Determining the Effectiveness of Computer-Based National Exams

By Leonardus Edwin Hartono Limbri

April 6, 2019

I. Introduction

1.1 Background

Indonesia is the largest archipelago in the world with 18,000 islands in total and over 200 million people reside. The city has 34 provinces with national schools spread out all over the countries (One World Nations Online, 2019). As stated by the United Nations (2014), Indonesia is classified as a country with a developing economy and therefore has a high economic growth rate compared to a developed one (Pettinger, 2017).

Education could affect the economic growth rate by affecting the human capital of a country (The World Bank, 2013), which therefore is an important aspect for a developing country such as Indonesia. One way to measure the level of education in the country is through a standardized National Exams, which have been done for years in Indonesia. These results are then used by the government to gage the education credibility of each province and determine the budget on education.

1.2 Problem

From 2015, the exams are done using a computer, which discourages student to cheat on the exam. This results in a declining trend of average score on a national scale; does this means that there are no more cheaters among the students? This question is intended to be answered in this paper.

1.3 Objective

By determining effectiveness of this type of national exams, the government could benefit by having the data regarding these exams to be reliable source to measure the level of education of the country, as described in the background on how important education is to a developing country such as Indonesia. However, this study could be used to other similar country to help incentivise the government to promote education in the country.

II. Data

2.1 Data Description

GRP is short for the Gross Regional Product, which is usually used in economic study to determine the economic level of certain region. In this study, I will be using the GRP per capita per province of Indonesia. Moreover, National Exam average score are just the average test score from a certain province. However, not all students went to the exams that is optional (such as Chatolic and Protestan tests), but all needed to undergo the same test with subjects ranging from Indonesian Language (Average_IL), Natural Science (Average_NS), Social Science (Average_SS), and Religion (Average_Religion). Lastly, it needed to be taken into account that the data used is from 2017.

2.2 Data Sources

- Indonesia' GRP per capita per province which is scraped from Wikipedia (https://en.wikipedia.org/wiki/List_of_Indonesian_provinces_by_GRP_per_capita).
- Indonesia's National exam average per provinces, which are scraped from the official websites by the Ministry of Education and Culture of Indonesia (https://puspendik.kemdikbud.go.id/hasil-un/).

2.3 Data Cleaning

Since the data that needed from the Wikipedia is only the GRP per Capita and its respective province, other columns are dropped, that is listed as below:

Columns Dropped	Reasons
Rank per Capita	Index used by Wikipedia to rank the GRP, which is not needed in this study.
Region	Since we focus on each province, we could neglect the region.
GRP Nominal in USD	Similar feature with the GRP per Capita
Nominal GDP Per Capita in IDR	
PPP Per Capita in USD	
GRP Nominal in IDR	
GRP PPP in USD	

Table 2.1

Since the data from the Ministry of Education and Culture of Indonesia are in Indonesian Language, and the province's name is all in Indonesian Language as well, I need to translate each of the province name back to its English, as what the GRP data showcase.

As mentioned previously, not all students are required to take the Catholic and Protestan religion test, therefore, I've dropped these datas as well, as no sufficient data is recorded.

The National Exam data also shows the number of test taker in the respective region, which is also dropped, because since it's all more than thousands, we could assume that the average score is enough to represent the overall distribution.

I also dropped/altered some of the rows as listed below:

Rows	**Dropped/Altered**	**Reasons**
North Kalimantan (Whole Row)	Dropped	No data reflected in the Nominal GRP data.
South Sumatera Column: Average_IL	Altered	No data given, therefore altered so that the means Total Average doesn't change.
Maluku Column: Average_Religion	Altered	
North Maluku Column: Average_Religion	Altered	

Table 2.2

III. Methodology

3.1 Target Variable

The target variable in this study is the Nominal GRP per Capita, the idea is that if the average test score data is regressed with the Nominal GRP per Capita and shown to have a positive coefficient (that is they both has a positive correlation), then the Computer-Based National Exams are effective enough, since it reflects the fact that

a higher level education (that is in this case higher test score) reflects a higher GRP per Capita or its respective level of economy.

3.2 Machine Learning

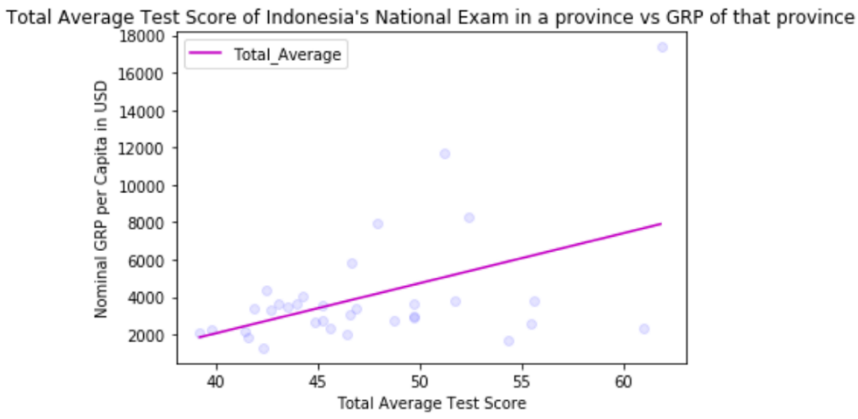
I've used a Regression algorithm which is a type of a supervised machine learning. Specifically, I've used the Linear Regression algorithm. I used this because it could reflect the correlation between the two datasets, namely GRP per Capita and the Average Test Scores. I also used this to find the overall trend of the datasets not to predict the model, that describes the level of error shown.

3.3 Methods

Firstly, the average test score is plotted with the Nominal GRP per Capita for its each respective province. Which then could be seen, each coefficient and each overall trend of the regression. If all of the regression has a positive coefficient with an overall trend of up, then the computer-based national exam is effective enough, in the case that there are no cheaters.

IV. Results

4.1 Total Average Test Score vs Nominal GRP per Capita



Coefficient = 267.7571107,

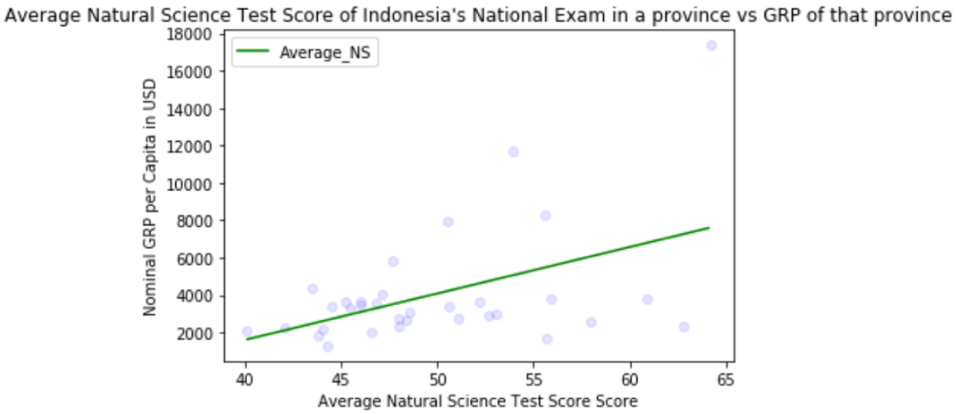
With a positive coefficient, it results in an upward trend, therefore this result corresponds to the fact that higher education level implies higher economic level (in this case GRP per Capita).

4.2 Average Indonesian Language Test Score vs Nominal GRP per Capita

Coefficient = 198.09922167,

With a positive coefficient, it results in an upward trend, therefore this result corresponds to the fact that higher education level implies higher economic level (in this case GRP per Capita).

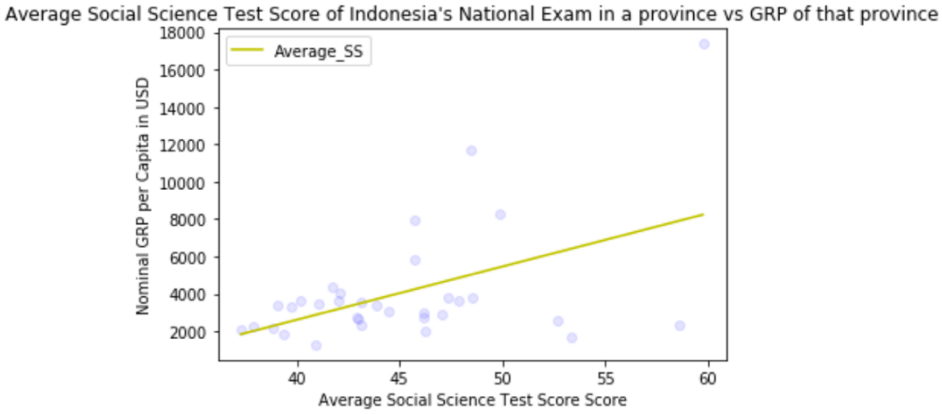
4.3 Average Natural Science Test Score vs Nominal GRP per Capita



Coefficient = 248.05090117,

With a positive coefficient, it results in an upward trend, therefore this result corresponds to the fact that higher education level implies higher economic level (in this case GRP per Capita).

4.4 Average Social Science Test Score vs Nominal GRP per Capita



Coefficient = 284.12356779,

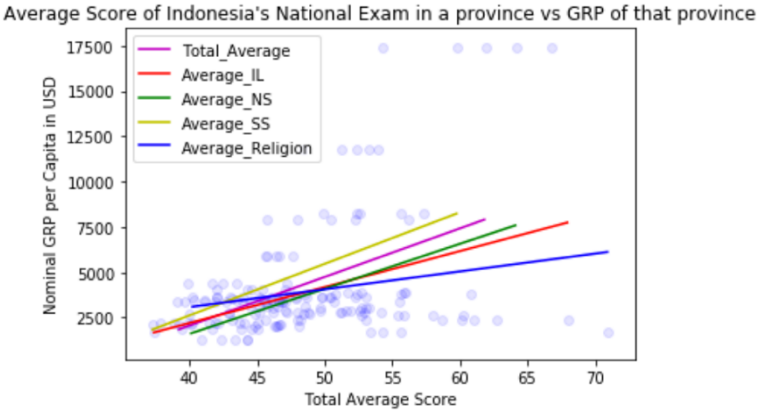
With a positive coefficient, it results in an upward trend, therefore this result corresponds to the fact that higher education level implies higher economic level (in this case GRP per Capita).

4.5 Average Religion Test Score vs Nominal GRP per Capita

Coefficient = 98.25103321,

With a positive coefficient, it results in an upward trend, therefore this result corresponds to the fact that higher education level implies higher economic level (in this case GRP per Capita).

4.6 All Regression Plotted Together



As shown from above plots, it all responds the same result, that is a positive coefficient and an upward trend.

V. Discussion

As the results suggest, test scores are positively correlated with the GRP per capita for each province. This tells us that it supports my first hypothesis that if it does positively correlate, then by increasing the quality of education, one could have a higher test score, which will result in an increase in GRP. This study assumes one particular year (2017) due to lack of data in computer-based national exams. Moreover, there are some instability during the first time this system was implemented, that is not all schools have access to computer, therefore the data doesn't represent all data required during year before 2017.

One thing that is important to notice is that, there is a significant differences between the scores, where some provinces are doing great by having the average score above the hurdle for passing the exam (50), there are a lot of unprivileged students that due to their economic condition, unable to have the best quality of education.

VI. Conclusion

The average test scores do reflect the province's GRP per capita, that is it correlates each other positively, with a positive correlation. Therefore, we can conclude that the computer-based national exam is an effective way compared to prior method of examination (by paper), where students are known to cheat along the way.

VII. Recommendations

Since the data used in this study is only for a year 2017, it is good to consider repeating the study when more data is available in the future, to get another look at different level of trend to see. Moreover, using the same data ones could also create a clustering of which regions are in need of attention by the government to improve their education services.