Meta

Daylight Earth Tables

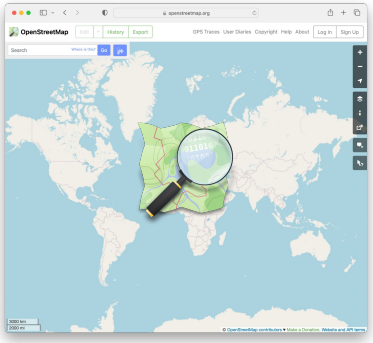
Jennings Anderson, Jonah Adkins, Jacob Wasserman

REALITY LABS

Maps at Meta

19/October/2022

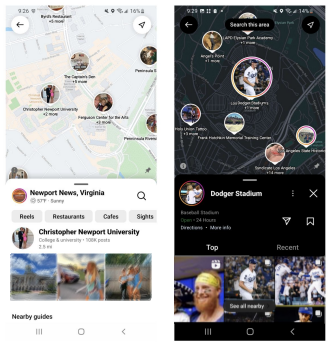
Meta



OpenStreetMap

Daylight Map Distribution

Earth Table



Maps in Meta Products

- Instagram
- Facebook Marketplace
- And more…

Meta

Overview

What is the Daylight Map Distribution?

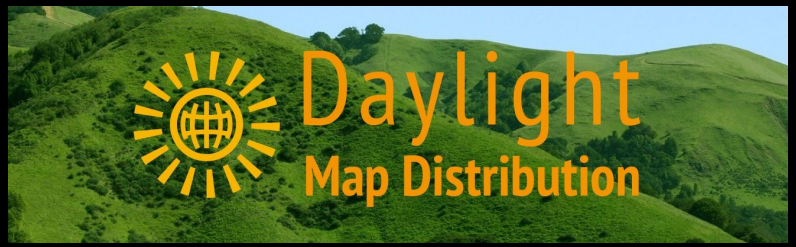
What is the Earth Table?

Introducing: Daylight Earth Table

Working with the Daylight Earth Table

Meta

Daylight

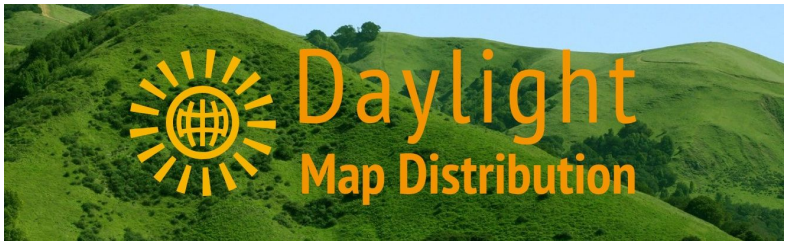


A monthly distribution of OpenStreetMap that undergoes a series of quality control and vandalism checks to ensure a degree of map quality and integrity.

REALITY LABS

Meta

Timeline of Daylight OpenStreetMap Distribution



< 2019

State of the Map 2018 & 2019

Mobius Logical Changesets (LoChas)

2020

March 2020

Announcing Daylight v0.1

2021

April 2021

Daylight v1.0 released

2022

January 2022

Daylight published on AWS Registry of Open Data as PBF and *Analysis-Ready* parquet files

Today

October 2022

Daylight v1.18

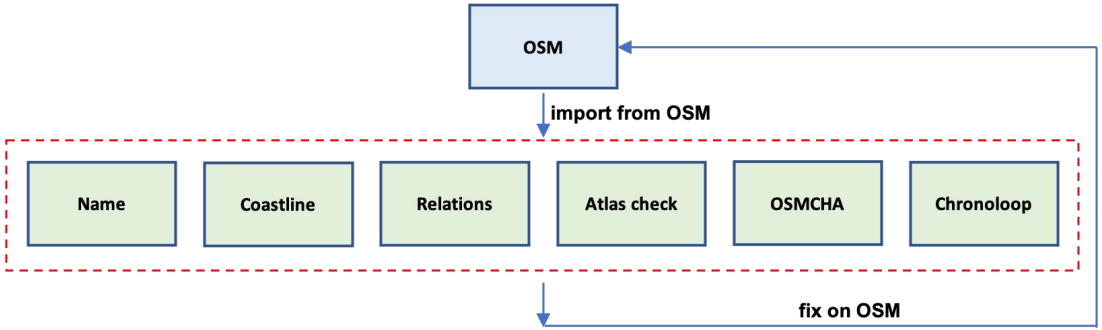
https://daylightmap.org

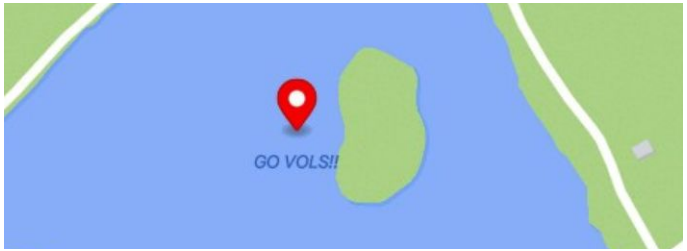
Daylight Process:
Find-Fix-Import Loop

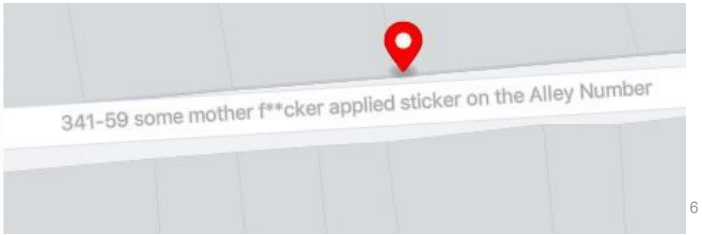
Find: Discover errors or other issues anywhere on the map

Fix: Submit fixes on live OSM, not in an internal database

Import: Apply fixes from OSM into the Daylight map







Meta

Daylight v1.18

525 M

Buildings

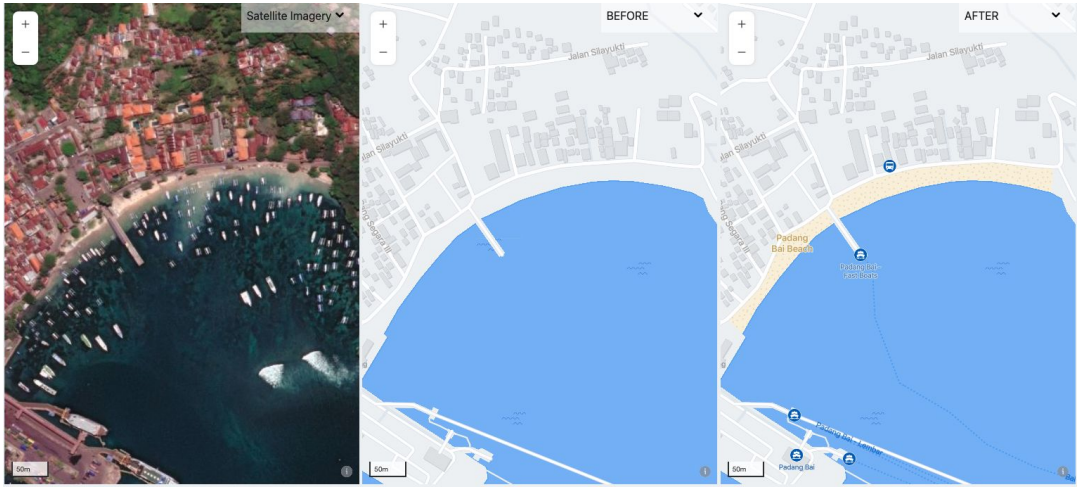
78 M

Kilometers of roads / paths

100%

OpenStreetMap Data

Think of Daylight as a snapshot of OSM where each feature might not be from the same snapshot.



Example: Repaired beach relation for Padang Bai Beach in Bali, Indonesia

Daylight v1.18

Changelog and summaries of fixes made to OSM availabl on daylightmap.org

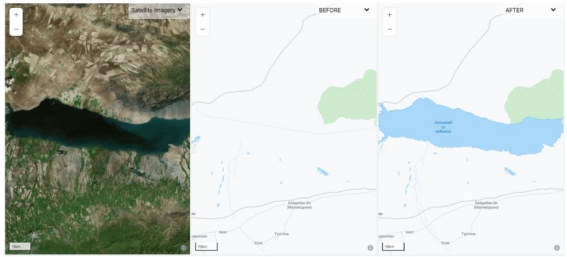
daylightmap.org

v1.18 Summary

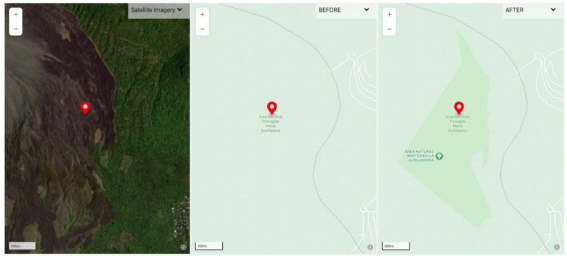
Most Notable 4

- 122 features were fixed from relation check.
- 172 features were fixed from coastline check.
- 4648 features were fixed from various atlas checks.
- 207 features were fixed from nlp check.

Fixed broken relation for Kapchagay Reservoir, in Almaty Oblysy, Kazakhstan



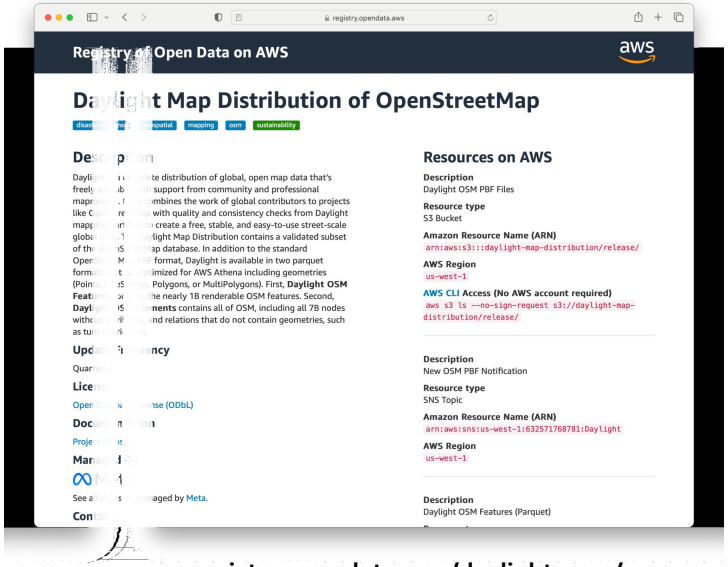
Fixed a broken relation for Área Natural Protegida La Auxiliadora in Departamento de Sonsonate, El Salvador



Fixed the Area de la Biosfera de Transicion del Bosque Mbaracayu, in Región Oriental, Paraguay nature reserve

Where can I find Daylight?

- **OSM PBF Format**
- **Analysis-Ready Cloud-Optimized Parquet Files**



Meta

Earth Table

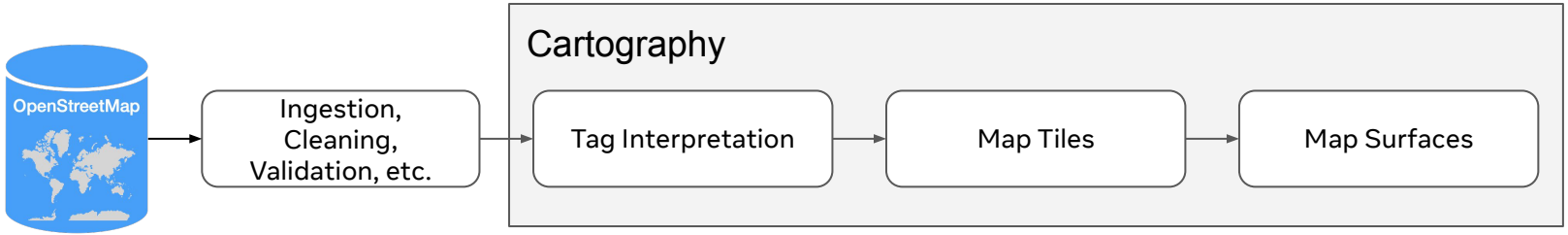
Meta

Where did the Earth Table come from?

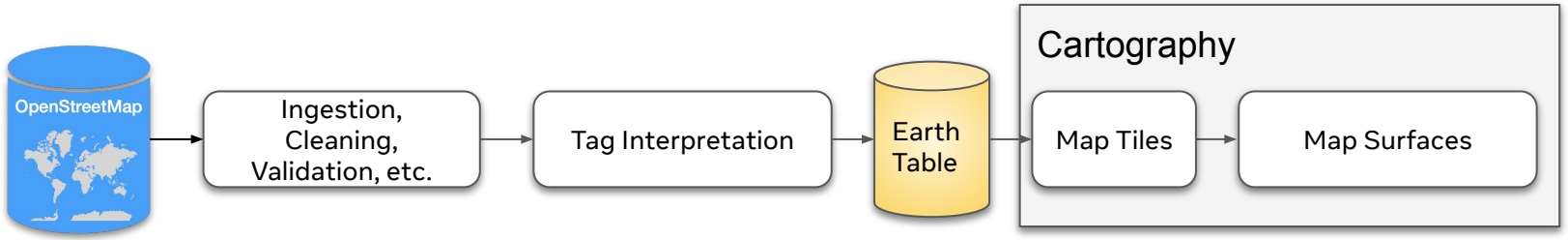
- Growing internal use cases to query the earth - “I need all the parks in the world to do x”
- Using basemaps is a cross-functional tool for projects across the org.
- Needs to be simple - tagging complexities of OSM are a high barrier to entry
- Needs to be in a single table, with easy to understand schema & hierarchy

Meta

Before:
Cartography team did all of the OSM tag interpretation and created map tiles.
Needed to look inside tiles to find well-formatted, translated map data

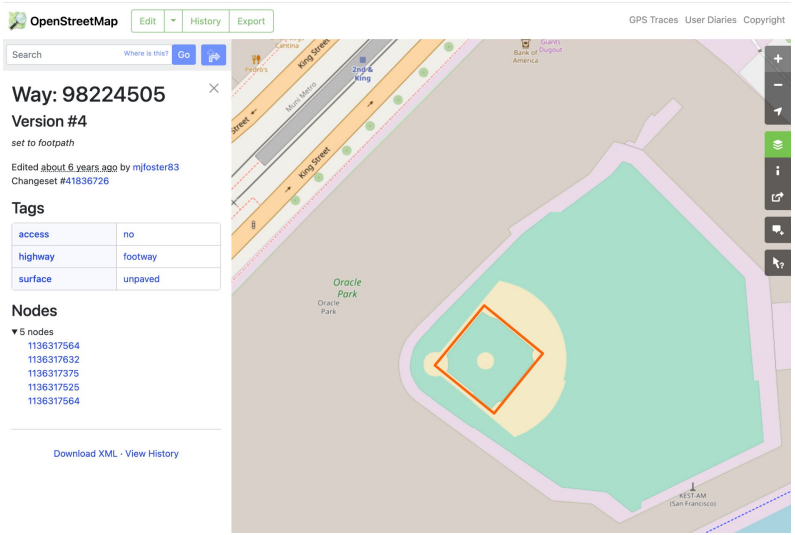


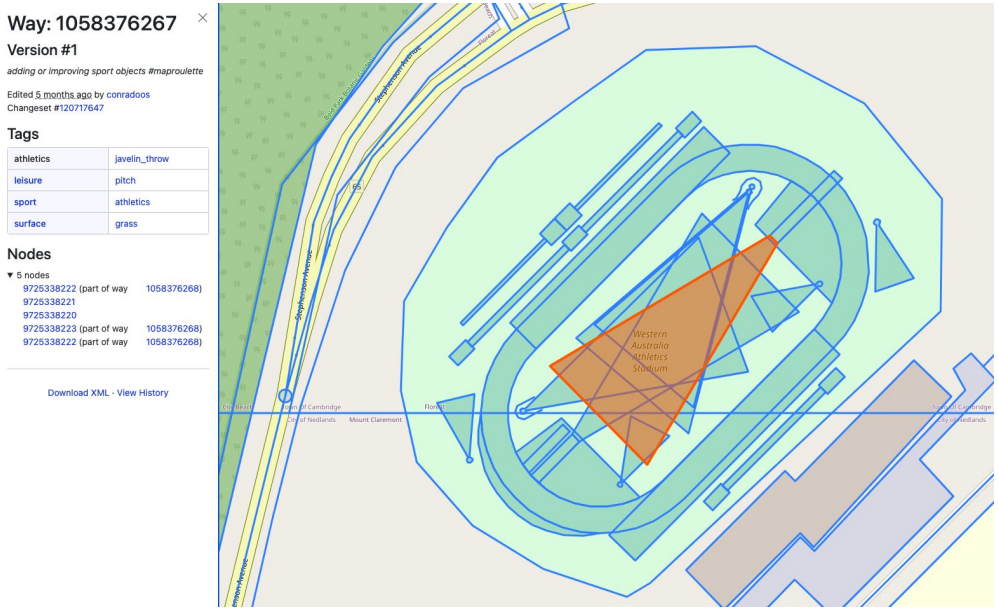
After:
The earth table is created by interpreting OSM tags into a simplified 3-level ontological schema.
Anyone can access this table, especially cartography



Meta

It doesn’t solve everything…





Healthcare			
amenity	baby_hatch		A place where a baby can be, out of necessity, anonymously left to be safely cared for and perhaps adopted.
amenity	clinic		A medium-sized medical facility or health centre.
amenity	dentist		A dentist practice / surgery.
amenity	doctors		A doctor's practice / surgery.
amenity	hospital		A hospital providing in-patient medical treatment. Often used in conjunction with `emergency=*` to note whether the medical centre has emergency facilities (A&E (brit.) or ER (am.))
amenity	nursing_home		Discouraged tag for a home for disabled or elderly persons who need permanent care. Use **`amenity`**`=social_facility` + `social_facility=nursing_home` now.

Value
`healthcare=alternative`
`healthcare=audiologist`
`healthcare=birthing_centre`
`healthcare=blood_bank`
`healthcare=blood_donation`
`healthcare=centre`
`healthcare=clinic`
`healthcare=community_health_worker`
`healthcare=counselling`
`healthcare=dentist`
`healthcare=dialysis`
`healthcare=doctor`
`healthcare=hospice`
`healthcare=hospital`
`healthcare=laboratory`
`healthcare=midwife`

Way: Riverside Regional Medical Center (298161573)

Version #12

Roads and Sidewalks

Edited almost 2 years ago by pardthemonster
Changeset #96688275

Tags

addr:city	Newport News
addr:housenumber	500
addr:postcode	23601
addr:state	VA
addr:street	J. Clyde Morris Boulevard
amenity	hospital
beds	450
emergency	yes
healthcare	hospital
name	Riverside Regional Medical Center
opening_hours	24/7
operator	Riverside
phone	+1 757 594 2000
website	https://www.riversideonline.com/rrmc/

theme	class	subclass	metadata	wkt
poi	medical	hospital	{"amenity":"hospital","is_area":true,"quadkey":"032010233003132"}	POINT (-76.4
landuse	medical	hospital	{"quadkey":"032010233003132","surface_area_sq_m":267697.27}	POLYGON (

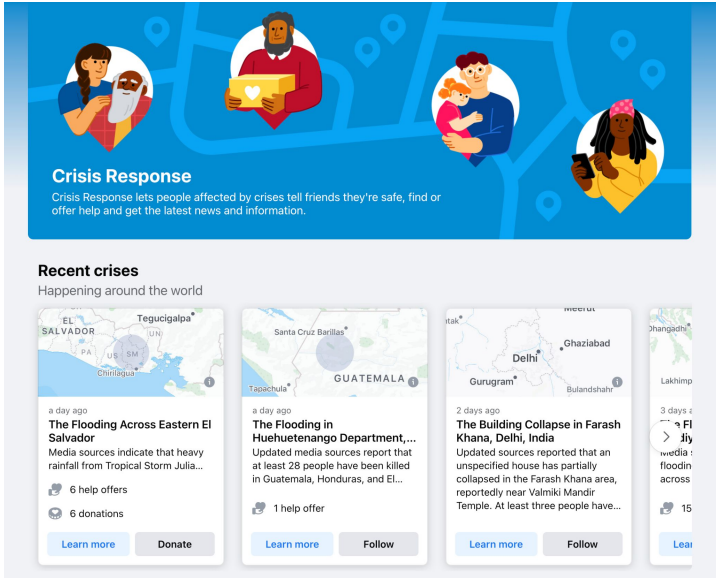
building	525,240,408
building_detail	2,232,081
infrastructure	11,226,443
land	40,831,400
landuse	39,397,340
placename	3,430,404
poi	38,777,950
road	216,890,017
transit	6,716,006
water	36,401,694

Cartographic Data Improvements:

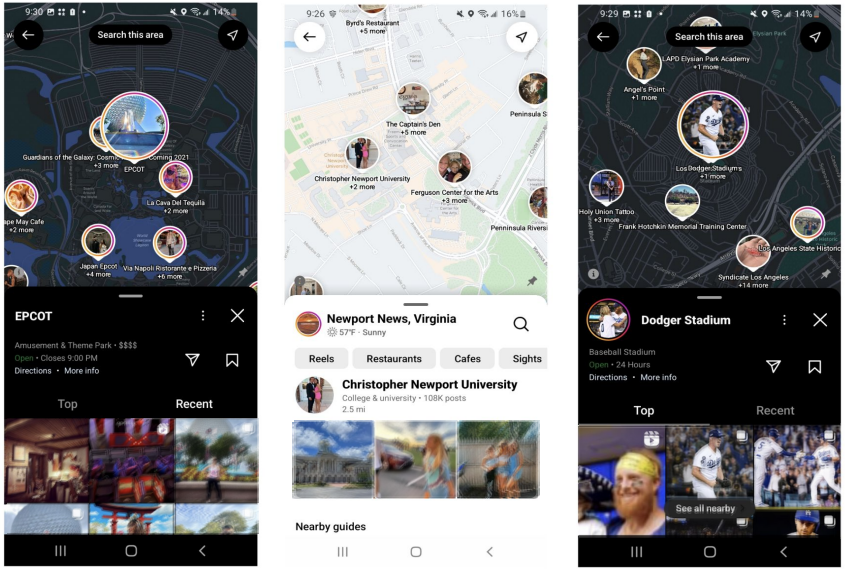
- land theme includes processed global coastlines
- building height information is validated and normalized
- placename theme includes simplified classification -> all places grouped into 3 classes: urban, settlement, local
- lengths and areas calculated as attribute columns
- Bing Tile Quadkey applied to each feature
- Lots of boolean evaluations for easy filtering:

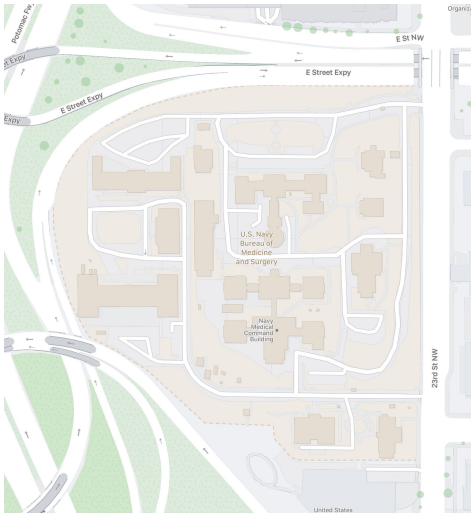
  “Is_indoor”, “is_intermittent”, “is_bridge”, “is_area”, etc
- Buildings include “landuse” class they are within
- Building_detail includes building id they are part of

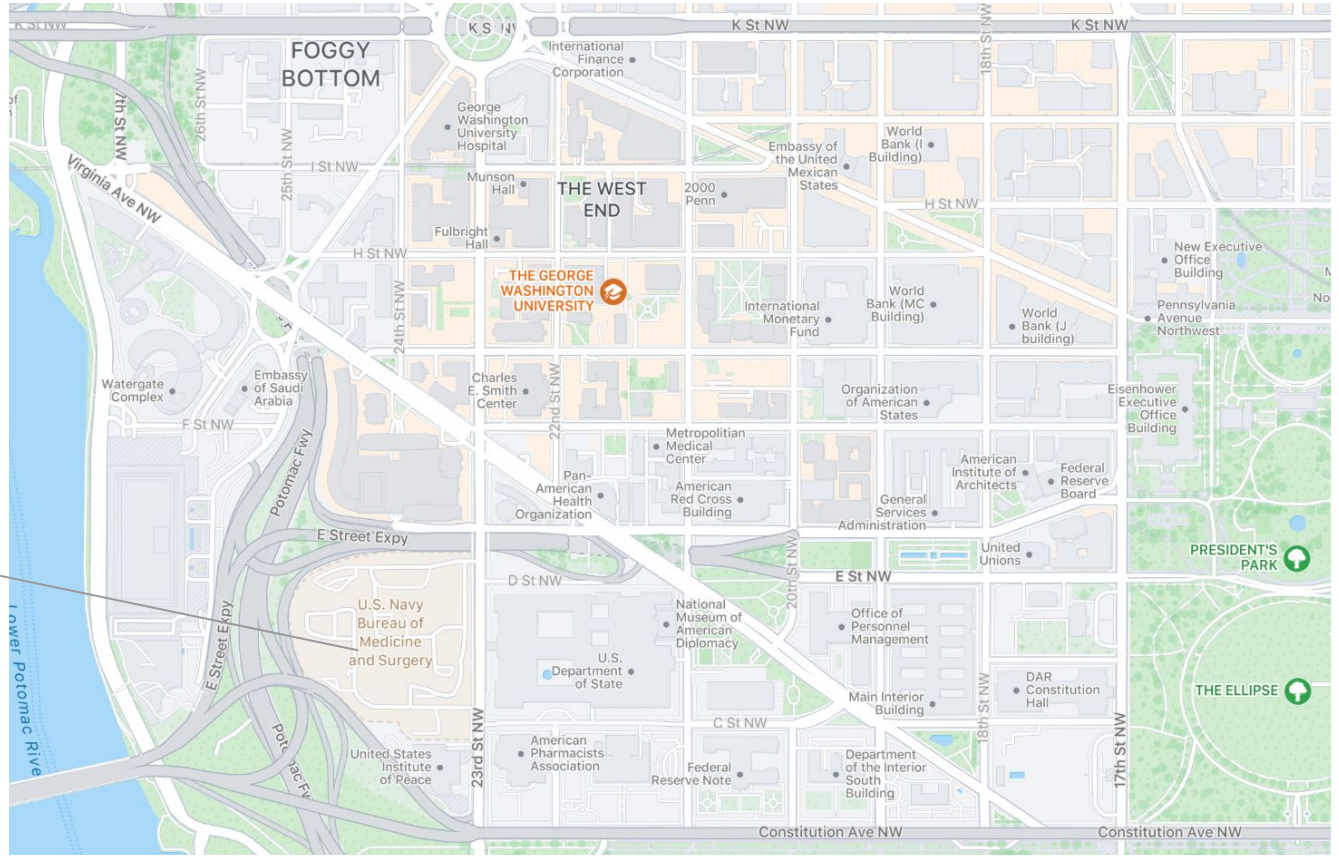
Meta



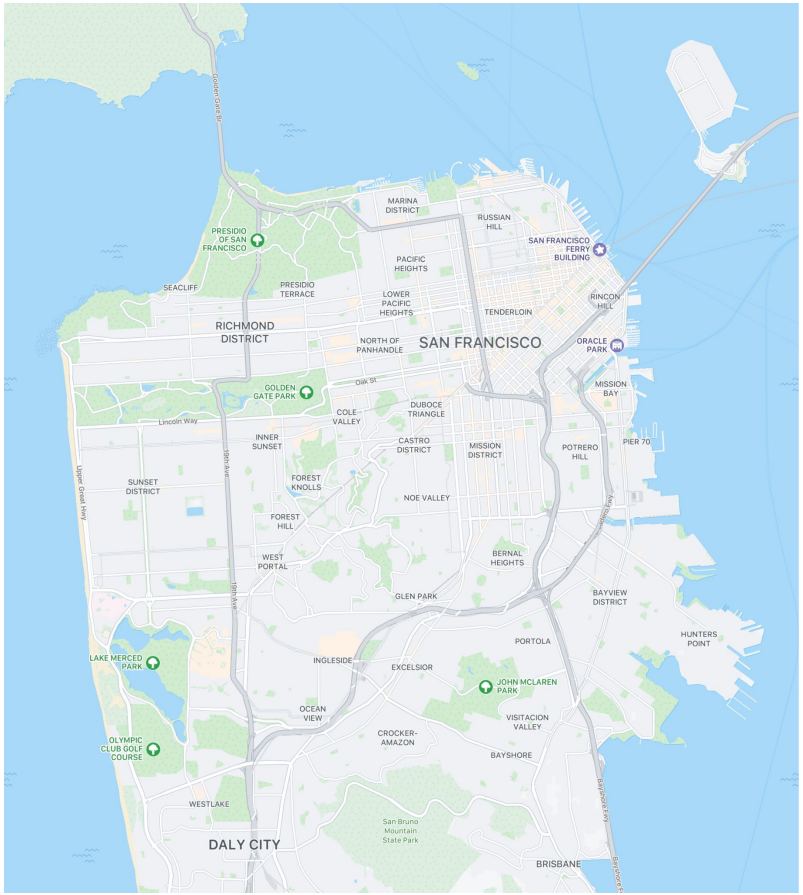
83%	Earth Table Data Source
16%	Internal Data Sources that are “Earthified”
1%	Curated Cartographic Map Features

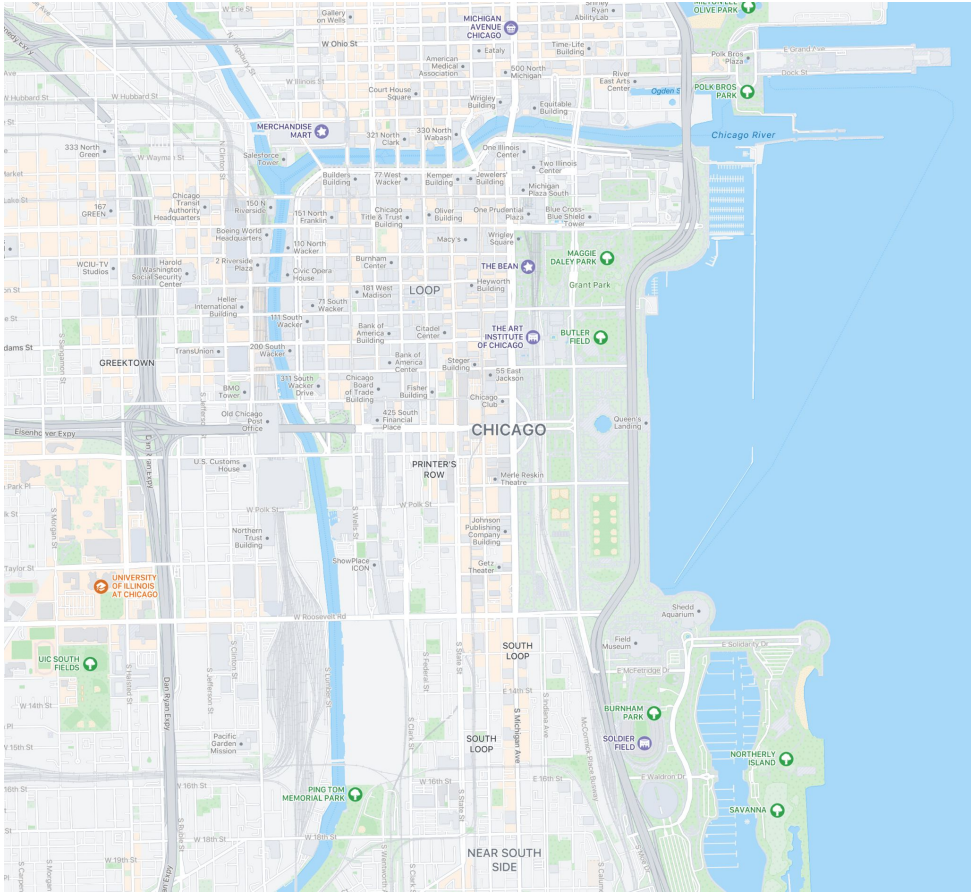




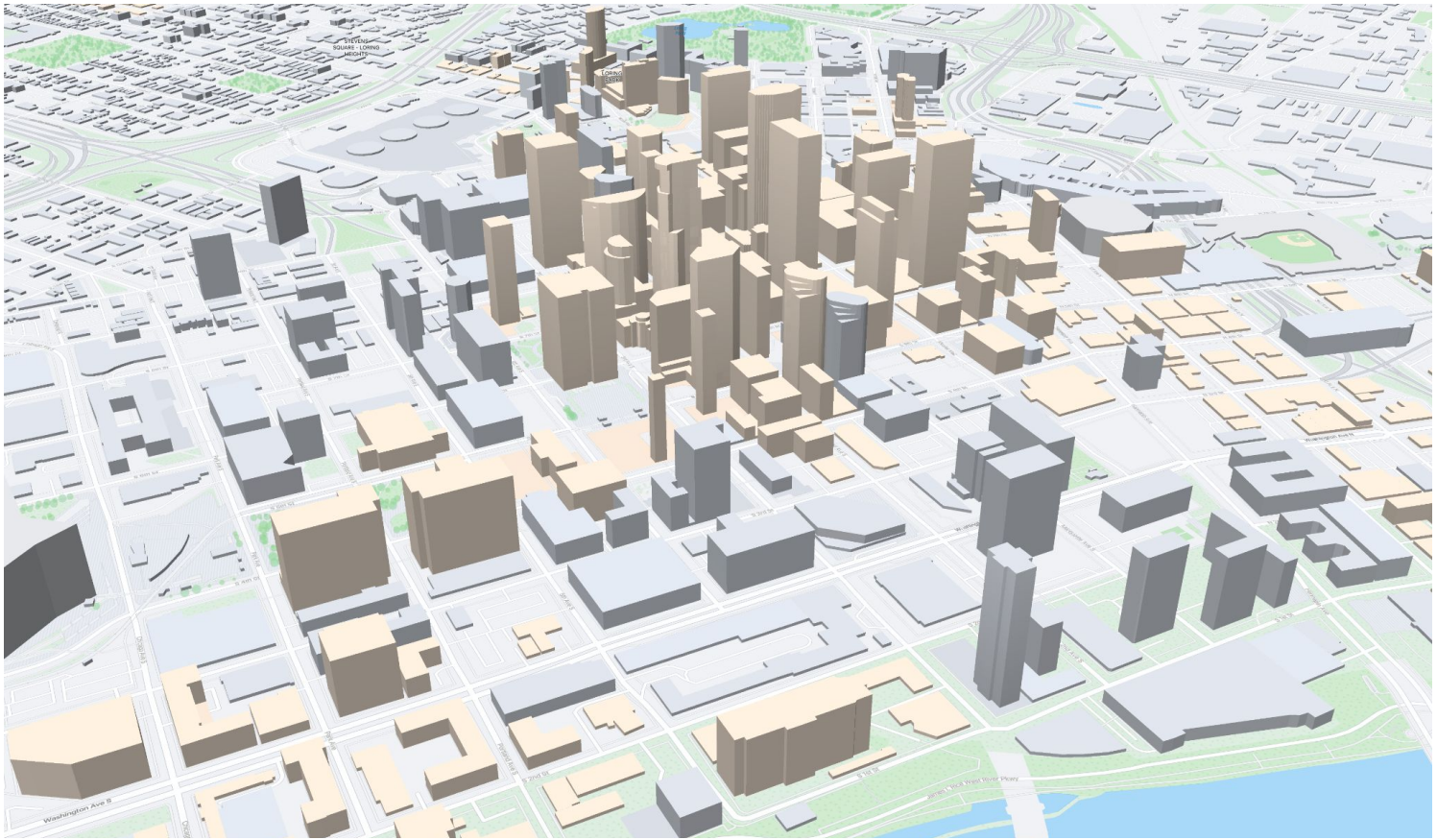


Meta





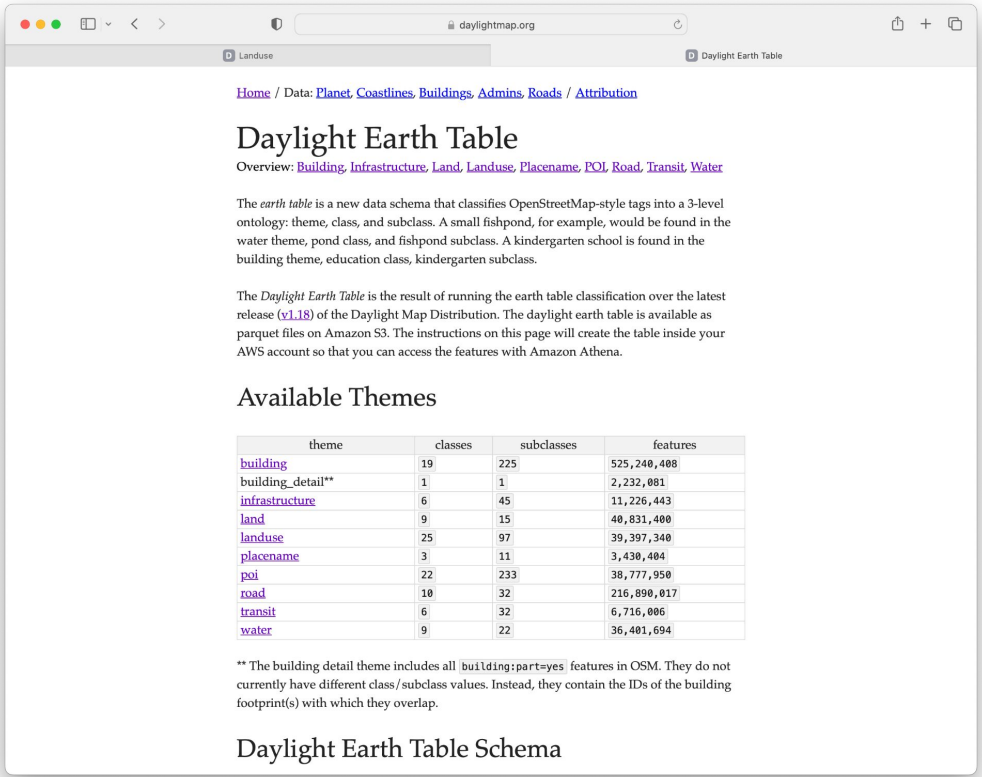
Meta



Meta

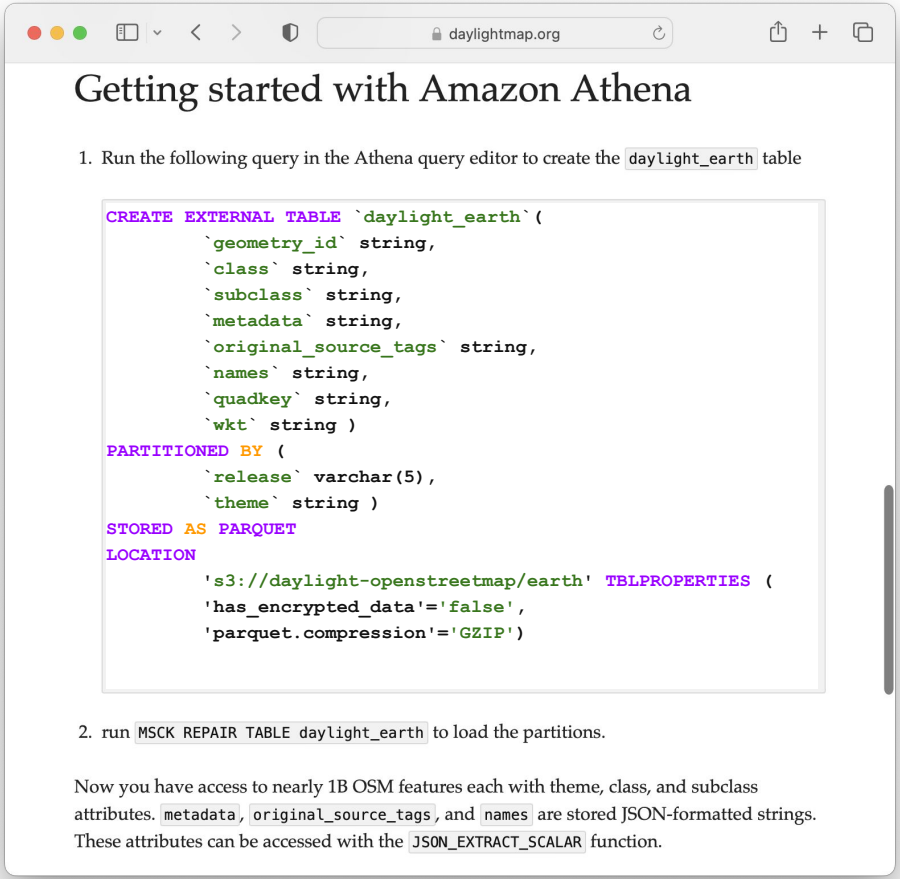
Accessing Daylight Earth Tables

s3://daylight-openstreetmap/earth

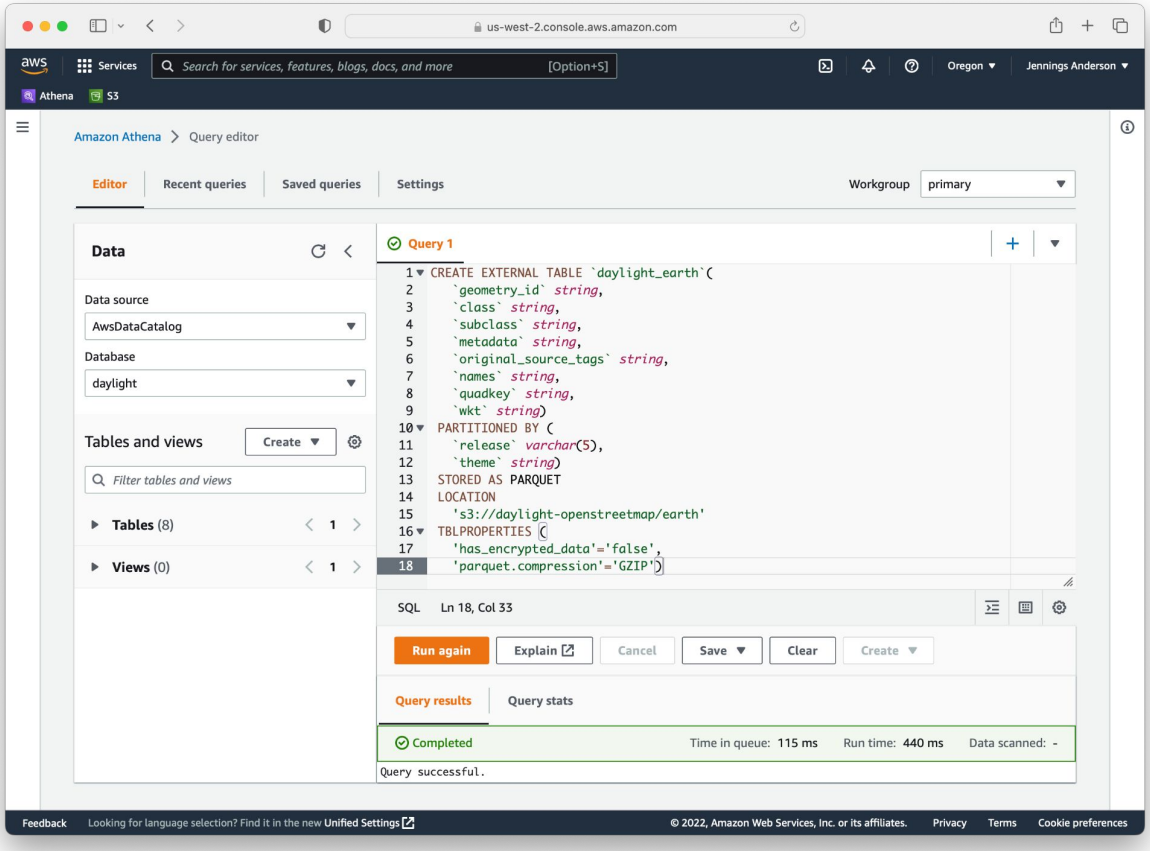


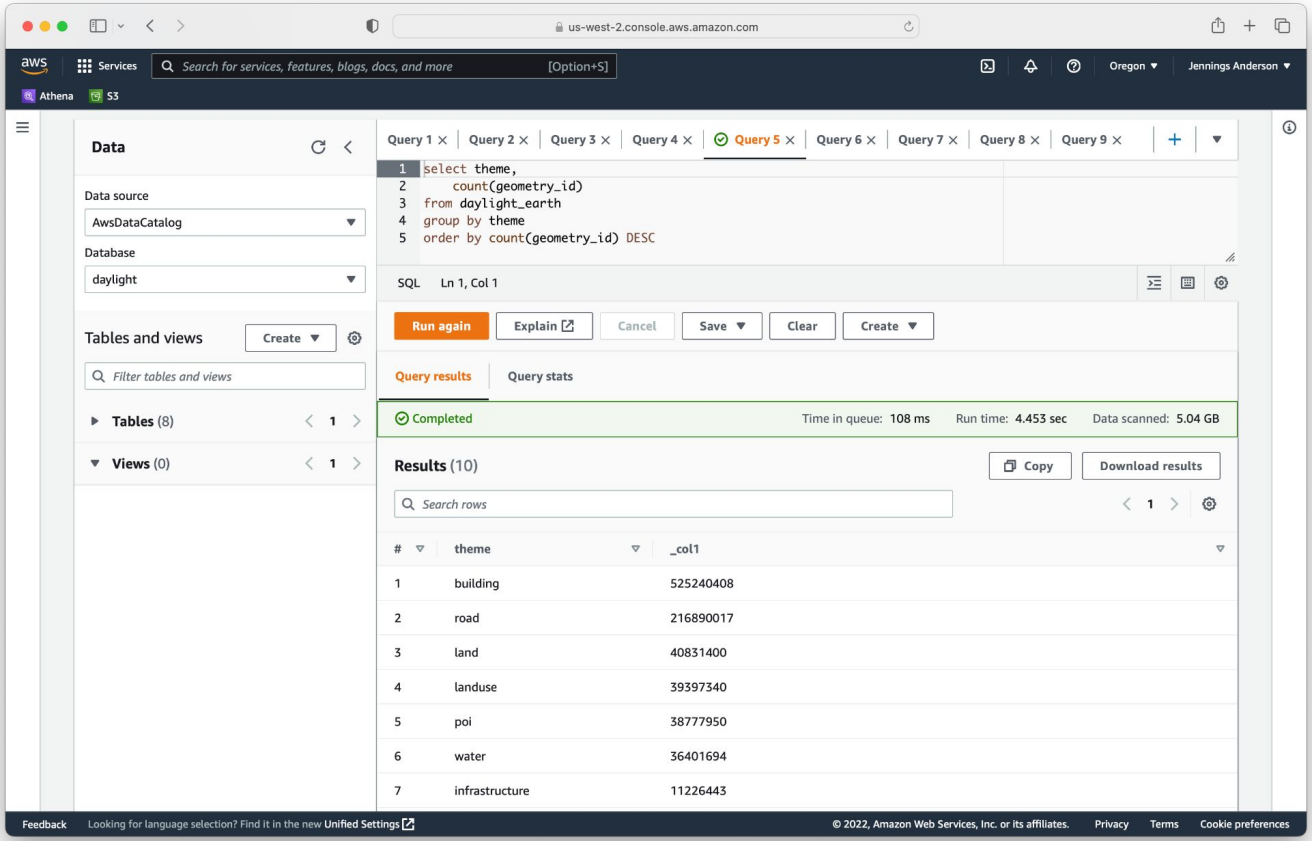
daylightmap.org/earth

Using Amazon Athena



daylightmap.org/earth


us-west-2.console.aws.amazon.com
aws
Services
Search for services, features, blogs, docs, and more [Option+S]
Oregon
Jennings Anderson
Athena
S3
Amazon Athena > Query editor
Editor
Recent queries
Saved queries
Settings
Workgroup primary
Data
Data source
AwsDataCatalog
Database
daylight
Tables and views
Create
Filter tables and views
Tables (8)
Views (0)
Query 1
1 CREATE EXTERNAL TABLE `daylight_earth`(
2 `geometry_id` string,
3 `class` string,
4 `subclass` string,
5 `metadata` string,
6 `original_source_tags` string,
7 `names` string,
8 `quadkey` string,
9 `wkt` string)
10 PARTITIONED BY (
11 `release` varchar(5),
12 `theme` string)
13 STORED AS PARQUET
14 LOCATION
15 's3://daylight-openstreetmap/earth'
16 TBLPROPERTIES (
17 'has_encrypted_data'='false',
18 'parquet.compression'='GZIP')
SQL Ln 18, Col 33
Run again
Explain
Cancel
Save
Clear
Create
Query results
Query stats
Completed
Time in queue: 115 ms
Run time: 440 ms
Data scanned: -
Query successful.
Feedback
Looking for language selection? Find it in the new Unified Settings
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences


us-west-2.console.aws.amazon.com
aws
Services
Search for services, features, blogs, docs, and more [Option+S]
Oregon
Jennings Anderson
Athena
S3
Data
Data source
AwsDataCatalog
Database
daylight
Tables and views
Create
Filter tables and views
Tables (8)
Views (0)
Query 1
Query 2
Query 3
Query 4
Query 5
Query 6
Query 7
Query 8
Query 9
select theme,
count(geometry_id)
from daylight_earth
group by theme
order by count(geometry_id) DESC
SQL Ln 1, Col 1
Run again
Explain
Cancel
Save
Clear
Create
Query results
Query stats
Completed
Time in queue: 108 ms
Run time: 4.453 sec
Data scanned: 5.04 GB
Results (10)
Copy
Download results
Search rows
theme _col1
1 building 525240408
2 road 216890017
3 land 40831400
4 landuse 39397340
5 poi 38777950
6 water 36401694
7 infrastructure 11226443
Feedback
Looking for language selection? Find it in the new Unified Settings
© 2022, Amazon Web Services, Inc. or its affiliates.
Privacy
Terms
Cookie preferences

Query 25 | Query 2 | Query 3 | Query 4 | Query 5 | Query 6 | Query 7 | Query 8 | Query 9 | Query 10

```
1 SELECT *
2 FROM daylight_earth
3 WHERE theme = 'building'
4 AND JSON_EXTRACT_SCALAR(metadata, '$.landuse') = 'developed:industrial'
5 LIMIT 100
```

SQL Ln 5, Col 10

Run again | Explain | Cancel | Save | Clear | Create

Query results | Query stats

Completed Time in queue: 168 ms Run time: 6.46 sec Data scanned: 15.19 GB

Results (100+)

Copy | Download results

Search rows

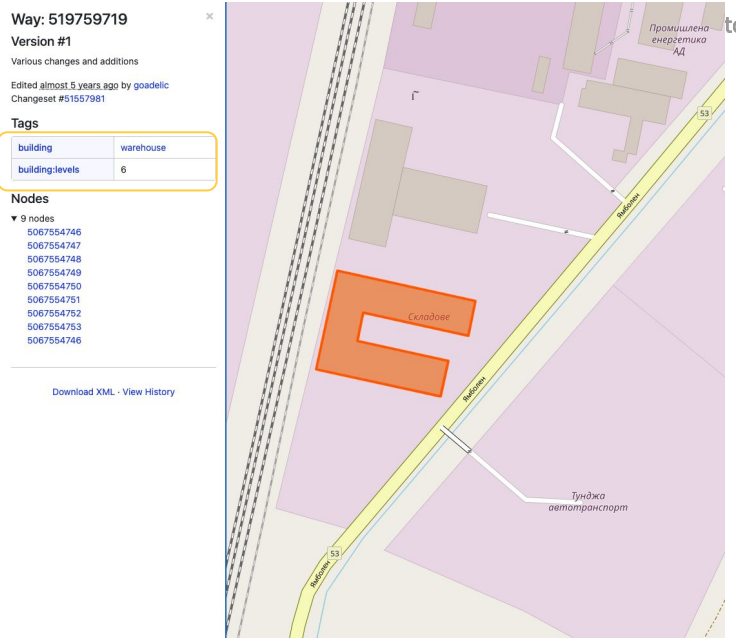
1 ...

#	geometry_id	class	subclass	metadata	original_source_tags
1	w496485510@1	general	building	{"building":"yes","landuse":"developed:industrial","quadkey":"031333211101132","surface_area_sq_m":268.9}	{"building":"yes","source":"Eusko Jaurlaritza-Gobierno Vasco. GeoEuskadi."}

Earth Table Metadata

Normalize attributes into metadata such as height

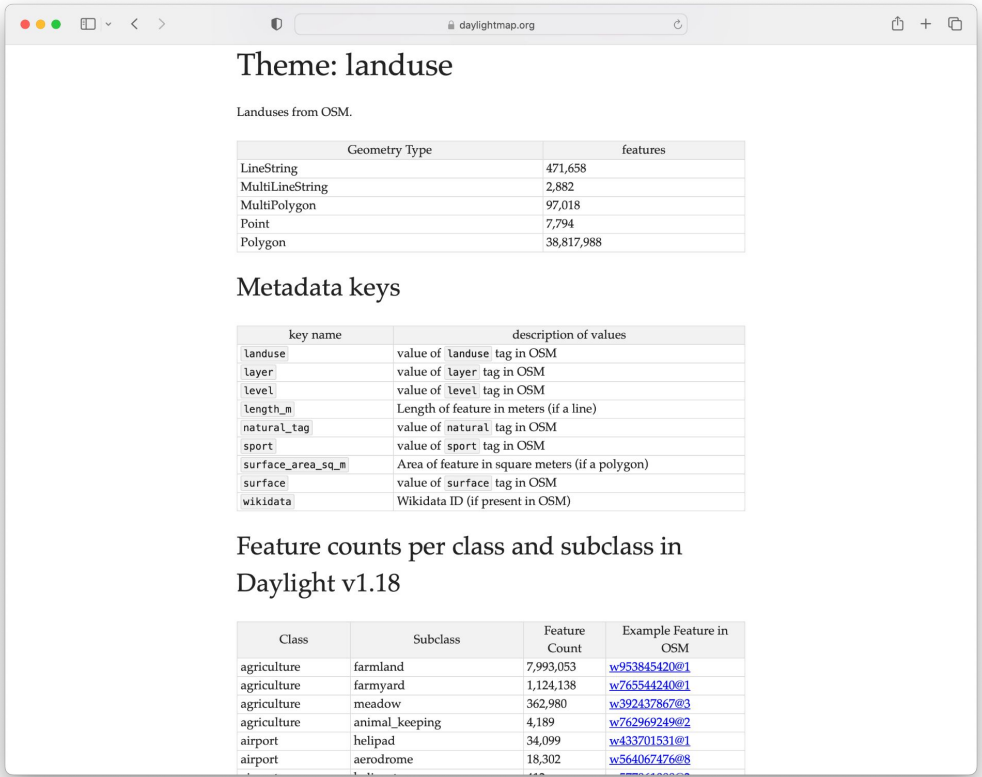
Append additional information from other themes for overlapping features



`earth_table`

source: **osm**
theme: **building**
class: **industrial**
subclass: **warehouse**

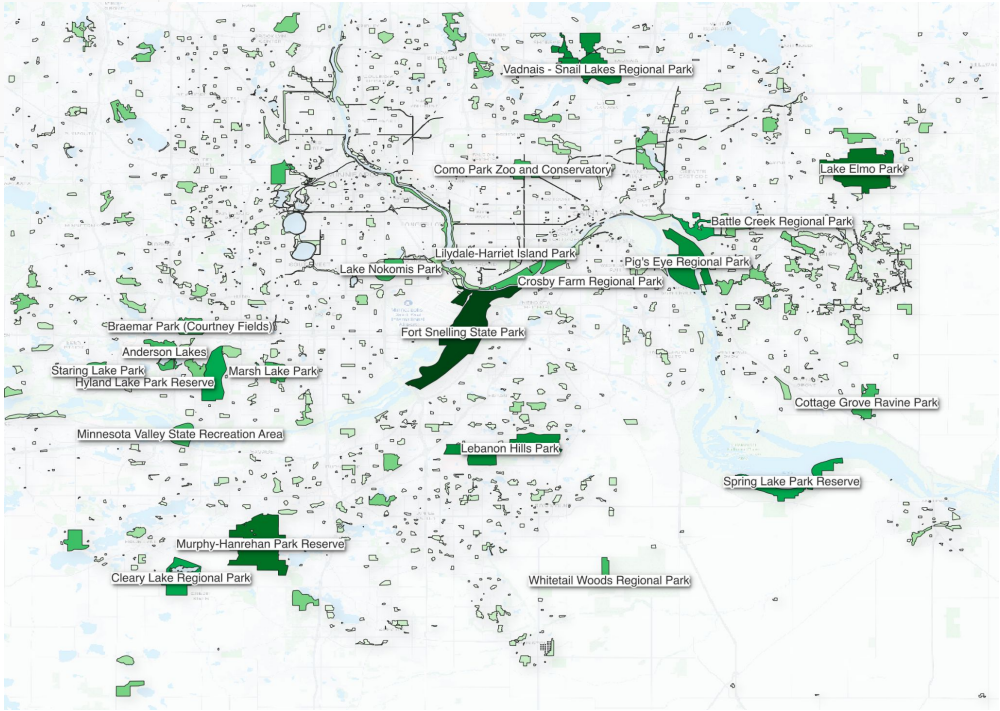
height: **20.52 meters**
area: **4962 square meters**
landuse class: **developed,**
landuse subclass: **industrial**
wkt: POLYGON ((26.483607 ….))



daylightmap.org/earth/landuse

```sql
SELECT
    geometry_id,
    class,
    subclass,
    JSON_EXTRACT_SCALAR(names, '$.local') as name,
    CAST(JSON_EXTRACT_SCALAR(metadata, '$.surface_area_sq_m') AS double) as area,
    wkt
FROM daylight_earth
WHERE release = 'v1.18'
  AND theme = 'landuse'
  AND class = 'park'
  AND subclass <> 'grass'
  AND quadkey like '021333011%'
```

geometry_id	class	subclass	name	area	wkt
w43377655@2	park	park	Snail Lake Marsh Park	170479.69	POLYGON ((-93.1
w522449122@5	park	park	Bunker Hill Park	62740.8	POLYGON ((-93.2
w216244611@2	park	park		22752.91	POLYGON ((-93.2
w47652678@7	park	park	East Phillips Park	29728.08	POLYGON ((-93.2
w336407468@1	park	park		1591.29	POLYGON ((-93.1
w35079781@1	park	park		19078.45	POLYGON ((-93.3
w329050291@5	park	park	Erickson Park	180152.37	POLYGON ((-93.1
w1085852524@1	park	park		200599.63	POLYGON ((-93.1
w314677340@2	park	park		44831.66	POLYGON ((-92.9
w912557268@1	park	park		7700.07	POLYGON ((-93.0
w635018766@1	park	park	Wargo Court	1806.61	POLYGON ((-93.2
w960891048@1	park	park		5114.3	POLYGON ((-93.5
w36657367@8	park	park	Cedar Manor Lake Park	60402.83	POLYGON ((-93.3
w893835529@4	park	park		115032.29	POLYGON ((-92.8
w894770103@3	park	park		35738.3	POLYGON ((-92.9
w1003253264@2	park	park	Dean Parkway	17493.68	POLYGON ((-93.3
w886329369@2	park	park	Edgerton Park	71072.74	POLYGON ((-93.0



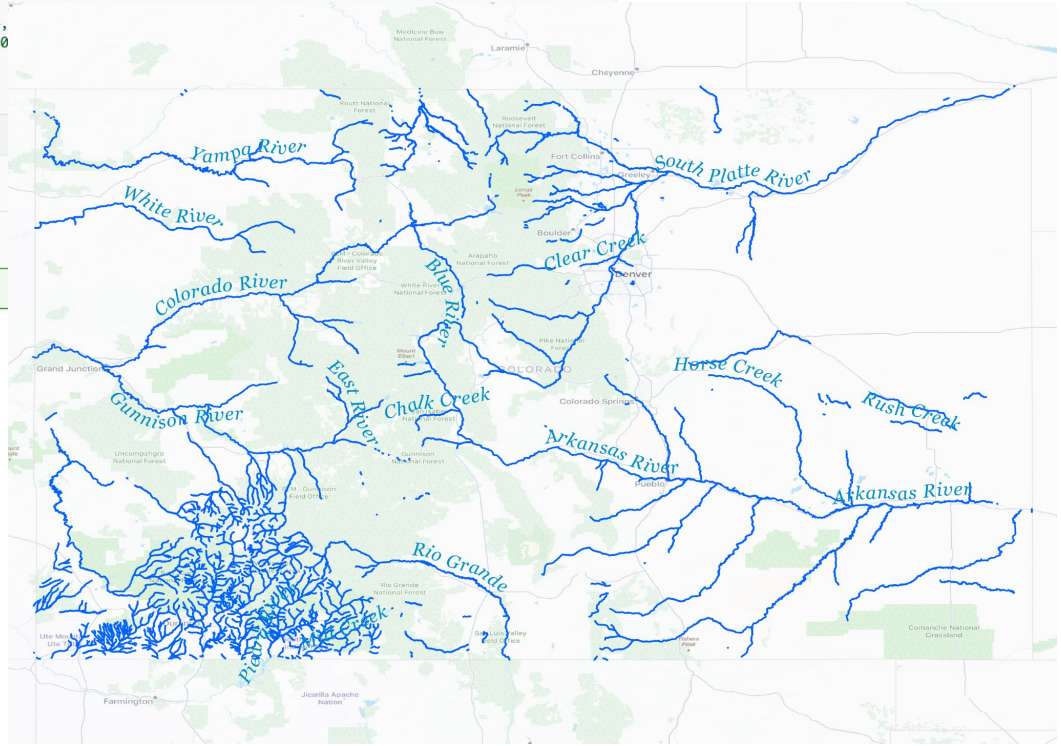
```sql
SELECT
    geometry_id,
    class,
    subclass,
    JSON_EXTRACT_SCALAR(names, '$.local') as name,
    wkt
FROM daylight_earth
WHERE theme = 'water'
  AND class = 'river'
  AND ST_CONTAINS(
    ST_GEOMETRYFROMTEXT('POLYGON((-109.06409916585731 41.008769739828494,
        ,-109.07238952929733 37.000860715568436,-109.06409916585731 41.00
    ST_GEOMETRYFROMTEXT(wkt)
    )
```

SQL Ln 5, Col 51

Run again | Explain | Cancel | Save | Clear | Create

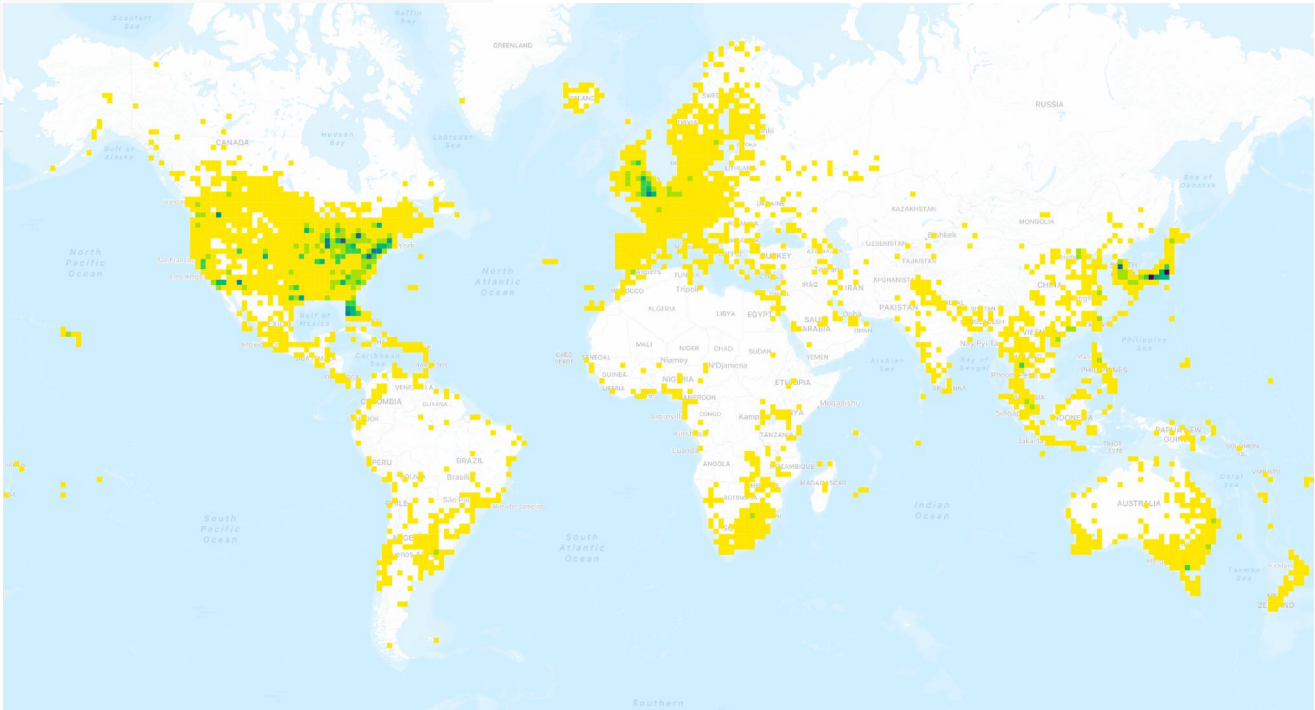
Query results | Query stats

Completed



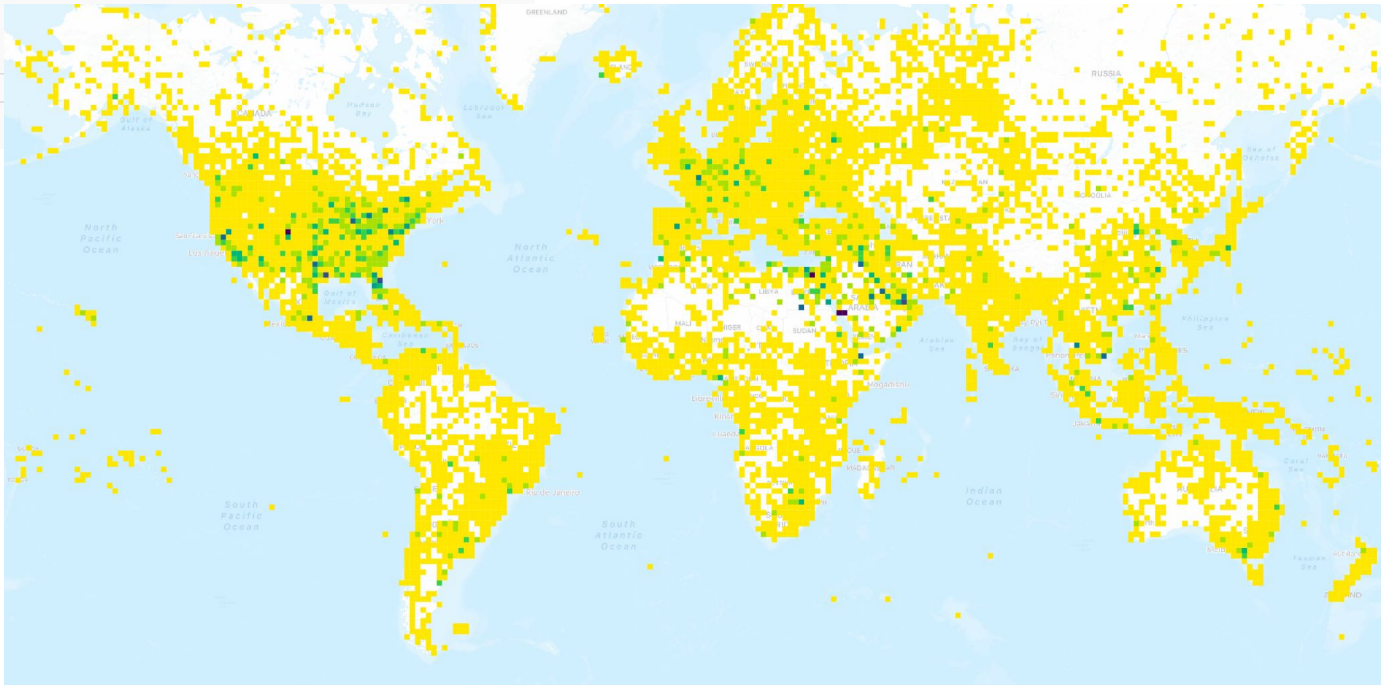
Query 3

```
1  select BING_TILE_POLYGON(BING_TILE(substr(quadkey, 1, 8))),
2      sum(
3          cast(
4              JSON_EXTRACT_SCALAR(metadata, '$.surface_area_sq_m')
5           as double)
6      ) / 1000000 AS sq_km_golf
7  from daylight_earth
8  where theme = 'landuse'
9      and class = 'golf'
10 group by substr(quadkey, 1, 8)
```



Query 3

```
1  select BING_TILE_POLYGON(BING_TILE(substr(quadkey, 1, 8))),
2      sum(
3          cast(
4              JSON_EXTRACT_SCALAR(metadata, '$.surface_area_sq_m')
5           as double)
6      ) / 1000000 AS sq_km_airport
7  from daylight_earth
8  where theme = 'landuse'
9      and class = 'airport'
10 group by substr(quadkey, 1, 8)
```



Meta

Thank you

daylightmap.org/earth

jenningsa@meta.com
jonahadkins@meta.com