Kyle Weng

IDS 126

Final Project Proposal

Introduction

My favorite part of my high school experience was the time I spent participating in the FIRST Robotics Competition (FRC) via my school's robotics team. Each season, FIRST (the organization behind the competition) unveils a thematic "game," in which robots would have to perform a variety of tasks. Teams have six weeks* from the game reveal in early January (here is an example of a "kickoff" video) to late February to design, build, and program a robot to the best of their ability. After that, there are six to seven weeks of competition in which teams travel to regional competitions, where there are various ways to qualify to the world championships. (A few states in the US have a district qualifier system: teams accumulate district points from regional performances and then participate in a playoff after the regular season to determine who advances to worlds.)

In the typical competition, each team plays about ten times. Each match is played by two alliances of three teams each. Matches typically begin with an "autonomous" period during which robots operate only with pre-written code, followed by a "teleop" period during which human players control the robots. Each game is different, but at a high level, robots complete as many objectives as possible during the two to two and a half minutes for each match. Here's a video of the final match of one of the two** 2019 worlds championships.

Why it's important/interesting

FRC is essentially a sport. Matches are commentated on by shout casters, the same generic music plays over and over during events, and most importantly, there are a lot of things happening on the field. As such, there are a lot of statistics associated with each match. Here's the detailed score breakdown from The Blue Alliance (a community-made hub for FRC match data) for the video I linked above. At a micro level, there are a lot of specific questions that can be asked, but a lot of these ultimately boil down to **"Which individual teams contribute the most to their alliances' victories?"**.

As with many activities, people constantly discuss who the best teams are. However, in an activity where statistics aren't aggregated per team, it's hard to discern what each team does*** in any given game. Official event rankings (ex. Prelim seeds) aren't useful because they use alliance-wide match results. There have been attempts to answer this question with community-made statistics, but these aren't particularly accurate. The central focus of this project revolves around creating a more accurate rankling list of teams, using data from multiple events and potentially even multiple years. This project could also be used to answer tangentially related questions—for example, **"How much do individual teams improve during the course of a single season?"**

* Starting in 2020, the six-week restriction actually got removed (for equality reasons), but there actually hasn't been a complete season without "bag-and-tag" day due to a certain pandemic.
** This started in 2017. It does seem strange at first, but similarly, this is for inclusivity/access reasons. As a side note, this actually allows us to see differences between the best teams from different regions—each team (with a handful of exceptions) is locked to attending either Houston or Detroit based on their location.
*** This is possible if people watch rounds as they happen, but not every team has someone watching every competitor in every round. There is also a lot of activity on the field at any given time, complicating the process of manually tracking a team.

Data

The data source used will be thebluealliance.com. I will be scraping this website as part of my data collection process.