Arunkumar Navaneethan

arunkn79@gmail.com 520-369-1363 | 1255 E University Dr #341 AZ 85281



Scan QR code to visit http://arunethan.com



- IT/Services
- IT/Development IT/Research

LinkedIn network

Regular blogger on Business Intelligence and enterprise innovation specific topics at http://arunethan.com/

Result driven technology professional with refined business skills and a successful track record in application development, database management and information systems management

- MS in Management Information Systems; University of Arizona; Eller College of Management; (12/2013)
- B.Tech in Information Technology; Anna University, Sri Sairam Engineering College; Chennai India; (4/2010)

Crestbrook, Nationwide Insurance (through Sogeti)Consultant, Business Information managementSQL Server, Report Management, Tableau, Master Data Management(May 2014 to Present)

- Rationalizing approaches to provide a single version of the truth for enterprise reporting
- Creating T-SQL, PL/SQL scripts to interface data between enterprise applications
- Developing reports in Tableau and Excel to support internal financial compliance and sales departments growth assessment
- Evaluating business intelligence applications to create standardized approach towards report development

American Realty Capital Properties (through RIIM)Developer, Business IntelligenceSQL Server, Tableau, Microsoft Cubes, Dimensional modeling(January 2014 to May 2014)

- Analyzed transactional and dimensional data present in SQL server and Microsoft CUBE using Tableau
- Created standard reports for property management and internal strategy group
- Enhanced the reporting environment performance by reviewing the Microsoft Cube and Tableau server configurations

Sanofi (through University of Arizona) Innovation management Application, SharePoint

Consultant, Information System Management

(June 2013 to December 2013)

Developer, Application Development

(July 2010 to July 2012)

- Interacted with cross functional managers from USA and Europe to create a web application to manage innovation within the organization using Spigit application platform
- Explored the functionalities of in-house data management and business intelligence applications, capitalized the out of the box features to match the business needs
- Pitched a proof of concept using technologies like SharePoint, Yammer, Denodo data virtualization and inhouse text analytics platform to manage internal events activities and engagements

University of Arizona Graduate student consultant, Information system management MySQL, Oracle, Data mining, Social media analytics, Data visualization (August 2013 to December 2013)

- Reviewed and enhanced the database system of the marketing department, updating the stored procedures, triggers, views of the Oracle DB system
- Performed predictive analysis using classification, clustering, association rule mining and network analysis; descriptive analysis using MySQL for back-end and Google visualization API for front-end
- Analyzed user behavior, demographics, device access patterns to enhance new acquisition and current members ease of web experience for StartupTucson.org venture company

Inautix, Bank of New York Mellon Mainframe, J2EE, DB2, Oracle DB, IBM Lombardi BPM

- Involved actively throughout the software development life cycle; planning, requirement gathering end-to-end support, individually managed the application documents, recorded project progress
- Created reporting using Fusion chart API and parsed data from Oracle DB to JSON format using Java programs
- Handled change request, bug fixes, testing and test region code deployments. Lead the production support activities during the vendor transition and the internal group training sessions
- Developed new workflows, migrated and redesigned code from Java to Lombardi authoring environment
- Interacted with stakeholders like client, business analyst, business process management architect and database architect

Project Samples

Social Media Analytics





Articles trends and performance Article's temporal analysis Number of twee scattered Mediatic and Topological Topological Topological Topological Category based analysis Mashab 3.42 Average degree 6.31 11.21 Average weighted degre 25.6 % of hashtag 33.2 Trending topics NYCC, iphone, cor ews, iphone5s, endag 5c, iphone4s, iphone

Article's temporal analysis







Prezi

Number of tweets and creation date (day level) Mashable and Engadget tweets



Record level analysis

View Data: Sheet 1						_ 6 ×
Show Aliases 🗜 Show	all fields					Copy
coordinates_latitude	coordinates_longitude	created_at	text	favorited	filter_level	id str
0.0	0.0	9/10/2013 5:00:01 PM	On @ Engedget	false	nedum	37747656564041267
0.0	0.0	9/10/2013 5:00:01 PM	@engadgetSitting in Malawi watching your iPhone liveblog on 2G (edge).	false	nedun	37747656869388288
0.0	0.0	9/10/2013 5:00:02 PM	Engadget's Apple Phone Linvel Line Bioghttp://t.co/LHirr5100tV	false	nedun	37747657076129382
0.0	0.0	9/10/2013 5:00:02 PM	@Maksin_Shemarov.http://t.co/3UnhgBy1UN	false	nedun	37747657220031283
0.0	0.0	9/10/2013 5:00:03 PM	Formula E details Spark-Renault's SRT_01E electric race #car http://t.co/D9qU/EDHZ	false	nedun	37747657399969382
0.0	0.0	9/10/2013 5:00:03 PM	Best liveblog in town for Apple's event starting up right now: http://t.co/M86e82c20u	false	nedun	37747657595002880
0.0	0.0	9/10/2013 5:00:05 PM	@troyellerbroek Check out Engadget's coverage this go-around: http://t.co/BEG/L2sOq2	false	nedun	37747658253926400
0.0	0.0	9/10/2013 5:00:05 PM	Apple's #Phone 55 #VOS LIVE announcement, follow any of the 3:1 http://t.co/GPW/QcGFz/	false	nedun	37747658309253939
0.0	0.0	9/10/2013 5:00:05 PM	El CEO de Intel muestra un SoC de 14 nm para laptops y promete tablets de menos de 100 d	false	nedun	37747658500971315
0.0	0.0	9/10/2013 5:00:06 PM	Following engadget, cnet, macrumors	false	nedun	37747658677933670
0.0	0.0	9/10/2013 5:00:06 PM	Thanks for the best live blog on Internet! #engadget #engadgetiphone #Apple	false	nedun	37747658844866969
0.0	0.0	9/10/2013 5:00:07 PM	RT @engadget: We're minutes away from the Phone event. Get up to speed with all the rumors to date: http://t.co/By(B7rcRdd. #engadgetPhone	false	nedun	37747659074712371
0.0	0.0	9/10/2013 5:00:07 PM	RT @mashable: Apple's Phone announcement begins shortly. Join our event pre-game here: http://t.co/AGDeu90Yz	false	nedun	37747659398932480
0.0	0.0	9/10/2013 5:00:08 PM	@engadget 6-inch PhonePad Constrictor that cols around your wrist #predictions #engadgetphone	false	nedun	37747659499596185
-37.78245	144.9968	9/10/2013 5:00:09 PM	Ølpstubbs nope. It's 3am and I'm up with Engadget and the Verge. #dedication #nerd #lphone #letsgo	false	nedun	37747660216822169
0.0	0.0	9/10/2013 5:00:10 PM	Apple's next-generation iPhone liveblogi Exciting https://t.co/WW/gP/tipe	false	nedum	37747660377045401
0.0	0.0	9/10/2013 5:00:10 PM	13 Ethical Ways to Increase Your Site's Search Traffic http://t.co/cAU7xe7xS5 via @mashable	false	nedun	37747660577533542
0.0	0.0	9/10/2013 5:00:10 PM	m, mmhttp://t.cojUm/FOlipp	false	nedun	37747660657682022
0.0	0.0	9/10/2013 5:00:11 PM	Apple's next-generation iPhone liveblogi at http://t.co/mH44US2C8g	faise	nedun	37747660767993036
0.0	0.0	9/10/2013 5:00:11 PM	http://t.co/HYOTQLikZF	false	nedun	37747660778060185
0.0	0.0	9/10/2013 5:00:11 PM	13 Ethical Ways to Increase Your Site's Search Traffic http://t.co/HHUFF7NuA3 via @mashable	false	nedun	37747660959252480
0.0	0.0	9/10/2013 5:00:11 PM	RT @MacStroke: 111111111111111111111111111111111111	false	nedum	37747661009586995
0.0	0.0	9/10/2013 5:00:12 PM	RT @iphoneros: Faitan s	false	nedum	37747661197450444
0.0	0.0	9/10/2013 5:00:12 PM	How to ruin a good thing? Instagram Plans to Introduce Ads Within the Next Year http://t.co/sMtQinh2535 via @mashable	false	nedun	37747661213008691
0.0	0.0	9/10/2013 5:00:12 PM	@/avierocasio liveblog: http://t.co/1yd0Q5ViSIC	false	nedun	37747661277562880
0.0	0.0	9/10/2013 5:00:12 PM	RT @revolucion_mobi: Puedes seguir el Evento de Apple en vivo en http://t.co/21gt/ML/IM (ingles) http://t.co/3d2VWIIycG	false	nedun	37747661491890995
0.0	0.0	9/10/2013 5:00:12 PM	@FabianCalvo ya engadget es ni fiel compa	false	nedun	37747661513322086
0.0	0.0	9/10/2013 5:00:13 PM	My Tweeted Times http://t.co/itbf2nOvnZ - top stories by mashable, MarkClayson, TheMrBlueprint	false	nedum	37747661790104780
0.0	0.0	9/10/2013 5:00:14 PM	App Anyone?http://t.co/NWWSGusGL2	false	nedun	37747662283354931
0.0	0.0	9/10/2013 5:00:15 PM	7777777 7 @hone 777777 http://t.co/itrORQM/SBI @engadgetp7777	false	nedun	37747662454064332
0.0	0.0	9/10/2013 5:00:15 PM		false	nedun	37747662639074099
0.0	0.0	9/10/2013 5:00:15 PM	http://t.cs/qUyoClyz5	false	redun	37747662647421337 🗸

Prezi

Engadget's Apple iPhone Unveil Live Blog Best liveblog in town for Apple's event starting up right now Apple's #iPhone5S #ios LIVE announcement, follow any of the 3:1

Category based analysis



Heatmap of sources used to tweet Mashable and Engadget tweets



())Prezi



User network (retweet relationship)

Single mode directed weighted network; size: degree; Color: degree; Layout: openord



କ ଭ ଭ











	INNOVATION DYNAMICS Use Brander Versite Persone Landscord	To Ween eventure transmitter by Print reg Sp-04	
	Customer Engagement	المحمد بالترا	
	mHealth Data & Analytics	Social Media	
	Big Data		
	Mobile Platform	Security	
Prezi	Misc 00000		







Avg path length 2.631 (lowest)





Tableau Experience Samples

Youtube demo of Tableau JavaScript API

https://www.youtube.com/watch?v=_91cpew085Q



⊯H ∎() 0.51

Sorting in Tableau

By Arunkumar Navaneethan

Basic sort

Right click on the discrete¹ field in the rows or columns shelf



 $^{1}\ \mathrm{Sort}\ \mathrm{can}\ \mathrm{be}\ \mathrm{performed}\ \mathrm{only}\ \mathrm{on}\ \mathrm{discrete}\ \mathrm{fields}\ \mathrm{or}\ \mathrm{measures}\ \mathrm{which}\ \mathrm{is}\ \mathrm{converted}\ \mathrm{to}\ \mathrm{discrete}$

	Columns					
	Rows	Mar	ket	F	Product Type	F
	Market Central	F C F	Filter Show (Sort	Quick Filter		
Abc 123 Text	West		Clear S Forma Show I Include Edit Al	ort t Header e in Tooltip iases		
p t)	East		Dimen Attribu Measu	sion Ite		
	Last		Subtot	als		
		Total	Remov	/e \$59.217]	
	South	Coffee		\$11,702		
		Espres	so	\$15,005		
		Herbal	Tea	\$5,771		
		Total		\$32,478		
	Grand Total			\$259,543		

- 1. Data source (if the field is pulled from a SQL data, this order is based on sort orders applied to this relation)
- 2. Alphabetically
- 3. Based on another Field across different aggregation level

Sort [Market]		×
Sort order		
Descending		
Ascending		
Sort by		
Data source order		
O Alphabetic		
Field		Aggregation:
Profit		▼ Sum ▼
Manual		
Central		Up
West East		Down
South		
Clear	OK Cance	Apply

Non-Nested Sort

The nature of the sort function in Tableau is that sorting is done at the content level.

When a sort is done based on the SUM(profits) descending, you can notice that CENTRAL market is at the top. This because the aggregate SUM of CENTRAL is the maximum. Similarly COFFEE product type as a MAX of SUM(PROFIT) arranged descending. Hence it is at the top across all the Markets. The results are not sorted based on the Markets but to the entire context.

This is the default behavior within Tableau.



Nested Sort

If the results is expected to be sorted based on the markets, we have to create a nested sort using a work around.

In the dimensions pane, press the CTRL key, and then select the dimensions. Right-click on e of the selected fields and click on Combine Fields option. The newly combined field, Market & Product Type is created. Drag this field to the Rows panel and set the sort options. We will get a nested sort output.



Selected Sort

With the previous sort you can add more measures but the results will still be still sorted by PROFIT

	Market	Product Type & Mar	Profit	Sales
	Central	Herbal Tea, Central	\$24,757	\$67,888
		Espresso, Central	\$23,501	\$59,701
Arunkumar Navaneethan © 2014. All Rights		Coffee, Central	\$23,264	\$69,080
Reserved.		Tea, Central	\$22,330	\$68,376
	East	Coffee, East	\$30,992	\$56,641
		Tea, East	\$15,558	\$32,177
		Herbal Tea, East	\$6,423	\$41,361
		Espresso, East	\$6,244	\$48,397
	South	Espresso, South	\$15,005	\$44,990
		Coffee, South	\$11,702	\$33,259
		Herbal Tea, South	\$5,771	\$25,677
	West	Herbal Tea, West	\$26,303	\$72,288
		Espresso, West	\$23,870	\$69,908
		Tea, West	\$15,098	\$72,220
		Coffee, West	\$8,725	\$57,848

Tableau combination chart

By Arunkumar Navaneethan

My Use-case

Step 1:

To display a chart with Number of stores across years, differentiating each year by category of the store. In addition, have an overlay that would show the total for the given year based on user selection. This should not modify the underlying chart.

 Durber of Opening Over Train
 Durber of Opening Over Train

 Output
 Output
 Output

 Output
 Output
 Output



1. Create a Stacked chart



Rows: Aggregate of number of stores

Columns: Years

Filter: The category of interest (Apparel specialty, supermarket, value market, department store) *Color:* Category

2. Overlaying chart

2.1. Create a parameter using Year field and select years from 2000 to 2007

Pages	Columns	YEAR(Date Started	()		
	Rows	AGG(Number of Stor	es)		
Filters			Year of Date Started		
Category 🚊	4000				
Marks	2 Create D	arameter			x
🖬 Bar 👻		arameter			
	Name:	Parameter 1			Comment >>
Color Size Label	30 Proper	ties			
	Dat	a type: Float	•		
Detail Tooltip	24				
	e Cur	rent value:	-		
😓 Category 🛓	Disp	lay format: Automatic	-		
	20 20				
Category	Allo	wable values: 🔘 All 🔘 I	List 🔘 Range		
Apparel Specialty	List of	values			
Supermarket	15	due.	Dicelay Ac		Add from Parameter
Value Market	va D	aue	Display As		
Department Store	2,	000	2,000		Add from Field 🕨
	10 2,	001	2,001		
	2,	002	2,002		Paste from Clipboard
	2,	004	2,003		
	4 2	005	2,005	E	
	2.	006	2,005		
	2,	007	2,007		
	A	bb		+ [Clear All
			i		
Sheet 1 🗉 🖽					OK Cancel
SUM of AGG(Number of Store	s): 9,38				

2.2. Create a calculated field Highlight Year (IF [Year]=[Parameter 1] THEN 1 END)



2.3. Display the parameter in the sheet and make it a slider, move the Highlight Year to rows shelf



2.4. Make the Highlight Year row as a dual Axis



2.5. Right click on the right hand side axis and synchronize the Axis



2.6.In the Sum (Highlight Year) marks; remove category from color



2.7. Change the color to white and transparency to 40%



2.8. Change the size to lesser size using the slider



2.9. Add the number of stores to the label



Resources:

Initial tableau workbook Combination Chart (Right-click and "Save link as...")

Original creation:

http://public.tableausoftware.com/shared/2QCDD7BPR?:display_count=no

Tableau exploration with Portfolio Management System

By Arunkumar Navaneethan

arunethan.com

Gathered and analyzed requirements for a portfolio management system following an iterative development approach. Created an OLTP in MySQL and star schema, OLAP for the data warehouse, followed the standard design convention. The OLAP was created accommodating slowly changing dimensions recording historical data of employees and time stamped status of project information. Create facts and dimension table in MYSQL using complex SQL join statements on the OLTP tables. The dimensions and facts are created with indexes on foreign keys



to enhance query performances. Created a view of the MYSQL data and the data extracts were loaded to the tableau server for query performance.

The tableau dashboard and workbooks are hosted in the tableau public server. Created calculated fields to find the margin based on the revenue gained and expense incurred. Structured the data as hierarchies with portfolio as the highest level followed by application and project details. Created groups for the application data to do category based analysis of the data. Created sets to display project data based on conditions like the top 3 projects for the given portfolio.

Created parameters to set working hours goals by checking if the working hours is more than a certain definite average. Created animated visualization to understand the growth of margin with respect to revenue and expense over time using the pages card. Created symbol maps to understand the geo-spatial distribution of projects with custom symbols made to represent the application category. Manually organized the dimensions and measures like moving the surrogate keys, date, descriptions, flags and categories as dimensions.

Customized the tooltips and text to end user understandable format, the numerical measure was changed to its appropriate formats like representing the revenue, expense and margins as currency values. Created dashboards in Tableau to analyze the revenue, expense over various attributes of employee, project, time and location dimension. The dashboard is built with floating objects and capitalizing the various dashboard features like URL, images, web integration's. The portfolio data is represented as tree map based on the margin, the detailed application data is represented as column graph with highlight table distributing the margin across employee roles. The detailed cross tab data of each employee is provided to understand the underlying data across attributes. Created the appropriate filter for project level and employee level data. The project level filter was single select and employee level filter was a wild card search. Experimenting with social media analytics to provide application related data from twitter REST API and streaming API.

The next step is to create a portal style dashboard using the existing dashboard and workbooks. Exploring tableau's javascript API features to implement this.

Tableau filter excluding All

By Arunkumar Navaneethan

The default filter will start from "All"





Change the custom filter settings



Filter excluding "All"



My Use-Case:

When a data-set consist of data across multiple years or multiple categorical separation. The visualization is expected to display diversity across years or categories and not the aggregate of "All" the data. Then we can follow this procedure.

Tableau fixed axis length

By Arunkumar Navaneethan

My use case:

The underlying data is percentage turnover for different regions. The maximum value was around 21.

Tableau default axis length settings is automatic. It created a chart with a range of 0 - 23, but the percentage representation of the data is lost.





Before we start

The axis setting can only be modified when the column or row value is continuous and not discrete.

Steps

1. Right click on the axis and select "Edit Axis"



2. Change the "Fixed" range to 0 - 100

Pages	111 C	olumns	Year	
	I Re	ows	Edit Avis [Ava. Turnover district]	
Filters	1(00.0	General Tick Marks	
Marks	9	90.0-	Automatic	
Automatic -	ŧ	30.0-	Uniform axis range for all rows or columns Independent axis ranges for each row or column	
Color Size Label	7	70.0-	© Fixed Start: End:	
Detail Tooltip	ي و	50.0-	0	
Abc AVG(Turnover distri	nover distr	50.0-	0.0 100.0	
	Avg. Tur	40.0-	Scale Titles Reversed Title: Avg. Turnover district	
	:	30.0-	Logarithmic Subtitle:	
	:	20.0-1	17.8	21.9
	1	10.0-	Clear OK Cancel Apply	
		0.0 20	04 2005 2006 2007 2008 2009 2010 2011 2012	2013



Dimensions © Zees Code Totas Networks Network Totas Network Totas Network Type © Robe Network Type © Robe Network Totas Network Network Totas Network Network Totas Network Totas Network Totas Network

Menures # Budger COSS # Budger Tote # Steps # Ste Coller San

🔇 arunethan.com

Tableau maps background colors

By Arunkumar Navaneethan

My use case:

Tableau maps visualization for data from a Single state

I have data for teacher turnover in schools within Austin, TX (2004 – 2013). I used the user defined latitude and Longitude values, plotting school name as details.



The map can be customized using the "Map Options" which is under "Map" tool bar.



Background color options: gray, dark and normal

The default map background color is "Gray" it is typically suitable for symbol maps and filled maps. It is best to use when map layers like "zip code labels", "place names" are displayed



The dark background is more suitable for the symbol maps.



The normal background is rarely used but it does highlight the geographical features like rivers and landscape natural.

Map Options ×	Pages	Columns AVG(Longitude)
Background		
Style: Normal Washout: 27%	Filters	
Repeat Background		
Map Layers	Marks	
Country/Region Borders Country/Region Names	Automatic	
 State/Province Borders State/Province Names 	Color Size Abc Label	
US County Borders	Detail Tooltip	
Zip Code Boundaries Zip Code Labels	SUM(Turnover rate)	
Area Code Boundaries Area Code Labels	School name	
US Metro Boundaries (CBSA) US Metro Labels (CBSA)	SUM(Turnover rate)	
Place Names	0.0 382.1	
Data Layer		
Layer: No Data Layer 🔹		

Data source:

http://www.mystatesman.com/interactive/aisd-teacher-turnover/

Tableau Mobility – Approach

By Arunkumar Navaneethan

In Tableau, you do not have to do anything special to make a dashboard mobile. Simply publish to Tableau Server like you always have, and Tableau will detect if you're using the Mobile app. You even get the native touch experience if you go to Tableau Server from your mobile browser, without the app at all.

Ease of use is the single most important aspect of mobile business intelligence. When you're on the go, you need to be able to get to the data you need with a few taps.





Filters: Tableau's controls such as filters, parameters, sliders, scrolling, and zoom & pan, are specially built to interact with your fingers. For example, tapping a filter pops a large, touch-optimized quick filter. And for long filters, there is scrolling inside the filter.

Views: Views themselves are touch-optimized with dynamic scrolling. Simply swipe to scroll through a long customers list, for example. Or pinch & zoom in a map.

The two major approaches for Mobility in Tableau



Security

Mobile business intelligence must be secure. With the Tableau mobile solution, security and metadata continues to be managed by Tableau Server. This means you can enforce your existing security protocols and integrate with ActiveDirectory via Tableau Server.

And if an employee loses their iPad or Android tablet, simply disable their Tableau Server account and give them a new one. No data other than descriptive data about a workbook (like the publisher, data modified and name) are stored on the device, so you can keep your data secure even while it's mobile.

Limitation

Write to a smaller form factor ["author-once, distribute everywhere"]

Smaller screen? You've got to use your space better. Place the most important view at the top of your dashboard. Be stingy with legends and filters, and for that matter views– limit your dashboard to 2-4 important views. 10 views in a dashboard doesn't work on the go!

Reference

 "5 Best Practices for Mobile Business Intelligence." *Tableau Software* 1 (2011). http://www.tableausoftware.com/sites/default/files/whitepapers/whitepaper_5-best-practices-mobile-bi.pdf (accessed April 8, 2014).

Tableau parameter as a filter

By Arunkumar Navaneethan

My Use-case

They are three different workbooks with data across years. Must link the workbooks together based on user selection of a year using parameters.

Steps



1. Create the Year Parameter

	u <mark>lx</mark> * ¶∥ * i≎ * ∎⊅ *∎			
Data ÷	Pages	Columns		
🍺 Retail Profile Data.xls [Data] C				
Dimensions IIII P 👻		📃 Ro Edit Parameter [Para	ameter 1]	×
Abc Category Abc Category (copy 2) Abc Category (copy) Date Started Latitude Latitude Longitude MSA Region State State State ID Abc Store Name	Filters Category Year Marks Marks Automatic Color Size Label	Name: Parameter Properties Data type: Current value: Display format: Allowable values	Float 2,001 Automatic C All List Range	Comment >>
=# Year =# Year1 Abc Measure Names	Detail Tooltip	List of values	Display As	Add from Parameter
Measures # Distance =# Highlight Year =# Number of Stores # Sales =# Sales	SUM(Sales per Sq f Category Abc Store Name	2,000 2,001 2,002 2,003 2,004 2,005 2,006	2,001 2,001 2,002 2,003 2,004 2,005 2,006	Add from Field Paste from Clipboard
 # Sales per Sq Teet # Sq# Feet # Latitude (generated) # Longitude (generated) # Number of Records # Measure Values 	Apparel Specialty Department Store Supermarket Value Market	Add	2,007	Clear All OK Cancel
Darameters		Super Markets Inc	Save Mart	
# Parameter 1	PieChart Combination	Chart Treemap 🖽	Dashboard 1 🔳 🖽	

2. Drag the year dimension to filter box and open the condition tab

)ata ≑ ⊐ pitipici pitipitic	Pages	Columns	
Retail Profile Data.xis [Data] C		I David	Filter [Year]
Dimensions III P ▼ Abc Category Abc Category (copy 2) Abc Category (copy) Date Started ⊕ Latitude ⊕ Longitude # MSA # Region ⊕ State # Store ID Abc Store Name =# Year # Year	Filters Catogrey Year Marks Automatic Color Size Label	Rows	General Wildcard Condition Top None By Field: Sales per Sq feet C O Range of Values Min: Load Max:
<pre>*# Year1 Abc Measure Names # Distance *# Highlight Year *# Number of Stores # Sales *# Sales # Sales # Sales per Sq feet # Sq# Feet @ Latitude (generated) @ Longitude (generated) # Number of Records # Measure Values</pre>	Detail Tooltip SUM(Sales per Sq f Category Abs Store Name Category Apparel Specialty Department Store Supermarket Value Market Value Market	Wingda Supern Inc. Ali, Inc. #N/A Town 8 Super I	By Formula: I Reset OK Cancel Apply
arameters			
# Parameter 1	PieChart Combination	Chart Treema	p 🗉 Dashboard 1 🔳 🖽

3. Use the formula ([Year] = [Parameter 1])

Data ÷	Pages	Columns	
🔂 Retail Profile Data.xls [Data] C			x
Dimensions 💠 🖓 👻			
Abc Category	Filters	Formula:	
Abc Category (copy 2)	Category	[Year] = [Parameter 1]	
Abc Category (copy)	Category		
Date Started	Year		n
latitude			
Longitude	Marke		
# MSA	IVIdTKS		
# Region	Automatic	The calculation is valid.	ear
State H C: ID			
# Store ID	S S S S S S S S S S S S S S S S S S S	Fields: Parameters: Functions: Help	
ADC Store Name	Color Size		
=# Veerl		Enter Text to Search Enter Text to Search Enter Text to Search	le
Abs Measure Names	Detail Tooltip	=# Number of Perce 4 # Datameter 1 APS	
noe measure mariles		=# Number of Stores Humileter ACOS Current value: 2,001	
	(SUM(Sales p	# Region ASCII	
Measures	le Category	# Sales ASIN	
# Distance	Abc Store Name	=# Sales per Sq feet ATAN	
=# Highlight Year	123	# Sq# Feet ATAN2	
=# Number of Stores		AVG	
# Sales	Category	# Store ID E CASE	
=# Sales per Sq feet	Apparel Special	Abc Store Name CHAR	
# Sq# Feet	Department Sto	CONTAINS	
Latitude (generated)	Supermarket		
=# Number of Pecerds	Value Market		
# Measure Values		OK Car	ncel
# measure values			
		Super Markets inc. Save Mark	
Parameters			
# Parameter 1			
	PieChart Co	ombinationChart Treemap Tashboard 1	

4. Follow the step 2,3 for the other workbooks



Resources

Initial tableau dashboard Parameter filter (Right-click and "Save link as...")

Original Creation

http://public.tableausoftware.com/shared/2QCDD7BPR?:display_count=no

Tableau Pie chart on a symbol map

By Arunkumar Navaneethan

Columns

Rows

My Use-case

Display pie chart over every state, sized by the total number of stores. Size the slices of the pie chart by store categories. This will help us understand which states have more number of store opening and detailed by category.

> Longitude (generated) Latitude (generated)

Steps

Pages

Filters

Marks Automatic 0

0

Abc 123

1. Add the state to the chart by double clicking on state dimension or add the state dimension to details and move the longitude measure to columns and latitude measure to rows.

2. Move the Number of store measure to the size mark







Move the Category dimension to the color mark



3. Change the Mark type from automatic to Pie and use the size mark to resize the pie size. The map can be zoomed into a closer view if required



Resources

Initial tableau workbook Pie Chart (Right-click and "Save link as...")

Original Creation

http://public.tableausoftware.com/shared/2QCDD7BPR?:display_count=no

Tableau URL action

By Arunkumar Navaneethan

My Use-case

We can carry out various external interaction using the URL action.

We can use this to redirect users to google search results, intranet sites, REST methods or predefined URLs

1. Google search

https://www.google.com/#q=<Store Name> store

2. Predefined URL's

<Store URL>

3. REST methods

https://www.arunethan.com/updateStoreInfo.jsp?StoreName=<Store Name>

4. Email with subject

mailto:info@arunethan.com?subject=Request for more information about <Store Name>

For a dashboard, open the Dashboard menu -> then Actions.. -> Add Action, button -> URL..

The URL can be created with values from tableau <dimension> or <measure> using the triangle button

Examples

ctions [As-is]				×
Connect sheets to external and Highlight actions.	web resources using URL action	ons, or to other sheets	in the same w	orkbook using Filter actions
Name	Run On	Source		Fields
Add Action >			<u>E</u> dit	<u>R</u> emove
Show actions for all shee	ets in this workbook		ОК	Cancel



Add URL Action	X
Name: Hyperlink1	Þ
Source Sheets	
🖬 As-is 👻	Run action on:
CombinationChart	₿ <u>H</u> over
PieChart Treemap	🚯 <u>S</u> elect
	🕄 Men <u>u</u>
URL	
	
Test Link	U
URL Options	
URL Encode Data Values	Item Delimiter;
Allow <u>M</u> ultiple Values	Delimiter <u>E</u> scape:
	OK Cancel

Add URL Action	×
Name: Hyperlink1	
Source Sheets	
📅 As-is 👻	Run action on:
CombinationChart	k} <u>H</u> over
PieChart	🎊 <u>S</u> elect
(e) neemp	R Men <u>u</u>
URL	
https://www.google.com/#q= <store name=""> store</store>	
Test Link https://www.google.com/#q=<	Store Name> store
URL Options	
URL Encode Data Values	Item <u>D</u> elimiter:
Allow Multiple Values	Delimiter Escape:
	OK Cancel

Add URL Ad	tion	X
<u>N</u> ame:	Hyperlink1	
Source She	ets	
📅 As-	is 🔻	Run action on:
	CombinationChart	k <u>H</u> over
Г Р Г Г Т	rieChart Treemap	🎊 <u>S</u> elect
		🖏 Men <u>u</u>
URL		
<store< th=""><th>URL></th><th>►</th></store<>	URL>	►
Tes	t Link http://arunethan.com/	
URL Option	IS	
	Encode Data Values	Item <u>D</u> elimiter:
Allow	Multiple Values	Delimiter <u>E</u> scape;
		OK Cancel

Using Range/Relative filter for a discrete Date field in Tableau

By Arunkumar Navaneethan

Different Date filters

Relative Date - This filter helps to analysis the data with respect to year, quarter, month, weeks, days relative to the current date or an user created anchor date.

Range of Dates - If the data has to be analyzed based on certain user defined interval

Start Date and End Date - End date is helpful for As-of Date analysis

Browse Period – Data can be analyzed on various date granularity

My Use-case

To have a range filter in my dashboard to filter crosstab data, which consist of sales values aggregated by months.

Steps

Create a dashboard with Date field in it

←→■■1■150-000	a	+ + - Abc		ormal 👻	+ 2.						 db. Shaw Ma
Data a	WX ull ~ th		94 000s 14								 IIII Show Me
Sample - Coffee Chain (Access)	Pages	Columns		ate)							
Dimensions III Q *		Rows	Product	E.							
Area Code Date Date Abc Market	Filters	Product Green Tea Mint	01-Jan-12 (\$44) \$690	01-Apr-12 (\$3) \$815	01-Jul-12 (\$78) \$488	01-Oct-12 \$35 \$518	01-Jan-13 (\$73) \$1,038	01-Apr-13 (\$5) \$1,154	01-Jul-13 (\$117) \$683	01-Oct-13	
Abc Product Abc Product Line Abc Product Type	Marks Abc Automatic •	Amaretto Caffe Latte Decaf Irish Cream	\$593 \$1,153 \$1,359	\$446 \$1,231 \$1,534	\$400 \$1,159 \$1,528	\$544 \$1,093 \$1,287	\$896 \$1,740 \$2,047	\$631 \$1,741 \$2,167	\$572 \$1,635 \$2,153	\$808 \$1,623 \$1,914	
de State de Type Abc Measure Names	Color Size Text	Regular Espresso Caffe Mocha Earl Grey Darjeeling	\$908 \$1,697 \$2,283 \$2,838	\$986 \$1,734 \$2,522 \$2,998	\$814 \$1,883 \$2,642 \$3,201	\$1,384 \$1,887 \$2,404 \$2,807	\$1,369 \$2,556 \$3,444 \$4,281	\$1,396 \$2,454 \$3,561 \$4,237	\$1,154 \$2,667 \$3,737 \$4,522	\$2,054 \$2,800 \$3,571 \$4,169	
	Abc SUM(Profit)	Chamomile Decaf Espresso Lemon	\$2,695 \$2,963 \$2,544	\$2,723 \$3,079 \$3,231	\$2,855 \$3,149 \$3,355	\$2,820 \$2,834 \$3,065	\$4,066 \$4,468 \$3,836	\$3,850 \$4,356 \$4,568	\$4,036 \$4,448 \$4,721	\$4,186 \$4,205 \$4,549	
Aessures # Budget COGS # Budget Profit # Budget Profit # Budget Sales # COGS # Lorentory # Margin # Marketing # Profit # Sales # Total Expenses @ Lotitude (generated) @ Longitude (generated)		Colombian	35,096	35,692	36,583	33,206	57,690	58,319	59,292	37,726	
*# Number of Records # Measure Values	Sheet 1 🖽 Dashboard	11 E E									

104 marks 13 rows by 8 columns SUM(Profit): \$259 543

Right-click on the Date field and select "show quick filter"



	\$	Dager	-		(m.a.)									
amp	le - Coffee Chain (Access)	rayes			Columns	QUARTER(D	ate)							
nsior	ns 💷 🔎 🔻				Rows	Product	E.							
rea	Code	Filters			Product	01- Jan-12	01-Apr-12	01- Jul-12	01-Oct-12	01- Jan-13	01-Apr-13	01. Jul. 13	01-Oct-13	
	Add to Sheet				Green Tea	(\$44)	(\$3)	(\$78)	\$35	(\$73)	(\$5)	(\$117)	\$54	
	Show Quick Filter				Mint	\$690	\$815	\$488	\$518	\$1,038	\$1,154	\$683	\$768	
					Amaretto	\$593	\$446	\$400	\$544	\$896	\$631	\$572	\$808	
	Сору				Caffe Latte	\$1,153	\$1,231	\$1,159	\$1,093	\$1,740	\$1,741	\$1,635	\$1,623	
	Paste		utomatic	•	Decaf Irish Cream	\$1,359	\$1,534	\$1,528	\$1,287	\$2,047	\$2,167	\$2,153	\$1,914	
	Duplicate			Abr	Regular Espresso	\$908	\$986	\$814	\$1,384	\$1,369	\$1,396	\$1,154	\$2,054	
	Rename		Size	123 Text	Caffe Mocha	\$1,697	\$1,734	\$1,883	\$1,887	\$2,556	\$2,454	\$2,667	\$2,800	
	Hide		JIZE		Earl Grey	\$2,283	\$2,522	\$2,642	\$2,404	\$3,444	\$3,561	\$3,737	\$3,571	
			Tooltin		Darjeeling	\$2,838	\$2,998	\$3,201	\$2,807	\$4,281	\$4,237	\$4,522	\$4,169	
	Create Calculated Field				Chamomile	\$2,695	\$2,723	\$2,855	\$2,820	\$4,066	\$3,850	\$4,036	\$4,186	
	Create Group		M(Profit)		Decaf Espresso	\$2,963	\$3,079	\$3,149	\$2,834	\$4,468	\$4,356	\$4,448	\$4,205	
٥	Create Set				Lemon	\$2,544	\$3,231	\$3,355	\$3,065	\$3,836	\$4,568	\$4,721	\$4,549	
	Create Parameter				Colombian	\$5,096	\$5,892	\$6,583	\$5,206	\$7,690	\$8,319	\$9,292	\$7,726	
	Create Hierarchy													
	Create Custom Date													
	Convert to Continuous													
	Change Data Type													
	Default Properties	•												
	Replace References													
	Describe													
les														
tal	Expenses													
titi	ude (generated)													
na	itude (generated)													

In the Filter option of the Date field, select "Continuous" instead of "Discrete"

<u>F</u> ile <u>D</u> ata <u>W</u> orksheet Dash <u>b</u> oar	d <u>A</u> n	nalysis <u>M</u> ap F <u>o</u> rmat	<u>S</u> erver Wi <u>n</u> do	w <u>H</u>	<u>l</u> elp									
← → 🖪 🛢 🖬 • 🔐 🛛	ц <mark>х</mark> -	₽ .• @ - ₽ 4	- + F @ - A	bc	💭 💷 - No	ormal 👻	₩ ∠•						ilii She	ow Me
Data ÷	Page	es	Columns	E	QUARTER(Da	ite)								
Sample - Coffee Chain (Access)														
Dimensions IIII P +			Rows	P	roduct	-								
Area Code Data	Filte	rs	Product		01-Jan-12	01-Apr-12	01-Jul-12	01-Oct-12	01-Jan-13	01-Apr-13	01-Jul-13	01-Oct-13	YEAR(Date)	
Abc Market	YE	AR(Date)	Green Tea		(\$44)	(\$3)	(\$78)	\$35	(\$73)	(\$5)	(\$117)	\$54	🗹 (All)	
Abc Market Size		Filter			\$690	\$815	\$488	\$518	\$1,038	\$1,154	\$683	\$768	2012	
Abc Product	1	Show Quick Filter			\$593	\$446	\$400	\$544	\$896	\$631	\$572	\$808	2013	
Abc Product Line		Clear Filter			\$1,153	\$1,231	\$1,159	\$1,093	\$1,740	\$1,741	\$1,635	\$1,623		
Abc Product Type		Add to Context	e	eam	\$1,359	\$1,534	\$1,528	\$1,287	\$2,047	\$2,167	\$2,153	\$1,914		
State		Apply to Worksheets	•	\$\$0	\$908	\$986	\$814	\$1,384	\$1,369	\$1,396	\$1,154	\$2,054		
Abc Type		hippij to monunced			\$1,697	\$1,734	\$1,883	\$1,887	\$2,556	\$2,454	\$2,667	\$2,800		
Nuc Meusure Numes	۲	Create Set			\$2,283	\$2,522	\$2,642	\$2,404	\$3,444	\$3,561	\$3,737	\$3,571		
		Year	2011		\$2,838	\$2,998	\$3,201	\$2,807	\$4,281	\$4,237	\$4,522	\$4,169		
		Quarter	02		\$2,695	\$2,723	\$2,855	\$2,820	\$4,066	\$3,850	\$4,036	\$4,186		
		Month	May	10	\$2,963	\$3,079	\$3,149	\$2,834	\$4,468	\$4,356	\$4,448	\$4,205		
		Davi	iviay		\$2,544	\$3,231	\$3,355	\$3,065	\$3,836	\$4,568	\$4,721	\$4,549		
Measures	-	Day	° .		\$5,096	\$5,892	\$6,583	\$5,206	\$7,690	\$8,319	\$9,292	\$7,726		
# Budget COGS		More												
# Budget Margin		Year	2011											
# Budget Profit		Quarter	Q2 2011											
# Budget Sales		Month	May 2011											
# Loventon/		Week Number	Week 5, 2011											
# Margin		Day	May 8, 2011											
# Marketing		More												
# Profit		Wore												
# Sales		Exact Date												
# Total Expenses # Latitude (aenerated)		Measure	+											
Longitude (generated)	•	Discrete												
=# Number of Records # Measure Values		Continuous												
104 marks 13 rows by 8 columns		Remove												

Result

E → ■ 🔒 🜆 - 🚇				- Normal		-							IL Show	Mo
ata é													JIII SHOW	IVIE
ata - Coffee Chain (Access)	Pages	Columns	DAY(Date)											
imensions III P 🔻		Rows	Product	E.										
Area Code	Filters												YEAR(Date)	
Date Naskat	YEAR(Date)	Product	January 1, 2012	February 1, 2012	March 1, 2012	April 1, 2012	May 1, 2012	June 1, 2012	July 1, 2012	August 1, 2012	September 1, 2012	October 1, 2012	2012	201
n Market Size		Green Tea	(\$51)	\$3	\$4	(\$1)	(\$7)	\$5	\$15	(\$15)	(\$78)	\$17		
c Product		Mint	\$220	\$231	\$239	\$281	\$265	\$269	\$294	\$100	\$94	\$168		
Product Line	Marks	Amaretto	\$218	\$200	\$175	\$152	\$154	\$140	\$99	S114	\$187	\$179		
Product Type	Abc Automatic 🗸	Caffe Latte	\$365	\$387	\$401	\$406	\$421	\$404	\$436	\$370	\$353	\$351		
State		Decaf Irish C.,	\$427	\$463	\$469	\$497	\$509	\$528	\$590	\$518	\$420	\$432		
Туре		Regular Espr	\$267	\$357	\$284	\$320	\$268	\$398	\$263	\$252	\$299	\$515		
Measure Names	Color Size lext	Caffe Mocha	\$554	\$582	\$561	\$551	\$583	\$600	\$588	\$671	\$624	\$667		
		Earl Grey	\$744	\$768	\$771	\$799	\$844	\$879	\$892	\$888	\$862	\$801		
	Detail looitip	Darjeeling	\$935	\$949	\$954	\$944	\$995	\$1,059	\$1,072	\$1,123	\$1,006	\$993		
	Abr. Country, Co.	Chamomile	\$897	\$902	\$896	\$893	\$895	\$935	\$942	\$1,000	\$913	\$975		
	123 SUM(Profit)	Decaf Espres	\$955	\$997	\$1.011	\$1.006	\$1.030	\$1.043	\$1.078	\$1.077	\$994	\$972		
		Lemon	\$800	\$864	\$880	\$1.029	\$1.082	\$1,120	\$1,337	\$1,155	\$863	\$898		
asures		Colombian	\$1,710	\$1,666	\$1,720	\$1,793	\$1,908	\$2,191	\$2,299	\$2.313	\$1,971	\$1,706		
Budget COGS														
Budget Margin												,		
Budget Sales														
COGS														
Inventory														
Margin														
Marketing														
Profit														
Sales														
Total Expenses														
Latitude (generated)														
Longitude (generated)														
Number of Percenter														
Measure Values	Sheet 1 🕀 Dashboar	d1 🗖 🖽												
# measure values														

OR

Drag and drop the Date field to the filters mark and you will get all the option related to Date filter

Data Pages DAY(Date) Pages DAY(Date)	
Dimensions III D + III Constant Product =	_
Filters January 1, February 1, March 1, Ac March 1, Control 1, 2012 August 1, September 1, 2012 October 1, November 1, 2012 October 1, 2012<	nber 1, , 2012 (\$23)
Abc Product Marks Mint S22 Fifter Field [Date] S269 S294 S100 S94 S168 S168 Abc Product Marks Amounta S24 S100 S94 S168 S168	\$182
Add Product Line Amaretto S21 How do you want to filter on [Date]? S140 S399 S114 S167 S179 S192 Abc Product Type Abc Automatic 	\$383
State Decaf Irish C S42 Relative date S528 S590 S518 S420 S432 S402	\$453
Acc Measure Names Color Size Text Regular Expr. 526 4 Years 5398 5263 5252 5299 5515 5486	\$383
Carte Mocha 555 # Quarters 5879 5892 5888 5862 5801 5756	\$594
Detail iootrip mons state Darjeeling \$93 # Days \$1,059 \$1,072 \$1,123 \$1,006 \$993 \$887	\$927
M25 SUM(Profit) Chamomile \$89 # Week numbers \$935 \$942 \$1,000 \$913 \$975 \$930	\$915
Decaf Espres. S99 # veexa3ys 51,043 \$1,073 \$1,077 \$994 \$372 \$320 Emon san # Month / Year \$1120 \$137 \$1155 \$883 \$988 \$379	\$942 \$1.188
Measures Colombian # Month / Day / Year Colombian 51/11 # Month / Day / Year 51/12 51/1	\$1,847
<pre># Budget Margin # Budget Margin # Budget Profit # Budget Sales # CoOS # Inventory # Margin # Marketing # Profit # Sales # Total Expenses @ Lotitude (generated) @ Longitude (generated) # Number of Records # Number of Records</pre>	Þ
	N NI +

Other Experience Samples

🔇 arunethan.com

Big Data AchieveMint

By Arunkumar Navaneethan

I had an opportunity to talk to the CEO of AchieveMint, Mikki Nasch about big data and data analytics approaches. The discussions were inclined more towards the healthcare industry. It was definitely a great learning experience.

Before I start, some background about Achievemint. It is company which provides reward points for tracking personal health activities which is already tracked in application. These applications various from wearable fitness monitors like fitbit to social media postings like in Tweeter, foursquare etc.

The company uses this information to understand various patterns and trends with relation to the healthcare industry and the user in specific.

In this blog I will try to highlight the various learning's I had through this conversation.

1. **Business problem definition** must always be the first step. It would be effective to do all the data processing after we know what we are looking for from the data we have.

2. Usefulness of data, when data is analyzed its fine to have

data which doesn't give any signals even that is an essential learning

3. Interesting data may not always mean returns, at times a lot of investment is done on the tools to analysis the data which might provide interesting signals but the returns provided by the signals to the business might be less 4. Raw data is valuable, it is important to store and manage all the raw data. The raw data is generally processed and the processed data is stored separately and utilized. They are possibility that certain parts of the raw data were not important during the previous processing, which might be useful for us in the future. Hence, it is essential to store this raw data.

5. **Unsupervised learning model**, unsupervised learning model will not work effectively. It is better to go with the heuristic approach where the domain knowledge drives the analysis to the right and effective direction.

6. **R&D and real-world coexistence**, it is important that R&D and the real world coexist. The results from R&D must be verified time to time with the real world happenings(social media data)

This conversation definitely helped me to understand the real world difficulties of implementing of big data and the challenges an organization would face to gain returns.





Big Data in fraud detection and online marketing – Phoenix Meetup

By Arunkumar Navaneethan

Presenter: Raz Yalov, CTO, 41st parameter

This was one of the best, simple and clear presentation on Big data I have attended.

The presenter distinguishes big data clearly as a challenge of data of size greater than petabytes and data, which we cannot eyeball. We need complex infrastructures only when the regular server cannot handle the processing complexity.



It was interesting to know the difficulties most of the big data analytics companies have with data privacy concerns. The company goes to extend of giving special offers to customers who can provide access to their logs or even realtime sample data to improve the predictive abilities of the analytics system.

Once the customers accepted to sharing the data, the company receives the data using API's, which can be stored into two different storage sources. In-house Hadoop clustered system or Amazons S3 system. The recommended compression formats was .lzo + .lzo.index.

The presenter verdict on various big data technology. The company uses technologies like Pig, Hadoop map reduce, Presto, excel and R programming. Pig and map reduce is used mostly to create the predictive models and excel has been the best resource for eyeballing purposes. The team consist of data scientist who loves R programming, using it as much as possible even if it is not built for that purpose. There was definitely a big dislike towards Hive for its performance issues and connectivity management abilities.



The biggest challenging with Big Data is getting the data itself from the customers and once collected, the next big challenge is to extracting value from it. The sampling processes is one of the major challenges since the level of garbage data present is high.

Looking forward to see how visualization can be injected to such a complex challenging environment.

Big learning from the INSITE big data symposium 2013

By Arunkumar Navaneethan

Conducted by INSITE, Business Intelligence & Analytics Center, University of Arizona

October 10, 2013

Big Data symposium an event organized in University of Arizona, MIS Department, and INSITE center. It had a great list of speakers from diverse backgrounds representing different industrial backgrounds. The symposium covered a huge group of big data use cases from health care to sports.



Started with a great speech from Brian Gentile, Chairman and CEO, jaspersoft, he gave a simple and a clear overview of how big data would influence the industry and discussed the various topics around big data. His speech was about the "New factors of production and rise of big data". Explained the abundance of data and the value big data brings in by analyzing the data. It was interesting to learn that how big data had been a transitional technical revolution. The transition of the economy from the traditional pattern of land, labor and capital to time, information and capital was an interesting topic. The impact of the four V's Volume, Velocity, Varity and the fourth new V that I learnt was Varsity. It is essential to understand the importance of the varsity of the analytics achieved from the big data. If the truthfulness of the data is lost, there is no use of the other V's.

"Applications, Analytics and Aspirations" Brenda Dietrich, VP Strategy and CTO for Business Analytics, IBM. Brenda gave us insights extracting values from the big data analysis. IBM added another V to the stack Value. It is important for industry to understand the value derived before implementing a big data implementation. Big data is not just about social media analytics it is more about extracting value from enormous data available in which we are interested. It is important to understand the internal data before exploring external data sources like the web to conduct your business well. It is safe to start within a controlled boundary where results varied by real experience. With the knowledge gained from that, efforts taken to expand the data sources. It is important that big data analytics provides business value to the organization.

"Driving your business through Big Data", Darren Stool and Kerem Tomak, marketing analytics, macys.com. It is interesting to see how customer analytics have grown. Importance have changed from understanding the customers implicit feedbacks and shopping preference to the explicit values like how people tend to correlate two different products and similar analysis. Retail industry has been using customer information effectively and the interest appears been growing since the last five years. Macys has been improving the information infrastructure to capture various customer and inventory data to focus and serve its customers better. The retail industry has been focusing on using data science and automation in analytics, which appears to be a great indication of big data and analytics growth.

"Unlocking business value with big data and predictive analytics", Tim Hood, Global VP Chief Solutions Architect, Retail industry, SAP AG. This session explained the importance of big data to carry out predictive analytics. The business value big data could bring through predictive analytics was enormous. Retail industry is one segment where it could be safely and effectively be tested. SAP had taken clear advantage of this using its product HANA. It was interesting to see how predictive analytics and big data was able to function successfully in the gaming world. The ability to use the gaming logs to create retail opportunity is next level of using predictive analytics. It was interesting to see how HANA functions and how it is able to provide real-time analytics from the big data which helps in realizing the business value of data in short interval of time.

Below you can find my real-time opinion as I was part of the symposium through Twitter. Looking forward for the next

years symposium. Will the focus move from Big data to a new concept before 2014? Looking forward for that advancement in analytics.

Business Intelligence and Analytics internship

By Arunkumar Navaneethan

I did my summer internship in Business Intelligence and analytics. It was a career defining experience. Met some interesting and motivating people. Worked very close to senior technology management resources in the company. Every interaction had led to great learning.

Worked in innovation management projects.

The goal was to support innovation within the company seamlessly and

also in a simple straightforward fashion. Worked in two projects to support this goal. Created a strawman proposal and a working prototype for data virtualization system. The system will integrate various data sources to form a single point system to search for internal events.

In the second project was able to setup an innovation management system according to business needs. Used a research approach in defining the ideation life cycle process. Considered all the pro's and con's associated with the various ideation approaches and mapped it to the business requirements. Conducted various sprints with the business users to create the system. Along the development process identified various enterprise technical issues and worked with the corresponding teams to get it resolved.

This internship helped me grow professional and come out of the shell of what a technical person must know or do within an organization. It is important to understand the companies mission and align our actions towards the business goals.





Machine data analytics, Big Data meetup

By Arunkumar Navaneethan

Machine data analytics, Big Data meetup

Phoenix, Arizona Feb 26th 2014



An event conducted by IBM for developers. Presented about machine data analytics and their big data product Big Insight.

Big insight is more of the customized Hadoop system to fit into an enterprise environment. It has various interfaces in it making it an enterprise ready system for big data challenges.

IBM has configured the Hadoop system to be more efficient by utilizing IBM's standard approach like replacing the event pooling system with message pooling system. Using adaptive MapReduce approach using UDP instead of the TCP messaging.

IBM utilizes various accelerators, which is an interface over its data sources like data warehouse, stream and Hadoop systems.

These accelerators helps in creating better insights, normalizes the data, indexes the data for efficient usage by other interfacing systems like visualization suites.

One of these accelerators is the Machine Data Adapter (MDA). The machine data adapter configured with templates of various standard log file structures. Out of the box, it can extract content from log files like web server log, web log and it as capabilities to create and modify these templates for the given machine data.

These templates are the configuration files, the templates can be in different file formats. Other inbuilt data source types are

- Apache Webaccess
- Delimited Separated Values, or CSV files
- Data Power®
- Generic (used when your data source type is not represented by one of the other log types)
- Hadoop Data Node
- Hadoop Jobtracker
- Hadoop Name Node
- Hadoop Secondary Name Node
- Hadoop Task Attempt
- Hadoop Task Tracker
- Syslog, or system log files
- WebSphere® Application Server

It was interesting to see how big data an ambiguous approach is now getting into a more defined enterprise solution suite. It is evident that big data is more about data in rest being utilized effective that just exploring unstructured data. This use case also demonstrates the importance of semi-structure data. When data is semi-structured, the possibility of extracting values from it is high.

I am looking forward for upcoming sessions on AQL, stream computing (real-time data analytics) sessions.



It is interesting that prediction made from data at rest used in real-time using stream analytics approach.

All the observation above, based on my understanding.

Social media analytics – Online news industry

By Arunkumar Navaneethan

Social media has become a major platform that business use to evaluate customer relationships.

For this project, we used Twitter feeds to understand how online news industry is function.

The goal for this project was to find ways by which the online news industry can improve their revenue.



We used a balance scorecard approach to support the goal. Conducted analysis concentrating on perspectives like finance, customer, internal process, and learning and growth opportunities.

The analysis pointed out that customers twitter communications containing reference to the news agency could be helpful.

Customers interaction can be used to understand which kind of articles are been well conceived. Carried out analysis taking into account various interesting parameters, which twitter records.

Article trends and performance helped us understand how events could influence users. The articles temporal information was limited but we were able to use language properties to classify users.

There was presence of group of articles and user by the category of the article. The users were interested in certain types of articles, which drilled down even to the topic.

The devices the users used also affected what there interest was. The devices could help us classify users and there interest on the type of articles. The device level analysis opened up opportunities, which the companies can capitalize.

The propagation of the articles from online sources, social media platforms and twitter users helped us some interesting patterns.

It was interesting to see how our social media behaviors could be used for making business decisions.

Web the laboratory of human impulse

By Arunkumar Navaneethan

(A topic inspired from a lecture by Professor Sudharam, MIS 587; Business Intelligence)

With online activities becoming an essential part of our day-to-day activity, we are leaving behind a lot of information about us in the web. Different companies have been leveraging the data available in the web, majorly the data present in social media sites. This article is about an news agency and how they leverage the web.



News agency New York Times have been doing research on Tweets. It uses an analysis tool called Cascade which analysis the pattern in which New York Times online articles are shared on the web.

Cascade creates a structure which depicts how the published articles propagates in the social media space. This gives them a multi-dimensional view of how the published articles as been conceived by the public. This would be a great platform for them to understand their audience. This gives the news agency the opportunity to understand how long a story published lives in the social media space.

The tool captures the activities like retweets, replies to the tweets posted by the editors or writers. It also captures tweets which contains the URL to articles published in the New Yorks times website. The tool can capture more interesting data as public share the articles along with their opinion in it. This tool would be available as a Software as a Service, part of the New York Times enterprise initiative; which would capture enterprise bit.ly links shared in through twitter and help organisation capture their social presence in twitter.

References: http://nytlabs.com/projects/cascade.html







With large data comes great responsilibilites

By Arunkumar Navaneethan

Working with data as always fascinated me. I worked in a financial services company first and was part of the transaction process group. It was amazing how each transaction data was process by multiple teams and the level of value the data had. Each team would process the same data in different ways and the value the data had for every group was different. Understood the various data processing carried out on the data in much better way. Why the data is cleansed, formatted, how the integrated of the data mattered a lot, how a mistake in processing the data can



influence the business. Enjoyed this experience of learning data processing in a real corporate environment. Worked in a wide range of database technologies like DB2, Oracle DB and MS SQL. The challenge of understanding a data driven application was very interesting. The value of creating a conceptual schema, relational modeling and designing a relational database with all the integrity constraints made sense.

Now as part of my master's course work, have stepped into a much challenging atmosphere; managing the unstructural data. The days where knowing your data structure before modeling the environment which manages it as slowly being changed. The importance of leveraging textual data like the conversation between people in social media space as grown. Therefore, I am in an interesting place of working with such data. Worked on web and text mining projects and it was quite a challenge. Used Application Programming Interface (API) to extract data from the web related to cities in US. Processed and cleansed this data. Used various data mining algorithms like similarity analysis, clustering and sentiment analysis. Utilized visualization APIs to represent the results in an easily accessible web application. It was a nice learning experience, but now working on much bigger challenge. When we want to extract value from this enormous web, we will need powerful computing abilities. Therefore, the industries buzz word "Big Data" came into picture. My journey with big data has begun where working on analyzing about online tech and social news agencies. Hope my journey with data takes me to better places.

Let data make the world a better place to live in.

Recognitions

Presenter for the Tableau Phoenix User Group – May 2014 on Tableau as a web application



Recommendations





Arunkumar, would you like to recommend Arunkumar? Recommend Arunkumar >

Applications Developer

iNautix Technologies, Bank of New York Mellon



Naveen Jesudas

Vice President at iNautix Technologies, a Bank of New York Mellon company

C Arunkumar Navaneethan was self-motivated and a self-learner in my team. He was able to learn the IBM Lombardi Teamworks BPM tool in a very short period of time on his own. He led the vendor transition efforts in taking over the production support activities of the WebSTP BPM application. He was also able to train the other team members on Lombardi. This was invaluable since recruiting Lombardi resources from the market can be quite challenging.

Arunkumar showed strong initiative in organizing activities for the corporate social responsibilities events within iNautix. He was also an outgoing and invaluable member of my team. **less**

January 17, 2013, Naveen managed Arunkumar at iNautix Technologies, Bank of New York Mellon