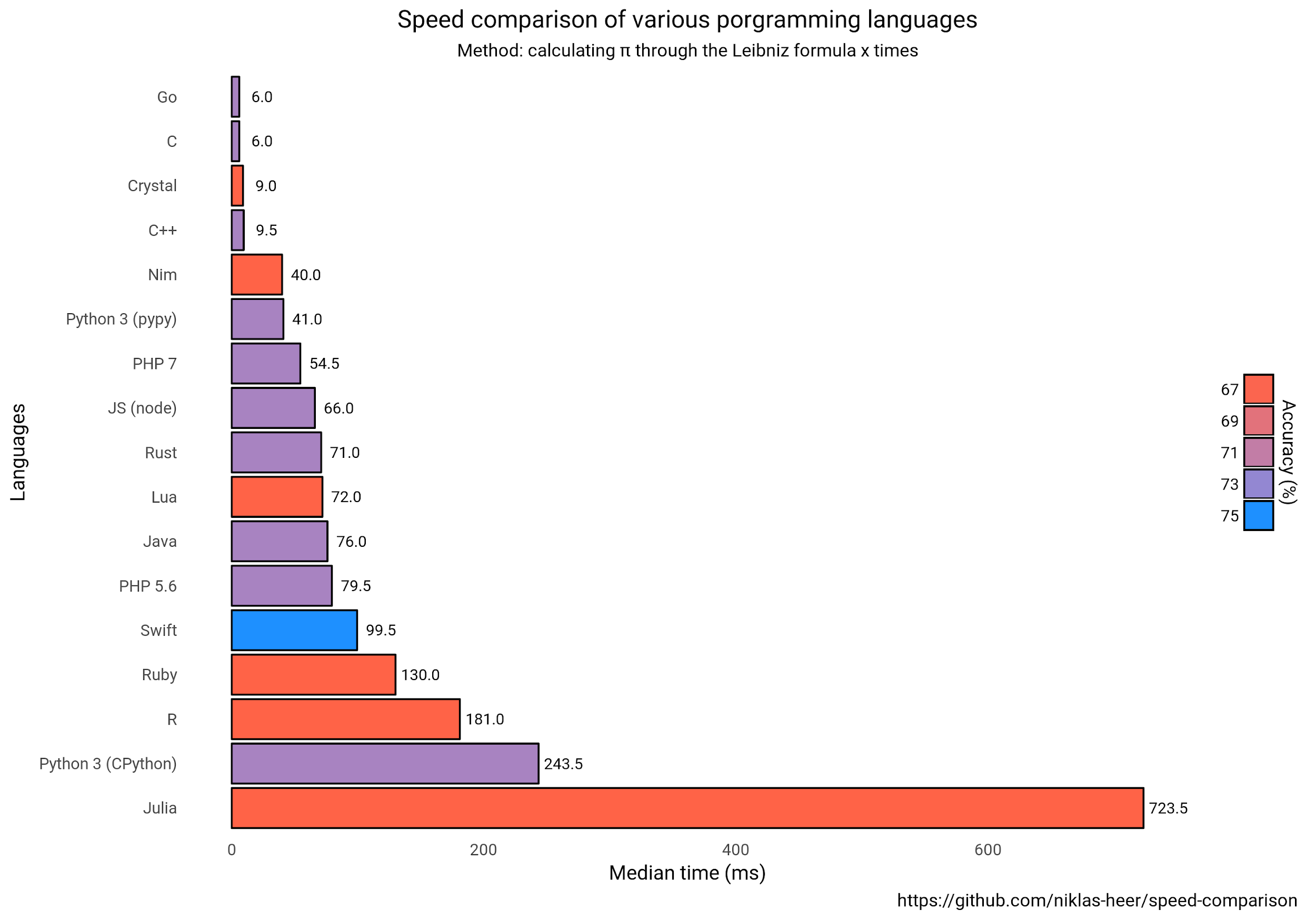
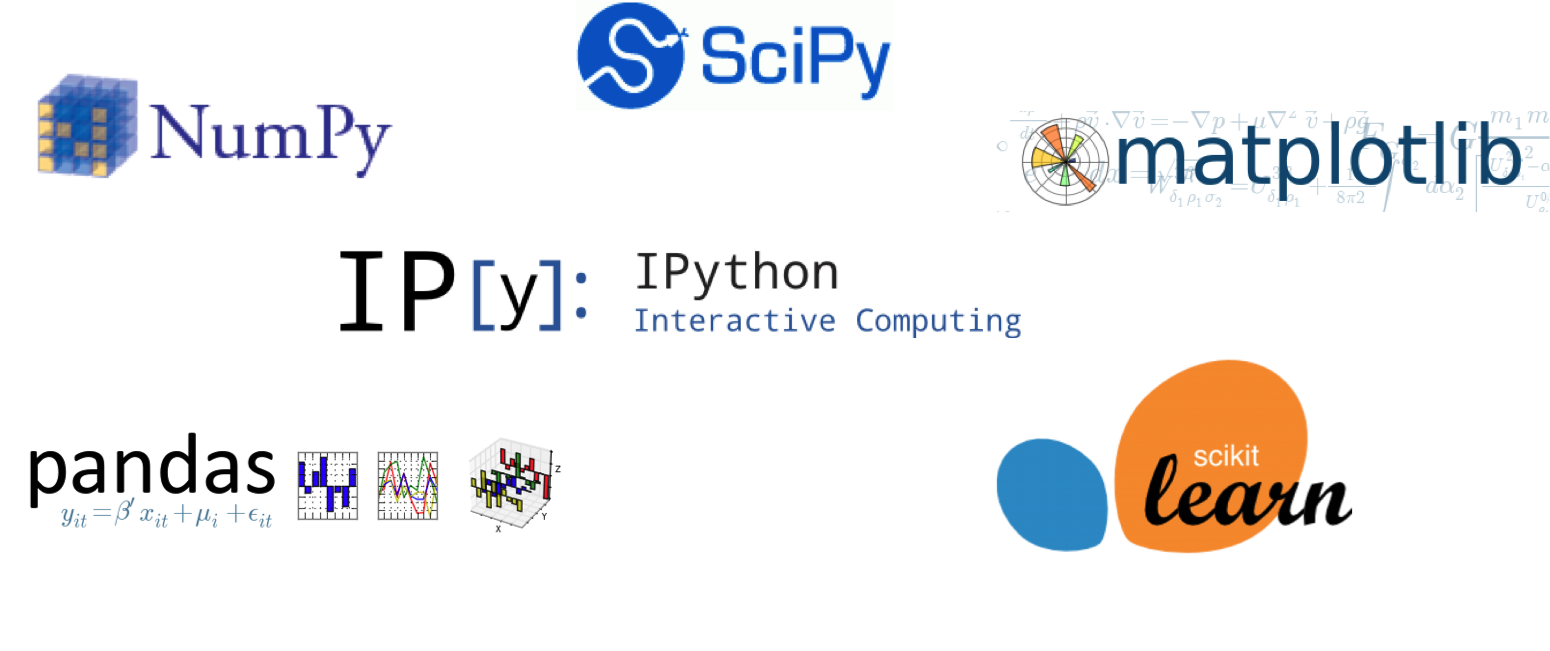
When deciding the language used for this project, we also considered languages speed. The algorithms we use require a lot of computation and this factor is very important.



This chart is comparing [Julia](http://julialang.org/) - JIT, [R](https://www.r-project.org/) - interpreted, [Python](https://www.python.org/) - interpreted (CPython), [Ruby](https://www.ruby-lang.org/) - interpreted, [Swift](https://swift.org/) - compiled (in this test interpreted due to Linux Swift limitations), [Java](http://www.oracle.com/technetwork/java/index.html) - compiled, VM, [Rust](https://www.rust-lang.org/) - compiled, [Javascript](https://www.ecma-international.org/publications/standards/Ecma-402.htm) using [Node.js](https://nodejs.org/) - interpreted, JIT, [Lua](https://www.lua.org/) - interpreted, [Nim](https://nim-lang.org/) - compiled, [PHP](https://secure.php.net/) - interpreted, [C++](https://isocpp.org/) - compiled, [Crystal](https://crystal-lang.org/) - compiled, [Go](https://golang.org/) - compiled, [C](https://en.wikipedia.org/wiki/C_(programming_language)) - compiled languages on calculating [Leibniz formula for π](https://en.wikipedia.org/wiki/Leibniz_formula_for_%CF%80). On a quick look we notice that Python has a very good median time. The faster languages are only C, C++, Nim, Crystal and Go.

Even if Python is a general purpose language, often used for other things than data science and analysis, some things make it very usefull for working with data.

Libraries give users essential functionality when managing data. Python has a lot of popular and usefull libraries for maching learning. After you get familiaried with them, they become a staple in a machine learning developing kit.



I will enumerate few of these libraries:

* **Pandas**: used for data manipulation and analysis. It’s used frequently on data preparation and munging. It is added recently to python and founds usage by data scientists.
* **Scikit-learn**: this library for machine learning has been built on NumPy, SciPy and matplotlib, it contains a lot of usefull tools for maching learning, clustering, regression, classification and reduction.
* **Seaborn**: Used for data visualization, mostly statistical. It is based on matplotlib and aims to make attractive graphics.
* and many more!

One more thing we kept in mind, was that we need a programming language that is easy to pick up. Our project has enough difficulity on understanding (business) so we don’t want to add extra problems on understanding the language we decide upon. According to a survey made of hundreds of developers from US, women and men, the results are:

1. HTML (13.3%)

2. Python (9%)

3. Javascript (6.2%)

4. PHP (4.9%)

5. Java (4.6%)

6. R (4.4%)

7. Shell (4.4%)

8. Ruby (4.1%)

9. Erlang (3.8%)

10. Go (3.6%)

This survey confirms that Python is an language with high readability and very simple syntax. It’s being surpassed by only HTML witch is a markup-language.