1. Provide an example for how each type of anomaly in your data tables (semantic, syntactic, coverage) can impact the interpretation of the relationship between a response variable (Y) and a set of predictor variables (X).
2. How can data cleansing pipelines potentially harm the mapping between the "real world" and your data table?