

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

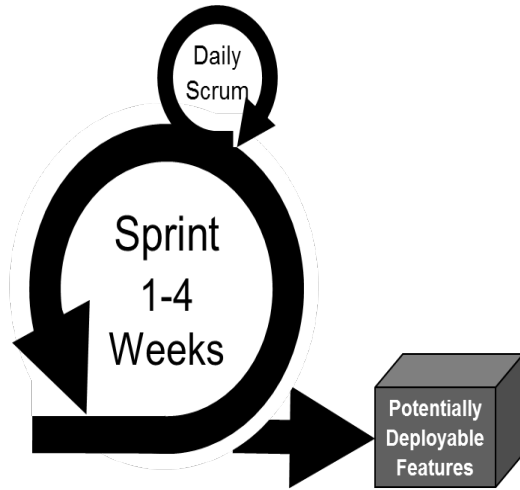
- Why is test automation important for agile data teams?
- Why aren't all data teams automating their tests today?
- What is the path to test automation?
- What does simple DW test automation look like?



Why is test automation important to agile data teams?



Agile Demands Something Different

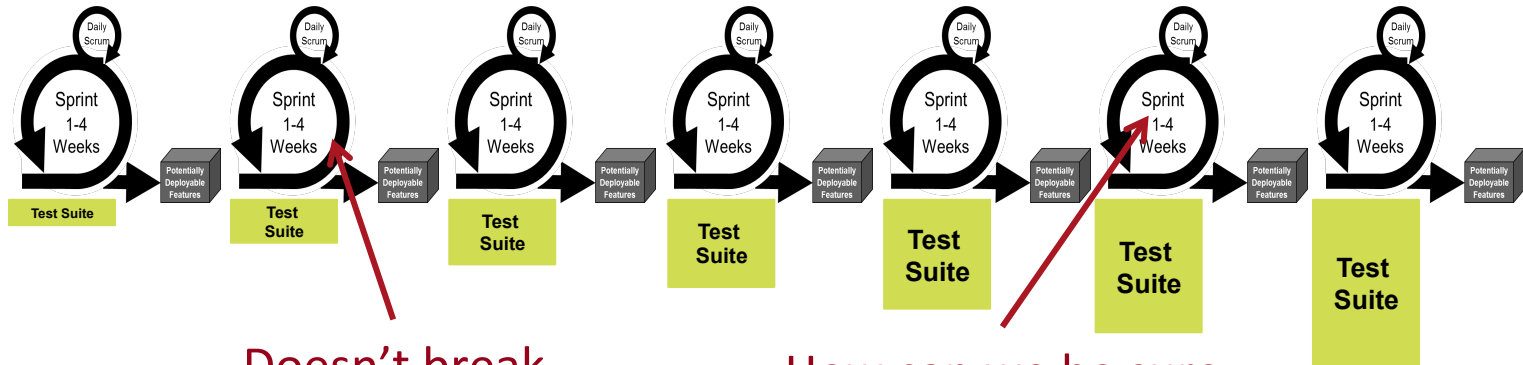


Agile calls for small increments of “potentially shippable code”.

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?

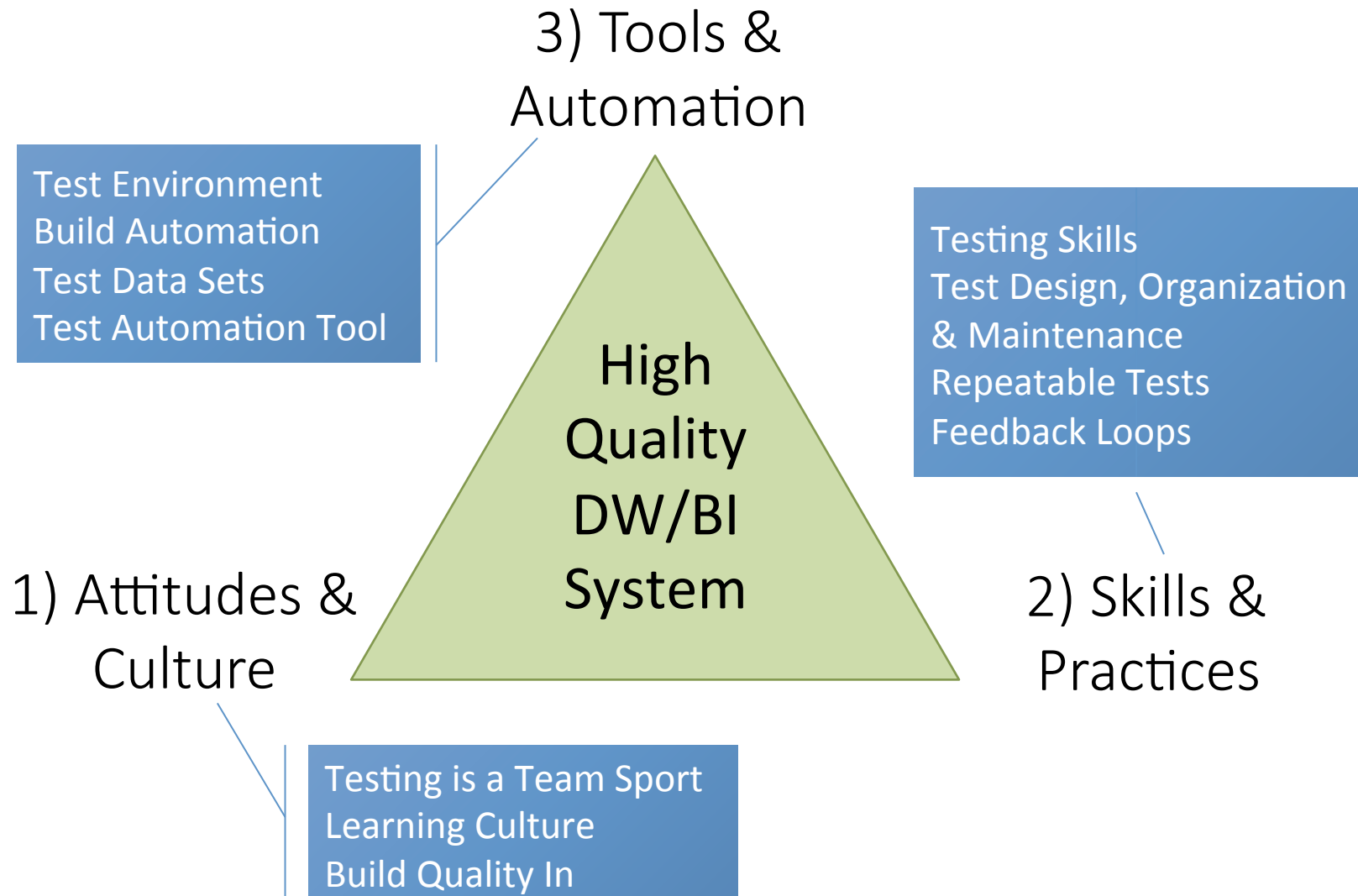


Exercise: Path to test automation

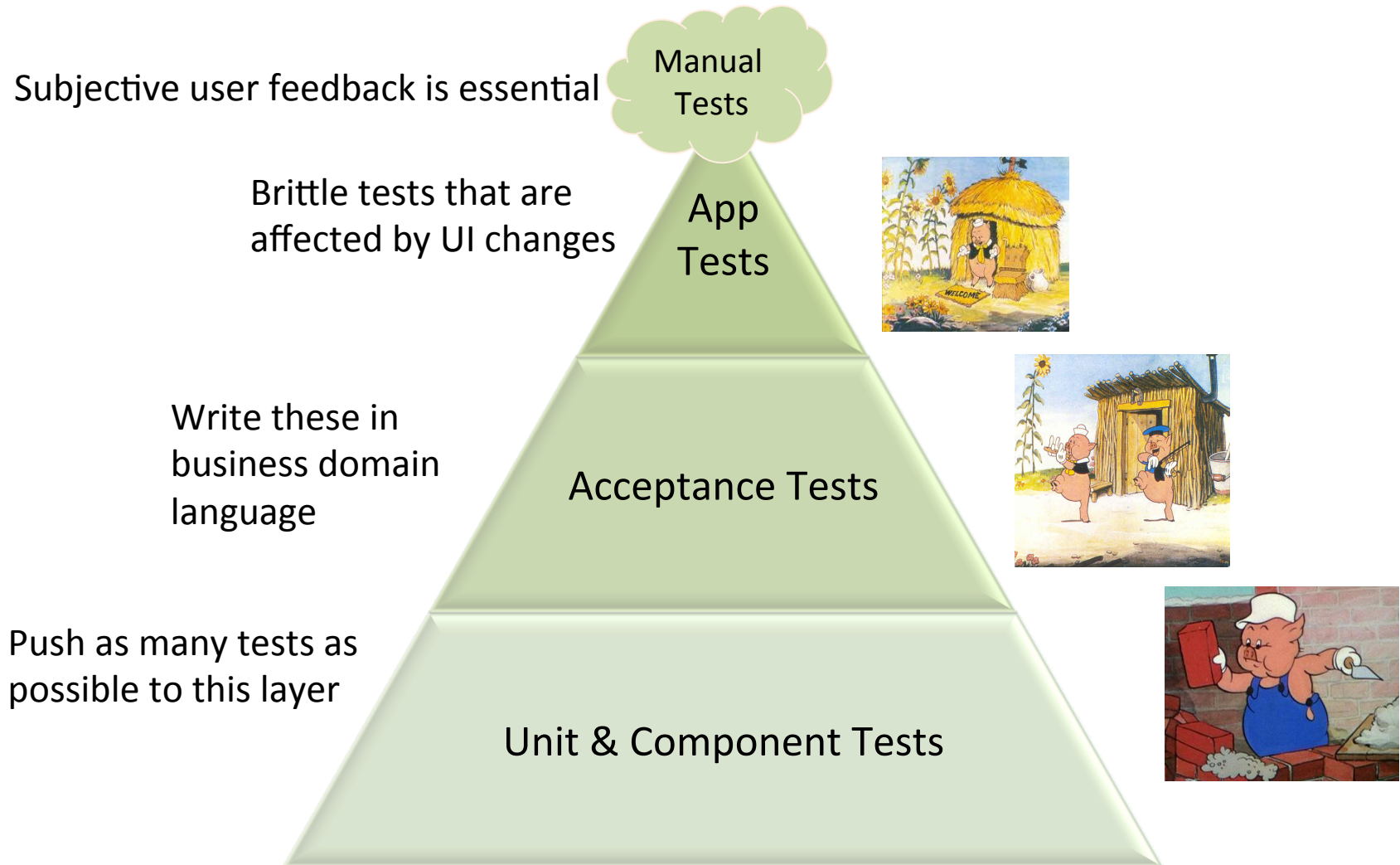
In small groups, discuss the following pre-requisites to test automation and decide on the first three steps a DW/BI team should take toward test automation:

Business Domain Knowledge	Learning Culture	Test Maintenance	Repeatable Tests
Test Design	Test Feedback Loops	Dedicated Testers	UI Tests
Exploratory Testing	Testing Skills	Test Data Sets	Unit Tests
Test Environment	Automated Builds	Testing is a Team Sport	Test Organization
Manual Tests	Build Quality In	Acceptance Tests	Test Automation Tool

Agile Testing Perspectives



Test Automation Pyramid



Good Places to Start

- Unit tests on new development

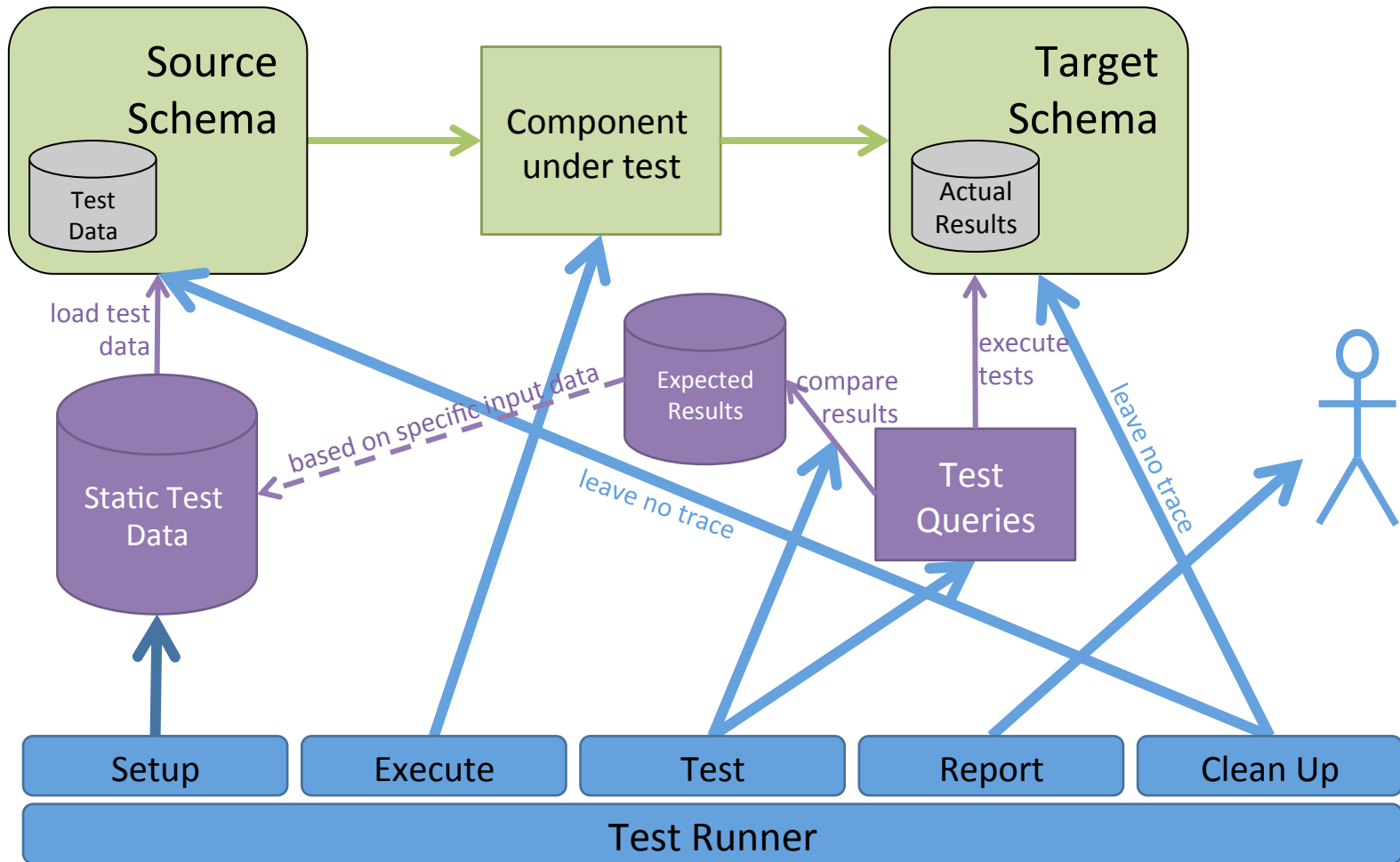
AGILE TEAMS WRITE TESTS BEFORE WRITING ANY CODE

- 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

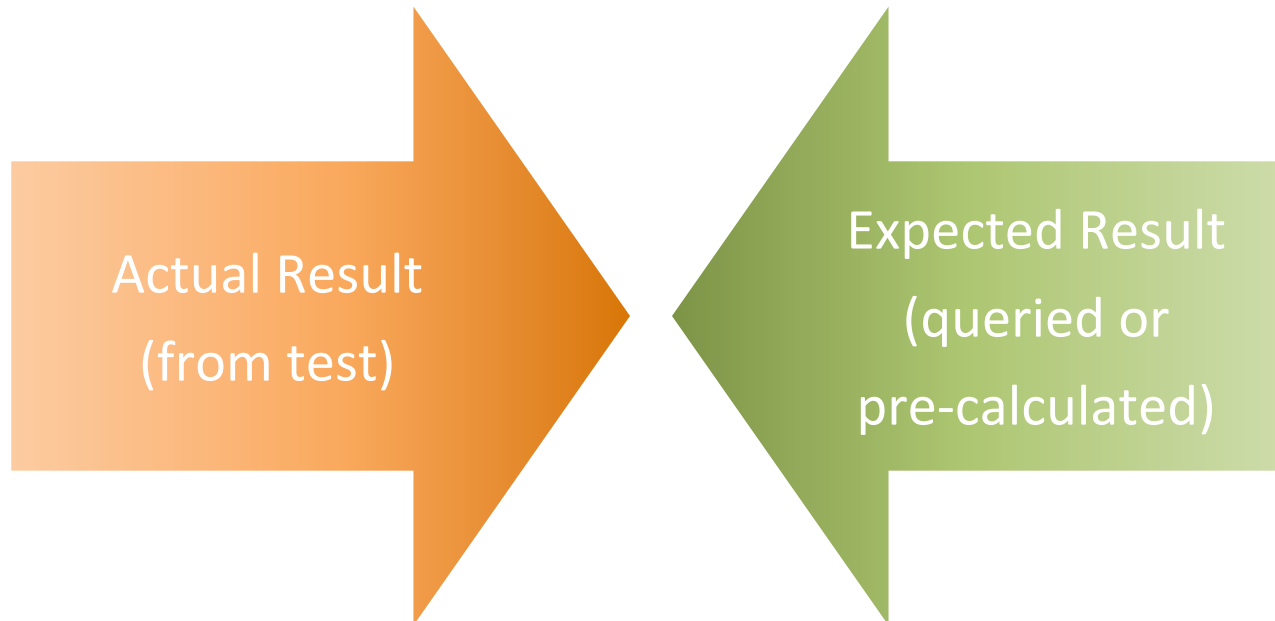
What does simple data warehousing test automation look like?



Basic Approach



What Constitutes a Test?



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)

The screenshot shows the 'dbo_D_TEST_DEF' form in MS Access. The form contains the following fields and values:

- ID: [Empty]
- CREATE_DT: 11/30/2015 4:23:51 PM
- TEST_DEF_DESC: SalesOrderHeader.SalesOrderID Not Null
- TEST_DEF_CD: SOHL_PK_NOT_NULL
- TEST_TYPE: Unit
- TEST_GROUP: Source to Stage
- TEST_QUERY: Select Count([SalesOrderID]) From [Stage].[SalesOrderHeader] Where [SalesOrderID] is Null
- EXPECTED_RESULT_QUERY: [Empty]
- TEST_DB: [Empty]
- EXPECTED_RESULT_DB: [Empty]
- RET_COLUMN_CNT: 1
- EXPECTED_RESULT: 0
- COMPARE_OPERATOR: =
- ACTIVE_FLG: Y
- VERSION_NUM: 1
- CURRENT_VERSION: Y

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

The screenshot shows the SQL Server Management Studio interface. The query editor contains the following SQL code:

```

SELECT ST.NAME, sum(SOH.TotalDue)
FROM Stage.dbo.SalesTerritory ST
INNER JOIN Stage.dbo.SalesPerson SP ON ST.TerritoryID = SP.TerritoryID
INNER JOIN Stage.dbo.Employee E ON SP.BusinessEntityID = E.BusinessEntityID
INNER JOIN Stage.dbo.Person C ON E.BusinessEntityID = C.BusinessEntityID
INNER JOIN Stage.dbo.SalesOrderHeader SOH ON SP.BusinessEntityID = SOH.SalesPersonID
where year(SOH.ShipDate) = 2007
GROUP BY ST.Name
ORDER BY ST.Name

exec TestFramework.dbo.SP_LOAD_Test_Results @DebugFlag = 'Y', @TestDefCD = 'TOT_SLS_2014_ALL_TER'
    
```

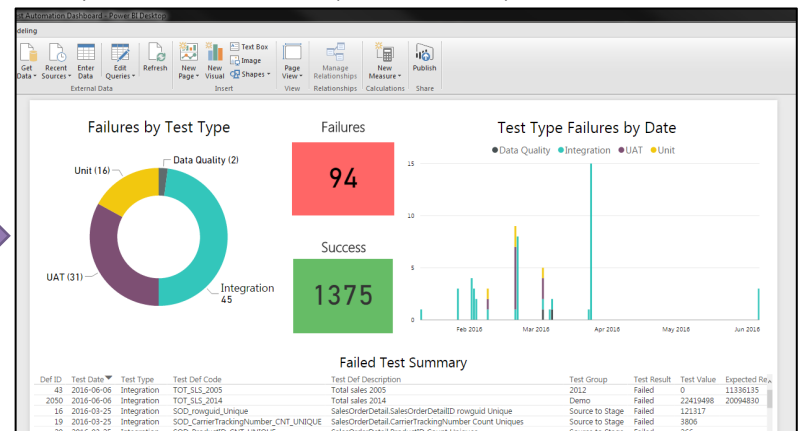
2b. Execute a batch of tests (PowerShell scripts)

The screenshot shows a Windows PowerShell terminal window with the following output:

```

<10 rows affected>
Loading EDW..D_SalesPerson
<17 rows affected>
Loading EDW..D_Employee
<298 rows affected>
Loading EDW..D_Person
<19972 rows affected>
Loading EDW..F_SalesOrderDetail
<121317 rows affected>
Loading EDW..P_SalesOrderHeader
<31465 rows affected>
Changed database context to 'EDW'.
Loading EDW..FA_SALES_ORDER_AGG
<3686 rows affected>
PS C:\Users\CH\Documents> sql Server Management Studio\testautomation) .\RunIntegTests.ps1
Changed database context to 'TestFramework'.
Running tests for type: Integration
Beginning Test SalesOrderDetail.SalesOrderDetailID OrderQty Not Null
***PASSED***
Beginning Test SalesOrderDetail.SalesOrderDetailID LineTotal Not Null
***PASSED***
Beginning Test SalesOrderDetail.SalesOrderDetailID UnitPrice Not Null
***PASSED***
Beginning Test SalesOrderDetail.SalesOrderDetailID UnitPriceDiscount Not Null
***PASSED***
Beginning Test SalesOrderDetail.SalesOrderDetailID rowguid Not Null
***PASSED***
Beginning Test SalesOrderDetail.SalesOrderDetailID rowguid Unique
***PASSED***
Beginning Test SalesOrderDetail check for orphaned records
***PASSED***
Beginning Test SalesOrderDetail.CarrierTrackingNumber Count Uniques
***PASSED***
Beginning Test SalesOrderDetail.ProductID Count Uniques
***PASSED***
Beginning Test Checksum on the customer table Test Stage to EDW
***PASSED***
    
```

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

Reference	Description
<p><i>Agile Analytics: A Value-Driven Approach to Business Intelligence and Data Warehousing</i> by Dr. Ken Collier</p>	Chapter 7: Test-Driven Data Warehouse Development
<p><i>Agile Data Warehousing for the Enterprise: A Guide for Solution Architects and Project Leaders</i> By Ralph Hughes</p>	Part V: Agile EDW Quality Management Planning
<p><i>Automated Data Warehouse Testing: Beginner's Step-by-Step Guide</i> by G. Suden</p>	Java-based test automation for simple data warehousing tests
<p><i>Agile Testing: A Practical Guide for Testers and Agile Teams</i> by Lisa Crispin and Janet Gregory</p>	Great introduction to agile testing for any type of development
<p><i>More Agile Testing: Learning Journeys for the Whole Team</i> by Lisa Crispin and Janet Gregory</p>	Includes several chapters on test automation as well as one dedicated to agile testing for DW/BI



DW/BI Test Automation Tools

Tool	Description
DBFit dbfit.github.io/dbfit/	Open source database testing tool
iCEDQ icedq.com QuerySurge QuerySurge.com Zuzena Zuzena.com	Test automation tools designed specifically for Data Warehousing and related projects.
Informatica Data Validation www.informatica.com/etl-testing	Accelerate and automate Informatica ETL testing in both production environments and dev/ test
Analytix Data Services analytixds.com WhereScape wherescape.com TimeXtender timeXtender.com	DW automation tools that include test automation capabilities.
Tricentis Tosca www.tricentis.com/tricentis-tosca-testsuite/	DW-friendly test automation tool



Questions?



Lynn Winterboer

Agile Analytics Educator & Coach

www.WinterboerAgileAnalytics.com
lynn@WinterboerAgileAnalytics.com

@AgileLynn



Cher Fox

DW/BI Solution Architect

www.foxconsulting.co
cfox@foxconsulting.co

@TheDatanista

Special thanks for contributing to this presentation:

Ken Collier

Brad Ewald

Deborah Krinitzsky

Joe Bernardini

