Categorize each of the following research questions as "good" or "bad", and provide alternative formulations for the bad ones. Save your responses in a document of some kind, submit a link below, and discuss your reasoning with your mentor.

1. What is the 1994 rate of juvenile delinquency in the U.S.?

Good! This could easily be expressed in a mathematical way. I assume Juvenile delinquency has a specific definition and that there are certain data sources and statistics at national level regarding this topic.

1. What can we do to reduce juvenile delinquency in the U.S.?

This is a perfect example of countermeasure policies. For doing a countermeasure analysis, we need to model the dependent variable (here a mathematical representation of juvenile delinquency) and identify contributing factors. Some factors encourage the dependent variable while some discourage. Countermeasures focus on strengthening the discouraging factors and eliminating the positive factors.

1. Does education play a role in reducing juvenile delinquents' return to crime?

Good! We are focusing on a specific variable “Education” which could probably be explained in certain terms and categories, e.g. elementary school, middle school, 9th grade, high school grad, college, etc. A sample data can be accessed and broken down into different subsamples based on the Education level. Certain statistical tests such as t-test and ANOVA could be applied to measure the difference in the dependent variable (here identified as “return to crime among young people”.

1. How many customers does AT&T currently serve in Washington, DC?

Good! Data source is available probably through AT&T services. The research focuses on a specific mathematical expression (number of customers).

1. What factors lead consumers to choose AT&T over other service providers?

Looks like we need to model the propensity of customers towards AT&T. Usually companies like this do periodic consumer surveys which reflect certain aspects of the market including socio-economic and demographic attributes, personal attitudes, and several other questions that could be used in market segmentation for their products. This is a perfect data source that can be used to model AT&T usage based on several attributes. Such modeling leads to identification of significant contributing factors.

1. How can AT&T attract more customers?

Same as No. 6.

1. Why did the Challenger Shuttle explode?

So complicated. As a person with no specific knowledge on this topic, First I need to review different studies and cases regarding shuttle explosions, to grab an idea about the potential contributing factors. The factors then need to be thoroughly checked in the specific case of challenger shuttle to find the reason or reasons.

1. Which genes are associated with increased risk of breast cancer?

First, we need to do a thorough literature review to see what the contributing factors are. Then we can use certain medical datasets to model breast cancer (either as a binary variable or as a continuous frequency variable) based on certain genes. Lots of interaction combinations with age, gender, and other personal attributes should also be considered. This is a complicated model.

1. Is it better to read to children at night or in the morning?

First, we need to define what “better” means here. There should be a metric to compare the results of reading. For example, we may compare based on what they learnt through the reading, and this could be on a short-term or long-term basis. Once the metric is defined, we can do a two-sample t-test to see whether the metrics’ differences in the two samples are statistically significant or not.

1. How does Google’s search algorithm work?

Bad! This is a very vague question. I don’t understand what exactly it means. Should we look into the procedure from a user’s perspective or from a computer programmer’s point of view. From a user’s perspective, we need to grab a general picture of google’s functionality without knowing the coding details (very much like what a manual does). From a programmer’s perspective, we need to dig into the details of all algorithms that are used in the algorithm structure.