the row.

- **Values:** If a variable is measured on a 10-point scale, then specify each value and its meaning. Click PERCEP and label each point (1 = very bad up to 10 = very good).
- **Measure:** Click on the Measure column, to specify the scale of the variables (e.g. categorical / scale etc.).
- **Missing:** It is important to specify what symbol/number is going to represent missing values within your data set so that SPSS can treat the value as missing rather than as a 0.

After describing all the variables, you need to switch to the **Data View** window (See Fig. 1.6) and input the data you've collected in the survey.

	ID_No	DATES	PERCEP	INCOME	var	var	var	var	var
	101	3	1	30000					
	102	3	10	32000					
	103	8	9	70000					
	104	1	8	25000					
	105	1	66	25000					
	106	1	2	25000					
	107	3	.	35000					
	108	4	5	40000					
	109	2	7	35000					

Figure 1.6: SPSS Data View

Workshop 2: Descriptive Statistics

How to prepare for this task?

- Read the related lecture material and book chapters.
- Read this brief on ‘Descriptive Statistics’ and understand the learning objectives.

Session Objectives

1. Learn how to generate descriptive statistics and interpret the SPSS output
2. Understand when to use descriptive statistics

STEP 1: Get to Know your Data Using Descriptive Statistics

Now you have set up the data in SPSS, you can start finding out information about your sample of friends. A good way to start ‘getting to know’ your data is to run descriptive statistics. Descriptive statistics is a global name for a range of general statistics drawn from a sample and include Mean, Standard Deviation and Minimum/Maximum. Let’s compute these statistics from the sample. See Fig. 2.1.

- (1) Go to the **Analyze** pull down menu. Select **Descriptive statistics > Descriptives**.
- (2) Select the variables that you need to analyse (all variables except ID) and move them to the Variable(s) box.

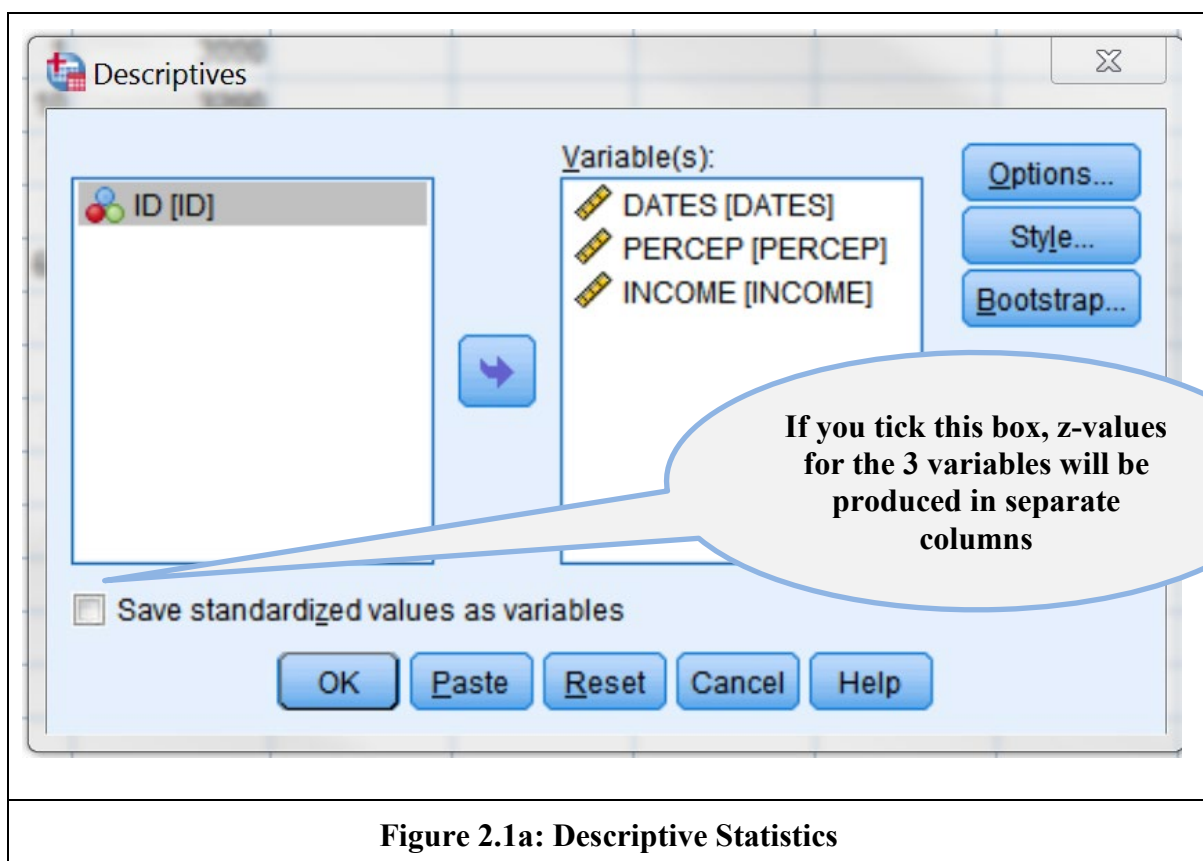


Figure 2.1a: Descriptive Statistics

- (3) Click on Options.

A separate box will open (see Fig. 2.2)

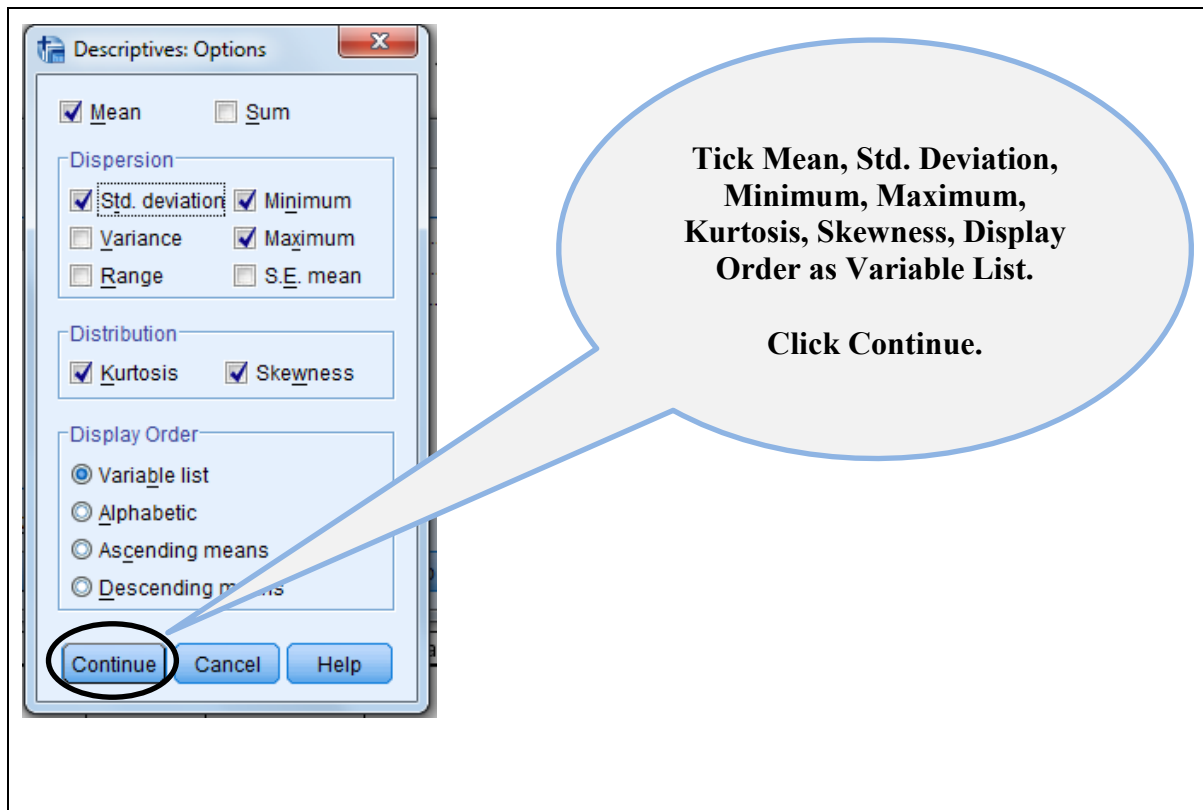


Figure 2.2: Options Window View in Descriptive Statistics Editor

- (4) Check that the boxes against all the statistics you want to compute are ticked.

- (5) Click continue. Click OK. The results will pop up in the Output screen (see Fig. 2.3).

Descriptive Statistics					
	N	Minimum	Maximum	Mean	
	Statistic	Statistic	Statistic	Statistic	
How many dates have you had in the past year?	10	1	20	4,60	
How well did your last go? 1=very bad/10=very good	9	1	66	12,56	
What is your annual income?	10	25000	70000	35700,00	
Valid N (listwise)	9				

	Statistics				
	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
How many dates have you had in the past year?	5,797	2,511	,687	6,670	1,334
How well did your last go? 1=very bad/10=very good	20,268	2,871	,717	8,440	1,400
What is your annual income?	13350,406	2,136	,687	5,454	1,334
Valid N (listwise)					

Figure 2.3: SPSS Output - Descriptive Statistics

How do these numbers (Fig. 2.3 Descriptive Statistics) help you get to know your data?

N The number of people who replied to the question (n = sample size)

Mean The mean is the average score calculated across the sample for a **continuous** variable. This can help you understand whether people on average are high or low on a variable or the frequency of an event across the population. It is the **sum** of all scores **divided** by the **number** of scores.

Std Deviation The standard deviation is a measure of how much variation occurs in the sample. The larger the standard deviation is, the more variation is in respondents responses to a variable. A smaller standard deviation suggests people are very similar in relation to a variable.

Min/Max The Minimum and Maximum scores provide information about the range in the data (how much of the scale was used). If the minimum and maximum are not the ends of the scale it may mean that none of the respondents held the extreme view or that design flaws may be present in the scale.

Skewness The measure of symmetry in the distribution of scores. **Skewed** distributions are **not symmetrical** and instead most frequent scores are clustered at one end of the scale.

Kurtosis The degree to which scores cluster at the ends of the distribution (known as the tails of the distribution) and how **pointy** a distribution is.

How to generate a histogram?

In a normal distribution kurtosis and skewness values are 0. It is very convenient to use **histograms** to examine the distribution of scores in the data. To produce a histogram in SPSS you may do the following.

- (1) Click **Graphs** in the SPSS menu and choose **Legacy Dialogs** → **Histogram**.
- (2) We might want to produce a histogram for our INCOME variable. Select INCOME and move it to the Variable box (see Fig. 2.4).

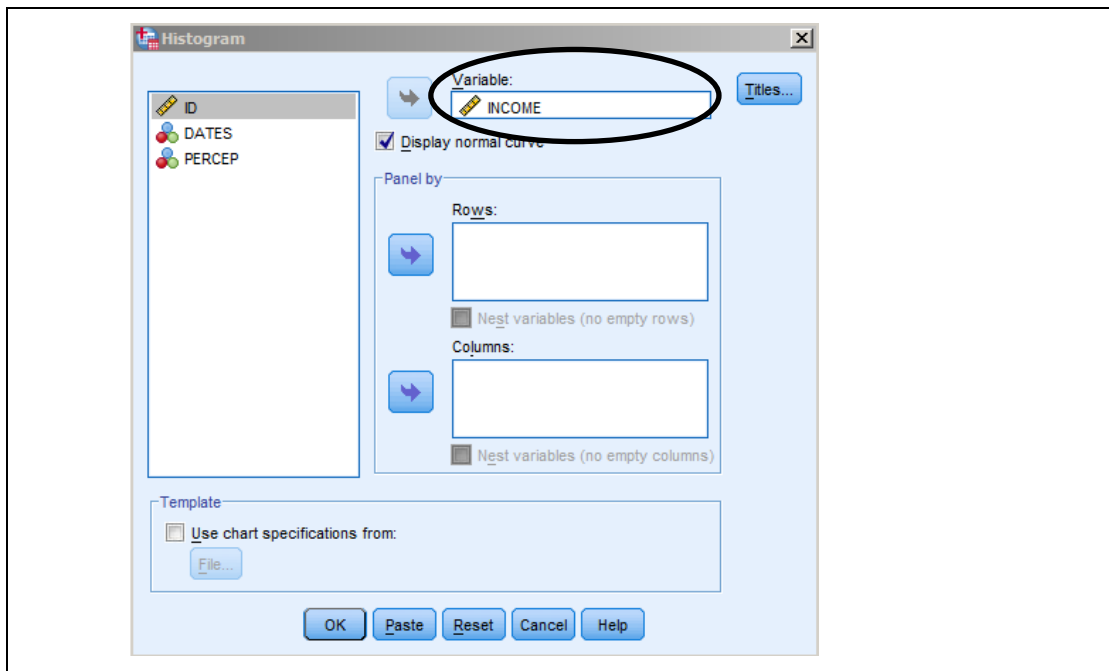


Figure 2.4: Histogram Window in the SPSS Graphs Menu

(3) We would like to see if the distribution of the variable INCOME is approximately normal. To do this we want to click the “Display normal curve” box before clicking OK. The following graph will appear in the Output window (see Fig. 2.5)

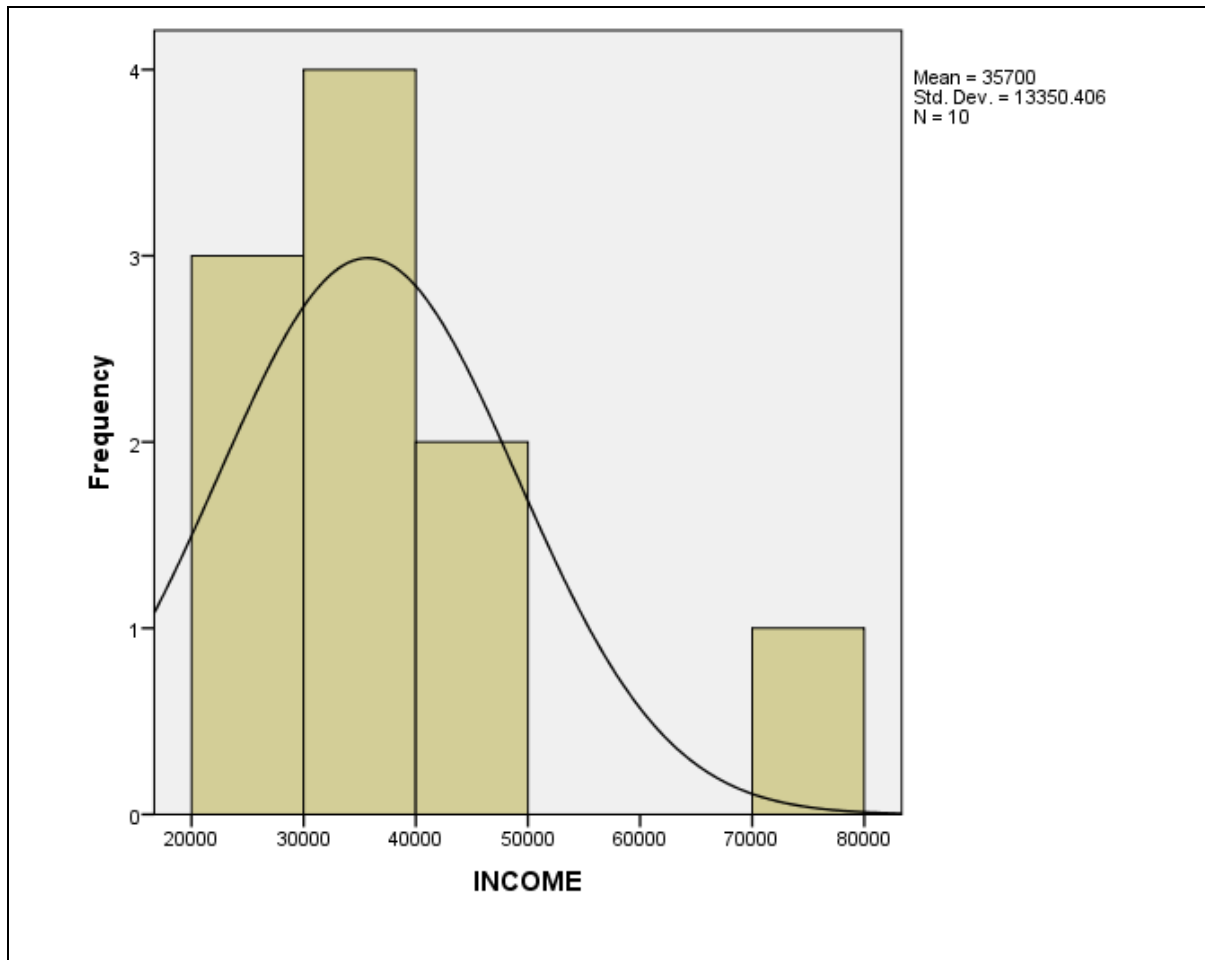


Figure 2.5: Histogram SPSS Output

Because we only have ten observations, the histogram (see Fig. 2.5) is not very “beautiful”, but you can still analyse the distribution based on it. The scores are positively skewed, since the frequencies are clustered in the lower range of the scale and the distribution tail points to the more positive end of the x-axis. The distribution has positive kurtosis, with more scores in the right tail than in the left tail.

STEP 2: Data Screening Using Descriptive Statistics

All of the above statistics are also very important to alert you to any issues within the data. Mistakes happen in the design of surveys and in the keying in of results. Did you pick up on the mistake in the data set?

Whether you look at the Mean, Standard Deviation or Minimum and Maximum scores it is clear to see that there has been a keying mistake in the PERCEP variable. When you look through the data, you can see that the respondents were asked to rate their perception on a scale of 1 to 10, meaning that the score of 66 was a typing error. Screening the data using descriptive statistics is a **vital step** to make sure that there are no errors in the data before proceeding to more complex analysis.

You might also notice that the tenth respondent had 20 dates in the past year, while all the others the maximum was 8. This extreme score is called an **outlier** because it is very different to the rest. You can spot outliers by producing a histogram or by just screening the data visually (applies for small data sets). It is very **important** to spot outliers as they bias the sample statistics such as **mean** and inflate the **standard deviation**.

Important!

- **Save data and output files to a USB or your drive. You will use it in future sessions.**

CLASS EXERCISE

- Discuss with the students around you what you can learn about the dating habits of your friends from these statistics.

HOMEWORK EXERCISE

- Fix the mistakes in the data sheet and rerun your descriptive statistics. You may also want to delete the response with the **outlier** score in the **DATES** variable and check how that change affected the mean and standard deviation of this variable.

Where can I use descriptive statistics?

- To understand the nature of data. To examine whether the data meet certain assumptions. Whatever research you may conduct you will be expected to provide descriptive statistics. That is, the mean, standard deviation, variance, skewness, kurtosis and range of the data.

Additional Resources

The following books are useful for developing a clear understanding of the theory underlying the data analysis techniques covered in the practicals.

1. Allen, Peter and Kellie, Bennett, and Heritage, Brody (2014), *SPSS 22: A Practical Guide* (3rd edition), Cengage Learning; Melbourne.
2. Field, Andy (2013). *Discovering Statistics Using IBM SPSS Statistics*, 4th edn, London: Sage.
3. Ho, Robert (2006). *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*, Boca Raton, FL: Taylor and Francis.
4. Howell, David C. (2013). *Statistical Methods for Psychology*, 8th edn, Belmont, CA: Thomson Wadsworth.

Symbols Frequently Used in Statistics

Symbol	Name	Meaning
α	Alpha	The probability of making Type I error
β	Beta	The probability of making Type II error
β_i	Beta	Standardized regression coefficient
χ^2	Kai Square	Chi-square statistic
F	F	Friedman's ANOVA test statistic
ε	Epsilon	Error
η^2	Eta-squared	An effect size measure
μ	Mu	The mean of a population scores
ρ	Rho	The correlation in population
σ^2	Sigma squared	The variance in a population data
σ	Sigma	The standard deviation in a population data
$\sigma_{\bar{x}}$	Sigma X bar	The standard error of the mean
τ	Tau	Kendall's tau (non-parametric correlation coefficient)
ω^2	Omega squared	An effect size measure

Source: Field, Andy (2013). Discovering Statistics Using SPSS, 4th edn, London: Sage

This page has been intentionally left blank

Workshop 3: Correlation Analysis

How to prepare for this task?

- Read the related lecture notes and book chapters.
- Read the Correlation Analysis tutorial and understand the learning objectives.
- Read through the Workshop Questions.

Learning Objectives

Now that you have a basic understanding of SPSS, this workshop will provide you with an opportunity to use SPSS to investigate some interesting problems.

Session Objectives

1. Become familiar with correlations and scatterplots.
2. Understand how to generate a scatterplot and interpret the results.
3. Understand how to generate a correlation matrix and interpret the results.
4. Understand the difference between association/correlation and causality.

Assumptions of Correlation Analysis

- **Pearson** correlation coefficient r requires that the **data are at least interval/ratio**;
- The relationship between the variables is **linear**; and
- Additionally, to establish that the correlation coefficient is significant we need to check that the bivariate distribution is **normally distributed**.

STEP 1: Problem and Data



Imagine you are working for the health insurance firm ‘Bupa’. You have been asked to do a survey of consumers’ attitudes towards health insurance. Specifically, your job is to evaluate if the Bupa Insurance adverts “Find a Healthier You” are positively influencing perceptions of Bupa.

You’ve done a telephone survey of 100 non-Bupa health insurance buying households and found 20 people that were home, remembered your Bupa adverts, purchased health insurance, and would answer your survey. All questions were answered on 1 to 100 scale, where 100 means very likely/satisfied/good. Following are the questions you asked the 20 households.

1. How likely are you to ring Bupa Insurance concerning your health cover needs?
[RING]
2. How satisfied are you with the rates offered by your current health insurance company? [CRATE]
3. How satisfied are you with the service offered by your current health insurance company? [CSAT]
4. In a few days, how well do you think you’ll remember the Bupa Insurance phone number? [MEMORY]
5. How well do you perceive the rates of Bupa health insurance? [RATES]
6. How well do you perceive the service of Bupa health insurance? [SERVICE]

See Table 2.2 for the BUPA Health Insurance Data.

STEP 2: Correlation Matrix

To examine the association between variables we can conduct correlation analysis (See Fig. 2.8).

- (1) Go to the **Analyze** pull down menu. Click on **Correlate** → **Bivariate...**
- (2) Click on all the variables included in the survey and click the arrow to move them into the **Variables** box. Click **OK**. An output window should pop up with the results.

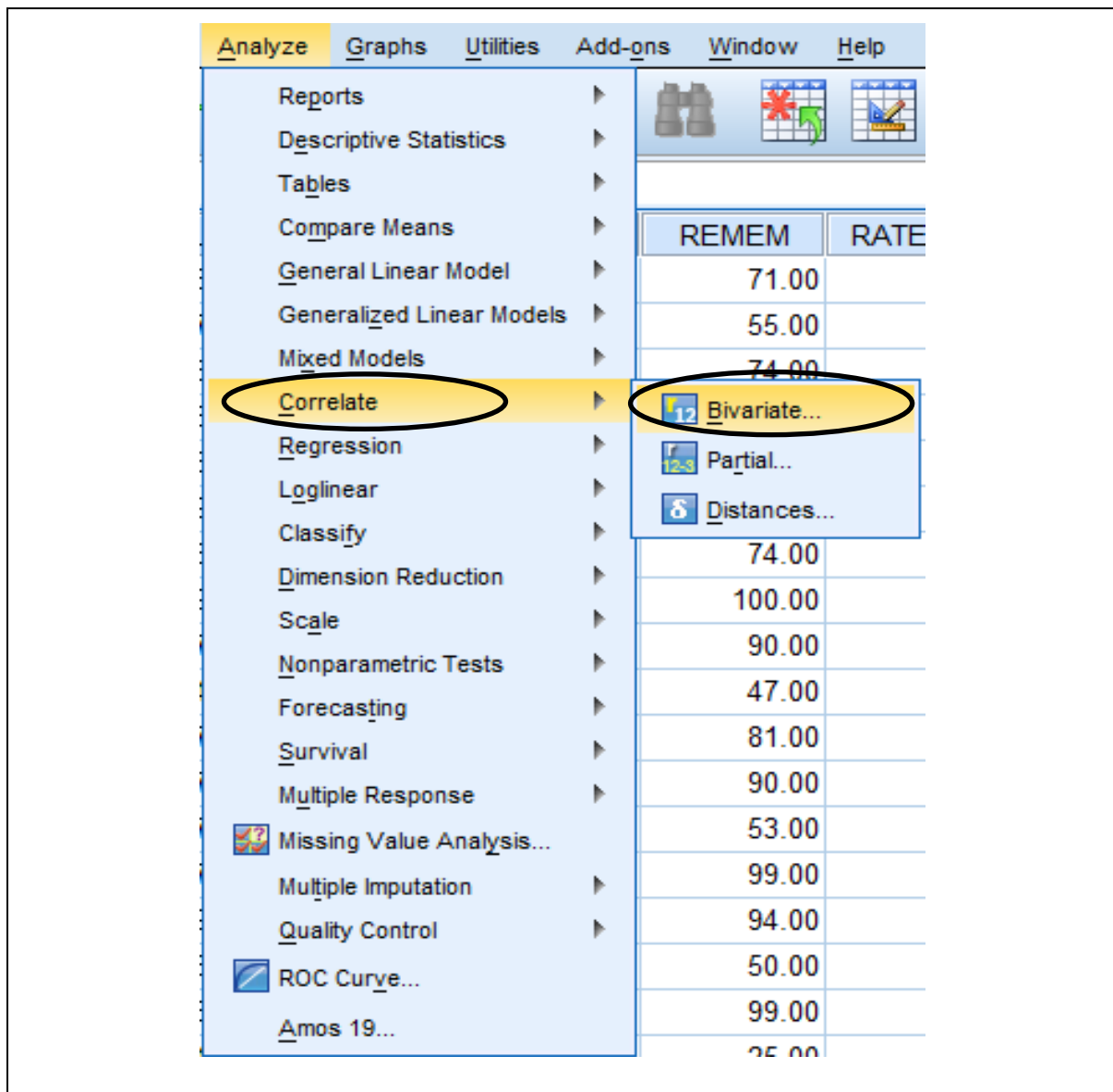


Figure 2.8: Analyze Menu - Bivariate Correlation Command

The output window should display a Correlation Matrix (Correlation Table) (see Fig. 2.9).

Correlations								
		RING	CRATE	CSAT	MEMORY	RATES	SERVICE	
RING	Pearson Correlation	1	-.056	-.822**	.869**	.207		.258
	Sig. (2-tailed)		.814	.000	.000	.380		.272
	N	20	20	20	20	20		20
CRATE	Pearson Correlation	-.056	1	.305	-.088	.050		.111
	Sig. (2-tailed)	.814		.191	.712	.835		.642
	N	20	20	20	20	20		20
CSAT	Pearson Correlation	-.822**	.305	1	-.732**	-.334		-.403
	Sig. (2-tailed)	.000	.191		.000	.150		.078
	N	20	20	20	20	20		20
MEMORY	Pearson Correlation	.869**	-.088	-.732**	1	.399		.497*
	Sig. (2-tailed)	.000	.712	.000		.081		.026
	N	20	20	20	20	20		20
RATES	Pearson Correlation	.207	.050	-.334	.399	1		.799**
	Sig. (2-tailed)	.380	.835	.150	.081			.000
	N	20	20	20	20	20		20
SERVICE	Pearson Correlation	.258	.111	-.403	.497*	.799**		1
	Sig. (2-tailed)	.272	.642	.078	.026	.000		
	N	20	20	20	20	20		20
**. Correlation is significant at the 0.01 level (2-tailed).								
*. Correlation is significant at the 0.05 level (2-tailed).								

Figure 2.9: SPSS Output Correlation Matrix

The correlation matrix includes information regarding (i) the correlation coefficients between each pair of variables, the significance of the correlation and (iii) how many observations are included in the results. It is important to mention here, that we calculated the **Pearson** correlation coefficient r , which is appropriate only for interval or ratio variables. If we have categorical variables, we need to use **Spearman's** correlation coefficient (**Spearman's rho**).

INTERPRETATION 1:

What does the correlation matrix tell you? Which of the variables in the collected data are associated with perceived service (SERVICE) of Bupa?

STEP 3: Scatterplots

For a visual representation of the level of association between variables we can use scatterplots.

- (1) Go to the **Graph** pull down menu. Click on **Legacy Dialogs** → **Scatter/Dot** (see Fig. 2.10).
- (2) Define a **Simple Scatter**. Choose a variable for the **Y Axis** and another variable for the **X Axis**.

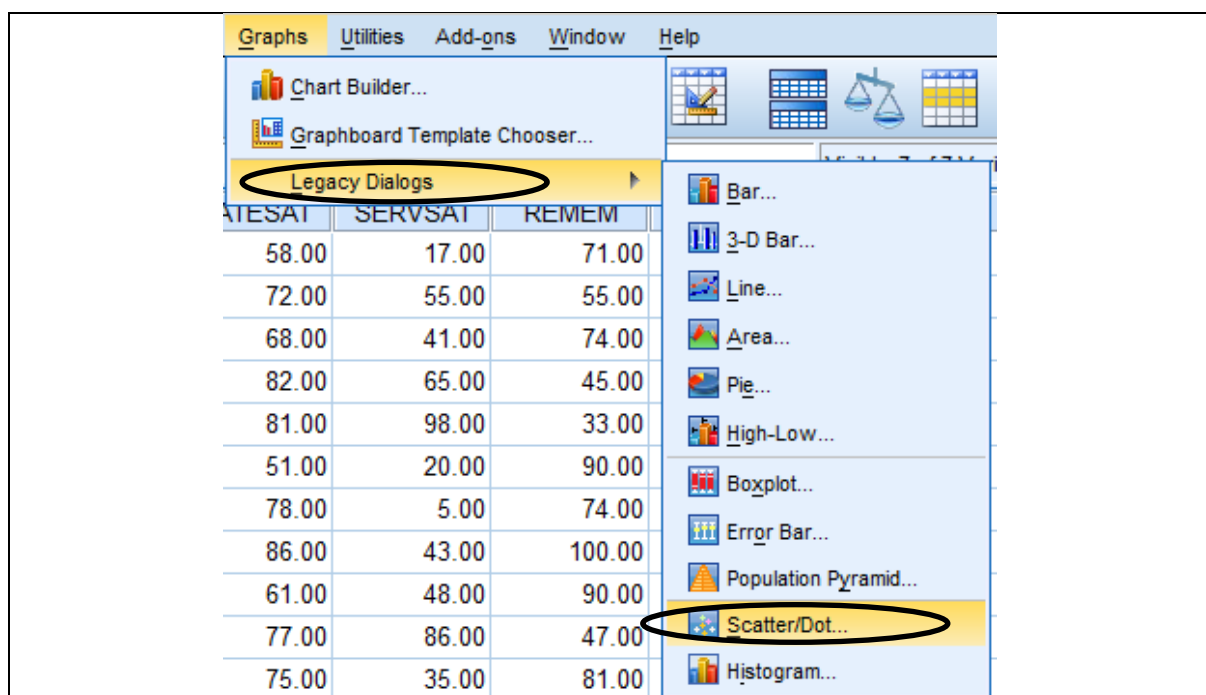


Figure 2.10: Graphs Menu - Legacy Dialogs Scatter/Dot Plot Command

CLASS EXERCISE:

Generate a scatter plot for two variables that were highly correlated in the results from Step 2. Complete a scatter plot for two variables that were NOT highly correlated.

INTERPRETATION 2:

What is the difference between the graph for two variables with a high association (correlation) and the graph for two variables with a low level of association (correlation)?

STEP 4: Interpreting Results

After you have a good idea of what is associated with what, we can now explore the concept of causality. When researchers speak of causation, they are often referring to a kind of prediction. Predictions can be considered to reflect cause only when all the relevant information is considered. We can never know all the relevant information, so our predictions are presumptive (or tentative).

We have to be careful when implying causation. While things that cause other things are associated with one another, **sometimes things are associated with one another do not have a causal relationship.**

First consider the relationships that have strong associations. Try to think about how one might cause the other. You may automatically think that X causes Y . However, a **correlation coefficient** of any magnitude or sign, regardless of statistical significance, **does not imply causation**. Several alternative explanations may be provided for correlation results. The following example uses the relationship between advertising and sales to demonstrate:

- X (advertising) causes Y (sales)
- Y (sales) causes X (advertising)
- X (advertising) and Y (sales) are both caused by another variable Z (profits)
- X (advertising) and Y (sales) have a reciprocal causal relationship

While no one can ever be certain that variable A causes variable B to occur, one can gather some evidence that increases the belief that A does, in fact, lead to B or that a dependence relationship exists. We seek four types of

evidence in support of causation (these are essential for making causal claims):

- Is there a predicted covariation between A and B?
- Is the time order of events moving in the hypothesized direction?
- Is it possible to eliminate other possible causes of B?
- Is there theoretical support for the predicted relationship?

Few studies in business (or the social sciences) achieve these four criteria. Typically, researchers provide evidence of predicted covariation (via some measure of association) and theoretical support (a rational explanation) in support of their causal claims.

STEP 5: Model Diagram

An important way to summarise how we think the variables fit together is using a “model diagram”. The basic idea is that we visually draw out what we propose causes what. **A relationship is represented with an arrow. Positive and negative relationships are represented with “+” and “-”.** Your job will be to do the same for the Bupa Insurance data based on the important relationships you’ve found. Remember your model diagram should include **all** the variables under study. Have a look at the example in the Figure.4.

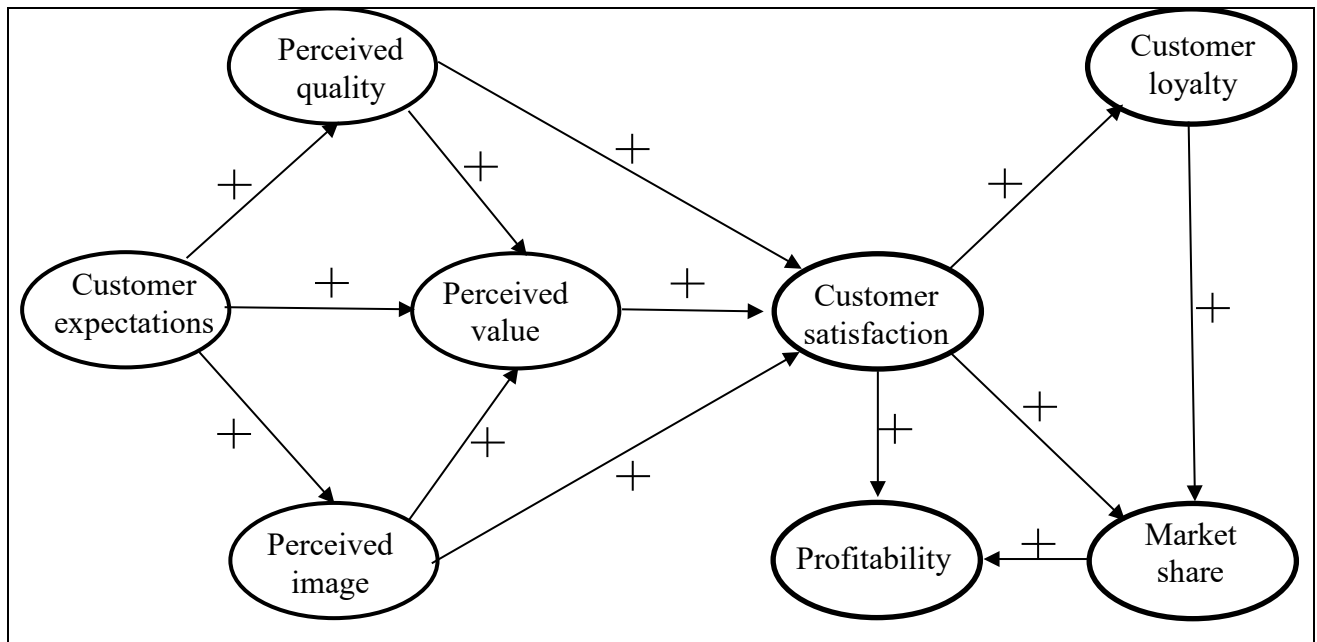


Figure 4: Example Model Diagram - Positive Causal Relationships

CLASS EXERCISE

1. Generate a correlation matrix;
2. Complete a scatterplot for a set of 'highly associated' variables and a set of 'not highly' associated variables; and
3. Draw a model diagram representing the results from the study.

Use the results to answer the questions following Table 2.2

.

Table 2.2: BUPA Health Insurance Data

ID	RING	CRATE	CSAT	MEMORY	RATES	SERVICE
107	81	58	17	71	52	37
122	72	72	55	55	62	49
126	82	68	41	74	50	33
131	52	82	65	45	67	81
134	51	81	98	33	66	60
140	90	51	20	90	68	55
142	87	78	5	74	61	83
143	92	86	43	100	75	80
144	72	61	48	90	59	86
149	40	77	86	47	62	40
158	78	75	35	81	57	75
161	74	67	23	90	87	98
165	71	80	43	53	58	42
172	79	68	23	99	86	100
174	93	69	40	94	46	35
194	60	66	50	50	47	49
196	88	75	26	99	70	85
197	44	64	73	25	48	23
198	45	55	60	40	67	75
199	91	63	1	85	89	90

Where can I use Correlation Analysis?

- If you want to find whether two variables are related to each other.
 - Use Pearson's correlation coefficient only if the **data are at least interval/ratio; There are other types of correlation coefficients, say if two variables are categorical.**
- If the relationship between the variables is **linear**; and
- If the bivariate distribution is **normally distributed**

Important!

Save the data and output files to a USB or your drive. You will be using it in future sessions.

HOMEWORK EXERCISE

1. Which two pairs of variables were the most associated? How can you tell?
2. Using the scatterplot explain the relationship:
 - a) between two variables that are highly associated and
 - b) between two variables that are **not** highly associated.
3. Draw a model diagram for the Bupa Insurance data reflecting the results of the study.

This page has been intentionally left blank

Workshop 4: Cross Tabulations

How to prepare for this practical?

- Read the related lecture notes and book chapters.
- Read this brief on ‘Cross Tabulations’ and understand the learning objectives.
- Read through the questions included towards the end of this note.

Learning Objectives

Cross tabulations are very commonly applied procedures used to summarise discrete information. Cross tabulations can be used to understand association between categorical variables such as respondent characteristics (e.g. location and gender) as well as responses to questions (particularly yes/no questions).

Session Objectives

1. Learn how to input crosstab data;
2. Learn how to correct and re-code variables using the pull down menus; and
3. Learn how to analyse data via cross-tabulations and interpret the output. You should be able to tell what the numbers mean.

Cross Tabulations: Assumptions

- Variables are categorical (nominal/ordinal – discrete variables).
- Observations are independent.
- None of the cells has a value of zero.
- Less than 20% of the cells have values below 5.

Let's say that you do a survey of 25 people passing through Queen Street Mall. You ask four questions:

1. Have you seen a 'Woolworths' advertisement in the last month? Yes - No
2. Do you think you will go shopping at 'Woolworths' in the next month? Yes - No
3. Do you like the Toyota car advertisements? Yes - No
4. Will you buy a Toyota car this year? Yes – No

The answers are listed in the Table 3.1 at the end of this tutorial. Five people became tired of your questions and would not finish your survey. Open SPSS, click on 'Type in data' to get to a blank **Data Editor** screen.

Save your file onto a USB or your student drive so you can access the file later. Remember to save the project as you go along so that you don't lose any of your actions. At the end of the exercise you should save your output file so that you don't need to redo your analysis every time.

STEP 2: Structuring a Data Set and Inputting Data

Think about the structure you are going to require in the data. Click the **Variable View** (see Fig. 3.1) tab and create your variables. You will need a **Person** variable to give each respondent a number and four question variables to mirror the four questions that were asked in the survey.

You will notice that once you have input a variable name, other details automatically fill themselves into the other columns. You can change these details depending on the characteristics of your variables. For this case we need to give each of the variables a **label** so that we can keep track of the

questions asked. Give a label to each variable that describes what the variable is (i.e. the question in the survey).

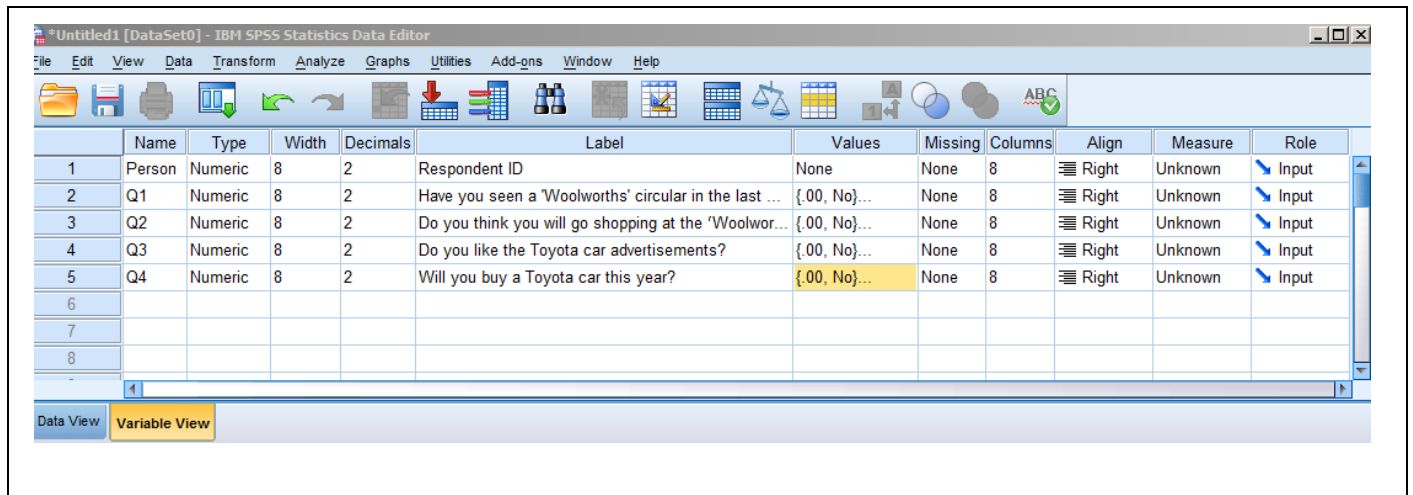
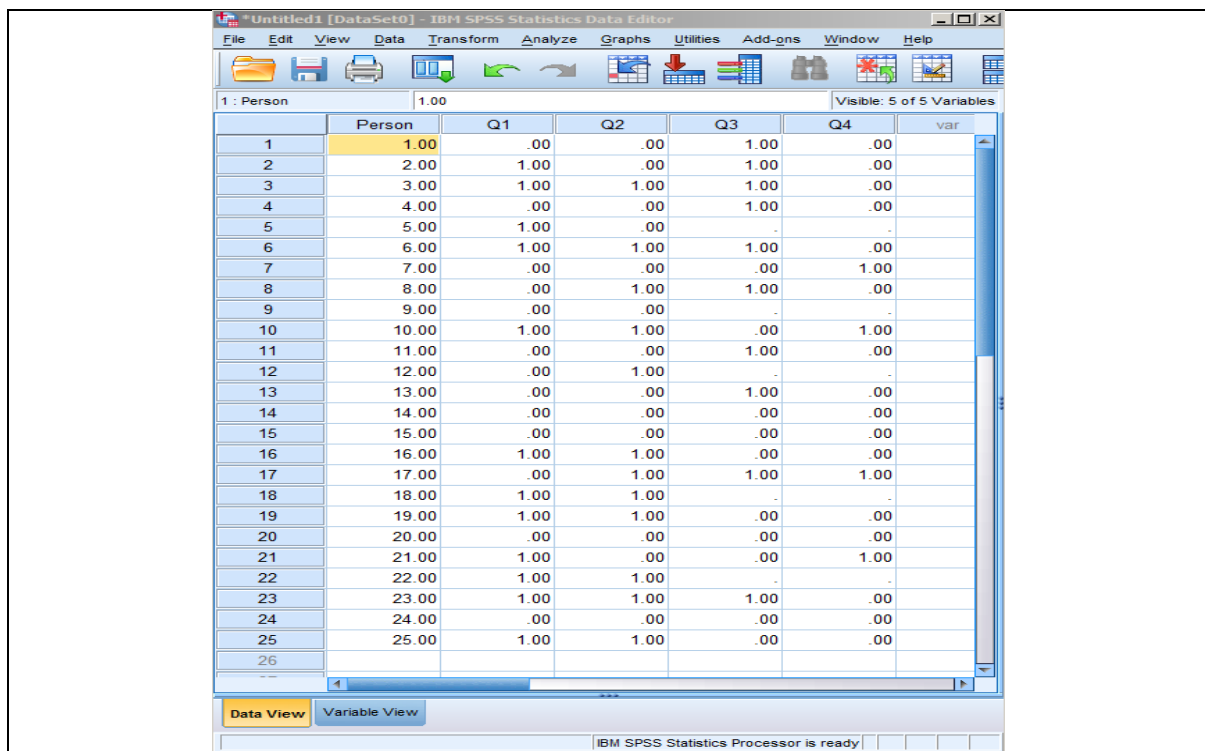


Figure 3.1: SPSS Variable View - Add Values and Labels

The data are in ‘Yes’ /‘No’ format rather than a numerical format. To conduct analysis, the data needs to be transformed so that ‘Yes’ and ‘No’ are represented by numbers. As per Fig. 3.1, the “Type” column has “Numeric” in it, indicating numerical variables. If you try to enter words or letters for “Numeric” variables, SPSS will not let you.

To enter text data you need to change the Type of variable to “String”. In the **Variable View** there is a column called “Values”. In this column you can assign values to the data, or ‘Code’ the data. Code a “Yes” answer as a “1” and a “No” answer as a “0”. Leave all other columns in the **Variable View** unchanged.

Go to the **Data View** (see Fig. 3.2) and begin to type in the data. Remember to type in the codes (1 and 0) rather than Yes and No. **Leave the blanks as blanks.** Missing values are automatically coded as “.” in SPSS. There should be five columns filled with the data from the survey.



The screenshot shows the IBM SPSS Statistics Data Editor window. The 'Data View' tab is active, displaying a dataset with 26 rows and 6 columns. The columns are labeled 'Person', 'Q1', 'Q2', 'Q3', 'Q4', and 'var'. The 'Person' column contains numerical IDs from 1 to 26. The 'Q1' through 'Q4' columns contain binary values (0.00 or 1.00) representing 'Yes' and 'No' responses. The 'var' column is currently empty. The status bar at the bottom indicates 'IBM SPSS Statistics Processor is ready'.

	Person	Q1	Q2	Q3	Q4	var
1	1.00	.00	.00	1.00	.00	
2	2.00	1.00	.00	1.00	.00	
3	3.00	1.00	1.00	1.00	.00	
4	4.00	.00	.00	1.00	.00	
5	5.00	1.00	.00	-	-	
6	6.00	1.00	1.00	1.00	.00	
7	7.00	.00	.00	.00	1.00	
8	8.00	.00	1.00	1.00	.00	
9	9.00	.00	.00	-	-	
10	10.00	1.00	1.00	.00	1.00	
11	11.00	.00	.00	1.00	.00	
12	12.00	.00	1.00	-	-	
13	13.00	.00	.00	1.00	.00	
14	14.00	.00	.00	.00	.00	
15	15.00	.00	.00	.00	.00	
16	16.00	1.00	1.00	.00	.00	
17	17.00	.00	1.00	1.00	1.00	
18	18.00	1.00	1.00	-	-	
19	19.00	1.00	1.00	.00	.00	
20	20.00	.00	.00	.00	.00	
21	21.00	1.00	.00	.00	1.00	
22	22.00	1.00	1.00	-	-	
23	23.00	1.00	1.00	1.00	.00	
24	24.00	.00	.00	.00	.00	
25	25.00	1.00	1.00	.00	.00	
26						

Figure 3.2: Recoding ‘String’ Data From ‘Yes’ & ‘No’ to Numeric Data 0 & 1

STEP 3: Analyse Data

Now the data are typed in we can start our analysis. The first analysis is a CROSSTAB of Questions 1 and 2.

- (1) Go to the **Analyze** pull down menu. Select **Descriptive Statistics** → **Crosstabs...**(see Fig. 3.3).

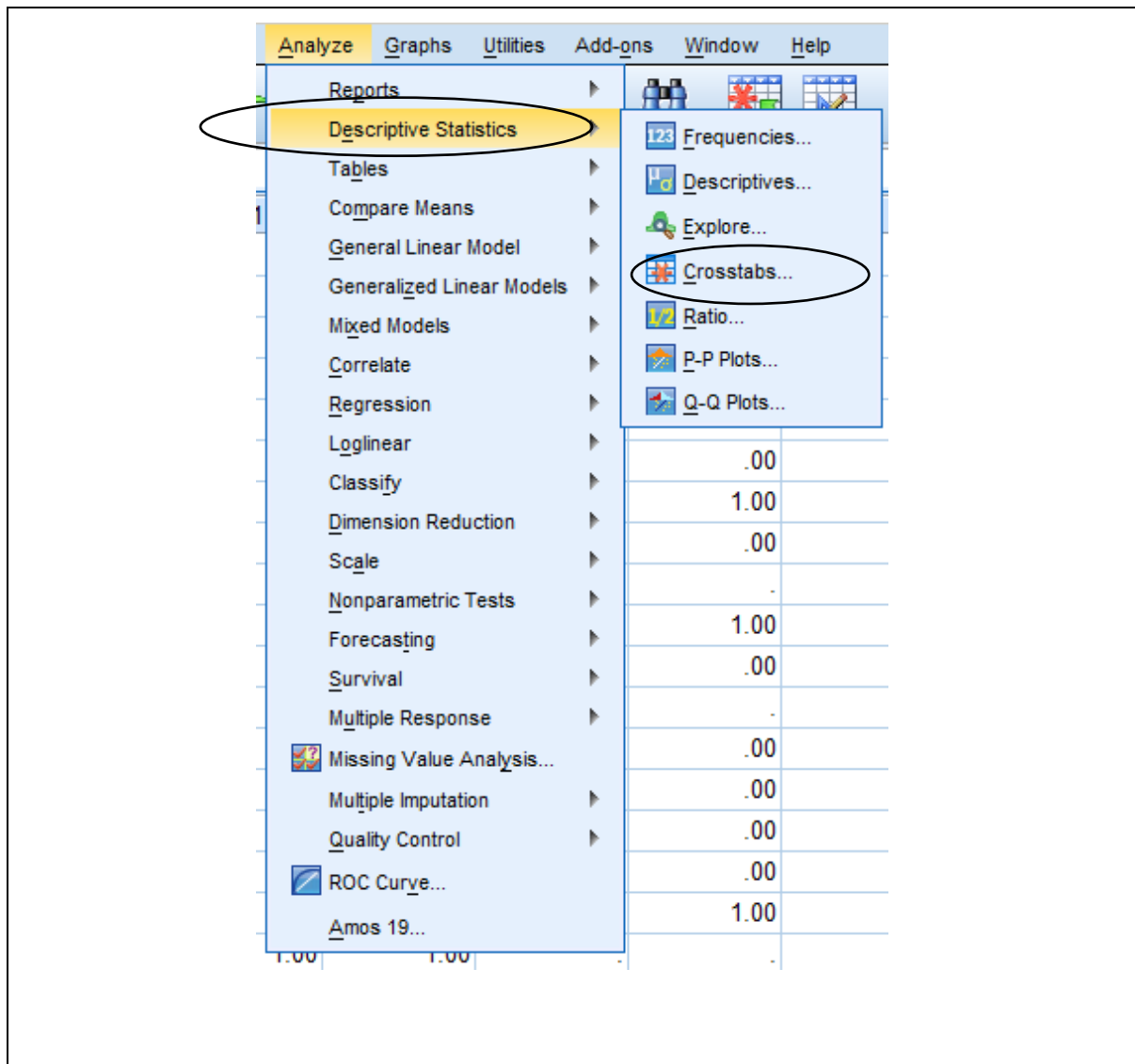


Figure 3.3: Cross Tabulations Command

- (2) This will open the **Crosstabs** dialogue box (see Fig. 3.4). Select a row variable (Q1) and click on the arrow button to move it across. Select a column variable (Q2) and click on the arrow button to move it across.
- (3) Click on the button 'Cells'. A radio box will open.
- (4) Check the box 'Row' under Percentages. Click **OK**.

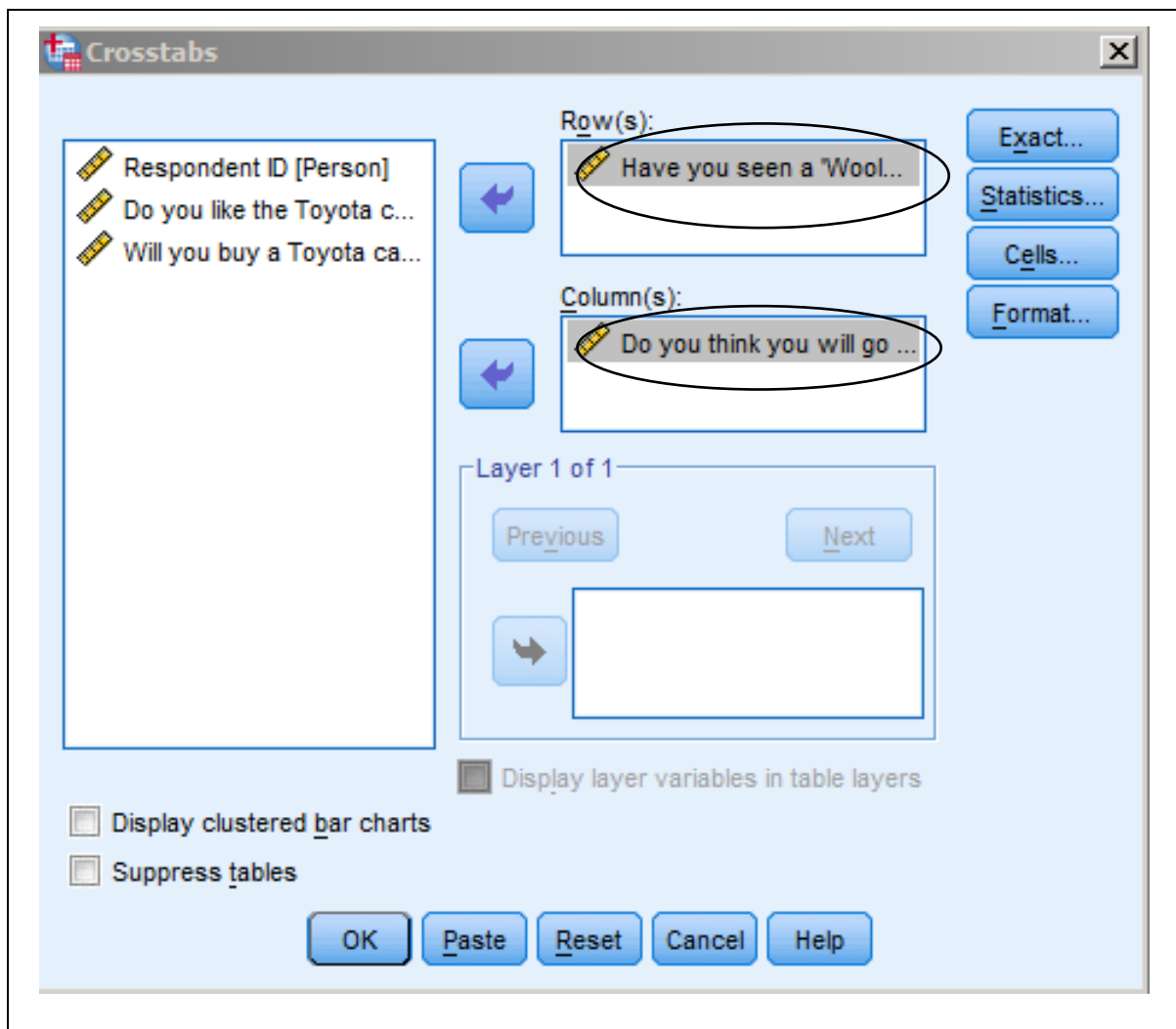


Figure 3.4: Cross Tabulations Dialogue Box

Once you have clicked OK an Output screen should appear showing the command syntax and your results. A **Crosstab** output has two tables (see Fig. 3.5). The **Case Processing Summary** table contains information about how many respondents answered the question (therefore how many data points were included in the analysis). Your tables will also have row percentages.

The second table is the **Cross Tabulation** results. This contains ‘counts’ for how many respondents answered each combination of ‘Yes’ and ‘No’ across the two questions.

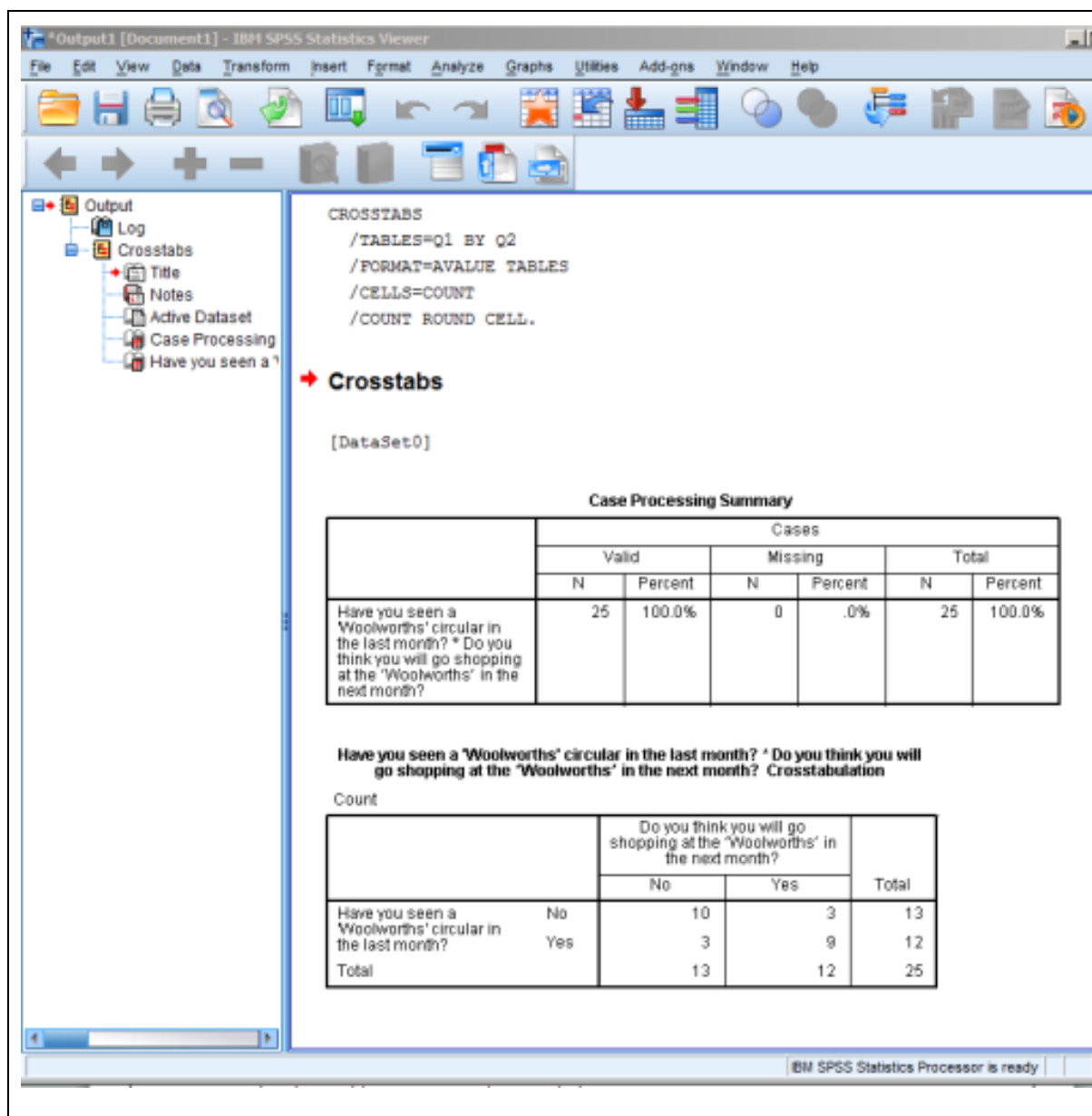


Figure 3.2: SPSS Cross Tabulations Output

INTERPRETATION

What do these results mean? Is there an association between seeing the Woolworths circular and intention to shop at Woolworths?

Note: Keep in mind that nothing in the results output will tell you what the numbers mean - you have to do it on your own.

CLASS EXERCISE

Complete a CROSSTAB for Questions 3 and 4.

STEP 4: Making Corrections to the Data and Recoding

Correction:

Before you go any further, you've noticed an error in the way that you inputted the data. For person #15, you noticed that the person put an X over "No" four times. You took that to mean that the person answered "No" four times. After checking with a colleague, you became convinced that they really meant, "Yes" four times and therefore your data needs correcting.

Go ahead and correct this data in the Data Editor.

Recoding:

Although we've inputted 0s and 1s for No and Yes respectively, you have been told by your boss that they want Questions 3 and 4 (Columns 4 & 5 in Table 3.1) to be coded as 1s and 2s for No's and Yes's. There is a way to do this without having to manually recode the variables.

- (1) Go to the **Transform** pull down menu (see Fig. 3.6). Select **Recode into Same Variables**.

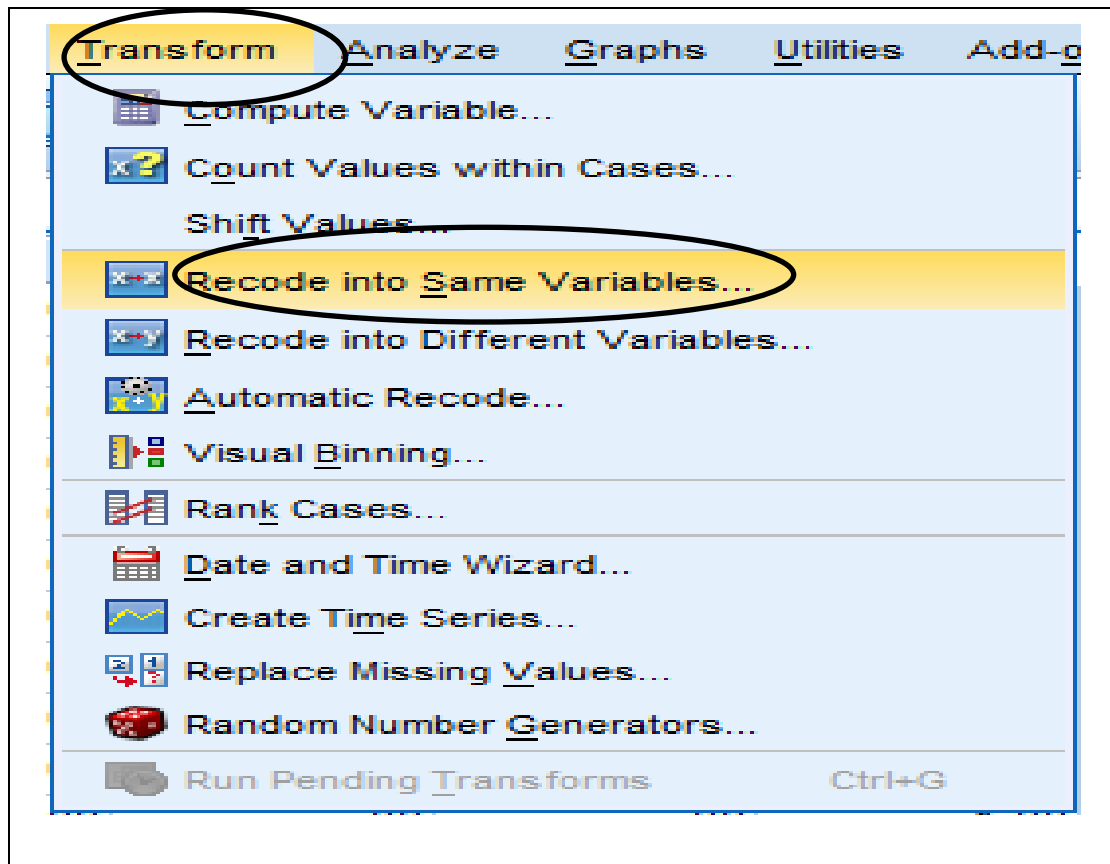


Figure 3.6: Transform Menu - Recode into Same Variables Command

- (2) Select the variables that you wish to recode and use the arrow to bring them into the **Numeric Variables** column.
- (3) Click **Old and New Values**. To change No from 0 to a 1, input 0 in the Old Value section and 1 in the New Value section. Click **Add** to submit the recoding instruction.
- (4) Repeat to change Yes from 1 to 2. After the codes appear in the Old → New box, click **Continue** and then **OK**.

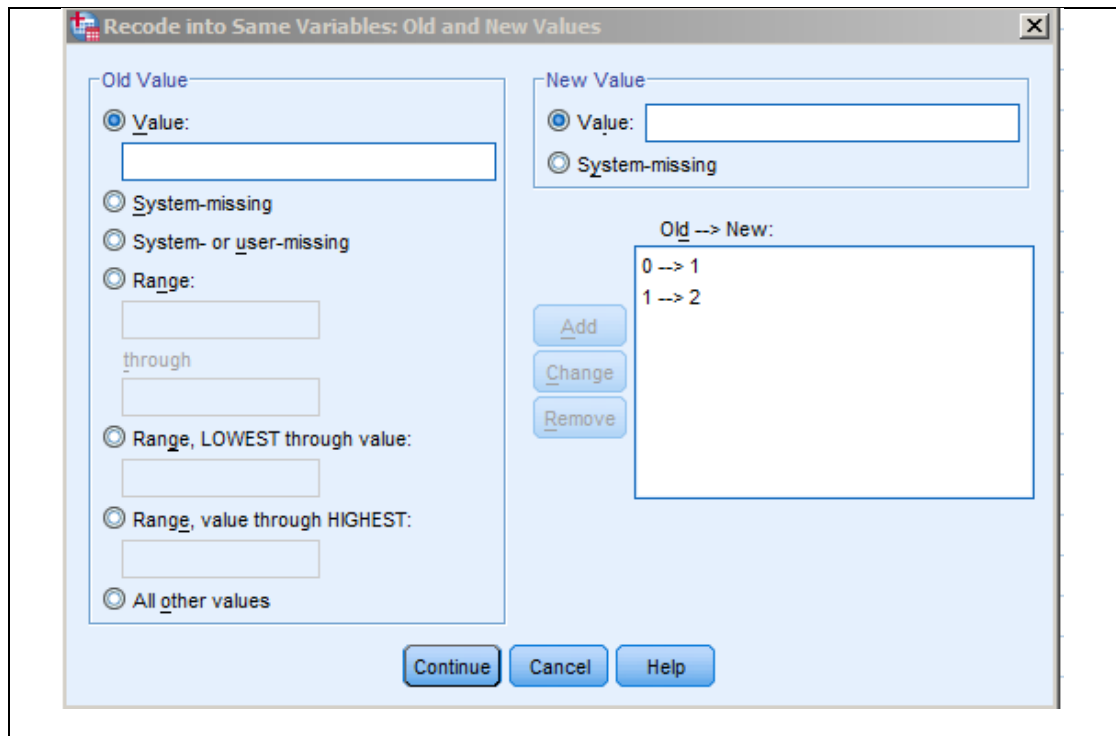


Figure 3.7: Recode into Same Variables - Old and new Values Command

- (5) An output window will appear with the command syntax. Go back to your **Data Editor** window and check that the recoding has been completed.

CLASS EXERCISES

After the data in Table 3.1 has been corrected and recoded, complete the following.

1. Complete a CROSSTAB for Q1 and Q2;
2. Complete a CROSSTAB for Q3 and Q4;
3. Complete a CROSSTAB for Q1 and Q4; and
4. Complete a CROSSTAB for Q2 and Q3.

Use the results to answer the following questions.

1. Interpret cross-tabulations of Questions 1 and 2. Comment on the nature of relationship between the two variables. Is it association or causality? Justify your answer.
2. Interpret cross-tabulations of Questions 3 and 4. Comment on the nature of relationship between the two variables.
3. Generate a cross-tabulation between Questions 1 and 4.
 - Comment on the nature of relationship between the two variables.
4. Generate a cross-tabulation between Questions 3 and 2.
 - Comment on the nature of relationship between the two variables.
5. Has the correction of data changed **the relationships** (not the numbers) between the variables Q1 and Q2? How? How did the re-coding of the data change the relationships (not the numbers) between the variables Q1 and Q2?

Table 3.1: Woolworth's & Toyota Advertisement Data (Before Correction)

Person	Q 1	Q 2	Q 3	Q 4
1	No	No	Yes	No
2	Yes	No	Yes	No
3	Yes	Yes	Yes	No
4	No	No	Yes	No
5	Yes	No	----	----
6	Yes	Yes	Yes	No
7	No	No	No	Yes
8	No	Yes	Yes	No
9	No	No	----	----
10	Yes	Yes	No	Yes
11	No	No	Yes	No
12	No	Yes	----	----
13	No	No	Yes	No
14	No	No	No	No
15	No	No	No	No
16	Yes	Yes	No	No
17	No	Yes	Yes	Yes
18	Yes	Yes	----	----
19	Yes	Yes	No	No
20	No	No	No	No
21	Yes	No	No	Yes
22	Yes	Yes	----	----
23	Yes	Yes	Yes	No
24	No	No	No	No
25	Yes	Yes	No	No

This page has been intentionally left blank

Workshop 5: t-test

How to prepare for this practical?

- Read the related lecture notes and book chapters.
- Read this note on the '*t*-test' and understand the learning objectives.
- Read through the questions included towards the end of this note.

Learning Objectives

This week's practical will introduce you to *t*-test analysis, when it is appropriate to use it and the two main types of *t*-test that you will be using; the independent samples *t*-test, and the dependent samples *t*-test (also known as a paired samples *t*-test).

Session Objectives

1. Become familiar with the concept of *t*-test;
2. Learn about the assumptions of *t*-tests; and
3. Practise at running two types of *t*-tests.
4. Learn how to conduct a *t*-test and interpret the results.

Assumptions of t-tests

- The sampling distribution is normally distributed. For *paired-samples t*-test the **difference** between scores should be normally distributed;
- Data are measured at least on an interval level;
- Additionally, for the independent-samples *t*-test, you need to check that:
 - Variances in the groups are roughly equal (homogeneity of variance);
 - The scores in the groups are independent.

STEP 1: Problem background

Apple is going to launch a new version of an iPod called iPod “ZEPTO”¹ and wants to estimate if there would be any difference in willingness to purchase the device between men and women, so the company can accordingly develop its distribution and marketing strategy and possibly the design of the mp3 player.

They asked the consumers to rate their willingness to purchase a new iPod on a scale from 1 (absolutely not willing to purchase) to 75 (definitely willing to purchase); the data are presented in **Table 4.1**. Let us call this study “Willingness to purchase a ZEPTO”.

Apple also wants to test whether a pilot advertisement can influence the level of preference for the product. To do that, the company first measured the preference for the product before consumers watched the advertisement, then made them watch the ad and collected their preference for the iPod straight after the ad was shown.

The data from this experiment is in that **Table 4.2**. We will call this study “Preference for ZEPTO”.

¹ Fun Fact: Did you know that Zepto is a prefix in a metric system meaning 10^{-21} ?

STEP 2: Two Types of t-tests

i. Study 1

In the “Willingness to purchase a ZEPTO” study Apple collected data from two different groups (men and women) on one variable: their willingness to purchase the new product. Because the company used **two different or independent** groups of people to collect the data on **one variable**, we will use an **independent-samples t -test** to estimate whether there is any difference in the willingness to purchase ZEPTO, between men and women.

ii. Study 2

In the second study, “Preference for ZEPTO”, Apple collected the data on one variable (preference for the product) from the **same group of people** but in **two different conditions**: before watching the advertisement and after watching the advertisement. To measure the difference between preferences for ZEPTO in this study we will use **paired-samples t -test (dependent--samples t -test)**, because we sampled the same group of people.

STEP 3: Checking the Assumptions of t-tests

Remember to check the following assumptions for *t*-tests?

- The sampling distribution should be normally distributed. For *paired-samples t*-test the **difference** between scores should be normally distributed.
 - **Kolmogorov-Smirnov (K-S)** and **Shapiro-Wilk (S-K)** tests or P-P Plots are used to check that;
- Data are measured at least on interval level.

Additionally, for the independent-samples *t*-test, you need to check that:

- Variances in the groups (in our case: men and women) are roughly equal (homogeneity of variance). We use **Levene's test** to do that;
- The scores in the groups are independent.

i. Normality of distribution

You may use P-P plots to visually estimate the normality of distribution. We can also use Kolmogorov-Smirnov and Shapiro-Wilk tests to check the normality of distribution. They compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation. If the tests are **not statistically significant**, it tells us that our variable distribution is not significantly different from a normal distribution.

To check data for the independent-samples t-test (study “Willingness to purchase a ZEPTO”):

- (1) Go to **Analyze** → **Descriptive Statistics** → **Explore**
- (2) Drag the variable WILLINGNESS into **Dependent list**
- (3) Drag the variable GROUP into **Factor list**

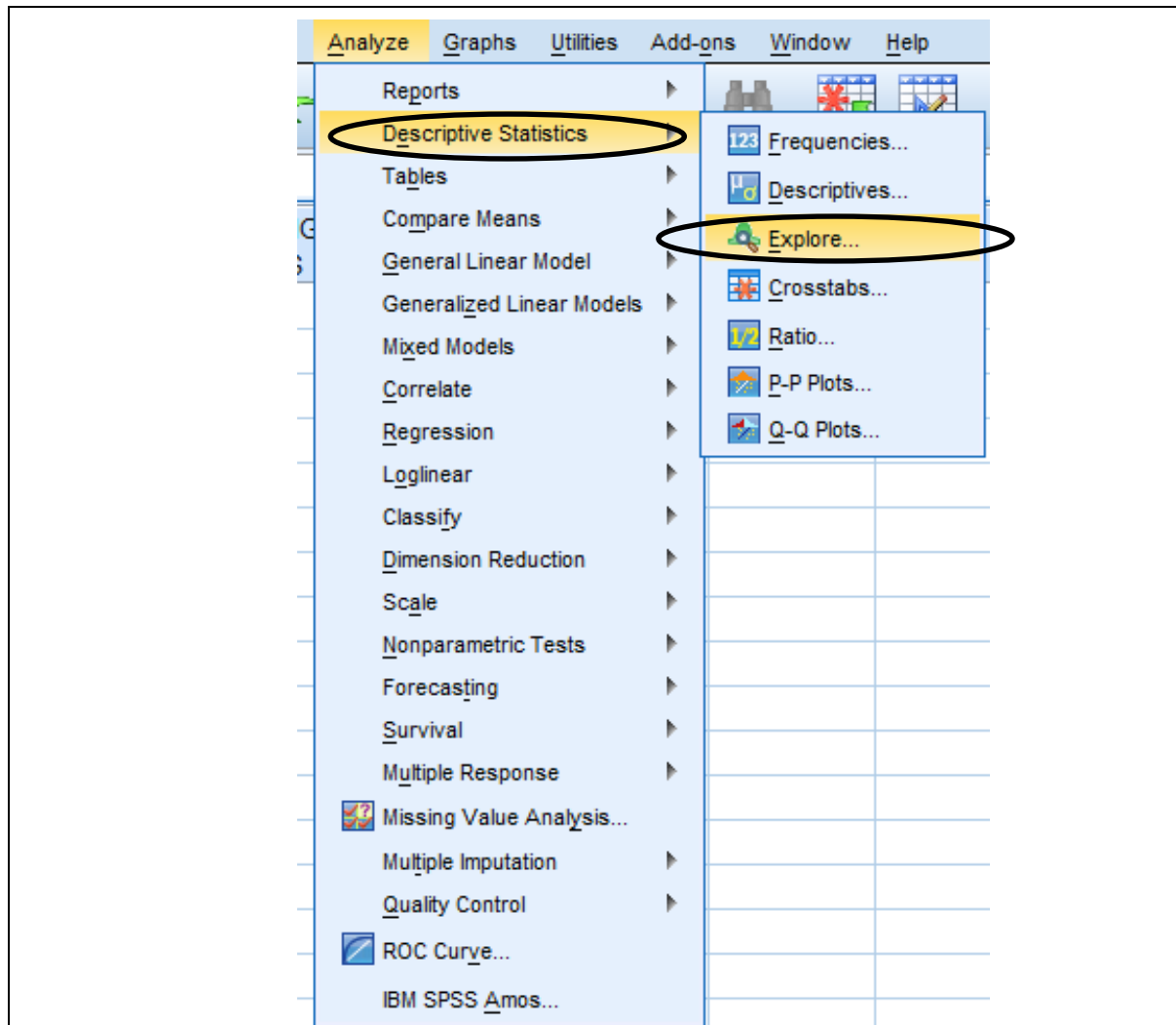


Figure 4.1: How to Produce Descriptive Statistics

- (1) Click on **Statistics** menu and choose “**Normality tests with plots**” and Click **Continue**
- (2) Click **OK**. Scroll to the “**Tests of Normality**” in your **Output window** (see Fig. 4.2).

Tests of Normality						
GROUP	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
WILLINGNESS 0	.116	15	.200 [*]	.944	15	.433
1	.151	15	.200 [*]	.958	15	.654

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 4.2: SPSS Output - Tests of Normality

CLASS EXERCISE

Is the distribution of the variable “Willingness to purchase a ZEPTO” normal?

Now, check the normality of distribution for the second study, bearing in mind that it is a dependent-samples study. Keep in mind also, that you will not need to drag anything into the Factor list in SPSS; you will only have one column of data.

ii. Homogeneity of variance

As introduced to you, we can use Levene’s test to check the assumption of homogeneity. Go back to remind yourself of the interpretation of the results of the test. We will not conduct a separated Levene’s test here. Levene’s test statistic is produced together with the procedure of actual *t*-test, however make sure you do not forget to as for it.

STEP 4: Running the t-tests

i. Independent-samples *t*-test

Let's run an independent-samples *t*-test for the “Willingness to purchase a ZEPTO” study.

- (1) Go to **Analyze** pull down menu. Choose **Compare Means**→**Independent-Samples *t*-test...**
- (2) Drag the variable WILLINGNESS to the **Test Variables** and the variable GROUP to the **Grouping Variable**.
- (3) Click **Define Groups**. Set **Group 1: 0**, **Group 2: 1** (as we coded male as 0 and female as 1 in our data sample). Click **Continue**.
- (4) Click **OK**.

The results will appear in the “Independent Samples t-test” table. First **check Levene's Test for Equality of Variances** column.

Independent Samples Test									
		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
WILLINGNESS	Equal variances assumed	.960	.336	4.746	28	.000	4.933	1.040	2.804 7.063
	Equal variances not assumed			4.746	27.303	.000	4.933	1.040	2.801 7.065

Figure 4.3: SPSS Output – Results of Independent Samples *t*-test

INTERPRETATION

Do the data violate the assumption of homogeneity of variance?

After checking the homogeneity of variance assumption, you can move to analysing the results of the *t*-test. If the *t*-test is statistically significant, it means that there is a statistically significant difference between the means of our two groups. In terms of our study, that would show that there are significant differences in the willingness to purchase ZEPTO between men and women.

ii. Paired-samples t -test

Now, conduct a paired-samples t -test. To do this, you need to perform the following steps:

- (1) Go to the **Analyze** pull down menu. Choose
 - **Compare Means**→**Paired-Samples t -test...**
- (2) Drag PREF_1 under **Variable 1** and PREF_2 under **Variable 2**.
- (3) Click **OK**.

“The table “Paired Sample t -test” shows you the most important results of the test. SPSS calculates the probability that a value of t statistic as big as the one obtained could occur if the null hypothesis was true (i.e. there was no difference between the two means). In our case, the probability of the result being due to chance is very small. Such a significant result means you can be confident to reject the null hypothesis. The positive sign of the t -statistic tells us that the mean of the group on the first occasion (before watching the ad) was smaller than the mean of the group on the second occasion (after watching the ad). In our study that would indicate that the preference for ZEPTO increased after watching the advertisement.

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PREF2	39.82	22	17.77	3.789
	PREF1	30.23	22	17.90	3.818

Paired Samples Test								
	Paired Differences			95% CI for the Difference in Means				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1 PREF2 - PREF1	9.59	1.79	.381	8.79	10.38	25.12	21	<.001

Figure 4.4: SPSS Output – Results of Paired Samples *t*-test.

Table 4.1: Willingness to Purchase a Zepto

INDEPENDENT		
ID	WILL	GROUP
201	67	0
202	69	0
203	71	0
204	75	0
205	69	0
206	72	0
207	68	0
208	71	0
209	73	0
210	74	0
211	75	0
212	66	0
213	70	0
214	73	0
215	66	0
216	66	1
217	66	1
218	66	1
219	66	1
220	69	1
221	67	1
222	67	1
223	70	1
224	65	1
225	65	1
226	63	1
227	63	1
228	69	1
229	62	1
230	61	1

Table 4.2: Preference for Zepto

DEPENDENT		
ID	BEFORE (PREF_1)	AFTER (PREF_2)
101	18	28
102	37	48
103	12	24
104	42	53
105	7	20
106	31	40
107	59	69
108	21	33
109	8	17
110	56	64
111	20	30
112	39	50
113	14	22
114	44	53
115	9	19
116	33	41
117	61	72
118	23	33
119	10	17
120	58	65
121	22	28
122	41	50

Where can I use t -tests?

- If you want to examine differences in means between two groups.
- If you want to test if a group mean differs from a standard value.
- If you have two independent groups (independent samples t -test)
- If you have two dependent groups (dependent samples t -test)
- Your independent variable does not have more than two levels.

Important!

- **Save the data and output files to a USB or your drive. You will be using it in future sessions.**

HOMEWORK EXERCISE

1. Two types of t -tests were conducted on the data. Which type is the most correct to use given the experimental design in each of the two ZEPTO scenarios and why?
2. Provide an interpretation of the paired sample t -test results. What would you tell Apple regarding gender differences in preference for ZEPTO? What could you tell Apple about the effectiveness of their advertising campaign?

This page has been intentionally left blank

Workshop 6: Analysis of Variance

How to prepare for this practical?

- Read the related lecture notes and book chapters.
- Read this note on 'Analysis of Variance' and understand the learning objectives.
- Read through the questions included at the end of this note.

Learning Objective

How to test for the difference between three or more means using the analysis of variance (ANOVA) procedure.

Session Objectives

- Understand the assumption of homogeneity of variance (equal variance or homoscedasticity), scale of the variable and independence;
- Become familiar with ANOVA analysis and experimental designs; and
- Compare and contrast causal diagrams, correlation matrices, experimental design, and ANOVA analysis.

Assumptions of ANOVA

- Variances in each experimental condition should be fairly equal (homogeneity of variance).
- Observations should be independent.
- The dependent variable (Y) should be measured on at least an interval scale.
- The variables are normally distributed.

STEP 1: Problem Background

This session deals with the introduction of a new product. Toothpaste GOOD CLEAN is being developed. It is believed that the attributes “colour” as well as the “flavour” of toothpaste influence consumers’ purchase decision. An experiment conducted for testing the preference levels of the product is described below.

GOOD CLEAN can be offered in blue, red or white. The toothpaste can be produced with ‘Mint’ or ‘Jasmine’ flavour. It is important to assess whether “colour” and “flavour” are important to the consumers.

STEP 2: Understand Experimental Design

A study was designed to test for differences in consumer’s preferences (WILL) in the composition of the toothpaste. Each subject was given one of twelve possible combinations of the toothpaste. The consumers were offered one combination and their willingness to purchase was measured. Subjects also indicated their knowledge (KNOWL) of the functional properties of any toothpaste. There are twelve combinations because we have three factors; one factor (colour) has three levels and the other two (flavour and price) have two levels each. That is, we could set up a combination such that the product is blue in colour, with no flavour and is priced at \$2.50 apiece. Or, we could have a combination with no colour (white), mint flavour, priced at \$3.00 apiece, etc. This design is called a “between-subjects design” because each respondent encounters only one combination of the product.

The variables are defined as follows:

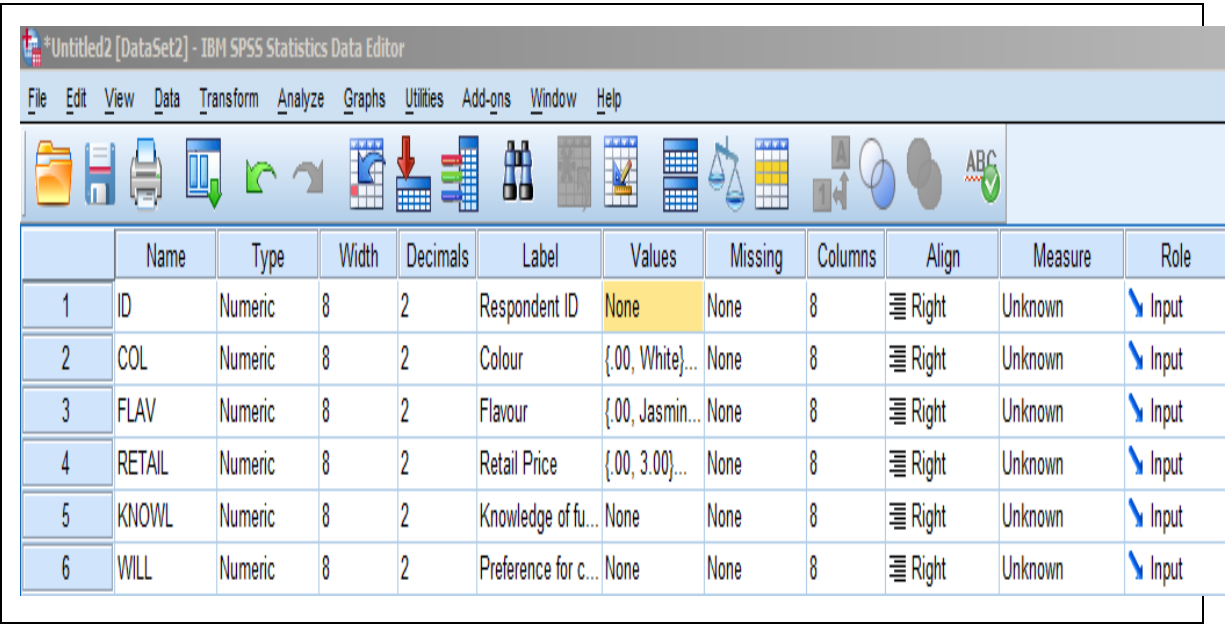
- Let COL = 2 if Red, COL = 1 if Blue, COL = 0 if No colour (white);
- Let FLAV = 1 if Mint, FLAV = 0 if Jasmine flavour;
- Let RETAIL = 1 if \$2.50, RETAIL = 0 if \$ 3.00;

- WILL (preference) is on a scale of 1 to 100 with 100 indicating strong preference;
- KNOWL (knowledge) is on a scale of 1 to 100 with 100 indicating strong knowledge of the functional properties of the toothpaste.

Subjects are given a particular product combination to use and their preference (WILL) for the product is measured. The data are shown in **Table 5.1** included at the end of this note.

STEP 3: Analysis

The data are in raw form and needs to be structured and labelled so that SPSS can understand what the variable levels mean. Go to the **Variable View** (see Fig. 5.1) and label the variables and the levels of the variables in the experimental design accordingly.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	2	Respondent ID	None	None	8	Right	Unknown	Input
2	COL	Numeric	8	2	Colour	{00, White}...	None	8	Right	Unknown	Input
3	FLAV	Numeric	8	2	Flavour	{00, Jasmin}...	None	8	Right	Unknown	Input
4	RETAIL	Numeric	8	2	Retail Price	{00, 3.00}...	None	8	Right	Unknown	Input
5	KNOWL	Numeric	8	2	Knowledge of fu...	None	None	8	Right	Unknown	Input
6	WILL	Numeric	8	2	Preference for c...	None	None	8	Right	Unknown	Input

Figure 5.1: SPSS Variable View – How to Assign Value Labels

Once the data are structured correctly and assumptions have been confirmed, we can analyse it using ANOVA.

To run ANOVA in SPSS you need to do the following (see Fig. 5.2):

- (1) Go to **Analyze** in the pull down menu. Click on **Compare Means > One-Way ANOVA...**
- (2) Pick **WILL** as the dependent variable and **COL** as the factor variable.
- (3) Click on **Options** and tick the **Descriptives** box. This will display the descriptive statistics for the variables along with your ANOVA output.
- (4) Click **OK**. Your results will pop up in the output window (see Fig. 5.2).

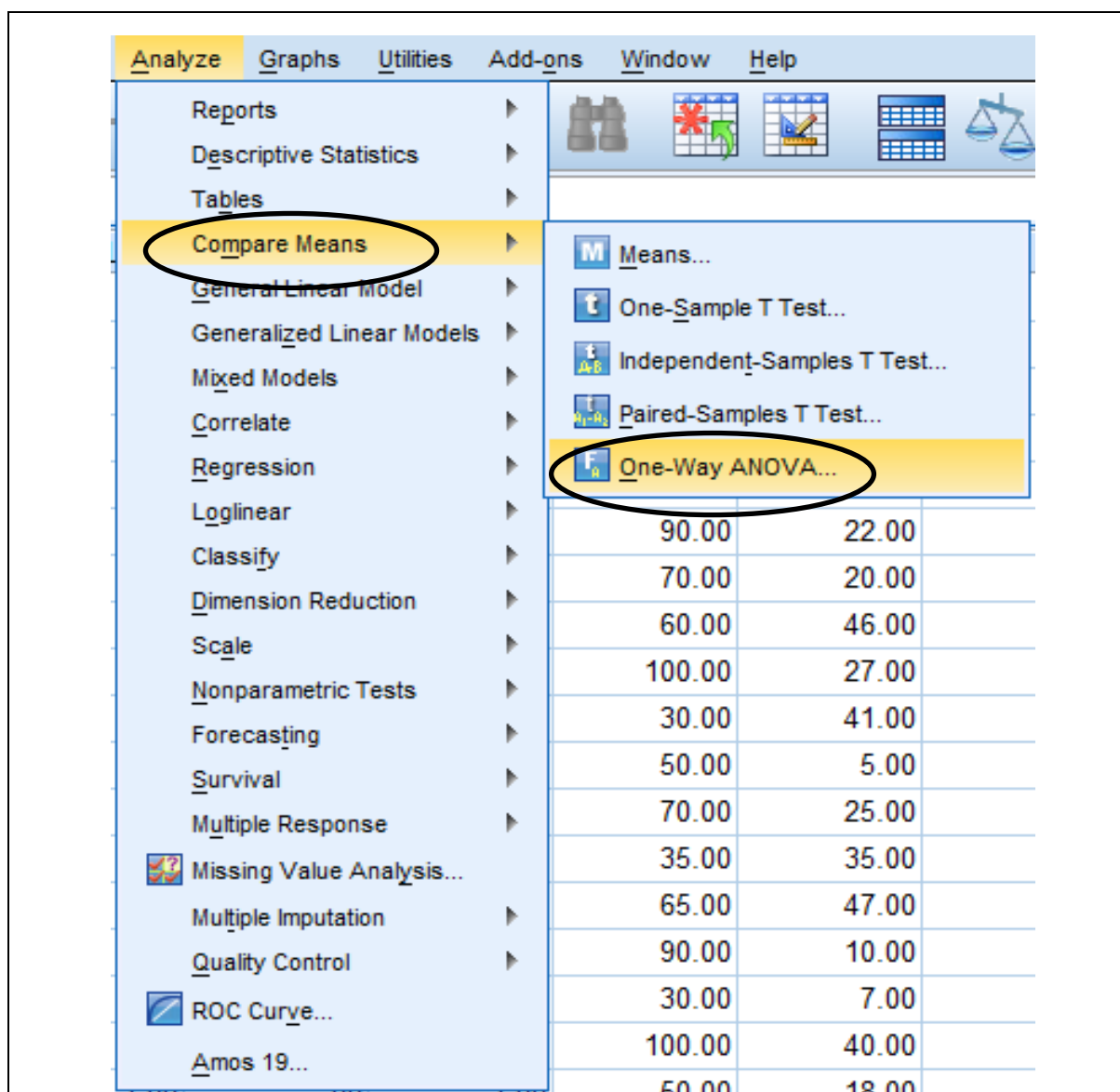


Figure 5.2: Generating One-Way ANOVA - Command

STEP 4: Interpretation of ANOVA output

Step 4.1: Test for significant difference between groups.

The first step is to identify whether there is a significant difference between the experimental groups. To do this we need to look at the F statistic and its corresponding significance value. Although the significance level of .05 is used most commonly, it is up to **you** to choose which level of significance you would like to have: **.01 or .05. or .001**. Is the significance value below .05 (or .01, or .001)? If so, you can be confident (at a 95% level or at a 99% level, or at 99.9%) that there is a true difference between the White, Blue and Red experimental groups. If there was no difference between the experimental groups in the population, you would only have 5% (or 1% - depending on the significance level you chose) of obtaining such a value of the 'F' statistic.

In other words, if the significance level reaches the satisfactory value (the one **you** have decided upon) **the null hypothesis** can be rejected and we can conclude that there is a significant difference between the means of the groups.

Descriptives								
Preference for composition								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
White	11	32,00	13,719	4,136	22,78	41,22	12	50
Blue	11	31,36	27,642	8,334	12,79	49,93	1	80
Red	10	11,40	11,683	3,694	3,04	19,76	1	34
Total	32	25,34	20,931	3,700	17,80	32,89	1	80

ANOVA					
Preference for composition					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2830,273	2	1415,137	3,817	,034
Within Groups	10750,945	29	370,722		
Total	13581,219	31			

Figure 5.3: Results of ANOVA

Step 4.2: Compare the means to understand which condition is preferred.

While we can be confident that there is a difference between groups we don't yet understand which group is preferred. To ascertain this we need to look at the descriptive statistics (specifically the means) to understand which condition is preferred. In this case the White colour is significantly more preferred than the Blue colour and the Red is the least preferred. Pay attention to all the statistics to see if there is further information about the structure of the preferences to inform your managerial implications.

It is also important to look at the confidence intervals for the mean. How big are they? Would we prefer to have bigger or smaller confidence intervals?

Step 4.3: Calculate the effect size

You will notice that SPSS does not produce the effect size, however it can be easily calculated using the figures from ANOVA output tables. We can simply calculate r^2 using the between-group effect (SS_M) and the total amount of variance in the data (SS_T) - for one-way ANOVAs it is called **eta squared, η^2** . We then take a square root to get the effect size r :

$$r^2 = \eta^2 = SS_M / SS_T = 2830.273 / 13581.219 = .208; r = \sqrt{.208} = .46$$

Step 4.4: Managerial Implications

A very important step in interpretation that is often forgotten is the relation of the results back to the original management problem that prompted the research. The following are key areas to explore when considering the managerial implications of ANOVA results:

- (1) Which factors have significant differences in preference?** Here you can report which of the experimental factors have differences in preference for the levels. Non-significant results can also be important because they suggest that preference does not differ across the levels and the manager may be able to cost cut in this area.

- (2) Which is the most preferred level of each of the significant factors?** Here you need to tell the manager which of the levels is most preferred for each of the significant factors. You can also discuss what the ideal combination of factor levels is, according to the results.

(3) What else can descriptive statistics show about preference levels for factors? Using the example of Colour in the ANOVA output, we can gain other information from the descriptive statistics. Look at the means for the colour levels. While White may be more preferred overall than Blue, its mean is still 32 out of 100. Also, Blue has a large standard deviation which suggests people differ in their preference for Blue. As such, some people have a high preference for Blue while others have a low preference for Blue.

Perhaps a new colour needs to be developed that has a higher overall preference than both the current ones available. Perhaps Blue can be maintained for those with a strong preference while another colour can be developed for those with a low preference for Blue.

CLASS EXERCISE

Perform a ONE-WAY ANOVA on:

1. the factor COL on the variable WILL;
2. the factor FLAV on the variable WILL; and
3. the factor RETAIL on the variable WILL.

Use the results to answer the homework exercises.

HOMEWORK EXERCISE

1. Draw the 2 x 2 x 3 experimental design in a table. Make sure all the possible combinations are present.
2. Present the one-way ANOVA for each of the three factors. Do they have significant effects on preference? How can you tell?
3. Which is the most preferred combination of toothpaste? What would you say to a manager of Toothpaste Company that produces a White/Jasmine/\$2.50 combination?

Table 5.1: Data from the New Product Testing Experiment

ID	COL	FLAV	RETAIL	KNOWL	WILL
1	1	1	1	47	80
2	1	1	1	75	60
3	1	1	1	60	76
4	2	1	1	100	2
5	0	1	0	45	50
6	1	1	0	90	22
7	1	1	0	70	20
8	0	1	0	60	46
9	1	1	1	100	27
10	0	1	1	30	41
11	2	1	1	50	5
12	0	1	1	70	25
13	0	1	0	35	35
14	0	1	0	65	47
15	2	1	0	90	10
16	1	1	0	30	7
17	0	0	1	100	40
18	0	0	1	50	18
19	2	0	1	75	24
20	2	0	1	95	8
21	1	0	0	40	1
22	0	0	0	65	19
23	1	0	0	85	26
24	2	0	0	50	34
25	0	0	1	60	19
26	2	0	1	80	24
27	1	0	1	70	11
28	0	0	1	40	12
29	2	0	0	80	1
30	2	0	0	95	1
31	2	0	0	60	5
32	1	0	0	40	15

This page has been intentionally left blank

Workshop 7: Regression Analysis

How to prepare for this task?

- Read the related lecture notes and book chapters.
- Read this note on 'Regression Analysis' and understand the learning objectives.
- Read through the Workshop Questions.

Learning Objective

In this session you will use a data set to conduct multiple regression analysis that explores the likelihood of donation to a charity.

Session Objectives

- To understand when to use regression.
- To check statistical assumptions underlying assumptions.
- To analyse data using linear regression analysis.
- To interpret linear regression output.

Assumptions of Regression Analysis

- All predictor variables (Xs) should be continuous and the outcome variable should be continuous;
- The predictor variables should have some variation in value;
- There should be no perfect linear relationship between two or more predictors;
- At each level of the predictor variables, the variance of the residuals should be constant (assumption of homoscedasticity);
- For any two observations the residuals should be uncorrelated; and
- All residuals in the model should be normally distributed.

STEP 1: Basics of Regression Analysis

Regression analysis is a powerful and flexible procedure for analysing associative relationships between a metric dependent variable (outcome) and one or more independent variables (predictors). There is a tight link between the model diagrams you have been drawing and regression. The variable that has the **arrows pointing at it** is the “**dependent**”, “**response**”, and/or “**Y**” **variable**. The variables that have the **arrows point away from** them are called “**independent**”, “**predictor**” and/or “**X**” **variables**.

Mathematically, regression is represented as an equation where the value of **Y** (dependent variable) is a sum of the values of all the **X**’s (independent variables) + a constant. The **X** values are calculated as the product of **X** and **b** (parameter) which is a weighting of the contribution of the **X** to the overall value of **Y**. A regression equation is below:

$$Y = b_0 + b_1 * X_1 + b_2 * X_2 + b_3 * X_3 + + \quad + \quad + \quad b_n * X_n$$

In many ways, regression is very much like correlation analysis. However, there is something very special about regression that makes it unique, and that is its ability to do “conditioning”. What conditioning means is “controlling for”; it removes the effect of a variable allowing you to focus on relationships that are of more interest to you. Here’s an example:

Women with more children were found to have higher blood pressure than women with fewer children. The correlation between a woman’s blood pressure and their number of children was high. On first sight it appears that having more children is the cause of blood pressure. However, as we know, association is not causality.

It was also observed that women with more children are older than women with fewer children. When these effects are included in a regression model of a woman's blood pressure, the effect of the number of children goes away. To view this another way, old age causes one to have health problems. In turn this causes an increase in the blood pressure.

Regression has a special way of cutting through the mess of complex correlation matrices and figures out what variable/s have what association with what other variables.

STEP 2: Problem and Data Screening

You will use a new data set. As regression requires a larger sample size an existing file called “Charity Regression dataset.sav” is available. Please find the file on the shared drive as per your tutor’s instructions. Before you run any analysis you will need to screen the data to ensure there are no keying errors in your file.

CLASS EXERCISE 1

Run Descriptive Statistics on the data set to screen the data.

CLASS EXERCISE 2

Another good exercise before conducting analysis is to draw a model diagram. This allows you to visualise what you are testing and understand the hypothesised directions of your relationships.

Draw a model diagram of what you are testing. Remember to include a + or – to indicate the direction of the hypothesised relationships.

STEP 3: Regression Analysis (see Fig. 6.1 & 6.2)

- (1) Go to the **Analyze** pull down menu. Click **Regression → Linear...**
- (2) Choose your DV (dependent variable) and click the arrow to move it to the **Dependent:** box. Select your IVs and move them across to the **Independent(s):** box.
- (3) Click **Statistics**; tick **Estimates**, **Confidence Intervals**, **Model fit**. Click **Continue**
- (4) Click **OK**. The output window will pop up with the results. (see Fig. 6.2)

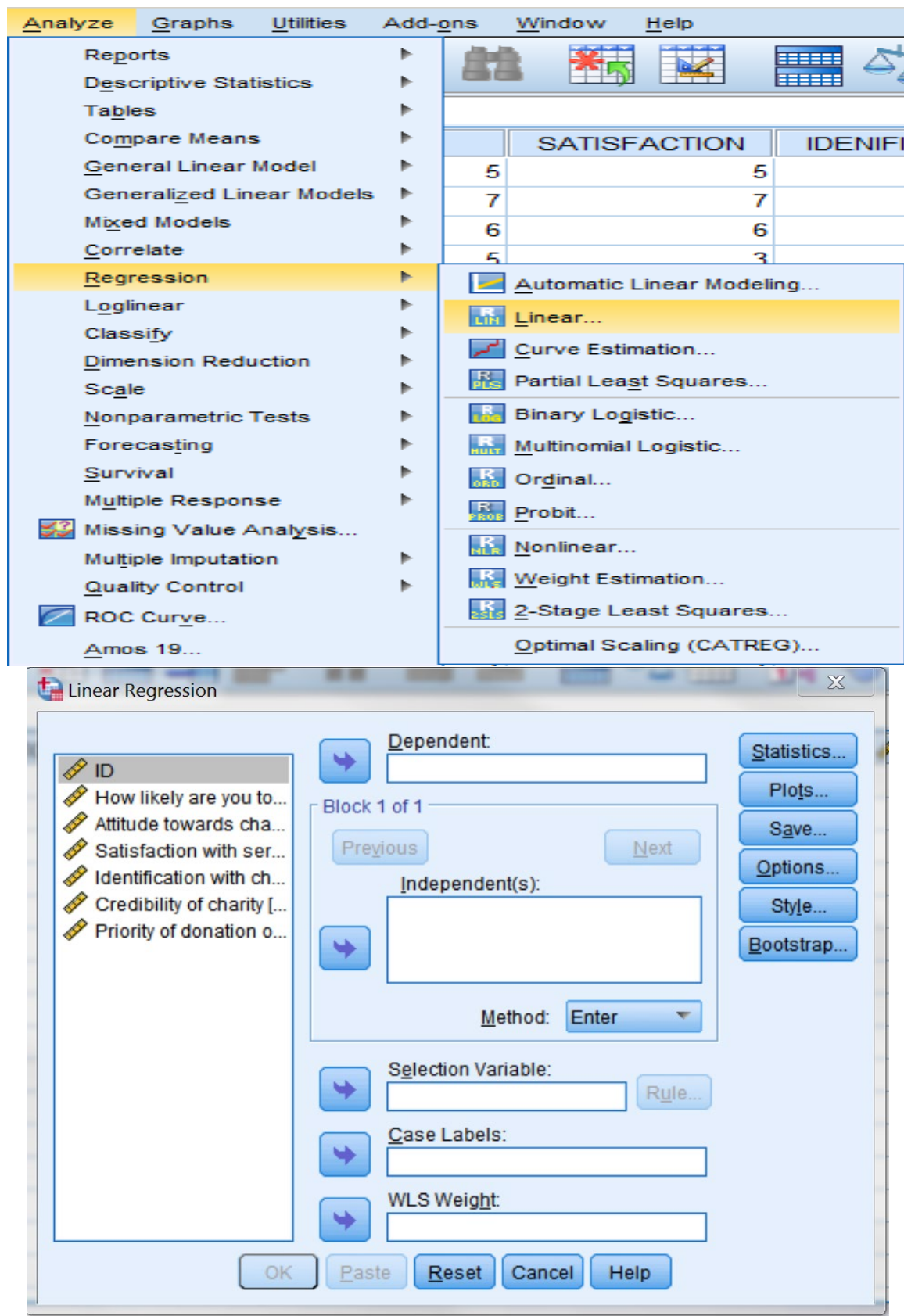


Figure 6.1: Command for Linear Regression

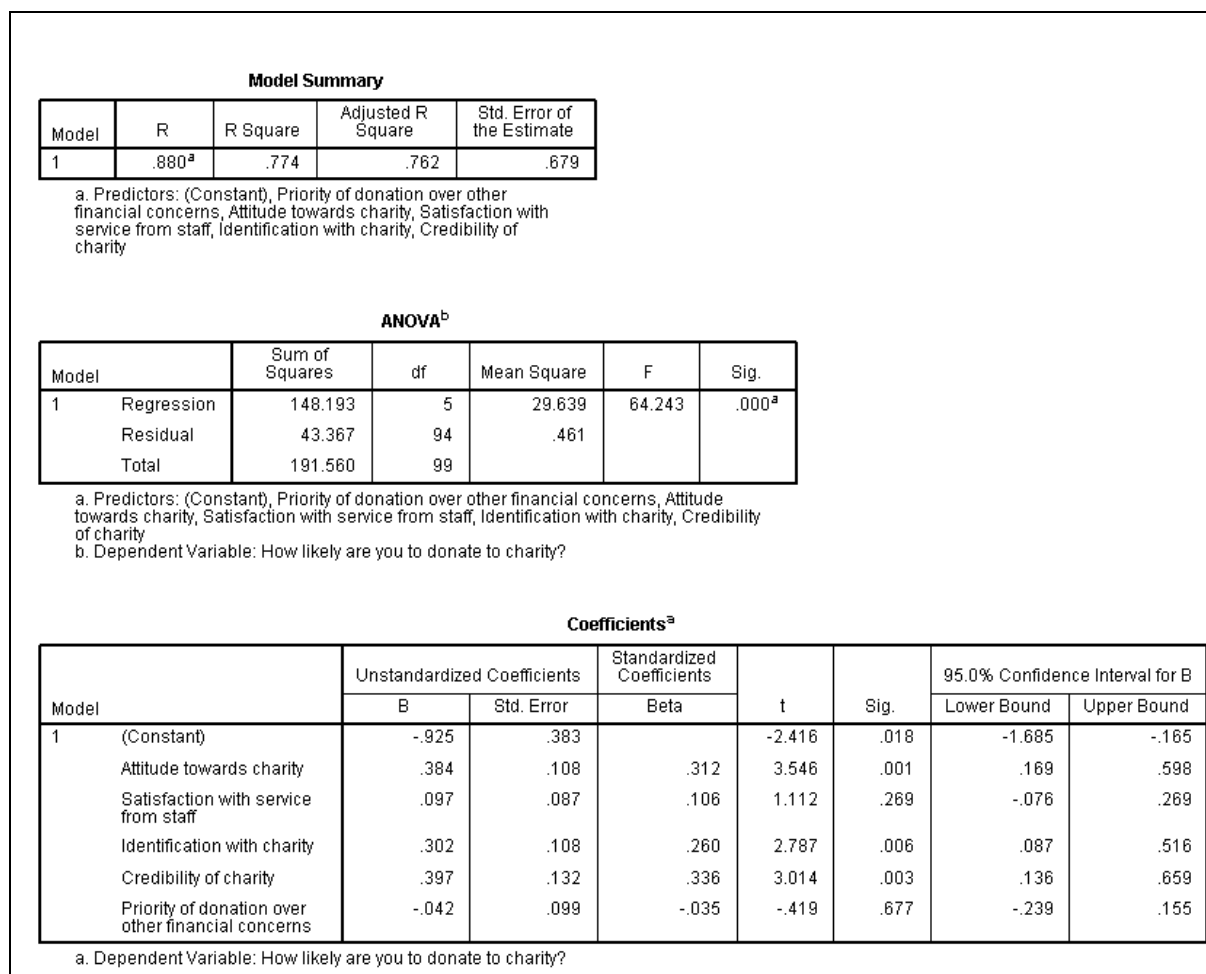


Figure 6.2: SPSS Regression Output

STEP 4: Interpreting Regression Output

Understanding regression output is another matter. Essentially, there is a great deal of redundant or useless information in a regression output that usually scares off a lot of students. Statistics packages report every conceivable number that one can report in regression. Only a few measures are really important. A recommendation is to look at only four types of numbers in a regression output and then ignore the rest. The four important measures are “p-values”, “R²”, “parameters” and, “standardised parameters” (Betas).

There are five key steps to interpreting a regression. The following discussion will use an example of a regression on a set of data trying to explain the key contributors to Customer Satisfaction (Figure 6.2).

Step 4.1: Interpret the Significance of the Model (see Fig. 6.2)

The first step is to find out whether the overall regression model is statistically significant. If the overall regression model is not statistically significant then none of the individual relationships will be significant and there would be no point in interpreting further (that's why we check this first)! To do this we need to look at the Sig value (p-value) in the ANOVA table. The ANOVA tests the R^2 against 0 to see if they are significantly different and whether the null hypothesis can be rejected. In the example the Sig value is <.001. This is below our significance value of 0.05 therefore the regression model is statistically significantly different from 0 and further interpretation is required.

Step 4.2: Interpret the Strength of Association (see Fig. 6.2)

The strength of association is measured by the square of the multiple correlation coefficients, R^2 , which is also called the coefficient of determination. As you can see in the Model Summary table there are two R^2 values (R^2 and Adjusted R^2). Unlike R^2 , adjusted R square allows for the degrees of freedom associated with the sums of the squares. This in effect prevents the design considerations such as sample size or the number of variables interfering with the calculation of R (large sample sizes and numbers of variables falsely inflate R). For this reason, adjusted R^2 is generally considered to be a more accurate goodness-of-fit measure than R^2 . Adjusted R^2 is interpreted as the percentage variance explained in Y (Dependent variable) by the X's (Independent variables). The adjusted R^2 in the example is .762. This means that 76.2% of the variation in Y is explained by the variables in the model.

Step 4.3: Interpret the Partial Regression Coefficients (see Fig. 6.2)

Now we understand that the overall model is statistically significant and how much variation is explained in the DV, we can move onto finding out which individual IV → DV relationships are significant. In the Coefficients table there is a Sig column that displays the p-values for each of the relationships. Again, the significance cut off is 0.05. We do not interpret the constant. (Note: theoretically, the constant is the value of Y where X's are zero)

In the example, three variables are statistically significant: Attitudes, Identification and Credibility. Satisfaction and Priority did not reach the required level of statistical significance. As these three relationships are statistically significant we can represent the regression model using their parameters (b values) as:

Likelihood of donating = $-.925 + .384 \times \text{Attitude} + .302 \times \text{Identification} + .397 \times \text{Credibility}$

Step 4.4: Interpret the Standardized Regression Coefficients (see Fig. 6.2)

Now we have determined which relationships are statistically significant we need to assess the strength of each relationship. To do this, we stay in the Coefficients table and look at the Standardized Beta column. Betas are the estimates that have been standardized so that their means are zero and variances are 1. Therefore, standardized coefficients refer to how many standard deviations a dependent variable will change, per standard deviation increase in the predictor variable. Standardization of the coefficient is usually done to answer the question of which of the independent variables have a greater effect on the dependent variable, when the variables are measured in different units of measurement (for example, income measured in dollars and family size measured in number of individuals).

Betas range from -1 to 0 to 1. Betas (β) are interpreted in absolute terms, so the direction of the relationship is not taken into account when considering strength (so -.567 and .567 are as strong as each other even though one is negative and one is positive). In the below example Credibility has the strongest relationship ($\beta = .336$) followed by Attitude ($\beta = .312$), followed by Identification ($\beta = .260$).

Now have a look at the two last columns in the “Coefficients” table: Lower and Upper Bound of confidence intervals for B. The confidence intervals represent the boundaries of unstandardized coefficients B computed in such a way that if we collect a number of different samples of data, in 95% of these samples these boundaries will contain the true value of b. That means that we can be fairly confident that the true value of b for the whole population will also be within these boundaries. Therefore, in a good model, confidence intervals should be small, indicating that the value of b in this sample is close to the true value of b in the whole population. We should be really wary of the models in which

confidence intervals cross zero (for example they range between -10.12 to 5.23), because that would mean that in some samples the predictor has a negative relationship with the outcome, in some samples it has positive relationship and there is also a chance that it has a zero relationship in the other samples. In our model, the confidence intervals for Satisfaction and Priority cross the zero mark, which emphasises that we cannot use these variables to predict the outcome. For those three significant variables though (Credibility, Attitude and Identification) the confidence intervals include the value of the B and do not cross the zero mark.

Step 4.5: Assess the managerial implications of the results (see Fig. 6.2)

A very important step in interpretation that is often forgotten is the relation of the results back to the original managerial problem that prompted the research. There are three key areas of managerial interpretation that should be explored:

Can the null hypothesis be rejected?

Here you can make a comment on the statistical significance of the overall model. If the overall model is significant, you can reject the null hypotheses.

How much variance is explained?

Here you can make a comment about how much of the variation in the DV is explained by the IVs in the model. Whether the model explains 20% or 80% of the variation will have a significant impact on how much change can occur in the DV following the regression model.

What relationships are statistically significant?

Here you can outline to managers which of the relationships in the model are significant using the p-values (i.e. which variables have an effect on the DV). Keep in mind that only statistically significant IVs explain the variance in the model.

Which are the strongest relationships?

Here you can outline which of the relationships are strongest using the Beta values. It is important to mention here, that statistical significance alone does not make much difference unless the effect size of the strength of relationships is big enough. That is why you need to assess R^2 for the overall model and Beta-coefficients for the IVs. Emphasising strongest relationships in the model helps to prioritise the most important issues to focus on, as time and money spent in areas with strong relationships should enact the most change on the DV.

CLASS EXERCISE

1. Check the descriptive statistics to ensure there are no errors; and
2. Perform Regression Analysis on the data set.

Use the results to answer the homework exercises.

Where can I use Regression?

To understand the relationship between metric independent variables and a metric dependent variable. To predict a metric variable using one of more metric predictor variables.

HOMEWORK EXERCISE

1. What is the model diagram that would accompany the above regression model? Does it differ from the diagram you drew before your analysis?
2. Provide an interpretation of the regression results. What would you say to a manager of a charity?

This page has been intentionally left blank

Workshop 8: Factor Analysis

How to prepare for this task?

- Read the related lecture notes.
- Read the Factor Analysis task and understand the learning objectives.
- Read through the Workshop Questions.

Learning Objective

In this session you will use a data set to conduct factor analysis.

Session Objectives

1. To learn how to run a Factor Analysis;
2. To understand Varimax Rotation; and
3. To apply Factor Analysis as a data reduction technique.

Assumptions of factor analysis

- The variables should be measured on at least an interval scale.

STEP 1: Problem and Data

Kay Sealey is the news director for KASI-TV, the local NBC affiliate for a large Southwestern city in USA. Sealey believes that the most important quality of an air newsreader is credibility in the eyes of the viewer. Accordingly, surveys are taken every six months that attempt to evaluate the credibility of the newsreader. One of the survey instruments used by the station is given below.

Questionnaire		
Evaluate the newsreader on the news broadcast that you reviewed by completing the following series of scales. Place a check mark on the scale position that most nearly matches your feelings about this newsreader. For example, if you thought that in this newsreader was extremely likeable, you would place a check mark in the blank nearest “likeable” (in this case, the far left blank).		
1. not knowledgeable	- - - - -	knowledgeable
2. unintelligent	- - - - -	intelligent
3. not believable	- - - - -	believable
4. not likeable	- - - - -	likeable
5. unattractive	- - - - -	attractive
6. bad looking	- - - - -	good looking

The questionnaire was administered to 40 individuals. **Table 1** contains the responses of the 40 people surveyed. The data was coded between 1 and 7. For example, a check mark closest to likeable would be coded as 7 and not likeable as 1.

It is important to mention here that to conduct factor analysis, the variables should be measured at least on ordinal scale and the number of observations should be minimum 100.

STEP 2: Interpreting the correlation matrix

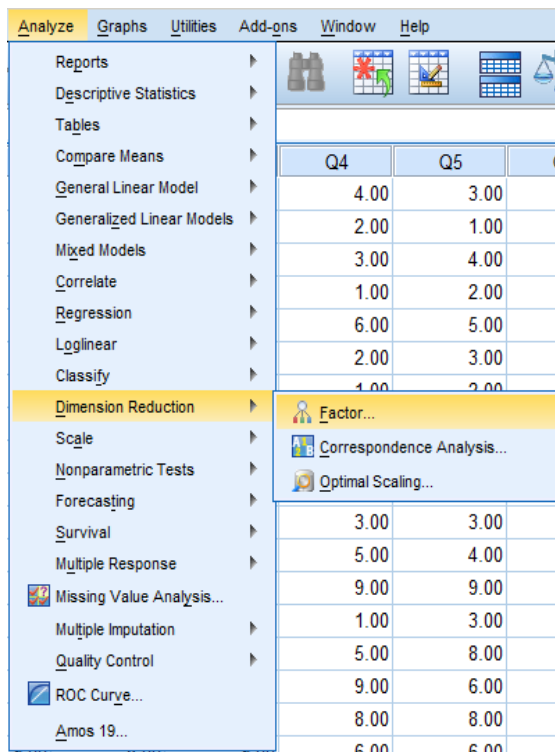
Produce a correlation matrix on the scale variables using your skill from the correlation tutorial.

INTERPRETATION:

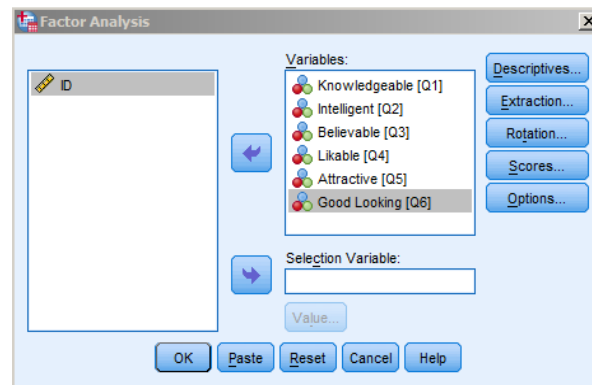
Which variables are correlated? Is factor analysis appropriate for this data?

STEP 3: Principal component factor analysis

(1) Go to the **Analyze** pull down menu. Click **Dimension Reduction** → **Factor...**



(2) Select your scale variables and put them in the **Variables:** box.



(3) Click **Extraction...** and tick the **Scree Plot** box.

(4) Click **Continue**. Click **OK**. Your results should pop up in an Output screen.

Communalities

	Initial	Extraction
Knowledgeable	1.000	.935
Intelligent	1.000	.883
Believable	1.000	.886
Likable	1.000	.885
Attractive	1.000	.824
Good Looking	1.000	.821

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.506	58.435	58.435	3.506	58.435	58.435
2	1.728	28.793	87.228	1.728	28.793	87.228
3	.302	5.040	92.267			
4	.246	4.099	96.367			
5	.124	2.062	98.429			
6	.094	1.571	100.000			

Extraction Method: Principal Component Analysis.

Component Matrix^a

	Component	
	1	2
Knowledgeable	.848	-.465
Intelligent	.771	-.537
Believable	.782	-.525
Likable	.758	.557
Attractive	.743	.521
Good Looking	.674	.606

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Figure 8.1: Running factor analysis in SPSS and factor analysis output

STEP 4: Interpretation of Factor Analysis

There are many different methods that have been developed to understand how many factors to place on scale variables. While books may differ on the best approach, the general consensus is that multiple rules should be used. There are four common approaches to determining how many underlying factors there are in scale data:

Determination based on Eigenvalues. In this approach, only the factors with eigenvalues greater than 1.0 are retained. An eigenvalue represents the amount of variance associated with the factor.

Determination based on Scree Plot. A scree plot is a plot of eigenvalues against the number of factors in order of abstraction. The shape of the plot is used to determine the appropriate number of factors. Typically, the plot has a distinct break between the steep slope of the factors (with large eigenvalues) and a gradual trailing off associated with the rest of the factors (the *scree*). The break between the steep slope and the gradual slope indicates the maximum number factors. The break is commonly one factor more than the eigenvalues cut off.

Determination based on Percentage of Variance. In this approach, the number of factors is determined so that the cumulative percentage of variance extracted by the factors reaches a satisfactory level. What the level is depends on the problem but as a general rule of thumb a minimum of 60% variance should be explained. Additionally, if a factor contributes more than 10% variance, than it should be considered for inclusion. In order to test this, factor

analysis is often run forcing a factor solution above and below the initial number of factors found to ensure the correct structure for the data is identified.

Determination based on Theory. While this is not mentioned a great deal in analysis textbooks, theory should also be taken into account when assessing the number of factors in the data. If theory suggests a certain number of facets of a variable than this should be used as a tie breaker should two factor structures be equally appropriate.

INTERPRETATION:

How many factors are there? How many factors should be imposed on the data and why? (Make your determination based on multiple sources of evidence from the common approaches mentioned above).

STEP 5: Varimax rotation

Once the number of factors has been ascertained, VARIMAX rotation can be used to make the structure of the data (which questions align with which factor) more evident. For this another factor analysis must be run.

- (1) Go to the Analyze pull down menu. Click Dimension Reduction → Factor...**
- (2) Select your scale variables and put them in the Variables: box.**
- (3) Click Extraction... and tick the Scree Plot box. Also click the Fixed number of factors button and input the number of factors you believe should be in the data according to the criterion. Click Continue.**
- (4) Click Rotation... Click the Varimax button. Click Continue.**
- (5) Click OK. Your results should pop up in an Output screen.**

INTERPRETATION:

Have a look at the rotated factor matrix to see the structure of the data. Which items factor together to form each variable? Can you give the variables a name?

STEP 5: Creating Variables

As the factor analysis suggests there are two variables underlying the six item scale. SPSS can be used to create new composite variables which can then be used in further analysis. In order to compute new variables you need to pay attention to the structure in the data. Note down which questions factor together to form each of the two variables.

To create a new composite variable:

- (1) Go to the **Transform** pull down menu. Click **Compute Variable...**
- (2) Give the new variable a name in the **Target Variable** box. Use the **Numerical Expression** box to calculate the variable.

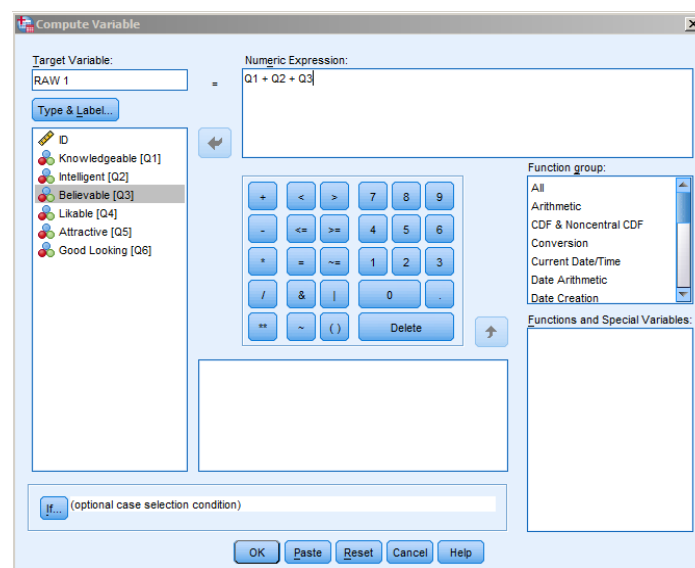


Figure 8.2 - Creating a composite variable in SPSS

- (3) In order to create an average of the scores across the items in a variable you can select each variable and + them together, then divide them by the number of items (i.e. $Q1+Q2+Q3/3$) or you can use the average function in the **Function Group**.
- (4) Click **Continue**. The new variable will appear as an additional column in your data editor.

STEP 6: Checking Reliability

However, after creating a new variable we might also want to check how reliable the new scale is. To do this, we can calculate a Cronbach's alpha, α . Usually the value of .7 or .8 is an acceptable value of Cronbach's alpha; values substantially lower than that indicate an unreliable scale.

To compute the Cronbach's alpha in SPSS:

- (1) Click on **Analyze** pull down menu. Choose **Scale → Reliability Analysis**.
- (2) Drag Q1, Q2 and Q3 into the **Items:** box.
- (3) The **Model** is set by default to **Alpha** (which is what we need).
- (4) Click **OK**.

The output will appear, now you need to find the **Reliability Statistics** table and interpret the result.

CLASS EXERCISE

1. Create a correlation matrix;
2. Run a FACTOR ANALYSIS;
3. Determine the number of factors underlying the data;
4. Re-run a FACTOR ANALYSIS using VARIMAX rotation; and
5. Create the variables suggested by the data.

Use the results to answer the homework exercise questions.

Table 8.1: Data from the Newsreader Survey

ID	Q1	Q2	Q3	Q4	Q5	Q6
101	7	7	7	4	3	5
103	2	2	3	2	1	2
105	8	7	9	3	4	1
106	4	4	3	1	2	5
107	8	7	9	6	5	6
109	3	2	3	2	3	2
112	2	1	2	1	2	3
113	1	2	2	4	2	5
114	6	6	8	4	2	5
116	2	2	1	2	2	4
119	6	9	6	3	3	5
120	4	4	1	5	4	2
121	7	8	7	9	9	8
123	2	3	3	1	3	2
124	4	2	3	5	8	9
125	5	1	4	9	6	8
126	8	9	5	8	8	6
128	9	8	9	6	6	8
129	2	1	1	5	9	6
131	7	7	4	8	8	9
132	2	1	1	7	6	6
133	8	8	9	8	8	5
135	7	9	5	5	7	8
136	5	2	4	3	3	3
138	9	7	8	6	5	7
139	8	8	9	8	8	6
140	7	8	8	9	9	9
141	8	8	5	5	5	8
143	9	8	9	8	6	9
144	1	2	1	2	3	5
145	8	9	9	7	8	6
146	9	8	9	5	3	5
147	2	3	5	4	3	4
151	5	8	6	2	5	3
152	7	9	6	1	3	1
153	4	4	3	1	3	2
155	8	9	8	2	5	4
156	9	8	8	1	2	2
159	6	5	7	9	8	8
160	9	7	8	2	5	4

HOMEWORK EXERCISE

1. Why was factor analysis used for this data?
2. How many factors are there? Use multiple criteria to determine the number of factors and justify your answers. Provide a figure for this scree plot.
3. How reliable are the composite variables?
4. Use descriptive statistics to describe how the station's newsreaders are performing in relation to the two composite variables?

Part III – Class Exercises

Exercise 1: Building Blocks of Research

Task: Answer the following questions based on this scenario. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Scenario

The attitudes and behaviours of frontline employees can impact on firm performance. That is, interactions with frontline employees can shape customers' satisfaction ratings and perceptions of service quality. Recognising this, the customer service director at a large service firm is seeking to improve the firm's management of its frontline employees. In particular, the manager is interested in identifying the factors that shape frontline employee performance. The manager has speculated that the firm's human resource policies and practices, and frontline employees' perceptions of the work environment are important determinants of performance.

Challenge Question 1

Give a market research example of a potential dependence relationship. (Specify the DV and the IV.) What is the rationale for their relationship?

Challenge Question 2

Describe the difference between a mediating variable and a moderating variable with the help of diagrams and marketing examples.

What is the practical significance of identifying mediating and moderating variables?

Challenge Question 3

What is a hypothesis? How can you determine whether a hypothesis is a 'good' hypothesis?

How are descriptive hypotheses different from relational hypotheses? Provide marketing examples.

– End –

Exercise 2: Problem Formulation

Task: Answer the following questions based on this scenario. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Briefly describe the notion of the question hierarchy. How does the management question differ from the research question (and how are they linked)?

Identify and state the management question and the research question in the following scenarios.

Scenario 1

Creativity (the production of new ideas) and innovation (the implementation of new ideas) may be the only enduring sources of competitive advantage. The chief knowledge officer at a global professional services firm is seeking ways to enhance employee creativity and innovation. Some of its offices have produced a series of new ideas and practices, while other offices have been consumers rather than producers of knowledge. There is a belief that the nature of the work itself (standardised versus non-standardised), work practices (individual versus group), and the attributes of individual employees impacts on creativity and innovation.

Scenario 2

Characteristics of the store environment can shape consumers' patronage choices. Common characteristics that can influence store patronage include the use of music, colour and scent, and store crowding. The marketing director of a national retailer is aware of wide variation in the success of its many retail outlets. The director is convinced that the differences in store environment explain at least part of the variation in store patronage. In particular, the director is seeking to understand how consumers' perceptions of the retailer's employees, the design of its stores, and store characteristics impact on their perceptions of merchandise value and patronage intentions.

– End –

This page has been intentionally left blank

Exercise 3: Research Design

Task: Answer the following questions based on this scenario. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Briefly discuss the potential benefits of exploratory studies. How do descriptive studies differ from explanatory studies? How are explanatory studies different from exploratory studies?

Challenge Question 2

Briefly describe the notion of research design (what is included, how can they be described, classified). For each of the following examples, identify which type of research design (exploratory, descriptive or causal/explanatory) would be more appropriate. Provide justification for your choice.

- a. You want to investigate the reasons behind people's choice of marketing major
- b. A local bakery wants to know if their new advertising is increasing sales.
- c. You are creating a profile of frequent shoppers of discount stores.
- d. You want to know whether the choice of university by one's peers influences an individual's own choice of university.
- e. A company wants know how satisfied its employees are.
- f. You want to investigate whether people who work more hours per week are more likely to eat out at restaurants.
- g. You want to understand the factors that influence people's choice of department stores.
- h. You were told that the organisational culture at a large mining company had largely contributed to its strong productivity and market performance. You have decided to investigate this issue.

Challenge Question 3

What are secondary data (list some specific examples)? What are the main advantages and disadvantages of using secondary data? Describe a marketing research problem that could be addressed using secondary data.

This page has been intentionally left blank

Exercise 4: Research Design- II-Ethics

Task: Answer the following questions based on this scenario. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question

Listed below are eight scenarios describing research situations that may involve questionable decisions. For each of the scenarios, do the following:

- (i) Identify the ethical issues that have been violated (if any), and
- (ii) Suggest a solution for the problem. Is there a better way to conduct this research?

Scenario 1

You are conducting a survey on behalf of a large telecommunications company. The company has requested that 1000 of their customers be surveyed. The deadline for the project is quickly approaching and you have only surveyed 990 customers. You decide to complete ten surveys yourself to complete the project on time.

Scenario 2

You have recently conducted interviews with 20 employees at an accounting firm about the amount of stationary they steal from the office. The employees are unaware that, in exchange for allowing you to conduct the research, you have agreed to give the officer manager a copy of your findings, including the names of the employees.

Scenario 3

You are interested in researching the organisational culture of a local real estate agency. You apply for a job at the real estate agency to gain a true insight into their day-to-day activities and interactions. You conceal your motives from the management and the employees. You make friends with some of the employees and intend to quote them in your research.

Scenario 4

You are a university tutor who wants to test the effect of unfair punishment on exam performance. You administer an exam to two different groups of students. You then artificially lower the grades for one group and proceed to berate the group for their poor performance. You then administer the exam again, and you discover that the performance of the unfairly punished group is substantially lower. Your hypothesis is confirmed and the results are published.

Scenario 5

You promise participants in your study a summary of the results. Later, due to a budget cut, the summary is not sent out.

Scenario 6

On the day that a research project is due for a client, you realise there is a fatal error in your data analysis. However, the report is already completed and you know that the company will never notice the mistake. Due to time pressures, you decide to deliver the report without changing the mistake.

Scenario 7

You are testing perceptions of product quality for store brands. You ask customers to taste two different kinds of canned corn, peas and tomatoes. You ask them to identify which one is the brand name product and which is the store brand product. In actuality, both products are store brands. Eighty percent of customers stated a quality difference in the products.

Scenario 8

You keep repeatedly calling a respondent to complete a survey over the phone, even though they have explicitly told you that they do not wish to participate.

— End —

This page has been intentionally left blank

Exercise 5: Survey Methods

Task: Answer the following questions based on this scenario. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

What are the key applications of survey research? Briefly comment on the advantages and disadvantages of survey research. Provide an example of the use of key informants. How are key informants selected by researchers (what is the screening criterion)?

Challenge Question 2

What is response error? Provide examples of various types of response error that may occur in the following scenario. What can researchers do to reduce/eliminate this type of error?

Scenario 1

The Australian Tax Office wants to conduct a survey on the level of tax evasion in the population. They conduct a mall-intercept survey, with members of the tax office intercepting tax payers in major shopping centers. The interviewers identify themselves as tax office employees and the organization who is conducting the survey is undisguised.

Challenge Question 3

What is nonresponse error? How might nonresponse error affect the following study? How does a researcher address this?

Scenario 2

A surf clothing company wants to understand how the shopping habits of customers differ across age groups. Specifically, they want to understand the shopping habits of their key market - 18-25 year and 26-35 year old males. To achieve this aim they employ a research agency to conduct telephone interviews using a random sample of 2000 names out of the phone book.

Challenge Question 4

What is the total design method (the Dillman method)? How is it useful for survey researchers? Discuss how this method can be applied in the following context.

Scenario 3

A software company conducts research on the needs of consumers with regard to their office software products. They conduct a survey every year in order to keep pace with the change in consumer preferences. Previous studies have provided vital information, however the response rate has been low (10% response rate). This year they have requested their research agency address this issue in their upcoming survey.

— End —

This page has been intentionally left blank

Exercise 6: Scaling

Task: Answer the following questions based on the given scenarios. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Briefly describe each of the following four scale types (what are their possible limitations?): nominal, ordinal, interval, and ratio.

Challenge Question 2

In each of the following examples, identify and state whether the measurement scale is a nominal, ordinal, interval or ratio scale. Explain why?

(i) Where in Queensland do you live?

- a. Brisbane
- b. Gold Coast
- c. Sunshine Coast
- d. Other: _____

(ii) How many times did you go to the movies this month?

(iii) How would you rate your experience at the hotel?

Terrible 1 2 3 4 5 Excellent

(iv) How old are you?

- a. Under 18
- b. 18-25
- c. 26-35
- d. 36-45
- e. 46-55
- f. Over 55

(v) How old are you?

(vi) What is your preferred brand of consumer electronics?

- a. Sony

- b. Panasonic
- c. Teac
- d. Akai
- e. Other: _____

(vii) Please write a number from 1 to 7 (where 1 is your most preferred) next to each genre of music listed below to indicate your relative preference.

- a. Hip Hop
- b. Pop
- c. Rock
- d. Classical
- e. Dance
- f. Techno

– End –

This page has been intentionally left blank

Exercise 7: Questionnaire Design

Task: Answer the following questions based on the given scenarios. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Evaluate and comment on the following questions:

- a) To make up for past discrimination, do you favour or oppose programs that make special efforts to help minorities get ahead?

☐ Favour
☐ Oppose

- b) Do you believe that private citizens have the right to own firearms to defend themselves, their families, and their property from violent criminal attack?

☐ Yes
☐ No

- c) How would you rate the room facilities and the attitude of the staff?

☐ Poor
☐ Satisfactory
☐ Good
☐ Excellent

Challenge Question 2

Consider the following draft questionnaire for a telecommunications company. What possible problems are there associated with this questionnaire, and what changes and improvements would you suggest? (There are at least ten areas for potential improvement!)

Continued...

Household Telecommunications Provider Survey

1. Your name? _____

2. Age _____

3. Marital status _____

4. Income _____

5. Which of the following Telecommunications providers are you aware of?

Telstra

☐

Optus

☐

Internode

☐

Others

6. How frequently do you make international calls?

Infrequently

Very Frequently

1

2

3

4

5

6

7

7. What do you think of the customer service offered by Telstra?

8. Suppose your household were to select a Telecommunications provider.

Please rate the importance of the following factors in selecting the Telecommunications provider.

	Not Important				Very Important
Cost of local calls	1	2	3	4	5
Ease of use	1	2	3	4	5
Cost of long distance calls	1	2	3	4	5
Rebates and discounts on calls	1	2	3	4	5
Quality of telephone calls	1	2	3	4	5
Quality of customer service	1	2	3	4	5

9. How important is it for a Telecommunications provider to offer broadband internet?

Not Important

Very Important

1

2

3

4

5

6

7

10. Do you have children living at home? _____

Thank You for Your Help

This page has been intentionally left blank

Exercise 8: Sampling Design

Task: Answer the following questions based on the given scenarios. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Briefly describe the notion of sampling and state the basic rationale for sampling. What are the advantages and disadvantages of sampling versus conducting a census? How could the quality of a study be better using a sample rather than a census?

Challenge Question 2

List the different types of sampling designs. How do nonprobability sampling techniques differ from probability sampling techniques (and what are their relative advantages and disadvantages)?

Challenge Question 3

For each of the scenarios listed below, identify the population, the sampling frame and the most appropriate sampling method.

- (a) A private health insurer wants to conduct research on their customers' price sensitivity. Their customer base can be split into three different groups: singles, couples and families with children. The health insurer suspects that these three groups have varying levels of price sensitivity.
- (b) A large shopping centre would like to survey its customers about the layout of their car park facilities.
- (c) An art exhibition would like to display a varied sample of an artist's paintings.
- (d) A law firm would like to interview employees regarding their satisfaction in the workplace.
- (e) A children's clothing manufacturer has produced a batch of 1000 t-shirts of a particular model. They would like to test the fabric strength (e.g. how much force is required to rip a t-shirt). The calculation of fabric strength requires physically tearing up the cloth.

– End –

This page has been intentionally left blank

Exercise 9: Data Analysis

Task: Answer the following questions based on the given scenarios. This task is NOT marked. However, completion of this task should immensely help you in achieving the course learning objectives and doing well in the assessment tasks.

Challenge Question 1

Why might a questionnaire be unacceptable for analysis (use examples to illustrate your answer)? What options do researchers have for handling unsatisfactory responses (and how might this affect the results)?

Challenge Question 2

Briefly explain the concept of coding (and provide examples of how nonmetric and metric variables might be coded). How could the researcher code (and measure!) the following variables?

Gender

Age

Income

Attitudes towards imported products

Voting intentions at a Federal election

Brand loyalty (attitudinal)

Brand loyalty (behavioural)

English speaking status

Challenge Question 3

What are the broad differences between univariate, bivariate and multivariate techniques?

How do dependence techniques differ from interdependence techniques?

This page has been intentionally left blank