**University Z Test Mark Analysis**

This audit trail has been produced by A. N. Actuary on 10th June 2018. It relates to the spreadsheet ‘UniZ Test Marks Spreadsheet Model.xlsx’.

**Objective**

This spreadsheet has been produced for Professor Kingston, a Maths lecturer at University Z. The Maths department was recently rated ‘poor’ and the Professor has been trying to improve performance. He has tested 100 students at 2 week intervals to measure performance. The purpose of the spreadsheet is to analyse the performance of the students over 12 consecutive tests.

The pass mark is set for each test by calculating the median test mark and rounding up to the nearest whole number.

As an incentive to study harder, the Professor decided to offer prizes to any student whose average percentage increase per test is greater than 10%. This spreadsheet also calculates how many students would win the prize.

The model performs the following calculations:

* The average mark for male and female students has been calculated and shown on a chart.
* The distribution of students across three year age bands has been calculated and shown on a chart.
* The average and median test marks and the standard deviation of the test marks has been calculated for each of the 12 tests.
* The pass mark has been calculated for each of the 12 tests.
* The pass mark has been compared to the marks the Professor was hoping to achieve, and this has been shown on a chart.
* The number of students who would receive the prize has been calculated.

**Data**

Professor Kingston has provided data showing the age and gender of 100 students and their test marks (out of 100) for the last 12 consecutive tests.

The following validation checks have been performed on the data:

* There are 100 marks for each of the 12 tests.
* The maximum mark for each test is less than or equal to 100.
* The minimum mark for each test is greater than or equal to 100.
* 100 genders and ages have been provided.
* The maximum age has been calculated as 33. This seems reasonable for a mature university student.
* The minimum age has been calculated as 18. This seems reasonable for a university student.
* All of the genders are either M (for male) or F (for female).

The checks all give satisfactory results and so I have assumed the data is accurate.

**Assumptions**

The following assumptions have been made in the model:

* The same 100 students sit each test.
* A student passes if their test mark is greater than or equal to the pass mark.
* The age of each student remains constant across all tests.
* The minimum and maximum possible test marks are 0 and 100 respectively.

**Parameters**

This sheet sets out the parameters used in the model:

* The number of students is 100.
* The number of tests is 12.
* The pass rate is 50%.
* Professor Kingston hopes that the first test will have a pass mark of 40, and the final one will have a pass mark of 66.
* The average increase in test mark required to get a prize is 10%.

This sheet also shows the colour coding used in the model.

**Data Statistics**

This sheet first calculates the average test mark across all tests for males and females. The gender and age for each student is linked to the data sheet. For each student, the average mark across all of the tests is calculated using the AVERAGE function. The average mark for males is calculated as follows:

$$\frac{Sum of male average marks}{Number of males}$$

These values are calculated using SUMIF and COUNTIF functions. The same calculation is performed for females. The results are shown on a bar chart.

The following checks have been performed:

* There are 100 gender, 100 ages, and 100 average marks.
* The average mark across all students is calculated in two ways:
	+ A weighted average using the male and female results from above.
	+ An average across all of the students.
	+ These calculations give the same result.
* The maximum and minimum average marks are calculated. These numbers look reasonable.

The next set of calculation on this sheet shows the number of students in three year age bands. The maximum and minimum ages are calculated as 18 and 33. This means 6 three year age bands are required to cover the 15 year age gap. For each age band I have calculated the total number of students aged less than or equal to the upper age. I have then used simple subtraction to calculate how many students are in each band. I have converted the numbers to proportions and plotted these in a pie chart. I added data labels to the pie chart.

I have checked that the total number of students across all of the age bands is 100 as expected.

**Test Mark Analysis**

This sheet calculates the average, median and standard deviation of test marks across all 100 students for each of the 12 tests. The functions used are AVERAGE, MEDIAN and STDEV.P.

The pass mark has been calculated for each test using the method described by Professor Kingston. This is to round up the median to the nearest whole number (using the ROUNDUP function). The pass rate has then been calculated by using COUNTIF to count the number of students with marks greater than or equal to the pass mark. This is equal to 50 for each test as expected.

Professor Kingston hoped that the pass mark work increase linearly from 40 for the first test to 66 for the last test. The estimated pass mark for each test is therefore calculated as (66-40)/(12-1)=2.36 more than the previous test mark. Note that the first and last cells in this row link to the parameter sheet.

The actual and estimated pass marks have been plotted using a line chart. The estimated pass mark is linear, as we would expect.

The following checks have been performed:

* The pass rate for each test equals 50% (or 50 students).
* The estimated pass mark increases by 2.36 for each test.
* The average mark across all tests matches that calculated on the ‘Data Statistics’ sheet.
* The maximum and minimum across the average and median test marks looks reasonable (e.g. is it around the middle of the 0-100 range).

**Prize Calculations**

In this sheet I have first calculated the total increase in test mark across all the tests using the following calculation for each student:

$$i\_{total}=\frac{Final test mark}{First test mark}-1$$

The average increase per test is then calculated using the following formula:

$$i\_{average}=\left(1+i\_{total}\right)^{\frac{1}{11}}-1$$

The COUNTIF function is then used to calculate the number of students whose average increase per test is greater than or equal to the threshold of 10%. This shows that 15 students would get the prize.

The following checks have been performed:

* There are 100 total increases and average increases.
* $\left(1+i\_{average}\right)^{11}=(1+i\_{total})$ as expected.