"By submitting this homework, I attest that the homework is my own work and I have not copied any text verbatim from any online or offline source, or from other students (past or present). I also attest that I have not used any paraphrasing tool or any homework solution websites. In addition, I attest that any information taken from course slides and handouts or any websites is properly attributed and referenced. I understand that any violation of the UAB Academic Integrity Code for any portion of the homework will result in a zero score for the entire homework and will be reported to the academic integrity coordinator per university policy."

1. With the Massachusetts health insurance dataset situation, the issue here was unique sensitive information being made public. Each variable alone would make it highly unlikely for someone to track someone down. All three together narrows down the list exponentially. Cross-referencing this data with another dataset with just one of these variables and their names will narrow down the list even more if not pinpoint the individual.
   With the Netflix Prize Database situation, due to only the names being replaced with random IDs, all other information remains and that information is attached to the IDs, the same information that is attached to the names in the imDB data. Cross-reference the data, match IDs of Netflix to the other information in imDB data.

2. The Airavat framework is ran on the cloud infrastructure. When the data provider uploads their data on Airavat, Airavat runs the code sent from outside, checking to see if it is malicious, while protecting the privacy of the data, The code can be audited, which is hard to do, or it can be confined. The MapReduce can be split into untrusted mapper and trusted reducer, then mandatory access control, which prevent leaks through storage channels like network connection, and differential privacy, which prevent leaks through the output of the computation, is implemented.
   Differential privacy is used to prevent leakage of data through extreme mapper output values. It makes it where the range mapper outputs must be declared in advance. This helps see how much does a single input influence the output. It also determines how much noise is added to outputs. Later on, JVM was modified to enforce mapper independence so each object is assigned an invocation number so it prevents reuse of objects from previous invocation.