

Why based page based format is superior to storing records using a CSV file?

CSV file is expensive in terms of data management. Typically traversing through a CSV file will take a lot of time since we have to go through the CSV file linearly(no guarantee on data size on each slot) and read the whole CSV, there is no way to access the data at a given point using a index. However this gets easy in our page file since we can access the data at one point immediately. This gives a huge performance advantage to page file over CSV file.

What is the short coming of the way we organize the page?  
1. Our page file does not have order for the data we store in it. So even we can read the data at a certain slot, we still need to traverse the page linearly.

2.After deleting the data on a certain slot, it will leave an empty space there(there is no auto adjustment of the data in it). Which cause the case where even though there might not be a lot of data in the page but the page size is still huge(which explain the graph above).

3.The record size are fixed in our page file implementation which we have less flexibility to play with it.