Security Governance is deeply tied into data science. Governance as defined in the article is the creation of a cross-department committee focused on the intersection between data security and business strategies. It’s basically asking what data do we need for our business and how do we protect it?

Risk Assessment is the identification of assets which need protecting. This is, again, an intersection between data collected and the act of protecting it. The difference between Security Governance, where a Data Scientist would take a lead role, risk assessment is more of an analytical role. Using the formula in the lecture of *Asset \* Risk = Threat Level*, risk assessment immediately becomes an optimization problem.

Security Awareness also falls within the purview of Data Scientists. We can recognize that people are the weakest point of security and the IT Security team can implement several techniques, such as the principal of least privilege, to help avoid a massive attack. A Data Scientist can help direct the Security team by pinpointing specific employees who are abnormally weak with respect to technological security. Data Scientists can also research the trends of attacks so there can be a pre-emptive development of security features.

The final model I’ll address for this assignment is Disaster Recovery. Ensuring constant access to an up-to-date database and avoiding any service interruptions is just an ETL process on a massive scale. The first is the process of analyzing data – we don’t use the original dataframe for all our analysis; we often create a duplicate dataframe to avoid making an irreparable change to the original dataset. That’s the extraction. The transformation occurs when we make changes to the new dataframe. And finally, we load the changed dataframe into a target database.