

BI - Data Visualizations

BigData of 25 million records

Which DV (Data Visualization) Tool to use?

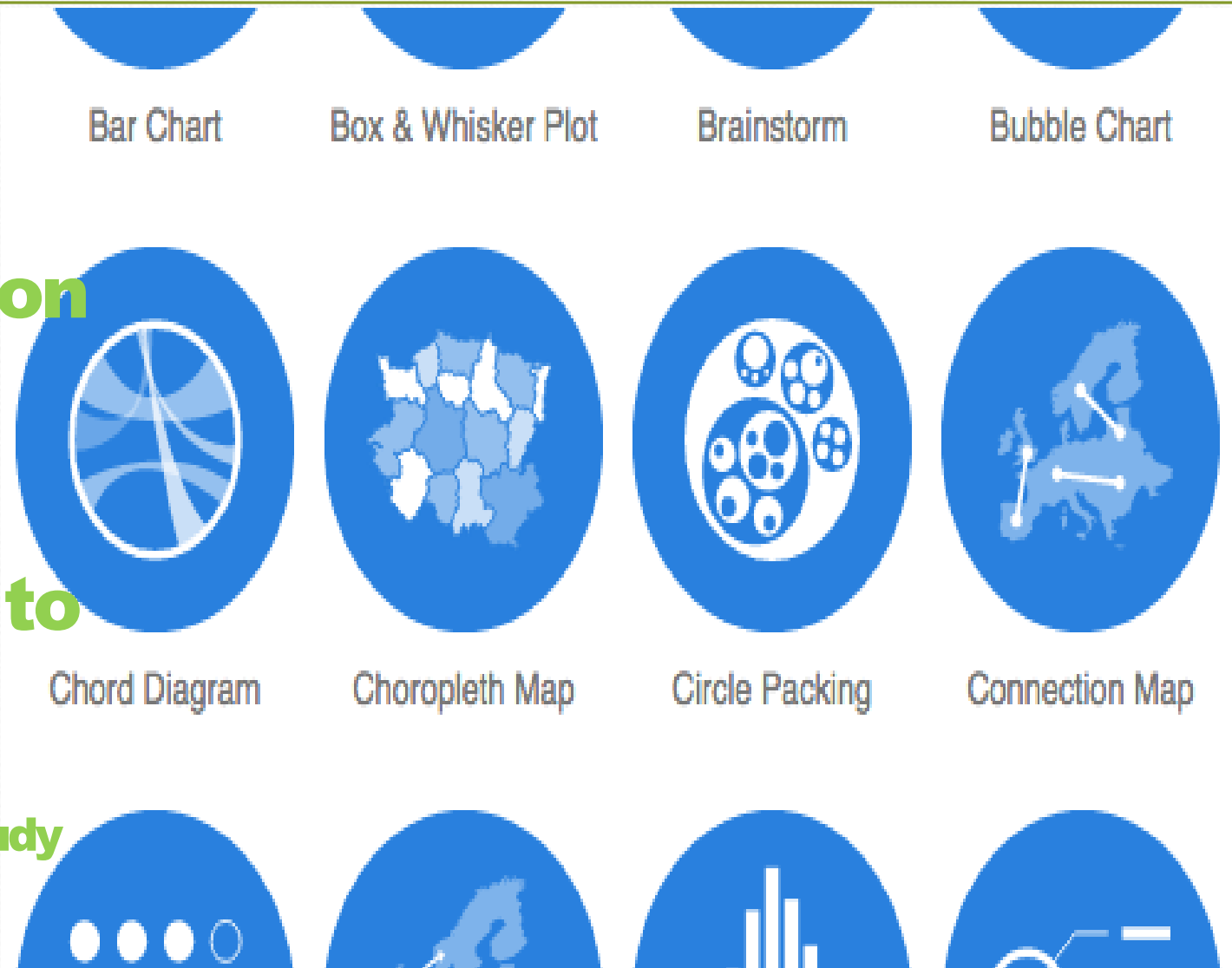
Quick Tool Evaluation Case Study

By Mohee Jarada (IT Project Manager)

Email: jarada.mohee@gmail.com

V1.0

Picture Internet Source: <https://twitter.com/infogram/status/667116103278206976>



BI Business Challenge: BigData of 25 or more Millions of CSV records

Which DV (Data Visualization) Tool can import and handle such data?

- *Facts/Inputs:*
 - *CSV: 25 Millions of CSV records – generated using my Java CSV fake data generator available at GitHub (free to use – you need just JRE v8+ to be installed priorly):*
[*https://github.com/mohijarada2017/CsvDataFileGeneratorRepo*](https://github.com/mohijarada2017/CsvDataFileGeneratorRepo)
 - *My notebook specs: 6 GB RAM with Intel i5 2.5Ghz machine*
 - *Ms-Windows 10*
 - *DV Tools installed are:*
 - *Ms-PowerBI v2.4 (desktop free edition)*
 - *Tableau v10.3 (desktop trial free edition for 1 month)*
 - *MicroStrategy v10.8 (desktop free edition)*
 - *Tibco Spotfire v7.10 (desktop free edition for 1 month)*

Let us examine each DV tool and see who's the Winner

Ms-PowerBI v2.4 – Is it able to handle that BigData?

The screenshot displays the Microsoft Power BI Desktop v2.4 interface. The ribbon at the top includes tabs for File, Home, View, and Modeling. The Home tab is active, showing various options like Paste, Cut, Copy, Format Painter, Get Data, Recent Sources, Enter Data, Edit Queries, Refresh, Solution Templates, Partner Showcase, New Page, New Visual, Text box, Image, Shapes, From Store, From File, Manage Relationships, Relationships, New Measure, New Column, and Publish. A 'Load' dialog box is open in the center, displaying a yellow warning banner that reads '1 of the loaded queries contained errors.' with a 'View errors' link. Below the banner, a green checkmark indicates 'SampleCSVFile_20million_records_mohi' and a red triangle indicates '12,496,468 rows loaded. 2 errors.' A 'Close' button is at the bottom right of the dialog. On the right side, the 'Visualizations' and 'Fields' panes are visible. The 'Fields' pane shows a table with columns labeled Column1 through Column9. The bottom status bar shows 'PAGE 1 OF 1' and the Windows taskbar at the very bottom.

*At importing 25 millions of CSV records,
data importing resulted in errors and only 12.5 millions of CSV
records were available to DV user*

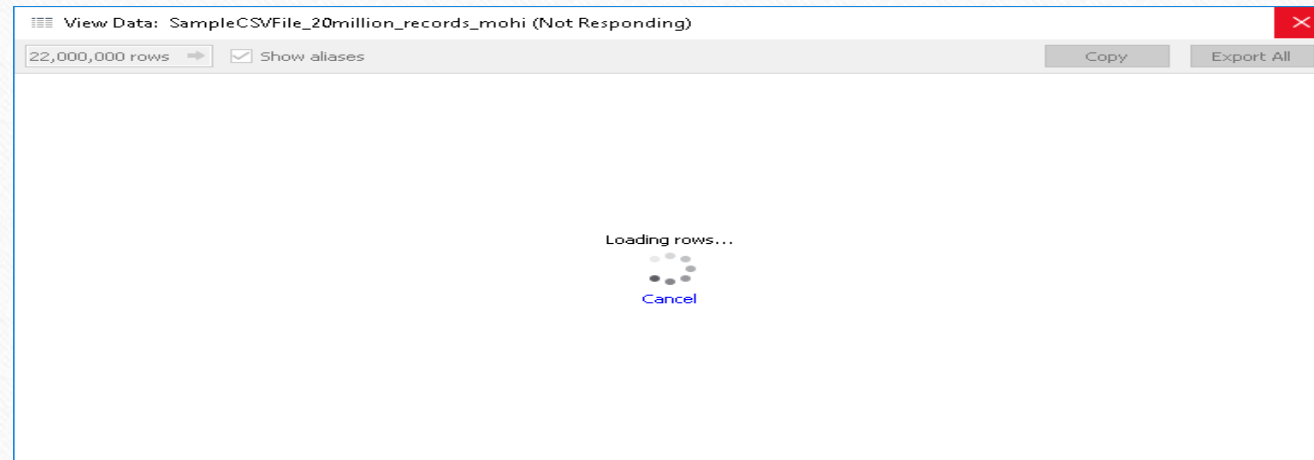
Tableau v10.3 – Is it able to handle that BigData?

*At importing 25 millions of CSV records,
data importing resulted in DV
Tool freeze and system's restart!*

View Data: SampleCSVFile_20million_records_mohi

12,496,468 rows → ☒ Show aliases Copy Export All

F1	F2	F7
Eldon Base for stackable storage shelf, platinum	Muhammed MacIntyre	Nunavut
1.7 Cubic Foot Compact "Cube" Office Refrigerators	Barry French	Nunavut
Cardinal Slant-D® Ring Binder, Heavy Gauge Vinyl	Barry French	Nunavut
R380	Clay Rozendal	Nunavut
Holmes HEPA Air Purifier	Carlos Soltero	Nunavut
G.E. Longer-Life Indoor Recessed Floodlight Bulbs	Carlos Soltero	Nunavut
Angle-D Binders with Locking Rings, Label Holders	Carl Jackson	Nunavut
SAFCO Mobile Desk Side File, Wire Frame	Carl Jackson	Nunavut
SAFCO Commercial Wire Shelving, Black	Monica Federle	Nunavut
Xerox 198	Dorothy Badders	Nunavut
Xerox 1980	Neola Schneider	Nunavut
Advantus Map Pennant Flags and Round Head Tacks	Neola Schneider	Nunavut
Holmes HEPA Air Purifier	Carlos Daly	Nunavut
DS/HD IBM Formatted Diskettes, 200/Pack - Staples	Carlos Daly	Nunavut
Wilson Jones 1" Hanging DublLock® Ring Binders	Claudia Miner	Nunavut
Ultra Commercial Grade Dual Valve Door Closer	Neola Schneider	Nunavut
#10-4 1/8" x 9 1/2" Premium Diagonal Seam Envelopes	Allen Rosenblatt	Nunavut



MicroStrategy v10.8 – Is it able to handle that BigData?

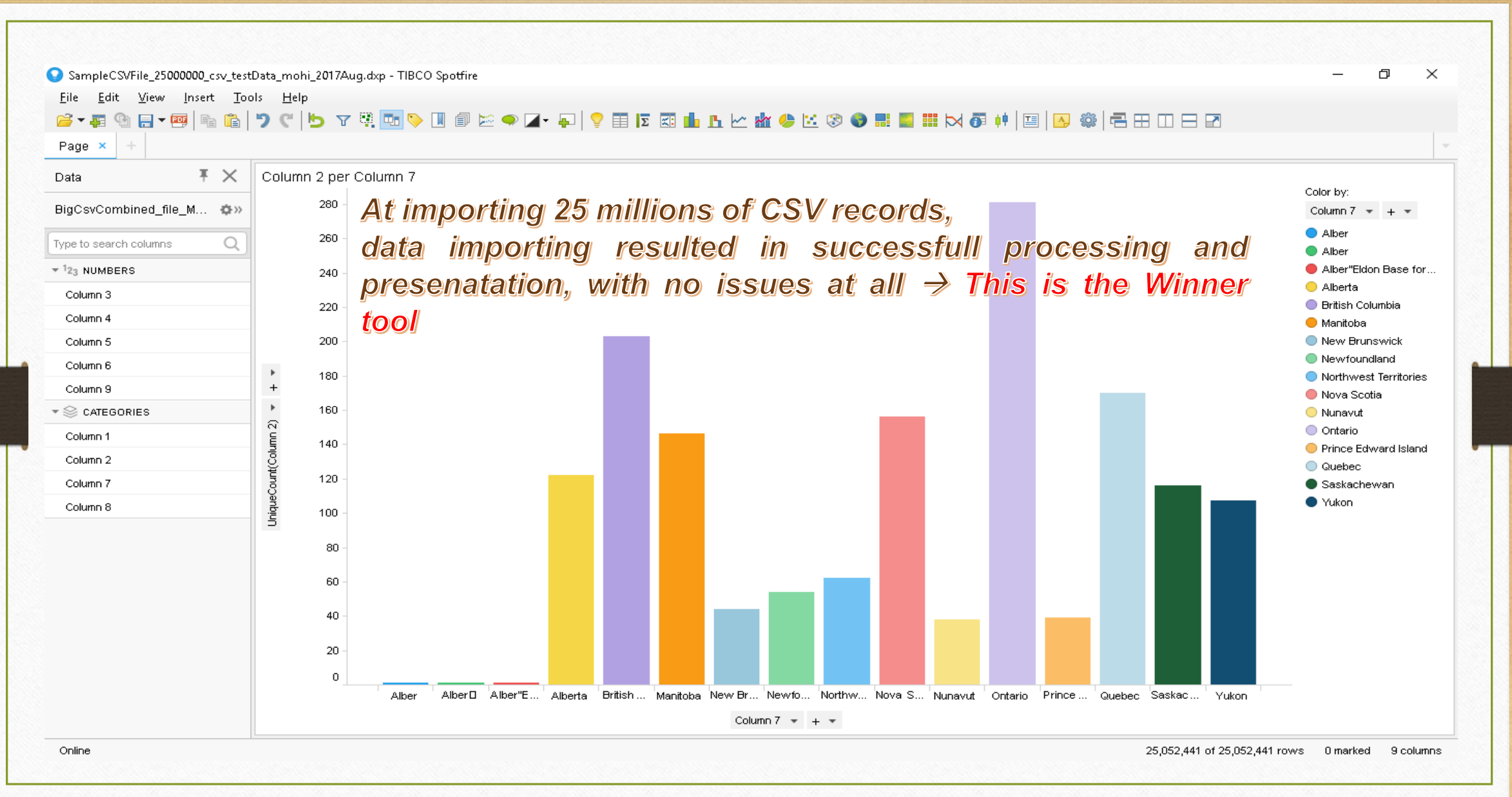
The screenshot displays the MicroStrategy Desktop application window. On the left, the 'DATASETS' pane shows a list of datasets, with 'SampleCSVFile_20million_record...' highlighted in green. A red arrow points from this dataset to the 'Row Count' metric in the 'Metrics' pane. The 'Row Count' metric is also highlighted in a red box in the 'Metrics' pane, and its value, '12,496,467', is displayed in the 'Visualization 1' pane. The 'Visualization 1' pane is currently empty, showing only the 'Row Count' metric.

1) I imported 1st CSV file: 12.5 million of CSV records: done in less than 2 mins and really good in Microstragy 10.8 trial version.

Now

2) I closed the program and try to import 25 million, but as you can see it failed! Looks to me the winner among all DV tools: Spotfire v7

Tibco Spotfire v7+ – Is it able to handle that BigData?



Thank you...