Tableau for Robotics: Collecting Data

INTRO TITLE

Collecting data is core to any scouting strategy - without properly collecting data, there is no way to get insights from your scouting work. This whitepaper focuses on collecting information efficiently by designing a form for scouts for eventual easy data input.

When people think of scouting they picture a row of students taking notes on paper and then entering it into a database. This phase, called “data collection” is just a small part of scouting, it is important to design it in a way that is fast, accurate and intuitive.

The quality of the data you collect is one of the biggest limiting factors in analysis. If your data is full of holes or inaccuracies, it will be hard to get enough credibility to convince people with your data. However, if you gain a reputation for accurate and comprehensive data, you will be valued as an alliance partner for it.

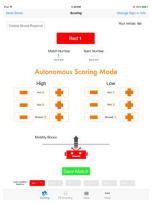
DATA COLLECTION TECHNIQUES

Just as there are a lot of different ways to take notes, there are many different ways to scout a match. Each method has pros and cons. Almost all scouting methods fall into one of these five categories.

Method	Cheap	Dependable	Easy to set up	Flexibility	Intuitivity	Power
Laptop Scouting	OK	OK	OK	OK	OK	Good
Scouting Apps (Tablet/Phone)	Bad	OK	OK	OK	OK	Good
Paper Scouting	Good	Good	OK	Good	Good	OK
Scouting App (Website)	OK	OK	Good	Bad	Bad	OK
Pure Notes	Good	Good	Good	Good	Good	Bad

Note: This table assumes you are using an app or website made by someone else. For these methods, dependability is lower because of batteries and internet connectivity issues. This graph should also change with time, as scouting is a relatively young part of FRC and is likely to become more intuitive and powerful later as teams develop better tools.

The table below goes into a little more detail about each one of the scouting methods previously mentioned.

Scouting Method	Example	**Description**	**Advantages**	**Disadvantages**
Scouting Apps		Using a mobile device (tablet or phone) with an app another team has developed to scout.	• Powerful • Already created • Can share data • Lots of choices • Saves you effort	• Not flexible • Hardware costs • Technical challenges (bluetooth, internet)
Paper Scouting		Using paper scouting sheets, scouts fill them out by hand each match. Data can be stored digitally or in a folder.	• Inexpensive • Easy to make • Customizable • Quick to make • Quick to train	• Lacks power of digital scouting • Hard to organize • Requires transcription
Scouting Website		Using any internet-connected device to access a website that acts like a scouting app	• Any device • Inexpensive • Secure data	• Requires internet • Not flexible
Laptop Scouting		Digital scouting, but on a laptop instead of a tablet.	• Easy device • Flexible • Customizable • Provides raw data	• Battery power • Cons of scouting apps
Pure Notes		Taking qualitative notes on other teams' robots.	• Very easy • Can be digitalized • Does not take many people	• Not powerful • Cannot run analytics on the information

Some of the less overt characteristics to look for may not directly impact score, but are important to strategy. For example, if the rules and space favor a blocking mechanism, this will not impact your roleplaying. When it comes time to face real teams, though, you may have to change your match strategy accordingly. Another example would be knowing that robot has trouble collecting a game piece or interacting with a human player makes them a prime target for defense.

DATA TYPES

Once you've worked through some preliminary analysis of how the game will be played, you can use a table to organize the important features you find.

Action	What to look for (data field)
Shooting a goal	• Did they hit? • Did they hit secondary goal? • Did they miss? • From where did they shoot?
Special feature	• Did they use it? • How many times did they use it?
Endgame	• Did they attempt endgame actions? • What did they attempt? • How successful was it?

To make it as simple as possible to enter this data once gathering it, it is also worth considering how to collect the data. All of the data in a game can be summarized in one of six ways:

- **Count**
  - How many times an event happened
- **Short list**
  - There are a limited number of outcomes, and one outcome must always happen
- **Boolean** (yes/no)
  - Whether something happened or not
- **Location** (coordinates)
  - Where something happened
  - Note: use only one location field per period to avoid problems
- **Open string**
  - Qualitative notes (use sparingly)
- **Ranking**
  - Ordinal notes (e.g., 1-3, good/average/subpar)

To get you started, here is a list of common features in most games and suggestions for data field types:

- **Autonomous**
  - Has autonomous phase (boolean)
  - Performed autonomous special action [blocking, moving, or collecting in auto] (boolean)
  - Autonomous strategy (short list)
  - Autonomous shots (count)
  - Autonomous starting location (location)
- **Tele-Op**
  - Tele scoring and accuracy (count)
  - Position (location)
  - Tactics/role played (short list)
  - Special actions (count)
- **End**
  - Endgame attempt (boolean)
  - End game scoring (count or short list)
  - Fouls (boolean, open string)
  - Qualitative
  - Robot design features(short list, boolean)
  - Comments

Your finished table should look somewhat like the example table below.

Action	What to look for (data field)	Data field type
Shooting a goal	Hit	Count
	Hit secondary goal	Count
	Miss	Count
	From where	Location
Special feature	Did they use it?	Boolean
	How many times did they use it?	Count
Endgame	Did they attempt endgame actions?	Boolean
	What did they attempt?	List
	How successful was it?	Boolean or count

WARNING SIGNS

It is easy to get carried away when identifying important game features, but there is only so much a scout can collect in a match. If you find a data field that fits into one of the categories below, consider changing how you are collecting it, or don't collect data on it.

- **Hard or distracting to collect**
  - Robot speed, for example. While having a fast robot is useful in a match, it is also virtually impossible for scouts to collect accurately and objectively. Having a scout focus on robot speed distracts them from the other more important data points (like scoring points).
- **Can be calculated from other data**
  - Cycle time, for example. Cycle time can be calculated by dividing the shots scored by the time of tele-op period, which is easier data to collect.
- **Subjective variables** (opinions, ratings)
  - Driver skill. While knowing that a bot has a bad driver is an important piece of information, the variance on skill evaluation from your scouting team will be significant.
  - If the game requires a rating or an opinion, you will get more accurate data by using a smaller scale. Bad, neutral, good is a much better system because there is less room for opinion.

This completed data collection table will be the basis of all future scouting. It serves as an outline and thesis for the development of the remainder of the scouting program for the competition. The data collection table is also a rationale for what change scouting will be able to affect. Completing these early steps outlined in this white paper should help you convince others why your scouting method is well thought out.

Once you have thought about how to break down the game, the next steps are to develop the scouting sheets you will use to collect data and to build the database you will use to store and organize the data.