

Assignment 3

1.6 Supplementary Exercises

1.

- a) This table shows how the 95% confidence interval of workers in the restaurant that answered 'YES'. Here we can see the interval at 57.8% to 77.0% (1 d.p.)

Binomial Proportion	
WORKSTRESS = YES	
Proportion	0.6800
ASE	0.0466
95% Lower Conf Limit	0.5886
95% Upper Conf Limit	0.7714
Exact Conf Limits	
95% Lower Conf Limit	0.5792
95% Upper Conf Limit	0.7698

- b) This measure of confidence shows that the national survey which was performed fits between the intervals and therefore does suggest that the majority of participants who answered work stress does have a negative impact on personal life.

Interestingly the range between the confidence intervals is a little over 20%. But from what we can draw from these values is that when individuals have a stressful day at work their troubles may stay on their mind which affects their personal lives.

As a separate thought, to evaluate the data ideally we could dissect the way in which the data was collected.

The job roles in the restaurant chain can vary for example waiter/chef/cleaner. If we see the 100 employees that were taken and understand the pressures of the different job roles and number of employees taken from each job category the data might not be representative of the whole restaurant chains. Perhaps another way this could've been performed is by choosing a defined number of people with respect to their job where the 100 would be divided into equal job roles and then the employees randomly sampled from their respected job categories. By doing this we can then see perhaps a more representative outcome rather than taking many or less individuals from a certain job category.

```
DATA RESTAURANT;
INPUT WORKSTRESS$ COUNT;
DATALINES ;
YES 68
NO 32
;
RUN;
```

```
PROC FREQ DATA = RESTAURANT ORDER=DATA;
WEIGHT COUNT;
TABLES WORKSTRESS / BINOMIAL;
RUN;
```

2.

a) Using the TTEST process I was able to generate the table showing averages and the 95% confidence intervals. This shows the intervals (in £'s) being between 4.14 and 7.00 (2 d.p.).

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
5.5667	4.1359	6.9975	2.2519	1.5952	3.8235

b) This sample data along with the distribution chart and Q-Q plot we can see visually the proportion of amount spent. Using the 95% confidence interval £5 does fall within the limits. The sample data does provide evidence that the amount spent by customers is around £5 and it shows the proportion spent was around £5 as people had also spent in excess and less than £5 this is shown in the distribution below. Yet the Q-Q plot and the boxplot displayed also identify that the median to be a little over £5.

We can notice how the data is centrally skewed and that the value of 10.50 seems to be an outlier where an individual has spent more than the average amount at the cinema complex. Therefore, fittingly the median as a measure of robustness represents the approximate average spent by individuals which was around £5

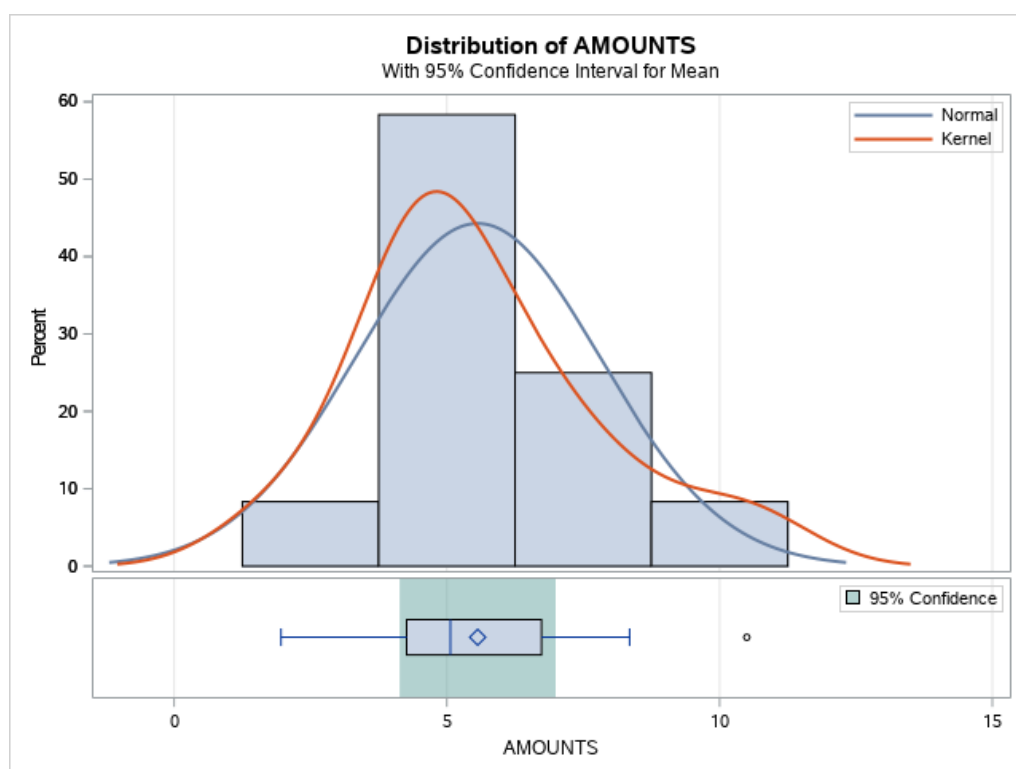


Figure 1: Distribution of amounts spent at the cinema complex.

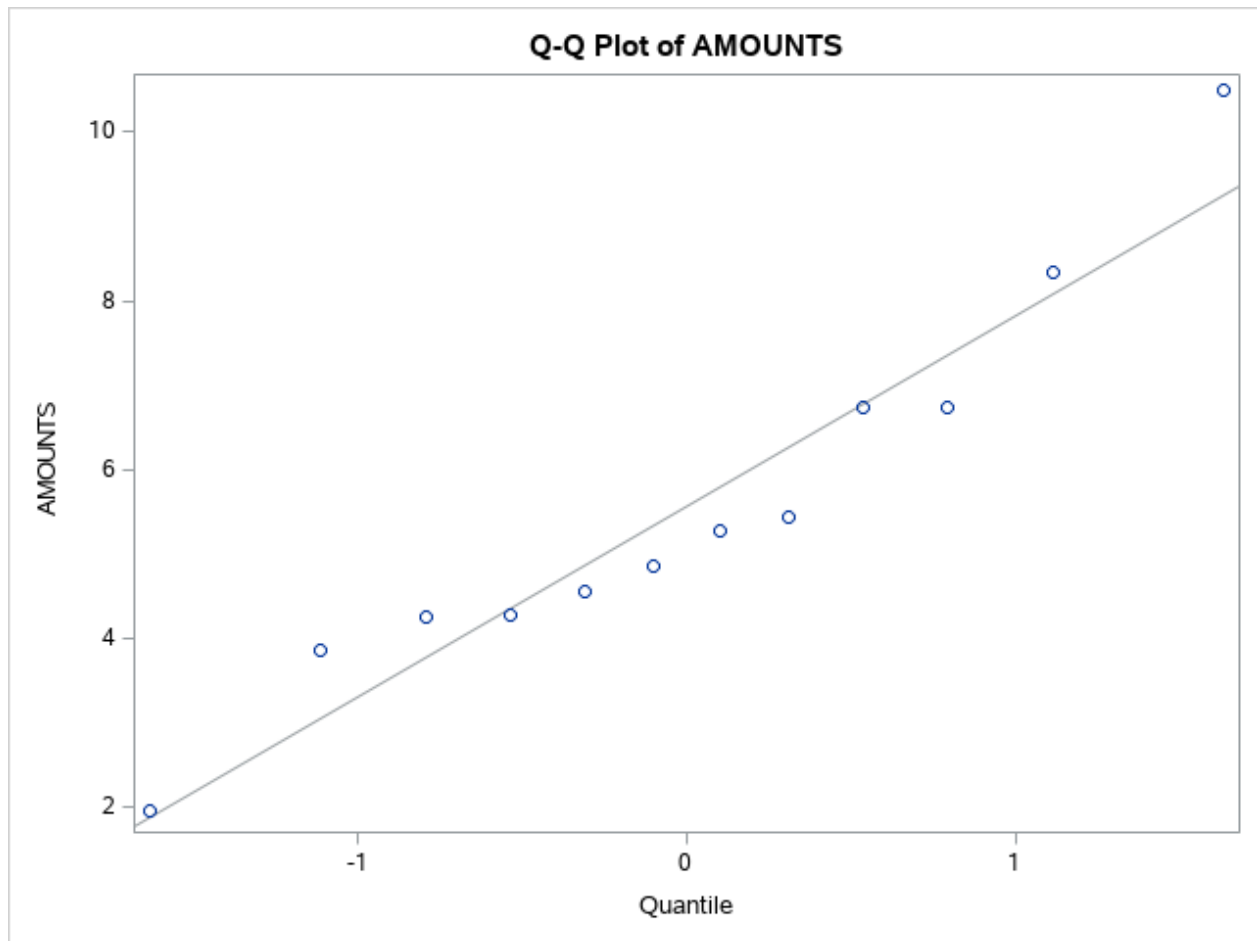


Figure 2: Q-Q plot of amounts spent at the cinema complex. As we can see the 0 quantile (median) lies around £5.

```
DATA CINEMA;  
INPUT AMOUNTS @@;  
DATALINES;  
3.85 5.28 6.74 1.96 4.85 4.28 6.73 4.56 5.45 8.35 10.50 4.25  
;  
RUN;
```

```
PROC TTEST DATA = CINEMA;  
VAR AMOUNTS;  
RUN;
```