Data Validation and Automatic Testing

All source code on GitHub: <https://github.com/grahamcrowell/TableCompare>

Zipped Source Code Direct Link: <https://github.com/grahamcrowell/TableCompare/archive/master.zip>

# Purpose

To avoid Excel based data testing.

To standardize and consolidate the data validation efforts of actute and community.

* Automate generation and execution repetitive testing queries in order to focus on intrepreting test results rather their creatino
* Standardize testing strategies between acute and community by developing tools that accomplish the union of requirements rather than the intersectino

# Use Cases

## Validate Type 1 Biz Rule

Compare column values before and after a biz rule is executed and be able to see high level comparison results as well drill down to a sample of records.

### Current Solution

1. Export sample of records into Excel
2. Execute Bizrule
3. Export same records into Excel again
4. Compare two samples to validate behavior of a biz rule

### Potiential Solution

1. Backup entire table (where? Maybe a sample?)
2. Execute Bizrule
3. Compare backup version to new version row by row and column by column with SQL to get high level aggregates of differences

## Validate new data source for existing strucutres

Compare an upstream table to a downstream table. Counts (natural) key (mis)matches, dim value (mis)matches/nulls/duplicates.

# Requirements

* Use ETLBizRuleAuditFact to validate changes; count proportion of changes accounted for in ETLBizRuleAuditFact
* Save test parameters for future rerunning
* Auto guess PK if not specified

# TODO

* Log params to a new table in TestLog database to facilitate a rerunning a past test
  + Create PROC (input KeyComparisonID) to rerun a test
* Infer/store data lineage from MD\_Object