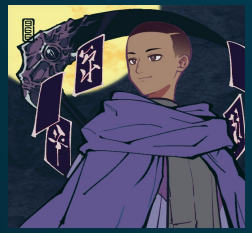
Tech Scholars Project

By Botong, Joji, Kamilah, Norbey & Terrence

Meet The Team

Botong Xu

Joji Kashimura



Terrence Lewis

Norbey Acevedo

Kamilah Guerrero

Project Overview

Utilizing what we learned from our recent lectures within the Tech Scholars Program, we were tasked to derive data from various job descriptions into a program where it would list individual keywords and their correlated frequencies while also sorting them in descending order. Furthermore, we had to prompt the user with a word-search feature, allowing them to find one particular word and its frequency count within the vast index of our already listed terms. Omitting generic words such as “the”, “it” & many more were also implemented as a means to narrow down the list to major keywords. Lastly after the program executes, a csv file would be created containing the program’s results listed accordingly.

In Simpler Terms

Derive data from CSV file

Omit/Find specific keywords

Map/Sort Keywords & Frequencies

Export results into CSV file

DEMO TIME!

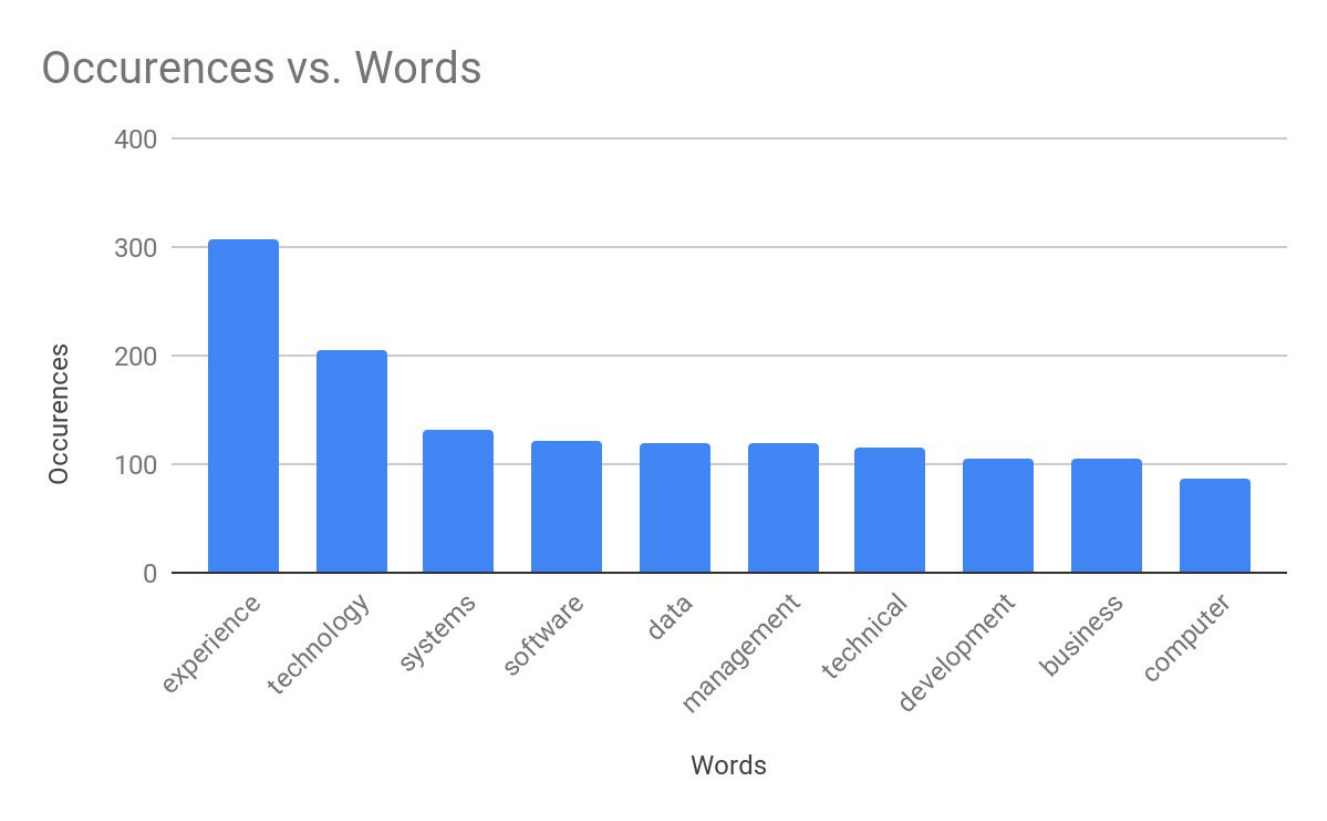
How Did We Do It?

Sorting Algorithm
Using the “algorithm” library

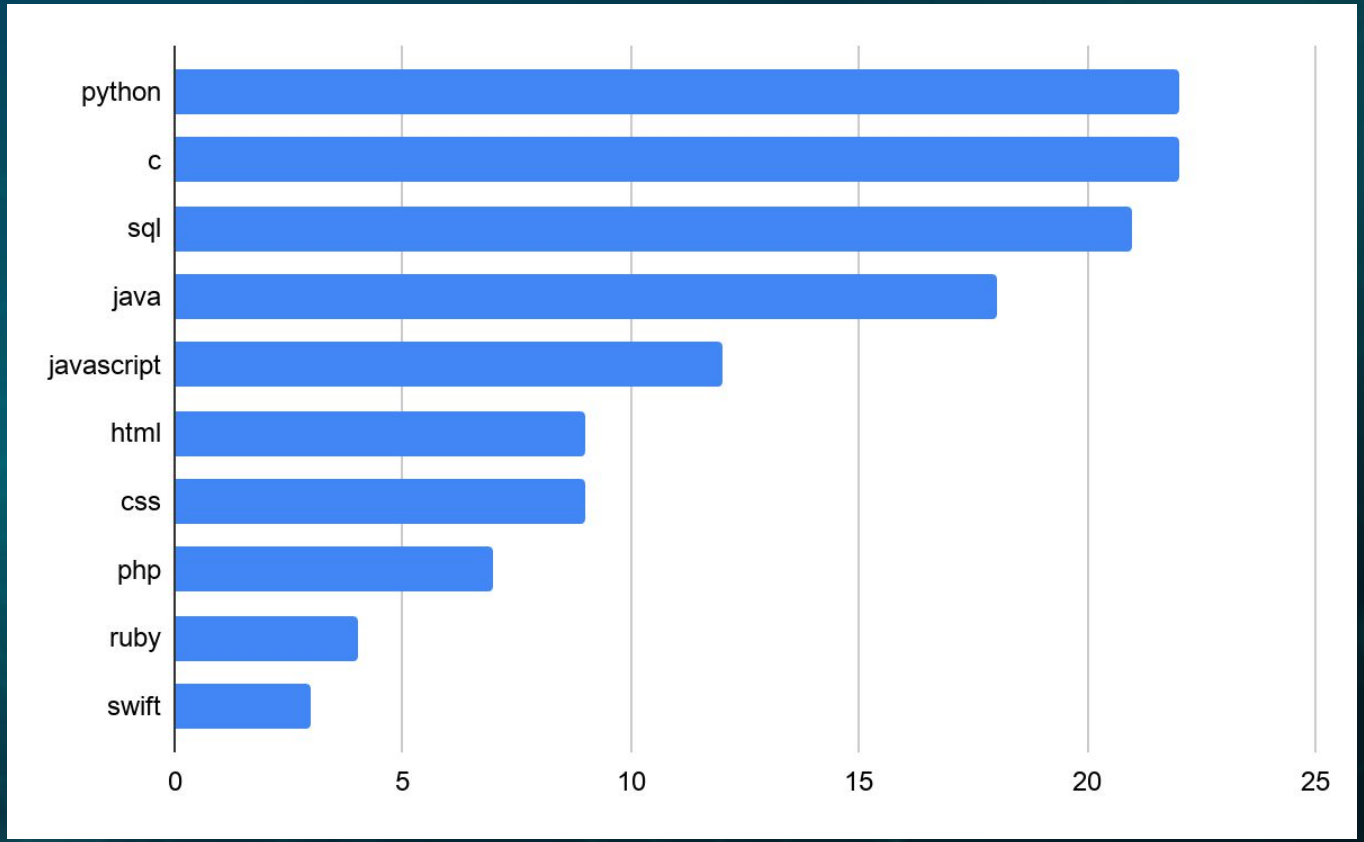
Frequency/Search Algorithm
Using the “unordered_map” library

Menu Algorithm
Using the choice menu

Top 10 Frequent Terms



Top Programming Languages



How Did We Map To Solve Our Problem?

```
std::transform(string.begin(), string.end(), string.begin(),
                    [](unsigned char c) { return std::tolower(c); });
    s.erase(
        std::remove_if(string.begin(), string.end(),
                        [](unsigned char c) { return std::isalpha(c) == 0; }),
        s.end());

      maps.erase("then");
      maps.erase("there");
      maps.erase("to");

            if (!string.empty()) {
                                ++maps[string];
      }
    }
```

How Did We Sort To Solve Our Problem?

```
partial_sort_copy(words.begin(), words.end(), sort.begin(), sort.end(),
        [](auto& a, auto& b) { return a.second > b.second; });

for (auto& pair : sort) {
        cout << ">> '" << pair.first << "' occurred " << pair.second << " time(s)\n";
}
```

SAMPLE OUTPUT:

```
>> 'experience' occurred 307 time(s)
>> 'technology' occurred 205 time(s)
>> 'systems' occurred 131 time(s)
>> 'software' occurred 121 time(s)...
```

How Did We Search To Solve Our Problem?

```
void search(MapType words){
        string s;
        cout << “Enter Keyword: ‘;
        cin >> s;
        transform(string.begin(),string.end(),string.begin(),::tolower);
        int count = maps[s]
        cout << “\n>> ‘“<< s << “‘ occurred  “ << count << “ time(s)\n”;
}
```

SAMPLE OUTPUT:

>> ‘data’ occurred 120 time(s)

How Could We Improve Our Program?

> Removing Plural and Past-Tense Word Extensions

> CSV File Formatting

> Automate a Set Table or Graph

> And Many More..

Our Challenges

As exciting as our project was during development, we did come across a few coding obstacles such as...

> Displaying multiple lines from a file

> Removing punctuations

> Formatting data to CSV file

Conclusion

We are very proud to share with you all our final project! It was honestly a very thrilling and fun experience. As a team, we made sure to support each other when any one of us had any questions or concerns. We established key roles and deadlines to certain phases within our project while also having everyone pitch out great ideas. Not only did we touch on Data Analysis, learned new libraries, mapping and sorting algorithms but we also learned how to work together as a team!

Thank you Tech Scholars.

Our Contact Info

Botong Xu
botong.xu@live.lagcc.cuny.edu

Kamilah Guerrero
kamilah.guerrero@live.lagcc.cuny.edu

Joji Kashimura
joji.kashimura@live.lagcc.cuny.edu
linkedin.com/in/joji-kashimura-765525196

Norbey Acevedo
norbeyace99@gmail.com
linkedin.com/in/Norbeyace

Terrence Lewis
terrence.lewis@live.lagcc.cuny.edu