Bureau of Statistics Data Analysis Competition 2023

based on

Multiple Indicator Cluster Survey

Simeon Chester

Faculty of Social Science

Department of Economics

University of Guyana

Bachelor of Social Science Economics

simeonachester@gmail.com

September, 26, 2023

Bureau of Statistics Data Analysis Competition 2023

based on

Multiple Indicator Cluster Survey

Introdution

This paper is based on an analysis of men and women between 15 to 49 years of age with emphasis on the highest level of school attended for the two groups.

Data and Descriptive Statistics

Data Collection

For this study, the dataset was based on a Multiple Indicator Cluster Survey (MICS6) conducted in 2018 by the Bureau of Statistics in collaboration with United Nations Children's Fund (UNICEF) with a sample size of 8700 households in Guyana distributed country wide in both urban and rural areas.

The datasets used for this particular study were:

1. wm_Women 15 to 49_Stat_Day Data Set for Surveyed Women in Reproductive age range 15 17
2. mn_Males 15 to 49_Stat_Day Data Set for Surveyed Men in Reproductive age Range 15 17
3. hh_Household_Stat_Day Data Set for Surveyed Households

Data Preprocessing

All data was processed using the Python programming language. The modules used in Python were seaborn, matplotlib and pandas. This document has also been typsetted using LaTeX.

For the men and women data, to get the region and area values, the cluster number and household number had to be searched for in the household dataset, with the region and area values extracted from the matched rows.

The total respondents per region for each region area had to also be calculated to get the averages of each school level in each region by area. However, before doing this, all null and not-available values had to be filtered so the total and averages were not skewed/distorted.

All data and source files used in this analysis have been uploaded to a public Github repository (Simeon Chester, 2023).

Final Processed Data

Tables

Region Number	School Level	Education Level per Region Area	Area Type	School Level Name
1	1	0.29	urban	Primary
1	2	0.29	urban	Lower Secondary
1	3	0.41	urban	Higher Secondary
1	4	0.02	urban	Higher
2	1	0.10	urban	Primary
2	2	0.41	urban	Lower Secondary
2	3	0.47	urban	Higher Secondary
2	4	0.02	urban	Higher
4	1	0.05	urban	Primary
4	2	0.19	urban	Lower Secondary
4	3	0.43	urban	Higher Secondary
4	4	0.32	urban	Higher
4	9	0.01	urban	Don't Know
6	1	0.11	urban	Primary
6	2	0.16	urban	Lower Secondary
6	3	0.59	urban	Higher Secondary
6	4	0.11	urban	Higher
6	9	0.03	urban	Don't Know
7	1	0.05	urban	Primary
7	2	0.18	urban	Lower Secondary
7	3	0.73	urban	Higher Secondary
7	4	0.04	urban	Higher

9	1	0.16	urban	Primary
9	2	0.19	urban	Lower Secondary
9	3	0.45	urban	Higher Secondary
9	4	0.19	urban	Higher
10	1	0.01	urban	Primary
10	2	0.19	urban	Lower Secondary
10	3	0.47	urban	Higher Secondary
10	4	0.33	urban	Higher
1	1	0.19	rural	Primary
1	2	0.32	rural	Lower Secondary
1	3	0.44	rural	Higher Secondary
1	4	0.05	rural	Higher
2	0	0.01	rural	Nursery
2	1	0.16	rural	Primary
2	2	0.30	rural	Lower Secondary
2	3	0.47	rural	Higher Secondary
2	4	0.07	rural	Higher
3	0	0.01	rural	Nursery
3	1	0.10	rural	Primary
3	2	0.29	rural	Lower Secondary
3	3	0.41	rural	Higher Secondary
3	4	0.18	rural	Higher
4	0	0.01	rural	Nursery
4	1	0.13	rural	Primary
4	2	0.22	rural	Lower Secondary
4	3	0.49	rural	Higher Secondary
4	4	0.16	rural	Higher

5	1	0.13	rural	Primary
5	2	0.22	rural	Lower Secondary
5	3	0.47	rural	Higher Secondary
5	4	0.16	rural	Higher
5	9	0.01	rural	Don't Know
6	0	0.00	rural	Nursery
6	1	0.20	rural	Primary
6	2	0.27	rural	Lower Secondary
6	3	0.41	rural	Higher Secondary
6	4	0.09	rural	Higher
6	9	0.02	rural	Don't Know
7	1	0.17	rural	Primary
7	2	0.25	rural	Lower Secondary
7	3	0.51	rural	Higher Secondary
7	4	0.07	rural	Higher
8	0	0.01	rural	Nursery
8	1	0.26	rural	Primary
8	2	0.15	rural	Lower Secondary
8	3	0.50	rural	Higher Secondary
8	4	0.08	rural	Higher
9	1	0.22	rural	Primary
9	2	0.29	rural	Lower Secondary
9	3	0.45	rural	Higher Secondary
9	4	0.03	rural	Higher
9	9	0.01	rural	Don't Know
10	1	0.03	rural	Primary
10	2	0.19	rural	Lower Secondary

10	3	0.61	rural	Higher Secondary
10	4	0.17	rural	Higher

Table 1

Table showing highest level of school attended by male respondents per Region Area

Region Number	School Level	Education Level per Region Area	Area Type	School Level Name
1	1	0.22	urban	Primary
1	2	0.20	urban	Lower Secondary
1	3	0.55	urban	Higher Secondary
1	4	0.03	urban	Higher
2	1	0.11	urban	Primary
2	2	0.22	urban	Lower Secondary
2	3	0.57	urban	Higher Secondary
2	4	0.09	urban	Higher
4	1	0.02	urban	Primary
4	2	0.12	urban	Lower Secondary
4	3	0.51	urban	Higher Secondary
4	4	0.35	urban	Higher
4	9	0.01	urban	Don't Know
6	1	0.11	urban	Primary
6	2	0.21	urban	Lower Secondary
6	3	0.52	urban	Higher Secondary
6	4	0.16	urban	Higher
6	9	0.00	urban	Don't Know
7	1	0.06	urban	Primary

7	2	0.17	urban	Lower Secondary
7	3	0.68	urban	Higher Secondary
7	4	0.08	urban	Higher
9	1	0.03	urban	Primary
9	2	0.23	urban	Lower Secondary
9	3	0.62	urban	Higher Secondary
9	4	0.12	urban	Higher
10	1	0.02	urban	Primary
10	2	0.11	urban	Lower Secondary
10	3	0.57	urban	Higher Secondary
10	4	0.30	urban	Higher
10	9	0.00	urban	Don't Know
1	1	0.21	rural	Primary
1	2	0.27	rural	Lower Secondary
1	3	0.47	rural	Higher Secondary
1	4	0.05	rural	Higher
2	0	0.00	rural	Nursery
2	1	0.15	rural	Primary
2	2	0.23	rural	Lower Secondary
2	3	0.52	rural	Higher Secondary
2	4	0.09	rural	Higher
2	9	0.01	rural	Don't Know
3	1	0.08	rural	Primary
3	2	0.26	rural	Lower Secondary
3	3	0.50	rural	Higher Secondary
3	4	0.16	rural	Higher
3	9	0.00	rural	Don't Know

4	1	0.09	rural	Primary
4	2	0.23	rural	Lower Secondary
4	3	0.48	rural	Higher Secondary
4	4	0.20	rural	Higher
4	9	0.00	rural	Don't Know
5	0	0.00	rural	Nursery
5	1	0.10	rural	Primary
5	2	0.19	rural	Lower Secondary
5	3	0.59	rural	Higher Secondary
5	4	0.11	rural	Higher
5	9	0.01	rural	Don't Know
6	0	0.00	rural	Nursery
6	1	0.17	rural	Primary
6	2	0.23	rural	Lower Secondary
6	3	0.51	rural	Higher Secondary
6	4	0.08	rural	Higher
6	9	0.00	rural	Don't Know
7	1	0.23	rural	Primary
7	2	0.19	rural	Lower Secondary
7	3	0.53	rural	Higher Secondary
7	4	0.05	rural	Higher
8	0	0.00	rural	Nursery
8	1	0.15	rural	Primary
8	2	0.19	rural	Lower Secondary
8	3	0.61	rural	Higher Secondary
8	4	0.04	rural	Higher
9	1	0.13	rural	Primary

9	2	0.28	rural	Lower Secondary
9	3	0.54	rural	Higher Secondary
9	4	0.05	rural	Higher
10	1	0.05	rural	Primary
10	2	0.19	rural	Lower Secondary
10	3	0.49	rural	Higher Secondary
10	4	0.27	rural	Higher

Table 2

Table showing highest level of school attended by female respondents per Region Area

School Level	Averages by Area			
	Urban Men	Rural Men	Urban Women	Rural Women
nursery	0.00	0.00	0.00	0.00
primary	0.09	0.16	0.06	0.13
lower secondary	0.22	0.26	0.16	0.23
higher secondary	0.50	0.46	0.56	0.52
higher	0.18	0.11	0.21	0.12
don't know	0.01	0.00	0.00	0.00

Table 3

Table showing highest level of school attended per each area by Gender

Figures

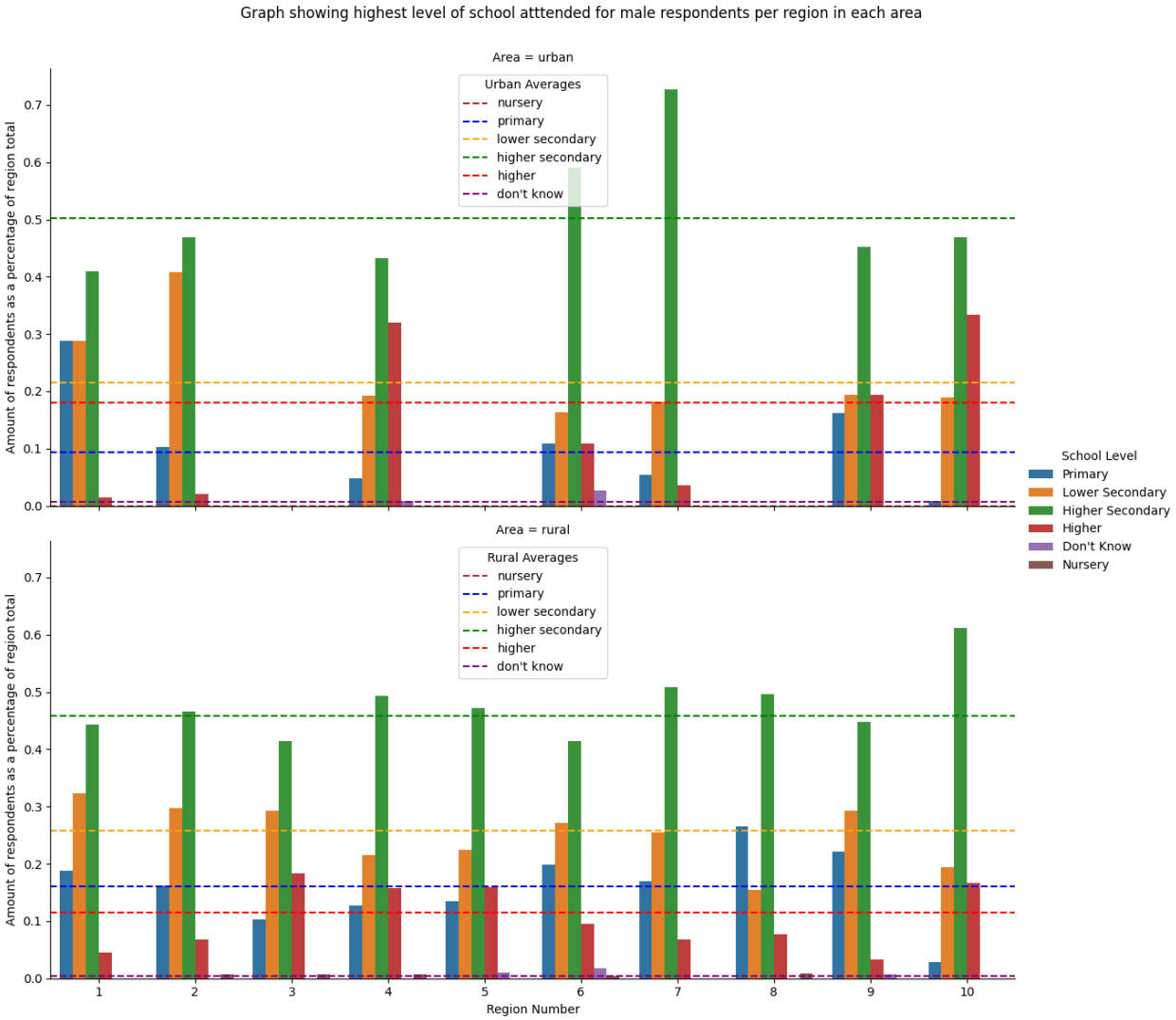

Figure 1

Graph showing highest level of school atttended for male respondents per region in each area

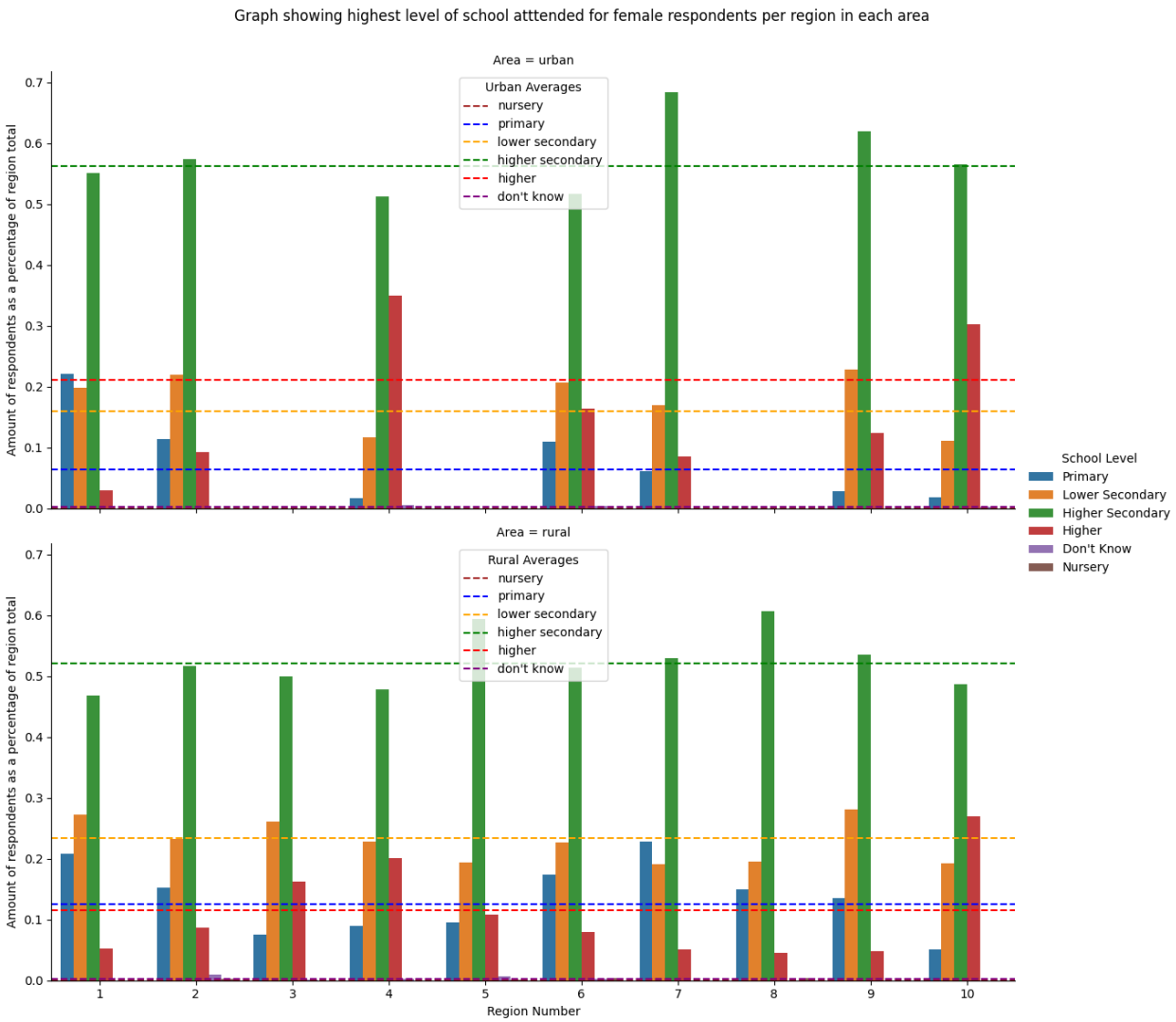

Figure 2

Graph showing highest level of school atttended for female respondents per region in each area

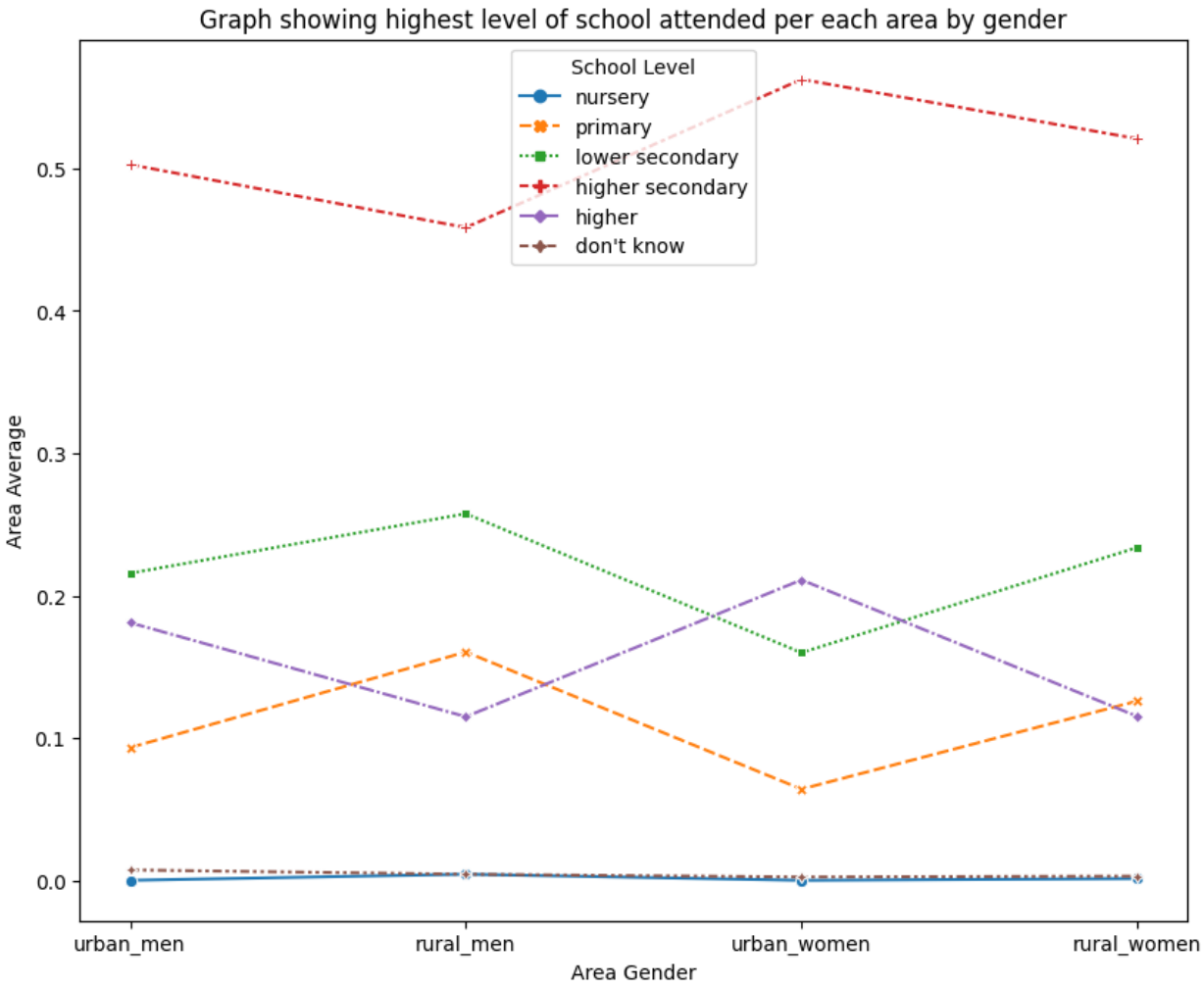

Figure 3

Graph showing highest level of school attended by area and gender per area total

Discussion

Female highest level of school attended per Area and Region

For both the urban and rural areas it was observed that the predominant school level was higher secondary with an average of nearly 50% of the respondents. In the rural area this was followed by lower secondary, primary and then higher with averages of 23%, 13% and 12% respectively. In the urban area, the higher secondary level was followed by higher, lower secondary then primary with averages of 21%, 16% and 6% respectively. This shows that in the urban area women are far more likely to attend institutions of higher learning after finishing their highschool education. This reason might be attributed to proximity of institutions from living area. This assumption can be supported by looking at the averages per region, where regions 3, 4 and 10 is observed to have the greatest portions of women in both rural and urban areas having attended a higher institution level.

Male highest level of school attended per Area and Region

For both the urban and rural areas it was observed that the predominant school level was higher secondary with an average of nearly 50% of the respondents. In the rural area this was followed by lower secondary, primary then higher with averages of 26%, 16% and 11% respectively while in the urban area this was followed by lower secondary, higher then primary with area averages of 22%, 18% and 9% respectively.

Limitations of analysis

Time has been a huge limitation in this analysis. While the code used to analyse the data was checked, extra time would have been used to further ensure data integrity. The code and by extension the data produced by the code has also not been peer-reviewed which further limits its integrity. This analysis also only focuses on a specific parameter i.e. highest level of school attended which makes the analysis relative myopic. If further analysis is done it would be interesting and wise to look at other paramters in the dataset, stratify and compare e.g. the actual education level of the respondent and age to see

whether

References

Simeon Chester. (2023). *Github Analysis Repository*. Retrieved September 26, 2023, from https://github.com/simeonachester/bos_data_competition_2023