Test Automation: Agile Enablement for Data Warehousing and Business Intelligence Teams

Presented to Agile 2016
July 26, 2016
Atlanta, Georgia

Lynn Winterboer, Agile Analytics Educator and Coach, @AgileLynn
Cher Fox, DW/BI Solution Architect, @TheDatanista
Agenda

• Why is test automation important for agile data teams?
• Why aren’t all data teams automating their tests today?
• What is the path to test automation?
• What does simple DW test automation look like?
Why is test automation important to agile data teams?
Agile Demands Something Different

Agile calls for small increments of “potentially shippable code”.

That means QA is essential on a regular, frequent basis.

The test suite grows larger each iteration.

Manual testing quickly becomes infeasible.

Doesn’t break something we did back here

How can we be sure that something we do here
Testing is Central to Agile

Agile BI development is driven by tests

Acceptance criteria are the definition of “done”

Passing tests are the measure of “done”

Regression tests are the measure of “still done”
Agile Teams Understand...

... that test automation is a key technical enabler to “being agile”
Why aren’t all data teams automating their tests today?
DW/BI Test Automation Challenges

Data industry does not focus on testing as much as other disciplines:

Education

Team Members

Skills & Discipline
DW/BI Test Automation Challenges

Existing tools are not easily adapted to the data world:

Development Languages

Web, Mobile and UI Focus
DW/BI Test Automation Challenges

Good test data sets are hard to come by:

- Large volumes (on small environments)
- Shared test environments
- Data sensitivity
- Small sets need extra planning and design
What is the path to test automation?
Exercise: Path to test automation

In small groups, discuss the following pre-requisites to test automation and decide on the first three steps a DW/BI team should take toward test automation:

<table>
<thead>
<tr>
<th>Business Domain Knowledge</th>
<th>Learning Culture</th>
<th>Test Maintenance</th>
<th>Repeatable Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Design</td>
<td>Test Feedback Loops</td>
<td>Dedicated Testers</td>
<td>UI Tests</td>
</tr>
<tr>
<td>Exploratory Testing</td>
<td>Testing Skills</td>
<td>Test Data Sets</td>
<td>Unit Tests</td>
</tr>
<tr>
<td>Test Environment</td>
<td>Automated Builds</td>
<td>Testing is a Team Sport</td>
<td>Test Organization</td>
</tr>
<tr>
<td>Manual Tests</td>
<td>Build Quality In</td>
<td>Acceptance Tests</td>
<td>Test Automation Tool</td>
</tr>
</tbody>
</table>
Agile Testing Perspectives

1) Attitudes & Culture
- Testing is a Team Sport
- Learning Culture
- Build Quality In

2) Skills & Practices
- Testing Skills
- Test Design, Organization & Maintenance
- Repeatable Tests
- Feedback Loops

3) Tools & Automation
- Test Environment
- Build Automation
- Test Data Sets
- Test Automation Tool

High Quality DW/BI System
Test Automation Pyramid

Subjective user feedback is essential

Brittle tests that are affected by UI changes

Write these in business domain language

Push as many tests as possible to this layer

Good Places to Start

- Unit tests on new development
  
  **AGILE TEAMS WRITE TESTS BEFORE WRITING ANY CODE**

- Regression tests you wish you had, before building something new

- Deployment tests to ensure each build migrated correctly

- Painful test setup or test execution prone to error
What does simple data warehousing test automation look like?
### Basic Approach

<table>
<thead>
<tr>
<th>Source Schema</th>
<th>Component under test</th>
<th>Target Schema</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Data</td>
<td></td>
<td>Actual Results</td>
</tr>
</tbody>
</table>

**Setup**

- Load test data
- Based on specific input data
- Leave no trace

**Test Runner**

- Setup
- Execute
- Test
- Report
- Clean Up

**Expected Results**

- Test Queries
  - Execute tests
  - Compare results
  - Leave no trace

**Actual Results**

- Static Test Data
  - Load test data
  - Based on specific input data
  - Leave no trace

**Test Data**

- Load test data
- Based on specific input data
- Leave no trace
What Constitutes a Test?

Actual Result (from test)

Expected Result (queried or pre-calculated)
Exercise: Test Data & Expected Results

In small groups, define at least one test, with a description of the test data set and how you would determine the expected results, for the following user story:

As a Sales Manager,

I want to know the $ total of orders that have successfully shipped, by sales rep and sales region,

So that I can manage the sales team’s progress toward quarterly revenue goals.
Demonstration
Testing Framework

Functionality and Components:

- Example source system
  - SQL Server and related model (“AdventureWorks”)
- Screens for describing and creating test cases
  - Microsoft Access form
- Test execution
  - SQL Server stored procedures (individual tests)
  - Power Shell scripts (batch execution)
- BI Reports and Dashboards for monitoring
  - MS Power BI
Demo Workflow

1) Define a test (MS Access form)

2a. Execute a single test (SQL Server stored procedure)

2b. Execute a batch of tests (PowerShell scripts)

3. Report on test results (MS Power BI)
Summary and Resources
Session Intent

1) Why:
   Test Automation is key to DW/BI agile success

2) How:
   Automate a solid testing practice; don’t automate chaos

3) Show:
   DW/BI test automation is not that hard
## DW/BI Test Automation References

<table>
<thead>
<tr>
<th>Reference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Agile Analytics</strong>: A Value-Driven Approach to Business Intelligence and Data Warehousing by Dr. Ken Collier</td>
<td>Chapter 7: Test-Driven Data Warehouse Development</td>
</tr>
<tr>
<td><strong>Automated Data Warehouse Testing</strong>: Beginner’s Step-by-Step Guide by G. Suden</td>
<td>Java-based test automation for simple data warehousing tests</td>
</tr>
<tr>
<td><strong>Agile Testing</strong>: A Practical Guide for Testers and Agile Teams by Lisa Crispin and Janet Gregory</td>
<td>Great introduction to agile testing for any type of development</td>
</tr>
<tr>
<td><strong>More Agile Testing</strong>: Learning Journeys for the Whole Team by Lisa Crispin and Janet Gregory</td>
<td>Includes several chapters on test automation as well as one dedicated to agile testing for DW/BI</td>
</tr>
</tbody>
</table>
## DW/BI Test Automation Tools

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBFit</td>
<td>Open source database testing tool</td>
</tr>
<tr>
<td><a href="dbfit.github.io/dbfit/">dbfit.github.io/dbfit/</a></td>
<td></td>
</tr>
<tr>
<td>iCEDQ</td>
<td>Test automation tools designed specifically for Data Warehousing and related projects.</td>
</tr>
<tr>
<td><a href="icedq.com">icedq.com</a></td>
<td></td>
</tr>
<tr>
<td>QuerySurge</td>
<td></td>
</tr>
<tr>
<td><a href="QuerySurge.com">QuerySurge.com</a></td>
<td></td>
</tr>
<tr>
<td>Zuzena</td>
<td></td>
</tr>
<tr>
<td><a href="Zuzena.com">Zuzena.com</a></td>
<td></td>
</tr>
<tr>
<td>Informatica Data Validation</td>
<td>Accelerate and automate Informatica ETL testing in both production environments and dev/ test</td>
</tr>
<tr>
<td><a href="www.informatica.com/etl-testing">www.informatica.com/etl-testing</a></td>
<td></td>
</tr>
<tr>
<td>Analyx Data Services</td>
<td>DW automation tools that include test automation capabilities.</td>
</tr>
<tr>
<td><a href="analytixds.com">analytixds.com</a></td>
<td></td>
</tr>
<tr>
<td>WhereScape</td>
<td></td>
</tr>
<tr>
<td><a href="wherescape.com">wherescape.com</a></td>
<td></td>
</tr>
<tr>
<td>TimeXtender</td>
<td></td>
</tr>
<tr>
<td><a href="timeXtender.com">timeXtender.com</a></td>
<td></td>
</tr>
<tr>
<td>Tricentis Tosca</td>
<td>DW-friendly test automation tool</td>
</tr>
</tbody>
</table>
Questions?

Lynn Winterboer
Agile Analytics Educator & Coach
www.WinterboerAgileAnalytics.com
lynn@WinterboerAgileAnalytics.com
@AgileLynn

Cher Fox
DW/BI Solution Architect
www.foxconsulting.co
info@foxconsulting.co
@TheDatanista

Special thanks for contributing to this presentation:
Ken Collier
Brad Ewald
Deborah Krinitzsky
Joe Bernardini