**War On Ice**

**Project Problem and Hypothesis**

* What's the project about? What problem are you solving?
* Where does this seem to reside as a machine-learning problem? Are you predicting some continuous number, or predicting a binary value?
* What kind of impact do you think it could have?
* What do you think will have the most impact in predicting the value you are interested in solving for?

**Datasets**

* Description of data set available, at the field level (see table)
* If from an API, include a sample return (this is usually included in API documentation!) (if doing this in markdown, use the javacription code tag)

**Domain knowledge**

* What experience do you already have around this area?
* Does it relate or help inform the project in any way?
* What other research efforts exist?
  + Use a quick Google search to see what approaches others have made, or talk with your colleagues if it is work related about previous attempts at similar problems.
  + This could even just be something like "the marketing team put together a forecast in excel that doesn't do well."
  + Include a benchmark, how other models have performed, even if you are unsure what the metric means.

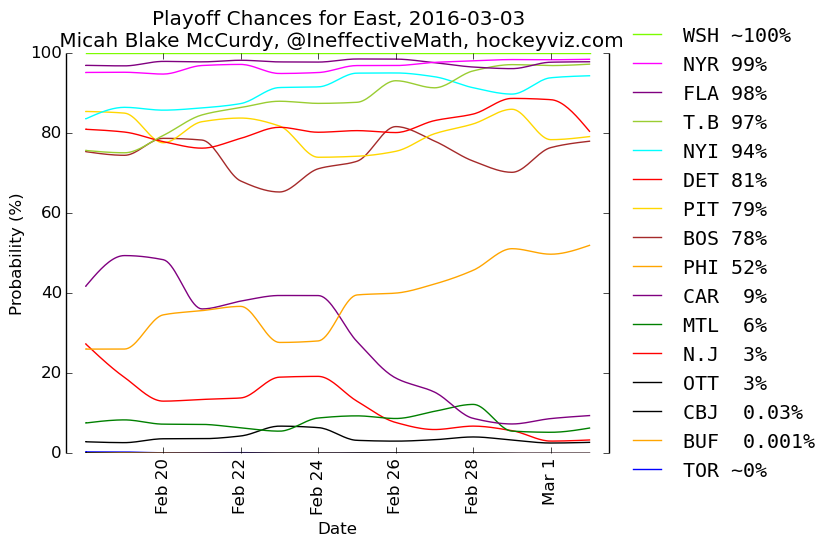
**Project Concerns**

* What questions do you have about your project? What are you not sure you quite yet understand? (The more honest you are about this, the easier your instructors can help).
* What are the assumptions and caveats to the problem?
  + What data do you not have access to but wish you had?
  + What is already implied about the observations in your data set? For example, if your primary data set is twitter data, it may not be representative of the whole sample (say, predicting who would win an election)
* What are the risks to the project?
  + What's the cost of your model being wrong? (What's the benefit of your model being right?)
  + Is any of the data incorrect? Could it be incorrect?

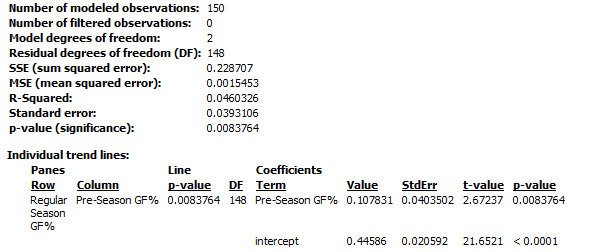
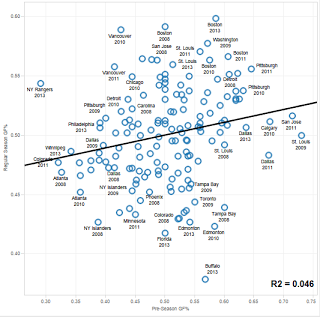
**Outcomes**

* What do you expect the output to look like?
* What does your target audience expect the output to look like?
* What gain do you expect from your most important feature on its own?
* How complicated does your model have to be?
* How successful does your project have to be in order to be considered a "success"?
* What will you do if the project is a bust (this happens! but it shouldn't here)?

**Similar Projects**

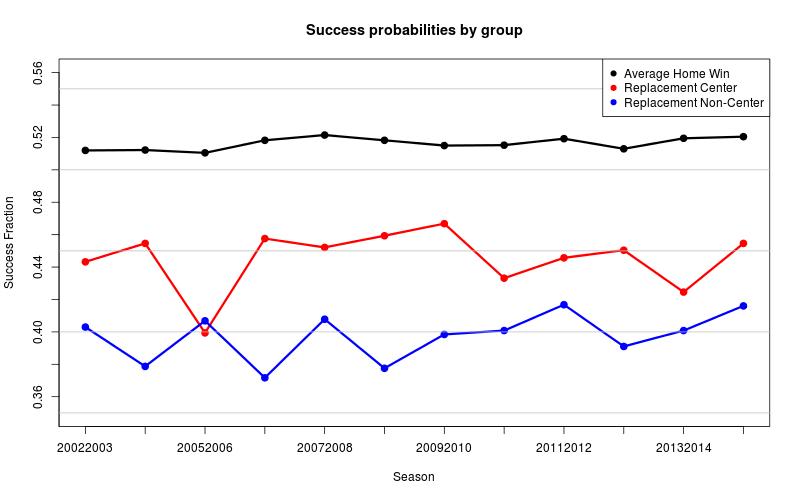
Logit Model for Probability of Making Playoffs

Pre-season impact on final standings:

Calculating “Wins Above Replacement” and anticipated “Success”





**SOURCES:**

[**http://blog.war-on-ice.com/sharing-is-caring/**](http://blog.war-on-ice.com/sharing-is-caring/)

[**http://war-on-ice.com/playertable.html**](http://war-on-ice.com/playertable.html)

[**http://www.hockey-reference.com/teams/NYR/1999.html#skaters::none**](http://www.hockey-reference.com/teams/NYR/1999.html#skaters::none)