Additional SPSS Activities. Access the relevant file in the SPSS Data Files folder to obtain the required data set.

1.Pizza Customers'

Part A

A national pizza company operates two hundred separate branches in the country. The owner of the company is interested in finding out just how effective the use of advertising signs on the side of the road approaching each store are in attracting customers. In order to do this, 14 branches in each of four regions record the number of customers that enter each branch over a two-week period with on-the-street advertising placed outside 14 branches in 2 regions only (Regions 1 and 2 in the data set). Branches in the other regions do not get this advertising (Regions 3, and 4).

Using the SPSS data file '**Pizza Customers'** on the website:
- Test the null hypothesis that there is no statistically significant difference between the number of customers entering the company's pizza shops in the four regions. Assume the groups are independent and set alpha at .01. Based on the printout what are your conclusions?
- Practise using SPSS to create different graphs, charts and tables to indicate any differences between the regions. Do these help to illuminate the variations in customers between the regions? .
- What are the strengths and weaknesses to this research design?

Part B

Following this two-week period, the owner decides to extend his study over another two-week period to investigate whether advertising the company's pizza shops in regions 1 and 2 on regional television is more successful. Shops in regions 3 and 4 would not have any regional television advertising.

Using the SPSS data file '**Pizza Customers'** on the website, now test the null hypothesis that television advertising has no statistically significantly effect on customer numbers between the regions.

- As well as obtaining a statistical printout, try using SPSS to create different graphs and tables to indicate any differences between the two stores.

- What are your conclusions?

After undertaking the analyses consider the following important points about the weaknesses and strengths to the above designs.

If you have had your 'thinking-caps' on, you would have realized that there are several weaknesses to both of these analyses!

In the first analysis, it would be hard to make any assertion about whether the on-the-side of the road really did contribute to an increase in the number of customers because we don't know whether or not this level of customer visitation is typical of those shops without the advertisement. In fact the evidence from comparing regions 1 and 2 suggests that there was possibly little effect from such advertising. How can we improve the design?

1 Firstly, we could have taken a baseline measure. That is, we could have taken records of the number of customers entering all the shops in the study before the on-the-side of the road advertising was introduced. We then want to test whether there is a significant increase over the baseline for each of the advertising regions and non-advertising regions.
2 Another possible way of overcoming this design weakness, would be to introduce on-the-road advertisements for Regions 3 and 4 after week 1. However, it still does not overcome the lack of baseline figures to support the effect of the advertising, which is the only way to confidently claim that the on-the-road advertisement had an effect.

In the second analysis, one possible weakness to the results could be attributed to practice effects. As Regions 1 and 2 had engaged in two-weeks of advertising already, customers who came during the second experiment could have done so as a result of seeing the on-the-road advertisement the previous week. This detracts from any conclusions we can draw about the effectiveness of the television advertisements in regions 1 and 2 as the customer numbers may partly have been due to the earlier on-the-road advertisement.

2 SPSS Dataset 'Organizational Climate'

Following recent changes in personnel at senior and middle management levels, there has been a significant increase in employee complaints with increased stress levels and decreased morale. Whilst workers blame the increase in negative emotions and decrease in positive emotions on various aspects of organizational climate such as a decline in supportive leadership, excessive work demands and removal of participative decision making that have occurred as a consequence of an influx of new managers, management believe that it is the workers themselves in terms of their personalities that determine their feelings of well-being.

As an external consultant, you have been tasked with identifying the factors that predict employee well-being within this company.

For this analysis you will need measures of organizational climate, employee personality, and employee well-being. The data for this exercise can be found in the SPSS data file '**Organisational Climate'**, on the website. The variables 'Supportive Leadership', 'Excessive Work Demands', and 'Participative Decision-Making' are used to measure 'Organisational Climate'. Employee Well-being is assessed by the variables 'Distress' and 'Morale', whilst Personality is assessed by the variables 'Extraversion' and 'Neuroticism'. High scores on each of these variables indicate high levels. For example, a high score of 10 on the 'Distress' variable indicates a high level of stress, whilst a score of 2 indicates a low level of Distress. Similarly, a score of 10 on the 'Supportive Leadership' variable indicates a high amount of Supportive Leadership, whilst a low score of 1 indicates very little "Supportive Leadership".

(a) Establish appropriate hypotheses to test the following questions:
1   Which organisational climate variables predict employee distress?
2   Which organisational climate variables predict employee morale?
3   Which personality variables predict employee distress?
4   Which personality variables predict employee morale?
5   Which best predict distress, organisational climate or personality?
6   Which best predict morale, organisational climate or personality?

(b) When answering these questions, create some graphs that will support the statistical analysis that you will do.

(c) Interpret your print out.

3 SPSS Datasets 'Internet Dating1' and 'Internet Dating2'

You are running an Internet dating service and you are interested in the possibility that physical appearance affects judgements about other characteristics that could impact on individual's success at being asked by a prospective partner out on a date. The importance of this may be that individuals who do not present well, may significantly reduce their chances of being selected for a date.

To test this hypothesis 10 women who had recently applied to join the dating service were offered free membership in return for their participation in the study.

Each of the women were shown video footage of 15 men, whom by consensus amongst the employees of the dating service had been designated into one of three groups based on physical appearance and grooming: *Well Kept, Average*, and *Untidy*. The female participants were then asked to judge each man's potential as a long-term partner on a 10-point scale.

The data file "**Internet Dating1**" can be found on the website to test the null hypothesis that physical appearance and grooming does not affect judgements about potential as a long-term partner.

To carry out the analysis, you will need to use the **TRANSFORM > COMPUTE VARIABLE** command in order to create three aggregate scores for the three groups: *Well Kept, Average,* and *Untidy*. Then you will be able to compare mean scores between the three groups. Do a repeated measures test because the three sets of judgements are being made by the same 10 women..

In addition, you may like to experiment and generate several graphs to determine whether some participants are more likely to alter their judgements between categories. It may be that one or two people do not change their judgements due to physical appearance; instead, they may not make such judgments till they know the individuals in the photos better before making an extreme judgement – either one way or the other – and scoring average scores in all conditions and also demonstrating very small standard deviations in their ratings across the three groups. Also, as it is video footage, it may be that some participants are influenced more by body language, vocal inflections, and the like. Obviously the effect of these factors would be assessed by other studies where observers were asked to rate on these.

Now that you have drawn your conclusions, you decide to replicate this study with 10 males. This will allow you to support the findings that you came to with your female sample. However, you should investigate whether there are gender differences, demonstrated by mean differences between the two groups. Also, was there greater uniformity in male responses? In the female sample above, you may have found that some participants were not affected by the conditions. Is this the case for participants in the male sample?

Using the data file "**Internet_Dating2**", test the following null hypotheses:
1 That there are no differences between gender levels in judgements of viability of long-term partner between conditions.
2 That there are no differences between gender levels in the degree of variability in responses to the judgements of viability of long-term partner between conditions.

Again, in order to carry out the analysis, you will need to use the **TRANSFORM > COMPUTE VARIABLE** command in order to create three aggregate scores for the three groups: *Well Kept*, *Average*, and *Untidy*

4 SPSS Dataset 'Beauty Contest'

The annual national beauty contest is under way. Each of the final 10 contestants receives a rank on both personality and appearance. Using the data file '**Beauty Contest'** on the website answer the following questions about the variables 'personality' and 'appearance':
(a) Calculate the correlation between personality and appearance.
(b) What conclusions do you draw about personality and appearance?

Using the same data file investigate using the variable 'height' to determine whether there is any correlation between any of the three variables. As 'height' is stated in metres, you may need to transpose the variables in some way in order to compare appropriately all three variables. Think how you could standardise them. What are your conclusions now?

5 SPSS Dataset 'Developing Nation1' and 'Developing Nation2'

(a) In a developing country the government is trying to identify whether there is a real gap in salaries between those who live in major cities, mixed rural/urban and rural regions of their country. Using the data file '**Developing Nation1'** on the website, determine whether there are significant differences between average salaries in three randomly chosen regions which represent one each of a major city, mixed rural/urban, and rural location.

What are your conclusions?
(b) Fearing that such research may lead to mass migration from rural and mixed areas into the cities, leading to increased poverty and overcrowding in the cities, you believe that the relationship between areas that people live and the salaries they receive is more complicated. Instead, you believe that the government should encourage students to stay in education by arguing that there is a relationship between years of post-primary education and salary. When discussing your ideas with a colleague, they further suggest that there might be a difference in the quality of educational institutions between

locations, with better facilities in the city than in the other regions. It may be that students do not stay in education because of their location – the expectation in rural areas is that they leave school to work on the farms and manual industry, whilst those in mixed rural/urban areas need further education to undertake the technical, white-collar and supervisory occupations. Also, because the numbers completing post-primary education is less in the rural areas and mixed areas, priority in funding school resources should be directed mostly to rural and mixed area schools.

Now using the data file 'Developing Nation2' from the website:
(a) undertake an exploratory investigation to determine to what extent the variables: 'Location', 'Years of Post-Primary Education', and 'Quality of School' may predict salary.
(b) Further, you may want investigate to see whether there is an even more complex relationship with 'Location' and 'Quality of School' predicting 'Years of Post-Primary Education'.
(c) Also, is there an association between a school's location and its quality?

Write up your conclusions so that you can prepare a report of results. Remember, figures, tables and flow charts showing relationships between variables are visually appealing and help to convince others of your arguments.

6 SPSS Dataset 'Employee Well-Being'

As part of an organizational health assessment of your corporation, you are unsure whether the scale you are using to measure employee well-being really does measure what it purports to measure. The scale of employee well-being consists of two dimensions: Positive Emotions and Negative Emotions. Whilst interviewing employees, several commented that they were unsure how to answer the question *"I often feel alert"* (question P5 in the data file). Whilst the authors of the scale classified it as a Positive Emotion, you believe that '*being alert*' should be classified as a Negative Emotion as some of the participants commented that they felt that they always had to be 'alert' because they believed middle management was always looking to catch them doing something wrong.

Using the data file '**Employee Well-Being**' from the website, analyse the data to determine how well the data supports the validity of the well-being scale and its two dimensions, and whether the question 'p5' has been correctly coded as an item measuring Positive Emotion.

Unsure how to do this?

Two hints: Factor analysis and consider the pattern matrix; and Cronbach's Alpha but remember the scale is already in the form of two subtests!

Recode p5 and conduct the analyses again. Explain your results.

7 SPSS Dataset 'Air Traffic Controllers'

Recent changes to the national workplace health and safety laws requires that smoking is banned in the workplace. Whilst not denying the possible health risks associated with smoking, there are concerns by some air traffic controllers, who work in a highly stressful occupation, that these changes may have considerable impact on their ability to maintain a high level of cognitive functioning whilst managing a high level of work demand and stress. Traditionally, workers were allowed to smoke whilst working, and it was assumed that nicotine controlled the high stress levels associated with this sort of occupation.

There was considerable debate during the last employee–managers meeting about complaints relating to this problem, as smokers expressed concern that reducing their frequency of smoking breaks would add to their stress and reduce their ability to concentrate and perform the problem-solving tasks that their work required. Obviously, sneaking out for a quick five minute smoke is not an option for controllers who must maintain a heightened state of vigilance at work.

You have been hired to identify whether smoking has an effect on the quality of performance of air traffic controllers.

You decide to sample 45 employees, 30 of whom smoke and 15 who are non-smokers (NS). Of the 30 smokers, you randomly allocate half to an Active Smoking (AS) group, meaning they are to have smoked during or within an hour of undertaking the experiment, whilst the other smokers are allocated to a Delayed

Smoking (DS) group, meaning they had not smoked for at least four hours before undertaking the experiment. The non-smokers were allocated to a Non-Smoking (NS) group.

Since much of the work within this organization requires higher-level cognitive skills, you decide to test these three groups on three different cognitive tasks: a pattern recognition task, a memory task, and a driving simulation. On each of the tasks, you have measured the number of errors the participants made (a higher number indicating a higher error rate). In addition, you believe that the number of errors may co-vary with an employee's level of distraction, in which case you also measured the total amount of time that an employee was distracted from the task at hand (a higher number indicating a higher rate of distraction).

Using the data file '**Air Traffic Controllers'** on the website, analyse the data to determine whether
- there are differences between smokers and non-smokers in error rates on the different cognitive tasks:
- there are differences between the AS and DS groups in error rates on the different tasks:
- rate of distraction co-varies between conditions, explaining differences between the AS, DS and NS groups.

8 SPSS Dataset 'Predicting Dementia''

A significant problem facing most governments in the developed world, as a consequence of successful treatments for life threatening illnesses such as cardiovascular disease and cancers, is markedly extending life expectancy. Conversely, treatments for problems associated with cognitive decline, such as age-related cognitive decline, mild cognitive impairment and dementias, such as Alzheimer Disease have shown little advancement.

It is believed that a number of individual and environmental factors are related to the rate of cognitive decline in later life. To test this hypothesis, you have interviewed a large sample of elderly retirees over 70 years of age, to determine which factors more significantly predict the clinical diagnosis of dementia. It is hoped that being able to predict those who are most at risk of dementia can lead to successful interventions much earlier in life, reducing the likelihood of dementia developing.

Using the data file '**Predicting Dementia'** from the website, identify which factors best predict the clinical diagnosis of dementia. Diagnosis was classified as either 'No Clinical Diagnosis' or 'Clinical Dementia'. You will need to consider which type of regression is suitable. The predictors in this analysis include Gender, Marital Status, Years of Education, and Frequency of Alcohol Consumption. The coding for these factors is included in the data file, and remember that SPSS needs to know which factors are categorical variables.

9 SPSS data set 'College Activities'

(a) Academic performance may well depend on many factors and not just intellect. Using the data set '**College Activities'** determine which combination of the variables: TV watching, computer time, alcohol consumption, sleep and exercise, best predict GPA among the students. Write out the regression equation. Using the **Data >Split File** procedure, find out if this combination is the same for each faculty. Is it the same for each gender. (Remember to reverse the file split before you try to undertake further analyses of the whole file below.)
(b) Are there any significant relationships between hours of exercise and hours spent on the computer, amount of alcohol drunk and hours of TV watched.
(c) Is there any faculty-gender interaction on average hours of computer use?
(d) Is there an association between gender and health status, and between faculty and health status?
(e) Are there any significant differences between health status levels and (i) amount of alcohol consumed, (ii) hours of sleep?
(f) Is there a gender-health status interaction on amount of exercise taken per week?