VALENCIA – 06/11/2017

Goal-Based Selection of Visual Representations for Big Data Analytics

MATTEO GOLFARELLI, TOMMASO PIRINI, STEFANO RIZZI
DISI – UNIVERSITY OF BOLOGNA

Business Intelligence Group

The Toreador Framework

The wide diffusion of big data technologies and of the increasing amounts of valuable data pushes companies of all sizes to exploit **big data analytics**

However, the **lack of in-house technical skills** often prevents companies from really benefiting of big data analytics, or even discourages them from taking this direction because of the outsourcing costs.

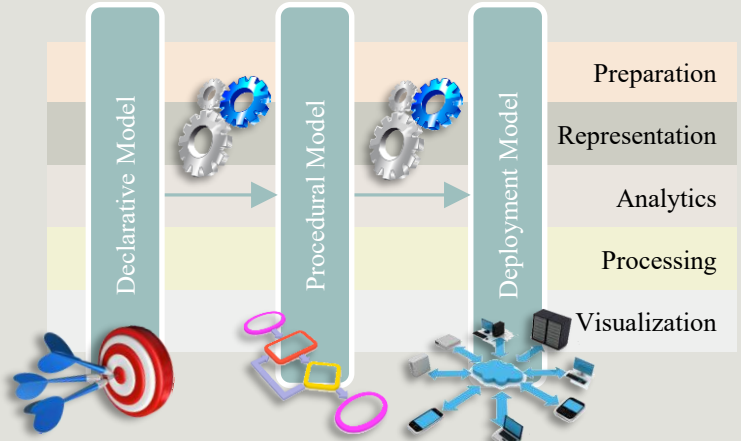
The H2020 TOREADOR Project adopts a **model-driven architecture** to streamline big data analytics and make it widely available to companies as a service

Toreador

The Toreador Framework

- **Declarative Model:** an abstract and platform-independent model that specifies the user goals *what* BDA should achieve in terms of data collection, preparation, analysis, and visualization.
- **Procedural Model:** a platform-neutral, vendor-independent model that specifies the algorithms for data preparation and for parallelizing and executing the analytics, as well as the way to present the results to users (*how* BDA should work).
- **Deployment Model:** the computational components and other resources for the process on a specific target execution platform (e.g., Hadoop-as-a-service).

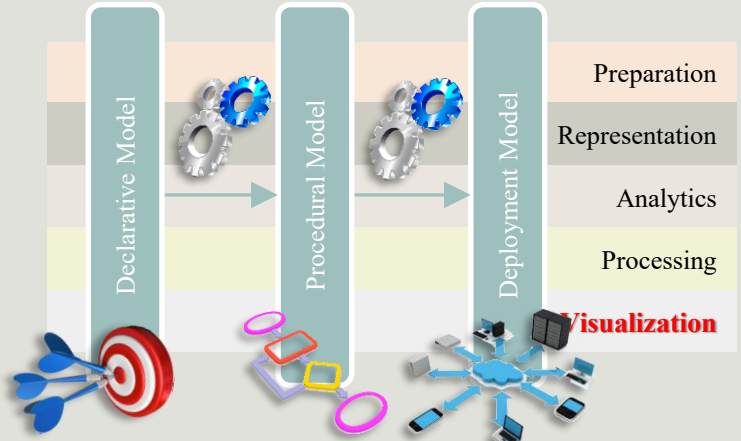
Remarkably, as required by the MDA paradigm, each model is (semi-)automatically derived from the previous one.



The Toreador Framework

In the context of Big Data Analytics the three models are grouped into five conceptual areas: *preparation, representation, analytics, processing*, and **visualization**.

- **how to specify the users objectives** and describe the dataset to be visualized within the declarative model (e.g., comparison-oriented visualization of 4-dimensional numerical data with low-cardinality domains),
- **how to translate this specification into a concrete platform-independent solution** (e.g., bar chart) within the procedural model, which will be eventually translated into a deployment model on the target execution platform (e.g., stacked-to-group bar chart in the D3 Java library).

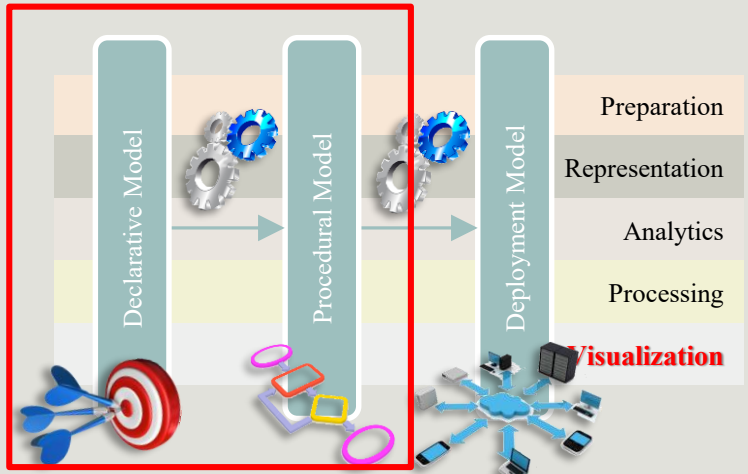


Declarative & Procedural Steps

Define the user desiderata through a **visualization context**

Evaluate visualization types against the Visualization contex through a **suitability function**

Find the **best visualization types** i.e. the undominated ones according to the **skyline** concept



Declarative Visualization Coordinates

Enable users to declare their objectives and describe the dataset to be visualized.

- **Goal:** enables users to declare their analysis goal.
- **Interaction:** enables users to declare the type of interactions to be supported by the visualization.
- **User:** enables users to declare their skill
- **Dimensionality:** which enables users to declare the number of variables they wish to visualize
- **Cardinality:** enables users to qualitatively declare the cardinality of the data to be visualized.
- **Type:** enables users to declare the type of each variable to be analyzed. We distinguish independent from dependent variables.

Most of the coordinate values are taken from the large data visualization literature

The Visualization Context

Let $O_1,\ldots, O_7$ be the sets of goals, interactions, users, dimensionalities, cardinalities, independent types, and dependent types, respectively. let $C=\{1,\ldots, 7\}$ and $O=\cup_{i \in C} O_i$. A **visualization context** is defined by a function **$c: C \rightarrow O \cup \{NULL\}$ (where $c(i) \in O_i \cup \{NULL\}$)** and by a weak order$_c$ $\succ$ on the set C that expresses the priorities between the seven coordinates.

c(1)=Comparison
c(2)=NULL
c(3)=Tech
c(4)=n-dimensional
c(5)=High
c(6)=Interval
c(7) Ratio

$$(3_c \backsim 6)_c \succ 1_c \succ (2_c \backsim 4_c \backsim 5_c \backsim 7)$$

Value	*Objective*	*Example*
	Goal	
Composition	highlighting the way in which distinct parts of data are composed to form a total	stacked column chart
Order	analyzing objects by emphasizing their ordering	alphabetical list of names
Relationship	analyzing the correlation between two or more objects or attribute values	scatter plot
Comparison	examining two or more objects or values to establish their similarities and dissimilarities	column chart
Cluster	analyzing data in such a way as to emphasize their grouping into categories	dendrogram
Distribution	analyzing how objects are dispersed in space	histogram
Trend	examining a general tendency of data variables	line chart
Geospatial	analyzing data values using a geographical map as a graphical context	choropleth map
	Interaction	
Overview	gain an overview of the entire data collection	dendrogram
Zoom	focus on items of interest	network map
Filter	quickly focus on interesting items by eliminating unwanted items	area chart
Details-on-demand	select an item and get its details	choropleth map
	User	
Lay	computer-literates who may have troubles in understanding complex visualizations	line chart
Tech	skilled users with a deeper understanding of BDA	tree map
	Dimensionality	
1-dimensional	a single numerical value or a string	gauge
2-dimensional	one dependent variable as a function of one independent variable	single-line chart
n-dimensional	each data object is a point in an n-dimensional space	bubble chart
Tree	a collection of items, each having a link to one parent item	dendrogram
Graph	a collection of items, each linked to an arbitrary number of other items	network map
	Cardinality	
Low	from a few items to a few dozens items	pie chart
High	some dozens items or more	heat map
	Type	
Nominal	qualitative, each data variable is assigned to one category	pie chart
Ordinal	qualitative, categories can be sorted	histogram
Interval	quantitative, it supports the determination of equality of intervals or differences	line chart
Ratio	quantitative, with a unique and non-arbitrary zero point	scatter plot

The Visualization Context

Let $O_1,\ldots, O_7$ be the sets of goals, interactions, users, dimensionalities, cardinalities, independent types, and dependent types, respec
visualization context is d
(where c(i) $\in O_i \cup$ {NULL
that expresses the priori

c(1)=**Comparison**
c(2)=NULL
c(3)=Tech
c(4)=n-dimensional
c(5)=High
c(6)=Interval
c(7) Ratio

$(3_c \backsim 6)_c \succ$

Value	*Objective*	*Example*
	Goal	
Composition	highlighting the way in which distinct parts of data are composed to form a total	stacked umn cha
Order	analyzing objects by emphasizing their ordering	alphabet list of na
Relationship	analyzing the correlation between two or more objects or attribute values	scatter p
Comparison	examining two or more objects or values to establish their similarities and dissimilarities	column c
Cluster	analyzing data in such a way as to emphasize their grouping into categories	dendrogr
Distribution	analyzing how objects are dispersed in space	histogra
Trend	examining a general tendency of data variables	line char
Geospatial	analyzing data values using a geographical map as a graphical context	choroplet map
	Interaction	
Overview	gain an overview of the entire data collection	dendrogr
Zoom	focus on items of interest	network
Filter	quickly focus on interesting items by eliminating unwanted items	area cha
Details-on-	select an item and get its details	choroplet

The Visua

Let $O_1,\ldots, O_7$ be the sets
dimensionalities, cardina
dependent types, respec
visualization context is d
(where $c(i) \in O_i \cup \{NULL$
that expresses the priori

c(1)=Comparison
c(2)=NULL
c(3)=**Tech**
c(4)=n-dimensional
c(5)=High
c(6)=Interval
c(7) Ratio

$(3_c \backsim 6)_c \succ$

Order	analyzing objects by emphasizing their ordering	alphabet list of na
Relationship	analyzing the correlation between two or more objects or attribute values	scatter p
Comparison	examining two or more objects or values to establish their similarities and dissimilarities	column c
Cluster	analyzing data in such a way as to emphasize their grouping into categories	dendrogr
Distribution	analyzing how objects are dispersed in space	histogra
Trend	examining a general tendency of data variables	line char
Geospatial	analyzing data values using a geographical map as a graphical context	choroplet map
Interaction		
Overview	gain an overview of the entire data collection	dendrogr
Zoom	focus on items of interest	network
Filter	quickly focus on interesting items by eliminating unwanted items	area char
Details-on-demand	select an item and get its details	choroplet map
User		
Lay	computer-literates who may have troubles in understanding complex visualizations	line char
Tech	skilled users with a deeper understanding of BDA	tree map
Dimensionality		
1-dimensional	a single numerical value or a string	gauge
2-dimensional	one dependent variable as a function of one independent variable	single-lin chart
n-dimensional	each data object is a point in an n-dimensional space	bubble c
Tree	a collection of items, each having a link to one parent item	dendrogr
Graph	a collection of items, each linked to an arbitrary number of other items	network
Cardinality		

The Visua

Let $O_1,\ldots, O_7$ be the sets o
dimensionalities, cardinal
dependent types, respect
visualization context is de
(where $c(i) \in O_i \cup \{NULL\}$
that expresses the prioriti

c(1)=Comparison
c(2)=NULL
c(3)=Tech
c(4)=**n-dimensional**
c(5)=High
c(6)=Interval
c(7) Ratio

$(3_c \backsim 6)_c \succ$

Comparison	examining two or more objects or values to establish their similarities and dissimilarities	column
Cluster	analyzing data in such a way as to emphasize their grouping into categories	dendrog
Distribution	analyzing how objects are dispersed in space	histogra
Trend	examining a general tendency of data variables	line cha
Geospatial	analyzing data values using a geographical map as a graphical context	choropl map
Interaction		
Overview	gain an overview of the entire data collection	dendrog
Zoom	focus on items of interest	network
Filter	quickly focus on interesting items by eliminating unwanted items	area cha
Details-on-demand	select an item and get its details	choropl map
User		
Lay	computer-literates who may have troubles in understanding complex visualizations	line cha
Tech	skilled users with a deeper understanding of BDA	tree ma
Dimensionality		
1-dimensional	a single numerical value or a string	gauge
2-dimensional	one dependent variable as a function of one independent variable	single-li chart
n-dimensional	each data object is a point in an n-dimensional space	bubble
Tree	a collection of items, each having a link to one parent item	dendrog
Graph	a collection of items, each linked to an arbitrary number of other items	network
Cardinality		
Low	from a few items to a few dozens items	pie char
High	some dozens items or more	heat ma
Type		
Nominal	qualitative, each data variable is assigned to one category	pie char

The Visualization Context

Let $O_1,\ldots, O_7$ be the sets of goals, interactions, users, dimensionalities, cardinalities, independent types, and dependent types, respectively. let $C=\{1,\ldots, 7\}$ and $O=\cup_{i \in C} O_i$. A **visualization context** is defined by a function c: $C \rightarrow O \cup \{NULL\}$ (where $c(i) \in O_i \cup \{NULL\}$) and by a **weak order**$_c$ $\succ$ on the set C that expresses the priorities between the seven coordinates.

$$(3 \sim_c 6) \succ_c 1 \succ_c (2 \sim_c 4 \sim_c 5 \sim_c 7)$$

User = Tech $\sim_c$ Independent types = Interval

↓

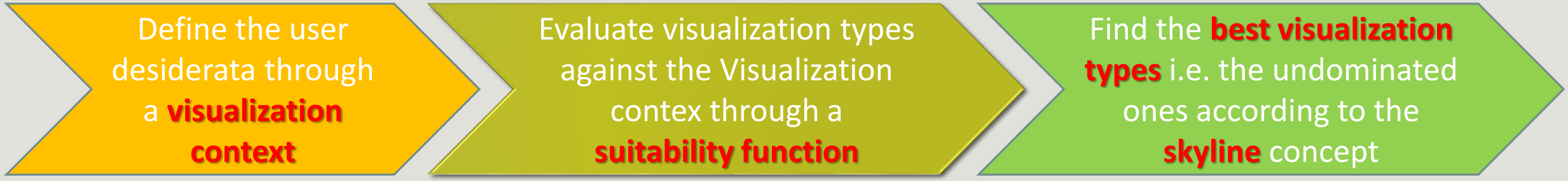
Goal = Comparison

↓

Interaction = NULL $\sim_c$ Dimensionality = n-dimensional $\sim_c$ Cardinality = High $\sim_c$ Independent types = Ratio

Value	*Objective*	*Example*
	Goal	
Composition	highlighting the way in which distinct parts of data are composed to form a total	stacked column chart
Order	analyzing objects by emphasizing their ordering	alphabetical list of names
Relationship	analyzing the correlation between two or more objects or attribute values	scatter plot
Comparison	examining two or more objects or values to establish their similarities and dissimilarities	column chart
Cluster	analyzing data in such a way as to emphasize their grouping into categories	dendrogram
Distribution	analyzing how objects are dispersed in space	histogram
Trend	examining a general tendency of data variables	line chart
Geospatial	analyzing data values using a geographical map as a graphical context	choropleth map
	Interaction	
Overview	gain an overview of the entire data collection	dendrogram
Zoom	focus on items of interest	network map
Filter	quickly focus on interesting items by eliminating unwanted items	area chart
Details-on-demand	select an item and get its details	choropleth map
	User	
Lay	computer-literates who may have troubles in understanding complex visualizations	line chart
Tech	skilled users with a deeper understanding of BDA	tree map
	Dimensionality	
1-dimensional	a single numerical value or a string	gauge
2-dimensional	one dependent variable as a function of one independent variable	single-line chart
n-dimensional	each data object is a point in an n-dimensional space	bubble chart
Tree	a collection of items, each having a link to one parent item	dendrogram
Graph	a collection of items, each linked to an arbitrary number of other items	network map
	Cardinality	
Low	from a few items to a few dozens items	pie chart
High	some dozens items or more	heat map
	Type	
Nominal	qualitative, each data variable is assigned to one category	pie chart
Ordinal	qualitative, categories can be sorted	histogram
Interval	quantitative, it supports the determination of equality of intervals or differences	line chart
Ratio	quantitative, with a unique and non-arbitrary zero point	scatter plot

Declarative & Procedural Steps



The Suitability Function

A **suitability function** is a function $\sigma: O \times V \rightarrow s$ where O is the set of all coordinate values, V is the set of all visualization types, and $s \in$ {unfit, discouraged, neutral, acceptable, fit} is a suitability score.

unfit < discouraged < neutral < acceptable < fit

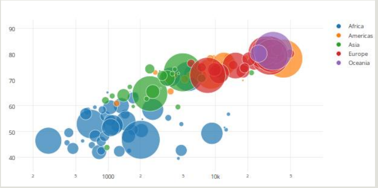
Our approach is general enough to be applicable to each possible visualization type v as long as a suitability evaluation is done for v based on our seven coordinates.

		pie chart	bubble chart	heat map
Goal:	Composition	fit	unfit	unfit
	Order	neutral	unfit	unfit
	Relationship	unfit	fit	unfit
	Comparison	neutral	fit	acceptable
	Cluster	unfit	acceptable	acceptable
	Distribution	discouraged	fit	fit
	Trend	unfit	fit	unfit
	Geospatial	unfit	unfit	fit
Interaction:	Overview	fit	acceptable	fit
	Zoom	unfit	acceptable	fit
	Filter	neutral	neutral	neutral
	Details-on-dem	acceptable	neutral	neutral
User:	Lay	fit	acceptable	acceptable
	Tech	acceptable	fit	fit
Dimensionality:	1-dimensional	unfit	unfit	unfit
	2-dimensional	fit	unfit	unfit
	n-dimensional	unfit	fit	fit
	Tree	unfit	unfit	unfit
	Graph	unfit	unfit	unfit
Cardinality:	Low	fit	acceptable	discouraged
	High	discouraged	discouraged	fit
Independent Type:	Nominal	fit	unfit	unfit
	Ordinal	acceptable	neutral	unfit
	Interval	discouraged	fit	fit
	Ratio	discouraged	fit	unfit
Dependent Type:	Nominal	unfit	fit	unfit
	Ordinal	unfit	fit	discouraged
	Interval	unfit	fit	fit
	Ratio	fit	fit	fit

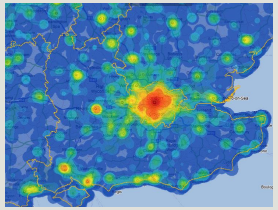
Sum of Amount	Product							
Country	Apple	Banana	Beans	Broccoli	Carrots	Mango	Orange	Grand Total
Australia	20634	52721	14433	17953	8106	9186	8680	131713
Canada	24867	33775		12407		3767	19929	94745
France	80193	36094	680	5341	9104	7388	2256	141056
Germany	9082	39686	29905	37197	21636	8775	8887	155168
New Zealand	10332	40050		4390			12010	66782
United Kingdom	17534	42908	5100	38436	41815	5600	21744	173137
United States	28615	95061	7163	26715	56284	22363	30932	267133
Grand Total	191257	340295	57281	142439	136945	57079	104438	1029734

Pivot Table

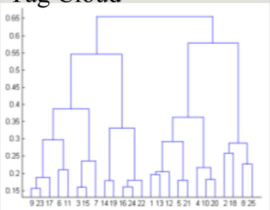
Tag Cloud



Bubble chart



Heat map



Dendrogram

The Suitability Function

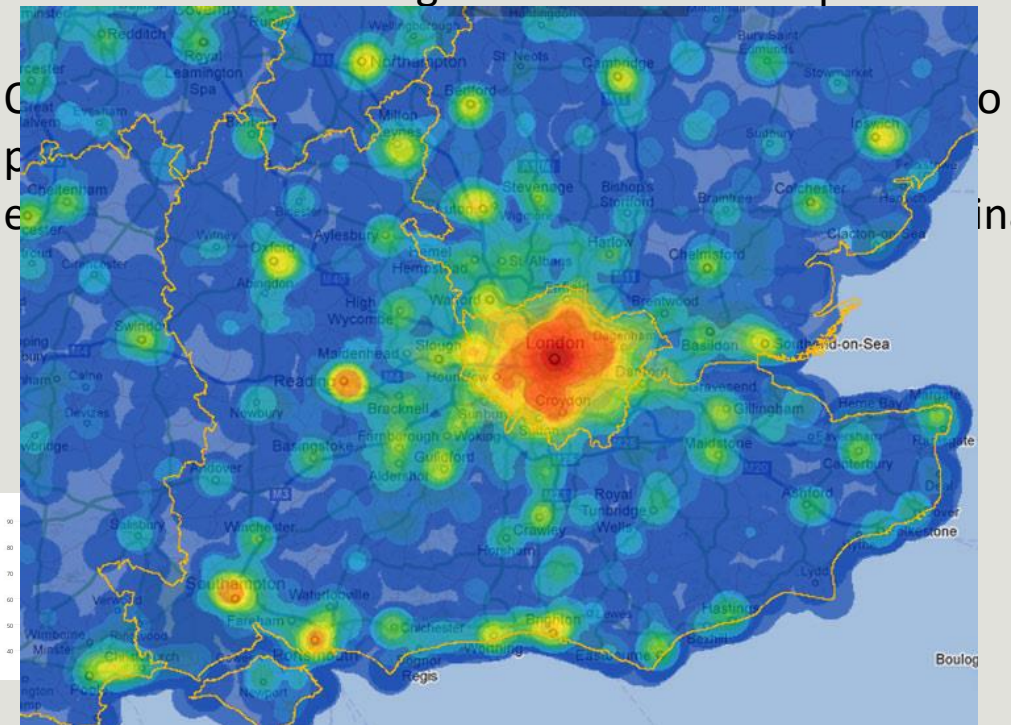
Composition: highlighting the way in which distinct parts of data are composed to form a total & stacked column chart

ral,

unfit < discouraged < neutral < acceptable < fit

o each

inates.



		pie chart	bubble chart	heat map
				unfit
		neutral	unfit	unfit
	Relationship	unfit	fit	unfit
	Comparison	neutral	fit	acceptable
	Cluster	unfit	acceptable	acceptable
	Distribution	discouraged	fit	fit
	Trend	unfit	fit	unfit
	Geospatial	unfit	unfit	fit
Interaction:	Overview	fit	acceptable	fit
	Zoom	unfit	acceptable	fit
	Filter	neutral	neutral	neutral
	Details-on-dem	acceptable	neutral	neutral
User:	Lay	fit	acceptable	acceptable
	Tech	acceptable	fit	fit
Dimensionality:	1-dimensional	unfit	unfit	unfit
	2-dimensional	fit	unfit	unfit
	n-dimensional	unfit	fit	fit
	Tree	unfit	unfit	unfit
	Graph	unfit	unfit	unfit
Cardinality:	Low	fit	acceptable	discouraged
	High	discouraged	discouraged	fit
Independent Type:	Nominal	fit	unfit	unfit
	Ordinal	acceptable	neutral	unfit
	Interval	discouraged	fit	fit
	Ratio	discouraged	fit	unfit
Dependent Type:	Nominal	unfit	fit	unfit
	Ordinal	unfit	fit	discouraged
	Interval	unfit	fit	fit
	Ratio	fit	fit	fit

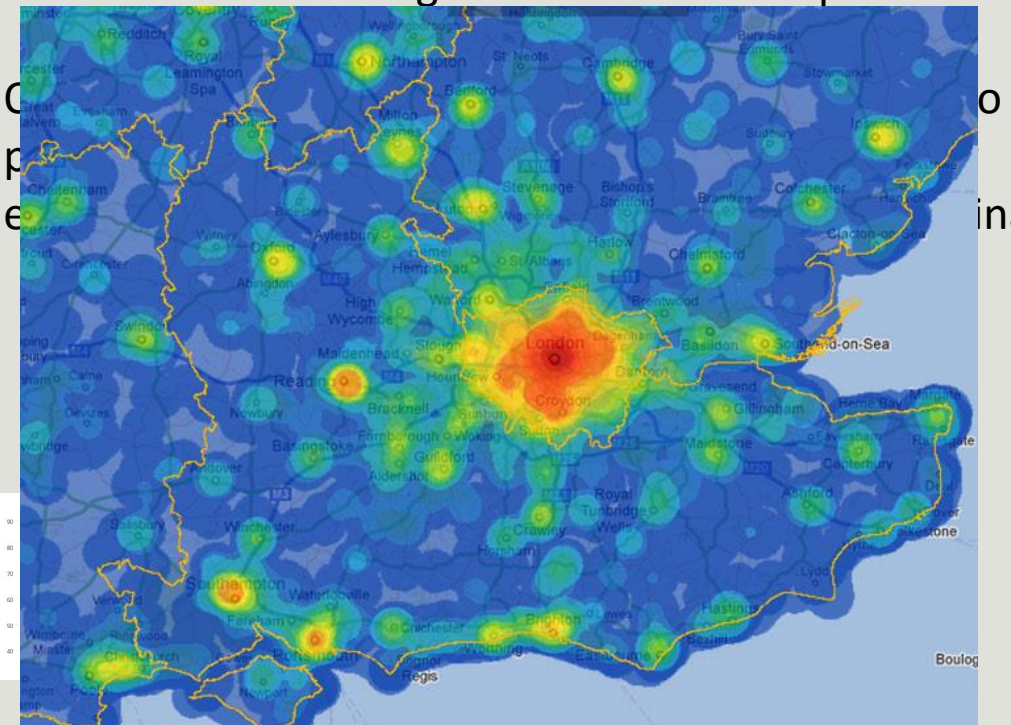
The Suitability Function

Comparison: examining two or more objects or values to establish their similarities and dissimilarities

e *O*

ral.

unfit < discouraged < neutral < acceptable < fit



o each

inates.

		pie chart	bubble chart	heat map
Goal:	Composition	fit	unfit	unfit
	Order	neutral	unfit	unfit
	Relationship	unfit	fit	unfit
	Comparison	neutral	fit	acceptable
		unfit	acceptable	acceptable
	Distribution	discouraged	fit	fit
	Trend	unfit	fit	unfit
	Geospatial	unfit	unfit	fit
Interaction:	Overview	fit	acceptable	fit
	Zoom	unfit	acceptable	fit
	Filter	neutral	neutral	neutral
	Details-on-dem	acceptable	neutral	neutral
User:	Lay	fit	acceptable	acceptable
	Tech	acceptable	fit	fit
Dimensionality:	1-dimensional	unfit	unfit	unfit
	2-dimensional	fit	unfit	unfit
	n-dimensional	unfit	fit	fit
	Tree	unfit	unfit	unfit
	Graph	unfit	unfit	unfit
Cardinality:	Low	fit	acceptable	discouraged
	High	discouraged	discouraged	fit
Independent Type:	Nominal	fit	unfit	unfit
	Ordinal	acceptable	neutral	unfit
	Interval	discouraged	fit	fit
	Ratio	discouraged	fit	unfit
Dependent Type:	Nominal	unfit	fit	unfit
	Ordinal	unfit	fit	discouraged
	Interval	unfit	fit	fit
	Ratio	fit	fit	fit

The Suitability Function

Geospatial: analyzing data values using a geographical map as a graphical context

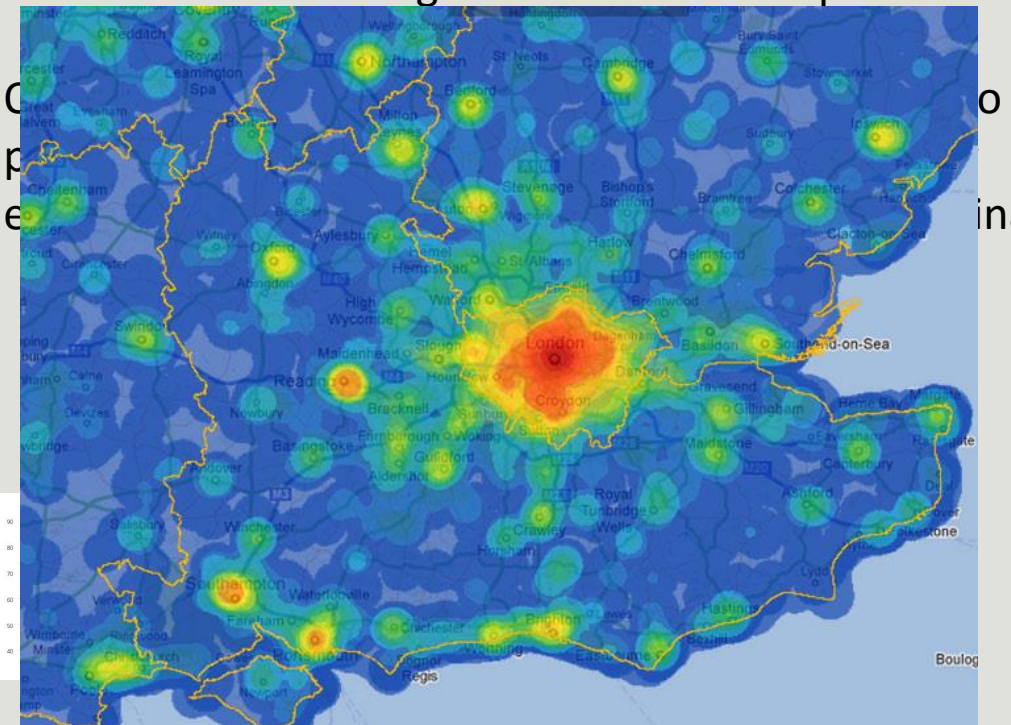
e *O*

ral,

unfit < discouraged < neutral < acceptable < fit

o each

inates.



		pie chart	bubble chart	heat map
Goal:	Composition	fit	unfit	unfit
	Order	neutral	unfit	unfit
	Relationship	unfit	fit	unfit
	Comparison	neutral	fit	acceptable
	Cluster	unfit	acceptable	acceptable
	Distribution	discouraged	fit	fit
	[illegible]	[illegible]	fit	unfit
	Geospatial	unfit	unfit	fit
Interaction:	Overview	fit	acceptable	fit
	Zoom	unfit	acceptable	fit
	Filter	neutral	neutral	neutral
	Details-on-dem	acceptable	neutral	neutral
User:	Lay	fit	acceptable	acceptable
	Tech	acceptable	fit	fit
Dimensionality:	1-dimensional	unfit	unfit	unfit
	2-dimensional	fit	unfit	unfit
	n-dimensional	unfit	fit	fit
	Tree	unfit	unfit	unfit
	Graph	unfit	unfit	unfit
Cardinality:	Low	fit	acceptable	discouraged
	High	discouraged	discouraged	fit
Independent Type:	Nominal	fit	unfit	unfit
	Ordinal	acceptable	neutral	unfit
	Interval	discouraged	fit	fit
	Ratio	discouraged	fit	unfit
Dependent Type:	Nominal	unfit	fit	unfit
	Ordinal	unfit	fit	discouraged
	Interval	unfit	fit	fit
	Ratio	fit	fit	fit

Declarative & Procedural Steps

Define the user desiderata through a **visualization context**

Evaluate visualization types against the Visualization contex through a **suitability function**

Find the **best visualization types** i.e. the undominated ones according to the **skyline** concept

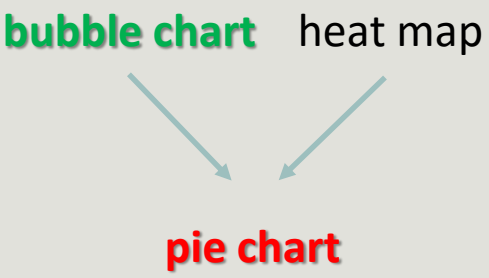
Flat (non prioritized) Dominance

Given visualization context *c* and two visualization types *v* and *v*', and given the set of coordinates C' $\subseteq$ C, we say that:

v is ***flat-substitutable*** to *v*' on *C*', denoted $v \underset{C'}{\sim} v'$$, iff $\sigma(c(j),v) = \sigma(c(j),v')$ for all $j \in C'$ such that $c(j) \neq$ NULL

v ***flat-dominates*** *v*' on *C*', denoted $v \underset{C'}{\succ} v'$, iff

- $\exists\, i \in C' : \sigma(c(j),v) > \sigma(c(j),v')$
- for all other $j \in C'$ such that $c(j) \neq$ NULL it is $\sigma(c(j),v) = \sigma(c(j),v')$



	pie chart	bubble chart	heat map
Goal: Comparison	neutral	fit	acceptable
Interaction: NULL	—	—	—
User: Tech	acceptable	fit	fit
Dimensionality: n-dim	unfit	fit	fit
Cardinality: High	discouraged	discouraged	fit
Independent Type: Interval	discouraged	fit	fit
Dependent Type: Ratio	fit	fit	fit

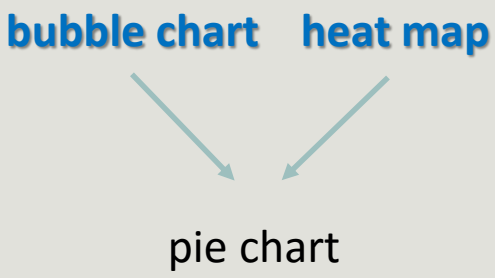
Flat (non prioritized) Dominance

Given visualization context *c* and two visualization types *v* and *v*', and given the set of coordinates C' $\subseteq$ C, we say that:

v is ***flat-substitutable*** to *v*' on *C*', denoted $v \underset{C'}{\sim} v'$$, iff $\sigma(c(j),v) = \sigma(c(j),v')$ for all $j \in C'$ such that $c(j) \neq$ NULL

v ***flat-dominates*** *v*' on *C*', denoted $v \underset{C'}{\succ} v'$, iff

- $\exists\, i \in C' : \sigma(c(j),v) > \sigma(c(j),v')$
- for all other $j \in C'$ such that $c(j) \neq$ NULL it is $\sigma(c(j),v) = \sigma(c(j),v')$



	pie chart	bubble chart	heat map
Goal: Comparison	neutral	fit	acceptable
Interaction: NULL	—	—	—
User: Tech	acceptable	fit	fit
Dimensionality: n-dim	unfit	fit	fit
Cardinality: High	discouraged	discouraged	fit
Independent Type: Interval	discouraged	fit	fit
Dependent Type: Ratio	fit	fit	fit

Dominance

Given visualization context c, $\succ_c$ and two visualization types v and v', we say that v ***dominates*** v' denoted $v \succ_c v'$, iff either

- $v_{\max(C)} \asymp v'$
- $v_{\max(C)} \succ v' \wedge v_{C \setminus \max(C)} \succ v'$

bubble chart

↓

heat map

↓

pie chart

$(3_c \asymp 6)_c \succ 1_c \succ (2_c \asymp 4_c \asymp 5_c \asymp 7)$

		pie chart	bubble chart	heat map
Goal:	Comparison	neutral	fit	acceptable
Interaction:	NULL	—	—	—
User:	Tech	acceptable	fit	fit
Dimensionality:	n-dim	unfit	fit	fit
Cardinality:	High	discouraged	discouraged	fit
Independent Type:	Interval	discouraged	fit	fit
Dependent Type:	Ratio	fit	fit	fit

Dominance

Given visualization context c, $_c \succ$ and two visualization types v and v', we say that v ***dominates*** v' denoted $v \,_c\!\succ v'$, iff either

- $v_{\max(C)} \asymp v'$
- $v_{\max(C)} \succ v' \wedge v_{C \setminus \max(C)} \succ v'$

bubble chart

↓

heat map

↓

pie chart

$(3 \,_c\!\asymp 6) \,_c\!\succ 1 \,_c\!\succ (2 \,_c\!\asymp 4 \,_c\!\asymp 5 \,_c\!\asymp 7)$

	pie chart	bubble chart	heat map
Goal: Comparison	neutral	fit	acceptable
Interaction: NULL	—	—	—
User: Tech	acceptable	fit	fit
Dimensionality: n-dim	unfit	fit	fit
Cardinality: High	discouraged	discouraged	fit
Independent Type: Interval	discouraged	fit	fit
Dependent Type: Ratio	fit	fit	fit

Dominance

Given visualization context c, $_c \succ$ and two visualization types v and v', we say that v ***dominates*** v' denoted $v \,_c\!\succ v'$, iff either

- $v_{\max(C)} \backsim v'$
- $v_{\max(C)} \succ v' \wedge v_{C \setminus \max(C)} \succ v'$

bubble chart

↓

heat map

↓

pie chart

$(3_c \backsim 6)\,_c\!\succ 1\,_c\!\succ (2\,_c\!\backsim 4\,_c\!\backsim 5\,_c\!\backsim 7)$

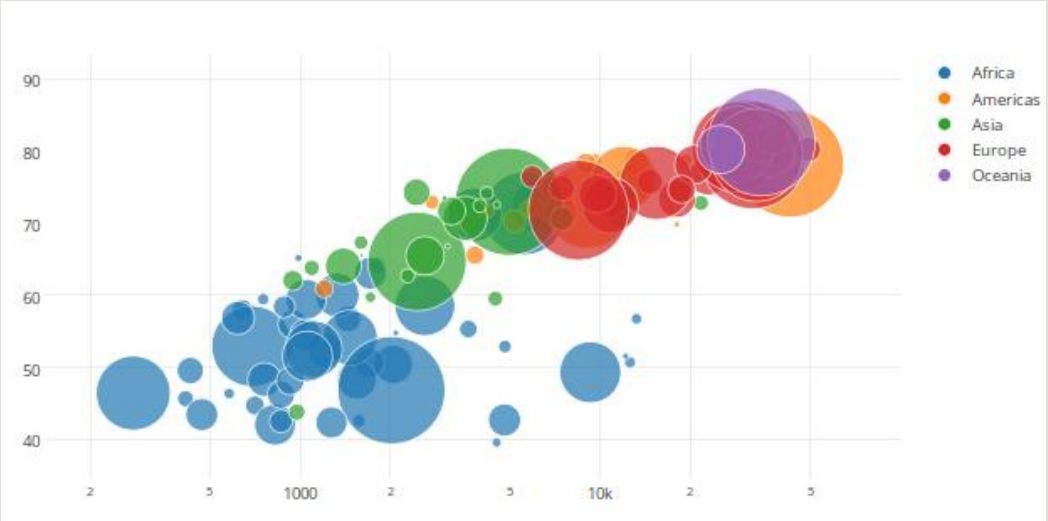
	pie chart	bubble chart	heat map
Goal: Comparison	neutral	fit	acceptable
Interaction: NULL	—	—	—
User: Tech	acceptable	fit	fit
Dimensionality: n-dim	unfit	fit	fit
Cardinality: High	discouraged	discouraged	fit
Independent Type: Interval	discouraged	fit	fit
Dependent Type: Ratio	fit	fit	fit

To close the loop

User = Tech $_c \backsim$ Independent types = Interval

Goal = Comparison

Interaction = NULL $_c \backsim$ Dimensionality = n-dimensional
$_c \backsim$ Cardinality = High $_c \backsim$ Independent types = Ratio



Future Works

Our future work mainly concerns the translation from the procedural to the deployment level of the TOREADOR platform.

1. **The user** chooses her preferred chart (e.g., bubble chart) among those in the skyline
2. **The system** will support the user in mapping each data variable onto a specific dimension of the chart
   - First variable onto X axis
   - Second variable onto Y axis
   - Third variable onto bubble color
   - Fourth variable onto bubble size

Prof. Matteo Golfarelli (Phd)
Computer Science & Engineering
University of Bologna
Tel: +39 0547 338 862
e-mail: matteo.golfarelli@unibo.it
skype: matteo.golfarelli

www: http://bias.csr.unibo.it/golfarelli/
BIG: http://big.csr.unibo.it/

Business Intelligence Group