Visual Data Science Tools

• Differentiate between
  • Charting libraries
  • Applications
Visual Data Science Tools

• Differentiate between
  • Charting libraries
    • Python Plotly
    • Python Matplotlib
    • D3
    • Highcharts
    • GGPlot
  • Applications
Visual Data Science Tools

• Differentiate between
  • Charting libraries
  • Applications
Visual Data Science Tools

• Differentiate between
  • Charting libraries
  • **Applications**
    • Standalone applications
    • Provide means for data handling & visualization
    • Usually no programming skills required
    • In many cases commercial
Visual Data Science Tools

• Differentiate between
  • Charting libraries
  • **Applications**
    • Excel
    • Tableau
    • Microsoft Power BI
    • Cognos
    • QlikView
    • ...

Data Science Applications
Example

• Tableau

Demo
Example

- QlikView

https://sense-demo.qlik.com/sense/app/fa0fb517-24d0-4bee-a82f-1ce5d8a7e614/sheet/cc8f73ae-f2e9-4e02-abee-2b89222ce7d9/state/analysis

Example

• Highcharts Cloud

https://www.highcharts.com/demo
Comparative Study

Commercial Visual Analytics Systems – Advances in the Big Data Analytics Field

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

Abstract—Five years after the first state-of-the-art report on Commercial Visual Analytics Systems we present a reevaluation of the Big Data Analytics field. We build on the success of the 2012 survey, which was influential even beyond the boundaries of the InfoVis and Visual Analytics (VA) community. While the field has matured significantly since the original survey, we find that innovation and research-driven development are increasingly sacrificed to satisfy a wide range of user groups. We evaluate new product versions on established evaluation criteria, such as available features, performance, and usability, to extend on and assure comparability with the previous survey. We also investigate previously unavailable products to paint a more complete picture of the commercial VA landscape. Furthermore, we introduce novel measures, like suitability for specific user groups and the ability to handle complex data types, and undertake a new case study to highlight innovative features. We explore the achievements in the commercial sector in addressing VA challenges and propose novel developments that should be on systems’ roadmaps in the coming years.

Index Terms—System Comparison, Commercial Landscape, Visual Analytics Research, Advances, Development Roadmap.

1 INTRODUCTION

In 1890, Herman Hollerith revolutionized the world of data analysis with a creative and innovative idea: he used punch cards to collect and analyze the US census data. Using punch cards saved two years and five million dollars over the manual tabulation techniques used in the previous census while enabling more thorough analysis of the data [1]. We currently face an analogous development in the Big Data Analysis field, where commercial Visual Analytics (VA) systems allow a faceted confirmatory or a data-driven exploratory analysis of large amounts of data in significantly less time than years ago. Today, the success of many businesses relies on efficient and effective analysis of massive quantities of data.

exploratory analysis scenarios. In particular, we see a scarcity of supportive environments where the domain expert and the machine work in an interplay towards formulating and validating hypotheses. This dearth is due to several reasons: (1) often users are left alone in finding a starting point in their analysis; (2) the communication of non-trivial hypotheses is challenging; (3) automatic algorithms for validating interesting findings are not scalable or even implemented in the systems. This survey counterbalances the efforts of the community against the needs and requirements imposed by the Big Data Era. Further, we ask which steps should be taken in the future by examining past directions to allow for exploratory data analysis in Big Data scenarios.
Comparative Study

• Repeated a study from 2012
• Field changed since then, more products/features available now
• More complete overview of commercial tools
• New study criteria (e.g., user groups)
Comparative Study

• Results made public

https://commercialtools.dbvis.de/home
Survey Pipeline

- **Commercial Reports**
- **Website Research**
- **Survey 2012**

1. **Vendor Longlist**
2. **Questionnaire Features**
3. **Questionnaire Replies / Shortlist**

**Feature Comparison**  
Criteria **Landscape**

**Performance Evaluation**  
Timing **Comparison**

**Case Study Evaluation**  
Findings **Target Groups**
Survey Pipeline

Feature Comparison
Criteria Landscape

Performance Evaluation
Timing Comparison

Case Study Evaluation
Findings Target Groups
Survey Pipeline

**Commercial**
- Reports

**Website**
- Research

**Survey**
- 2012

1. **Vendor Longlist**
2. **Questionnaire Features**
3. **Questionnaire Replies / Shortlist**

**Feature** Comparison
- Criteria **Landscape**

**Performance** Evaluation
- Timing **Comparison**

**Case Study** Evaluation
- Findings **Target Groups**
Available Tools

• Longlist

https://commercialtools.dbvis.de/systems
Survey Pipeline

**Commercial Reports**

**Website Research**

**Survey 2012**

**Vendor Longlist**

**Questionnaire Features**

**Questionnaire Replies / Shortlist**

**Feature Comparison**
Criteria Landscape

**Performance Evaluation**
Timing Comparison

**Case Study Evaluation**
Findings Target Groups
Vendor Questionnaire

Welcome to our 2017 Commercial Tools Survey.

*** Deadline: 3rd of March, 2017 ***

Aim of the survey:
- To collect data to inform a report on the top related topics and areas of interest.
- To identify trends and to highlight gaps in the market.
- To provide insights that can help decision-makers in their selection processes.

Survey Procedure:
- Please take a few minutes to complete this survey.
- You can skip questions if you choose.
- Your responses are anonymous and will be used for research purposes.

Thank you for your participation.

[Survey Link]
Survey Pipeline

- **Feature Comparison**
  - Criteria: Landscape
  - Graph: ○ ○ ○

- **Performance Evaluation**
  - Timing: Comparison
  - Graph: | | |

- **Case Study Evaluation**
  - Findings: Target Groups
  - Graph: $ ☑ ☑ ☑
Shortlist

• **Received replies from**
  - Advizor
  - IBM Cognos
  - SAS JMP
  - SAS Visual Analytics
  - Tableau
  - Jaspersoft
  - Spotfire
  - MS Power BI

• **Additionally added**
  - QlikView
  - SAP Lumira
Vendor Questionnaire – Structure

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

Diagram:
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
Survey Pipeline

**Commercial Reports**

**Website Research**

**Survey 2012**

**Vendor Longlist**

**Questionnaire Features**

**Questionnaire Replies / Shortlist**

**Feature Comparison**
Criteria **Landscape**

**Performance Evaluation**
Timing **Comparison**

**Case Study Evaluation**
Findings **Target Groups**
Survey Pipeline

**Feature Comparison**
Criteria: Landscape

**Performance Evaluation**
Timing: Comparison

**Case Study Evaluation**
Findings: Target Groups

**Commercial Reports**
- Website Research
- Survey 2012

**Vendor Longlist**
- Features
- Replies / Shortlist
<table>
<thead>
<tr>
<th>Data Sources</th>
<th>Tableau</th>
<th>Adizar</th>
<th>JMP</th>
<th>Cognos</th>
<th>Jaspersoft</th>
<th>Spotfire</th>
<th>Visual Analytics</th>
<th>PowerBI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Import</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSV files</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Excel-files (.xls/.xlsx)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>JSON APIs (REST)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>JSON-files</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>XML (REST)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>XML files</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>SQL-Databases (e.g. ODBC/Postgres)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>NOSQL-Databases (e.g. MongoDB)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Directory of text/image/video</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>OPC (UA)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>SAP</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Salesforce</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Export</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>CSV</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Excel</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>JSON</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>XML</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>SQL</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>OPC (UA)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Writeback</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Data Handling and Management

Feature Richness

Degree of Innovation

Subj. Assessment; No Questionnaire Response

Lumira
QlikView
Tableau
PowerBI
Visual Analytics

Spotfire
Cognos
JMP

Jaspersoft
Advizor
Complex Data Types

Feature Richness

Degree of Innovation

QlikView
Advizor
JMP
Tableau
Visual Analytics
PowerBI
Spotfire
Cognos
Lumira
Jaspersoft

Subj. Assessment; No Questionnaire Response
# Data Types

## Supported Data Types

<table>
<thead>
<tr>
<th>Data Types</th>
<th>Advizr</th>
<th>Cognos</th>
<th>Jaspersoft</th>
<th>JMP</th>
<th>PowerBI</th>
<th>Spotfire</th>
<th>Tableau</th>
<th>Visual Analytics</th>
<th>Lumira*</th>
<th>QlikView*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeric with error</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Numeric intervals</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Complex time series</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Item Sets</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Relational/Networks</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Geographic positions</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Geographic traces</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Geographic areas</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Text</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>String-like (e.g. DNA)</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Image</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Video</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>
Infrastructure

- Feature Richness
- Degree of Innovation

- JMP
- Jaspersoft
- Cognos
- PowerBI
- Tableau
- QlikView
- Lumira
- Spotfire

- Computation can be deferred to dedicated clusters/nodes
- Subj. Assessment; No Questionnaire Response
## Infrastructure

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Advisor</th>
<th>Cognos</th>
<th>Jaspersoft</th>
<th>JMP</th>
<th>PowerBI</th>
<th>Spoffire</th>
<th>Tableau</th>
<th>Visual/Anal.</th>
<th>Lumira*</th>
<th>QlikView*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand alone</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Client/Server</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cloud on premise</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cloud at</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>product company</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cloud in Internet</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Multi-CPU support</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Single-GPU support</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Multi-GPU support</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Deferrable dedicated compute clusters/nodes</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

[2]
Automatic Analysis

Feature Richness

PowerBI
Spotfire
JMP
Visual Analytics
Lumira
QlikView
Tableau
Cognos
Jaspersoft
Advizor

Bridge to external data analytics engine (R, Python, etc.)
Subj. Assessment; No Questionnaire Response

Degree of Innovation
User Guidance, Perception, Cognition

- Advizor
- Jaspersoft
- Cognos

- Tableau
- PowerBI
- Spotfire
- Visual An.

OLAP
- JMP
- Lumira

Subject Assessment; No Questionnaire Response

Degree of Innovation

Feature Richness
Survey Pipeline

Feature Comparison
Criteria Landscape

Performance Evaluation
Timing Comparison

Case Study Evaluation
Findings Target Groups
Performance Tests

• Large random datasets of 1 - 500 GB
• All tests performed on the same workstation
• Tested three aspects
  • Data loading
  • Data analysis
  • Data visualization (scatter plot)
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 Visualization

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

Feature Comparison
Criteria Landscape

Performance Evaluation
Timing Comparison

Case Study Evaluation
Findings Target Groups
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
Results
<table>
<thead>
<tr>
<th></th>
<th>Presentation</th>
<th>Confirmatory Analysis</th>
<th>Exploratory Analysis</th>
<th>Data Handling/-Mgmt</th>
<th>Automatic Analysis</th>
<th>Visualization Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advisor</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognos</td>
<td>●●●</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JasperSoft</td>
<td>●</td>
<td>●</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JMP</td>
<td>●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PowerBI</td>
<td>●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spotfire</td>
<td>●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tableau</td>
<td>●●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Analytics</td>
<td>●●●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumira</td>
<td>●●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QlikView</td>
<td>●●●●</td>
<td></td>
<td>●</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Feature Richness Comparison

<table>
<thead>
<tr>
<th></th>
<th>Data Handling and Management</th>
<th>Automatic Analysis</th>
<th>Complex Data Types</th>
<th>Visualization Capabilities</th>
<th>User-Guidance, Perception, Cognition</th>
<th>Infrastructure Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alik Q</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>Spotfire</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>Tableau</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>SAS</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>JMP</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>ADVIZOR Solutions, Inc.</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>Lumira</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>Cognos</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
<tr>
<td>TIBCO</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
<td>🌌 🌌 🌌 🌌 🌌 🌌 🌌</td>
</tr>
</tbody>
</table>
Degree of Innovation Comparison

<table>
<thead>
<tr>
<th></th>
<th>Data Handling and Management</th>
<th>Automatic Analysis</th>
<th>Complex Data Types</th>
<th>Visualization Capabilities</th>
<th>User-Guidance, Perception, Cognition</th>
<th>Infrastructure Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qlik</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>Spotfire</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>Tableau</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>SAS</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>JMP</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>Advizor Solutions, Inc.</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>Lumira</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>Cognos</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
<tr>
<td>TIBCO</td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
<td><img src="image" alt="Moon" /></td>
</tr>
</tbody>
</table>
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.
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
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
Comparative Studies

• Gartner
Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner (February 2018)
### Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

<table>
<thead>
<tr>
<th></th>
<th>Data Handling/Mgmt</th>
<th>Automatic Analysis</th>
<th>Visualization Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advisor</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>Cognos</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>JasperSoft</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>JMP</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>PowerBI</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>Spotfire</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>Tableau</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>Visual Analytics</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>Lumira</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
<tr>
<td>QlikView</td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
<td><img src="image" alt="Rating" /></td>
</tr>
</tbody>
</table>

Source: Gartner (February 2018)
Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner (February 2018)

As of February 2018

© Gartner, Inc

[3]
### Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner (February 2018)

<table>
<thead>
<tr>
<th>Data Handling/Mgmt</th>
<th>Automatic Analysis</th>
<th>Visualization Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advizor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cognos</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>JasperSoft</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>JMP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PowerBI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Spotfire</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tableau</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Visual Analytics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Lumira</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>QlikView</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

<table>
<thead>
<tr>
<th>Data Handling/-Mgmt</th>
<th>Automatic Analysis</th>
<th>Visualization Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advizor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JasperSoft</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JMP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PowerBI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spotfire</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tableau</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Analytics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumira</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QlikView</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Gartner (February 2018)
Figure 1. Magic Quadrant for Business Intelligence and Analytics Platforms

Source: Gartner (February 2017)

Figure 1. Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner (February 2018)
Comparative Studies

• What I Learned Recreating 1 Chart Using 24 Tools

https://source.opennews.org/articles/what-i-learned-recreating-one-chart-using-24-tools/
1 Chart, 12 Applications
Study comparison

- Advizor
- IBM Cognos
- SAS JMP
- SAS Visual Analytics
- Tableau
- Jaspersoft
- Spotfire
- MS Power BI
- QlikView
- SAP Lumira
Study comparison

- Advizor
- IBM Cognos
- SAS JMP
- SAS Visual Analytics
- **Tableau**
- Jaspersoft
- Spotfire
- MS Power BI
- QlikView
- SAP Lumira
Study comparison
Comparative Studies

• Comparison of Open Source Visual Analytics Toolkits

Comparison of Open Source Visual Analytics Toolkits

John R. Harger\textsuperscript{a,b} and Patricia J. Crossno\textsuperscript{a}

\textsuperscript{a}Sandia National Laboratories, PO Box 5800 Albuquerque, NM 87185-1323, USA;
\textsuperscript{b}University of New Mexico, Albuquerque, NM 87131, USA

ABSTRACT
We present the results of the first stage of a two-stage evaluation of open source visual analytics packages. This stage is a broad feature comparison over a range of open source toolkits. Although we had originally intended to restrict ourselves to comparing visual analytics toolkits, we quickly found that very few were available. So we expanded our study to include information visualization, graph analysis, and statistical packages. We examine three aspects of each toolkit: visualization functions, analysis capabilities, and development environments. With respect to development environments, we look at platforms, language bindings, multi-threading/parallelism, user interface frameworks, ease of installation, documentation, and whether the package is still being actively developed.

Keywords: Visual Analytics, open source, toolkits, comparison, evaluation
## Comparative Studies

<table>
<thead>
<tr>
<th></th>
<th>Bar Chart</th>
<th>Line Chart</th>
<th>Scatterplot</th>
<th>Box Plot</th>
<th>Pie Chart</th>
<th>Contour Plot</th>
<th>Stacked Bar Chart</th>
<th>Stacked Area Chart</th>
<th>Parallel Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Axiis</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>birdeye</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flare</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gephi</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Google Vis</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>GraphViz</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvise*</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>IVTK</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JIT</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JFreeChart</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>JGraph</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JUNG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NetworkX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prefuse</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Protovis</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>R</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Titan</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Talip</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>VisAD</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>WilmaScope</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zest</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[6]
"Computers are incredibly fast, accurate, and stupid; humans are incredibly slow, inaccurate, and brilliant; together they are powerful beyond imagination."

attributed to Albert Einstein
References


