# Cosine Data Engineering AI Bootcamp Midpoint Assessment

By now you have learned about statistics and hypothesis testing, and exploratory data analysis, and regression. We will be putting all three together into an assessment project.

## Overview

This assignment asks you to ask and answer a question about the 2019 FRC Robotics Competition. Data will be provided to you for results from two different districts as well as information about the teams.

## The Assignment

Use R and knitr to create a pdf asking a question, exploring the dataset, phrasing it as a hypothesis test or regression question, and then analyzing that hypothesis.

Instructions:

1. Download and unzip the project file.
2. Rename ProjectTemplate.Rmd to your project’s name
3. Create a new project in this directory.
4. Edit your .Rmd file.
5. When complete, render a pdf using the knitr button. To do this, you will need tinytex installed (below).
   1. You can render the output as HTML in the initial stages if you want to make it faster.
6. Upload your assignment here: <http://aka.ms/aib-assessment> .
7. Review the submissions of several peers (instructions coming later).

## FRC Overview

The data you are analyzing comes from the First Robotics Competition’s 2019 season, Destination: Space. In FRC, high school students are challenged to build robots to complete a series of tasks. They have wide latitude to build as they see fit and you may observe many different designs and materials used.

Matches are always played with 3 robots vs 3 robots. One team is designated red, the other blue. In the early games, matches are assigned randomly. Afterwards, there is a playoff where teams are chosen and then stay together for the rest of the competition.

The basic rules and scoring for Destination: Space are described in this [video](https://www.youtube.com/watch?v=Mew6G_og-PI). At the end of the match, the highest scoring team wins. Teams are then ranked in the competition based on Rank Points. 2 for a win, 1 for filling at least one rocket ship, and 1 for having enough hab climb points. After each robot gets 12 rounds of competition with random partners, there is a playoff system where the top 8 ranked robots pick teams and compete for overall victory.

It might be a good idea to view some matches to get a feel for the game. You can see some from the PNW Glacier Peak event here: <https://www.thebluealliance.com/event/2019wasno>. Here is a good example match: <https://youtu.be/6BTP7oGam0E> .

If you have any other questions about the competition, feel free to ask on the channel. My daughter competed for team 492 Titans out of International School.

## Ideas

You can analyze anything you want, but here are some ideas to get you started.

* Are teams from one region better than the other? From one state in the PNW vs another?
* Does the time of day a match is held affect the outcomes?
* What elements of the event are highly correlated with victory?
* Does the timing of the event (early vs late season) affect results?
* Does the age of the team (rookie year) relate to success? If so, is there a cutoff?
* Is it a good idea to risk fouls?

Additionally, here is some analysis done on last year’s competition that may inspire you: <https://blog.thebluealliance.com/2018/03/10/2018w2-match-analysis-using-the-tba-api-and-r/>

## Data

The data you are working with result data from the FRC competition season in the Pacific Northwest and Texas.

There are three files:

Events.csv: Data on each district event

Teams.csv: data on all teams in the two districts

Matchdetails.csv: data on each of the nearly 2,000 matches

### Field Definitions (selected fields only):

|  |  |
| --- | --- |
| Field | Meaning |
| Comp\_level | Qm = Qualifying match qf = quarterfinals  Sf = semifinals  F = finals |
| Event\_key | which event this match took place at |
| Match\_number | Matches are numbered 1..n in the order they took place |
| *Score\_breakdown.<alliance>* |  |
| autoPoints | Points scored during the 15 second “standstorm” period |
| BayX | At the end of the match, was there a hatch cover (panel) and ball (cargo) on that location? |
| completeRocketRankingPoint | Did the team get a ranking point for having fully filled a rocket? |
| endgameRobot# | What level of climb did each robot achieve? |
| habDockingRankingPoint | Did the team get >= 15 hab points? If so, they got a ranking point |
| preMatchBay# | How was this location preconfigured before the match? |
| preMatchLevelRobot# | What level did the robot start on? Higher == more points. |
| teleopPoints | Total points for the team during the human-controlled period |
| foulPoints | How many points worth of fouls did the team gain? |
| foulCount | How many times did this team foul the other? |
| District | Name of the district robot competed in |
| Blue.robot# | Team code for that robot |
|  |  |

This should be all the data you need, but if you want, more data is available here: <https://www.thebluealliance.com/apidocs>

I have put some code at access the results up on my github account: <https://github.com/steverowe/FRC-R>

## RMarkdown

We will be using RMarkdown to produce a PDF document. RMarkdown is a powerful way to mix text, code, and output from code into one document.

You will need *RStudio*. You should have already been using this, but if you have not been, you should install it now from here: <https://www.rstudio.com/products/rstudio/download/#download>. There is a new version (1.2) which has improved rmarkdown support which you may want to pick up.

You will also need a LaTeX installation. The current recommendation is *TinyTex*:

<https://yihui.name/tinytex/> (detailed instructions [here](https://bookdown.org/yihui/rmarkdown/installation.html#installation))

If you need a refresher on RMarkdown, go here: <https://rmarkdown.rstudio.com/lesson-1.html>