

Individual Project Ideas

Idea 1

Baseball has become the most statistical of America's sports. Every facet of the game is now measurable. I intend to create my own model for team wins, in order to learn more about the methods used by Major League Baseball (MLB) teams. It should utilize readily available data for a given team in order to predict the team's record.

My plan is to research various statistical models to find an appropriate one that can be used for this problem. It should take in statistical variables for a given team and output an approximation of their predicted wins. I assume the model will be more complicated than simple linear regression. If it is not, then I will shift my focus to modeling individual player's performance over time, which is considerably more difficult due to the variability of a single player. The variables can include factors like walks, runs scored, assists, Wins Above Replacement (WAR) and other statistics used to measure team performance. Explanations of these can be located on Baseball-Reference.com. I will ignore parameters like salary and attendance since I want this model to be applicable to any MLB team. Additionally, since it would make my model too cumbersome if I used all the available metrics for a team, any I do not use will be considered constant. Since most historical baseball data is very easily available, the sport's statistics have been exhaustively recorded, I should be able to locate any necessary data fairly easily online.

Idea 2

We live in a rapidly changing world. Due to shifting climates and global warming countries have had to switch sectors of their economies to be more climate-friendly. For instance, coal plants are being traded for solar panel farms. This has become controversial since people will lose their jobs in this transition. In order to answer questions about how these shifts will affect the economy, I will be modeling an economic shift to see its outcomes. I want to find out if the economy will be adversely, positively or neutrally affected by this change. In order to isolate this occurrence, I will be focusing on the shift from coal plants to solar panels.

The parameters used in this study will be economic: employment, unemployment, average² salary, jobs created, jobs lost and any other factor that is used in measuring the economy. Since the panels cannot usually be put in the same community as where the coal economy was based, I will assume constant populations across the two communities as well as ignoring any other community-specific parameters. The model will probably consist of two differential equations for the two communities, each measuring an economic growth factor like GDP (or some economic measure that is more specific to smaller populations). This way, it will be possible to see the change in the two communities over time. If this proves to be too complicated, I will construct two differential equations detailing jobs lost and jobs created in the two communities.

Idea 3

Bees are essential to the planet. Without them, there would be decreased pollination and many foods would become scarce or disappear. Currently, bee populations are decreasing. There are numerous theories about the reason for this and I would like to explore them. In order to see what factors affect bee population, I will model their population to see what elements might contribute to the current trend. The elements will be based on the various theories on the population decline. In that way, some light will be shed on what is causing the bee's deaths.

The model will use parameters like pesticide use, climate change, rate of deforestation, growth in competing species, rate of spread for any relevant bee-diseases and other factors to create the model. These will be quantified into rates of change, so that the model can output bee population over time. Factors specifically related to what bee-keepers spend on their hives will be ignored since it is assumed that all bee-keepers treat their bees the same. That assumption is made so that the model can be for all bee populations and not just specific ones for a given farmer. The model will be a differential equation for bee population over time, using the aforementioned parameters. It will be used to test the different parameters and see what may be causing the decline in their population.