**ECON1310**

**Tutorial 2 – Week 3**

**DESCRIPTIVE STATISTICS II**

At the end of this tutorial you should be able to

* Define measures of central tendency, variation and shape for ungrouped data.
* Calculate measures of central tendency and variation using a calculator or Excel.
* Draw and interpret a box and whisker plot.
* Calculate a coefficient of correlation using Excel and interpret the answer.



**Q1.** a) Suppose a distribution is symmetrical with mean μ.

 i) The median must equal the mean μ. T/F

 ii) The mode must equal the mean μ. T/F

 iii) The mode must equal the median. T/F

 iv) Approx 68% of the values must lie within one standard deviation of the mean T/F

b) An additional observation is added to a data set and this observation is larger than all the previous observations. The new observation

 i) always causes the median to increase. T/F

 ii) sometimes causes the median to increase. T/F

 iii) always causes the mean to increase. T/F

c) i) If the standard deviation is large, the data are less dispersed. T/F

 ii) A data set with more than one mode is said to be bimodal. T/F

**Q2.** Each of ten taste-testers rated a new brand of barbecue sauce on a 10-point scale where

 1 = awful and 10 = excellent. The ten ratings are 8, 7, 9, 6, 8, 10, 9, 9, 5, 7.

 Which measure of central tendency would be the most appropriate to use here. Why?

**Q3.** a) The following are fourteen responses to a graduate survey on the salaries offered to Bachelor of Business graduates (in $ thous).

29.7 26.5 18.0 27.0 30.2 37.7 32.9

22.0 38.6 25.3 19.6 36.0 31.9 38.6

Calculate the five number summary and draw a box plot by hand. How would you describe the shape?

Calculate the overall average salary in dollars.

Does it matter here whether you are using units of $thous or dollars?

Calculate the length of the whiskers. Discuss.

What is the absolute value of the difference between the median and mean of the sample?

**Q4.** a)Currently your inventory consists of 20 bottles of discounted wine that you are willing to sell at cost. You purchased 6 bottles for $12/bottle, 8 bottles for $14/bottle and 6 bottles for $6/bottle. If a single selling price (the weighted average price) is to be set for all 20 bottles, what price/bottle should be set in order to recover all of your costs?

b)A student's best attempt for each of their six CML quizzes at the end of the semester in ECON1310 were 14/24, 21/27, 22/30, 14/26, 11/26, and 17/30. Their mid semester exam mark was 34/50. The best five out of six CML quizzes contribute 20% to the end of semester grade, the mid semester exam 25%, and the final exam 55%. By referring to the ECON1310 course profile, what minimum percentage mark would be required on the final exam for the student to receive:

 i) a grade of 4? ii) a grade of 5?

 If this same student had not attempted quiz 3 at all (and so received a score of 0 instead of 22/30), while the scores for the other five quizzes remained the same, what minimum percentage mark would be required on the final exam for the student to receive:

 iii) a grade of 4? iv) a grade of 5?

**Additional Practice Questions (answers will be posted on Blackboard)**

**PQ1.** Use the data file **T2 data.xls** to answer the following questions. The data file has been posted in the *Tutorial Materials* folder on Blackboard.

1. Use Excel Data Analysis to compute descriptive statistics.
2. Calculate $\overbar{x}\pm s$ and $\overbar{x}\pm 2s$. Using the empirical rule, how many data points are expected to lie in these ranges? How many do? From these answers, do you consider that the distribution is approximately normal?
3. Which measure of central tendency is best? Why? Use the measures of central tendency to comment on the shape.
4. Compute the first and third quartiles and the interquartile range.
5. Compute the Coefficient of Variation. What does this tell us? How can the value be most appropriately used?
6. Construct a horizontal box-and-whisker plot and describe the shape. Compare with your answer in c). Use *Kaddstat/boxplot* to obtain a vertical plot.

**PQ2.** The number of cars that went through a car wash of a petrol station in the morning over a seven-day period is as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **6** | **3** | **9** | **6** | **7** | **5** | **1** |

a. Find the mean, median and mode.

b. Estimate the values of the 1st and 3rd quartiles and determine the interquartile range.

c. Construct a box plot.

d. On the basis of your results above, comment on the shape of this distribution.

**PQ3.** a. Construct a box plot from the statistics presented below, and from observation of the plot, describe the shape.

|  |  |
| --- | --- |
| Mean | 5.2035 |
| Standard Error | 0.0132 |
| Median | 5.2 |
| Mode | 5.19 |
| Standard Deviation | 0.0631 |
| Sample Variance | 0.0040 |
| First Quartile | 5.1800 |
| Third Quartile | 5.2400 |
| Range | 0.26 |
| Minimum | 5.09 |
| Maximum | 5.35 |
| Sum | 119.68 |
| Count | 23 |

b. Assuming the empirical rule may be applied to this data set, determine the range $\overbar{x}\pm s$ and specify how many data points from this dataset would be expected to be in this range.

**PQ4.**

 Data on two variables, X and Y, are listed below.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **i** | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* |
| **X** | 7 | 5 | 8 | 7 | 6 | 17 | 12 | 4 | 9 | 12 | 18 |
| **Y** | 21 | 15 | 24 | 9 | 18 | 30 | 36 | 8 | 27 | 45 | 36 |

 a) State the value of X6 and Y9

 b) Calculate the correlation coefficient, r, using:

 (i) the ***fx*** /*Correl* function in Excel

 and (ii) *Data / Data Analysis / Correlation*

 *(for Excel 2003 use tools / data analysis / Correlation)*

[Notice the difference in data entry and results presentation between the two Excel methods.]

 You should obtain r = 0.769. How would you describe the strength of the relationship