



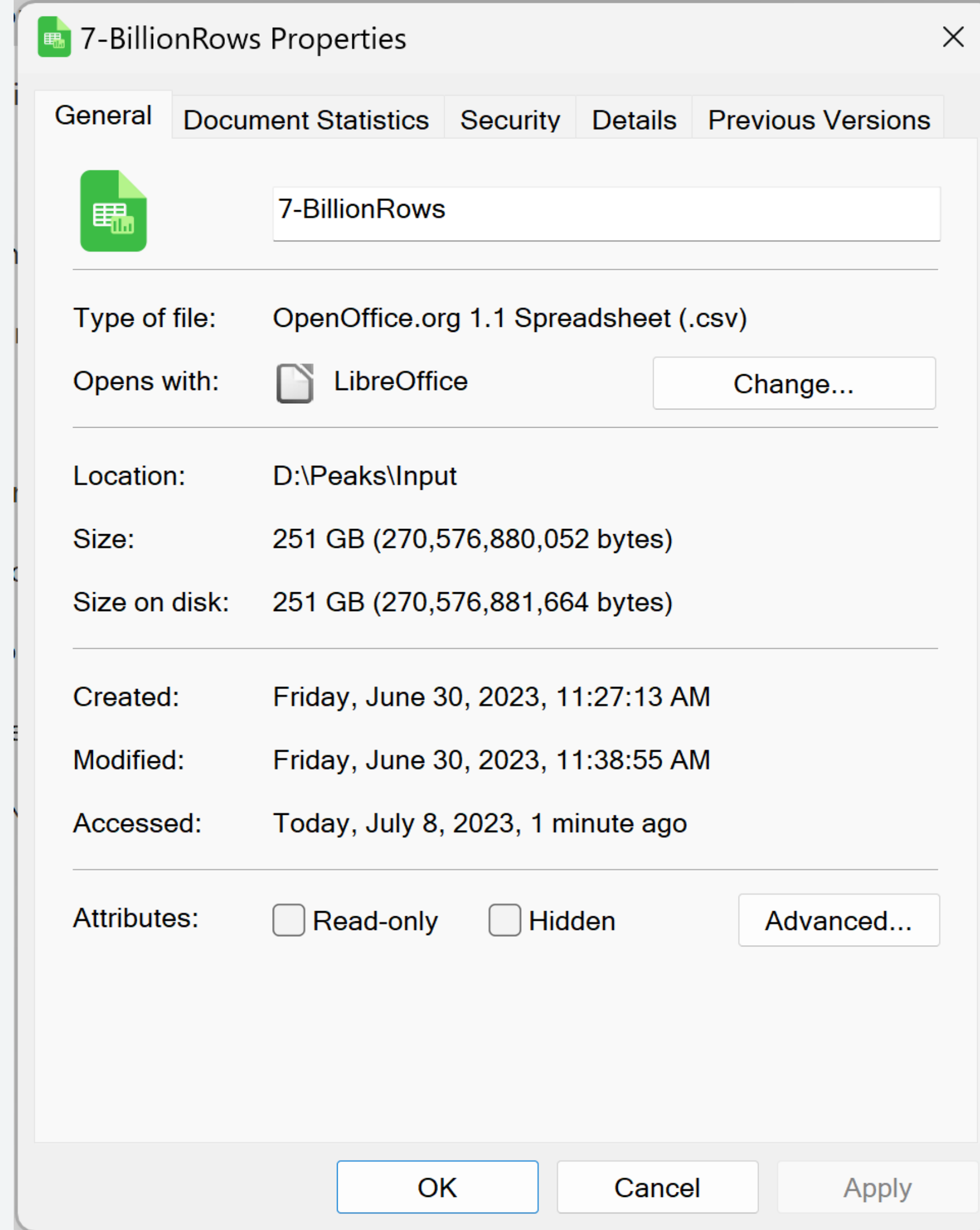
# Open & View 250GB CSV File Instantly

<https://youtu.be/IX2HKLDOfwk>

**Peaks Consolidation  
for Big Data Analytics**

# CONTENT

- 01 REQUIREMENTS
- 02 BIG 4 DATAFRAME SOFTWARE
- 03 TECHNICAL PROBLEMS
- 04 PEAKS DEMO 1
- 05 PEAKS DEMO 2
- 06 PEAKS DEMO 3
- 07 PEAKS DEMO 4





# Requirement

```
D:\Peaks\Input>dir
Volume in drive D is Data
Volume Serial Number is 8016-5DB2

Directory of D:\Peaks\Input

07/08/2023  09:45 AM    <DIR>          .
07/08/2023  09:44 AM    <DIR>          ..
06/30/2023  11:07 AM      38,653,894 1-MillionRows.csv
06/30/2023  11:14 AM    3,865,384,052 100-MillionRows.csv
06/30/2023  11:38 AM   270,576,880,052 7-BillionRows.csv
07/07/2023  09:09 PM      4,085,903 Car.csv
07/07/2023  09:09 PM      4,968,976 Catering.csv
06/30/2023  11:17 AM           130 Combination.csv
07/07/2023  09:09 PM      4,503,601 Credit.csv
07/07/2023  09:09 PM    12,499,993 FoodProduction.csv
07/07/2023  09:09 PM    11,453,862 Mushroom.csv
07/08/2023  09:45 AM    <DIR>          Output
07/07/2023  09:09 PM    18,543,697 SaleContract.csv
07/06/2023  08:47 PM      721,851 Sample.csv
07/07/2023  09:09 PM    15,024,278 SocialNetwork.csv
12 File(s) 274,552,720,289 bytes
3 Dir(s)  27,285,061,632 bytes free
```

- You would like to quickly preview sample rows of these CSV files.
- You would like to view all columns.
- Some of them have over one million rows.
- In the extreme case, one file has 7 billion rows.

**You may have tried some  
data frame software.**

- Pandas
- Polars
- DuckDB
- Spark

# Requirement

- You would ~~like~~ **\$** to quickly preview sample rows of these CSV files.
- You would like to view all columns.
- Some of them have over one million rows.
- In the extreme case, one file has 7 billion rows.

# Technical Problems

- CPU capability  
=> may be very slow
- Memory size  
=> out of memory
- View all columns  
=> screen size
- Text/number alignment issues

# Requirement

- You would ~~like~~ **\$** like to quickly preview sample rows of these CSV files.
- You would like to view all columns.
- Some of them have over one million rows.
- In the extreme case, one file has 7 billion rows.





**Preview CSV File**

**DEMO 1**

**7 Billion Rows 250GB**

[github.com/hkpeaks/peaks-consolidation](https://github.com/hkpeaks/peaks-consolidation)

# **Do FileName.csv**

**where Do is the file name of the Peaks runtime for  
Windows and Linux**





**Preview CSV File**

**DEMO 2**

**Many Columns**

[github.com/hkpeaks/peaks-consolidation](https://github.com/hkpeaks/peaks-consolidation)





**Preview CSV File**

**DEMO 3**

**Text/Number Alignments**

[github.com/hkpeaks/peaks-consolidation](https://github.com/hkpeaks/peaks-consolidation)



**Preview CSV File**

**DEMO 4**

**Preview Particular Position**

[github.com/hkpeaks/peaks-consolidation](https://github.com/hkpeaks/peaks-consolidation)



# **ReadSample for Coming 2nd Pre-release**

**Peaks Consolidation  
for Big Data Analytics**

