



Commercial Visual Analytics Systems

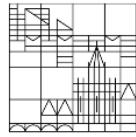


Michael Behrisch, Dirk Streeb, Florian Stoffel, Daniel Seebacher, Brian Matejek, Stefan Hagen Weber, Sebastian Mittelstaedt, Hanspeter Pfister, Daniel Keim



HARVARD
John A. Paulson
School of Engineering
and Applied Sciences

Universität
Konstanz



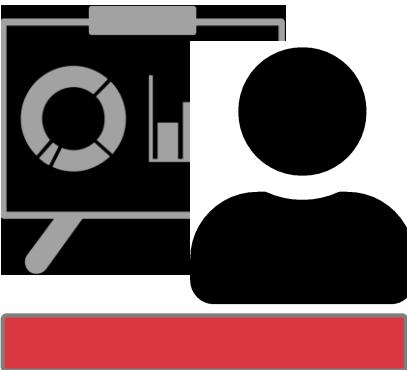
SIEMENS
Ingenuity for life



Upper **Management**



Reporting **Manager**



Data **Analyst**







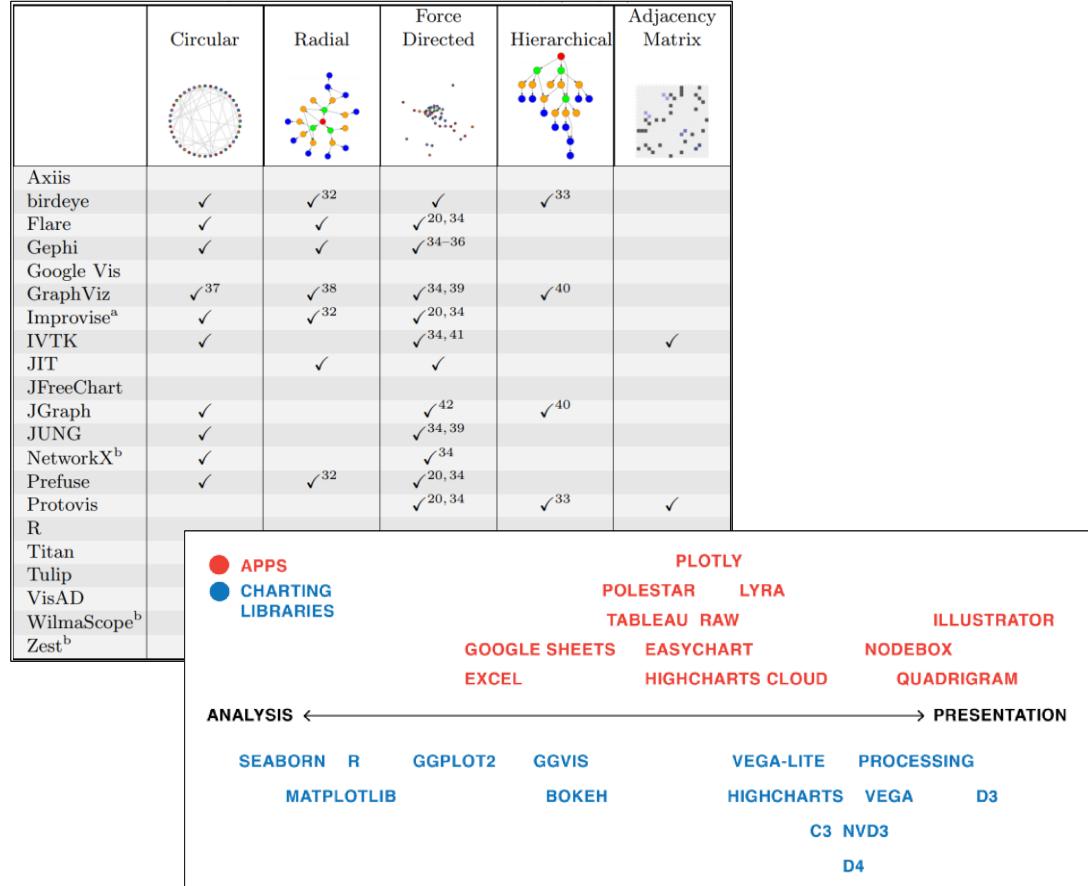
Existing Surveys

Survey of open source toolkits

“Comparison of Open Source Visual Analytics Toolkits” by John Harger et. al

Blog posts on BI/Infovis tools

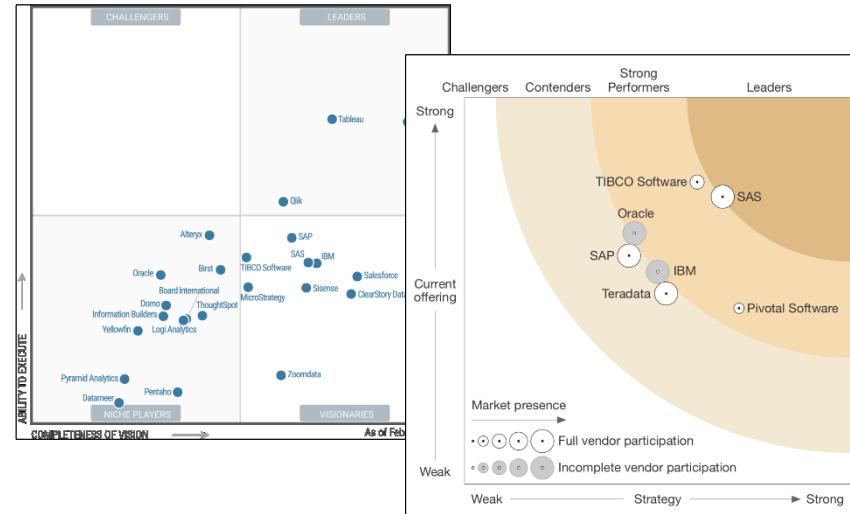
“What I Learned Recreating One Chart Using 24 Tools” by Lisa Charlotte Rost





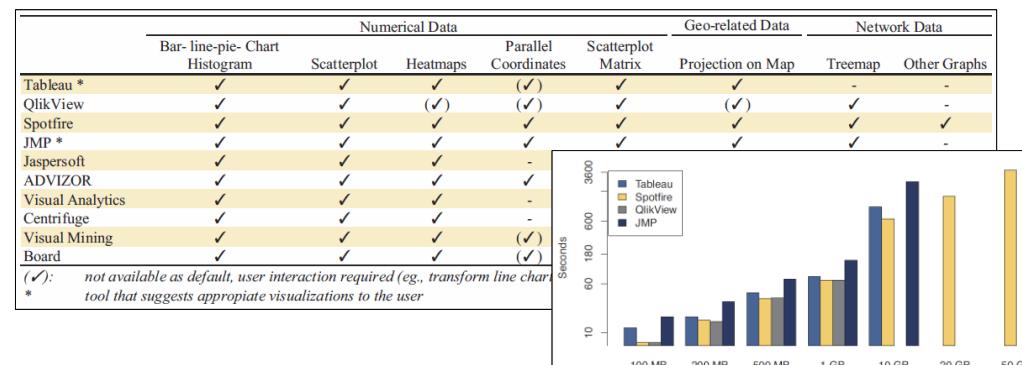
Existing Surveys

Surveys of BI systems
from *Gartner, Forester, Passionned, etc.*



Our 2012 Survey

*“Visual Analytics for the Big Data Era –
A Comparative Review of State-of-the-Art
Commercial Systems”*



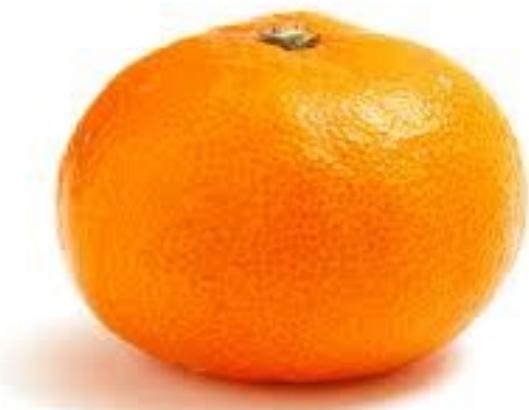


Survey Structure



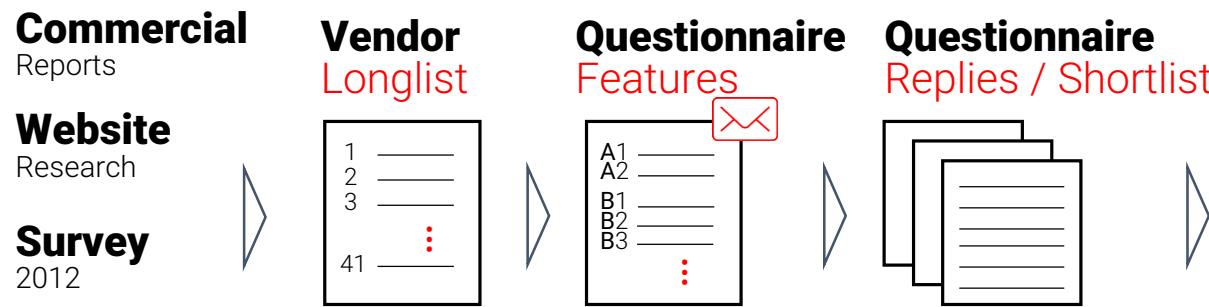
Challenge

how do we compare them?

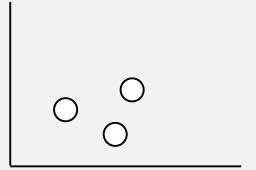




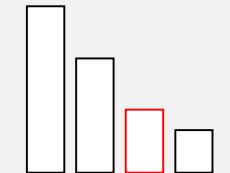
Survey Pipeline



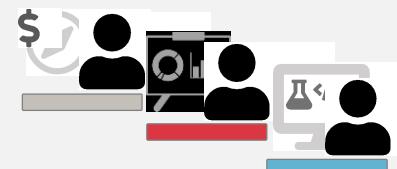
Feature Comparison
Criteria Landscape



Performance Evaluation
Timing Comparison



Case Study Evaluation
Findings Target Groups





Survey Design

Feature Comparison

Questionnaire
Design

- D Data Handling and -Management
- A Automatic Analysis
- C Complex Data Types
- V Visualization
- P Guidance, Perception, Cognition
- I Infrastructure

Performance & Case Study

Data-driven
Evaluation Design

- a Loading Stress Test
- b Analysis Stress Test
- c Visualization Stress Test
- d Use Case Evaluation VAST 2015 Dataset

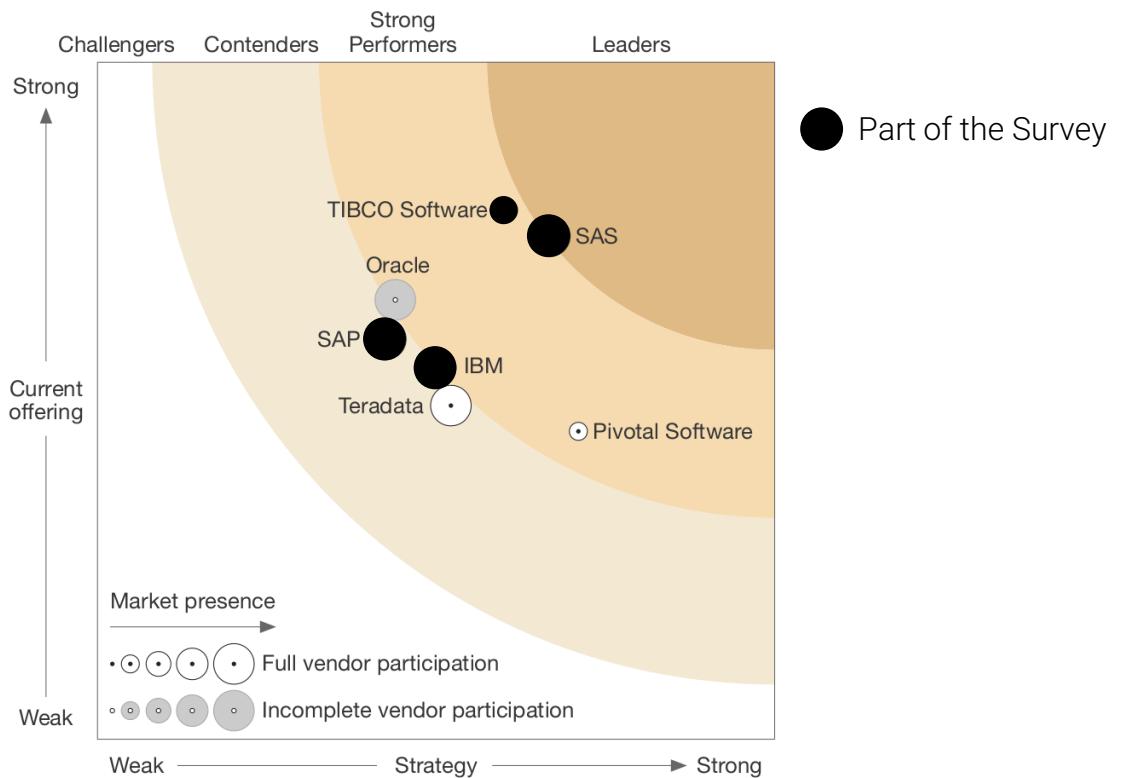


System Selection

Invited vendors of 46 systems

All respondents included

Added QlikView and SAP Lumira
later Microsoft PowerBi was added



Forrester 2016: The Forrester Wave™:
Enterprise Insight Platform Suites, Q4 2016, p. 13



System Selection

Invited vendors of 46 systems

All respondents included

Added QlikView and SAP Lumira
later Microsoft PowerBi was added



Gartner 2017: Magic Quadrant for Business Intelligence and Analytics Platforms, p. 6



User Categories

Upper Management

Present information (e.g. at board or shareholder meetings)

Storytelling capabilities



Reporting Manager

Confirmatory or hypothesis-driven analysis

Broad range of interactive analysis and visualization techniques



Data Analyst

Exploratory analysis

Extensibility, interactivity and data handling





Questionnaire



Vendor Questionnaire

commercialtools.dbvis.de/survey

The following surveys are available:
No available surveys
Please contact Administrator (commercialtools@dbvis.inf.uni-konstanz.de) for further assistance.

Language: English - English •

[Toggle legend](#)

Survey 2017

Exit and clear survey

Commercial Tools Survey 2017

Welcome to our 2017 Commercial Tools Survey.
*** Deadline 3rd of March, 2017 ***

Five years have elapsed since we published successfully our first commercial visual analytics tool survey "Visual analytics for the big data era - a comparative review of state-of-the-art commercial systems" in the Information Visualization and Visual Analytics community [1]. Now we are planning to update the Information Visualization and Visual Analytics Community about the current progress and emphasize future development directions. The planned survey will be published and presented in front of decision makers in large companies and academics that will soon join them.

Aim of the survey:

- Independent and competitive survey of Commercial Visual Analytics Systems
- Show the State-of-the-Art and highlight novel developments
- Show developments and future directions within the field
- Bring together the research and product vendors.

Survey Procedure:

- The survey system will be closed (hard deadline) on the **3rd of March, 2017**.
- The survey link **MAY BE DISTRIBUTED** in your company to the responsible persons.
- The survey is structured along the topics: General Questions, Data Import/Export, Data Preparation and Cleaning, General Analysis, Automatic Analysis, Visual Analysis, Presentation of Results, Extensibility/Integration in Existing Workflows, Security/Working with Sensitive Data, Collaboration, Support
- The survey system is session-based and allows saving and resuming.

In case you have any questions please do not hesitate to send us an email (commercial-star@dbvis.inf.uni-konstanz.de)

There are 53 questions in this survey.

[Next](#)

Survey 2017

Resume later Exit and clear survey Question Index •

General

To begin with, please answer some general questions about your product.

1 What is the name of your product?
My Commercial Visual Analytics Tool

2 Which version of your product do you refer to when answering the following questions?
1.0

3 Which operating systems does your product support to access and explore the data?
Check all that apply

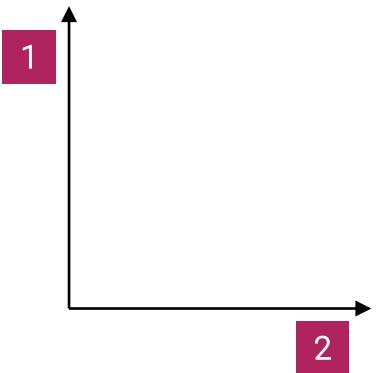
Windows
 Mac OS
 Linux



Vendor Questionnaire – Structure

- D Data Handling and -Management
- A Automatic Analysis
- C Complex Data Types
- V Visualization
- P Guidance, Perception, Cognition
- I Infrastructure

- 1 Feature Richness
- 2 Degree of Innovation





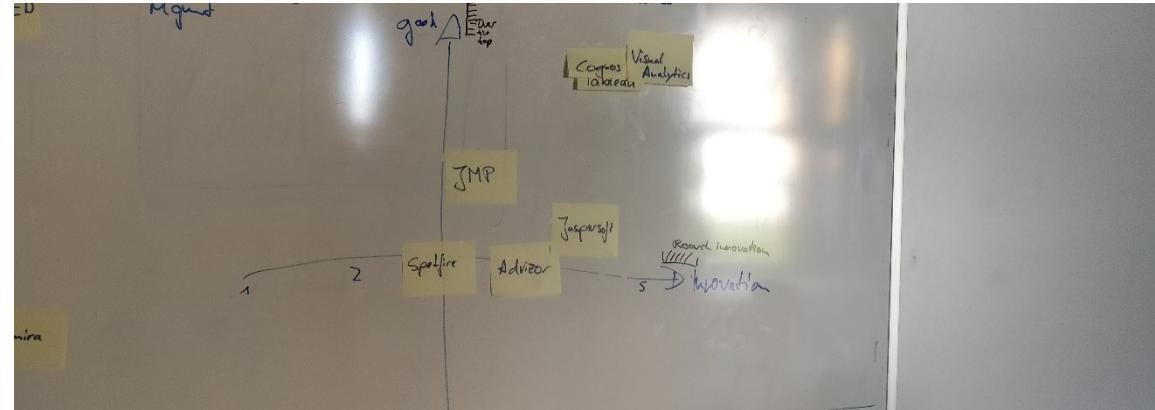
Vendor Questionnaire – Evaluation

Four visual analytics experts went **individually** through the answers

Mode: category by category,
comparative evaluation of the answers

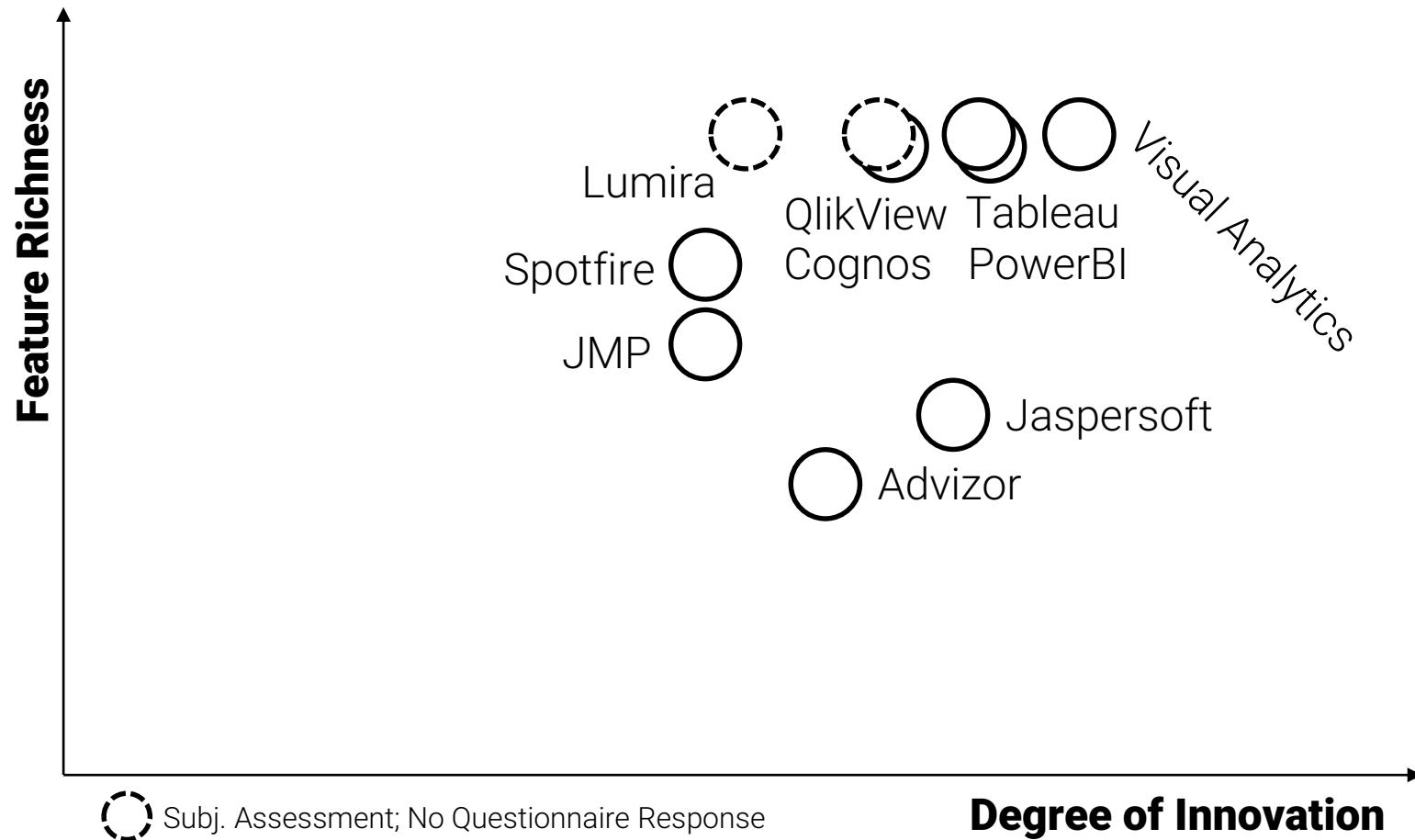
Consent must be reach among the experts
for each tool and the overall overview

Final **discussion** of the overall ratings





Data Handling and -Management





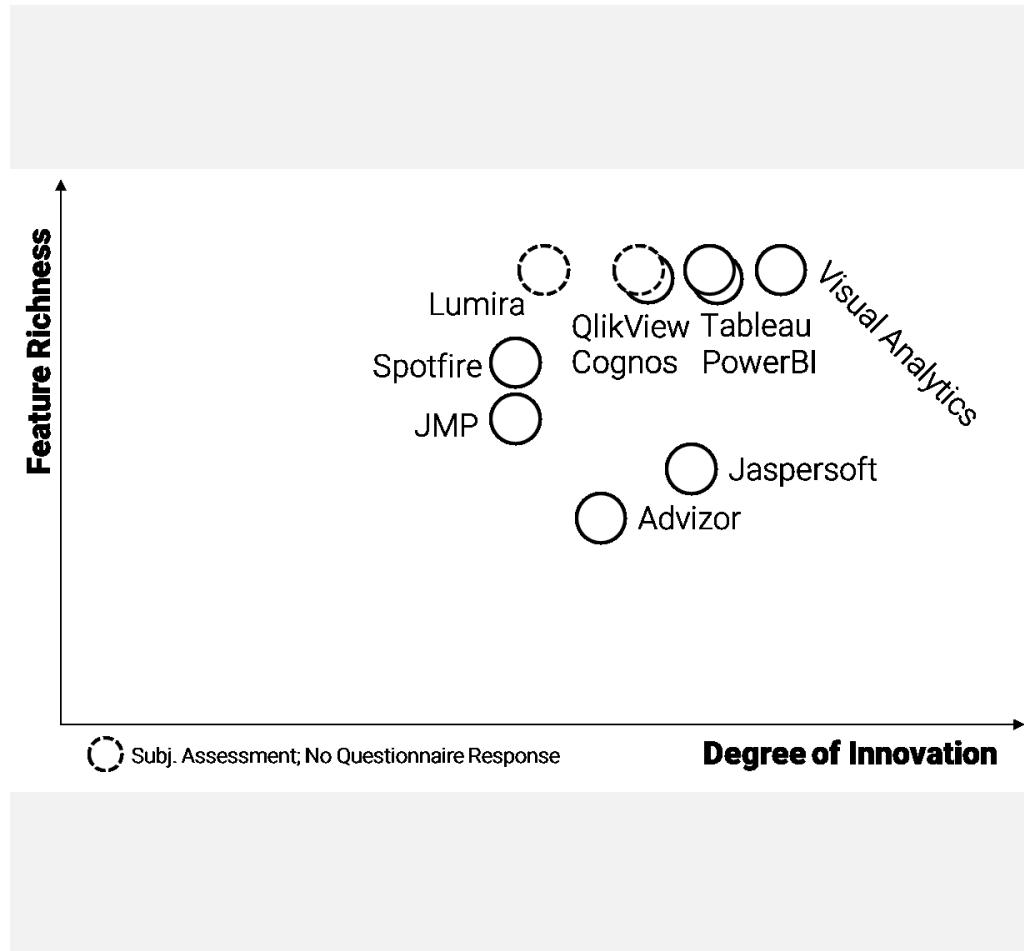
		Data Management, ETL and Preprocessing							
		Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Data Sources	Import	CSV files	✓	✓	✓	✓	✓	✓	✓
	Import	Excel-files (.xls/.xlsx)	✓	✓	✓	✓	✓	✓	✓
	Import	JSON APIs (REST)	✓	✓	✓	✓	✓	✓	✓
	Import	JSON-files	✓	✓	✓	✓	✓	✓	✓
	Import	XML (REST)	✓	✓	✓	✓	✓	✓	✓
	Import	XML files		✓	✓	✓	✓	✓	✓
	Import	SQL-Databases (e.g. ODBC/Postgres)	✓	✓	✓	✓	✓	✓	✓
	Import	NOSQL-Databases (e.g. MongoDB)	✓		✓	✓	✓	✓	✓
	Import	Directory of text/image/video		✓			✓	✓	✓
	Import	OPC (UA)					✓	✓	✓
Export	Export	SAP	✓		✓	✓	✓	✓	✓
	Export	Salesforce	✓	✓	✓	✓	✓	✓	✓
	Export	CSV	✓	✓	✓	✓	✓	✓	✓
	Export	Excel	✓	✓	✓	✓	✓	✓	✓
	Export	JSON			✓	✓	✓	✓	✓
	Export	XML			✓	✓	✓	✓	✓
Writeback	To data source	SQL			✓	✓	✓	✓	✓
	To data source	OPC (UA)							



		Data Management, ETL and Preprocessing							
		Tableau	Advizor	JMP	Cognos	Jaspersoft	Spotfire	Visual Analytics	PowerBI
Preprocessing	ETL (Extract, Transform, Load)	Complex relational queries/functions							
	(Semi-) Automatic Data Preprocess.	ETL before data is loaded in-memory	By scripting Via visual query interfaces Via wizards						
		Predefined data Data transformations implementable by user (e.g. in JAVA, R) Table operations like joins							
		Imputation of values (e.g. normalisation of dimensions) Filtering by values Sub-sampling Outlier detection							

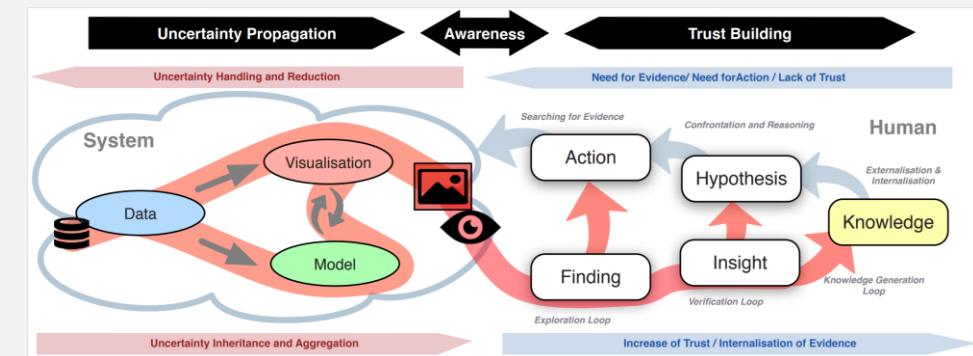


Data Handling and -Management



Developments & Research Topics

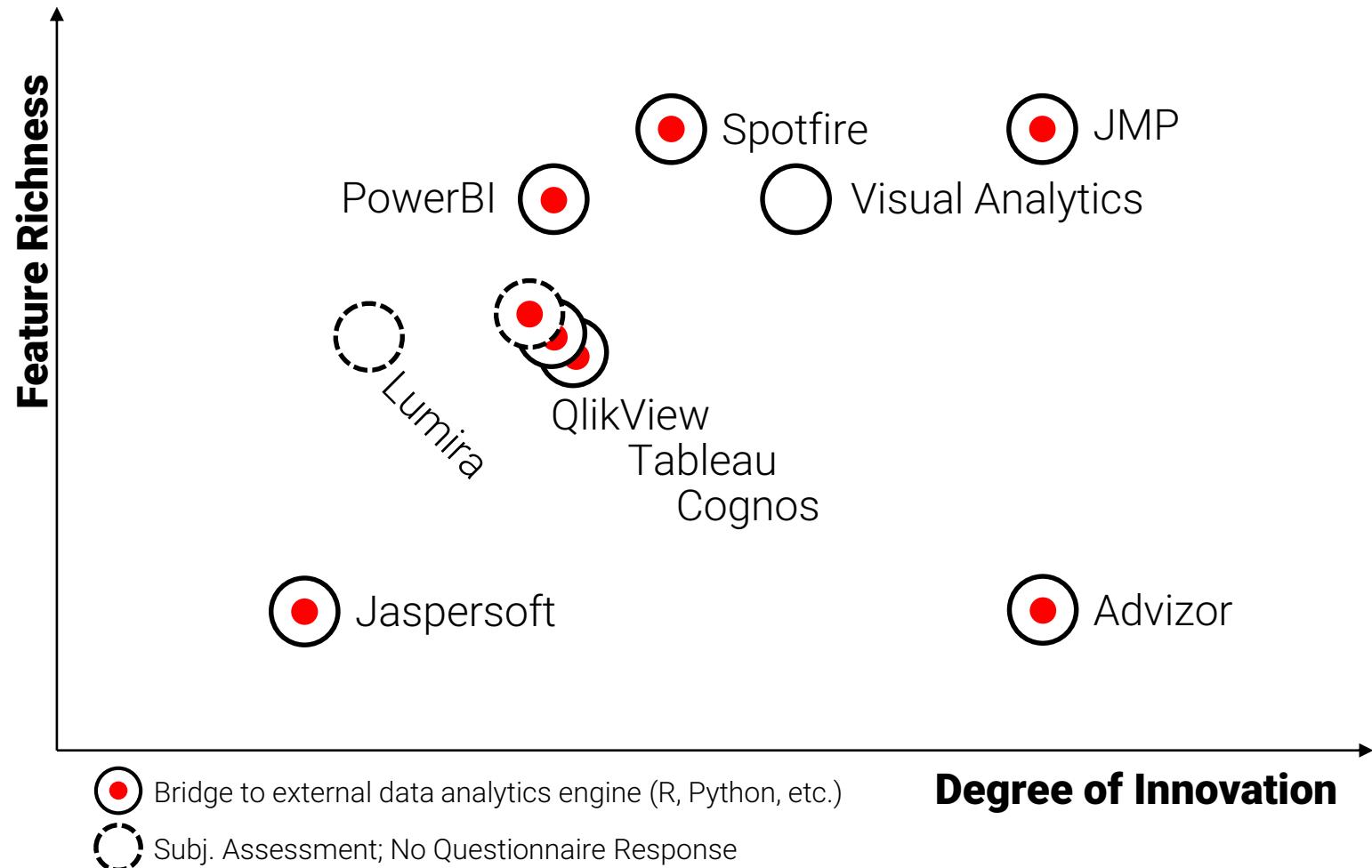
Data Uncertainty and Trust
Data (Pre-) Processing



D. Sacha, H. Senaratne, B. C. Kwon, G. P. Ellis, and D. A. Keim,
"The role of uncertainty, awareness, and trust in visual analytics"
IEEE Trans. Vis. Comput. Graph., vol. 22, no. 1, pp. 240–249, 2016.

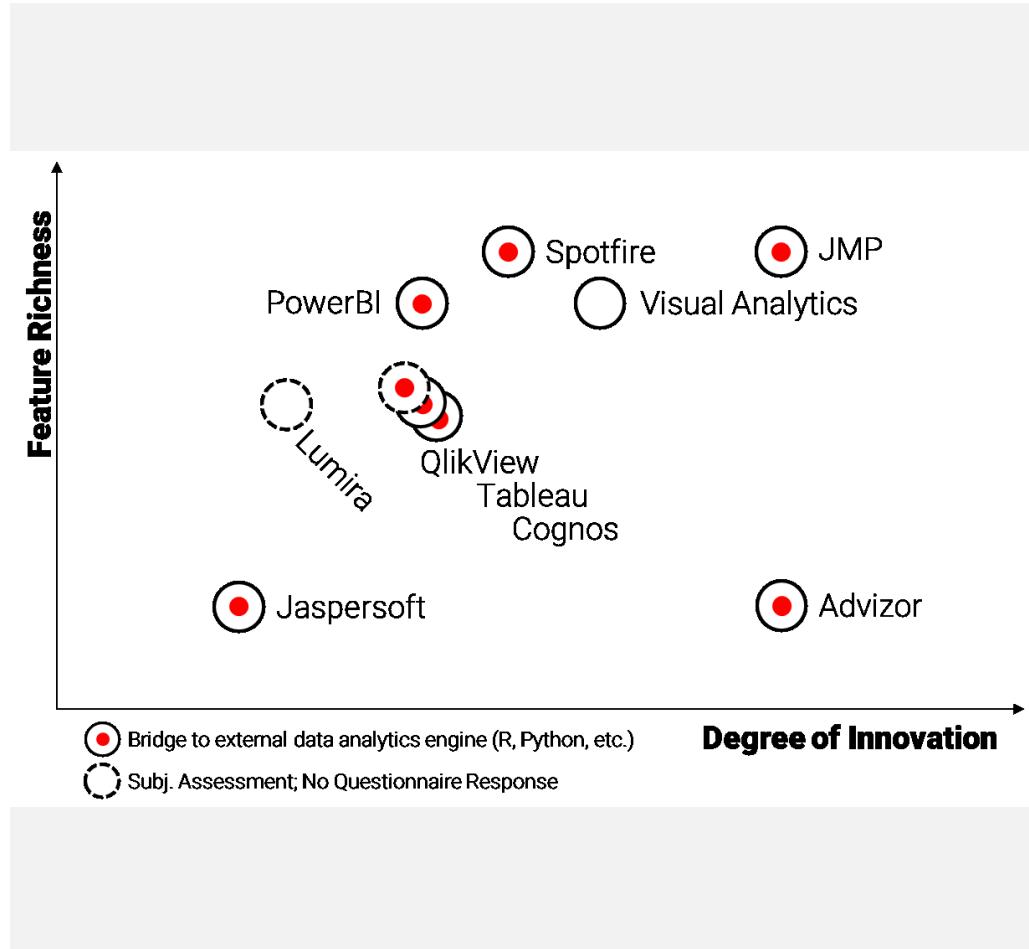


Automatic Analysis



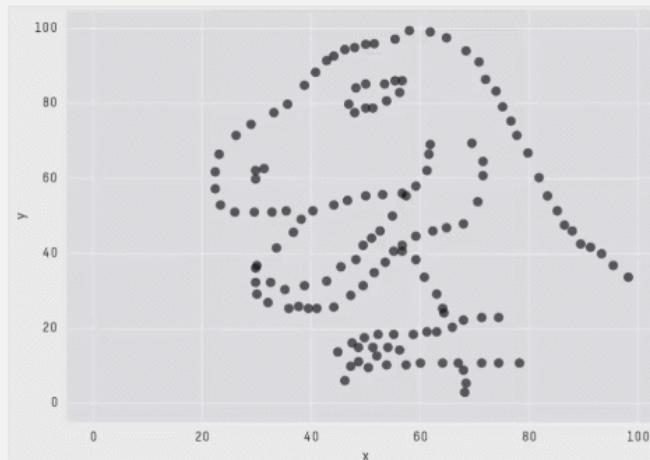


Automatic Analysis



Developments & Research Topics

Model and Parameter Space Exploration
High-Dimensional Data Analysis
Feature Encoding and Learning



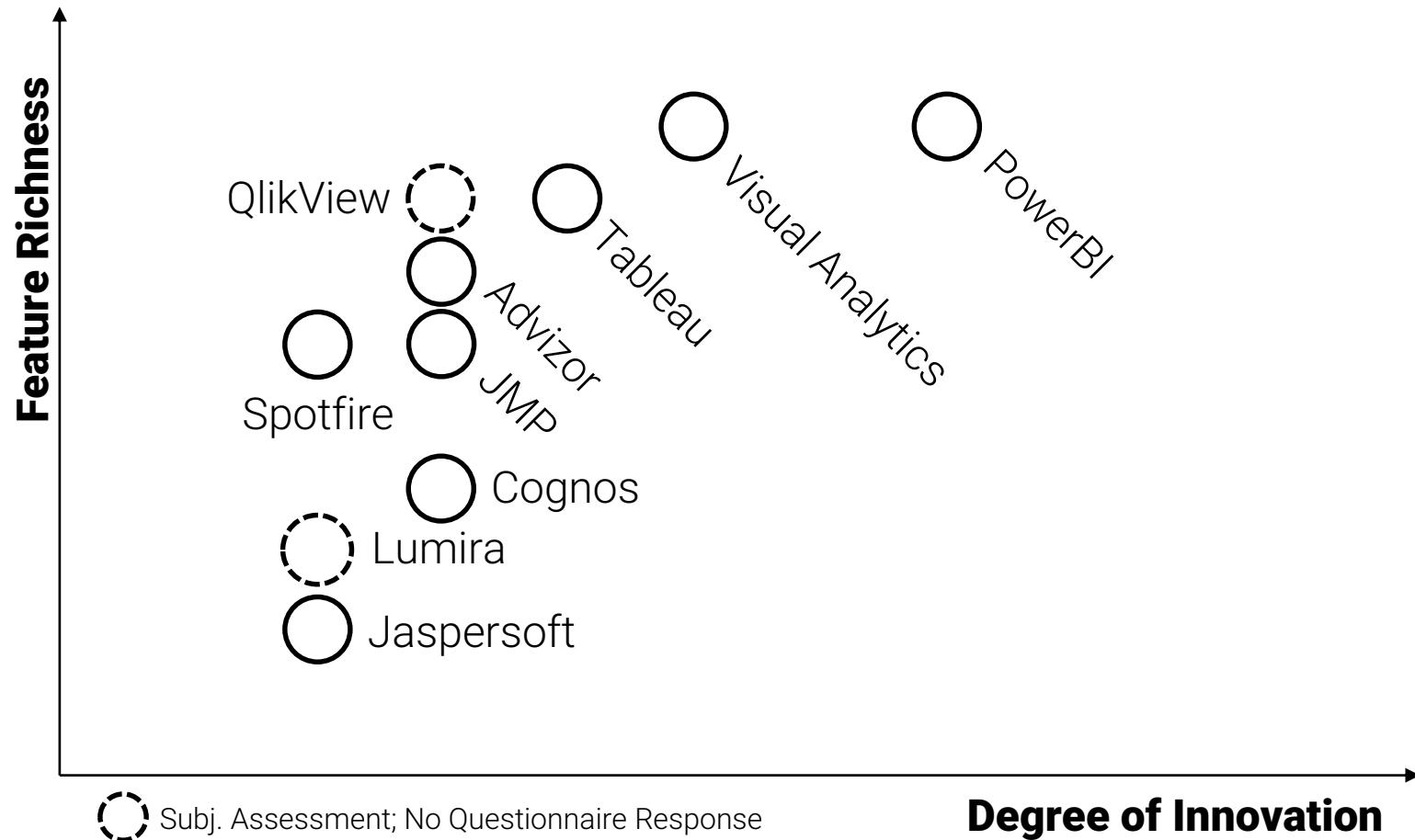
X Mean: 54.2659224
Y Mean: 47.8313999
X SD : 16.7649829
Y SD : 26.9342120
Corr. : -0.0642526



J. Matejka and G. Fitzmaurice, "**Same stats, different graphs: Generating datasets with varied appearance and identical statistics through simulated annealing.**"
in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.
ACM, 2017, pp. 1290–1294.

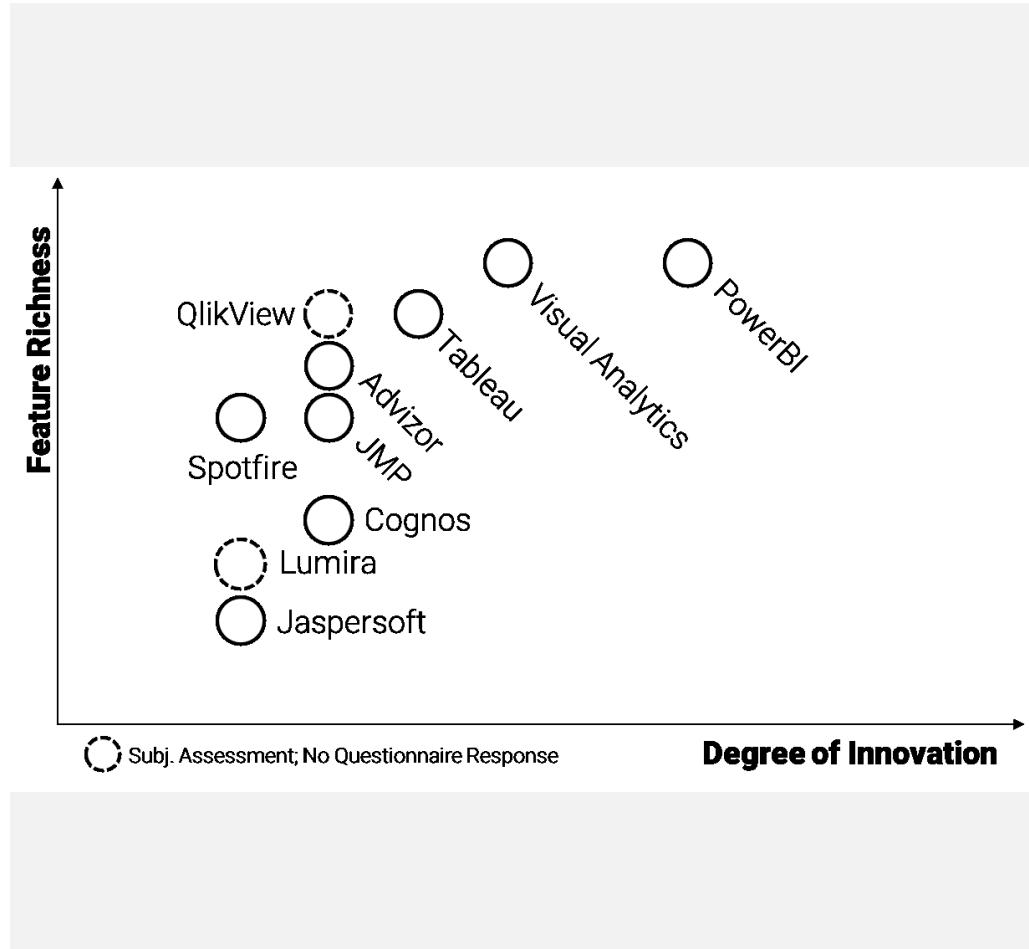


Complex Data Types



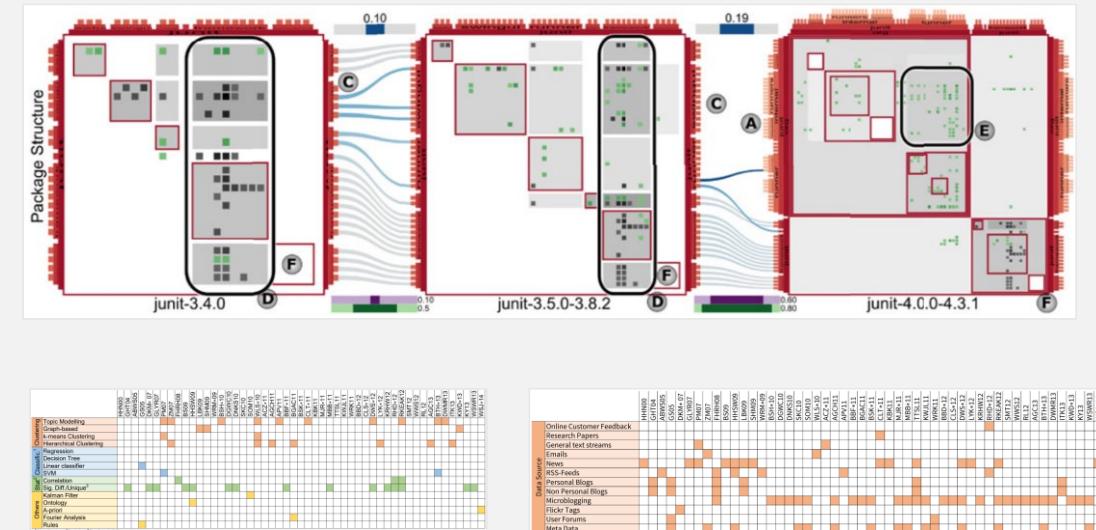


Complex Data Types



Developments & Research Topics

Complex Data Type Analysis and Streaming Data
Ranking & Quality Metrics



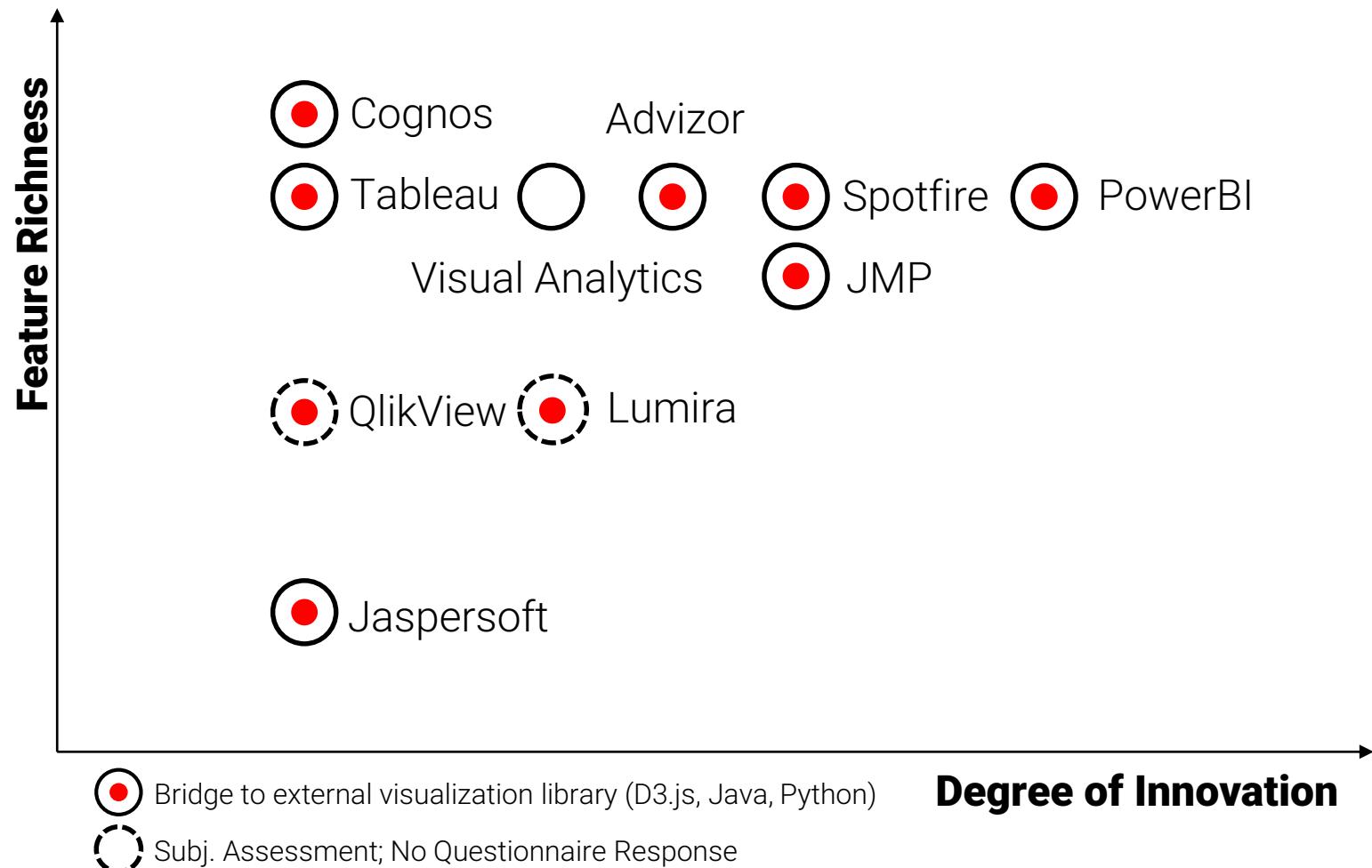
Wanner et. al [72]



Corinna Vehlow, Fabian Beck, and Daniel Weiskopf,
"Visualizing dynamic hierarchies in graph sequences,"
IEEE Trans. Vis. Comput. Graph., vol. 22, no. 10, pp. 2343–2357, 2016.

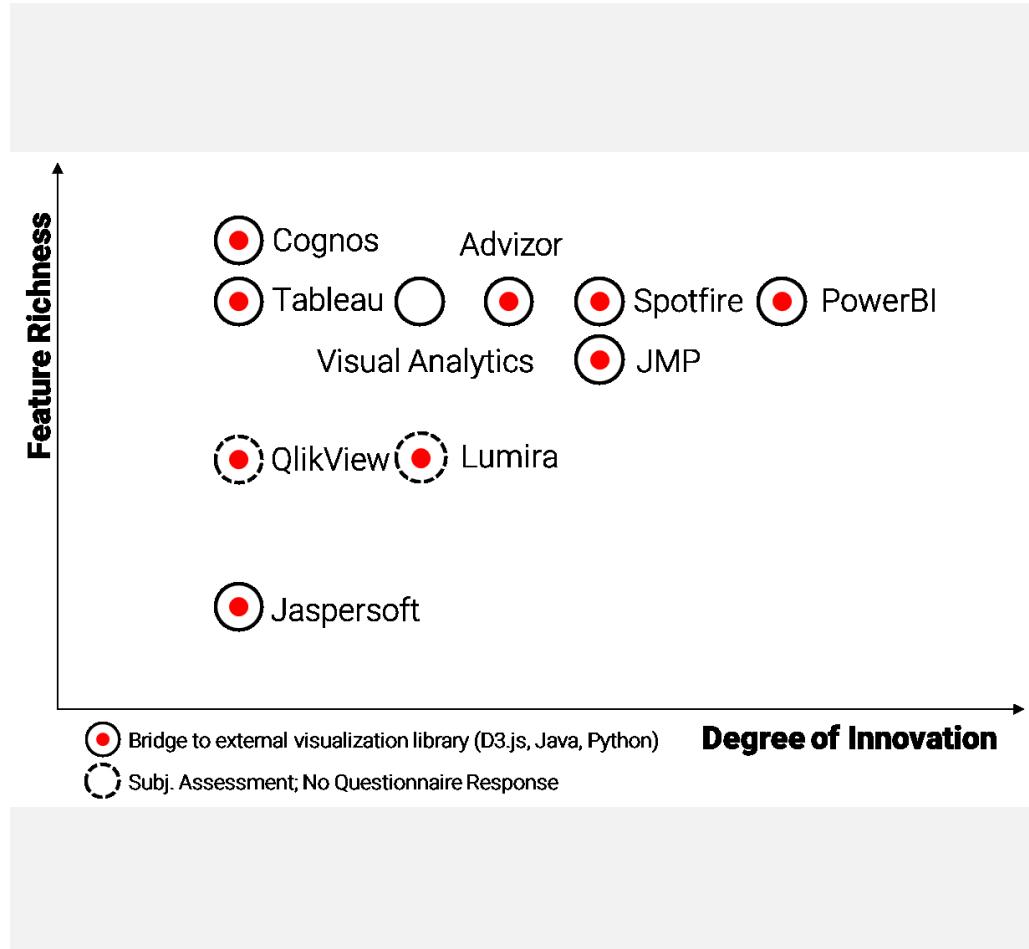


Visualization



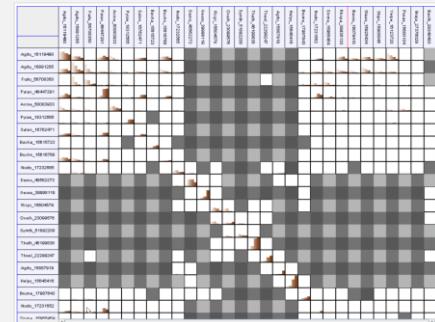


Visualization



Developments & Research Topics

Adaptive/Scalable Visualizations
Visualization Grammar
Interaction Design



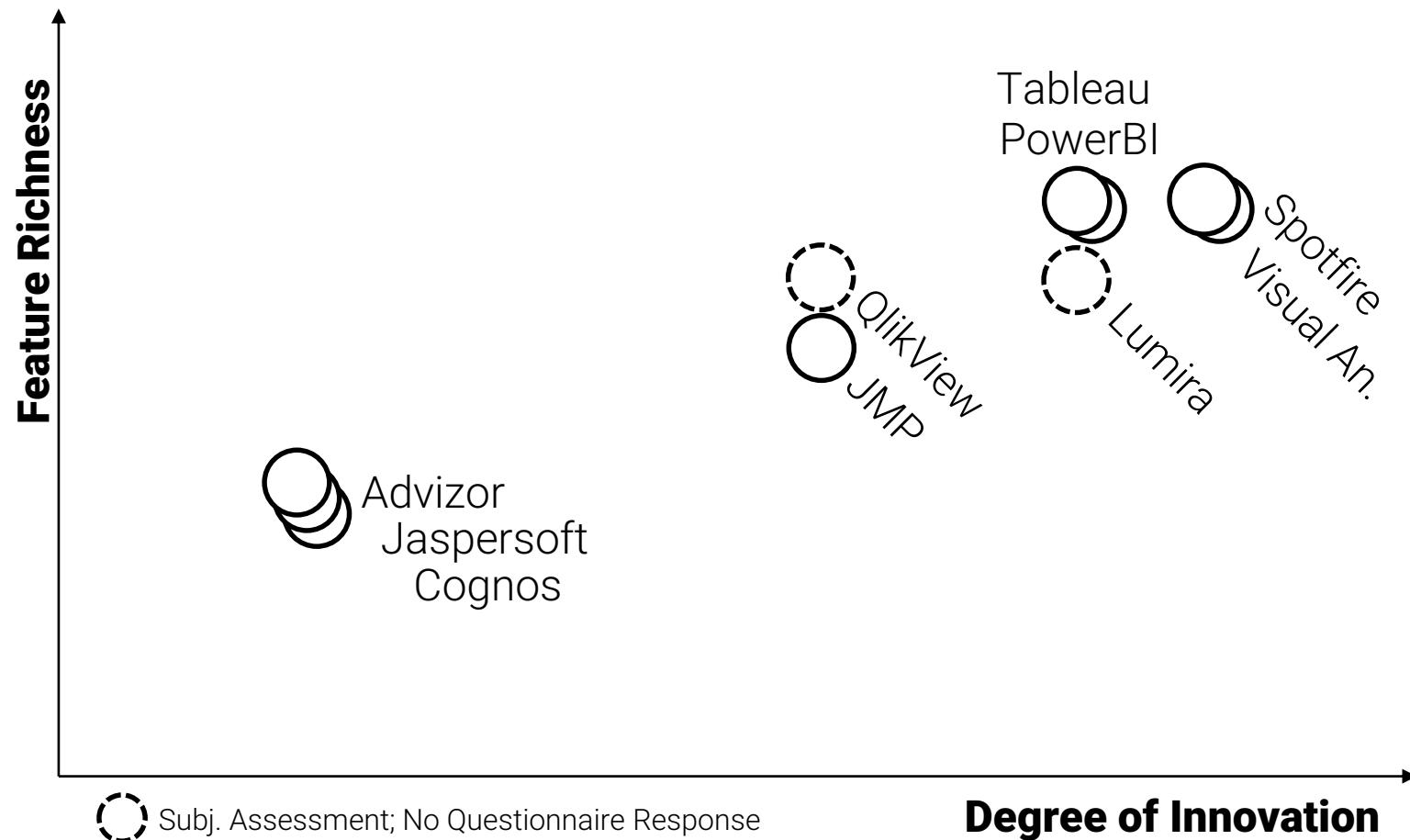
Elmqvist et. al [80]



F. Perteneder, E.-M. B. Grossauer, J. Leong, W. Stuerzlinger, and M. Haller,
"Glowworms and fireflies: Ambient light on large interactive surfaces,"
Conf. on Human Factors in Computing Systems, CHI '16. ACM, 2016, pp. 5849–5861.

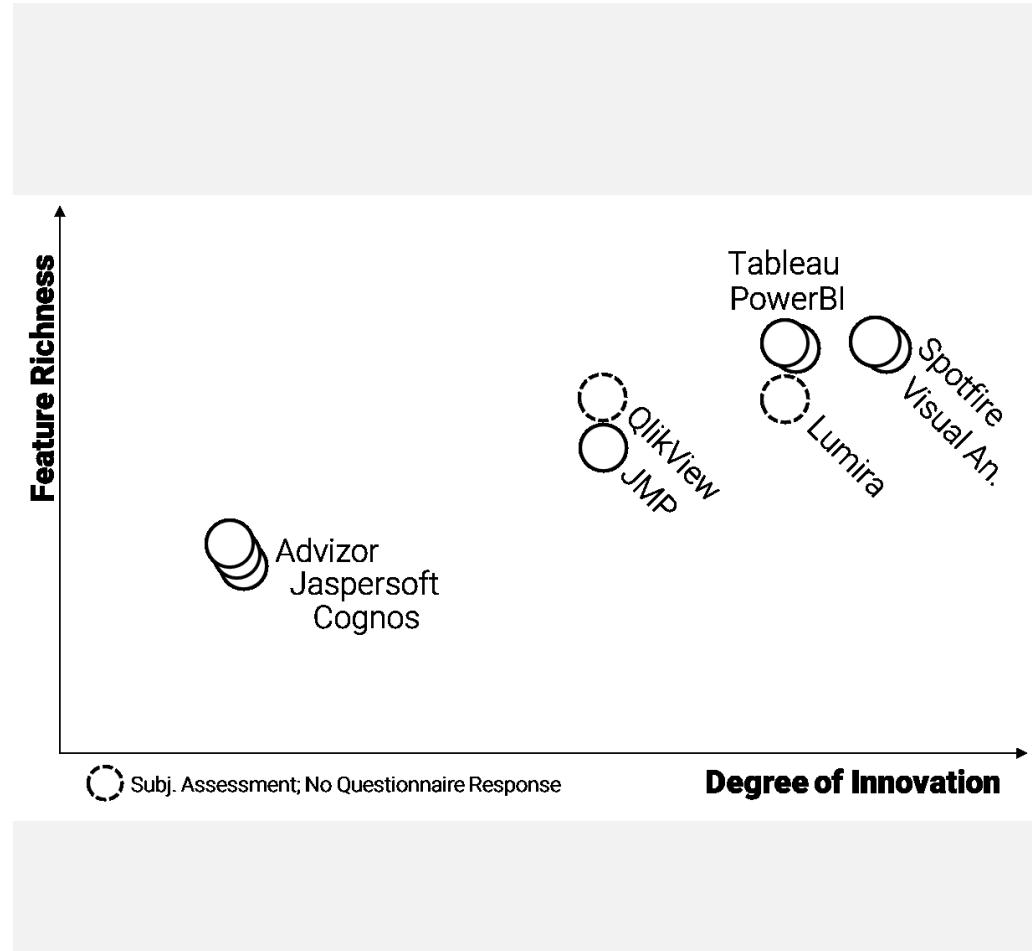


User Guidance, Perception, Cognition



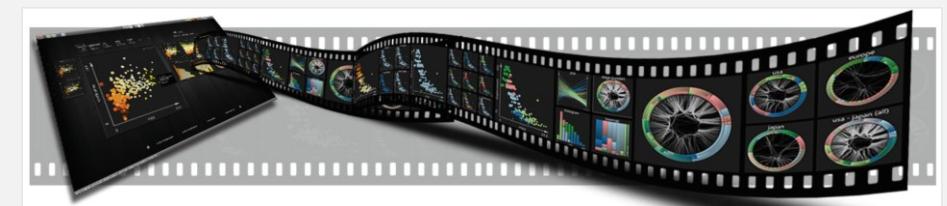
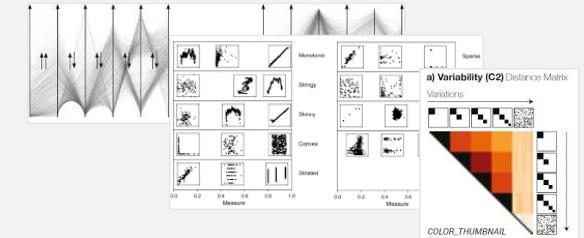
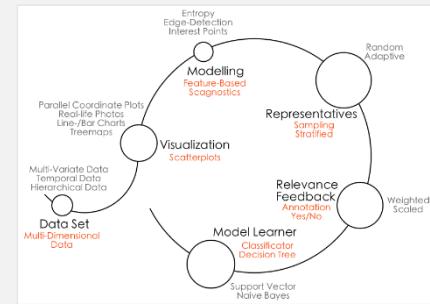


User Guidance, Perception, Cognition



Developments & Research Topics

Guided Exploration & Quality Assessment
Analytic Provenance



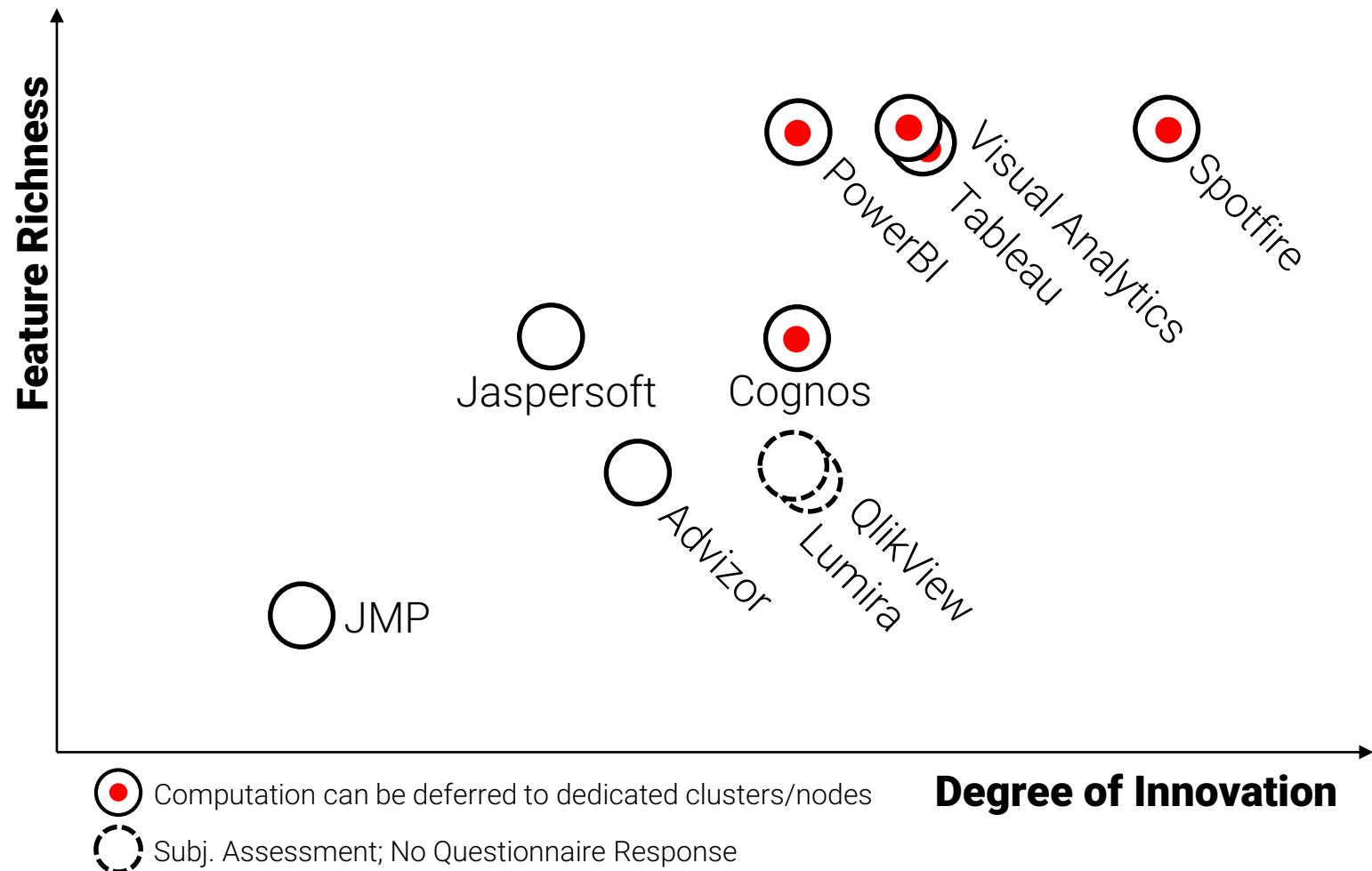
Stefan van den Elzen and J. J. van Wijk [84]



M. Behrisch, F. Korkmaz, L. Shao, and T. Schreck, “**Feedback-Driven Interactive Exploration of Large Multidimensional Data Supported by Visual Classifier,**” IEEE Conference on Visual Analytics Science and Technology. Oct. 2014, pp. 43–52..

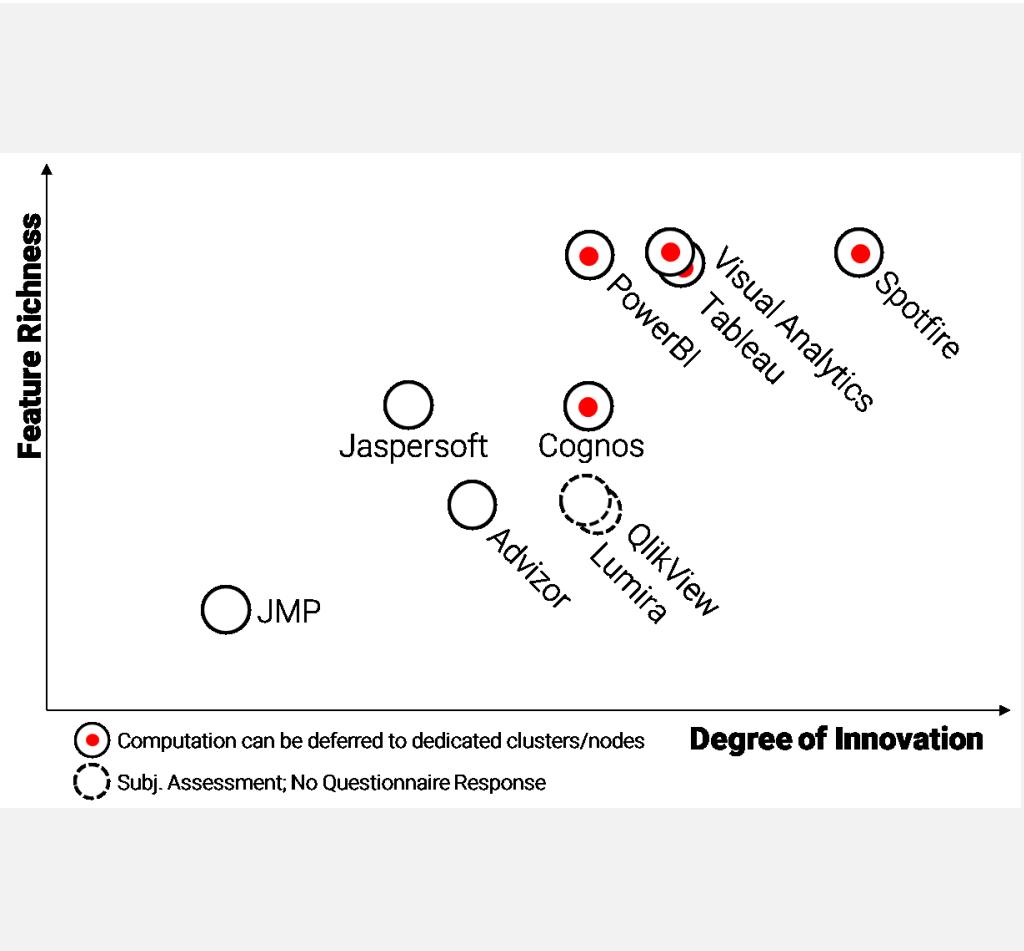


Infrastructure





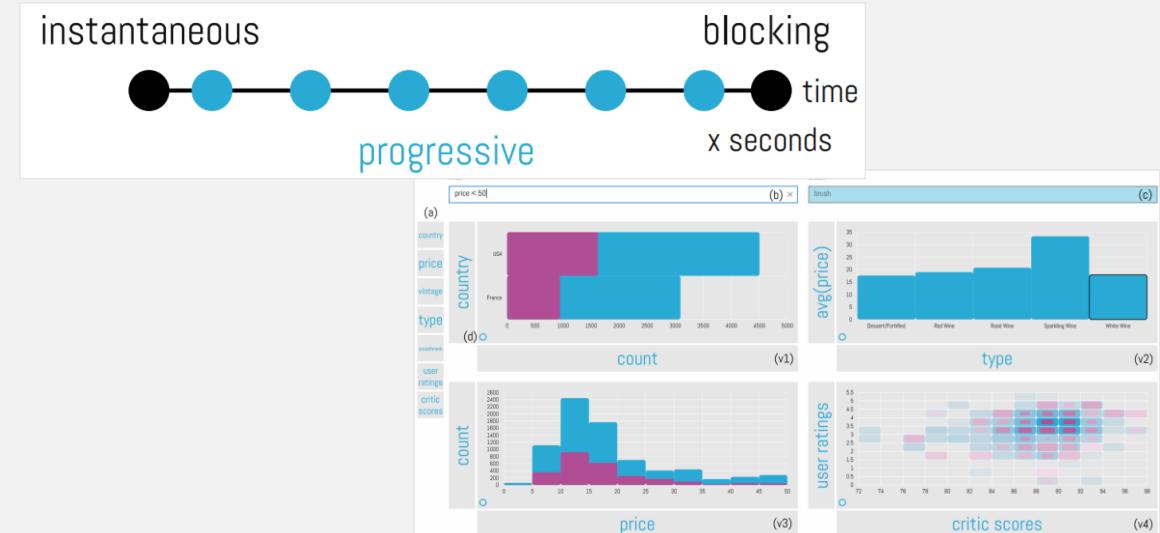
Infrastructure



Developments & Research Topics

Progressive Analytics

Dealing with larger datasets



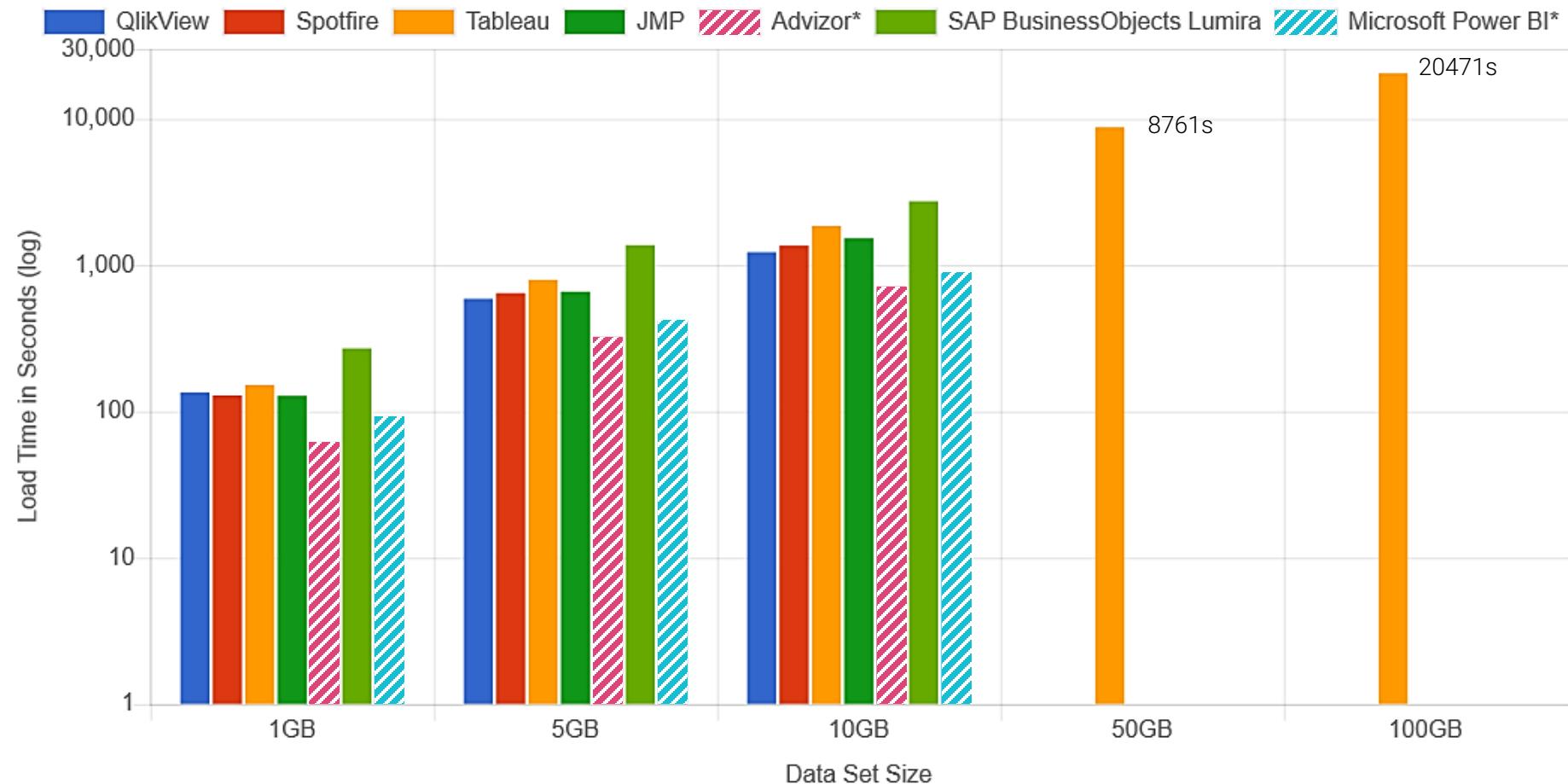
Emanuel Zgraggen, Alex Galakatos, Andrew Crotty, Jean-Daniel Fekete, Tim Kraska.
"How Progressive Visualizations Affect Exploratory Analysis."
 IEEE Transactions on Visualization and Computer Graphics, 2017



Performance Tests



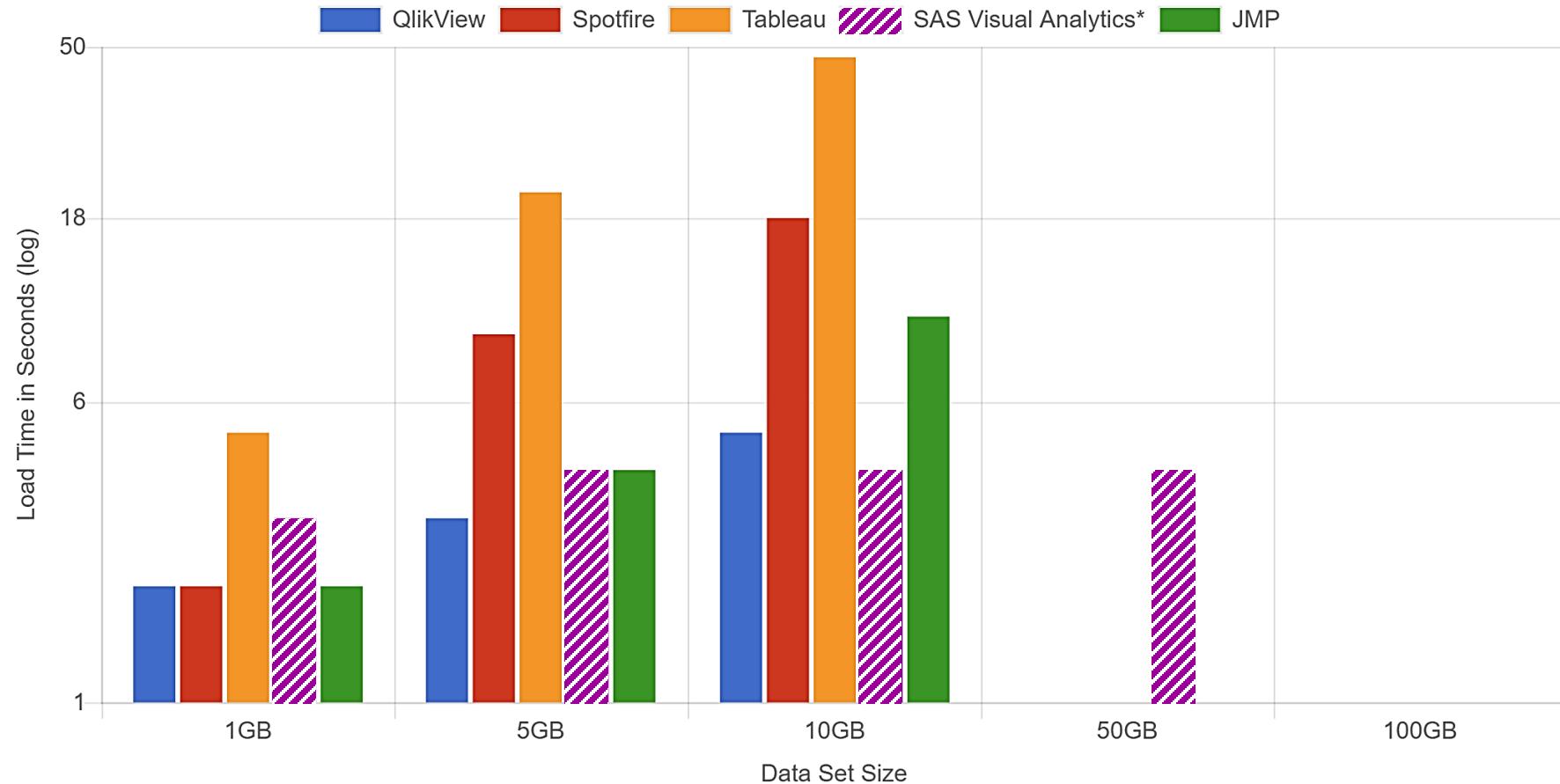
Performance Tests – Data Loading



Note: **SAS Visual Analytics is excluded**, as the test datasets have been imported by SAS for our tests.
SAS Visual Analytics has been hosted by SAS.



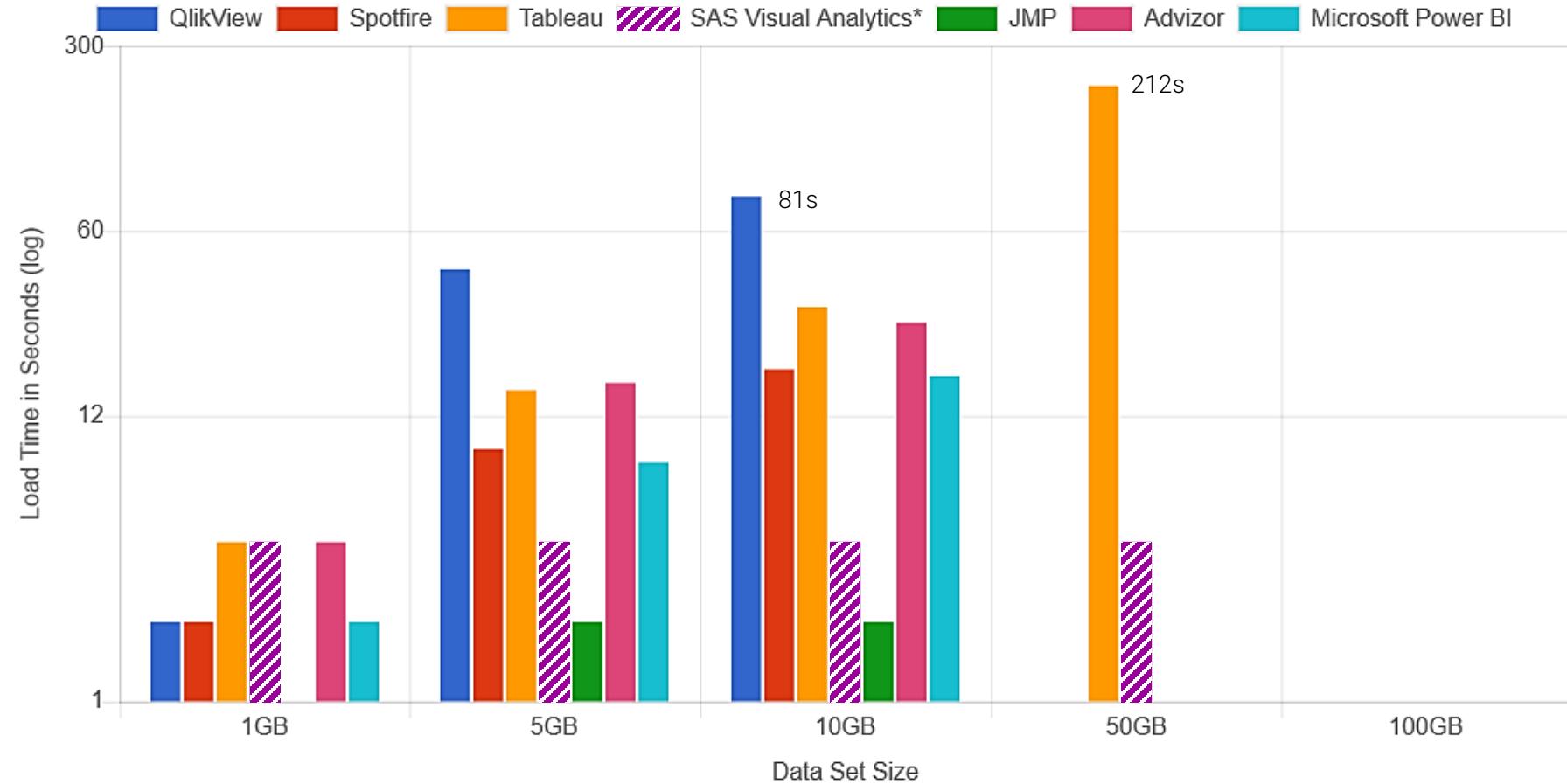
Performance Tests – Data Analysis



Note: **SAS Visual Analytics** has been hosted by SAS. The R bridge of **PowerBI** is limited to 150,000 rows
Advizor is missing analysis features required for this test and is therefore excluded.



Performance Tests – Data Visualization



Note: a larger data set (100GB) has been tested, but no system could visualize it.
SAS Visual Analytics has been hosted by SAS.



Case Study Evaluation



Use Case Study – Design

Data:

VAST Challenge 2015

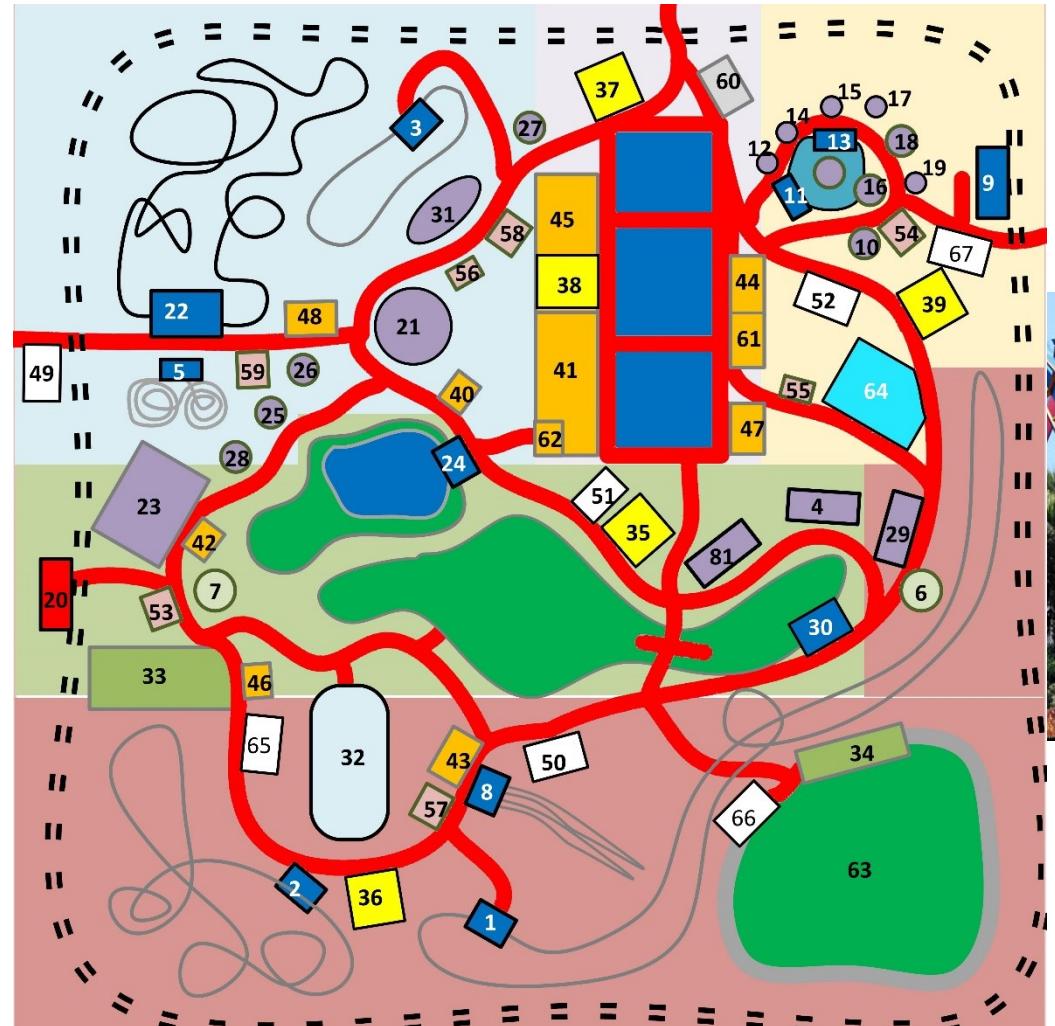
Three days of visitor movement and communication
in a fictitious amusement park
(3,000,000 rows, 1.3GB)

Strings, dates, timestamps and geo-coordinates

Tasks:

resemble findings for different visitor groups
patterns in time and space

find patterns of attractions not operating
mass-movement data analysis



Source: VAST Challenge 2015



Use Case Study – Data

Two data sets with movement data (right), and meta-data about individual communication.

First data set:

1.1 GB, almost 2 M rows of data

Second data set:

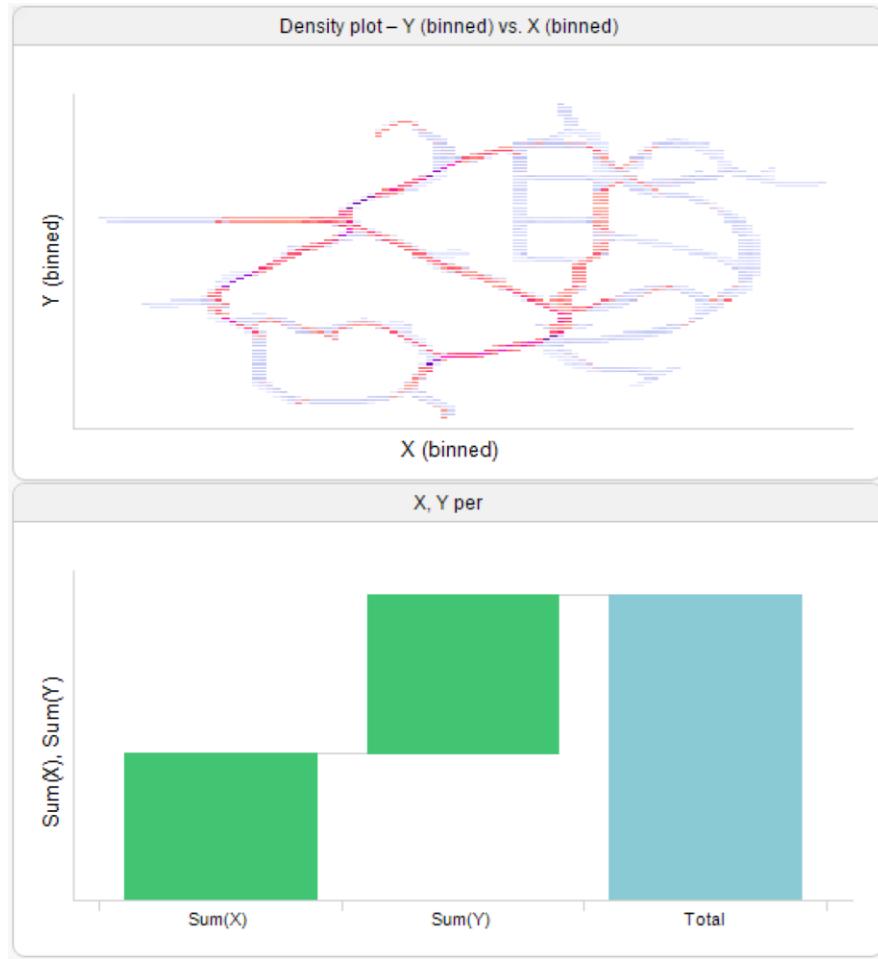
200 MB, 2.6 M rows of data

Format: CSV, plain text, split in different files

```
1 Timestamp,id,type,X,Y
2 2014-6-06 08:00:16,1591741,check-in,63,99
3 2014-6-06 08:00:16,825652,check-in,63,99
4 2014-6-06 08:00:19,179386,check-in,63,99
5 2014-6-06 08:00:19,531348,check-in,63,99
6 2014-6-06 08:00:31,1483004,check-in,0,67
7 2014-6-06 08:00:31,849521,check-in,0,67
8 2014-6-06 08:00:38,398396,check-in,99,77
9 2014-6-06 08:00:38,675561,check-in,99,77
10 2014-6-06 08:00:40,2007598,check-in,0,67
11 2014-6-06 08:00:40,455752,check-in,0,67
12 2014-6-06 08:00:40,1881823,check-in,0,67
13 2014-6-06 08:00:40,374745,check-in,0,67
14 2014-6-06 08:00:40,422300,check-in,0,67
15 2014-6-06 08:00:40,73216,check-in,0,67
16 2014-6-06 08:00:40,1109689,check-in,0,67
17 2014-6-06 08:00:40,1007202,check-in,0,67
18 2014-6-06 08:00:49,1948150,check-in,63,99
19 2014-6-06 08:00:49,996258,check-in,63,99
20 2014-6-06 08:00:53,950113,check-in,0,67
```



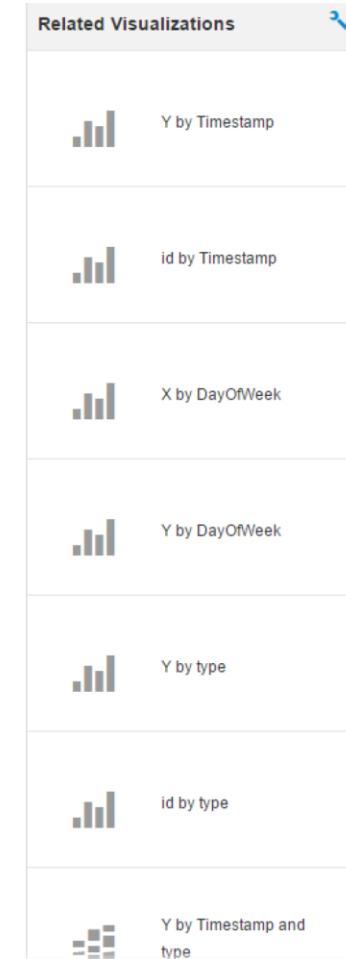
Use Case Study Example: Spotfire



Spotfire



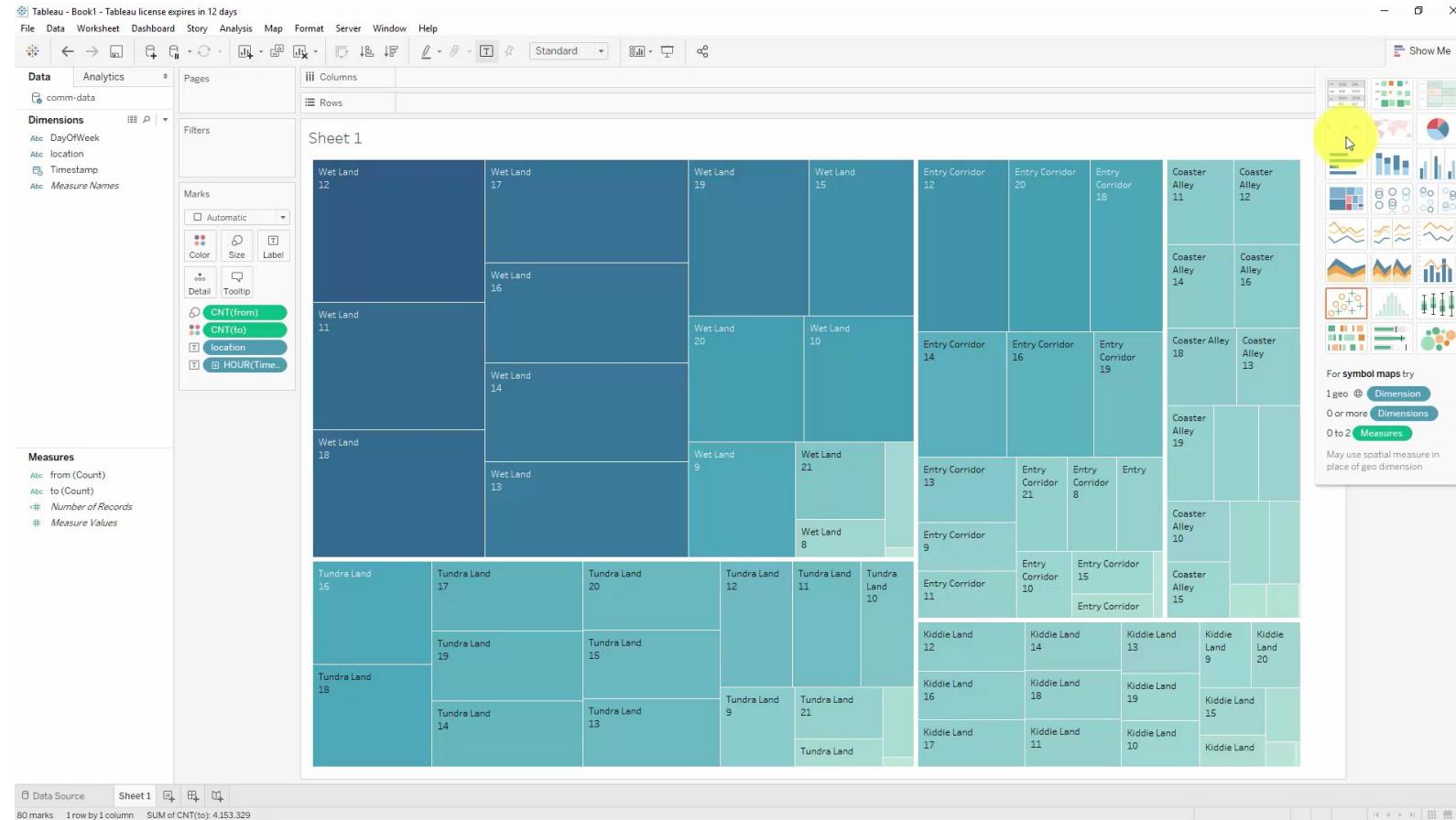
Tableau



Lumira

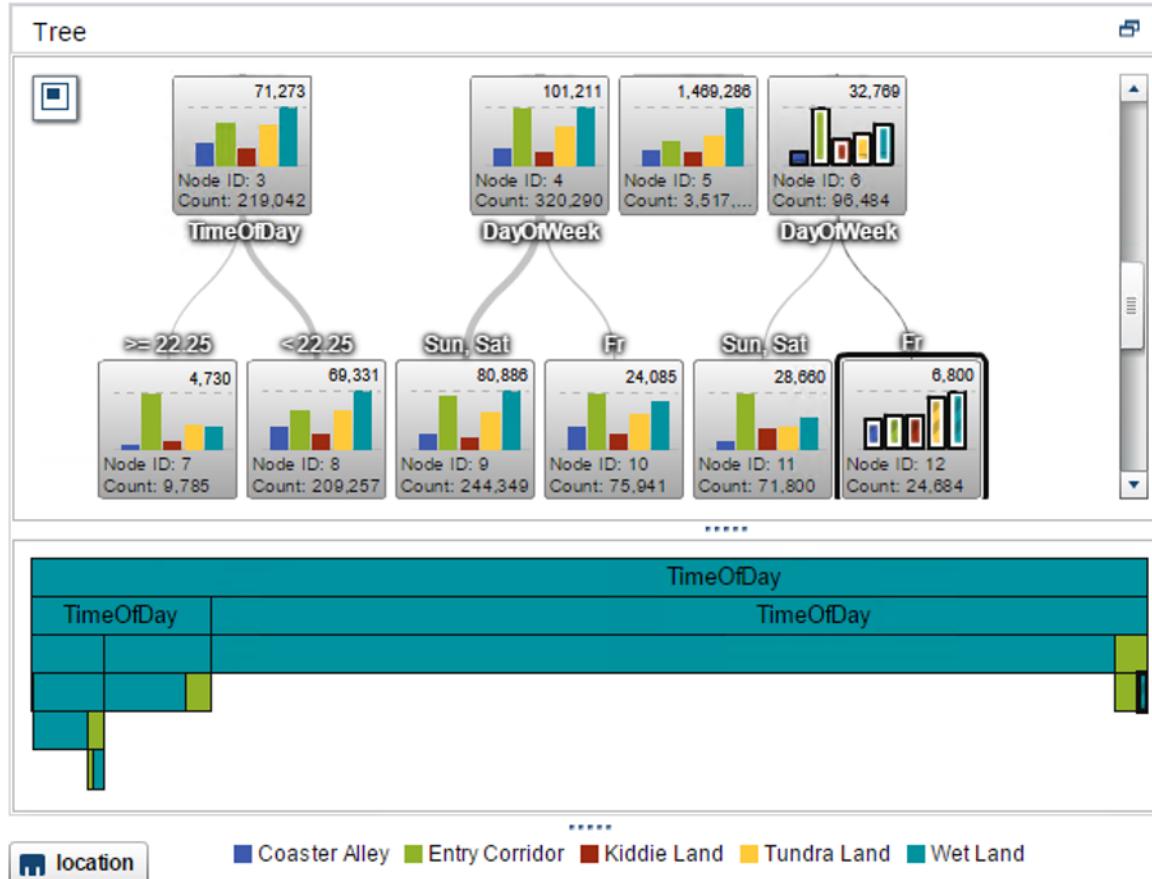


Use Case Study Example: Tableau

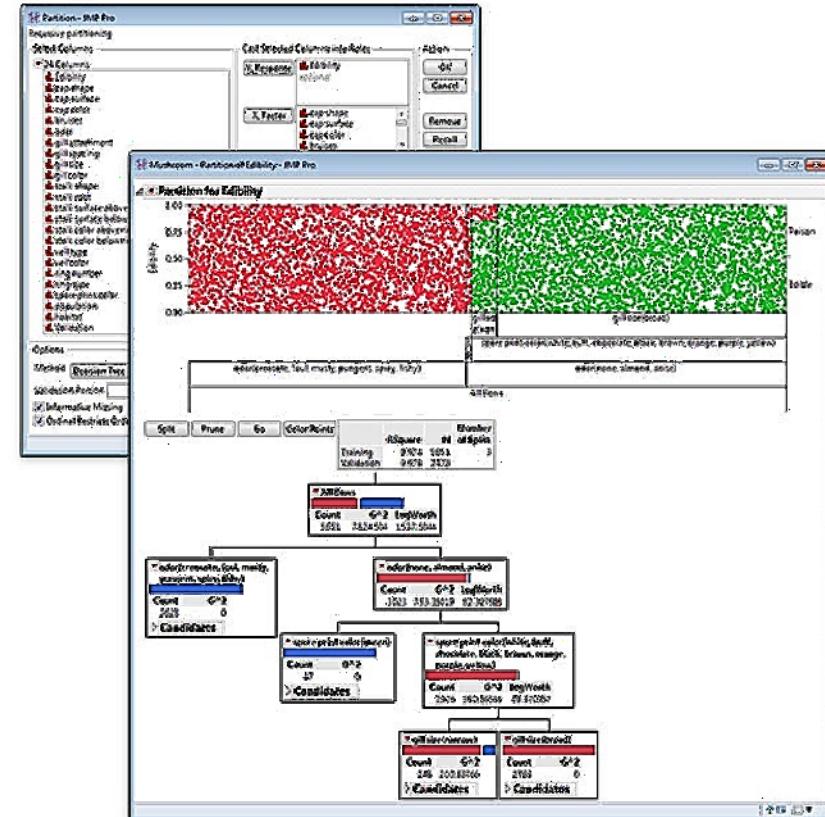




Use Case Study Example: SAS Visual Analytics

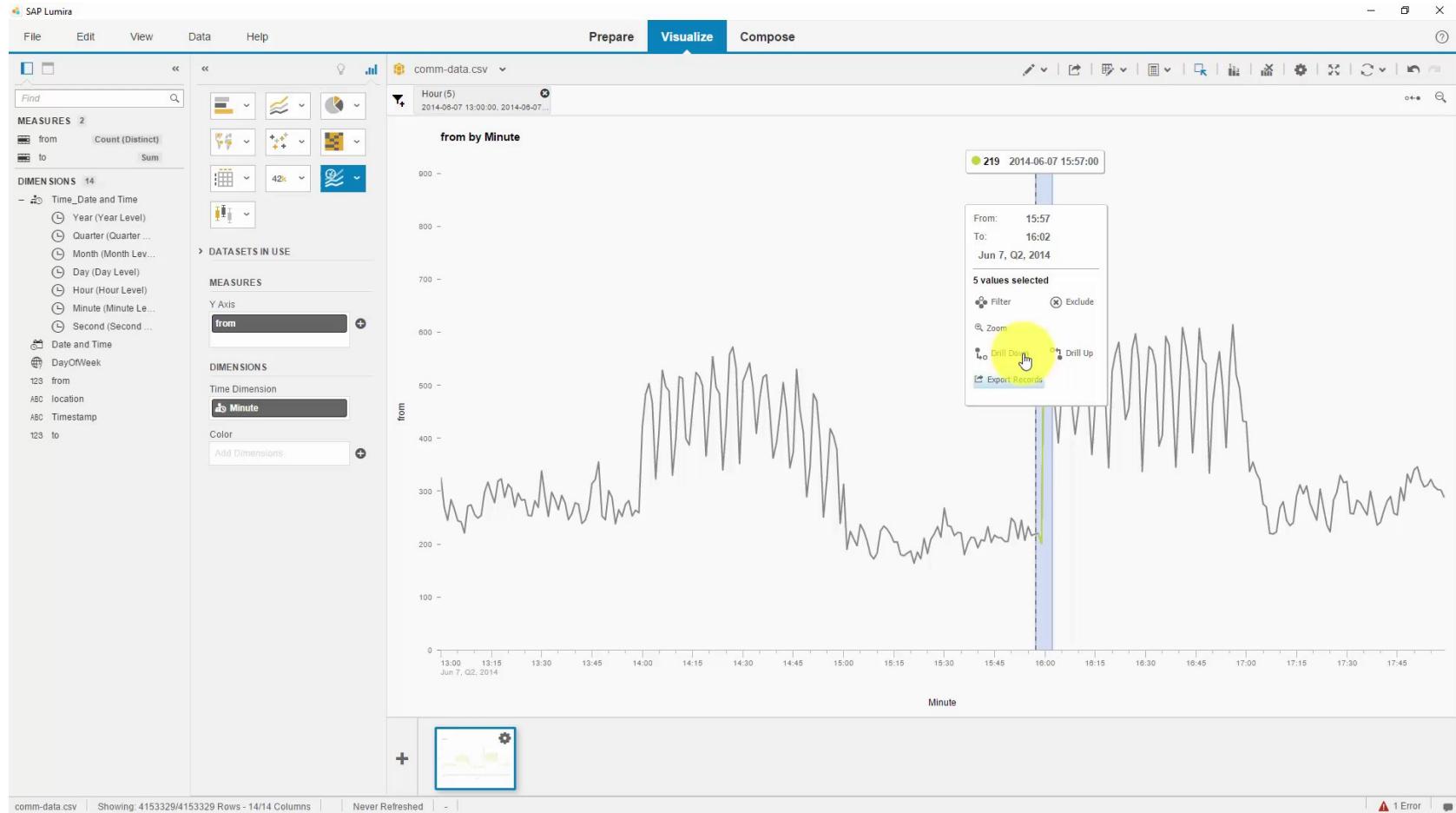


Visual Analytics



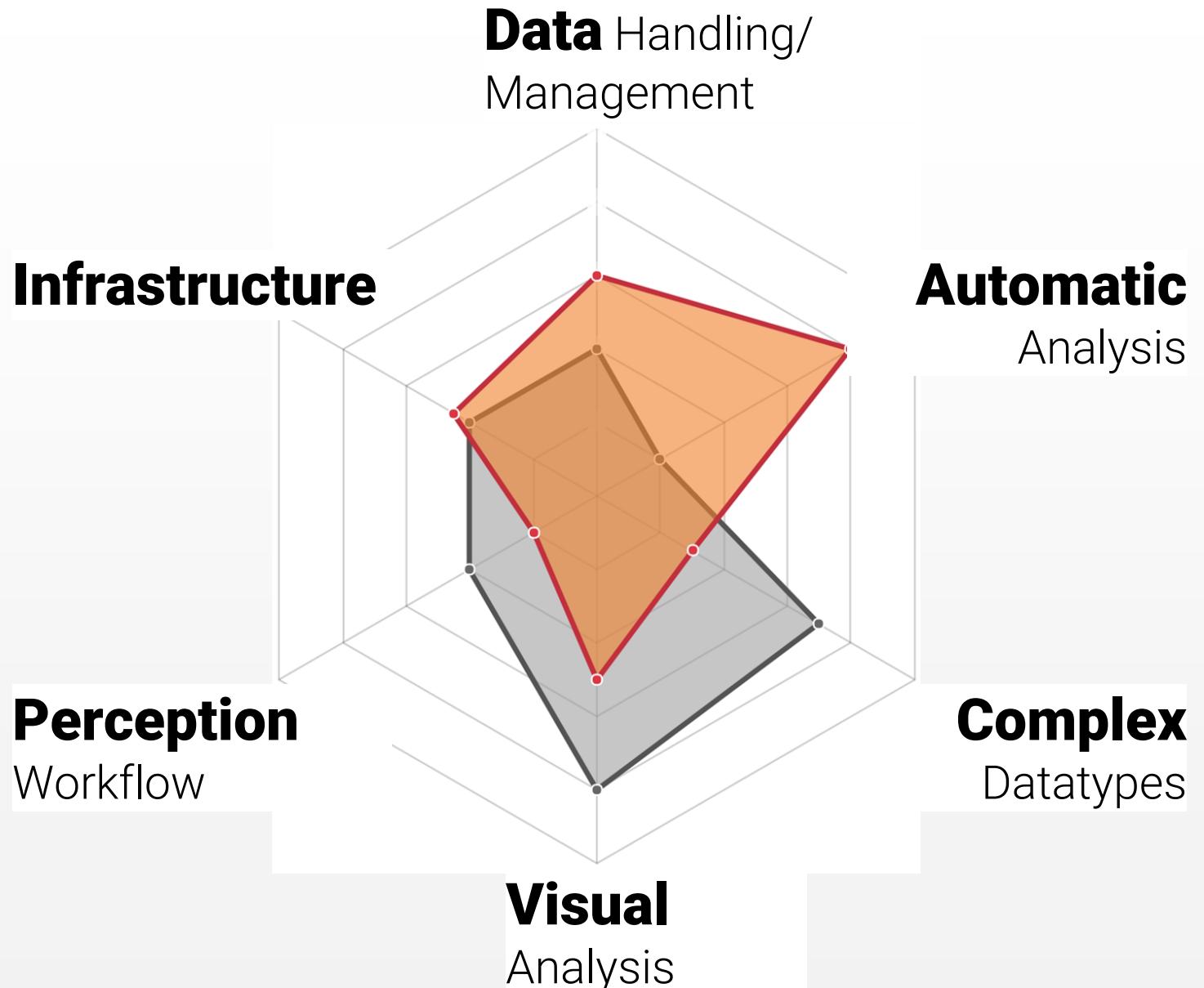


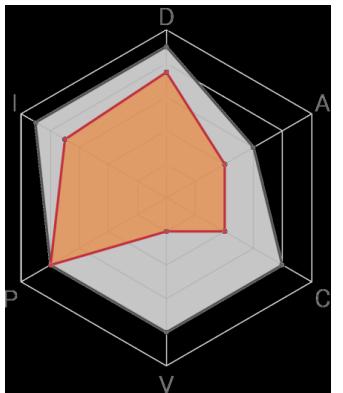
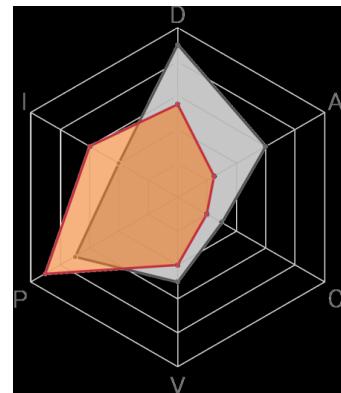
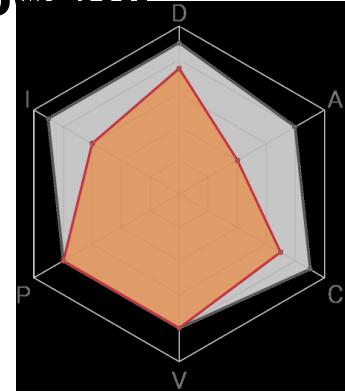
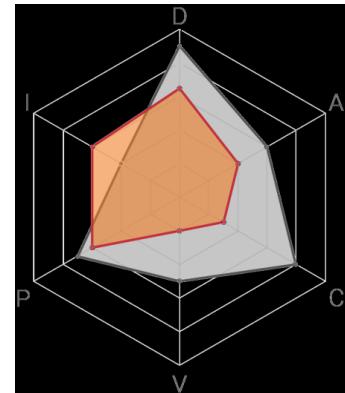
Use Case Study Example: Lumira

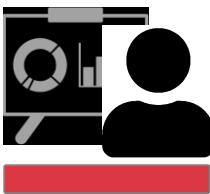
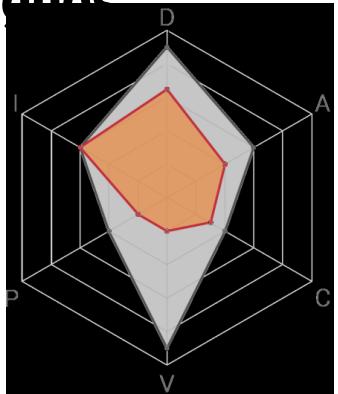
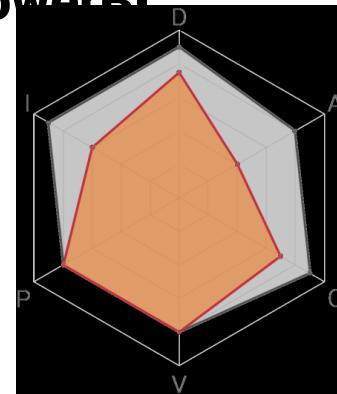
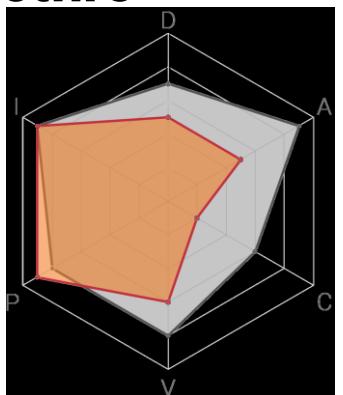
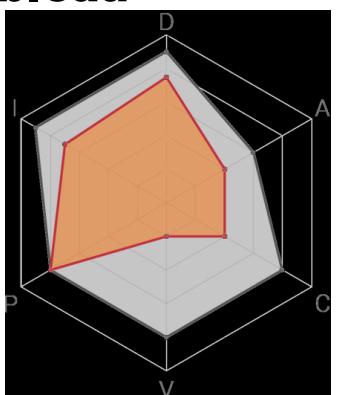
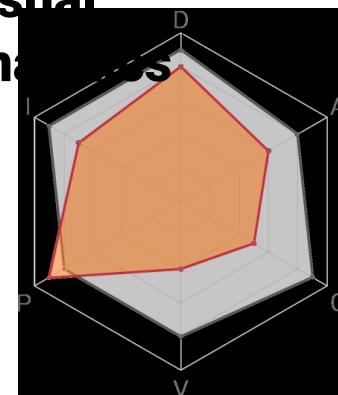
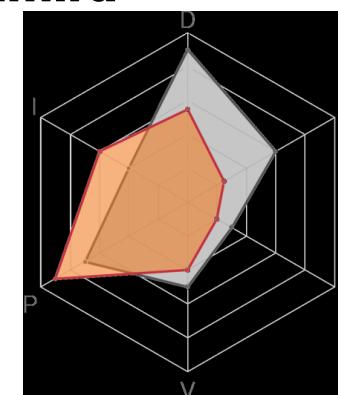
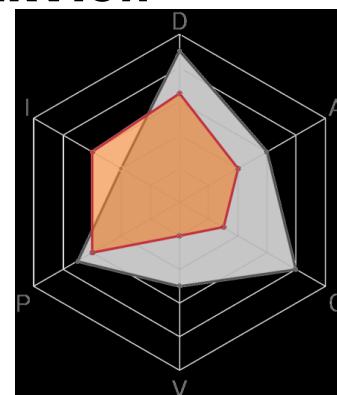


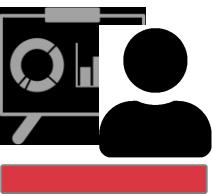
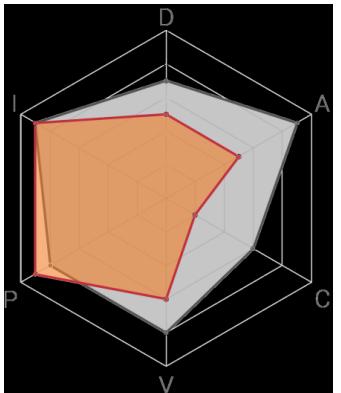
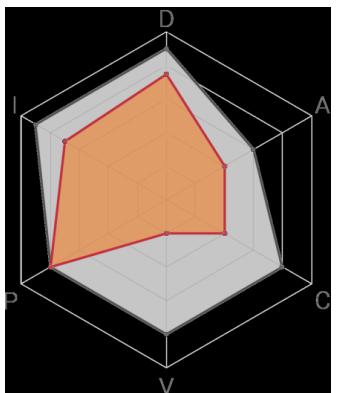
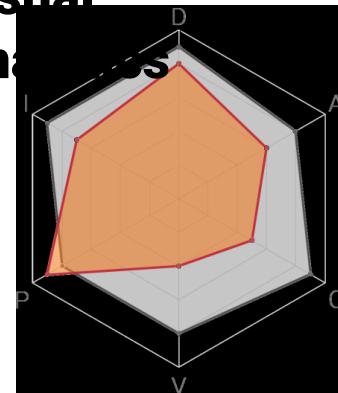
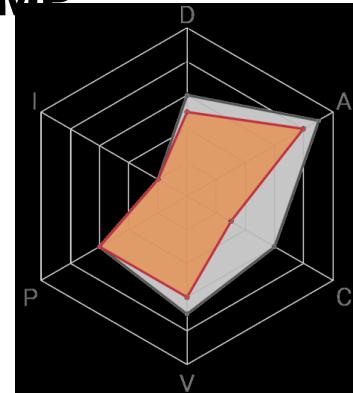
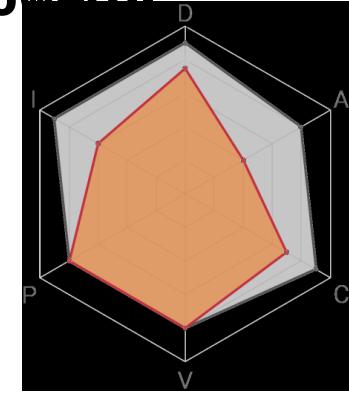
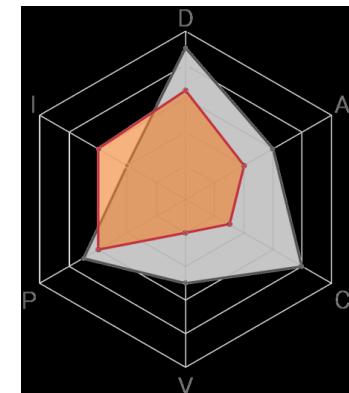


Conclusion



**Tableau****Lumira*****PowerBI****QlikView***

**Cognos****PowerBI****Spotfire****Tableau****Visual
Analytics****Lumira*****QlikView***

**Spotfire****Tableau****Visual
Analytics****JMP****PowerBI****QlikView***



Limitations and Conclusion

- 1 Survey based on vendor feedback
- 2 Evaluation of base systems without extensions
- 3 Performance difficult to compare across architectures
- 4 Use cases with known outcomes
- 5 Practical demands difficult to model generically



Take Home Messages

System choice depends on task and users.

Large variety is offered.

Strong Development on Extendibility Side.

From "scripting bridges" to "Add-in stores".

Research their way into products.

Companies are contributing & benefitting of the field.





Thank YOU



HARVARD
John A. Paulson
School of Engineering
and Applied Sciences

SIEMENS
Ingenuity for life





Backup



Spearman Rank Correlation

$$\rho = \frac{S_{xy}}{S_x S_y} = \frac{\frac{1}{n} \sum_{i=1}^n (R(x_i) - \overline{R(x)}) \cdot (R(y_i) - \overline{R(y)})}{\sqrt{\left(\frac{1}{n} \sum_{i=1}^n (R(x_i) - \overline{R(x)})^2 \right) \cdot \left(\frac{1}{n} \sum_{i=1}^n (R(y_i) - \overline{R(y)})^2 \right)}}$$

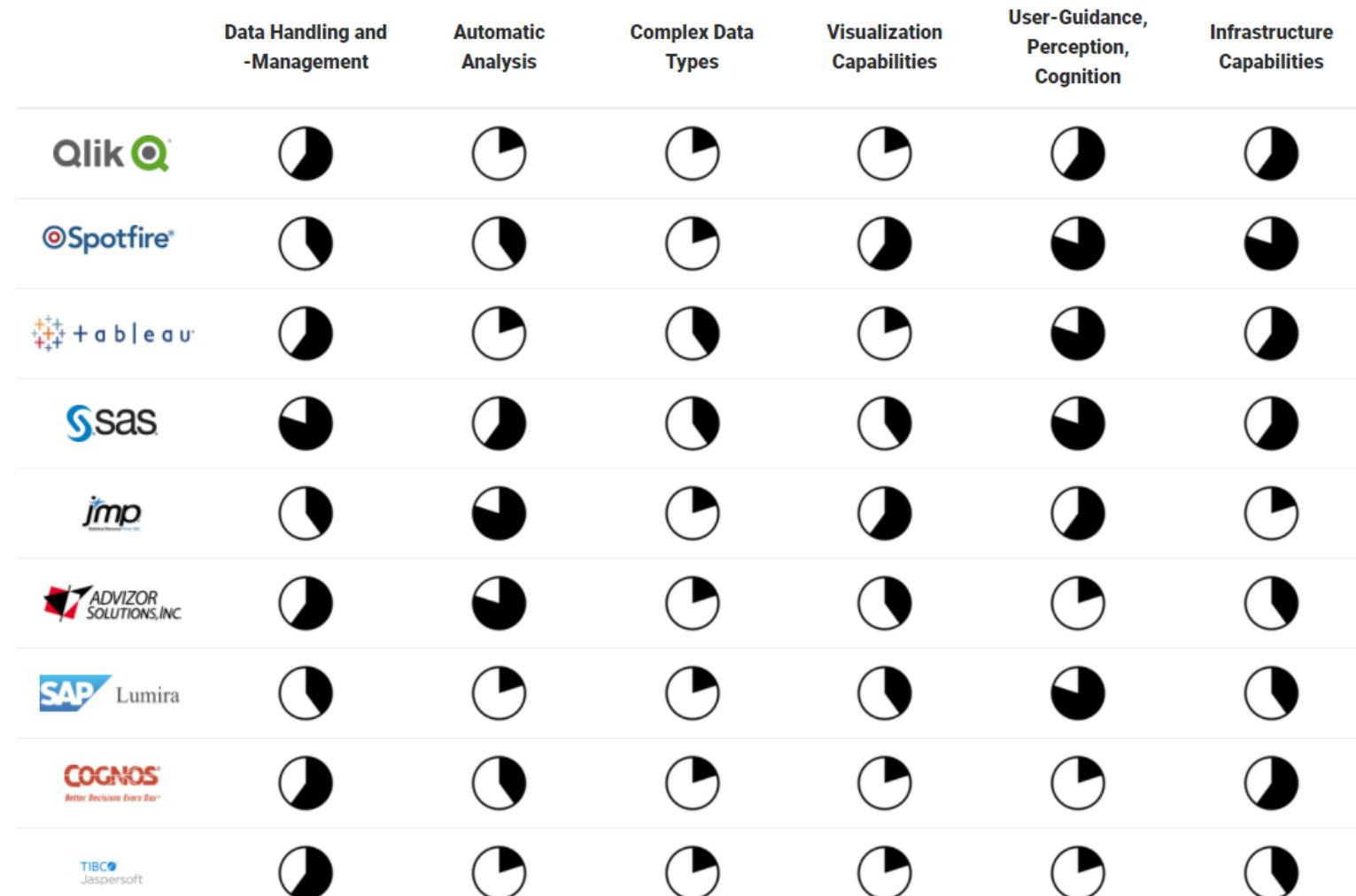


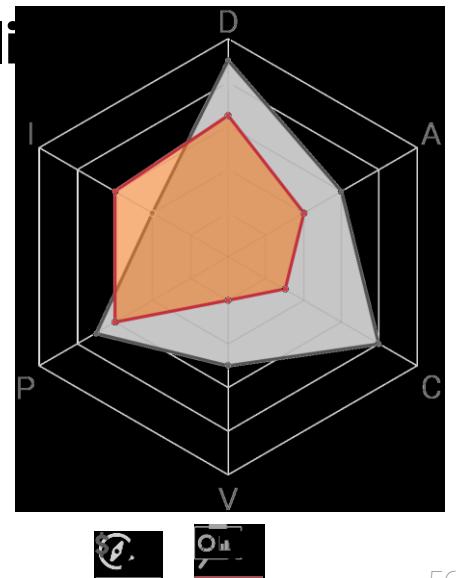
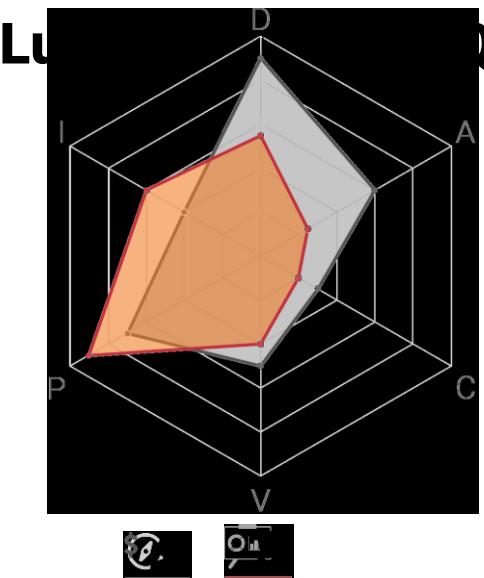
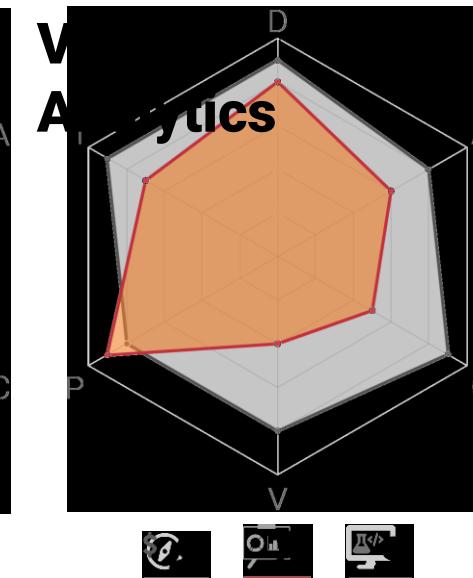
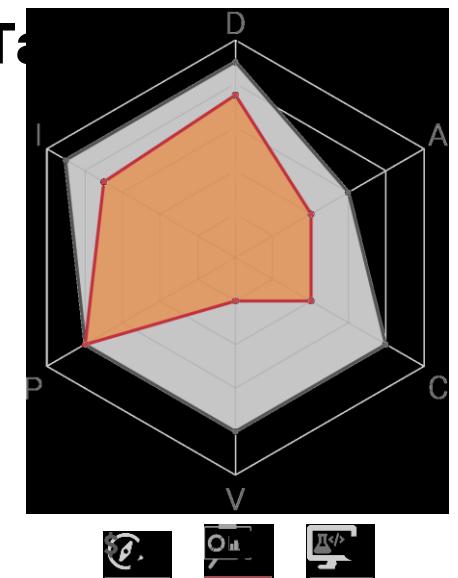
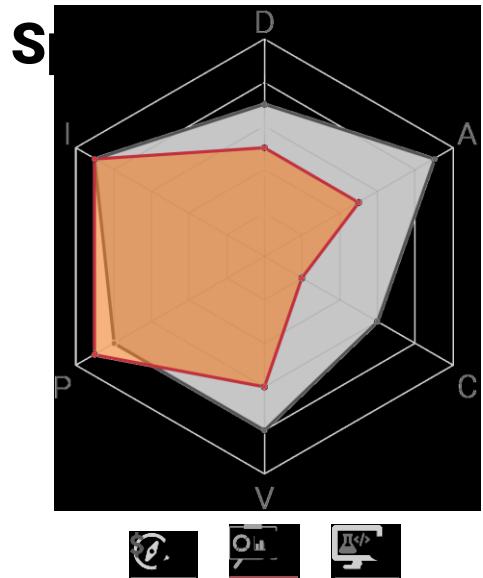
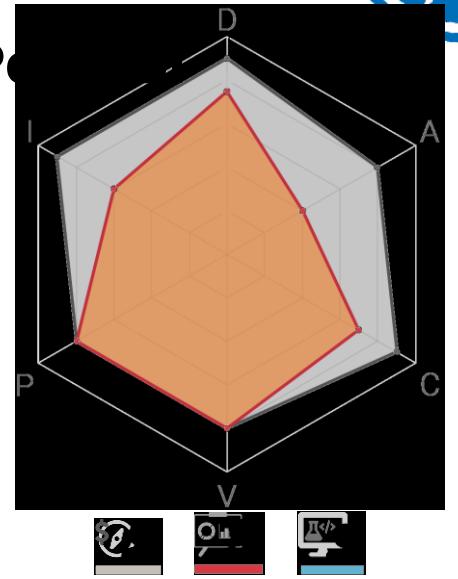
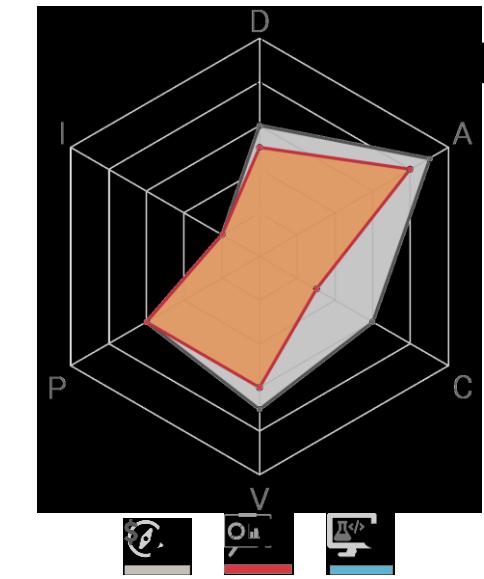
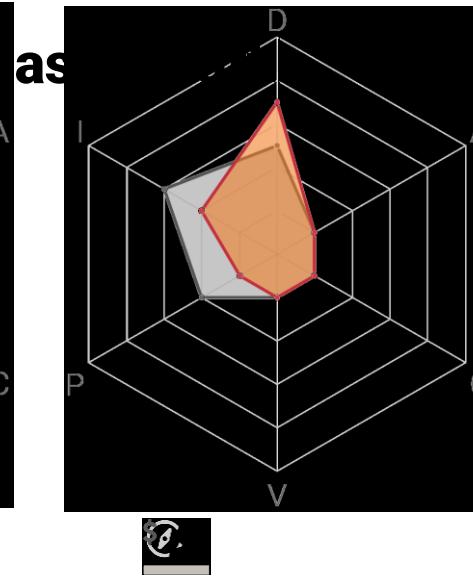
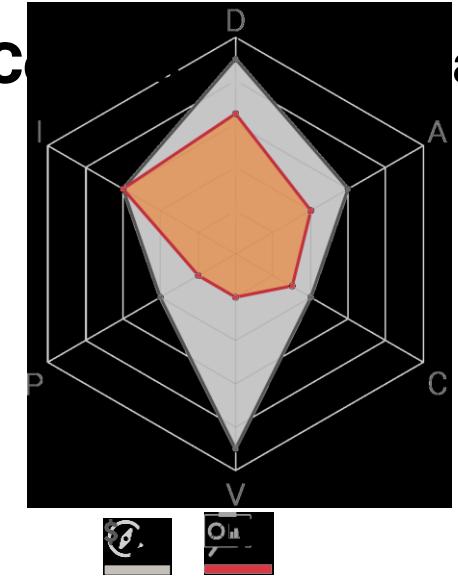
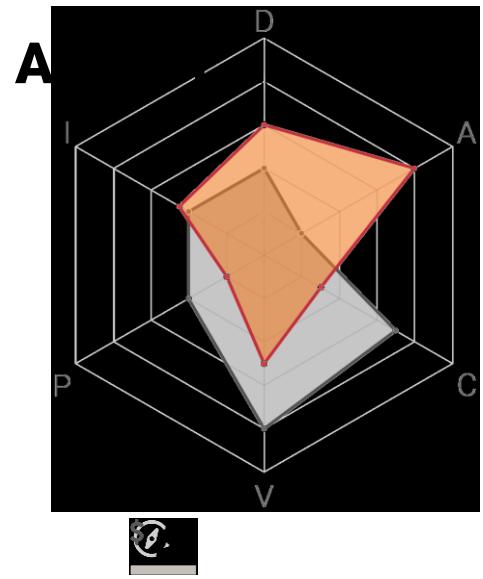
Feature Richness Comparison

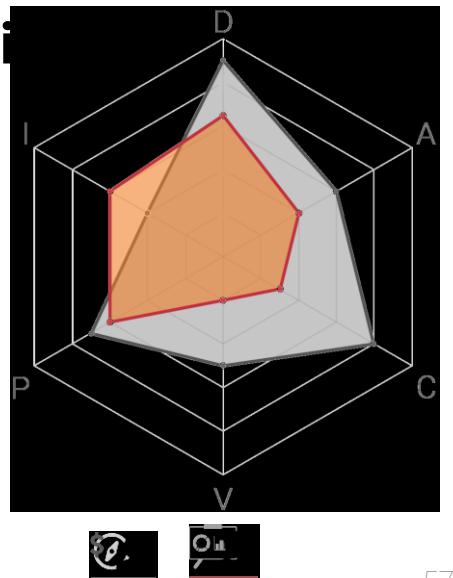
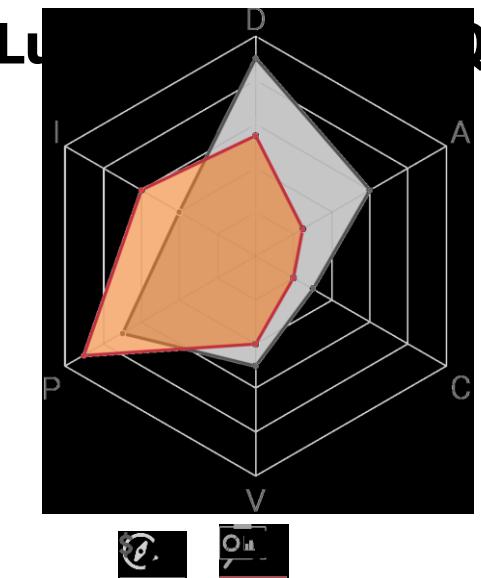
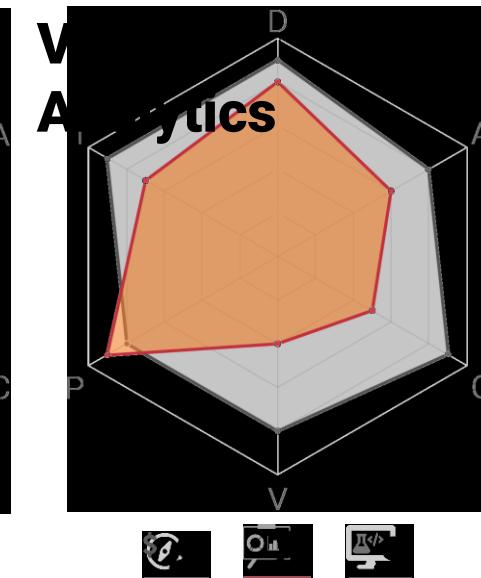
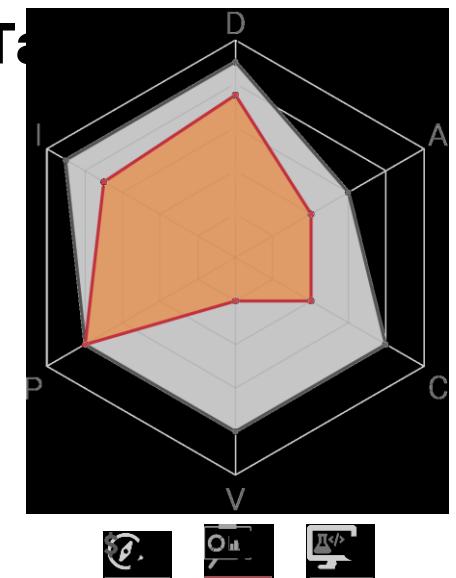
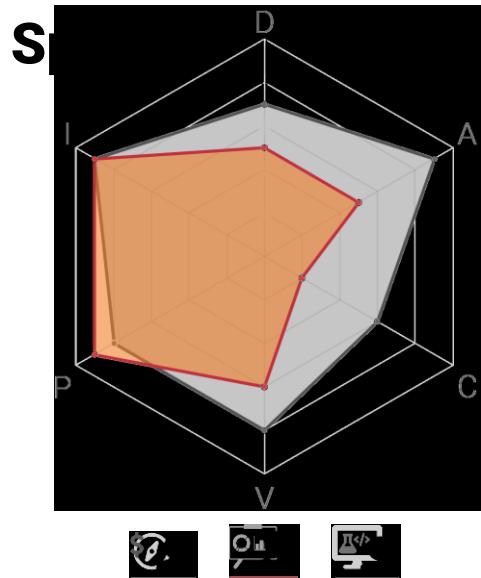
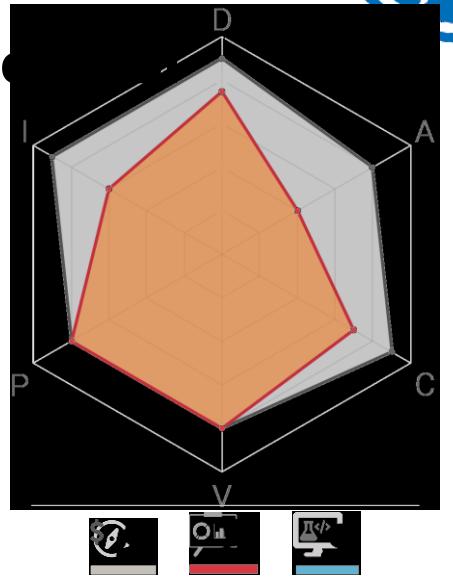
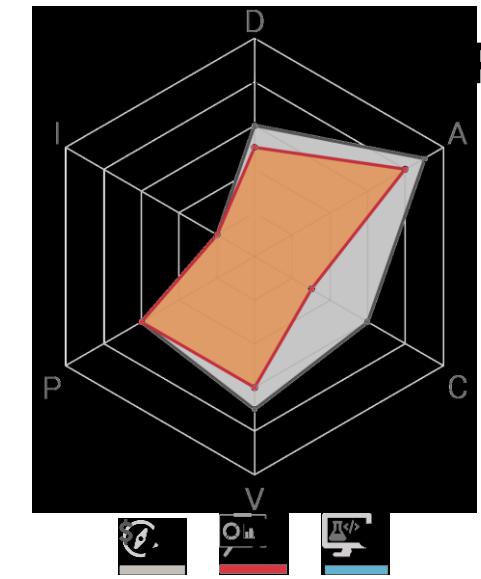
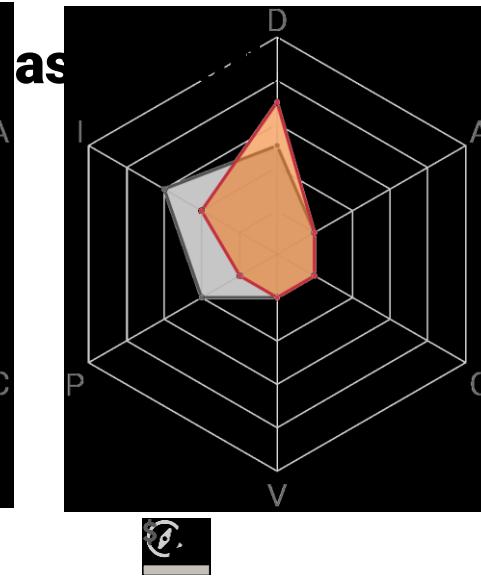
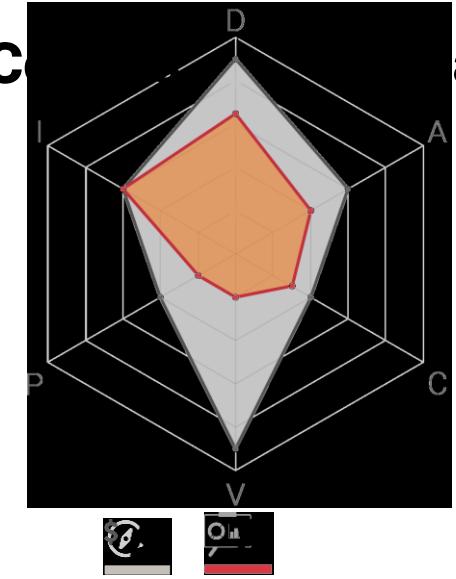
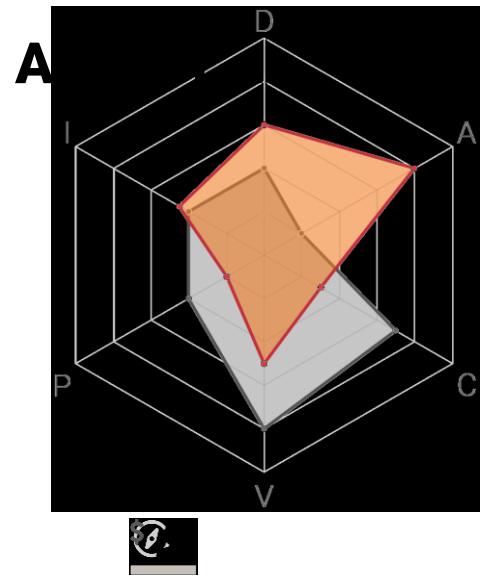
	Data Handling and -Management	Automatic Analysis	Complex Data Types	Visualization Capabilities	User-Guidance, Perception, Cognition	Infrastructure Capabilities
						
						
						
						
						
						
						
						
						



Degree of Innovation Comparison









	Presentation	Confirmatory Analysis	Exploratory Analysis		Data Handling/-Mgmt	Automatic Analysis	Visualization Capabilities
Advizor	●		●		■■■	■■■	■■■■■
Cognos	●—●		●		■■■■■	■■■■■	■■■■■
JasperSoft	●	●	●		■■■	■■■	■■■
JMP	●—●—●		●		■■■	■■■■■	■■■■■
PowerBI	●—●—●		●		■■■■■	■■■■■	■■■■■
Spotfire	●—●—●		●		■■■■■	■■■■■	■■■■■
Tableau	●—●—●		●		■■■■■	■■■■■	■■■■■
Visual Analytics	●—●—●		●		■■■■■	■■■■■	■■■■■
Lumira	●—●		●		■■■■■	■■■■■	■■■■■
QlikView	●—●		●		■■■■■	■■■■■	■■■■■



@michaelbehrisch

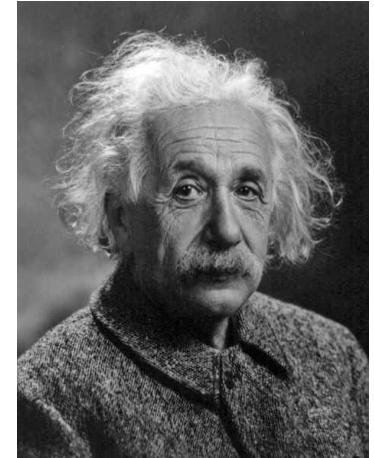


Data Analysis and Visualization

Visual Analytics



"Computers are incredibly fast, accurate, and stupid; humans are incredibly slow, inaccurate, and brilliant; together they are powerful beyond imagination."

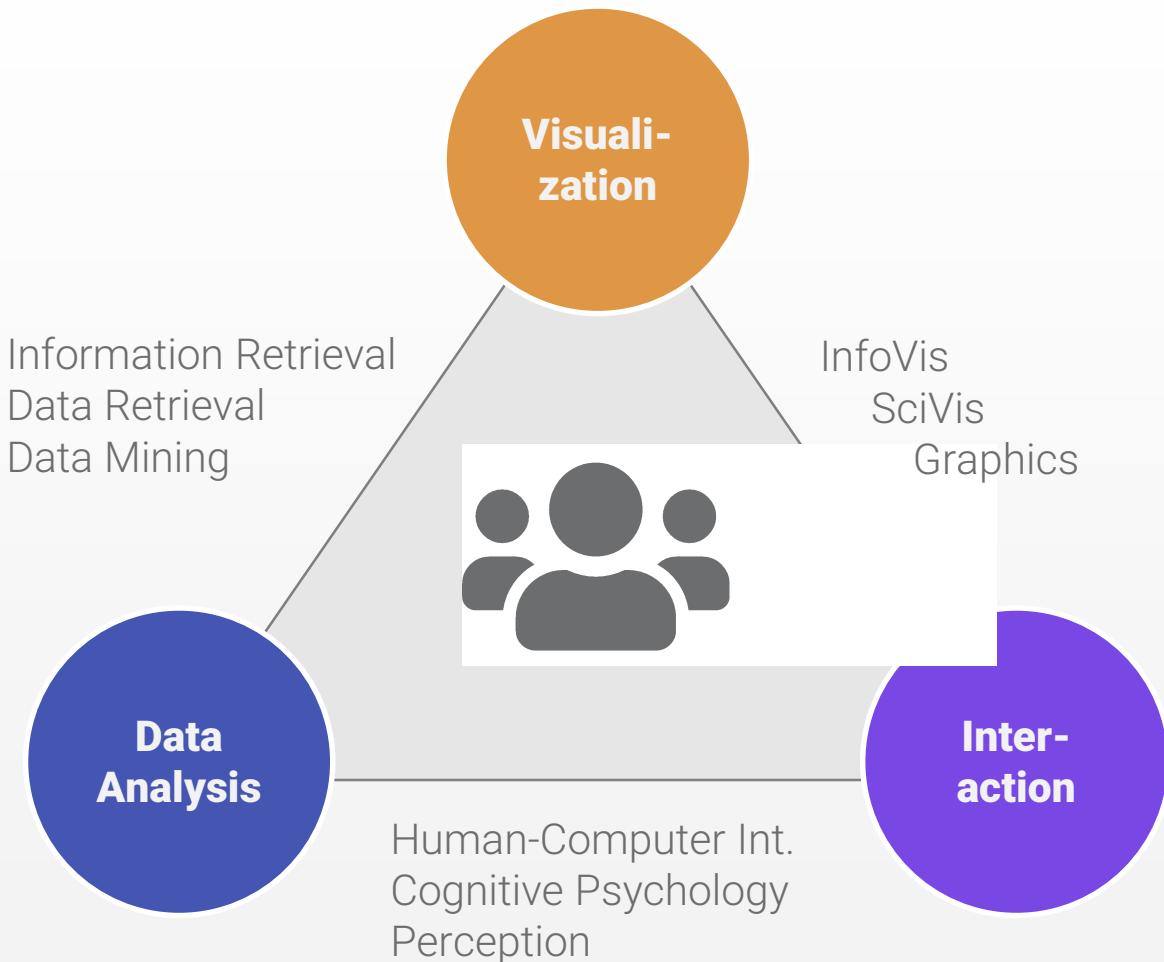


attributed to
Albert Einstein



Data Analysis and Visualization

Visual Analytics





Data Analysis and Visualization

Visual Analytics

User-Centric Exploration of Large Visual Pattern Spaces

Behrisch, M., Blumenschein, M., Kim, N.W., Shao, L., El-Assady, M., Fuchs, J., Seebacher, D., Diehl, A., Brandes, U., Pfister, H. and Schreck, T., 2018, June. Quality Metrics for Information Visualization. *Computer Graphics Forum* (Vol. 37, No. 3, pp. 625-662).

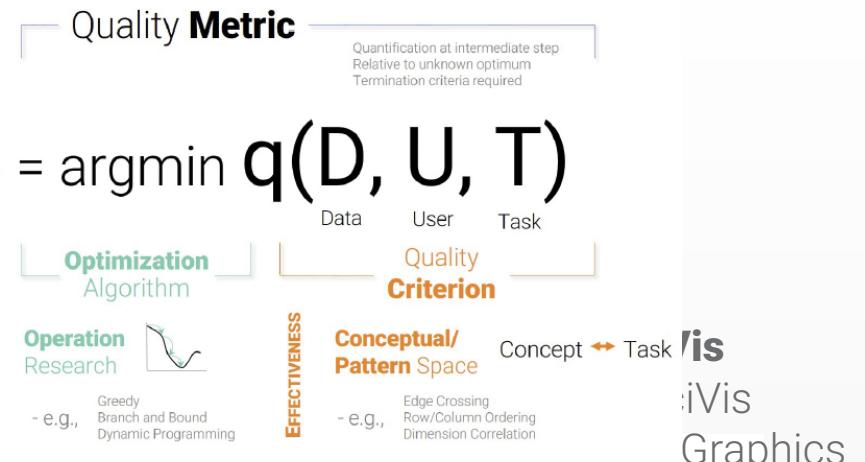


Figure 2: Quality Metrics Formalization

