

Rob Acheson

racheson@fas.harvard.edu

CS-171 - Homework 3

2/21/2013

Objectives

1. Which (primary) programming language is preferred by male and female students, respectively?
2. Which (primary) programming languages do students feel more comfortable with?
3. What relationships are discernible between programming experience (in years) and (primary) programming languages of choice?
4. How do OS choices relate to students' reported comfort as programmers?
5. What is the relationship between age, programming language, and programming experience, and how does it change over time?
6. How does the 2013 data compare to the data from years prior? Is there anything notably different about the data from 2013 relative to the past, given enrollment in the class has increased quite a bit? What trends can you pick out from the data, and what can you predict (if anything) about the future of CS171?

Required Data for Objectives

- Primary Programming Language - string
- Gender - male or female
- Programming Comfort - 1 to 5
- Experience Years– number
- OS – list of strings
- Age – number
- Enrollment Year – number

Merging and Cleanup

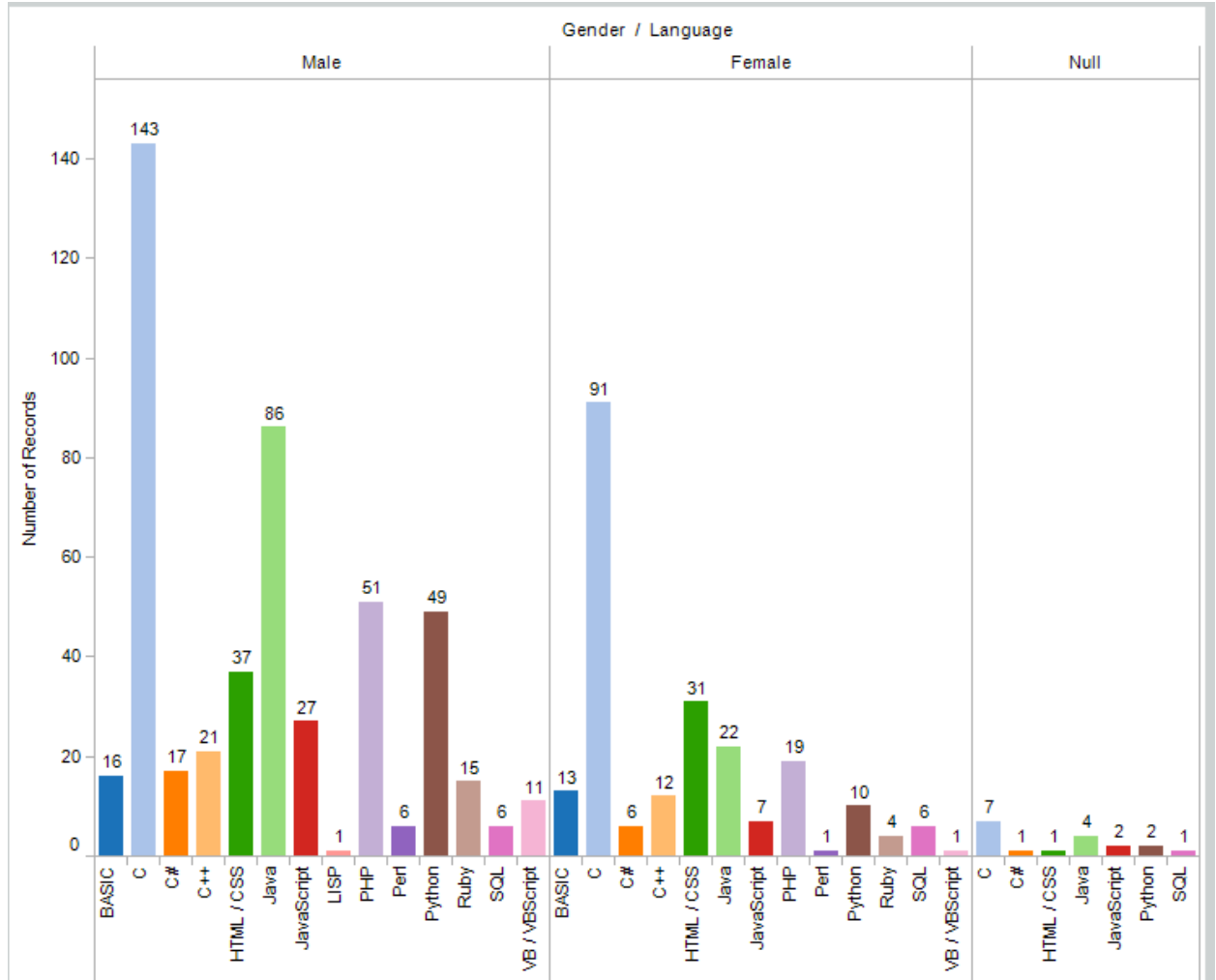
Through a series of edits with Google Refine, many columns were renamed for proper merging within Fusion Tables. For example Degree became Primary Concentration, Coding Skill became Coding Comfort. After gauging the required data for the current tasks and in essence of time, not every column was sanitized. Language and Programming Comfort were fine as-is. Within Gender and Age, N/A and I prefer not to disclose were combined and changed to blank. Enrollment year was added to the 2013 data before combining.

Answers

Which (primary) programming language is preferred by male and female students, respectively?

C was the primary programming language of choice across both genders. This was revealed using a bar graph. Languages, separated by gender, were plotted across the x-axis. The number of records found for each particular language was shown by the height of the bars. I added color coding to the languages to allow users to quickly draw

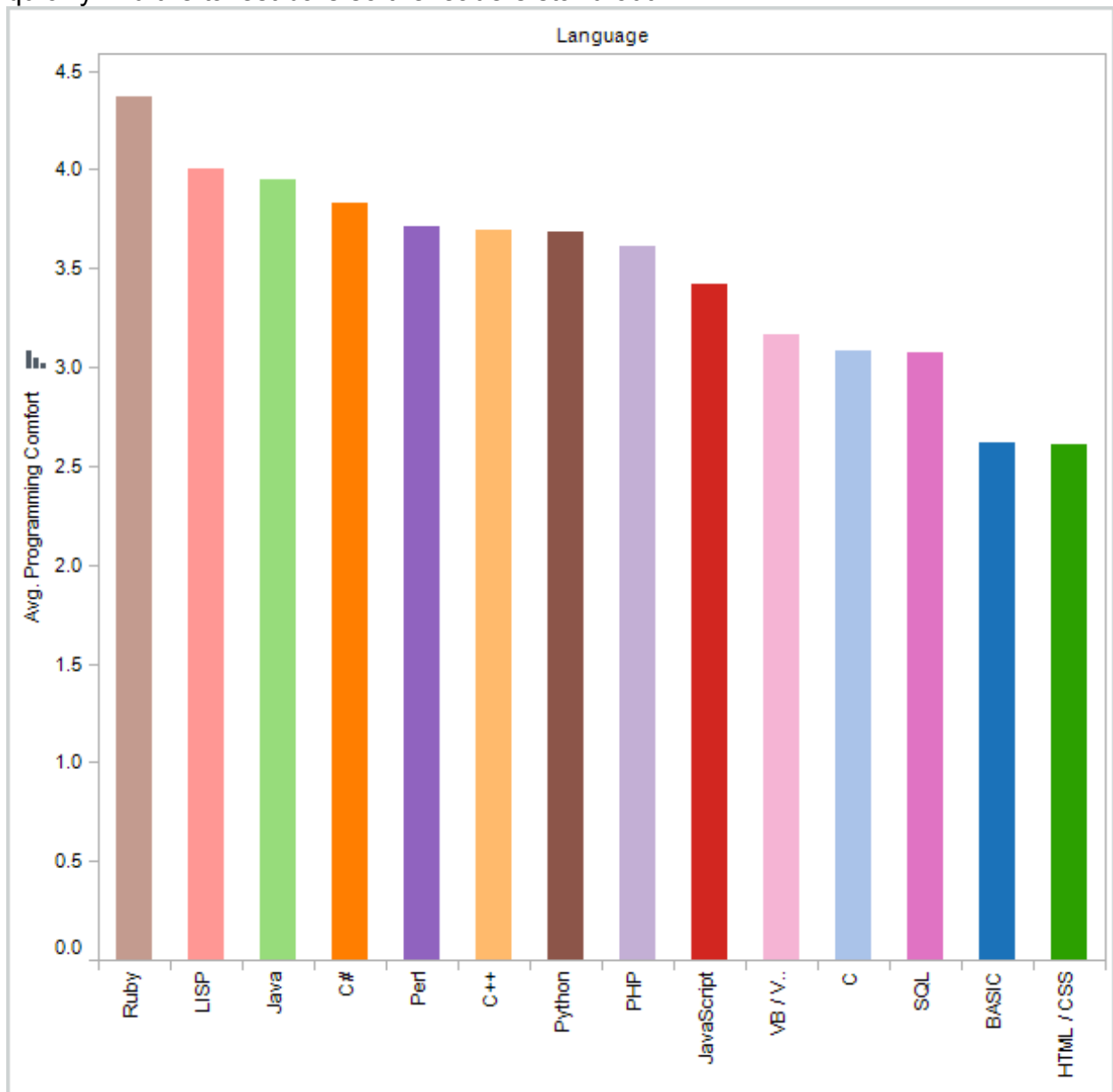
parallels between the genders. We can quickly find the tallest bars so the leaders stand out.



Which (primary) programming languages do students feel more comfortable with?

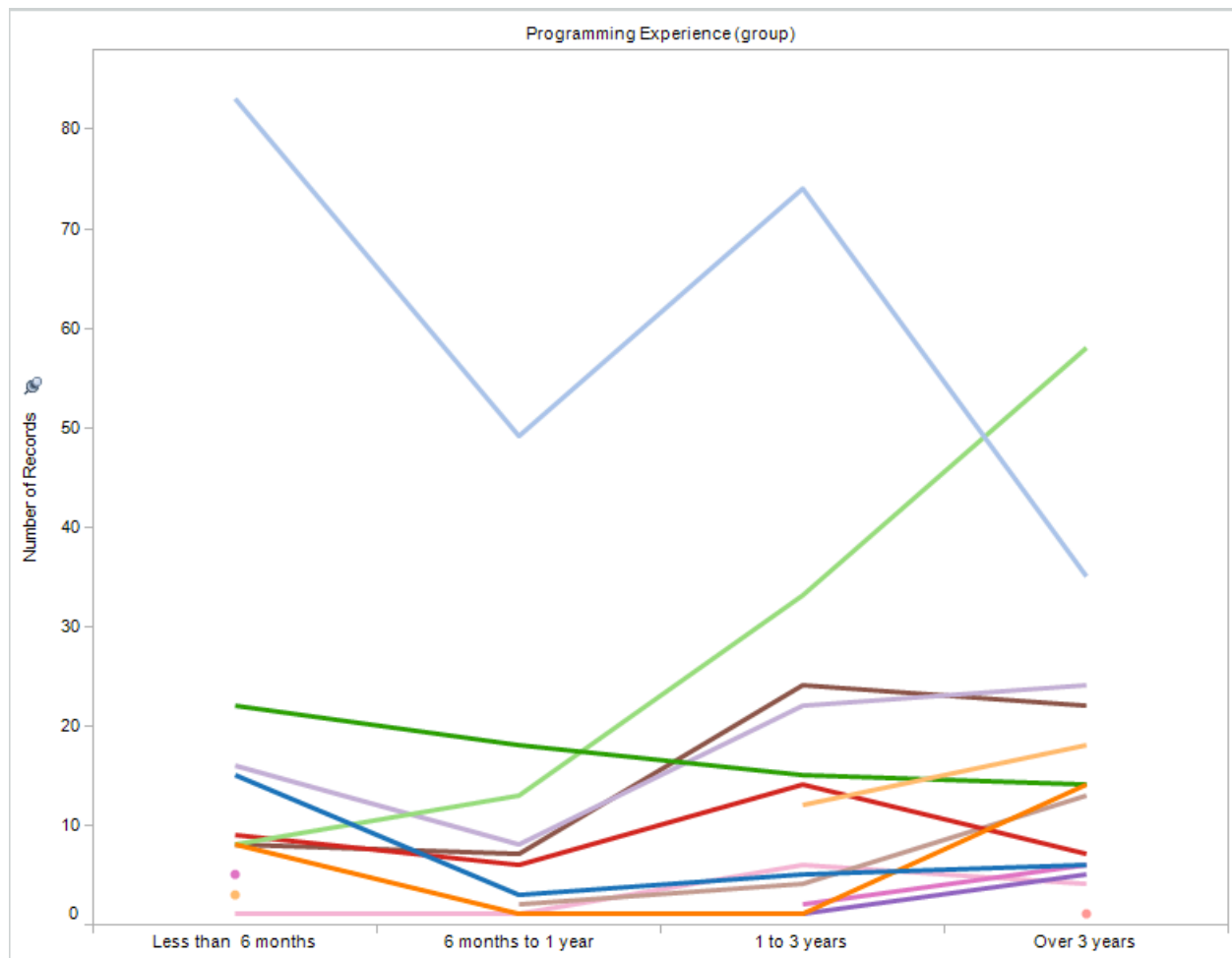
Users of Ruby, LISP, Java, and C# were the most comfortable in that order. To determine this all languages were plotted out on the x-axis in a bar graph. The height of the bars shows the average comfort level of each language's users. The graph was then organized in descending order of comfort to reveal the top performers. We can

quickly find the tallest bars so the leaders stand out.



What relationships are discernible between programming experience (in years) and (primary) programming languages of choice?

At less than 6 months, C is most popular programming language, with HTML, PHP and others distant alternatives. Over time however, it is evident that C loses its hold of the number one spot. Java, which starts out in the middle of the pack, shows the most consistent growth year-after-year and by the 3 year point has surpassed C as the number one language, although it shares the gains with Python and PHP. HTML declines over time and most of the others remain flat. As a bar graph, this visualization was useful, but it had room for improvement. Considering that experience was a measure of time, as a line graph one can quickly see trends for each particular language.

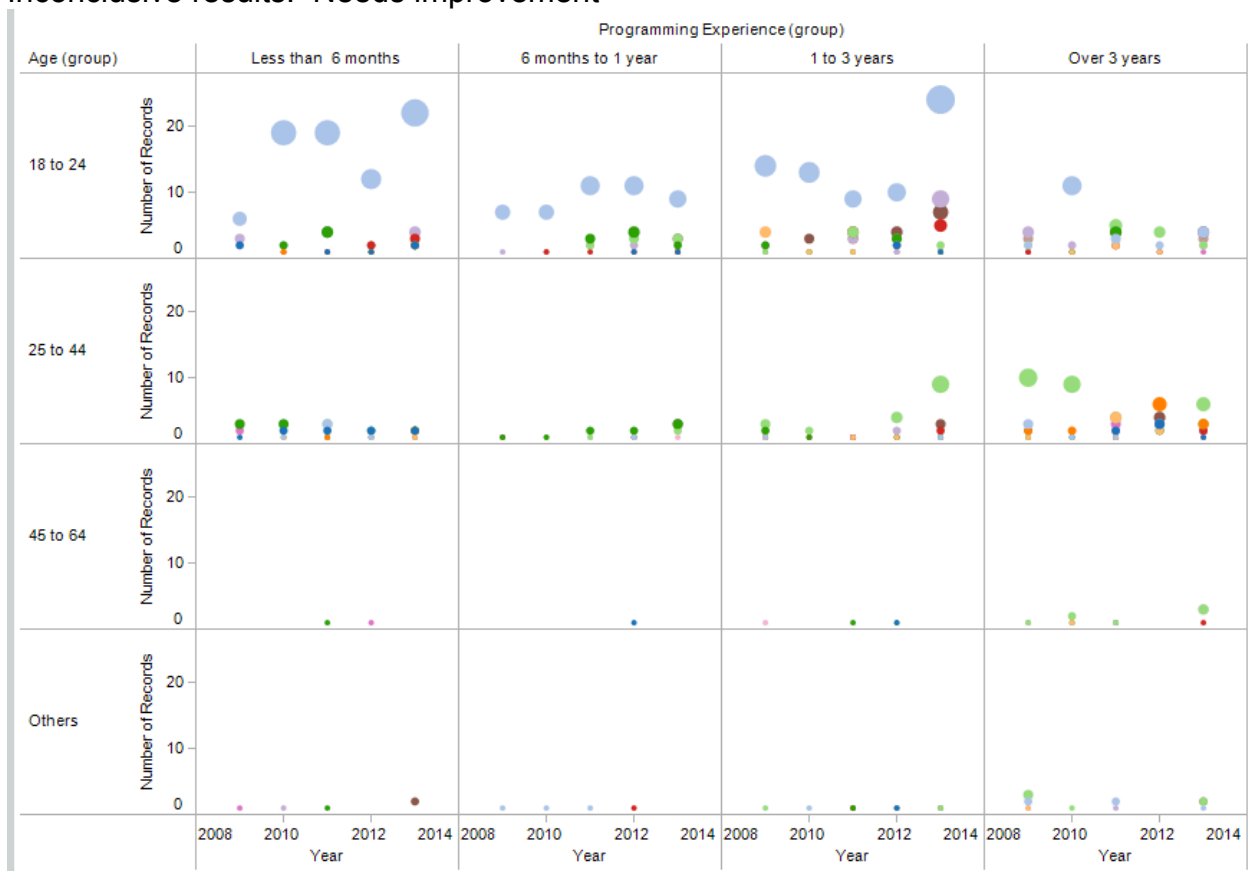


How do OS choices relate to students' reported comfort as programmers?

I was not able to properly display this data

What is the relationship between age, programming language, and programming experience, and how does it change over time?

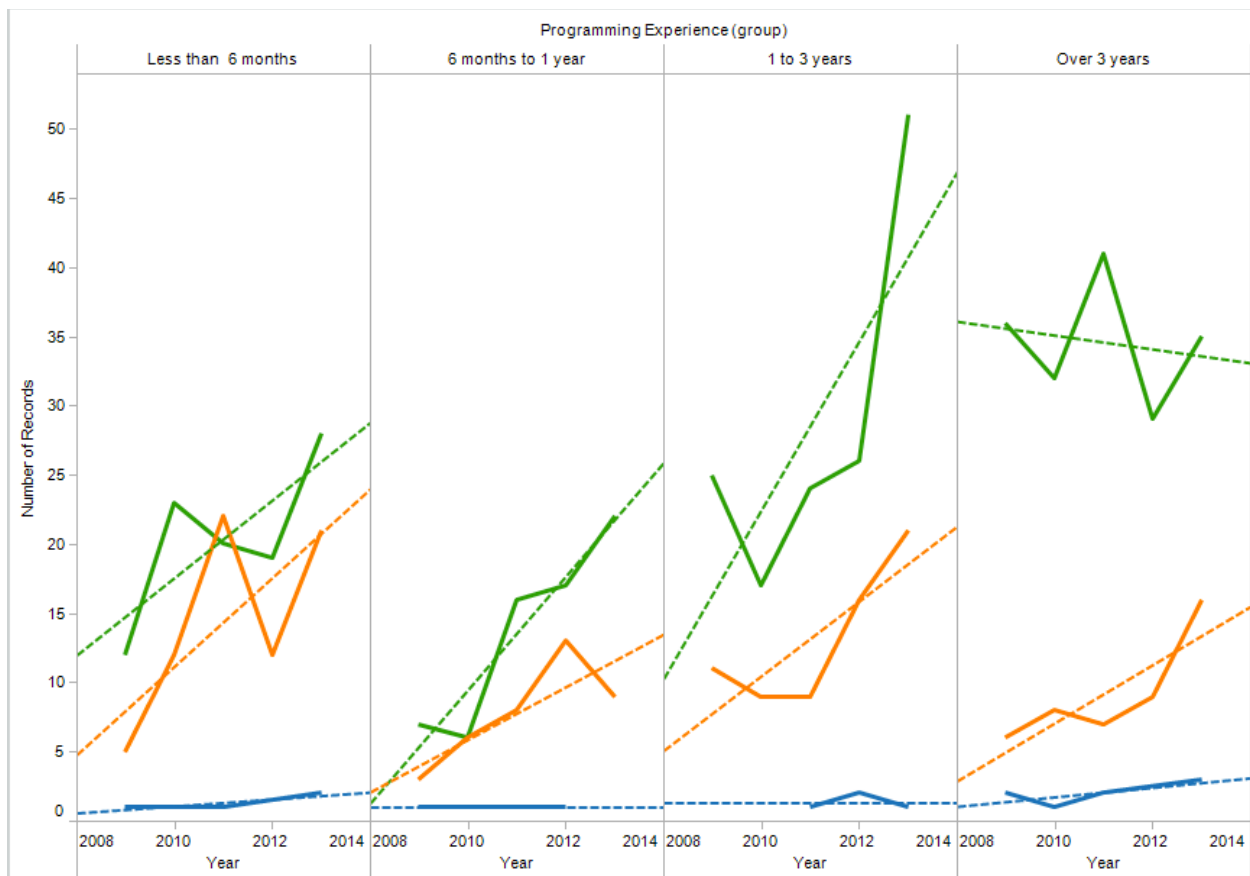
Inconclusive results. Needs improvement



How does the 2013 data compare to the data from years prior? Is there anything notably different about the data from 2013 relative to the past, given enrollment in the class has increased quite a bit? What trends can you pick out from the data, and what can you predict (if anything) about the future of CS171?

Since we know that enrollment is increasing I thought it would be interesting to investigate the groups that were growing the most rapidly. With less than 6 months experience, enrollment is growing evenly between the sexes. However between 6 months and 3 years of experience the growth rate of Male enrollment is greatly increased and Females, though still growing, do so at a slightly lower rate than their less experienced counterparts. This rate slows even more when Females have more than 3 years experience. What is very interesting, however is that Male enrollment with more than 3 years is declining slightly.

In the future, I would predict the majority of students to be male with 1 to 3 years experience. The overall number of females will continue to grow, but the ratio of women to men will decrease over time. Line graphs were chosen because it was easy to compare the heights between genders and to perceive change over time.



Dashboard

The first two sheets at the right are the source sheets that can filter the rest of the data. This allows users to separate results by gender, language, and years of experience. With this one can quickly see how a the trends at the right detailing languages and comfort over time

