Advanced Analytics and R/Python Integration

Ugo Giaquinto
Senior Solution Architect
Qlik® roadmaps provides a general overview of our anticipated product direction as of the date hereof. Qlik undertakes no intention or obligation to update or revise any forward-looking statements contained herein, whether as a result of new information, future events or otherwise.

The information contained in Qlik roadmaps is the proprietary information of Qlik and is not to be used for any reason other than to provide our customers, prospective customers, partners and strategic partners with a general overview of our anticipated future offerings. Our roadmaps are not a commitment, promise or other obligation on the part of Qlik to deliver any particular code or functionality. Furthermore, Qlik makes no commitment that any future functionality may be made available as part of maintenance and support.

Roadmap information should not be relied on in making a purchasing decision, as the development, release, and timing of any features or functionality described for Qlik’s products remains at our sole discretion.
Introduction to Advanced Analytics
Advanced Analytics Goals

Happy Customers

Eliminate Bad Behaviors

Run the Business Better

<table>
<thead>
<tr>
<th>Goal</th>
<th>2011</th>
<th>2013</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improving understanding of customers</td>
<td>33%</td>
<td>45%</td>
<td>46%</td>
</tr>
<tr>
<td>Retaining customers</td>
<td>30%</td>
<td>36%</td>
<td>37%</td>
</tr>
<tr>
<td>Improving customer experiences</td>
<td>22%</td>
<td>36%</td>
<td>36%</td>
</tr>
<tr>
<td>Selling products / services to existing</td>
<td>23%</td>
<td>33%</td>
<td>35%</td>
</tr>
<tr>
<td>Customers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market research / survey analysis</td>
<td>29%</td>
<td>36%</td>
<td>34%</td>
</tr>
<tr>
<td>Acquiring customers</td>
<td>23%</td>
<td>32%</td>
<td>32%</td>
</tr>
<tr>
<td>Improving direct marketing programs</td>
<td>22%</td>
<td>27%</td>
<td>30%</td>
</tr>
<tr>
<td>Sales forecasting</td>
<td>39%</td>
<td>27%</td>
<td>27%</td>
</tr>
<tr>
<td>Fraud detection or prevention</td>
<td>21%</td>
<td>23%</td>
<td>26%</td>
</tr>
<tr>
<td>Risk management / credit scoring</td>
<td>22%</td>
<td>26%</td>
<td>25%</td>
</tr>
<tr>
<td>Price optimization</td>
<td>14%</td>
<td>22%</td>
<td>23%</td>
</tr>
<tr>
<td>Manufacturing improvement</td>
<td>10%</td>
<td>15%</td>
<td>17%</td>
</tr>
<tr>
<td>Medical advancement / drug discovery /</td>
<td>12%</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>biotech / genomics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supply chain optimization</td>
<td>7%</td>
<td>11%</td>
<td>15%</td>
</tr>
<tr>
<td>Investment planning / optimization</td>
<td>11%</td>
<td>13%</td>
<td>14%</td>
</tr>
<tr>
<td>Software optimization</td>
<td>7%</td>
<td>9%</td>
<td>11%</td>
</tr>
<tr>
<td>Website or search optimization</td>
<td>8%</td>
<td>12%</td>
<td>10%</td>
</tr>
<tr>
<td>Human resource applications</td>
<td>4%</td>
<td>8%</td>
<td>9%</td>
</tr>
<tr>
<td>Collections</td>
<td>6%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>Language understanding</td>
<td>4%</td>
<td>7%</td>
<td>8%</td>
</tr>
<tr>
<td>Criminal or terrorist detection</td>
<td>4%</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>Information security</td>
<td>4%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>Natural resource planning or discovery</td>
<td>3%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>Fundraising</td>
<td>3%</td>
<td>3%</td>
<td>3%</td>
</tr>
<tr>
<td>Reducing email spam</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Common Use Cases

Manufacturing & High Tech
- Demand Forecasting
- Quality & Machine Breakdown
- Warranty Analytics

Healthcare
- Patient Admission Forecasting
- Readmission Analysis
- Clinical Outcomes

Communications
- Customer Segmentation/Churn
- Quality/Process Control
- Next Best Offer/Action

Financial Services
- Credit Risk Analysis
- Loan Default
- Fraud Detection

Public Sector
- Crime Prevention
- Transportation Optimization
- Recidivism

Retail & Consumer
- Market Basket Analysis
- Retail Store Selection
- Product Recommendation

Life Sciences
- Territory Alignment
- Campaign Management
- Drug Discovery & Effectiveness

Energy & Utilities
- Customer Churn
- Proactive Failure
- Peak Demand
What our customers are asking

We have invested hundreds of hours in developing algorithms in R, Python, or other engines to support their business…

Can I leverage this in Qlik Sense?

We are evaluating Qlik Sense and one of our main requirements is to include advanced analytics with visual discovery…

Is Qlik Sense right for my business?

We have an advanced analytics use case, such as fraud detection, sales forecasting, inventory management, or price optimization…

Can Qlik Sense support this?
Key questions to consider

• How will you deliver advanced analytics to users who make decisions?
• How will these decision makers explore and interact with advanced analytics calculations?
• How can your organization get the most out of the data scientists you have invested in?
A broad spectrum of analytics requirements

- **Corporate Reporting**
  - Consumers
  - Business User
  - Ad-hoc Analysis
  - Business Analyst
  - Advanced Calculations
- **Self-Service Analytics**
  - Data Preparation
- **Advanced Analytics**
  - Data Scientist
  - Scenario Analysis

**BI / Reporting**
- Descriptive
- Diagnostic
- Predictive
- Prescriptive
- Optimization

- Prompts
- Static Output
- Table / List
- Dashboard
- Selection
- Linkage
- Benchmarking
- Trends
- Growth
- Anomaly
- Bar / Line Chart
- Filtering
- Sorting
- Pivot Table
- Data Export
- Visualization
- Aggregation
- Association
- Comparison
- Search
- Data Stories
- Set Analysis
- Subtotals
- Rolling
- Aggregation
- Classification
- Segmentation
- Scripting
- Data Generation
- Randomization
- Relative Time
- Distribution
- Regression
- Forecasting
- Correlation
- Clustering
- Histogram
- Box Plot
- Sentiment
- Sampling
- Decision Trees
- Time Series
- Text Mining
- Neural Nets
- Random Forests
- Ensemble Models
- Factor Analysis
- Monte Carlo

- Business User
- Guided Analytics
- Advanced Calculations
- Statistics / Algorithms
- Scenario Analysis
- Data Preparation

- Business Analyst
- Visual Exploration

- Ad-hoc Analysis
- Advanced Calculations

- Self-Service Analytics

- BI / Reporting

- Consumers

- Producers

- Descriptive
- Diagnostic
- Predictive
- Prescriptive
- Optimization

- Business User
- Guided Analytics
- Ad-hoc Analysis
- Business Analyst
- Visual Exploration
- Advanced Calculations
- Statistics / Algorithms
- Scenario Analysis

- Data Preparation
Qlik Sense offers native and third party capabilities

- **Native capabilities** include statistical functions, interactive visualizations, and scripting for many advanced analytics use cases.

- **Advanced Analytics Integration** offers engine-level sharing between Qlik Sense and third party tools such as R and Python, for incorporating advanced calculation and machine learning into analyses.
The old way

Data Acquisition
Data Preparation
Model Creation
Data Selection
Model Selection
Result Generation
Result Presentation

Results Interpretation

Action Steps
Roles are changing

Data Scientist
- Model Creation
- Governance & Oversight

Business Analyst
- Data Acquisition
- Data Preparation
- Automated Model Selection
- Advanced Parameterization

Business User
- Data Selection
- Basic Parameterization
- Result Generation
- Result Presentation
- Action Steps

Roles are changing
- Business Analyst
- Data Scientist
- Business User

Business Analyst

Data Scientist

Business User

Advanced Parameterization

Automated Model Selection

Data Preparation

Data Acquisition
The benefit – advanced insights for everyone

• Advanced analytics in Qlik Sense deliver powerful insights to business users
  – Data scientists build advanced models and calculations
  – Business decision makers can utilize them in the context of associative exploration
  – Analytics are calculated and visualized in real-time as the user explores, based on selected context
Advanced Analytics Integration
Introducing Advanced Analytics Integration

- Direct integration with 3rd party advanced analytics engines through server-side extension APIs
- Allows data to be directly exchanged between the Qlik engine and external tools during analysis
  - Leverages Qlik’s Associative Engine to pass relevant data based on user context
- Full integration with Qlik Sense expressions and libraries
- Connectors can be built for any external engines
- Open source Analytic Connections are available in GitHub for R and Python (https://github.com/qlik-oss)

Leverage the power of advanced analytics calculations in Qlik Sense
AII Demonstration – UK Crime Statistics
How it works

1. User interacts with app, making a selection or a search
2. Hypercube recalculated by Qlik Engine to the new context
3. In-context data and script sent to external engine
4. External engine runs and sends results to Qlik engine
5. Qlik engine combines hypercube with new data
6. Combined hypercube is visualized for the user in the app
R.ScriptEval('library(TTR);decompose(ts(q$sumBirthsPerMonth, frequency=12, start=c(1946,1)))$seasonal', Sum([Births per month]) as sumBirthsPerMonth)
Developers
• Build and customize extensions and connectors
• Contribute to open source projects

Administrators
• Deploy and secure extensions and connectors
• Manage environments and infrastructure

Data Scientists
• Design and develop third party algorithms for R, Python, etc.
• Support development of Qlik Sense apps with advanced analytics

Business Analysts and Users
• Build Qlik Sense apps with advanced analytics for business use cases
• Perform business analysis
What makes our approach unique
Qlik’s Associative Engine working with all third party engines

- Connectors can be built for any third party engines, through open APIs
- As the user explores, only a small set of chosen and relevant data is sent
- Results are instantly visualized for the user, allowing for further exploration

- Third party engines quickly process smaller, user-specific data sets
- Far more speed than conventional batch techniques
- Results for each user are sent back to Qlik Sense in real-time
Advanced Analytics Use Cases
Financial Services

Fraud Detection

- Leverage pattern analysis comparing customer activity with peer group behavior
- Utilize customer’s own past behavior to identify outlying transactions
- Bayesian learning, neural networks, fuzzy neural nets, and combinations of neural nets and rules, have been extensively explored

Expert knowledge is integrated with statistical power and rule-learning programs to uncover indicators of fraudulent behavior from a large database of customer transactions.
Knowing how many patients are likely to be admitted provides a forward look on bed availability and flags potential capacity issues in the days or weeks to come.

- Qlik solution has taken customer from one of the worst performing ER's to one of the best in the country.
- All the forecasting calculations are performed inside Qlik and get ~85% accuracy.

Forecasting is an important aid in many areas of hospital management, including elective surgery scheduling, bed management, and staff resourcing.
Financial Services

Credit Loss Forecasting

- Credit managers need to predict their expected future credit losses.
- This is mandatory for regulations like iFRS9.
- Qlik’s solution enables teams to ask complex questions without notice – and answer them in real time.
- Managers can predict the expected losses for specific products – loans written in any region, in any time period, filtered by employment type or credit grade.

Enable Credit Managers to forecast with accuracy – understanding what factors influence credit losses, and optimizing the credit book to minimize future losses.
Retail

Market Basket Analysis

- Driving actionable insight from customer purchase patterns hidden within the mountain of data
- Leverage advanced analytics and algorithms to produce statistics like support, confidence, and lift for a deep dive into product affinities
- Quickly consume but also allow business users to ask the inevitable “next question” that goes beyond simple static reporting

Retailers must understand consumer preferences and purchase patterns to help provide more targeted offers and tailored assortments, resulting in larger basket sizes and locking in higher margins
Market Basket Analysis
Consumer

Customer Churn

• Classification (Decision Tree) algorithms to segment customers as at-risk

• Analyze different ‘What-if’ scenarios by modifying the input variables and assess the overall impact on customer models

• Evaluation and processing of call records for sentiment analysis and categorization

Focus retention campaigns and deliver preferential services to at-risk high-value subscribers who have been predicted to have a propensity to churn
Customer Sentiment Analysis
For More Information

• Qlik Community Resources – Advanced Analytics Integration
  – https://community.qlik.com/community/advanced-analytics-integration

• Qlik Product Innovation Blog
  – https://community.qlik.com/blogs/qlikproductinnovation
Thank You