Sean Morton
MDS 395
Reading Assignment 1

Mechanistic Data Science is the use of mathematical analysis to discover patterns and relationships between variables in the real world. With MDS, data on a real-world phenomenon or mechanism is collected into a large database with a variety of parameters. In the case of the Moneyball problem, Billy Beane and crew gathered metrics on players' batting like OBP and SLG (among other metrics) to understand the factors that produce a team with a high number of runs scored.

Once data is gathered, dimensions of the data are reduced to find the most relevant and novel relationships between data. Data is "cleaned" to remove irrelevant information and properly format charts and tables. The most important step is to find the relationships between variables, which is often carried out using either regression analysis or neural networks. Conclusions can be gathered, or the regressions and neural networks can be refined via the scientific method.

Three main types of problems exist in Mechanistic Data Science:

1. Data-driven problems without underlying scientific principles. The prime example that we talk about in class is the relationship between diamond properties and their prices.
2. Problems with minimal data and minimal scientific guiding principles. AlphaGo was trained on how to play games like Go and chess without the raw numerical data provided by the diamonds problem, for example, and without any "science" behind winning chess.
3. Problems where the science is known, but the parameters of the problem are unknown. This summer, Owen Huang's project on the spin of ping pong balls was an example of this: the motion and spin of spheres in air is known science, but the value of spin is unknown until the computer vision code passes the video of the ball through a FFNN.

Principles of Mechanistic Data Science can be used to understand physical phenomena. In the category of materials design, one application of the scientific method was when Professor Fleming carried out a durability test with his daughter on different mixtures of ice and other substances. The finding was that the ice that behaved most like a composite material (fibers in a matrix) had the highest durability. While simple, this experiment relates to an MDS framework for researching more durable tires: variables of temperature, filler quantities, and $\tan(\delta)$ are analyzed in industry to find the parameters that create the best long-lasting tires.

More examples of MDS throughout history are provided in the text, including Newton's studies on inertia, forces and reaction forces, as well as Galileo's experiments on gravitation. These examples can help motivate our studies as MDS students on how to analyze physical phenomena and create mechanistic explanations for the phenomena.