Impact Factor (SJIF): 3.632



International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Big Data Visualization Using Excel Features

Ruchita Jani¹

P.G. Student, Department of Computer Engineering, Atmiya Institute of technology and science College, Rajkot,

Gujarat, India¹

ABSTRACT: Now a days, data is growing in every field with a way faster rate due to growth of social media. So, visualization is an important approach to helping big data get a complete view of data that can help decision making. There are many tools for visualizing data in graphical format. Among all, Excel is the most used tool by industry. Now a days it has been added with lots of new features so that it can work with big data as well. So, the most used tool for visualization, its challenges and improvements in in context of big data is introduced in this paper.

KEYWORDS: Big Data Visualization, Visualization Methods, Microsoft Excel, New Features, Limitations.

I. INTRODUCTION

In this technology trend, Data is growing way faster than before. People are sharing their information easily in every field on internet via social media. To process this huge amount data, there are lots of tools developed. Big data [1] Visualization is a representation of data in pictorial or graphical format. It enables us to evaluate things visually so that we can identify new patterns. Now, the main issue is visualizing massive data in order to understand and extract knowledge from stored data and give efficient decision making through its reports. In simple words, [1] the collection of information is no longer a problem but extracting of valuable knowledge from the collected or available information is the main thing.

Visualization methods are very important as it provides mental model of data to users. It makes huge and complex information intelligible. [1] The basic purpose of visualization is to create interactive visual representations of the huge information that increase capability of problem solving of human. User should be able to easily understand and interpret huge and complex set of information.

There are lots of tools for visualizing information. Visualization Approaches [2] are the techniques used to create tables, images, diagrams and other intuitive display ways to understand data. Big Data visualization is not that easy like traditional relative small data sets because of the complexity in big data. The extension of traditional visualization approaches are already emerged but far away from enough. Among all, here we are going to cover one tool that is Microsoft Excel, the most popularly used tool by IT industries and individuals as well. Microsoft Excel is very easy to use but it is not that smaller we think. It is having many features inside it. It has launched with effective and powerful new features to make better representation of massive data. Still it has some problems that leads to a challenge and need extension. The main objective of this paper is to provide basic knowledge about various methods and progress of visualization techniques. This also includes the future enhancements needed to improve current scenario.

II. VISUALIZATION METHODS

A number of visualization techniques have been developed in last few years due to need of representing huge amount of data and providing better decision making from these reports. All the tools mostly provide some basic kind of conventional techniques that make user easily understand the data in brief. So, we are going to discuss few known traditional techniques and give its example in order to get initial idea about features with Microsoft Excel tool.

1. Table

Table is a very basic, simple and easy to implement, easy to understand and interpret data representation technique. Table [1] is structured format including rows and columns. Row has different synonyms like tuple, vector etc. Columns are also referred as field, attribute etc. The role of table is key part in research and analysis. Example of excel table is as given in below figure.

100	A Cut	ne mise	n Page	Layout P	ormulas	Data	ceview .	View			
Conur			Calibri • 11 • A* A* = = 😸 🗞								
Paste	e d Form	at Painter	BIU	B Z Ų · ⊞ · 🎝 · A · ≣ ≡ ≡							
	Clipboard	in Parricer III		Font	5		All	anme			
	J8		(m	fx				a free free free free free			
14	A	в	С	D	E	F	G	T			
1	ID	Temp	Distance	Gender	Age						
2	1	24	32.0	male	28						
3	2	28	27.0	female	22						
4	3	19	35.5	male	33						
5	4	25	28.2	female	42						
6	5	21	32.6	female	21						
7	6	25	28.8	female	28						
8	7	26	29.1	female	33						
9	8	25	29.5	female	41						
10	9	18	37.2	male	29						
11	10	26	27.9	female	36						
12	11	28	26.5	male	35						
13	12	26	28.1	male	18						
14											
15											
16											

Fig 1: Showing simple table

2. Pie Chart

Pie chart is also called the circle graph. [1] Pie chart circle include number of sectors, each sector or portion describe a proportion in a whole quantity. It is mainly used to determine size of particular part of common data with other parts of data. Mostly sector is shown by percentage as shown in below figure.



Fig 2: Standard pie chart

3. Bar Chart

Bar chart is one of the most commonly used method for representing data mostly for discrete data rather than common data. Bar chart is the diagram in which numerical values of variables are represented by the height or length of lines or rectangles of equal width as shown in below figure [3].

International Journal of Advance Research in Engineering, Science & Technology (IJAREST) Volume 3, Issue 10, October 2016, e-ISSN: 2393-9877, print-ISSN: 2394-2444





4. Line chart

A line chart or line graph is a type of chart which displays information as a series of data points called 'markers' connected by straight line segments. It is a basic type of chart common in many fields. A line chart is mostly used to represent a trend in data over intervals of time as shown in below figure.



Fig 4: Line chart

5. Area chart

This is used to display quantitative data graphically. It is mainly based on the line chart. The area between axis and line are commonly emphasized with colors and textures. It uses the area chart for showing trends over time among related attributes as shown in below figure.



Fig 5: Area chart

6. Scatter plot

Scatter plot is graphical display of set of data in Cartesian coordinate, shows the relationship between two variables, one variable represent horizontal distance (independent variable) and second variable vertical distance (dependent variable) of data point from the coordinate axis [4]. Scatter plot shows the how strong the relationship are between the variables, and determines whether their exit any outlier in the data or not. It is use to look how the data is dispersed. Example of plot is shown in the following figure [5].



Fig 6: Scatter plot

III. POWER FEATURES IN EXCEL

There are many more basic techniques to represent data in excel like tree map, Entity relationship diagram, flow charts, venn diagrams, etc. These kind of techniques help user understand the data in quick way. Now, the focus is on massive data. Every field is now coming with big data. Many tools and techniques have been developed to process and analyze this big data. Excel was unable to give that much effective performance when analyzing and representing big data. It was getting slower to process the huge data to analyze and represent.

So, to overcome performance issues with massive data, Excel has been launched with 4 new features which are Power Pivot, Power query, Power view and power map. Power Pivot [6] is an Excel add-in that user can use to perform powerful data analysis and create sophisticated data models. With Power Pivot, user can mash up large volumes of data from various sources, perform information analysis rapidly, and share insights easily. Any data that is imported into Excel is available in Power Pivot, and vice versa. Power Query [7] for Excel can be found as an add-in in Excel 2013 which is used to simplify data discovery and allow easier access to different data sources. Power View [7] and Power Map are developed to improve the visualization of Microsoft Self-service BI. Power View can help users to create interactive reports and analytical views. Power Map enables the 3D map experience in Excel, with this feature, users can explore and navigate data in a particular geographical area.

1. Power Pivot

By using Power Pivot for Excel, we can transform enormous quantities of data with better speed into meaningful information to get the answers of queries. With Power Pivot for Excel, we can import millions of rows of data from multiple data sources into a single Excel workbook, create relationships between heterogeneous data, create calculated columns and measures using formulas, build PivotTables and Pivot Charts, and then further analyze the data so that you can make timely business decisions all without requiring IT assistance [8]. It overcomes existing limitations of traditional techniques for massive data analysis and representation.

International Journal of Advance Research in Engineering, Science & Technology (IJAREST) Volume 3, Issue 10, October 2016, e-ISSN: 2393-9877, print-ISSN: 2394-2444



Fig 7: Reports with Power Pivot add-in [10]

As shown in above figure excel is with the add-in power pivot. It can be viewed in power pivot window by clicking on 'Manage'. By this, a separate window of power pivot will be opened that provide all the features i.e. formatting, sorting, filtering, creating relationships between tables, etc. We can add data from multiple sources by its button Get External Data. Then data can be shaped, analyzed and visualized.

🍋 💷 😓 🖘 🗟 = =		Power Pivot for Excel - Cross Platform Users Report (Weekly) - 12.28.								
File Home Design Advar	nced									
Paste Append Paste Replace Paste Copy	rom Data From O Service • Source	ther Existing es Connections	Refresh	PivotTable	Data Type : Text ▼ Format : Text ▼ \$ ~ % ?	2↓ Sort A to Z ↓ Sort Z to A ↓ Clear Sort Filters Column →	Find			
Clipboard	Get External Da	ita			Formatting	Sort and Filter	Find			
[TitleFriendly Ass	assin's Creed IV	: Black Flag								
TitleFriendlyName	Platform 🔽	TitlePlatformK	ey		🖥 🔽 Add Column					
1 Alien: Isolation	Both	Alien: Isolation	Both							
2 Assassin's Creed IV: Black Flag	Both	Assassin's Cree	d IV: Blac	k FlagBoth						
3 Battlefield 4	Both	Battlefield 4Bo	th							
4 Call of Duty: Advanced Warfare	Both	Call of Duty: Ad	dvanced \	NarfareBoth	1					
5 Call of Duty: Ghost	Both	Call of Duty: G	nostBoth							
6 Child of Light	Both	Child of LightB	oth							
7 Costume Quest 2	Both	Costume Ques	t 2Both							
8 Destiny	Both	DestinyBoth								
9 Diablo III: Reaper of Souls – Ul	Both	Diablo III: Reap	er of Sou	ls – Ultimat	e					
10 Dragon Age: Inquisition	Both	Dragon Age: In	quisition	Both						
11 Duck Dynasty	Both	Duck DynastyB	oth							
12 Fantasia: Music Evolved	Both	Fantasia: Musio	Evolved	Both						
13 Far Cry 4	Both	Far Cry 4Both								

Fig 8: Power Pivot window

2. Limitations

Although Power Pivot feature provide far better performances than traditional techniques with massive amount of data, it has some limitations that need to be overcome.

International Journal of Advance Research in Engineering, Science & Technology (IJAREST) Volume 3, Issue 10, October 2016, e-ISSN: 2393-9877, print-ISSN: 2394-2444

1. File size of excel worksheet increase: When huge amount of data is loaded in power pivot, size of the excel worksheet increases considerably. As shown in below figure, file size is 194MB for 2,576,500 records of data. So, it is not much suitable to have this much size for bigger data.

Date modified	Туре	Size
6/13/2016 7:35 AM	File folder	
9/27/2016 11:50 PM	Microsoft Excel W	194,204 KB
	Date modified 6/13/2016 7:35 AM 9/27/2016 11:50 PM	Date modified Type 6/13/2016 7:35 AM File folder 9/27/2016 11:50 PM Microsoft Excel W

Fig 9: File size increase

🕸 l 💷 🔙 🖘 🐡	• 1	F	owerPivot	for Excel - Eng	jagement K	Pls Dashboard (Monthly	y) - Aug 2016.xl	lsx			- 0	x
File Home De	sign Advanced										~	^ 🕜
Paste Append	From From Data From Other Database * Service * Sources	Existing Connections	PivotTable	Data Type : Who Format : Genera \$ • % > 3	al *	∮↓ Sort Smallest to Largest ⋠↓ Sort Largest to Smallest ⊘ Clear Sort	Clear All Sort by Filters Column •	Find	∑ AutoSum ▼	Data Diagram Show Hidden	fx Calculation Area	
Clipboard	Get External Data			Format	ting	Sort and Filt	ter	Find	Calculations	View		- Inc.
[Geographyid] •	33							_				٢
🔟 Monthid 💌 Geo	grap 🚯 💌 SubsMigrati 🕏	😼 🔽 UsageMigration	ntd 👘	Tenu 😘	Platfo	😘 🖬 UsageTypeId 🗖	TrxTypeId 💌	Lo 💌	Subscipti 🕫	Add Column		^
201506	0 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		_
201506	3 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	9 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		_
201506	9 Dual Paid	Both		91 + days	All	Game SP + MP	Paid	1		17		
201506	18 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		_
201506	33 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	27 Dual Paid	Both		91 + days	All	Game SP + MP	Paid	1		17		
201506	28 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	31 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	34 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	26 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	26 Dual Paid	Both		91 + days	All	Game SP + MP	Paid	1		17		
201506	46 Dual Paid	Both		91 + days	All	Game SP + MP	Paid	1		17		
201506	50 Dual Paid	Both		91 + days	All	Game SP + MP	Paid	1		17		
201506	47 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	54 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	86 Dual Paid	Both		91 + days	All	Game SP + MP	Non-Paid	1		17		
201506	84 Dual Daid	Roth		91 + dave	All	Game SD + MD	Daid	1		17		~
DimGeography DimPlat	form DimSubsOfferShort DimSubs	MigrationState Dimil	sageMigration	State DimTenu	re BaseVals	EngagementVals SystemA	nns Engagement	loggedilser				~
Record: H 6 of 2.57	6.500 H	ingrationatate Dimo	sagegrador	Store Shirleito	ie busevais	Lesbabeneretara Lakarenne	hina Crangagement	Copped Oser	<u> </u>			

Fig 10: Bigger Data

2. Data Refreshes stuck at some places: Due to heavy data, power pivot cannot refresh all tabs in one shot because it stuck. And sometimes excel also crashes. So, we have to refresh all tabs sequentially which is time consuming.

🚇 🛯 🖬 🛃 ನೀರೇ 🗸			PowerPivot for Excel - E	Engagement KPIs Dashboard	l (Monthly) - Aug 2016.x	dsx			_ 0	x
File Home Des	ign Advanced									^ ?
Paste Append	From From Data From Other Database * Service * Sources	Existing Connections	R Data Type : V	Whole Number ▼ A J Sort Smalles heral ▼ A J Sort Largest Data Refresh	to Largest to Smallest	Find	∑ AutoSum ▼	Data View View Hidden	fx alculation Area	
Clipboard	Get External Data		Data Refresh Progress			Find	Calculations	View		
[GeographyId] 🔻	33		Refreshing gets updated data fro	m the original data sources						¥
🔟 Monthid 🖬 Geog	rap 🚯 💌 SubsMigrati 🦄	😼 🔽 UsageM	lg		*	Lo 🔽	Subscipti ។	💵 Add Column		^
201506	0 Dual Paid	Both		10	Total O Cassellad	1		17		
201506	3 Dual Paid	Both		10	Success 9 Error	1		17		
201506	9 Dual Paid	Both				1		17		
201506	9 Dual Paid	Both	Details:			1		17		
201506	18 Dual Paid	Both	Work item	Status	Message	1		17		
201506	33 Dual Paid	Both	Dim Dim Distrom	Error	Error details	1		17		
201506	27 Dual Paid	Both	Dim SubsOfferShort	Error	Error details	1		17		
201506	28 Dual Paid	Both	DimSubsMigrationState	Error	Error details	1		17		
201506	31 Dual Paid	Both	Dim Usage Migration State	Error	Error details	1		17		
201506	34 Dual Paid	Both	Dim Tenure	Error	Error details	1		17		
201506	26 Dual Paid	Both	BaseVals	Error	Error details	1		17		
201506	26 Dual Paid	Both	EngagementVals	Error	Error details	1		17		
201506	46 Dual Paid	Both	EngagementLoggedUsers	Error	Error details	1		17		
201506	50 Dual Paid	Both	SystemApps	Success. 107 rows transferred.		1		17		
201506	47 Dual Paid	Both				1		17		
201506	S4 Dual Paid	Both				1		17		
201506	84 Dual Paid	Both				1		17		~
				Stop	Refresh Close					*
DimGeography DimPlatfo Record: I 4 6 of 2,576	orm DimSubsOfferShort DimSub	sMigrationState	DimUsageMigrationState DimTe	enure BaseVals EngagementVal	s SystemApps Engagemen	tLoggedUser	s			

Fig 11: Data Refresh stuck

3. Sometimes, it doesn't respond quickly when too heavy data is loaded i.e. gives slow performance.

XI	Engagement KPIs Dashboard (Monthly) - Aug 2016.xlsx - Excel (Not Responding) – 🗖 📑												×	
File		Home Insert	Page Layout	Formulas Dat	a Review	View Develope	LO4	AD TEST Pow	er Pivot TEAN	1 V Tell	me	Punita Jani (Tata Consulta	incy Services) 🔉 Share	₽_₩
Paste	Clip	Cut Calil Copy * Format Painter board 15	bri • 12 I <u>U</u> • EE • Font	2 • A • A • =	E =	Wrap Text	ter + 1	General \$ * % * 58 Number		al Format as • Table • Styles	Cell Insert Del Styles* Ce	► Format • Clear •	Sort & Find & Filter * Select *	^
E17	E17 • : × ✓ fr ==FERROR(E15/E13,***) •													٣
Ö	RURALSS Engagement KPIs Dashboard June 2015 Users Selection													
P	Check the Definition tab for known issues with some of the data. Geography													
	ull I			Xbox360			Xhoy	Ine			Area		Construction	
							, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	inc.			India	Japan	Greater China	
				Jun-2016		M-o-M	97 A	h 2045	Y-0-Y	97 A	MEA	UK	United States	
		Total Base		3.106	May-2016	-100	-3.1%	2 618	+488 **	+18.6%	Country			
		Total Active Users		855	972	-117 🗱	-12.0%	1,369	-514 🕽	-37.5%	country			_
		% of Base		27.5%	30.3%	-2.8% 🚦	-9.2%	52.3%	-24.8% 🕽	-47.4%	Germany			
		Total Engaged		802	904	-102 🗱	-11.3%	1,302	-500 🕽	-38.4%				
1.1	۲	% of Base		25.8%	28.2%	-2.4%	-8.4%	49.7%	-23.9% 🕽	-48.1%				
	ē	Games Users		712	788	-76 🖁	-9.6%	1,213	-501 🕽	-41.3%				_
	E.	% of Base		22.9%	24.6%	-1.7%	-6.7%	46.3%	-23.4% 🕽	-50.5%				
	8	% of Engaged		88.8%	87.2%	+1.6%	+1.8%	93.2%	-4.4%	-4.7%				_
	Ë,	Multiplayer Users		440	498	-58 🕱	-11.6%	646	-206 3	-31.9%		Xbox LIVE S	ubscription	
		% of Engaged		14.27	10.0%	-1.4% 8	-0.6%	24.7%	-10.5%	-42.0%	Tier	3= .T <u>x</u>	Payment Class	-
		Apps Users		468	509	-41	-8.1%	691	-223	-32.3%	Gold		Free	
		% of Base		15.19	15.9%	-0.8%	-5.1%	26.4%	-11.3%	-42.9%	Channel	5	Duration	
		% of Engaged		58.4%	56.3%	+2.0%	+3.6%	53.1%	+5.3%	+10.0%				
		Avg. Titles		5.1	5.1	+0.0	+0.0%	5.3	-0.2	-4.5%	Online	Unknown	1 3 12	
		Avg. Days Logged-In		10.5	11.0	-0.7 🐰	-6.4%	10.5	-0.1	-1.3%	Pricing			1
		Avg. Presence Hours		#GETTING_DATA	#GETTING_DATA	n/a n/	а	*****	n/a n	/a	Free	Paid Bonus	Month Paid Discou	int 👻
4	- F	Summary	Dashboard	Charts Defini	tions System	Apps (+)			a/a a	4	1 S			
Ready	0	a a a a a a a a a a a a a a a a a a a		single for the second	in a system							III III	四	► 90%

Fig 12: Not Responding

4. Basically, entire massive data is loaded to power pivot permanently so that it acquires memory

IV. CONCLUSION AND FUTURE SCOPE

There are number of tools for massive data visualizations in market. All these uses basic traditional approaches to import or dump all of the data to respective tool. In this paper, Excel has been described with its new features which are very effective but having some limitations. Although all these conventional visualization techniques are very useful for understanding the data, there is a need for new methods and approaches to improve the big data visualization scenario.

REFERENCES

- M. Khan, S.S. Khan, "Data and Information Visualization Methods and Interactive Mechanisms: A Survey", International Journal of Computer Applications, vol. 34, no. 1, 2011, pp. 1-14.
- [2] C.L. P. Chen, C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on Big Data", Information Sciences, 275 (10), August 2014, pp. 314-347.
- [3] JavaScript Column Charts | canvasjs. CanvasJs. N.p., 2016. http://canvasjs.com/html5-javascript-column-chart/. Accessed: 2016-10-01.
- [4] Utts, Jessica M. (2005). "Seeing Through Statistics", 3rd Edition, Thomson Brooks/Cole, 2005, pp 166-167
- [5] What is Scatter plot? : 2009. www.psychwiki.com/wiki/what_is_a_scatterplot%3F. Accessed: 2016-10-02.
- [6] Power Pivot: Powerful data analysis and data modeling in Excel: 2016. https://support.office.com/en-us/article/Power-Pivot-Powerful-dataanalysis-and-data-modeling-in-Excel-D7B119ED-1B3B-4F23-B634-445AB141B59B. Accessed: 2016-10-02.
- [7] Mengwei Lu (2014), "Discovering Microsoft Self-service BI solution: Power BI", Haaga-Helia University of Applied Sciences, 2014, pp 1-94
- [8] Power Pivot Add-in: 2016. https://support.office.com/en-us/article/Power-Pivot-Add-in-a9c2c6e2-cc49-4976-a7d7-40896795d045. Accessed: 2016-10-02
- Power Pivot: Powerful data analysis and data modeling in Excel: 2016. https://support.office.com/en-us/article/Power-Pivot-Powerful-dataanalysis-and-data-modeling-in-Excel-D7B119ED-1B3B-4F23-B634-445AB141B59B. Accessed: 2016-10-02.
- [10] Optimize a Sample PowerPivot Model for Power View Reports: 2016. https://technet.microsoft.com/enus/library/hh759324(v=sql.110).aspx. Accessed: 2016-10-02.