
Splicer Documentation

Release 0

triv.io

December 31, 2014

Contents

`splicer`, the In Place Data Querying Library, is a python module for working with data from disparate sources using commands to those familiar with SQL. It aims to make quick one off queries and *etl_def* scripts more declarative rather than procedural.

`splicer` enables:

- Analysts to create *dataset_def* linking various foreign tables together along with *udf_def*. Once defined the datasets can be queried via SQL Select statements to create new Views of the Data.
- Extension Developers to create extensions that make various data sources REST endpoints, log files, NoSQL Servers, traditional Databases, CSV Files to behave like tables.

`splicer` will take advantage of these various sources capabilities where appropriate and will compensate for sources that lack basic features.

For example if a database supports joins and you want to query two tables within that database, `splicer` will have that system perform the join for you. If however you work with a less sophisticated source, like a plain file, `splicer` will perform the operations for you locally.

Enough reading! *Try it out!*