

## Homework #5

### Team 2:

**Julian Carvajal Rico**

**James Platt Standard**

**Roberto Enriquez Vargas**

Table 1. summarizes the performance of different file formats for data handling across five matrices (A, B, C, D and, E). For matrices A, B, and C, NumPy is the fastest format for both writing (generation time) and reading (loading time), indicating high efficiency in both creating and accessing data with NumPy.

However, the smallest file sizes are different, matrices A and C are more compact in CSV format, while matrix B is smaller in NumPy format. For the much smaller matrices D and E, the HDF5 format is better than others in speed for both writing and reading, with matrix D also having the smallest file size in HDF5 format, while matrix E has a smaller size in NumPy format. This suggests that while NumPy is generally efficient for larger matrices, HDF5 may be more advantageous for smaller datasets in terms of speed and file size.

*Table 1 - Matrix Analysis*

Matrix	Fastest Write Format	Fastest Read Format	Smallest File Size Format
A	NumPy Format	NumPy Format	CSV
B	NumPy Format	NumPy Format	NumPy
C	NumPy Format	NumPy Format	CSV
D	HDF5 Format	HDF5 Format	HDF5 Dataset
E	HDF5 Format	HDF5 Format	NumPy

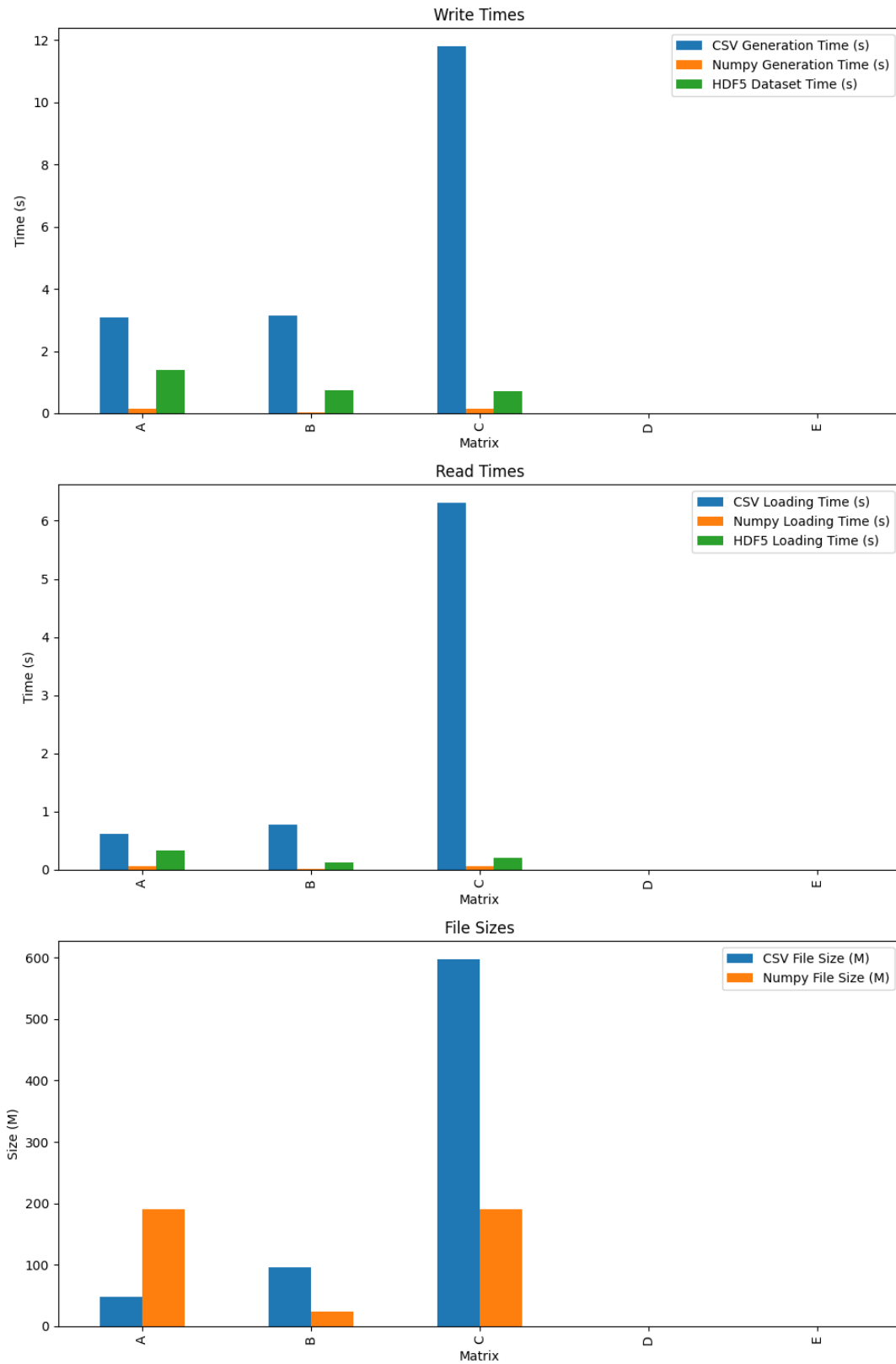


Figure 1 - Bar Plots Write Time, Read Times, and File Sizes