World of Computer Science and Information Technology Journal (WCSIT)
ISSN: 2221-0741
Vol. 9, No. 6, 28-33, 2019

The Challenges of Big Data Visual Analytics and Recent Platforms

Hoda A. Abdelhafez
1. Faculty of Computers & Informatics, Suez Canal University, Ismailia – Egypt
2. College of Computer & Information Sciences, Princess Nourah University, Riyadh, Saudi Arabia

Abeer A. Amer
Sadat Academy for management and Sciences, Alexandria, Egypt

Abstract— Visual Analytics plays an important role in discovering hidden information from massive, heterogenous and streaming data. Visual analytics using visual representations and interactive techniques that are combined with statistical and machine learning methods for analysis process. Big data visual analytics faces many challenges related to technological issues and human cognition. This paper's aim is to focus on the challenges of big data visual analytics and how the recent platforms including Knime, SAS visual analytics, Arcadia enterprise and TensorFlow could deal with these challenges. It also provides comparison between these platforms. The results show that these platforms can overcome most of the general challenges. TensorFlow is a promising platform for handling challenges related to interaction and user interfaces. Other specific challenges like uncertainty and human cognitive bottlenecks still need more efforts.

Keywords- Big data; Visual Analytics; Challenges; Visual Analytics platforms.

I. INTRODUCTION

Big data represents massive amounts, complex and growing datasets, which are generated from various sources like sensors and social networks [1], [2]. Big Data has many characteristics that differentiate it from traditional data which mentioned by set of V's characteristics. The main three characteristics are (volume, variety, velocity). Volume refers to the enormous amount of data being generated from various sources. Variety represents the different types and sources of data (relational databases, social media, e-mail, weblogs, sensors, photographs, and videos [3]. Velocity means the speed of the data generation, processing, and aggregation.

Some of researchers add other Vs, which are veracity, validity, variability, volatility, visualization and value. Veracity refers to the completeness and accuracy of data which make a leader trusts information in order to make decision [4]. Validity means that the prepared data is correct and accurate to use. Volatility refers to the stored data and how long is useful to the use. Visualization means complex graph of all the huge amount of data to be easy to understand and read [5] [6]. Variability represents the data arrives from different sources and how efficiently it differentiates between noisy data or important data. Value is an important feature derived from the results of data analysis to help in decision making [1][7].

Visual analytics is an important method used to analysis and visualize big data. It is defined as "visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces" [8]. Visual Analytics is more than just visualization it represents an integrated approach combining data analysis, visualization and human factors [9]. The aim of big data visual analytics is to take advantage of human's cognitive abilities in visualizing information while applying computer's capability in automatic analysis [10]. Therefore, decision makers can use visual analytics techniques to combine their human imagination, flexibility, and knowledge with the large storage and processing capacities to gain valuable insight to solve their complex problems [11], [12], [13].

The paper is organized as follows; Section 2 defines and illustrates the 5Ws dimensions and behaviors patterns. Section 3 introduces the five densities with additional axes and shrunk attributes in parallel coordinates. Section 4 shows the implementation of our model. Section 5 explains related works. Section 6 summarizes our conclusions and explores future works.

II. RELATED WORK

In big data visual analytics, few works focus on the visual analytics challenges for large dataset. Cui (2019) presented a comprehensive survey of visual analytics including its evolution from visualization and algorithmic data analysis and how it is applied in various application domains. The study also discussed the major challenges characterized by the scalability, interaction, infrastructure, and evaluation from both technical and application perspectives. The scalability challenge involves both human and machine limitations. Interaction challenge is to investigate its cognitive and perceptual impacts for integrating human judgment in the data-analysis process. Infrastructure challenge means the lack of a visual-analytics framework that works on high-performance computing platforms. There is a need of libraries and frameworks to support multiple levels of abstraction. The evaluation is a challenge because of the complexity of a visual-analytics applications. The study illustrated some future directions of the visual analytics [14].

Rohrer et al. (2014) illustrated that visual analytics is important for big data analysis. The authors discussed the challenges of visualizing big data as well as offering some approaches and strategies to overcome these challenges for effective analysis. They discussed research program at the National Security Agency covering both the problems and opportunities of big data visual analytics. They concluded that supporting analysis of big data is more important than technical issues of data management, storage, and computation [15].

Keim et al. (2008) presented an overview of big data visual analytics and its challenges. The authors focused on the scope of visual analytics and identifying the formal model of visual analytic process. This model includes data sets, hypotheses, visualizations and insight as well as transition functions from one concept to another. They also addressed the important application challenges and the significant technical challenges that are related to the field of visual analytics [16].

III. CHALLENGES OF BIG DATA VISUAL ANALYTICS

Visual analytics as a subset of data visualization, it deals with combining visualization, human factors and data analysis [8]. The challenges of big data toward visual analytics are divided into general and specific challenges [15]. The general challenges are:-

- In-memory analysis causes design and implementation challenges including algorithms, interactive analytics, memory, workflow, I/O, and threading.
- Databases and storage, the extreme-scale databases cause both a hardware and a software problem.
- Traditional algorithms do not deal with the scalability issue and unable to reduce the effort of data analytics and exploration
- Data movement and network infrastructure
- Parallelism, redesign visual analytics algorithms that allow parallel processing and enable interactive data analytics.
- Domain and development libraries, frameworks, and tools are challenge for visual analytics applications.
- Active engagements of society, community, government.

The specific challenges that are relevant to visual analytics include: Large-data visualization, uncertainty quantification and interaction and user interfaces. Large-Data visualization focuses on data presentation in visual analytics. This requires building visual representations at multiple levels of abstraction and scale as well as dimension-reduction techniques for handling extreme data scales. Uncertainty quantification is critical for any decision maker. Visual analytics tasks should cope with risks and incomplete data. It is important for visualization to accurately convey uncertainty to minimize misleading results and help decision makers understand risks. Interaction and user interfaces, the growing volume of data and the unchanged human cognitive abilities cause barrier in human computer interaction.

The challenge of interaction and user interfaces especially for extreme-scale visual analytics includes several issues. These issues are: (1) share cores in the hardware execution units and reduce the workflow disruption due to human computer interaction, (2) user-driven data reduction needs flexible mechanism, (3) navigating deep multilevel hierarchy and searching for optimal resolution are major issue for scalable analysis, (4) how to represent the evidence and uncertainty of large-scale data without causing significant bias through visualization, (5) Analyze the interrelationships among heterogeneous data objects and extract the right amount of semantics from large-scale data, (6) Data summarization and triage for interactive query especially if the data size exceeds petabytes, (7) Develop effective visual analytics techniques that can handle dynamic data exploit the humans' cognitive ability, (8) find alternative methods to compensate the weaknesses of human cognitive ability, (9) Establish standardized design and engineering resources for high performance computing systems, and (10) Information visualization mantra falls when applied to extreme-scale data [13].

IV. VISUAL ANALYTICS PLATFORMS

There are many visual analytics platforms that can handle big data. The Knime, SAS visual Analytics and Arcadia Enterprise are some of visual analytics platforms that could deal with the challenges of big data. Another open source platform is also discussed in this section called TensorFlow, which includes machine learning visualization library.

A. Knime Analytics Platform

KNIME, The Konstanz Information Miner is an open source analytics platform including multitude of nodes for machine learning and data analysis. It integrates with many scripting languages. It enables using big data connector for accessing and using Hadoop via Hive and provides Spark connector to deploy Apache Spark [17]. KNIME platform allows user to model workflows that consist of nodes to process data. The connections between those nodes are used to model, modify, transform, or visualize the data. In Knime, all data flowing between nodes is wrapped within a class called DataTable. These data are accessed through iterating over instances of DataRow. The data access using Row ID or index are avoided for scalability issue and ability to handle large amount of data. KNIME applies a powerful caching strategy that moves parts of a data table to the hard drive if it becomes too large. Knime analytics platform supports various data types: structured and unstructured or time series data [18]. It also sort, Aggregate, filter, and join data either on local machine, in-database, or in distributed big data environments.

KNIME analytics platform offers visualization nodes such as scatter and scatter matrix plots, boxplots, histograms and parallel coordinates. Knime provides interactive and normal views. The interactive view is used for data exploration, which requires to change the certain parameters of the view without reconfiguring and executing the view. Both principal components analysis (PCA) and multidimensional scaling methods (MSD) are available for calculating new coordinates such as scatter plot views using color, size and shape information. Color/size/shape manager nodes are used for adding additional visual dimensions. Hieratical clustering view node is also used to visualize hieratical clustering [19], [20].

B. SAS Visual Analytics

SAS Visual Analytics offers a complete platform for analytics visualization, identifying patterns and relationship in data. SAS visual analytics uses in memory data storage. Its memory component is SAS LASR Analytic Server in earlier release or cloud analytics server (CAS) in the new release. Both servers use random access memory to execute and accelerate analytic computations in-memory with high performance. SAS Visual Analytics platform includes SAS home/hub, data builder, explorer, designer, graph builder and administrator. SAS home/hub is an entry point to the tool. Data builder enables users to join and summarize data as well as enhancing the predictive power of their data. The explorer is highly visual, drag-and-drop data interface combined with the speed of the SAS LASR Analytic Server to accelerate analytic computations and enable organizations to derive value from large amounts of data. The designer enables users to create reports to view them on the web or on a mobile device. The graph builder enables users to create graph objects for use in the reports. The Administrator in SAS visual analytics supports administrative tasks [21], [22].

SAS visual analytics works with large volume of data through in-memory engine to speed the task of data exploration and visualization. It provides new visualization techniques based on core fundamentals of data analysis such as network diagram, word cloud, scatter plot and correlation matrix as well as the basic charts. It also uses correlation matrix to combine big data and fast response times to quickly identify variables and their relationship. SAS visual analytics also has autocharting and filtering capabilities to help the user to refine and present the most appropriate visualization of data based on the volume and the type of these data [23], [24].

C. Arcadia Enterprise

Arcadia Enterprise is native visual analytic platform for accessing, analyzing, designing, and sharing insights and data-centric applications. It applies In-cluster and in-cloud architecture for running analytics directly within cloud instances, Hadoop nodes, or other scale-out modern data platforms. Arcadia Enterprise was built using the latest web technology as well as native HTML5 to offer fast business insights with lower cost. It is not required data movement to get high-definition access to all granular data in its native state. It can explore data direct in-cluster distributed processing without having to start with extracts, cubes, or data marts to gain insights from big data [25].

Arcadia Enterprise dashboards is used to organize the visualization. Dashboards can display and link visuals that are based on different datasets across different connections. Arcadia Enterprise implements advanced analysis techniques for visuals, such as trellises, dimension hierarchies, and Dual Axis. Arcadia Enterprise offers several types of visuals such as network, interactive maps, word cloud, dendrogram and more. It also has real-time streaming visualizations with granular and time-based filters. The user can pull the most recent data from streaming sources like Apache Kafka and refresh the visuals without manual polling [26].

D. TensorFlow

TensorFlow machine intelligence platform is an open source Google's software used to allow users to utilize and control large-scale and heterogeneous environments. It provides deep learning with visual analytics. It has dataflow graphs to represent the computation in an algorithm and the state on which the algorithm operates [27], [28], [29]. Edges in TensorFlow graphs serve three different purposes. First, data dependency edges represent tensors, which are input and output data of the operations. Second, reference edges represent pointers to the variable allowing dependent operations to mutate the referenced variable. Third, control dependency edges indicate that their source operations must execute before their tail operations can start [30]. TensorFlow maps the nodes of a dataflow graph across many machines in a cluster, and within a machine across multiple devices to give flexibility to the application developers. A scripting interface wraps constructing dataflow graphs and allows users to experiment with different model architectures and new optimization algorithms [27], [28].

TensorBoard is machine learning visualization library based on Google TensorFlow. It is a graph visualization software included with any standard TensorFlow installation to understand debug and optimize TensorFlow programs more easily [25]. TensorBoard includes modules for monitoring scalar values, distribution of tensors, audio and images. The users could use TensorBoard to visualize some statistics to monitor the training process. The user can do certain operations in a TensorBoard-activated TensorFlow program, these operations are exported to event files. TensorBoard converts these event files to graphs that can give insight into a model's behavior. In fact, TensorFlow as a machine learning platform provides interactive exploration of the dataflow architecture behind machine learning models, but the weakness is the static dataflow graph, especially for algorithms such as deep reinforcement learning [31].

Table 1 shows the comparison between four visual analytics platforms. These platforms can deal with big data characteristics: volume, variety and velocity of big data. In general challenges of big data visual analytics, Knime, TensorFlow and Arcadia enterprise platforms could overcome most of these challenges. Although, SAS visual analytics uses in-memory analysis, it is one of the general challenges. The table also illustrates that four visual analytics platforms using abstraction and dimension-reduction techniques to overcome the data presentation as a specific challenge. The other specific challenges like uncertainty and human cognitive bottleneck still need more efforts to overcome them. Moreover, Tensorflow improves the interaction and user interfaces through an inbuilt module for deconvolutional layer. This module gives insight into what types of features deep neural networks are learning at specific layers.

Table 2 represents many visualization methods that are used in Knime, SAS and Arcadia Enterprise platforms bar, line and pie charts as well as scatter plot are used in these platforms as simple visualization methods. For multidimensional data visualization, parallel coordination, scatter plot matrix is used in Knime platform. Maps method are also used like treemap in Knime platform and interactive maps in Arcadia Enterprise. TensorBoard offers many views for the training set which include scalar values, computational graph of the model, histograms and distributions. It also visualizes audio, image and text data.

Table 2: VISUALIZATION METHODS

Visualization Method	**Knime Analytics Platform**	**SAS visual Analytics**	**Arcadia Enterprise**
Bar chart	x	x	x
Line chart	x	x	x
Pie Chart	x	x	x
Histogram	x	-	x
Scatterplot	x	x	x
Scatterplot Matrix	x		
Heatmap	x		x
Parallel coordinates	x		
Boxplot	x		
Dendrogram	x		x
Bubble plot	x	x	
Maps	x		x
Chord chart	x		x

I. CONCLUSION

The visual analytics of big data is used to help decision makers examine large volume of heterogenous, multi-dimensional and streaming data to gain insights, solve critical problems and take effective decisions. Visual analytics implements visual representations and interaction techniques during the analysis process. There are many challenges of visual analytics for handling big data. These challenges related to technical issues, data representation, interaction and user interfaces. Some of the recent and promising visual analytics platforms tried to overcome these challenges. These platforms are Knime, SAS visual analytics and Arcadia Enterprise and TensorFlow. The comparison between these platforms shows that some challenges such as uncertainty and human cognitive weaknesses still need more efforts in order to solve them. On the other hand, TensorFlow has its dataflow graph library through TensorBoard, which includes built-in components to help users understand, debug, and optimize TensorFlow models

REFERENCES

[1] S. Owais and N. Hussein, "Extract Five Categories CPIVW from the 9V's Characteristics of the Big Data," International Journal of Advanced Computer Science Applications, Vol. 7(3), pp. 254–258, 2016.

[2] P. Arockia, S. Varnekha, and K. Veneshia, "The 17 V's of Big Data," International Research Journal of Engineering Technology., Vol. 4(9), pp. 3–6, 2017.

[3] S. Sagiroglu, and D. Sinanc, "Big Data: A Review, Collaboration Technologies and Systems (CTS)" International Conference on Digital Object Identifier, PP. 42-47, 2013.

[4] I. El-Alaoui, Y. Gahi, R. Messoussi, A. Todoskoff, and A. Kobi, "Big Data Analytics: A Comparison of Tools and Applications", SCAMS 2017, LNNS 37, Springer International Publishing AG, part of Springer Nature, pp. 587–601, 2018

[5] H. Jasim, A. Shnain, S. Hadishaheed, and A. Haji, "Big Data and Five V'S Characteristics," International Journal of Advances in Electronics and Computer Science, Vol. 2(1), pp. 2393–2835, 2015.

[6] U. Sivarajah, M. Kamal, Z. Irani, V. Weerakkody, "Critical analysis of Big Data challenges and analytical methods", Journal of Business Research, Vol. 70, PP. 263-286, 2017.

[7] R. Barik, R. Lenka, S. Ali, N. Gupta, A. Satpathy, and A. Raj, "Investigation into the efficacy of geospatial big data visualization tools," Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2017, pp. 88–92, 2017.

[8] D. Keim, F. Mansmann, J. Schneidewind and H. Ziegler, "Challenges in Visual Data Analysis", IEEE 2006, PP. 9-16, 2006.

[9] J. Kohlhammer, D. Keim, M. Pohl, G. Santucci, G. Andrienko, "Solving Problems with Visual Analytics", The European Future Technologies Conference and Exhibition, Vol. 7, PP. 117-120, 2011.

[10] L. Ren, Y. Du, S. Ma, X. Zhang, G. Dai, "Visual analytics towards big data," Journal of Software, Vol. 25(9), PP. 1909-1936, 2014.

[11] J. Hilda, C. Srimathi and B. Bonthu "A Review on the development of big data analytics and effective data visualization techniques in context of massive and multidimensional data", Indian Journal of Science and Technology, Vol 9(27), PP. 1-13,

2016.

[12] P.C. Wong, H. Shen, and V. Pascucci, "Extreme Scale Visual Analytics." IEEE Computer Graphics and Applications, Vol. 32(4), PP. 23-25, 2012.

[13] P. C. Wong, H. Shen, C. R. Johnson, C. Chen, and R. B. Ross, “The top 10 challenges in extreme-scale visual analytics.” IEEE Computer Graphics and Applications, Vol. 32(4), PP. 63–67, 2012.

[14] W. Cui, “Visual Analytics: A Comprehensive Overview”, IEEE Access, Vol. 1, PP. 81555-81574, July 2019.

[15] R. Rohrer, C. Raul and B. Nebesh “Visual Analytics for Big Data”, The Next Wave”, Vol. 20(4), PP. 1-17, 2014.

[16] D. Keim, F. Mansmann, J. Schneidewind, J. Thomas, H. Ziegler, “Visual Analytics: Scope and Challenges”. In: S. J. Simoff, M. H. Böhlen, A. Mazeika (eds) Visual Data Mining. Lecture Notes in Computer Science, Vol 4404, Springer, Berlin, Heidelberg, 2008.

[17] A. Fillbrunn, C. Dietz, J. Pfeuffer, R. Rahn, G. Landrum, M. Berthold, “KNIME for reproducible cross-domain analysis of life science data”, Journal of Biotechnology, Vol. 261, PP. 149-156, 2017.

[18] M. R. Berthold, N. Cebron, F. Dill, T. R. Gabriel, T. Kötter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, “KNIME: the Konstanz information miner”, ACM SIGKDD Exploration Newsletter, Vol. 11, PP. 26-31, 2009.

[19] G. Bakos “Knime Essentials”, Packt Publishing, ISBN:1849699216, 2013.

[20] M. Berthold, C. Borgelt, F. Höppner, F. Klawonn, “Guide to Intelligent Data Analysis: How to Intelligently Make Sense of Real Data”, Springer, ISBN-10: 1848822596, 2010.

[21] SAS Institute Inc. “SAS® Visual Analytics 7.3: User’s GuideSAS”, SAS Institute Inc., Cary, NC, USA, 2015.

[22] T. Anderud, R. Collum and R. Kumpfmiller, “An introduction to SAS visual Analytics: How to explore numbers, design reports, and gain insight into your data”, SAS institute Inc., 2017

[23] J. Choy, V. Chawla, and L. Whitman, “Data Visualization Techniques: From Basics to Big Data with SAS® Visual Analytics”, White paper, SAS Global Forum, 2012.

[24] P. Chate, A. Karad and U. Patkar, “Big Data Visualization: Challenges and SAS Visual Analytics”, International Journal of Computer Science and Mobile Computing, Vol.4(12), PP. 44-48, 2015.

[25] Arcadia Data, “gaining insights from big data: Scaling Analytics on Apache Hadoop and the Cloud”, Whitepaper, Arcadia Data, Inc., 2017, http://go.arcadiadata.com/rs/627-XIL-022/images/Arcadia_Data_Native_Visual_Analytics_for_Big_Data_Whitepaper.pdf

[26] Arcadia Data, “Arcadia Enterprise Documentation: Visualization Guide”, Arcadia Enterprise 5.0.0.0 (ArcViz 5.0.0.0, ArcEngine 2.11.2), last access July, 2019. http://documentation.arcadiadata.com/latest/#pages/topics/visualization-guide.html#visualization-guide

[27] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Man´e, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Vi´egas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems”, Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI ’16), Savannah, GA, USA, 2016.

[28] K. Wongsuphasawat, D. Smilkov, J. Wexler, J. Wilson, D. Man´e, D. Fritz, D. Krishnan, F. Vi´egas, and M. Wattenberg, “Visualizing Dataflow Graphs of Deep Learning Models in TensorFlow”, IEEE Transactions on Visualization and Computer Graphics, 2017.

[29] F. Hohman, M. Kahng , R. Pienta, and D. H. Chau, “Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers”, IEEE Transactions on Visualization and Computer Graphics, VOL. 25(8), PP. 2674-2693, August 2019.

[30] K. Wongsuphasawat, D. Smilkov, J. Wexler, J. Wilson, D. Man_e, D. Fritz, D. Krishnan, F. B. Viegas, and M. Wattenberg, “Visualizing dataflow graphs of deep learning models in TensorFlow,” IEEE Trans. Vis. Comput. Graph., vol. 24(1), pp. 1–12, Jan. 2018.

[31] C. Huyen and D. Hafner “TensorFlow for Deep Learning Research” Lecture Notes, Stanford University, 2018, http://web.stanford.edu/class/cs20si/lectures/notes_02.pdf

TABLE 1: COMPARISION BETWEEN VISUAL ANALYTICS PLATFORMS

Description	Knime Analytics Platform	SAS visual Analytics	Arcadia Enterprise	TensorFlow
License	Open Source GPLv3 license	SAS Institute, Inc.	Arcadia Data, Inc.	Open Source Google's software
Data sources	Text formats like CSV, JSON and XML Unstructured data types like images, networks and molecules Time series data Databases and data warehouses Data from social media like twitter and facebook	Data from social media like Twitter streams, Google Analytics and Facebook Call center logs, Online comments text-based documents Data from operational systems	Historical resources such as relational data, delimited files, Excel documents, and business applications Real-time multi-structured data from NoSQL databases Cloud sources such as Amazon S3	Data from images, text files, CSV files, Audio, Video, Structured data
Architecture	In database processing The data analysis process consists of a pipeline of of nodes, each node processes the arriving data and/or models Able to add new processing nodes and distribute them	Combine in-memory technologies with easy-to-use exploration interface and drag-and-drop analytics capabilities. Its components are SAS home/hub, data builder, explorer, designer, graph builder and administrator	In-cluster and in-cloud architecture runs analytics directly within cloud instances, Hadoop nodes, or other scale-out data platforms Unifying real-time, batch and interactive analytics	Tensorflow is based on a multi-dimensional array TensorFlow uses dataflow graph, shared state, and the operations, which mutate that state. It maps the nodes of a dataflow graph across several machines in a cluster, and within a machine across multiple computational devices
Key features	- Visual workflows - Machine learning models - Code free setup and intuitive interface - Expose workflow fragment as RESTful web service	- Self-service, easy analytics - Robust report design, creation and viewing - Statistical analysis and machine learning capabilities - Deployment flexibility	- Open data access - Inherit privileges and authentication from existing systems - Multitude of visual types to build production-quality dashboards & applications - Single copy of data	- Interactive multiplatform programming interface - Build and train machine learning models - TensorBoard Visualizer - Using intuitive high-level APIs with eager execution
3 Vs of big data	- Variety: open platform using broad types of data sources and extensive tool integration - Volume: It has big data extensions that cover Hadoop based data integration and aggregation - Velocity: It has distributed execution of heavy workflows as well as high performance scoring of predictive models	- Variety: It provides visualization techniques for visualizing semi- structured and unstructured data. - Volume: It works with large volume of data through in-memory engine - Velocity: intelligent autocharting function provides correlation matrix that quickly identify which variables among billions are related and the strength of their relationship	- Variety: It can connect multiple data sources together into rich, interactive applications - Volume: It is a massive parallel processing (MPP) analytics platform running inside big data environments and modern storage engines. - Velocity: immense scale of analysis, with access to billions of records in near real-time via a simple, intuitive drag-n-drop interface	- TensorFlow is one of big data technologies used to provide deep learning models. It could run on multiple CPUs or GPUs and mobile operating systems. - Tensorflow uses large volumes of data to train the model, - It also uses a variety of data types
Abstraction & aggregation for visualization	- Aggregation & filtering in-database, or in distributed big data environments	- Binning(grouping) the data - Summarize the distribution of a dataset - Using hierarchies for drill-down	- Using granular path analysis to drill down to raw data - Using Dimensional hierarchies and hierarchical cross tabulation	Embedding projector offers methods of dimensionality reduction
Advanced analytics for visualization	- Offer modular framework enables integration of new modules or nodes, and allows for interactive exploration of trained models or analysis results - Advanced predictive and building machine learning models	- Using several analytic strategies such as graph or network analysis and stochastic process analysis - Interactive data visualization to explore and make sense of data.	- Using behavior-based segmentation and correlation across dimensions and metrics - Apply blending visualizations across sources through creating semantic relationships across multiple sources, and set hierarchies & logical datasets - Apply flow and funnel algorithms	- TensorFlow offers TensorBoard to visualize the TensorFlow graph, plot quantitative metrics about the execution of the graph and any additional data - Tensorflow has an inbuilt module for deconvolutional layer