FINAL CAPSTONE

Brent Bogan

# Overview

## Dataset

|  |  |
| --- | --- |
|  | For this capstone project, I am using the Education Statistics dataset from World Bank Open Data, which I discovered on Kaggle. The dataset holds internationally comparable indicators that describe education access, progression, completion, literacy, teachers, population, and expenditures, covering the education cycle from pre-primary to vocational and tertiary education. Unfortunately, because the dataset was unable to download from Kaggle, I had to go to The World Bank’s website in order to access the original database titled ‘Education Statistics – All Indicators’. However, when I attempted to download the database in its entirety, a popup appeared stating, “Your selection (79024356 cells) exceeds the maximum report limit (2500000 cells " or remove some variables to continue.” Because of this, I have to limit and preemptively cleanse the dataset in order to not exceed the report limit. The variables of the dataset are categorized into ‘Time’, ‘Series’, and ‘Country’. The rows are by country, alphabetically. The other columns in order are ‘Country Code’, ‘Series’, ‘Series Code’, and ‘YR’. Those columns are all categorical, specifically nominal, data types, apart from ‘YR’ being quantitative, specifically continuous. |

## Research Question

|  |  |
| --- | --- |
|  | What is the more advantageous environment for a child to better achieve a basic education in? |

## Hypotheses

|  |  |
| --- | --- |
|  | 1. There is a significant difference in the minimum proficiency level achieved for mathematics between rural and urban environments. 2. There is a significant difference in the minimum proficiency level achieved for reading between rural and urban environments. |

## Testing Method

|  |  |
| --- | --- |
|  | I’ll be conducting a t-test for each hypothesis on Python via Google Colab. |

## Affected Parties

|  |  |
| --- | --- |
|  | Parents who are looking to move to either environment and want to know if it’ll better their child(s)’s education.  Researchers wanting to know if environment plays a significant factor in the quality of basic education.  Governments’ focusing of where educational funding should be spent. |

*Kaggle:* [*https://www.kaggle.com/theworldbank/education-statistics*](https://www.kaggle.com/theworldbank/education-statistics)

## The World Bank: [https://databank.worldbank.org/source/education-statistics-%5e-all-indicators#](https://databank.worldbank.org/source/education-statistics-%5e-all-indicators)