Date: 09 February 2020

To: Miguel Marca, Superintendent Bolivia School System

CC: Dr. Kathleen Campbell Garwood

From: Team 8

Re: Initial Findings on Bolivian Poverty

## Introduction:

Bolivia, a landlocked country located in South America, has an approximated population of 11.51 Million as of 2019. Since the early 2000s, the Country has experienced economic growth of approximately 4.9 percent each year due to increased exports. As a result, the poverty rate has been steadily declining, from around 59 percent to about 40 percent. Although the average rate of poverty has decreased, the factors that contribute to an individual or family being considered impoverished are still unknown. To answer the question of what factor/factors most contribute to poverty, we obtained survey responses from 837 Bolivian Students. The results of our analysis will be discussed in proceeding paragraphs.

## About the Survey:

* 837 Bolivian Students were surveyed
  + 499 were in the Fourth Grade.
  + 338 were in the Fifth Grade.
* Survey had 28 questions.
  + 11 Questions were Yes/No
  + 17 Questions had multiple responses.

## Limitations within our analysis:

Although we attempted to perform an inclusive analysis, we noted certain limitations within the data.

* 55 students did not respond to any of the survey questions. As a result, these students were excluded from the analysis as to not skew the results.
* 423 students did not respond to survey questions 23-28. Given that this was approximately half the number of polled students, we did not exclude these students from the analysis.
* 36 Students did not respond to questions regarding Father’s occupation or education level. We assume that these students are without a father figure in their household. Given that most, if not all, other survey questions were answered, we did not exclude these individuals from our analysis.
* 15 Students did not respond to questions regarding Mother’s occupation or education level. We assume that these students are without a mother figure in their household. Given that most, if not all, other survey questions were answered, we did not exclude these individuals from our analysis.

## Assumptions of our Analysis:

Using the aggregated responses from the Fey Alegria Survey, we identified six (6) questions that we believe would have a correlation with those students in poverty. From the six selected factors, we derived the following three assumptions:

1. Students who have no access to running water have a strong probability of being impoverished (Question 7). To help us understand the different dimensions, we chose to examine if they had a shower in their home (Question 12).
2. If there are 3 or more occupants per room in a household (Question 5 & 6), or if a student shares their room with half or more of the occupants living in their home (Question 9), they are more likely to be impoverished.
3. Students who only consume one or two meals per day (Question 10) are more likely to be impoverished.

## Assumption 1: Access to Water

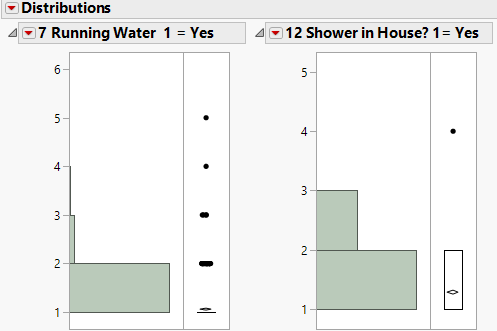


Figure : Histogram of Data Set for Running Water & a Shower

Both survey questions relating to assumption 1 required a Yes (1) or No (2) Response. Out of the 784 responses, only 774 students answered both questions 7 & 12. Figure 1 shows 2 Histograms that provide an overview of the responses based on student view. There are 8 students (0.01% of the total population) who responded with an incorrect option or skipped the question. To ensure that these responses did not skew the analysis, we removed them from the calculation of this assumption.

After removing the students who didn’t respond or provided an incorrect answer to either question 7 or 12, we re-examined the output. Figure 2 provides the corrected histogram results on a student basis, which is a total of 766 student responses. Based on these results we can observe that most of the students do have running water and a shower available to them. Although there are students who do not have access to running water or some with access to running water, but no shower.

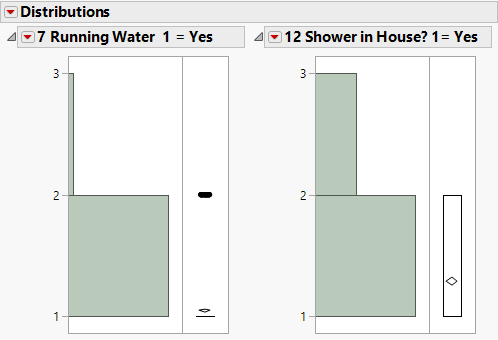


Figure : Histogram of cleaned Data set for Running Water & a Shower

Continuing looking at the survey results on a student view, not taking schools or classes into consideration, the Mosaic plot in figure 3, provides an overview of the number of students who have access to running water and a shower. With a chi-square independence test of 21.468, and a p-value less than 0.0001, we can conclude that having running water and a shower within your home are correlated. This would make sense, because a home cannot have a shower if it does not have access to running water.

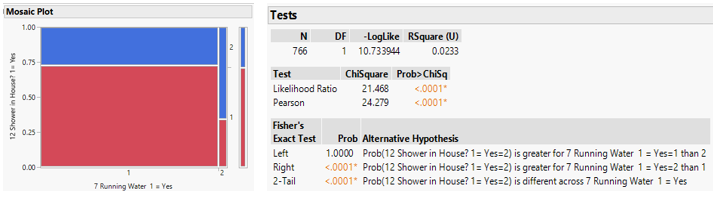


Figure : Mosaic Plot of Assumption 1

Based on these results we can determine that there are 3 Levels of indication for a student living in need with regards to running water:

* Level A (Represented by 0) – Students with no access to running water or a shower.
* Level B (Represented by 1) – Students with access to running water but no shower.
* Level C (Represented by 2) – Students with access to running water and a shower.

The different levels of impoverished with 0 being high and 2 being low, represents that 69.5% probability that the students are not living in impoverished conditions based solely on access to running water and a shower, with 27.5% of no shower access as an indicator for impoverished conditions presents and 3% of no access to running water or a shower which is a strong indicator that they are living in impoverished conditions.

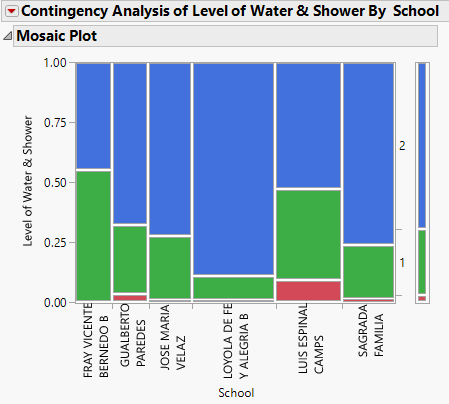


Figure : Mosaic Plot of Levels of Water Access by School

On the student level no consideration of schools, we can state that 30% of the student’s survey have an indicator of impoverished living condition with no access to running water. If we want to group the students based on the likelihood of the severity of their impoverished living conditions, not having a shower & no running water is 3% of the students surveyed have a higher chance of living in impoverished conditions.

Continuing with the grouping created at the student level we can indicate which schools are likely to be impoverished using a Mosaic Plot. The Mosaic Plot (Figure 5) shows that no school is without impoverished students, but Luis Espinal Camps & Gualberto Paredes have a higher likelihood of impoverished students based on the indicator of access to water level and severity with access to a shower.

## Assumption 2: Household Count

Given the data from survey questions five (5), six (6), and nine (9), we chose to analyze the students based on the school they attended. First, we looked at the average number of rooms in each house based on school. As stated previously, we believe that those who have the least number of rooms are the most impoverished. From our analysis, we saw that students at Jose Maria Velaz School had the least average number of rooms in their house (3.85), while those at Loyola De Fe Y Alergria B had the largest. (See Figure 6 for graphical output).

Figure 6: Graph Output of Average Number of Room per Household based on School.

In addition to examining the number of rooms in a house, we also looked at the number of people that the students share’s a room with. We felt that the greater the number of students in each room, the higher rate of poverty. Like Figure 6, Figure 7 shows students who attend Jose Maria Velaz have the greatest number of people sharing a room.

Figure 7: Graph Output of Average Number of Individuals in a Room based on School.

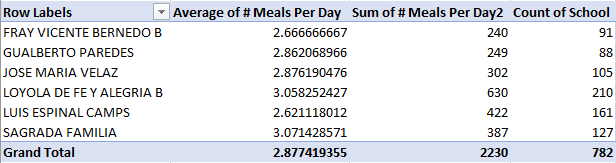
Based on the above two factors, we can conclude that the students who attend Jose Maria Velaz are the most impoverished, because they have the lowest average number of rooms and the highest average number of individuals sharing a room.

## Assumption 3: Access to Food

Prior to digging deep in our analysis, we first examined the distribution of the number of meals per day for all six of the schools. The distributions for 5 of the 6 schools appeared to be normal, with the outlier, Loyola de Fe, being left-tailed. The average number of meals for these six schools can be observed below in figure 8.

When we considered student’s access to food and whether they eat prior to school, we chose to analyze the information based on school. First we aggregated the information in to a pivot table (Figure 6), where we examined the average number of meals per day. This information ranged from a low average of 2.6 meals per day to the high of 3.07. Based on this analysis, we concluded that the students at Luis Espinal Camps School are the most impoverished, based on the assumption that average number of meals correlates to poverty.

Figure 8: Pivot Output of Meals per Day by School



Based on our analysis, we concluded that students attending Luis Espinal Camps are the most impoverished, due to the number of meals they eat per day, with the students attending Loyola De Fe Y Alegria B assumed to be the least impoverished.

## Conclusion:

Although we limited our analysis to a few specific areas, we can observe certain similarities between our three assumptions. Specifically, assumption 1, relating to water, and assumption 3, relating to food, both have the same conclusion. Both analyses showed Luis Espinal Camps as the most impoverished school, based on the respective factors. Assumption 2 produced a different result, with Jose Maria Velaz being the most impoverished. Further, Luis Espinal Camps appeared to be one of the better off schools in this analysis.

## Future Analysis:

Given this preliminary assessment only addressed with six (6) out of twenty-eight (28) questions, we may not have received the full picture for this data. Further, we first chose to examine our factors based on school and not on the individual student level. As we continue our analysis utilizing the same data set, our future analysis will focus on identifying additional areas that we believe will contribute to poverty. As we continue to work through the information, we will group the students into different rankings, using a cluster analysis, and ultimately be able to conclude on who we believe are the most impoverished students. In doing this, we hope to be able to identify which survey questions hold the most weight and meaning in determining which students are impoverished. Although the rates of poverty are declining in Bolivia, the use of this information would aid Fey Algria and other non-profits in their battle against poverty.