OwlDQ

Predictive Data Quality

Do you catch bad data? Or does bad data catch you by surprise?

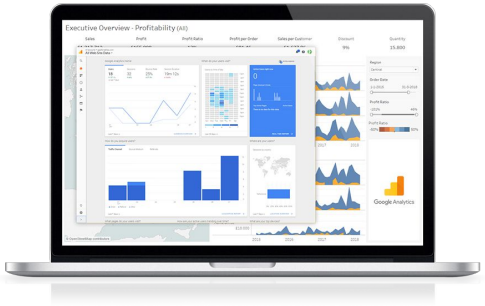
Spark's Scale **+** Owl's Data Science **=** Scalable Insights

Daily Reactive Scenario

The Model is not right!

Analyst

1



It’s hard to say what the model did, but I don’t think it has changed.

Data Scientist

2

I’m investigating the database table and checking the 300 data quality rules. We might have missed something.

Data Engineer

3

The Problem

10,000
rules

The average number of **rules** needed to be manually written and maintained with a traditional approach.

2,000
hrs

The number of **man hours** needed with a domain expert. Drafting rules, boundaries, and conditions that model the business.

30%
coverage

The amount of **coverage** and **trust** actually achieved with a conventional approach.

Predictive with Transparency

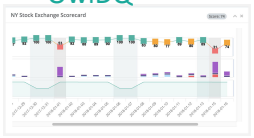
Data Engineer

Data Scientist

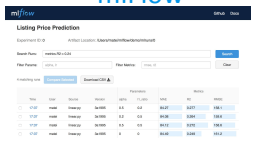
The Model isn't wrong, the data suddenly changed

Analyst

OwlDQ



mlFlow



Data Flow

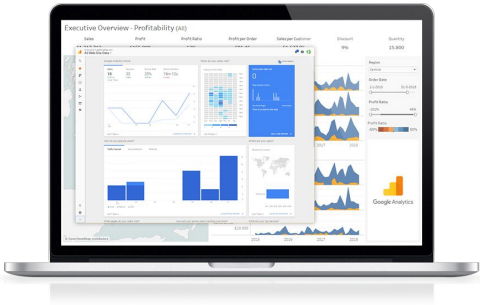
Dataset
Table/File/Kafka

→

Owl Scan
REST/CMDLine/Spark

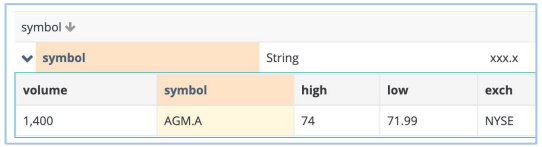
→

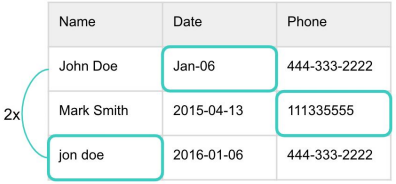
Reports
Models
Dashboards

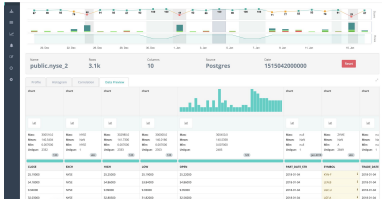


Data Loading | OwlDQ Suite | Trusted Analytics

A Complete DQ Suite







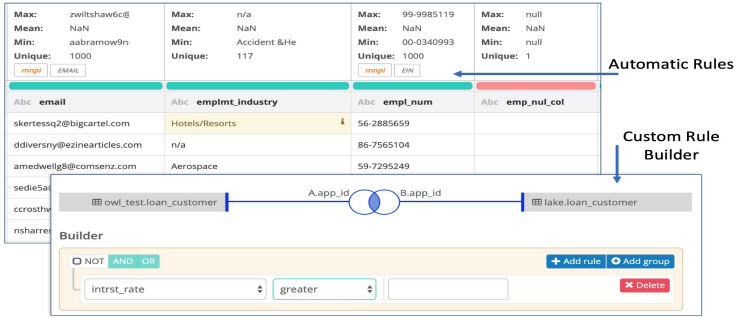
Profile & **Generate** thousand of rules

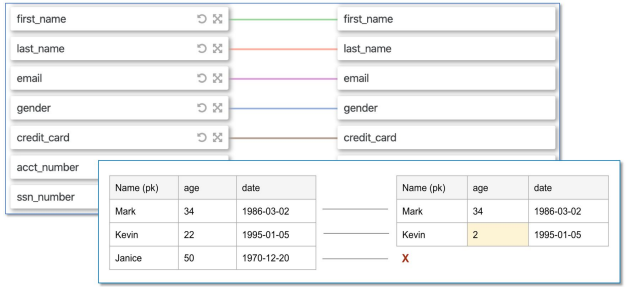
Schema Evolution & Data **Shape** issues

Reconciliation Source to Target

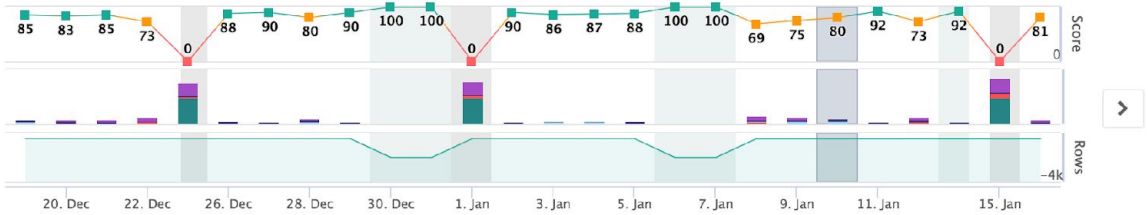
Duplicates, **Outliers**, **Pattern**

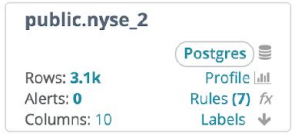
Discover data **Correlations**, **Segmentations** and **Behaviors**

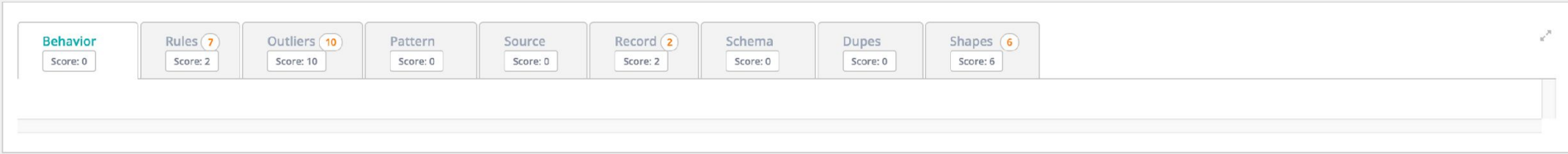




Owl automatically learns the patterns in your data, drill in for unsupervised discovery







Search: Copy CSV Excel PDF Print Whitespace OFF

CLOSE	EXCH	HIGH	LOW	OPEN	PART_DATE_STR	SYMBOL	TRADE_DATE	VOLUME	owl_id
Max: 308350.(Mean: 143.304 Min: 0.09000(Unique: 2397 123	Max: NYSE Mean: NaN Min: NYSE Unique: 1 123	Max: 308599.5 Mean: 143.788(Min: 0.10600(Unique: 2536 123	Max: 303930.(Mean: 141.509 Min: 0.09000(Unique: 2302 123	Max: 304625.(Mean: 142.198 Min: 0.09500(Unique: 2525 123	Max: null Mean: NaN Min: null Unique: 1 123	Max: ZYME Mean: NaN Min: A Unique: 2837 123	Max: null Mean: NaN Min: null Unique: 1 123	Max: 1075129 Mean: 1152843 Min: 0 Unique: 2533 123	Max: 3105 Mean: 1553.(Min: 1 Unique: 2834
1.60000	NYSE	2.95000	1.60000	2.35000	2018-01-10	KOD.W	2018-01-10	1190400	1628
21.01000	NYSE	21.75000	20.50000	21.28000	2018-01-10	AFS-A	2018-01-10	59600	58
21.07000	NYSE	21.33000	20.89000	21.25000	2018-01-10	AFS-F	2018-01-10	152500	63
22.18000	NYSE	22.56000	22.08000	22.30000	2018-01-10	AFS-B	2018-01-10	13500	59

OwlDQ Helps Companies to...

AI

Reach your AI goals within weeks by applying OwlDQ's proven machine learning approach that is purpose-built to find erroneous data.

Save

Achieve costs savings and lessen DQ efforts up to 70% after integrating OwlDQ's auto-discovery and adaptive rules.

Gain

Gain end-user trust, unified coverage, team collaboration through feedback loops, and derive deeper insights.