VR schools

**Which topic did you choose to apply the data science methodology to? (2 marks)**

Education (I just realised it should be emails, hospitals or credit cards, but I am way more passionate about this topic :D)

**Next, you will play the role of the client and the data scientist.**

**Using the topic that you selected, complete the Business Understanding stage by coming up with a problem that you would like to solve and phrasing it in the form of a question that you will use data to answer. (3 marks)**

**You are required to:**

1. **Describe the problem, related to the topic you selected.**
2. **Phrase the problem as a question to be answered using data.**

**For example, using the food recipes use case discussed in the labs, the question that we defined was, "Can we automatically determine the cuisine of a given dish based on its ingredients?".**

1. Poor people do not have access to great education like wealthy people do, thus keeping them poor. If poor people could have access to world class materials and access to VR technology, they could have a classroom inside their own home.
2. Can providing VR tech to low-income areas improve education?

**Briefly explain how you would complete each of the following stages for the problem that you described in the Business Understanding stage, so that you are ultimately able to answer the question that you came up with. (5 marks):**

1. **Analytic Approach**
2. **Data Requirements**
3. **Data Collection**
4. **Data Understanding and Preparation**
5. **Modeling and Evaluation**

**You can always refer to the labs as a reference with describing how you would complete each stage for your problem.**

1. Analytic approach: We need to find data that indicates if VR technology is correlated with an increase in academic performance, since grades could result from better quantity and quality of education.
2. Data requirements: We need data that shows if VR technology can be correlated with academic performance. We also need data on the poorest regions where the most impact could be made if this is the case.
3. Data collection: We could use the already available data from research by government and universities. We could use government census data to find how many people live in the lowest income areas.
4. Data understanding and preparation: Most data are already sorted in csv files, so we could use it as is to find correlation between academics and VR education, then use the data about population density in poor areas to predict the impact on academics in that those areas if the technology is deployed.
5. Modeling and evaluation: We could use descriptive analytics to find if there is truly a correlation and tune different models to find the best suited model. Then we could visualize it using a suitable graph/table. If there is a correlation we could use predictive analytics to estimate the impact of this strategy on poor areas.