Lindsay Sheppard

SOC-565: Demography

Demography Report 3

**Overview**

Urbanization is often considered the movement of populations from rural to urban areas; as a concept, it issomewhat difficult to define because there is no standard international definition of an ‘urban’ area or ‘urban’ population (Deuskar, 2015). Each country defines urban differently. To define a “city”, the most frequently used minimum threshold values amongst world nations are 2,000 or 5,000 persons (Deuskar, 2015). The operational definition for the current report will use the World Bank (2009) World Development Report’s approach, originally proposed by Chomitz, Buys, and Thomas (2005), who argued that rurality is a gradient and a multidimensional concept, not a dichotomy, with low population density and remoteness from large cities being the two most useful indicators of rurality.

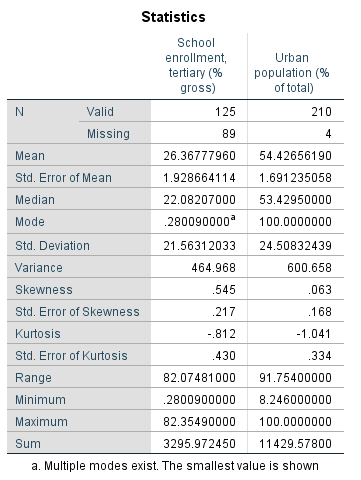
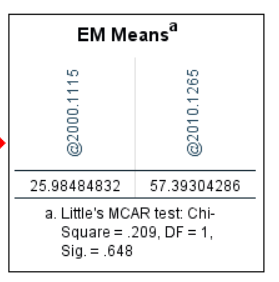
Based on this definition, many rural areas may experience reduced access to important resources such as education. Educational attainment is perhaps one of the most important factors influencing the social stability, economic success, and overall well-being of a location. Areas with more highly educated people have lower crime rates, lower rates of obesity and smoking, better overall health, higher incomes, and higher levels of economic development (Florida, 2018).

To understand the connection between urbanization and educational attainment, the data file “wdi.sav” analyzed the relationship between two variables: *@2010.1265*, or, *Urban population (% of total)* and *@2010.1115*, or, *School enrollment, tertiary (% gross);* the variables were referred to as *tertiary* and *urban*, respectively.It was assumed from the variable names that the data was collected in the year 2010.While the *urban* variable was rather self-explanatory, the *tertiary* variable contains information about how many people were enrolled in colleges and universities, regardless of an institution’s ‘urban’, ‘suburban’, or ‘rural’ designation. IBM SPSS Statistics (SPSS) version 24 analyzed the variables and provided results for the current report.

**Descriptive Statistics**

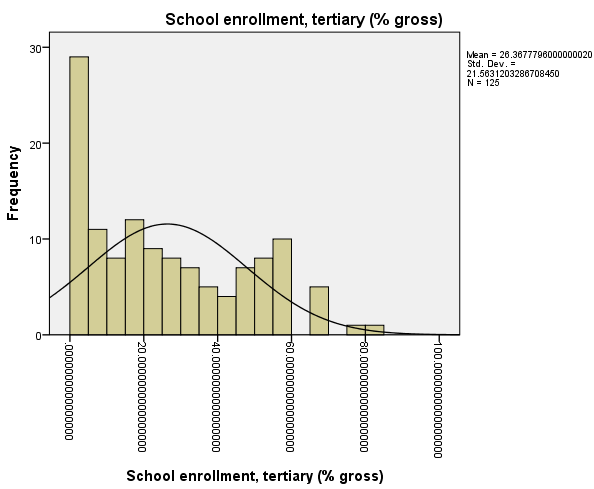
Using SPSS’s ‘Analyze’ function, a series of descriptive statistics was generated using the steps Analyze-> Descriptive Statistics -> Frequencies, which examined the two aforementioned variables independently. Figure 1 displays the results, which indicate that 89 entries are missing for *tertiary* but only 4 were missing from *urban*. By default, SPSS excludes system-missing values from analysis; deleting cases can result in a smaller sample and larger standard error. As the dataset is already small (N= 214), removal of the cases could return biased results. However, without metadata available on the dataset, imputation could be unwarranted or even detrimental to the analysis. With such a large amount of data missing from *tertiary* (42%), the imputation may render the results confirmative when they should not be, unless the data is considered Missing Completely at Random (MCAR) (Jakobsen, Gluud, Wetterslev, & Winkel, 2017).

The variables of interest were checked using Little’s MCAR Test (Figure 2). Significance was reported at 0.648, indicating that the missing values may be MCAR. If data are MCAR, either multiple imputation or ignorance of missing values are valid approaches (Jakobsen et al., 2017). As the current report focused on establishing a basic relationship between the two variables *tertiary* and *urban*, multiple imputation was not performed.

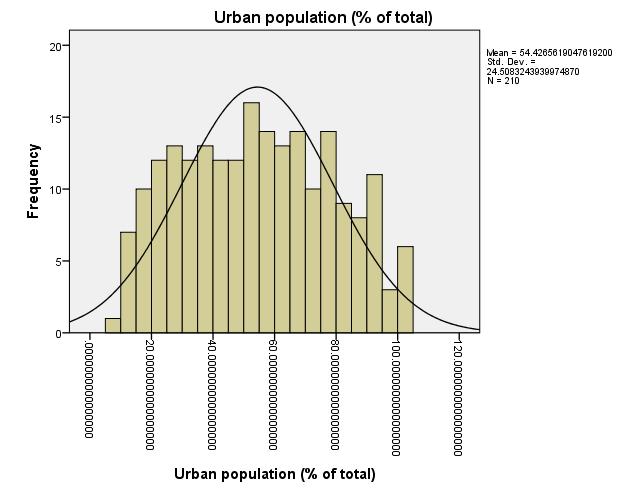
 

**Figure 1:** Descriptive Statistics for the variables **Figure 2:** Little’s MCAR test results *@2010.1115 (tertiary)* and *@2010.1265 (urban)*

Referring back to Figure 1, one notices that the average percentage of persons enrolled in a tertiary institution (26.4%) is less than half the average urban population percentage (54.4%), which is reasonable since not everyone who lives in the city will be enrolled in a tertiary institution. The median values for *tertiary* and *urban* were 22.08% and 53.43%, respectively. For the former, the median is somewhat smaller than the mean, with the reported mode even smaller (though, the variable is multi-modal.) The standard deviations of both variables are between 20 and 25 percentage points, indicating that the data points for both are spread over a large value range. Skewness values indicate that *tertiary* (.545) is slightly more skewed than *urban* (.063). For the latter, the median being so close to the mean indicates a more “normal” distribution, with smaller skewness and kurtosis values in concurrence). Figures 3 and 4 demonstrate the initial distribution of the two variables against a normal curve.

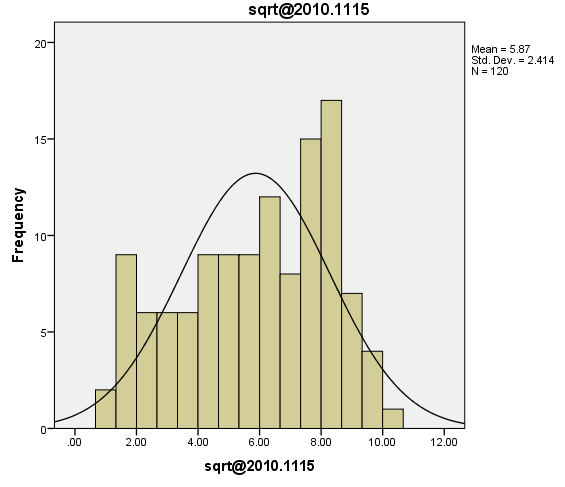


**Figure 3:** Initial histogram of *tertiary*

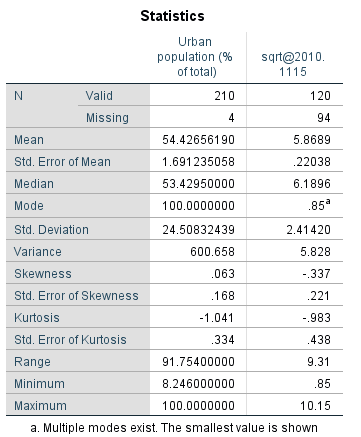


**Figure 4:** Histogram of *urban*

Skewness between 0.5 and 1 may be considered moderately skewed (Normality, n.d.).; moderate positive skewness may benefit from a square-root transformation (Tabachnick & Fidell, 2007). The transformed *tertiary* variable (*sqrt@2010.1115*) distribution may be viewed in Figure 5. The result is somewhat more normalized, at least visually. The reported skewness post-transformation was somewhat improved at -.337 (Figure 6).



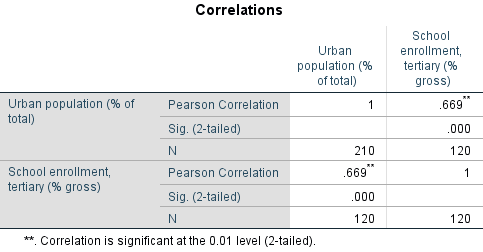
**Figure 5:** Histogram of the square-root transformed variable *tertiary* (sqrt@2010.1115)

****

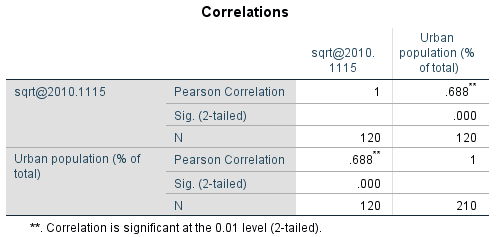
**Figure 6:** Descriptive statistics of the sqrt@2010.1115 variable. Due to only mild improvement in normalization using transformation, future analyses with the *tertiary* variable may or may not require transformation.

**Correlation**

A bivariate correlation with both the untransformed (Figure 7) and transformed (Figure 8) variable *tertiary* examined the effect of the transformation on the correlation. In either case, results of the Pearson correlation indicated that there was a significant positive association between tertiary enrollment and urban residence, with the untransformed variables (r=.669, p <.001) reporting only a slightly weaker correlation than the one with the transformed *tertiary* variable (r = .688, p < .001). Based on the results, tertiary school enrollment and urban living exhibit a statistically significant linear relationship; the variables are positively correlated, indicating that both tend to increase together, and the magnitude of the association is moderate approaching strong.



**Figure 7:** Pearson correlation for untransformed *tertiary* and *urban*



**Figure 8:** Pearson correlation for transformed *tertiary* and *urban*

**Discussion**

The particular variables analyzed in the current report are interesting because it is difficult to determine which one is the dependent variable. Similar to the case of the chicken and the egg, one may question which occurs first: do rural-dwelling people enroll in post-secondary education and subsequently relocate to urban areas, or are urban-dwelling people more likely to enroll in college because they are already located in an urban area? In this case, it is likely both, with a little more of the latter occurring due to proximity and expectations.

In fact, there have been several studies that suggest rural students are at a disadvantage; students in rural counties are only 85% percent as likely as their urban peers to enroll in higher education (Koricich, 2014). Rural students may not have the same knowledge about their college enrollment and financial options as urban peers, lack familial support or understanding, may be lower-income and less likely to attend for monetary reasons, and may have a lack of transportation or access to their preferred schools. Those that do enroll from rural areas may migrate to the city and remain there because that is where the jobs tend to be. On the other hand, students in urban or suburban areas are often visited by nearby colleges and universities, who tend to recruit them more than rural students (Gomes, Grotteria, & Wachter, 2018; Gettinger, 2019). They also are already familiar with city life and culture and may acclimate to the college environment better, grow up with the expectation of attending colleges or universities, and have increased access and opportunities due to their proximity as well as their socioeconomic status.

As rural America loses youth, rural areas become more economically depressed, become increasingly old, and experience ‘brain drain’. It is not that rural areas are undesirable; approximately one out of four urban residents would like to relocate, with a third of those stating they would like to move to a rural community (Picchi, 2019). The jobs and the opportunities are simply not there. The younger, working-age population remain in the cities, where higher paying jobs allow them to pay back their student loan debts and afford greater luxuries (Picchi, 2019). The youth left behind in rural areas are essentially ignored by tertiary institutions; rural schools do not receive the same recruiting efforts or have the same dual-enrollment programs that many city schools offer. This is detrimental to those youth who would like to better their lives by furthering their education, but do not have access to the information that they need to do so. Society suffers, then, from the loss of potential talent that could contribute to our overall economy and well-being.

**References**

Chomitz, K.M., Buys, P., & Thomas, T.S. (2005, June). Quantifying the rural-urban gradient in Latin America and the Caribbean. Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/8317/wps3634.pdf?sequence= 1

Deuskar, C. (2015, June 2). What does ‘urban’ mean? [Blog Post]. *World Bank Blogs.* Retrieved from <https://blogs.worldbank.org/sustainablecities/what-does-urban-mean>

Florida, R. (2018, October 31). The talent dividend in urban and rural areas. Retrieved from [https://www.citylab.com/life/2018/10/college-education-workforce-urban-rural-divide- jobs/574273/](https://www.citylab.com/life/2018/10/college-education-workforce-urban-rural-divide-%09jobs/574273/)

Gettinger, A. (2019, March 6). One reason rural students don’t go to college: Colleges don’t go to them. *NPR*. Retrieved from https://wamu.org/story/19/03/06/one-reason-rural-students-dont-go-to- college-colleges-dont-go-to-them/

Gomes, J.F., Grotteria, M., & Wachter, J.A. (2018, March 23). Rural America is losing young people: Consequences and solutions [Blog post]. *Wharton Public Policy Initiative.* Retrieved from https://publicpolicy.wharton.upenn.edu/live/news/2393-rural-america-is-losing-young-people-

Koricich, A. (2014, April 4). The effects of rurality on college access and choice. *American Educational Research Association.* Retrieved from <http://www.aera.net/Portals/38/Newsroom%20> %20Recent%20Research/The%20Effects%20of%20Rurality%20on%20College%20Access%20and %20Choice.pdf?ver=2014-04-04-074023-417

Normality testing: Skewness and kurtosis. (n.d.). Retrieved from https://help.gooddata.com/doc/en/reporting-and-dashboards/maql-analytical-query- language/maql-expression-reference/aggregation-functions/statistical-functions/predictive- statistical-use-cases/normality-testing-skewness-and-kurtosis

Picchi, A. (2019, January 21). Rural America’s “brain drain”: How student debt is emptying small towns. *CBS News.* Retrieved from https://www.cbsnews.com/news/rural-americas-brain-drain-how- student-debt-is-emptying-small-towns/

Tabachnick, B.G. & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.) Boston, MA: Allyn and Bacon.

World Bank. (2009). World development report: Reshaping economic geography. Retrieved from https://openknowledge.worldbank.org/handle/10986/5991