HIED 801 Introductory Predictions Assignment: Week 10
Emily Lane





First, I applied a time series plot to view the overall trend in Institution A, B, and C over years 1-7. If the institution shows a clear increase or decrease, I could use the average yearly change to predict the next year.

Institution A showed a downward trend, therefore an overall decrease. Instution B showed showed an overall increase, and Instituion C showed neither an increasing or a decreasing trend.



To calculate a precise prediction for year 8 enrollment at Institution A, I can calculate the average rate of decrease between each year, then calculate the overall average decrease between years 1-7. Then I can subtract the average decrease from the latest data to show the new prediction.

I repeated this process for Institutions B and C. I took into account that institution B had an overall increase and instituion C had no clear trend.

**The predicted enrollment for year 8 for Institutions A, B, and C respectively are 3,359 6,051 and 2,847.**

Calculations and specific data can be shown in the table below:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Year** | **Institution A** | **Rate of change A** | **Institution B** | **Rate of change B** | **Institution C** | **Rate of change C** |
| **Year 1** | 3,679 | NA | 5,867 | NA | 2,852 | NA |
| **Year 2** | 3,608 | -1.97% | 5,974 | 1.79% | 2,901 | 1.69% |
| **Year 3** | 3,512 | -2.73% | 6,085 | 1.82% | 2,853 | -1.68% |
| **Year 4** | 3,601 | 2.47% | 6,097 | 0.20% | 2,812 | -1.46% |
| **Year 5** | 3,489 | -3.21% | 6,034 | -1.04% | 2,888 | 2.63% |
| **Year 6** | 3,472 | -0.49% | 6,101 | 1.10% | 2,795 | -3.33% |
| **Year 7** | 3,404 | -2.00% | 6,025 | -1.26% | 2,848 | 1.86% |
| **Total Avg Rate of change** | NA | -1.32% | NA | 0.43% | NA | -0.05% |
| **Predicted Change in Enrollment** | NA | -44.973 | NA | 26.156 | NA | -1.355 |
| **Year 8**  | 3,359 | NA | 6,051 | NA | 2,847 | NA |



To calculate the predicted amount of undeclared students in their first year for year 6, I took the average of year 1 – 5 values and got a predicted 1st year value of 1286 students.

To calcuate the predicted amount of undeclared students in their second year for year 6, I calculated the average proportion of students still undeclared in their second year. For example:

$$Proportion of students still undeclared in year 2=\frac{Year 2 Undeclared 2nd year}{Year 1 Undeclared 1st Year }=\frac{618}{1274}=48.5\%$$

I repeated this process for each of the years, and I found that the average amount of students still undeclared in their second year is 49.3% of the previous year. If I multiply Year 5’s 1st Year undeclared students by 49.3% I get a predicted 2nd year undeclared count for year 6 of 744 students.

I repeated this process for 2nd and 3rd year students for each of the years, and I found that the average percent of students still undeclared in their 3rd year was 25.8% of the previous year’s 2nd year undeclared students. If I multiply year 5’s 2nd year undeclared students by 25.8% I get a predicted 3rd year of undeclared count for year 6 of 133 students.

The total amount of predicted students in year 6 is the sum of 1st year predicted, 2nd year predicted, and 3rd year predicted, and it is 21,64 students.

Detailed work is show in the tables below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **% Change** | **Year 1-2** | **Year 2-3** | **Year 3-4** | **Year 4-5** | **Year 5-6 predicted** |
| 1st year to 2nd year | 48.5% | 51.1% | 48.9% | 48.9% | 49.3% |
| 2nd year to 3rd year | 26.1% | 25.7% | 24.9% | 26.3% | 25.8% |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Student Type** | **Year 1** | **Year 2** | **Year 3** | **Year 4** | **Year 5** | **Year 6 Prediction** |
| 1st Year | 1,274 | 1,493 | 1,102 | 1,060 | 1,507 | 1287 |
| 2nd Year | 694 | 618 | 763 | 539 | 518 | 744 |
| 3rd Year | 185 | 181 | 159 | 190 | 142 | 133 |
| Total | 2,153 | 2,292 | 2,024 | 1,789 | 2,167 | 2164 |



$$Expected donation=\frac{Total donated}{Number of donations }\*\frac{Time since last donation}{Average time betweeen donations}$$

The above statistic is a rough measure of today’s expected donation which calculates the average donation and multiplies it by a ratio of time since last donation. I’ll use a couple examples below to explain how the formula works.

**Example 1**: A person usually donates once a year, and on average donates 10,000 dollars per year. It has been 6 months since the last donation. The expected donation value is:

$$Expected donation=10,000 dollars\*\frac{6 months}{12 months}=\$5,000$$

**Example 2**: A person usually donates once a every month, and on average donates 500 dollars per month. It has been 45 days since the last donation. The expected donation value is:

$$Expected donation=500 dollars\*\frac{45 days}{30.5 days}=\$737.70$$

Even though the person in example 1 donates large donations at one time, they are not due for another donation for some time. The institution could take this into account for their budget. This donor is not due to be contacted by the university, however due to the magnitude of this donation, the university should maintain regular contact with this donor.

In example 2, the donor is overdue for their donation, so this could signify an instance where the institution should reach out to the donor and remind them that their donations are appreciated. As their donations occur regularly, monthly contact could be scheduled to ensure that this donor is feeling involved and appreciated.

This formula aligns with Dawes’ theory because it states:

“Dawes observed that the complex statistical algorithm adds little or no value. One can do just as well by selecting a set of scores that have some validity for predicting the outcomes and adjusting the values to make them comparable” (Penn, n.d.)

Selecting simple scores such as amount and frequency to predict expected donations does exactly this. Furthermore, this formula could be set up in an excel spreadsheet to become a self-correcting model thus furthering the simplicity.