3. Based on the computational efficiency of implementations in Python and R, which one would you prefer? Based on a consideration of implementation (i.e., designing and implementing the code), which approach would you prefer? Taking both of these (run time and coding time), which approach would you prefer?

Method	Python Execution Time (ms)	R Execution Time (ms)
For Loop	5.51	0.0001013
Apply	1.19	0.0001856
Vectorized	0.226 (numpy)	0.00000656
Vectorized	1.37 (pandas)	N/A
iterrows	0.226	N/A

Table for Python and R Comparison

The comparison table of computation time for Python and R in basic for-loop, apply, and vectorization methods demonstrate that R considerably outperforms Python for all three comparable methods. While the dataset used is relatively small, with a size of (3, 30) after feature selection, the result might have been the opposite with large datasets due to Python's efficient memory management and optimized computation. For my personal preference, I much prefer Python to R. Python is the most popular programming language consisting of 31.19 percent of the market share while R accounts for only 0.87 percent. Although R indeed excels in statistical programming, Python also provides a great tool for statistical programming as well and shows high potential in data science and AI for its broad tool selection [1]. I have used Python much more than R for academic coursework and individual projects due to its vast eco-system and efficiency for large datasets. Therefore, I obviously prefer Python over R.

4. Identify and describe one or two other considerations, in addition to these two, in determining which of the two environments – Python or R – is preferable to you.

Aside from theoretical reasons, Python feels more comfortable to code than R. Its basic libraries such as pandas, numpy, matplotlib, and seaborn are simple yet powerful tools that are capable of many tasks in data analytics including EDA, which is one of the most critical processes of data analysis. They also provide a wide variety of necessary modules for data science. These beginner-friendly libraries and modules make Python more accessible than R in my case.
On top of that, Python has an advantage over R in multiple fields such as speed, scalability, and GPU/TPU utilization supported by powerful and highly optimized libraries such as Scikit-learn, PyTorch, and TensorFlow have made it the dominant programming language in Machine Learning and Deep Learning. Therefore, as my goal is to become a data scientist who excels in ML and DL tasks, Python is a preferred language for me.

Citations:

1. R vs Python for Data Science: An In-Depth Comparison for 2025 - Bomberbot. Bomberbot. Published January 5, 2025. https://www.bomberbot.com/data-science/r-vs-python-for-data-science-an-in-depth-comparison-for-2024/