Research Analyst Position Assignment

Education Insights

2024-10-04

## Data Definitions

The data set contained in the RetentionData.csv file lists the number of new students enrolled in each program at a hypothetical Alberta post-secondary in each fall term and the number of those students who were retained to the following fall term, broken down by national status of the student. Here are the definitions of the columns:

* **AdmitTerm** - The term in which the students were enrolled in their program.
* **School** - NAIT is organized into 7 schools, each of which are comprised of a varying number of programs. Here they have been labelled “School A” through “School G”.
* **Program** - The 3-letter alphabetical code that denotes each program.
* **NationalStatus** - Indicates whether the student is studying at NAIT on an international student visa (INTERNATIONAL) or is a citizen or permanent resident (DOMESTIC).
* **Enrolled** - The number of new students of a given national status enrolled in the given program in the given term.
* **Retained** - The number of enrolled students who remained enrolled in any NAIT program in the fall term of the following year.

## Part A

Describe how retention varies between schools, over time, and between domestic and international students. Submit a slide deck summarizing your findings and identifying any oddities, missing information, or questions for future analysis for this part. Your presentation should be 3-5 slides and consumable by a non-technical audience.

## Part B

Design and implement an algorithm to identify program/national status combinations that have a significant trend in their retention rates. Summarize this algorithm, outlining and justifying any statistical tests or methods used in detail (you may use either plain-language pseudo-code or well-commented code). Mention any assumptions made in the process. Report the combinations that have significant trends, describing the patterns observed for at least one example. Prepare a brief PDF summary of the algorithm, your findings, and identify any oddities, missing information, or questions for future analysis for this part. This should be a 2–3 page report for a more technical audience.

**Due Date**

* Due date: Monday, October 7th at 9:00am MST
* Deliver your response to [careers@nait.ca](mailto:careers@nait.ca)