Capstone Project Ideas

1. How people learn in informal environment?

I’m working on a project with my professor. The project is based on data from Scratch Online Community at MIT. There are almost 2 million projects’ data from 2007 to 2012. The main users are 8-16-year-olds who don’t have programming background but learn coding and programming basics through Scratch to create games, animations, interactive stories and etc. I’ve already spent a week or two to do data wrangling, recoding and transforming data. I still need time to fully understand the data, so I can figure out what to do next. The main problem to work on this project as my Capstone project is I need to get IRB approval so that I can share with you the results.

1. How to detect fake reviews from Amazon reviews?

If I cannot get an approval on time from above project, I’d like to work on amazon reviews data I find from Data is Plural. I haven’t looked at the data yet, but I like the idea of determining fake reviews from all the reviews.

1. Kaggle Competition: Sold! How do home features add up to its price tag?

I like the idea of predicting a home price using different variables. As the competition describes, there are 79 explanatory variables describing (almost) every aspect of residential homes in Ames, Iowa, to challenge to predict the final price of each home.