CSC-450, Senior Research Project Ideas

Your *research* project must be centered around a specific hypothesis that you will test or an objective you will meet by designing a method or model, collecting and analyzing data, and drawing conclusions using an appropriate and reproducible methodology.

The following is a list of possible research questions / projects, and is designed to help generate ideas. There are many, many more ideas that would be acceptable. Don't hesitate to contact me if you want to run any ideas by me.

Area of Interest	Question / Hypothesis / Need	Method
Data Mining*	What region of the U.S. is most interested in *cats*?	Based on Twitter data, compare the proportion of tweets that mention *cats* in the west, southwest, midwest, southeast, and northeast.
	Are Donald Trump's tweets "angrier" than those of his staffers (see: http://varianceexplained.org/r/trump-tweets/) What can we learn from Hillary Clinton's tweets?	Compare the times, word choices, and perform sentiment analysis on @realDonaldTrump tweets originating from an Android vs. iPhone
	Any individual can be connected to any other individual in at most 6 steps (see six degrees of separation; six degrees of Kevin Bacon).	Using (a subset of)Twitter data (or Linked In data), determine the average number of connections separating any two individuals.
	What computer programming languages are employers looking for?	Write a script (better yet, a web page) that downloads current programming job listings from a website such as Monster.com, and tallies results for a set of languages
Computer Security	A password stored using *md5* encryption is easier to crack than a password stored using *bcrypt* encryption. (inspired by: https://blog.avast.com/2015/09/07/taking-a-closer-look-at-cracked-ashley-madison-passwords/)	Choose a set of passwords to analyze. Use a password cracking tool and evaluate how long it takes to crack each password. How does the time change with the complexity of the password?

	Method A can be used to detect spam e-mails (or spam tweets) (and is better than Method B)	Develop a method (e.g., a machine learning algorithm) to detect whether an e-mail is spam or not. Calculate the accuracy of the method across 500 e-mail messages. Compare the accuracy of this method to the accuracy of another method.
Modeling / Simulation	Placement of a pillar that partially obstructs a door results in more people escaping safely in the event of a fire.	Use an agent-based model to simulate individuals inside a room, that either has or does not have a pillar. At a fixed time, a fire is initiated, and the individuals run out of the room. Compare the speed and number of people safely exiting the room across the two conditions.
Game Development	There are currently no (good) games for students to learn (or reinforce) C++ concepts	Create a fun, engaging game where students get points for demonstrating knowledge of C++ concepts. (Ideally, you would evaluate whether students learned more (or had greater enjoyment) from this game than others).
	There is no SDL library for drawing humans	Create an SDL library for drawing humans. The library should provide basic functions, such as zooming in and out, and moving left and right.
Any field	You may answer a question by conducting a survey, for example to evaluate how knowledgeable individuals are about security issues.	This research falls under Human Subjects Research (see the *HumanSubjects.pdf* file for more information)

*Many datasets can be found here: http://rs.io/100-interesting-data-sets-for-statistics/