

COP290 Assignment 1 Subtask 1

Abhinav Shripad(2022CS11596)

January 2024

1 Introduction

We measure the time taken by each file format and the space (MB) required by each file for different file formats. The following file formats are considered:

- csv
- txt
- binary
- parquet
- json
- feather
- xlsx

2 Results

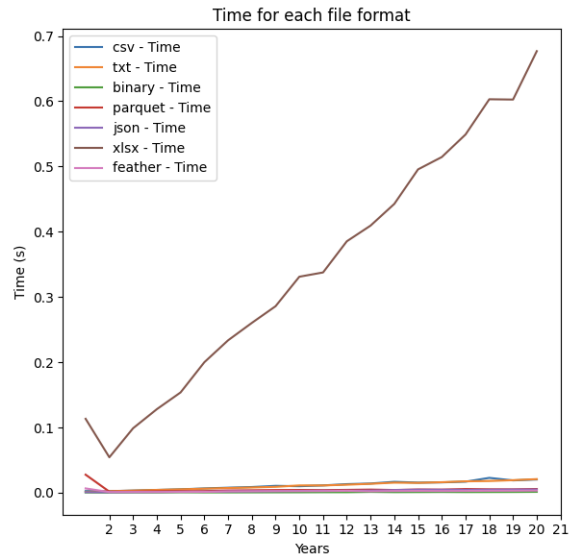


Figure 1: Time for each format

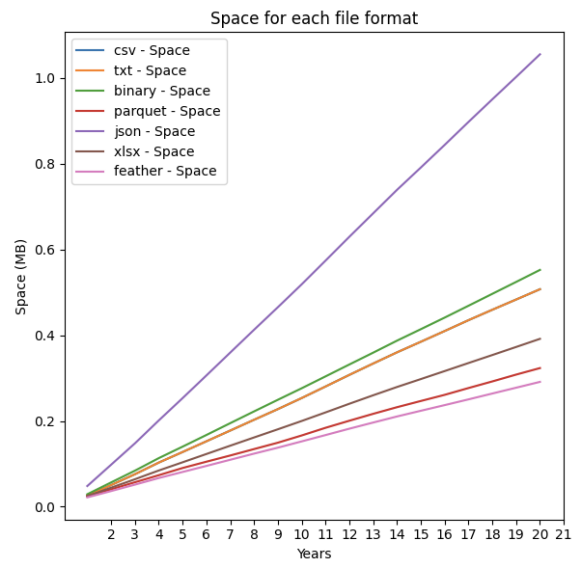


Figure 2: Space taken by each format

We graph the time and space for each format with the number of years. The results show an expected linear relationship, except for the peculiar behavior of the xlsx format.

To replicate the results, run the following command in terminal

```
python3 graph.py
```

This will create a `Format_performance.png` file.