

Latte-Q Cheat Sheet

Qualified data brings you a latte!



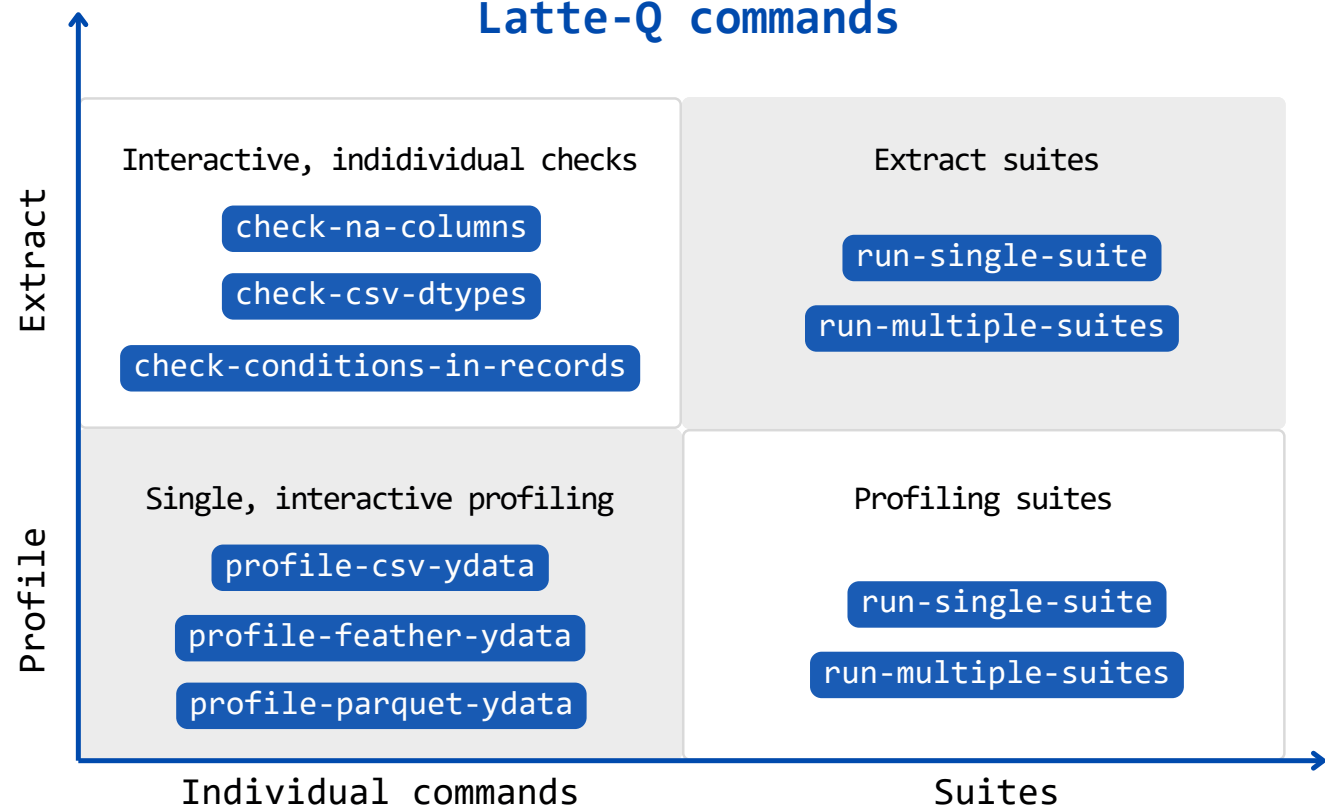
Linux Python>=3.9 Tabular Data Files

> Hướng dẫn cài đặt

- Tạo và sử dụng môi trường ảo (virtual environment) khi chạy Latte-q
- Chạy lệnh cài đặt/ cập nhật Latte-Q

```
pip install latteq-0.0.1-py3-none-any.whl --force-reinstall
```

Latte-Q commands



> Profiling: thống kê chất lượng trực quan

`profile-csv-ydata` `profile-feather-ydata`
`profile-parquet-ydata`

Trực quan hóa profile về dữ liệu thông qua tích hợp với *Ydata Profiling*

Alerts	Severity
RP_DATE has constant value "2024-04-10"	Constant
TRANG_THAI is highly imbalanced (51.2%)	Imbalance
DUNG_LUONG_BINH_ACCU has 211 (1.3%) missing values	Missing
THOI_GIAN_BACKUP has 236 (1.5%) missing values	Missing
SO_LUONG_BINH_ACCU has 237 (1.5%) missing values	Missing
SITE_ID has 229 (1.4%) missing values	Missing
KHO has 15863 (98.6%) missing values	Missing
SERIAL_ACCU has 863 (5.4%) missing values	Missing
PRODUCT_CODE_ACCU has 286 (1.8%) missing values	Missing
THOI_HAN_BAO_HANH has 1146 (7.1%) missing values	Missing
THOI_HAN_BAO_DUONG has 858 (5.3%) missing values	Missing
MA_QR has 1480 (9.2%) missing values	Missing
OBJ_ID has unique values	Unique
THOI_GIAN_BACKUP has 3263 (20.3%) zeros	Zeros

`run-single-suite` `run-multiple-suites`

Chạy nhiều tác vụ cho nhiều nguồn dữ liệu kiểm tra trực quan hóa thống kê kết quả, tích hợp *Great Expectations*

- `profile_column_conditions_not_null`
- `profile_column_conditions_in_set`
- `profile_column_conditions_between`
- `profile_column_of_type`

> Sử dụng --help cho hướng dẫn chạy lệnh

```
$latteq --help  
$latteq commandname --help
```

> Extract: trích xuất dữ liệu theo chất lượng

`check-na-columns`

Trích xuất các dữ liệu NA (missing data)

`check-csv-dtypes`

Trích xuất các dữ liệu sai định dạng (wrong data type)

`check-conditions-in-records`

Trích xuất các dữ liệu theo các điều kiện định cho các trường

`run-single-suite` `run-multiple-suites`

Chạy nhiều tác vụ kiểm tra dữ liệu đồng thời và trích xuất, lưu trữ các kết quả kiểm tra cho nhiều nguồn dữ liệu

- `extract_records_by_conditions`
- `extract_records_na_columns`
- `extract_duplicated_records`
- `extract_outlier_records`
- `extract_unmatched_pattern_records`

> Convert: chuyển đổi định dạng file

```
excel-to-csv  
excel-to-feather  
excel-to-parquet
```

> Contact us

- Website: www.daienso.com
- Email: contact@daienso.com