

Figure 1: Number of students for each programming language used in the class. The x-axis represents the number of students and the y-axis represents the specific tool obtained by the student. The number of students obtaining the specific programming language is denoted by the green bar and the number of students not obtaining the programming language is denoted by the blue bar. The name of the programming language is ordered according to the number of students who obtain that.

From the Figure 1, we can find that the most popular programming language in the EDAV class is Python and R, of course. While the familiarity of Excel and SQL also beyond 50%. Other programming languages are all below 50%.

From above programming language, the C/C++ language is much more computer science oriented language, and we can infer reasonably that the students who obtain C/C++ would have much more experience in doing programming. Then we can raise an interesting problem, that are the students who have experience in C/C++ more confident in programming than other students?

To solve this problem, we can use the survey results about the confident level in the programming and analytic skills to evaluate the confident level of students obtaining different kind of programming language. We use the students who obtain C/C++ to compare with the students who obtain Python and R, which is the most popular programming language in the class.

**The first problem in survey is “Programming and Analytical Experiences in Github.” The result is shown in the Figure 2.**

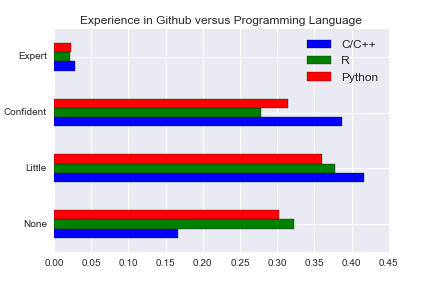


Figure 2: Experience of students in Github versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with the Github. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 2, we can easily find that the students with C/C++ would be more confident in Github than students with R and students with Python, because the proportion of ‘Confident’ and ‘Expert’ in Github of students obtaining C/C++ is explicitly higher than that those of students obtaining Python and R. At the same time, the proportion of students obtaining C/C++ with ‘None’ experience in Github is almost half of the students obtaining Python and R.

**The second problem in survey is “Programming and Analytical Experiences in Matlab.” The result is shown in the Figure 2.**

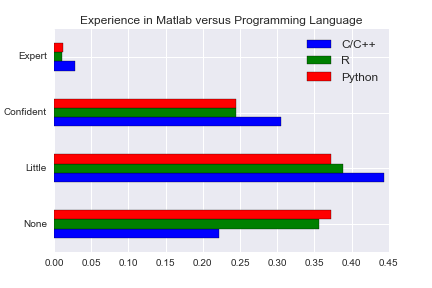


Figure 3: Experience of students in Matlab versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with the Matlab. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 3, we can easily find that the students with C/C++ would be more confident in Matlab than students with R and students with Python, because the proportion of ‘Confident’ and ‘Expert’ of students obtaining C/C++ is explicitly higher than that those of students obtaining Python and R. At the same time, the proportion of students obtaining C/C++ with ‘None’ experience in Matlab is significantly lower than the students obtaining Python and R.

**The third problem in survey is “Programming and Multivariate Analytical Experiences in R.” The result is shown in the Figure 3.**

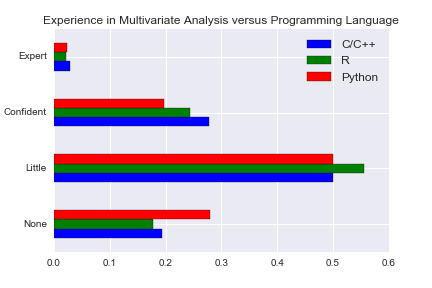
****

Figure 4: Experience of students in multivariate analysis in R versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with the multivariate analysis. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 4, we can easily find that the students with C/C++ would be more confident in multivariate analysis than students with R and students with Python, because the proportion of ‘Confident’ and ‘Expert’ of students obtaining C/C++ is also a little higher than that those of students obtaining Python and R.

**The forth problem in survey is “Programming and Data Manipulation and Modeling Experiences in R.” The result is shown in the Figure 5.**

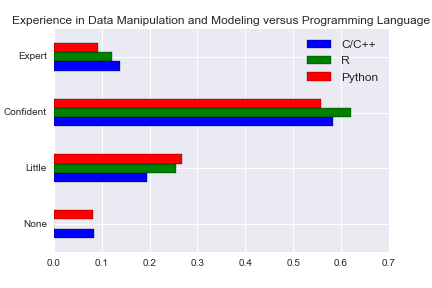
****

Figure 5: Experience of students in data manipulation in R versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with data manipulation. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 5, we can easily find that the students with C/C++ would be comparable confident in data manipulation to the students with R and students with Python, because the proportion of ‘Confident’ and ‘Expert’ of students obtaining C/C++ similar to that those of students obtaining Python and R.

**The fifth problem in survey is “Programming and Data Manipulation and Modeling Experiences in R.” The result is shown in the Figure 6.**

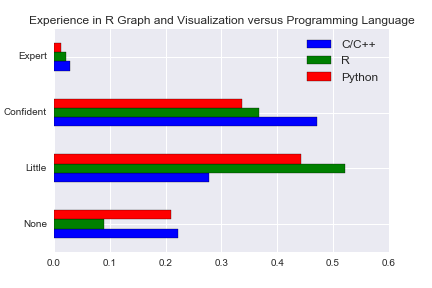
****

Figure 6: Experience of students in Graph and Data Visualization in R versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with the R-Graph. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 6, we can easily find that the students with C/C++ would be more confident in R-Graph than students with R and students with Python, because the proportion of ‘Confident’ and ‘Expert’ of students obtaining C/C++ is explicitly higher than that those of students obtaining Python and R. At the same time, the proportion of students obtaining C/C++ with ‘Little’ experience in R-Graph is almost half of the students obtaining Python and R.

**The fifth problem in survey is “Programming and Data Manipulation and Modeling Experiences in R.” The result is shown in the Figure 6.**

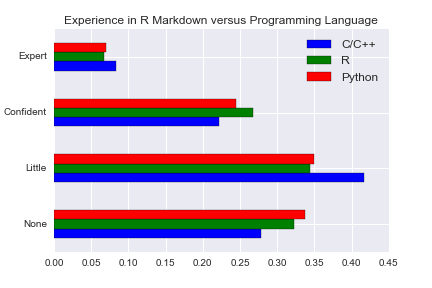
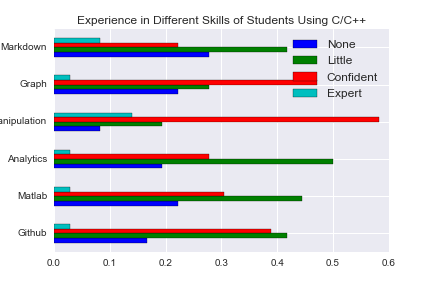
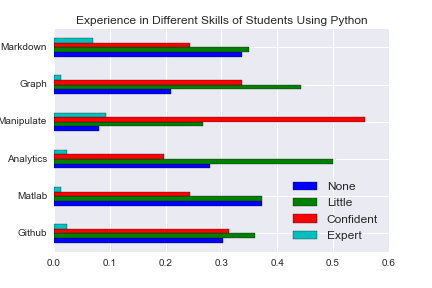


Figure 7: Experience of students in R markdown versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the confident level in familiarity with the R markdown. The blue, green and red histograms represent the students who obtained C/C++, R and Python. The number of students have been normalized to make the comparison easy.

From the Figure 7, we can easily find that the students with C/C++ would be comparable confident in R-markdown comparing to the students with R and students with Python. No explicit pattern to distinguish the student obtaining C/C++ and students obtaining R or Python.

****

****

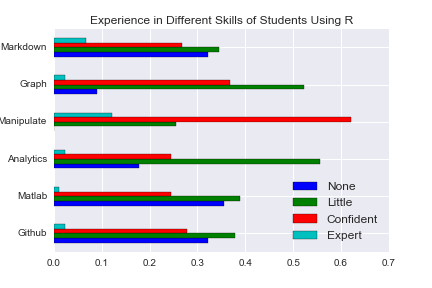
****

Figure 8: Experience of students in different skills versus the programming language obtained by the student. The x-axis represents the proportion of the students, and the y-axis represents the specific skills. The blue, green, red, and cyan histograms represent the students who have ‘None’, ‘Little’, ‘Confident’ and ‘Expert’ experience in the area. The number of students have been normalized to make the comparison easy.

Summary

Generally speaking, the students are more confident in data-manipulation and data-visualization tools with R. Specifically, the students who obtain C/C++ are much more confident than the students with R or Python in the experience of Github, Matlab and even data visualization with R explicitly. In other areas, the students with C/C++ are at least comparable confident comparing to the students with R or Python. It is an interesting pattern that we find the students with C/C++ experience always show more confident in their programming skills, no matter the skill is related to C/C++ language or not in the EDAV class.