Broadening the Semantic Web

Bryon Jacob - data.world

Briefest possible agenda

- What is data.world?
- What are our DB architecture challenges?
  - * and how do we address them?
- Where are we going next?

data.world

Search data.world

+ Add

Bryon Jacob
@bryon
View your profile >

Home

Library   473

Bookmarks   101

Requests   0

ORGANIZATIONS

Asymptote Analytics

BedandBreakfast.com

BKJ PLM DEMO

EXPLORE

Features Overview

Tutorials & Docs

data.world Forum

Who to follow

Integrations

Invite people you know

Recent   Featured   Feed

sya is now a contributor to **Chinese Provinces** by **ddw-ontology-team**.

Chinese Provinces   🔒 PRIVATE
Dataset • Updated Jun 6

@ddw-ontology-team   🔖 Bookmark   💬 Comment

🔥 Trending this week

2018/W23: U.K. Gender Pay Gap   🕐 OPEN
U.K. Gender Pay Gap
Dataset • Updated Jun 4

united kingdom, **gender**, **pay**, **gap**, **companies**, **makeover monday**, **hmrc**, **valuation office agency**, **voa**

@makeovermonday   🔖 108   💬 Comment

sya is now a contributor to **Russian Provinces** by **ddw-ontology-team**.

Russian Provinces   🔒 PRIVATE
Dataset • Updated Jun 6

@ddw-ontology-team   🔖 Bookmark   💬 Comment

RECENTLY UPDATED (473)   +

🔒 picturengine/Test 2

🔒 synthesis-data/IMF Country Informa...

🔒 bkj-plm-demo/Clean Data 2017-11

🔒 synthesis-data/Composer

🔒 synthesis-data/Products

Show 468 more

TOP TAGS

🏷 hxl   1029

🏷 education   1046

🏷 porites   1519

🏷 cso   3141

🏷 pocillopora   1709

RECENT DISCUSSIONS

2018/W23: U.K. Gender Pay Gap   41
2018/W23: U.K. Gender Pay Gap

Viz Review   30
2018/W23: U.K. Gender Pay Gap

Help/Questions   13
2018/W23: U.K. Gender Pay Gap

2000-2016 US County Eviction Esti...   1
2000-2016 US County Eviction Estimates

Significant Volcanic Eruptions   1
Significant Volcanic Eruptions

https://data.world/sya

Search data.world

linked-data/linkedmdb   🔄 OPEN
Updated Oct 17, 2017 · Version: fad38758 · Other License

+ Add to project   ⬇   ▦   **Launch workspace ▸**

Overview   Contributors 1   Discussion   Activity   Settings

🔖 32

A linked-data collection of movies, actors, directors, and the relationships between.

Edit

Linked Movie Database

Edit

An RDF graph containing a linked-data collection of movies, actors, directors, and the relationships between all of those entities.

This dataset contains a copy of the Linked Movie Database synced from this URL:

http://www.cs.toronto.edu/~oktie/linkedmdb/linkedmdb-18-05-2009-dump.nt

This data is the work of Oktie Hassanzadeh and Mariano P. Consens

Querying the data with SPARQL, you can uncover some interesting facts:

Cast of Pulp Fiction and Number of Movies Acted In

⌄ Show more

☁ **Drag and drop or connect to a data source.**
You can add up to 595.78 MB to this dataset.

Add data

QUERIES (9)   ❯

▪ Six degrees of Kevin Bacon via "Daze...
🔒 @bryon · last year

▪ Six degrees of Kevin Bacon (shorthan...
🔒 @bryon · last year

▪ Six degrees of Kevin Bacon via "Daze...
🔒 @bryon · last year

▪ Six degrees of Kevin Bacon
🔒 @bryon · last year

▪ Actors who've worked with Kevin Ba...
🔒 @bryon · last year

TAGS (2)   ⚙

movies   linked data

PROJECTS USING THIS DATASET (2)

my_love_data
@songshu · 5 months ago

exploring
@seafella · 7 months ago

CONTRIBUTORS (4)   ⚙

Add all Linked Data members

1 file

Sort ▾

linkedmdb-18-05-2009-dump.nt
Add a description

TOTAL TRIPLES
3,579,616

DISTINCT ENTITIES
> 1,000

DISTINCT SUBJECTS
> 1,000

Namespaces

8 COMMON
rdf, rdfs, foaf, owl, purl, laabs, dbpe...

> 1,000 CUSTOM
http://... , http://... **+ more**

41 Distinct Classes

148 Distinct Predicates

MOST COMMON
rdf:type (438151)

NEXT MOST COMMON
rdfs:label (395549)

OTHER FILES (1)   Hide

linkedmdb-18-05-2009-dump.nt

Six Degrees of SPARQL 🔄 SHARED

Save a copy   ▶ Run query   >

DOCUMENTS (2)   Hide

Dataset summary

Data dictionary

Dataset schema

🔍 Search the schema

QUERIES (9)   Hide

New query ▾

Six degrees of Kevin Bacon via "D...
🔒 @bryon · last year

Cast of Pulp Fiction and Number ...
🕐 @linked-data · last year

Six Degrees of SPARQL
🕐 @linked-data · last year

Six degrees of Kevin Bacon (short...
🔒 @bryon · last year

Six degrees of Kevin Bacon via "D...
🔒 @bryon · last year

Six degrees of Kevin Bacon
🔒 @bryon · last year

SPARQL Query 1
🔒 @bryon · last year

SPARQL Query 2
🔒 @bryon · last year

Actors who've worked with Kevin...
🔒 @bryon · last year

```
 1  PREFIX linkedmdb: <http://data.linkedmdb.org/movie/>
 2  PREFIX dc: <http://purl.org/dc/terms/>
 3
 4  SELECT ?movie1 ?actor1 ?movie2 ?actor2 ?movie3 ?actor3 ?movie4 ?actor4 ?movie5
 5  WHERE {
 6    # select the source and target nodes
 7    ?s linkedmdb:actor_name "John Goodman" .
 8    ?t linkedmdb:actor_name "Kevin Bacon" .
 9
10    # find the five movies and the connecting actors between (make sure to filter out dupes)
11    ?m1 dc:title ?movie1 ; linkedmdb:actor ?s ; linkedmdb:actor ?a1 .
12    FILTER(?s != ?a1 && ?t != ?a1)
13
14    ?m2 dc:title ?movie2 ; linkedmdb:actor ?a1 ; linkedmdb:actor ?a2 .
15    FILTER(?m1 != ?m2)
16    FILTER(?a1 != ?a2)
17    FILTER(?s != ?a2 && ?t != ?a2)
18
```

⊞ Results table   📊 Chart   1 query result   ⬇ Download ▾   ⊞ Open in app

movie1 ▾	actor1 ▾	movie2 ▾	actor2 ▾	movie3 ▾	actor3 ▾	movie4 ▾	
1	Dirty Work	Artie Lange	Beer League	Anthony DeSando	New Jack City	Allen Payne	The Tuskegee Airmen	Laur

Confidential

5 files

Sort ▾

beers.csv
Request more info

Explore

	id	brewery_id	name	cat_id	style_id	abv	ibu
1	1	812	Hocus Pocus	11	116	4.5	0
2	2	264	Grimbergen Blonde	−1	−1	6.7	0
3	3	779	Widdershins Barleywine	−1	−1	9.1	0
4	4	287	Lucifer	−1	−1	8.5	0
5	5	1056	Bitter	−1	−1	4	0

Showing 1-5 of 5,845 rows, 12 columns See all

Switch to column overview

breweries.csv
Request more info

Explore

Type	Column name	Empty	Distinct	Distribution	
	city ⓘ	22 (2%)	987	SAMPLE VALUE Aberdeen	esk Budjovice
	state ⓘ	182 (13%)	167	SAMPLE VALUE Alaska	Wyoming
	code ⓘ	635 (45%)	653	SAMPLE VALUE 953-0076	YO1 6JT

Showing 1-5 of 1,414 rows, 13 columns See all

Switch to data preview

breweries.csv

Request more info

Explore

Type	Column name	Empty	Distinct	Distribution

	city ⓘ			esk Budjovice
	state ⓘ			Wyoming
	code ⓘ			YO1 6JT

⚡ Showing 1-5 of 1,414 rows, 13 columns See all

state   ✕

T	string

SAMPLE VALUES

Alaska
Wyoming

DISTINCT	NON-EMPTY	EMPTY
167	1,232	182 (12.87%)

MIN LENGTH	MAX LENGTH	MEAN LENGTH
2	22	9.08

TABLE
breweries

📋 Copy column IRI to clipboard

⇄ Switch to data preview

breweries_geocode.csv

Request more info

Explore

# id	# brewery_id	# latitud		accuracy
1	1	1	3 ...4)	ROOFTOP
2	2	2	3 ...5)	ROOFTOP
3	3	3	50.7668	4.3081 POINT(4.3081 50.7668)	RANGE_INTERPOLATED

data.world

Mean ABV by Style 🔒 PRIVATE

Save ⋮ ▶ Run query

```
1   # Mean ABV by Style
2   PREFIX : <https://too-root.linked.data.world/d/openbeerdb/>
3   PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
4
5   SELECT ?style_name (AVG(xsd:float(?_abv)) AS ?abv) (COUNT(*) AS ?count) WHERE {
6     [ :col-beers-abv ?_abv ; :col-beers-style_id ?style_id ] .
7     FILTER(REGEX(?_abv, '^[0-9]*[.]?[0-9]*$'))
8     BIND(xsd:integer(?style_id) AS ?join_style)
9     [ :col-styles-id ?join_style ; :col-styles-style_name ?style_name ] .
10  }
11  GROUP BY ?style_name
12  ORDER BY DESC(?abv)
13
```

Dataset schema

🔍 Search the schema

▾ beers
- Ⓣ id ⓘ
- Ⓣ brewery_id ⓘ
- Ⓣ name ⓘ
- Ⓣ cat_id ⓘ
- Ⓣ style_id ⓘ
- Ⓣ abv ⓘ
- Ⓣ ibu ⓘ
- Ⓣ srm ⓘ
- Ⓣ upc ⓘ
- Ⓣ filepath ⓘ
- Ⓣ descript ⓘ
- Ⓣ last_mod ⓘ

▾ breweries
- # id ⓘ
- Ⓣ name ⓘ
- Ⓣ address1 ⓘ
- Ⓣ address2 ⓘ
- Ⓣ city ⓘ
- Ⓣ state ⓘ
- Ⓣ code ⓘ
- Ⓣ country ⓘ
- Ⓣ phone ⓘ

Documents (2)   Hide
- Dataset summary
- Data dictionary

Queries (8)   Hide

New query ▾

- ABV
  🔒 @bryon · 2 minutes ago
- WeMap breweries demo
  🔒 @bryon · 23 days ago
- SQL Query 2
  🔒 @bryon · last year
- SQL Query 1
  🔒 @bryon · last year
- Mean ABV by Style
  🔒 @bryon · 2 years ago
- ABV by Style
  🔒 @bryon · 2 years ago
- New query 1
  🔒 @bryon · 2 years ago
- beers

Files
- beers.csv
- breweries.csv
- breweries_geocode.csv
- categories.csv
- styles.csv

📋 Results table   📊 Chart

68 query results   ⬇ Download   Open in app

	style_name	abv	count
1	Belgian-Style Quadrupel	11.44	10
2	Strong Ale	11.25	2
3	American-Style Imperial Stout	10.212	75
4	American-Style Barley Wine Ale	9.8677	39
5	Imperial or Double Red Ale	9.2	1
6	Belgian-Style Dark Strong Ale	9.1857	14
7	Specialty Beer	8.75	1
8	American-Style India Black Ale	8.7	1
9	Belgian-Style Pale Strong Ale	8.5133	15

Mean ABV by Style (SQL) 🔒 PRIVATE

Save ⋮   ▶ Run query

```sql
1  SELECT style_name, AVG(CAST(beers.abv AS DECIMAL)) AS avg_abv, COUNT(*) AS count
2    FROM beers
3    JOIN styles
4      ON CAST(beers.style_id AS INTEGER) = styles.id
5  GROUP BY style_name
6  ORDER BY avg_abv DESC
```

DOCUMENTS (2)   Hide

- 📄 Dataset summary
- 📄 Data dictionary

QUERIES (9)   Hide

New query ▾

- **Mean ABV by Style (SQL)**
  🔒 @bryon · 40 seconds ago
- ABV
  🔒 @bryon · 8 minutes ago
- WeMap breweries demo
  🔒 @bryon · 23 days ago
- SQL Query 2
  🔒 @bryon · last year
- SQL Query 1
  🔒 @bryon · last year
- Mean ABV by Style
  🔒 @bryon · 2 years ago
- ABV by Style
  🔒 @bryon · 2 years ago

📋 Results table   📊 Chart

68 query results   ⬇ Download ▾   Open in app

	style_name	# avg_abv	# count
1	Belgian-Style Quadrupel	11.44	10
2	Strong Ale	11.25	2
3	American-Style Imperial Stout	10.212	75
4	American-Style Barley Wine Ale	9.8677	39
5	Imperial or Double Red Ale	9.2	1
6	Belgian-Style Dark Strong Ale	9.1857	14
7	Specialty Beer	8.75	1
8	American-Style India Black Ale	8.7	1
9	Belgian-Style Pale Strong Ale	8.5133	15
10	Imperial or Double India Pale Ale	8.4667	99
11	Belgian-Style Tripel	8.4084	43
12	American-Style Strong Pale Ale	8.3786	14
13	Foreign (Export)-Style Stout	7.5	1

Dataset schema

🔍 Search the schema

∨ beers
- 🔤 id ⓘ
- 🔤 brewery_id ⓘ
- 🔤 name ⓘ
- 🔤 cat_id ⓘ
- 🔤 style_id ⓘ
- 🔤 abv ⓘ
- 🔤 ibu ⓘ
- 🔤 srm ⓘ
- 🔤 upc ⓘ
- 🔤 filepath ⓘ
- 🔤 descript ⓘ
- 🔤 last_mod ⓘ

∨ breweries
- # id ⓘ
- 🔤 name ⓘ
- 🔤 address1 ⓘ
- 🔤 address2 ⓘ
- 🔤 city ⓘ
- 🔤 state ⓘ
- 🔤 code ⓘ
- 🔤 country ⓘ
- 🔤 phone ⓘ

- 📄 beers.csv
- 📄 breweries.csv
- 📄 breweries_geocode.csv
- 📄 categories.csv
- 📄 styles.csv

https://data.world/too-root/openbeerdb/workspace/query?queryid=40411bf2-514e-4a67-a1b9-f799

Mean ABV by Country 🔒 PRIVATE

Save ⋮ ▶ Run query ›

```
1   # Mean ABV by Country
2   PREFIX : <https://too-root.linked.data.world/d/openbeerdb/>
3   PREFIX c: <https://markmarkoh.linked.data.world/d/metadata-by-country/>
4   PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5
6   SELECT ?cc (AVG(xsd:float(?_abv)) AS ?abv) (COUNT(*) AS ?count)
7   FROM NAMED <https://data.world/markmarkoh/metadata-by-country/>
8   WHERE {
9     [ :col-beers-abv ?_abv ; :col-beers-brewery_id ?brewery_id ] .
10    FILTER(REGEX(?_abv, '^[0-9]*[.]?[0-9]*$'))
11    BIND(xsd:integer(?brewery_id) AS ?join_brewery)
12    [ :col-breweries-id ?join_brewery ;  :col-breweries-country  ?country_name ] .
13    GRAPH <https://data.world/markmarkoh/metadata-by-country/> {
14      [ c:col-countries-name ?country_name ; c:col-countries-iso_a2 ?cc ]
15    }
16  }
17  GROUP BY ?cc
18  ORDER BY DESC(?abv)
19
```

Dataset schema

🔍 Search the schema

∨ beers
- 🔲 id ⓘ
- 🔲 brewery_id ⓘ
- 🔲 name ⓘ
- 🔲 cat_id ⓘ
- 🔲 style_id ⓘ
- 🔲 abv ⓘ
- 🔲 ibu ⓘ
- 🔲 srm ⓘ
- 🔲 upc ⓘ
- 🔲 filepath ⓘ
- 🔲 descript ⓘ
- 🔲 last_mod ⓘ

∨ breweries
- # id ⓘ
- 🔲 name ⓘ
- 🔲 address1 ⓘ
- 🔲 address2 ⓘ
- 🔲 city ⓘ
- 🔲 state ⓘ
- 🔲 code ⓘ
- 🔲 country ⓘ
- 🔲 phone ⓘ

DOCUMENTS (2) Hide
- 📄 Dataset summary
- 📄 Data dictionary

QUERIES (10) Hide

New query ▾

- ▶ Mean ABV by Country
  🔒 @bryon · 7 minutes ago
- ▶ Mean ABV by Style (SQL)
  🔒 @bryon · 12 minutes ago
- ▶ ABV
  🔒 @bryon · 20 minutes ago
- ▶ WeMap breweries demo
  🔒 @bryon · 23 days ago
- ▶ SQL Query 2
  🔒 @bryon · last year
- ▶ SQL Query 1
  🔒 @bryon · last year
- ▶ Mean ABV by Style
  🔒 @bryon · 2 years ago
- ▶ ABV by Style
  🔒 @bryon · 2 years ago
- ▶ New query 1
  🔒 @bryon · 2 years ago
- ▶ beers

📊 Results table   📊 Chart   49 query results   ⬇ Download ▾   🔲 Open in app

	cc	abv	count
1	LK	8.4	2
2	EE	7.8	1
3	DK	7.0667	6
4	AR	6.5	4
5	PL	6.1333	9
6	LV	5.7333	3
7	LT	5.7	4
8	HU	5.575	4

styles.csv

data.world

SQL **remote join** 🔒 PRIVATE   Save ⋮   ▶ Run query   ›

```
1  SELECT c.iso_a2 AS cc, breweries.name
2    FROM breweries
3    JOIN markmarkoh.metadata_by_country.countries c
4      ON breweries.country = c.name
5
```

DOCUMENTS (2)   Hide
📄 Dataset summary
📄 Data dictionary

QUERIES (12)   Hide

New query ▾

▸ remote join
   🔒 @bryon · 4 minutes ago

Mean ABV by Country (SQL)
   🔒 @bryon · 10 minutes ago

Mean ABV by Country
   🔒 @bryon · 18 minutes ago

Mean ABV by Style (SQL)
   🔒 @bryon · 23 minutes ago

ABV
   🔒 @bryon · 31 minutes ago

WeMap breweries demo
   🔒 @bryon · 23 days ago

SQL Query 2
   🔒 @bryon · last year

SQL Query 1
   🔒 @bryon · last year

Mean ABV by Style
   🔒 @bryon · 2 years ago

ABV by Style

▦ Results table   📊 Chart   1,381 query results   ⬇ Download ▾   ▦ Open in app

	cc	name
1	US	(512) Brewing Company
2	US	21st Amendment Brewery Cafe
3	GB	Adnams & Co.
4	DE	Berliner Kindl Brauerei AG
5	JP	Pokka Beer Works
6	US	Pony Express Brewing
7	US	Port Brewing Company
8	US	Port Washington Brewing
9	US	Porterhouse Restaurant and Brewpub
10	US	Portland Brewing
11	US	Portsmouth Brewery
12	US	Prairie Rock Brewing - Elgin
13	US	Prescott Brewing Company

Dataset schema

🔍 Search the schema

▾ beers
   ⊤ id ⓘ
   ⊤ brewery_id ⓘ
   ⊤ name ⓘ
   ⊤ cat_id ⓘ
   ⊤ style_id ⓘ
   ⊤ abv ⓘ
   ⊤ ibu ⓘ
   ⊤ srm ⓘ
   ⊤ upc ⓘ
   ⊤ filepath ⓘ
   ⊤ descript ⓘ
   ⊤ last_mod ⓘ

▾ breweries
   # id ⓘ
   ⊤ name ⓘ
   ⊤ address1 ⓘ
   ⊤ address2 ⓘ
   ⊤ city ⓘ
   ⊤ state ⓘ
   ⊤ code ⓘ
   ⊤ country ⓘ
   ⊤ phone ⓘ

🦉 data.world

CSVW Namespace Vocabulary Terms

Latest editor's draft:
http://w3c.github.io/csvw/ns/

Editor:
Gregg Kellogg, Kellogg Associates

Repository:
We are on GitHub
File a bug

Changes:
Commit history

This document is also available in these non-normative formats: Turtle and JSON-LD

Abstract

This document describes the CSVW Namespace Vocabulary Terms and Term definitions used for creating Meta-data descriptions for Tabular Data. This document provides the RDFS [RDF-SCHEMA] vocabulary definition for terms defined in [tabular-metadata] and a description of the JSON-LD context definition for use with defining metadata documents.

Alternate versions of the vocabulary definition exist in Turtle and JSON-LD, which also includes the @context required for metadata descriptions. These versions may also be retrieved from http://www.w3.org/ns/csvw using an appropiate HTTP *Accept* header.

Published:
2017-05-23

HDT – Your binary format for RDF

"HDT compresses big RDF datasets while maintaining search operations" (Read more)

What is HDT

Discover HDT (Header, Dictionary, Triples), a compact data structure and binary serialization format for RDF

RDF
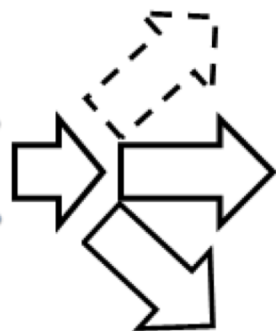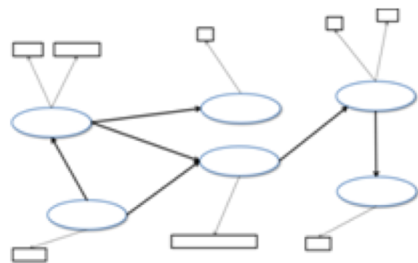
Header — Logical and physical metadata describing the RDF data set. It serves as an entrance point to the information.
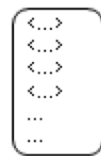
Dictionary — Mapping between elements in the data set and unique IDs, thus contributing to compactness.

<...>
<...>
<...>
<...>
...
...

Triples — Structure of the data after the ID replacement, in a compressed form.
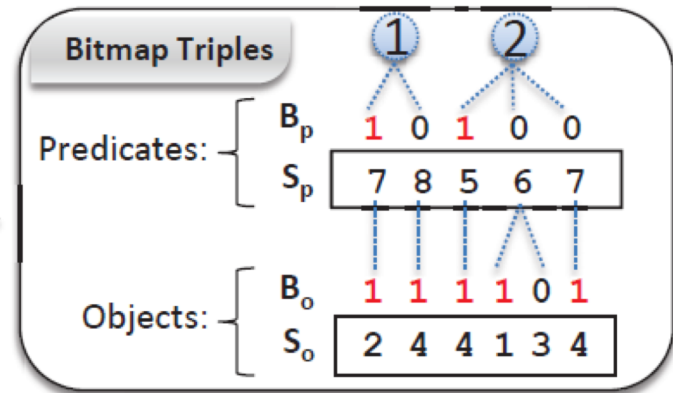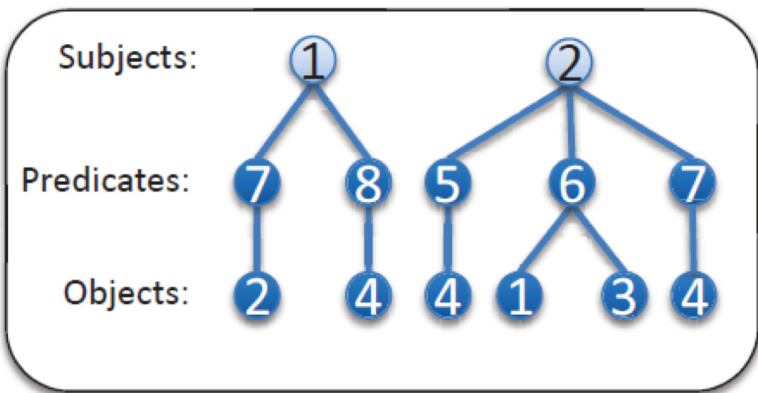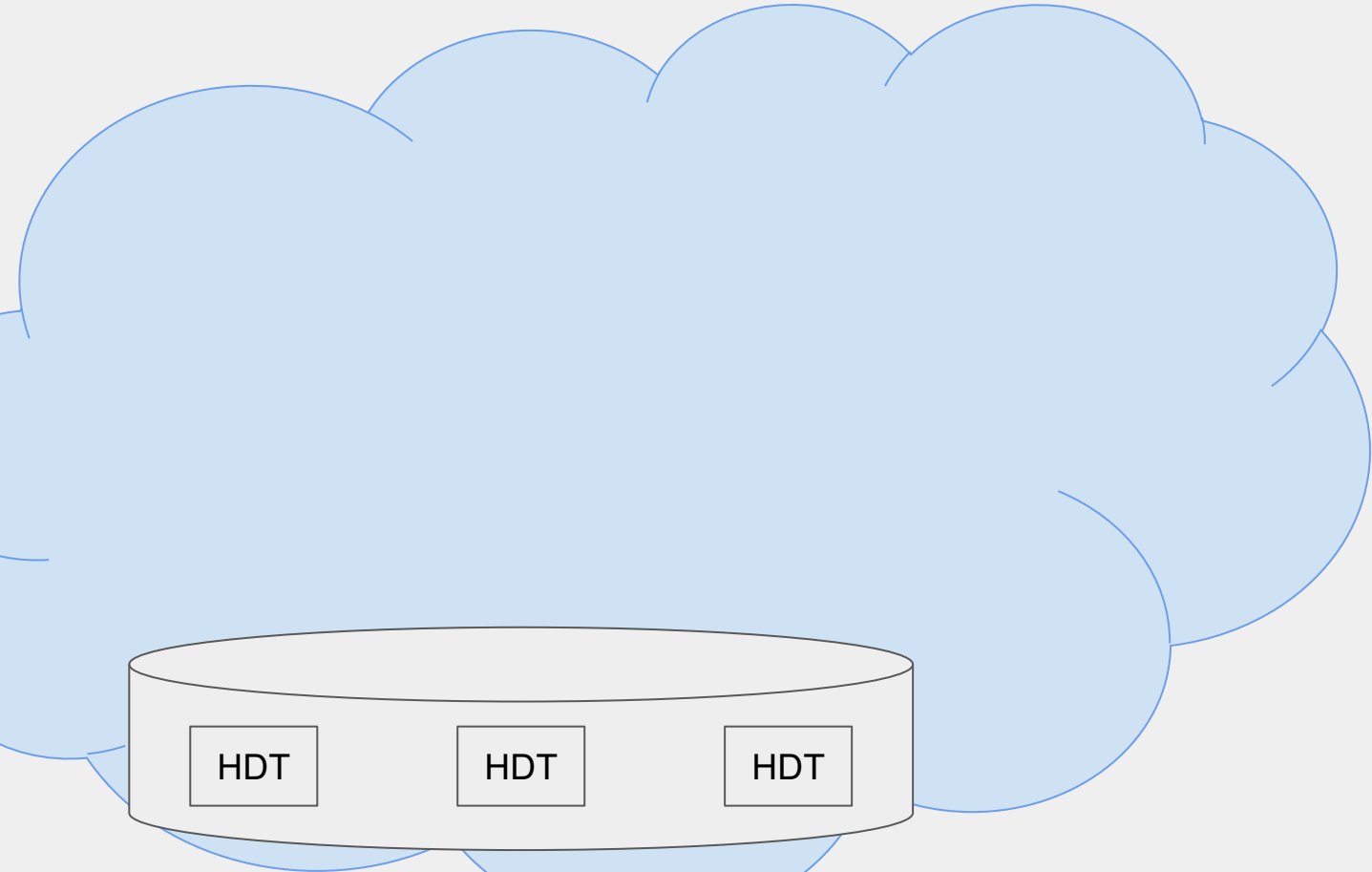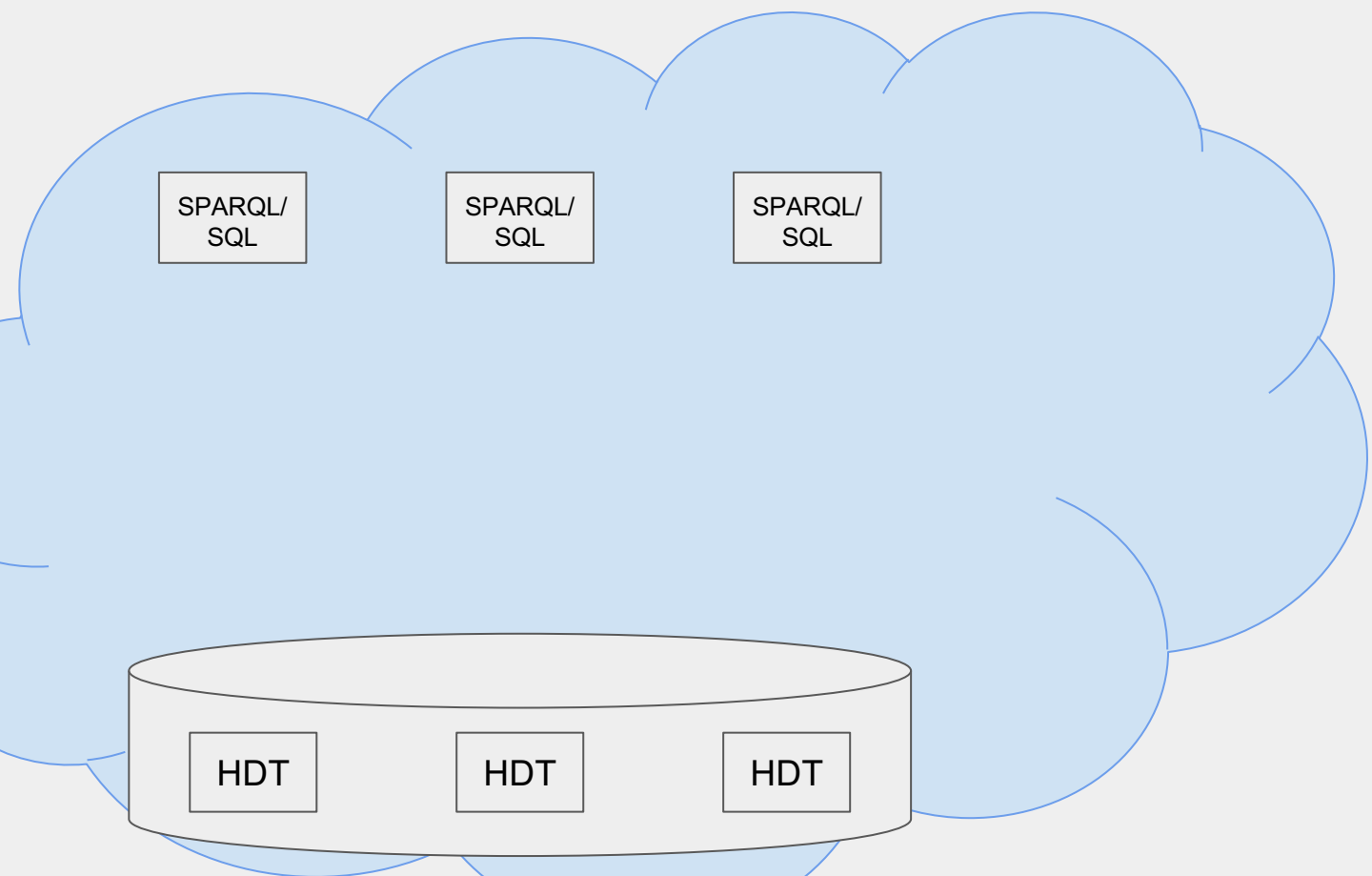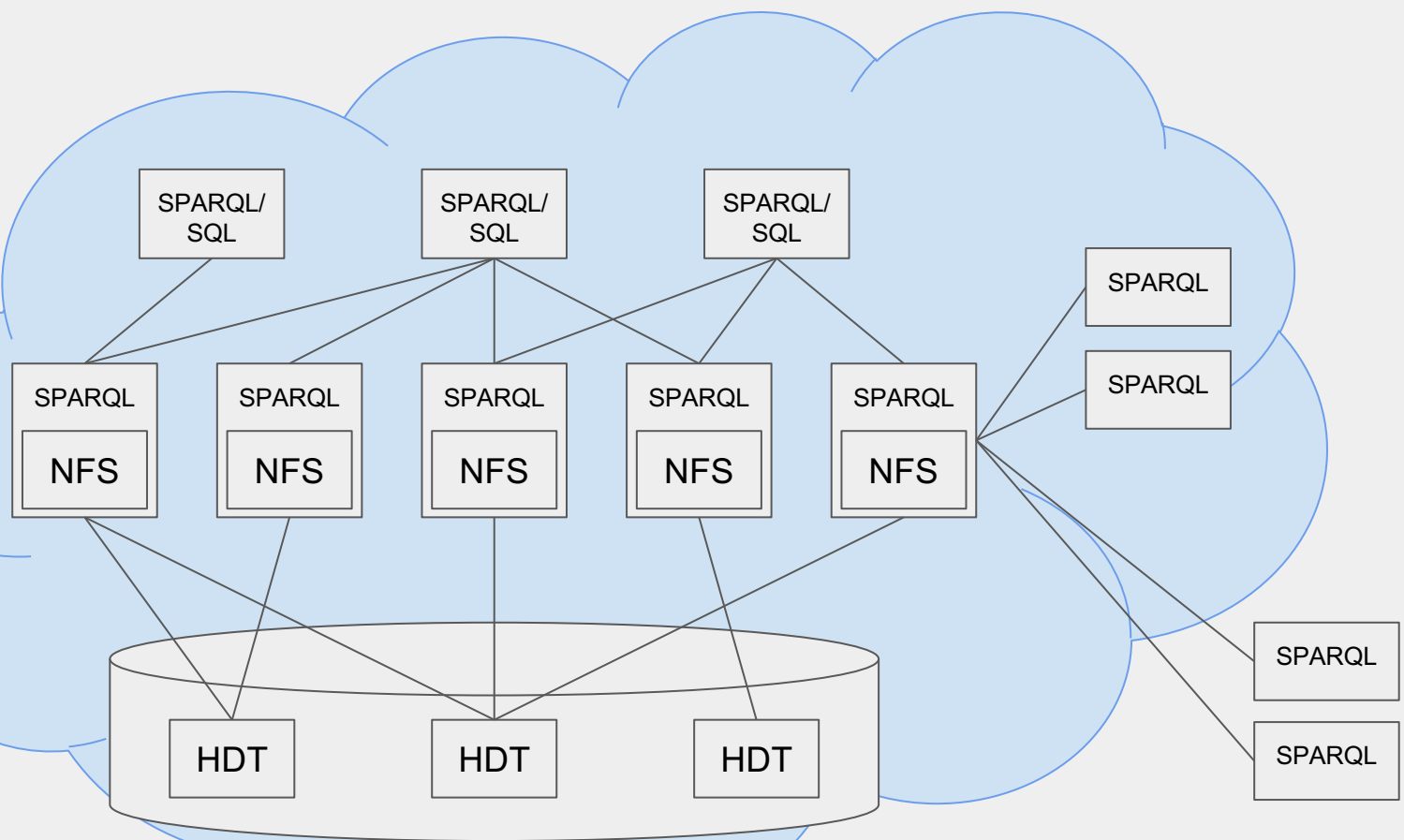
HDT   HDT   HDT

SPARQL/ SQL   SPARQL/ SQL   SPARQL/ SQL
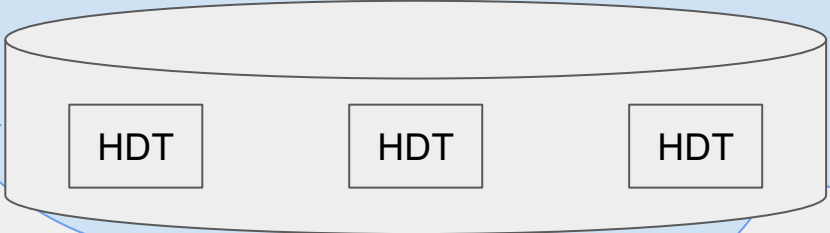
HDT   HDT   HDT
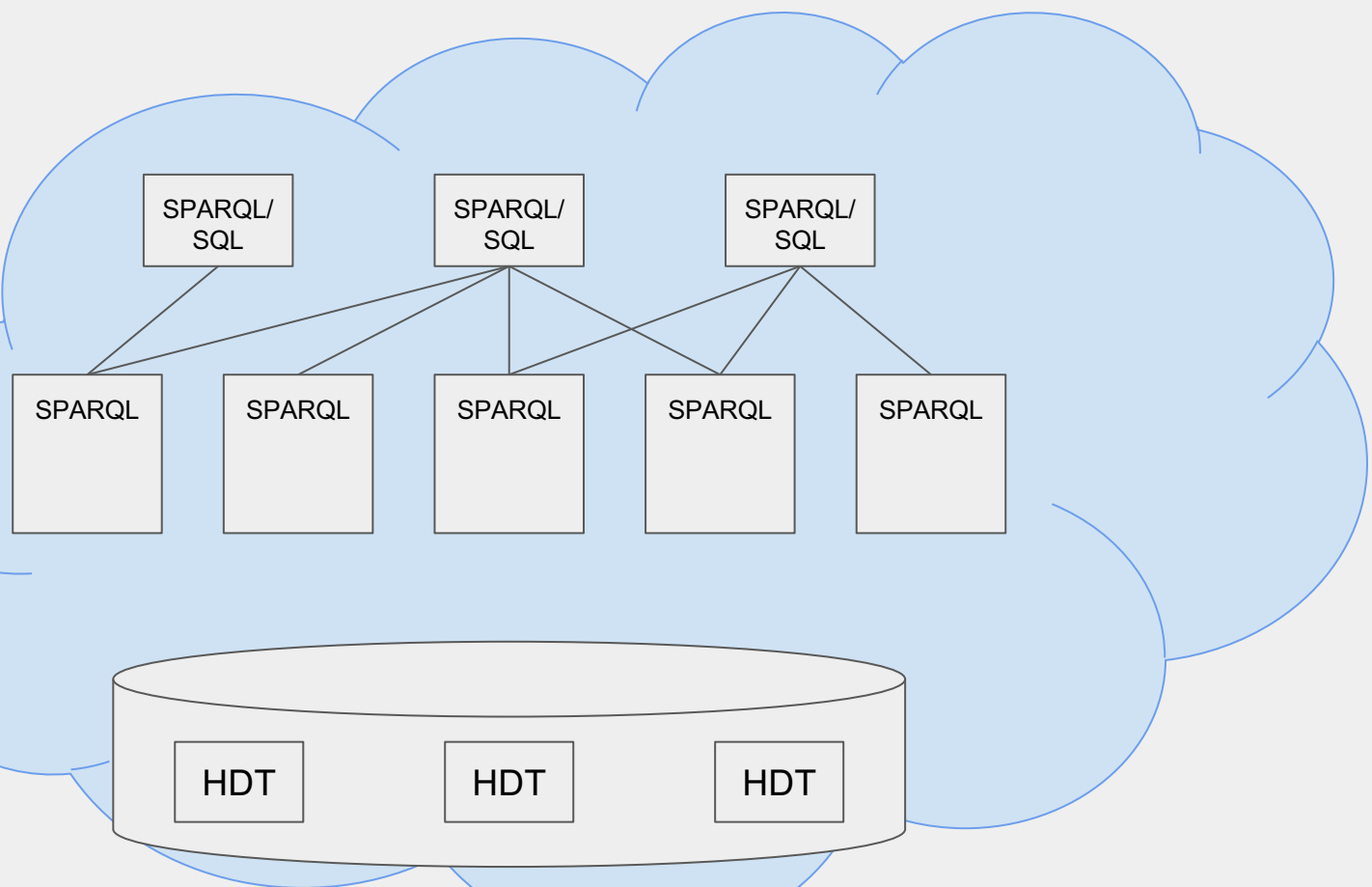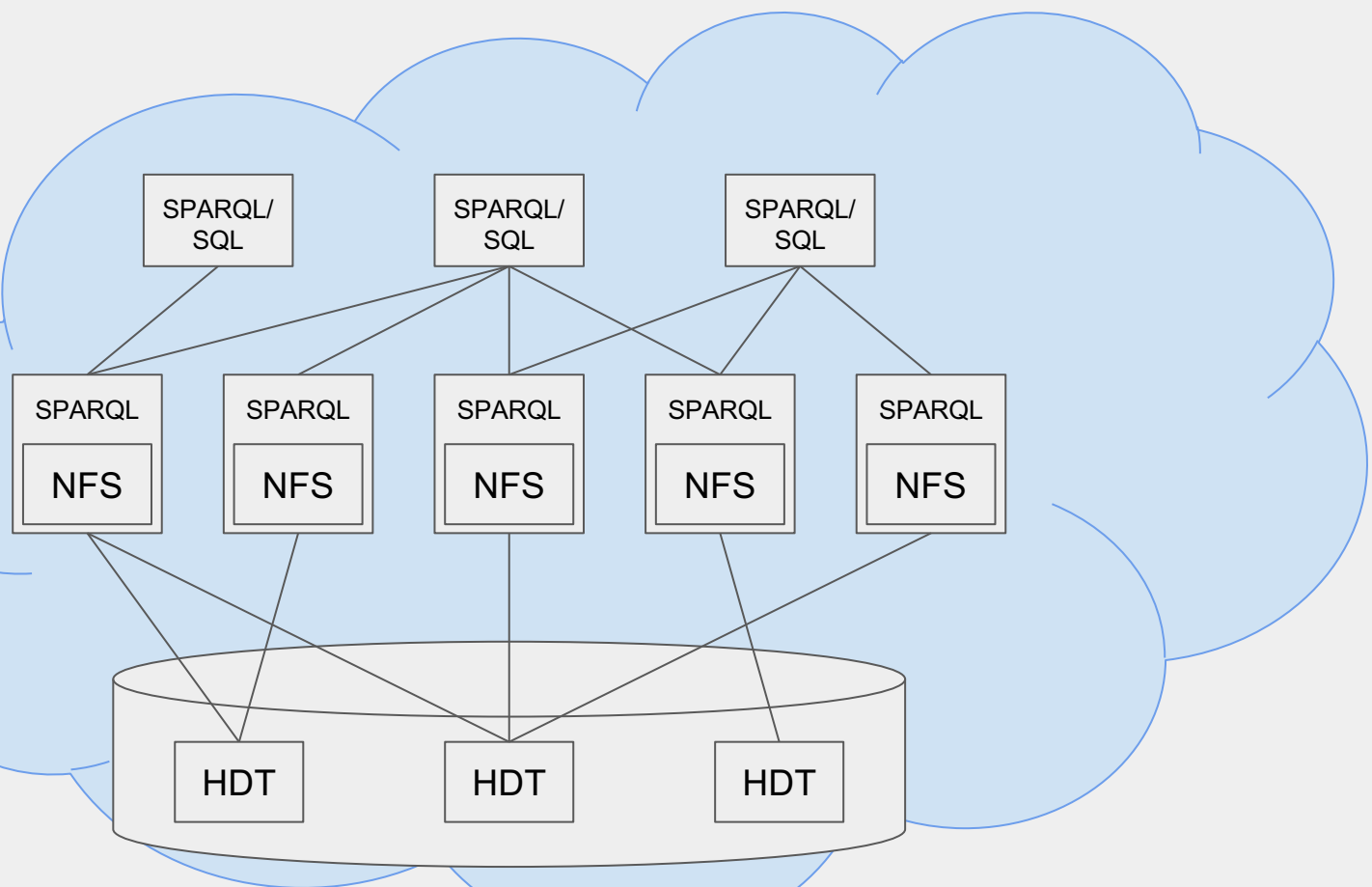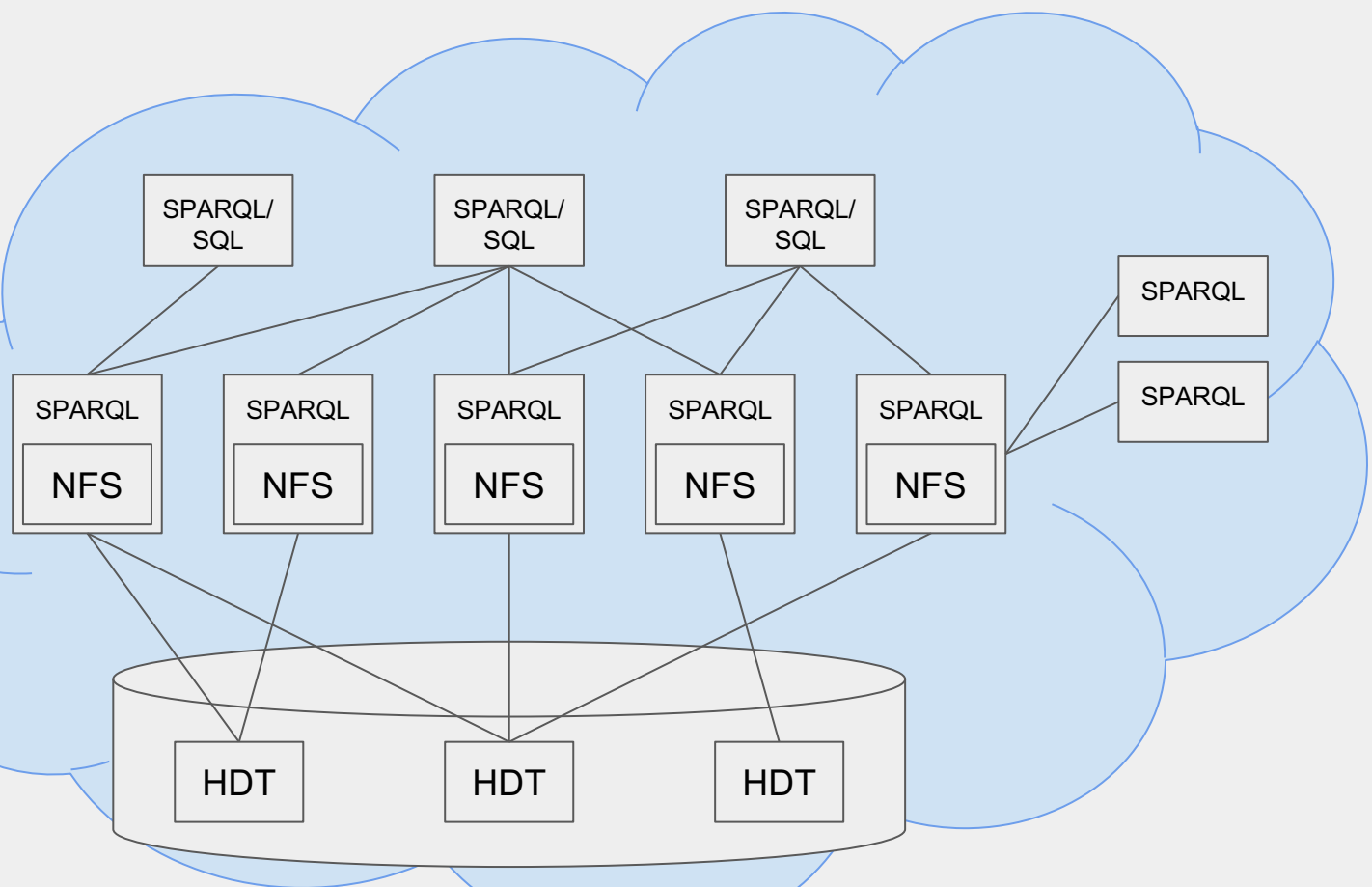
PopulationEstimates.xls
⚠ 99+ · Add a description

Explore

population_estimates_2010_2015 | variable_descriptions

#	fips	T state	T area_name	# rural_urban_continuum_code_2003	# rural_urban_continuum_code_2013
1	00000	US	United States	No data.	No
2	01000	AL	Alabama	No data.	No
3	01001	AL	Autauga County	2	
4	01003	AL	Baldwin County	4	
5	01005	AL	Barbour County	6	

Showing 1-5 of 3,273 rows, 101 columns See all

⤬ Switch to column overview

transactions.csv
Add a description

Explore

#	id	# zip	T county	# naics	T product
1	11,087,645	25285	Clay	211111	STONE-001
2	11,160,873	92780	Orange	236116	CATTL-001
3	12,219,509	33801	Polk	236115	CHERR-001
4	12,409,198	64134	Jackson	444190	CATTL-001
5	14,015,475	33811	Polk	333913	SHEEP-001

Showing 1-5 of 1,012 rows, 5 columns See all

⤬ Switch to column overview

Confidential

transactions.csv

	# id	# zip	T county ⌄	# naics	T product ⌄
1	11,087,			211111	STONE-001
2	11,160,			236116	CATTL-001
3	12,219,			236115	CHERR-001
4	12,409,			444190	CATTL-001
5	14,015,			333913	SHEEP-001
6	15,427,			444190	CHERR-001
7	15,724,426	20866	Montgomery	236116	CAMEL-001
8	16,512,336	02852	Washington	112120	CHERR-001
9	16,605,340	40203	Jefferson	236116	OAKWO-001
10	18,171,732	32807	Orange	236116	STONE-001
11	20,523,893	33805	Polk	237310	CAMEL-001
12	20,597,811	28301	Cumberland	327120	CLAYB-001
13	20,668,558	71378	Franklin	444190	SHEEP-001
14	22,265,107	47403	Monroe	236115	SHEEP-001
15	24,713,218	46227	Marion	111140	CLAYB-001
16	26,577,718	28390	Cumberland	335224	PINEW-001
17	27,486,735	83442	Jefferson	327120	STONE-001
18	29,812,464	40216	Jefferson	236115	PINEW-001

Menu overlay on county column:
- Sort ascending
- Sort descending
- Edit column info
- Show column info
- Match this column
  2 SUGGESTIONS

Bring in contextual data by matching **county** to other datasets.

⊤ county
ORIGINAL COLUMN

Clay
Orange
Polk

◉ county_1
MATCHING COLUMN

100% MATCH

County by data.world

Clay County, West Virginia
Orange County, California
Polk County, Florida

✓ Matched column

◉ us_state
RELATED COLUMN

BASED ON

US State by data.world

West Virginia
California
Florida

Add related column

REL

BASE

US S

We

Ca

Fl

Done

About th

LAST MODIFIE

OWNER

CREATED

SIZE

LABELS

DESCRIPTION

Displaying 5 c

transactio

⌄ # Id

+ Add a de

11,058,969.19

DISTINCT

> 1,000

15	24,713,218	46227	Marion	111140	CLAYB-001
16	26,577,718	28390	Cumberland	335224	PINEW-001
17	27,486,735	83442	Jefferson	327120	STONE-001
18	29,812,464	40216	Jefferson	236115	PINEW-001

Bring in contextual data by matching **county** to other datasets.

us_state_1
`RELATED COLUMN`

BASED ON

US State by data.world

| West Virginia |
| California |
| Florida |

Add related column

census_division
`RELATED COLUMN`

BASED ON

Census Division by data.world

| South Atlantic |
| Pacific |
| South Atlantic |

Add related column

census_region
`RELATED COLUMN`

BASED ON

Census Region by data.world

| South |
| West |
| South |

✕ Remove column

Done

#	id	#	zip	county	county_1	census_region	#	naics	product
1	11,087,645		25285	Clay	Clay County, West Virginia	South		211111	STONE-001
2	11,160,873		92780	Orange	Orange County, California	West		236116	CATTL-001
3	12,219,509		33801	Polk	Polk County, Florida	South		236115	CHERR-001
4	12,409,198		64134	Jackson	Jackson County, Missouri	Midwest		444190	CATTL-001
5	14,015,475		33811	Polk	Polk County, Florida	South		333913	SHEEP-001
6	15,427,096		64106	Jackson	Jackson County, Missouri	Midwest		444190	CHERR-001
7	15,724,426		20866	Montgomery	Montgomery County, Maryland	South		236116	CAMEL-001
8	16,512,336		02852	Washington	Washington County, Rhode Island	Northeast		112120	CHERR-001
9	16,605,340		40203	Jefferson	Jefferson County, Kentucky	South		236116	OAKWO-001
10	18,171,732		32807	Orange	Orange County, Florida	South		236116	STONE-001
11	20,523,893		33805	Polk	Polk County, Florida	South		237310	CAMEL-001
12	20,597,811		28301	Cumberland	Cumberland County, North Carolina	South		327120	CLAYB-001
13	20,668,558		71378	Franklin	Franklin Parish, Louisiana	South		444190	SHEEP-001
14	22,265,107		47403	Monroe	Monroe County, Indiana	Midwest		236115	SHEEP-001
15	24,713,218		46227	Marion	Marion County, Indiana	Midwest		111140	CLAYB-001
16	26,577,718		28390	Cumberland	Cumberland County, North Carolina	South		335224	PINEW-001
17	27,486,735		83442	Jefferson	Jefferson County, Idaho	West		327120	STONE-001
18	29,812,464		40216	Jefferson	Jefferson County, Kentucky	South		236115	PINEW-001
19	31,256,333		65803	Greene	Greene County, Missouri	Midwest		112120	CHERR-001
20	31,690,084		35214	Jefferson	Jefferson County, Alabama	South		112120	PINEW-001

⬇ Download ▾

population_estimates_2010_2015   variable_descriptions

#	fips	state	area_name	# rural_urban_continuum_code_2003	# rural_urban_
1	00000	US	United States	No data.	
2	01000	AL	Alabama	No data.	
3	01001	AL	Autauga County	2	
4	01003	AL	Baldwin County	4	
5	01005	AL	Barbour County	6	
6	01007	AL	Bibb County	1	
7	01009	AL	Blount County	1	
8	01011	AL	Bullock County	6	
9	01013	AL	Butler County	6	
10	01015	AL	Calhoun County	3	
11	01017	AL	Chambers County	6	
12	01019	AL	Cherokee County	8	
13	01021	AL	Chilton County	1	
14	01023	AL	Choctaw County	9	
15	01025	AL	Clarke County	7	
16	01027	AL	Clay County	9	

Bring in contextual data by matching **fips** to other datasets.

fips

ORIGINAL COLUMN

| 0 |
| 1000 |
| 1001 |

county

MATCHING COLUMN

96% MATCH

County by data.world

| null |
| null |
| Autauga County, Alabama |

Add matched column

Done

Bring in contextual data by matching **fips** to other datasets.

fips ✎

ORIGINAL COLUMN

| 0 |
| 1000 |
| 1001 |

county ✎

MATCHING COLUMN

96% MATCH

County by data.world

| null |
| null |
| Autauga County, Alabama |

✓ Matched column

us_state

RELATED COLUMN

BASED ON

US State by data.world

| null |
| null |
| Alabama |

Add related column

REL

BASE

US S

| nu |
| nu |
| Al |

Done

data.world

Confidential

population_estimates_2010_2015   variable_descriptions

# fips	county	state	area_name	# rural_urban_continuum_code_2003
1	00000	No data.	US	United States	No da
2	01000	No data.	AL	Alabama	No da
3	01001	Autauga County, Alabama	AL	Autauga County	
4	01003	Baldwin County, Alabama	AL	Baldwin County	
5	01005	Barbour County, Alabama	AL	Barbour County	
6	01007	Bibb County, Alabama	AL	Bibb County	
7	01009	Blount County, Alabama	AL	Blount County	
8	01011	Bullock County, Alabama	AL	Bullock County	
9	01013	Butler County, Alabama	AL	Butler County	
10	01015	Calhoun County, Alabama	AL	Calhoun County	
11	01017	Chambers County, Alabama	AL	Chambers County	
12	01019	Cherokee County, Alabama	AL	Cherokee County	
13	01021	Chilton County, Alabama	AL	Chilton County	
14	01023	Choctaw County, Alabama	AL	Choctaw County	
15	01025	Clarke County, Alabama	AL	Clarke County	
16	01027	Clay County, Alabama	AL	Clay County	
17	01029	Cleburne County, Alabama	AL	Cleburne County	

JOIN 🌐 SHARED

```
1  SELECT transactions.id,
2         transactions.product,
3         transactions.county_1,
4         population_estimates_2010_2015.census_2010_pop
5    FROM transactions
6    JOIN population_estimates_2010_2015
7      ON transactions.county_1 = population_estimates_2010_2015.county
```

⋯

📋 Results table   📊 Chart   1,011 query results   ⬇ Download ▾   ⊞ Open in app

#	id	product	county_1	# census_2010_pop
1	11,087,645	STONE-001	Clay County, West Virginia	9,386
2	11,160,873	CATTL-001	Orange County, California	3,010,232
3	20,523,893	CAMEL-001	Polk County, Florida	602,095
4	121,341,948	TEAKW-001	Madison County, Tennessee	98,294
5	986,208,161	CLAYB-001	Lincoln County, New Mexico	20,497
6	986,743,632	CAMEL-001	Orange County, Florida	1,145,956
7	986,837,245	CATTL-001	Franklin County, Ohio	1,163,414
8	989,426,677	TEAKW-001	Jefferson Parish, Louisiana	432,552
9	990,982,232	CATTL-001	Orange County, California	3,010,232
10	991,554,121	STONE-001	Orange County, California	3,010,232
11	992,286,947	CLAYB-001	Jackson County, Missouri	674,158

🐙 data.world

https://entities.data.world/country_us

United States of America

`dwo:Country`

Property	Value
rdf:type	• dwo:Country
dc:description	• federal republic in North America
dwo:geonameId	• 6252001
dwo:officialName	• United States of America
dwo:partOfContinent	• dwe:continent_north_america
dwo:wikidataLink	• http://www.wikidata.org/entity/Q30
dwo:wikipediaLink	• https://en.wikipedia.org/wiki/United_States
skos:altLabel	• America • U.S.A. • USA • United States • United States America
skos:prefLabel	• United States of America

Thanks!

Bryon Jacob

`bryon@data.world`

`https://data.world/bryon`