DuckDB in a Spreadsheet With WASM

Feb 20, 2025
Amsterdam

Chris Laffra

If it Quacks Like a Duck...

Chris Laffra

Feb 20, 2025

Amsterdam

Ducks fly Where?

Chris Laffra

Feb 20, 2025

Amsterdam

Ruddy Shelduck

Quack by Example

Winner

Chris Laffra

Feb 20, 2025

Amsterdam

File View Help

Arial 14 normal normal bottom left

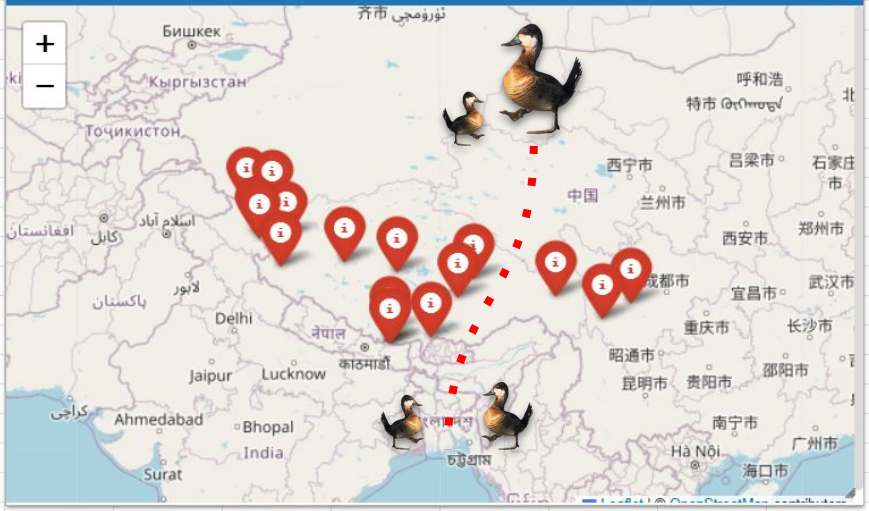
Ducks fly where?

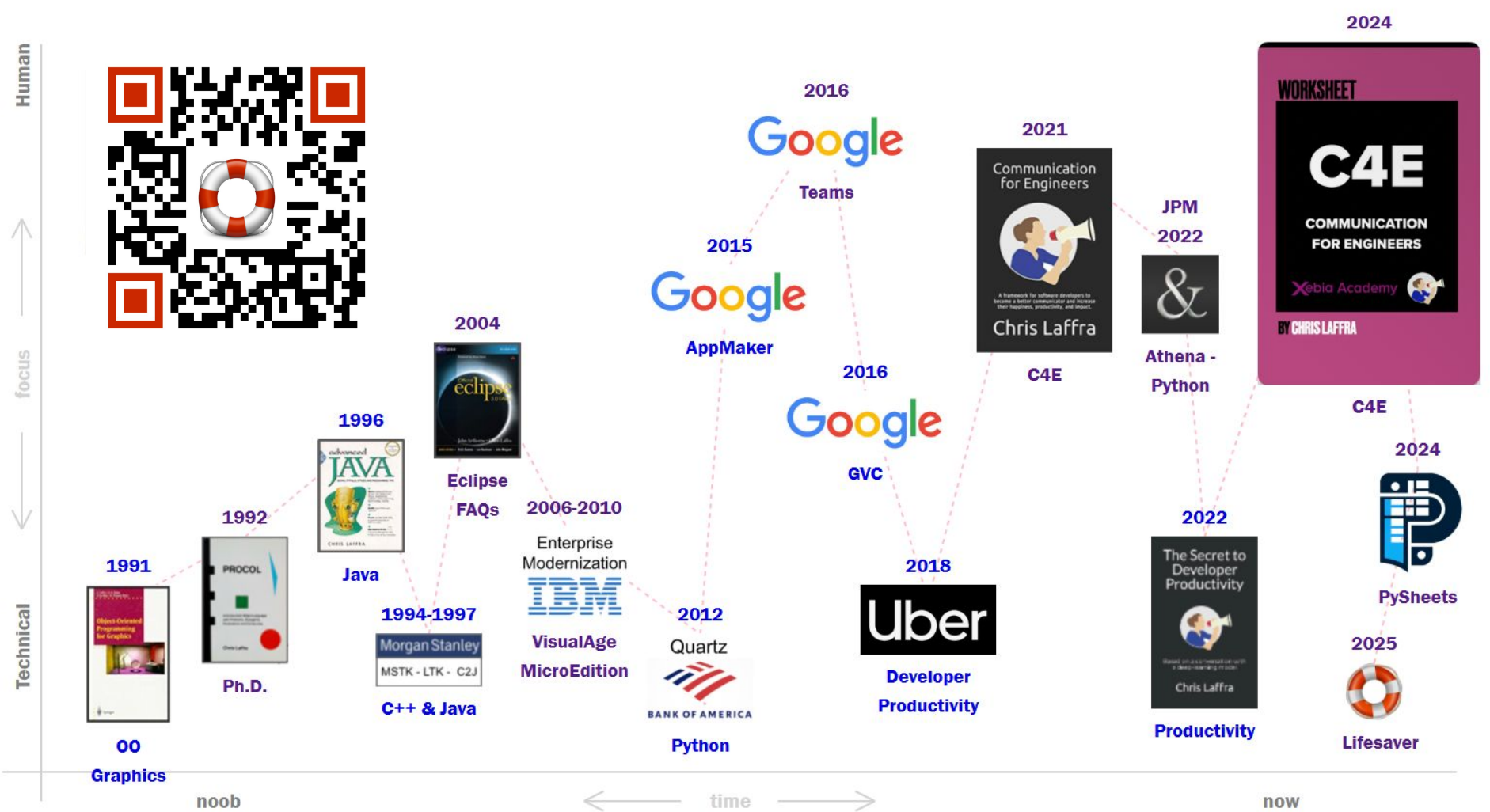
	A	B	C	D	E	F	G	H	I	J	K	L
1	DataFrame with 21 rows				Map							
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												

A1 - DataFrame with 21 rows

All	name	elevation_ft	coordinates
0	Damxung Air Base	14105	30.486273, 91.082561
1	Rikaze Dingri Airport	14108	28.604567, 86.798
2	Rutog Heliport	14859	33.65626, 80.45084
3	Gêrzê Heliport	14534	32.30204, 84.02739
4	Nyima Heliport	14997	31.78739, 87.29839
5	Seni Heliport	14730	31.414472, 91.983751
	Aksai Chin	15988	35.24569, 79.54054

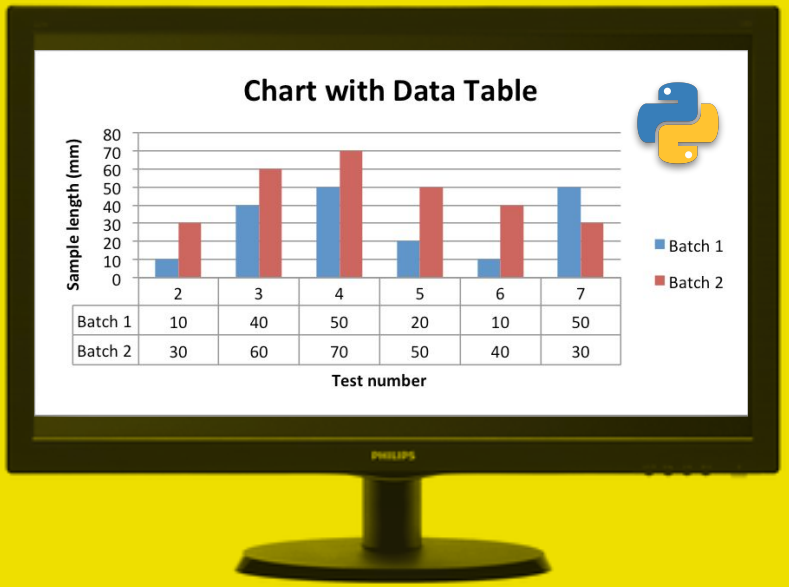
E1 - Map




Human
focus
Technical
noob
time
now
1991
Graphics
OO
1992
Ph.D.
1996
Java
1994-1997
C++ & Java
2004
Eclipse
FAQs
2006-2010
Enterprise Modernization
VisualAge MicroEdition
2012
Quartz
Python
2015
AppMaker
2016
Teams
2016
GVC
2018
Developer Productivity
2021
C4E
JPM
2022
Athena - Python
2022
Productivity
2024
C4E
2024
PySheets
2025
Lifesaver

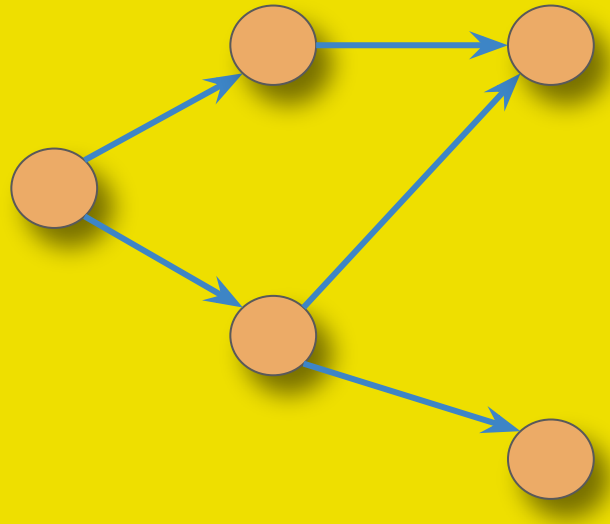
PySheets Motivation

1. Python In The Browser



	2	3	4	5	6	7
Batch 1	10	40	50	20	10	50
Batch 2	30	60	70	50	40	30

2. Dependency Graph



3 PySheets Demos

DuckDB

Dataframes

LTK

1

2

3

The LTK Kitchensink for PyScript running MicroPython took 0.084s to load

Styling | DOM | Inputs | Reactive | Tables | Custom | App | PubSub | Editor | Pitch | SVG | Canvas | Pizza | Splits | TicTacToe

Reactive LTK Demo

Product 1

Name: Wrench

Price: 50.0

Count: 10

Warranty: ☐

Delivery: 1-day

Service: With white gloves: ⬤

Summary: 10 * Wrench = $500 with service - Delivery: 1-day

Product 2

Name: Drill

Price: 500.0

Count: 6

Warranty: ☐

```
import ltk

DELIVERY_OPTIONS = ["1-day", "2-day", "pickup"]

class Product(ltk.Model):
    count: int = 10
    name: str = "Screwdriver"
    price: float = 50.0
    warranty: bool = False
    service: bool = True
    delivery: int = 1
    summary: str = ""

    def changed(self, name, value):
        if name == "summary":
            return
        self.summary = f"""
            {self.count} * {self.name} =
            ${round(self.count * self.price):,}
            {'including warranty' if self.warranty else ''}
            {'with service' if self.service else ''}
            - Delivery: {DELIVERY_OPTIONS[self.delivery]}
        """

product1 = Product(name="Wrench", delivery=0)
product2 = Product(name="Drill", count=6, price=500, service=False)

def increment_count(event):
    # Increment the default value on the class.
```

File View Help

Arial 14 normal normal bottom left

Pandas - Airports

	A	B	C	D	E	F	G	H	I	J	K
1	DataFrame with 82,274 rows			DataFrame with 21 rows			DataFrame with 21 rows		Map		
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											
32											
33											
34											
35											
36											
37											
38											
39											
40											
41											
42											
43											
44											
45											
46											
47											

A1 - DataFrame with 82,274 rows

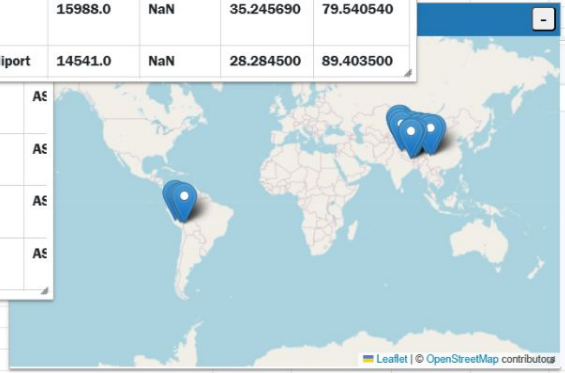
	ident	type	name	el
0	00A	heliport	Total RF Heliport	11
1	00AA	small_airport	Aero B	34
2	00A			
3	00A			
4	00A			
...	...			

D1 - DataFrame with 21 rows

	ident	type			
19113	CN-0033	small_airport	Da Ba		
19302	CN-0236	medium_airport	Ril Di Air		
19433	CN-0367	heliport	Rutog Heliport	14859.0	AS
19434	CN-0368	heliport	Gêrzê Heliport	14534.0	AS
19435	CN-0369	heliport	Nyima Heliport	14997.0	AS
19436	CN-0370	heliport	Seni Heliport	14730.0	AS

G1 - DataFrame with 21 rows

	name	elevation_ft	icao_code	lat	lon
19113	Damxung Air Base	14105.0	NaN	30.486273	91.082561
19302	Rikaze Dingri Airport	14108.0	NaN	28.604567	86.798000
19433	Rutog Heliport	14859.0	NaN	33.656260	80.450840
19434	Gêrzê Heliport	14534.0	NaN	32.302040	84.027390
19435	Nyima Heliport	14997.0	NaN	31.787390	87.298390
19436	Seni Heliport	14730.0	NaN	31.414472	91.983751
19438	Aksai Chin Heliport	15988.0	NaN	35.245690	79.540540
19439	Daklam Heliport	14541.0	NaN	28.284500	89.403500



LLM Prompt generate code | Import... | D1 | G1 | I1 | A1

```
Copy the dataframe stored in D1.

Add "lat" and "lon" columns.
Parse the comma-separated "coordinates" column.

Drop columns "ident", "type", "iso_country", "iso_region",
"municipality", "gps_code", "iata_code", "local_code",
"coordinates", and "continent".
```

Cell: G1 **Packages:** folium run script

```
1  =
2  # This code was generated by an AI. Please check it for errors.
3
4  # Copy dataframe
5  df = D1.copy()
6
7  # Add "lat" and "lon" columns
8  df["lat"] = df["coordinates"].apply(lambda x: float(x.split(",")[0]))
9  df["lon"] = df["coordinates"].apply(lambda x: float(x.split(",")[1]))
10
11 # Parse "coordinates" column
12 df["coordinates"] = df["coordinates"].apply(lambda x: x.split(","))
13
14 # Drop columns
15 df = df.drop([
16     "ident", "type", "iso_country", "iso_region", "municipality",
17     "gps_code", "iata_code", "local_code", "coordinates", "continent"
18 ], axis=1)
19
20 df
```

Console | Timeline

Filter the console... clear

```
13:19:28.400 [License] Buy Sponsor PySheets and support the PySheets team .
13:19:28.400 [Help] Learn Learn more about PySheets with our tutorials .
13:19:28.400 [UI] Running MicroPython; Python 3.4.0; UI startup took 0.040s.
13:19:32.567 [Github] Star PySheets is open-source. Give it a star .
13:19:33.377 [Worker] Running PyOdide; Python v3.12.7; Worker startup took 4.772s.
13:19:36.681 [Network] GET 200 https://raw.githubusercontent.com/datasets/airport-codes/m
13:19:36.682 [DAG] A1: Ran once in worker, taking 1.544s
13:19:38.028 [DAG] D1: Ran in worker 2 times, last run took 0.015s
13:19:39.176 [DAG] G1: Ran in worker 3 times, last run took 0.011s
```

built with LTK

File View Help

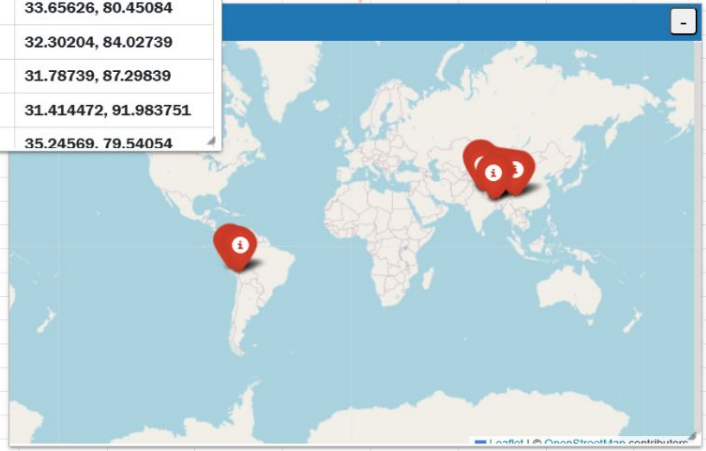
Arial 14 normal normal bottom left

Ducks fly where?

	A	B	C	D	E	F	G	H	I
1	DataFrame with 21 rows				Map				

A1 - DataFrame with 21 rows

All	name	elevation_ft	coordinates
0	Damxung Air Base	14105	30.486273, 91.082561
1	Rikaze Dingri Airport	14108	28.604567, 86.798
2	Rutog Heliport	14859	33.65626, 80.45084
3	Gêrzê Heliport	14534	32.30204, 84.02739
4	Nyima Heliport	14997	31.78739, 87.29839
5	Seni Heliport	14730	31.414472, 91.983751
	Aksai Chin	15988	35.24569, 79.54054



LLM Prompt generate code Import... A1 E1

Enter your prompt here...

Cell: A1 **Packages:** fsspec duckdb folium run script

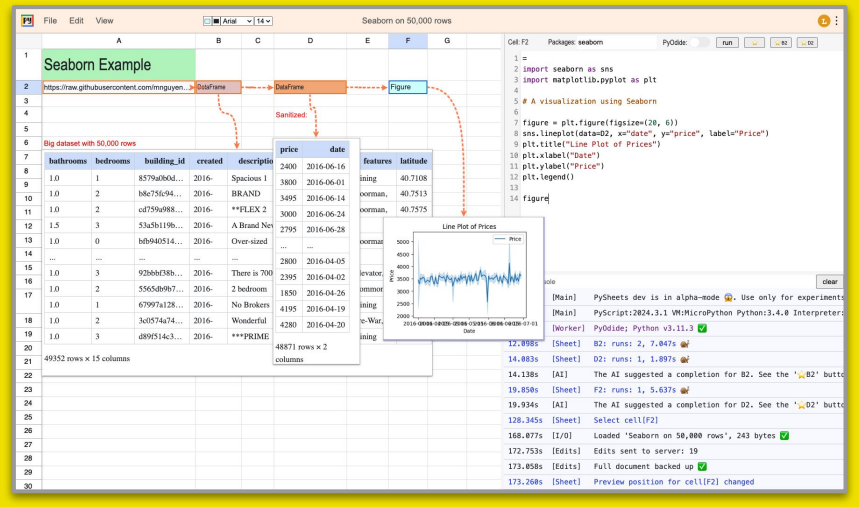
```
1  =
2  import io
3  import urllib.request
4  import duckdb
5
6  url = "https://raw.githubusercontent.com/datasets/airport-codes/master/data/airport-codes.c
7  csv_bytes = io.BytesIO(urllib.request.urlopen(url).read())
8  csv_sql = duckdb.read_csv(csv_bytes)
9  select = '''SELECT name,elevation_ft,coordinates FROM csv_sql'''
10 where = '''WHERE elevation_ft > 14000'''
11 duckdb.sql(f"{select} {where}").df()
12
```

Console Timeline

Filter the console... clear

```
13:24:02.752 [License] Buy Sponsor PySheets and support the PySheets team 🌷.
13:24:02.753 [Help] Learn Learn more about PySheets with our tutorials 🎓.
13:24:02.753 [UI] Running MicroPython; Python 3.4.0; UI startup took 0.043s.
13:24:06.259 [Github] Star PySheets is open-source. Give it a star ⭐.
13:24:06.720 [Worker] Running PyOdide; Python v3.12.7; Worker startup took 3.809s.
13:24:10.963 [Network] GET 200 https://raw.githubusercontent.com/datasets/airport-codes/mast
13:25:00.772 [DAG] A1: Ran in worker 5 times, last run took 0.654s
```

PySheets Technology



Quack by Example

Fig 1a. Examples of tabular QBE

Sailors	*sid*	*sname*	*rating*	*age*
P.			10	

1975

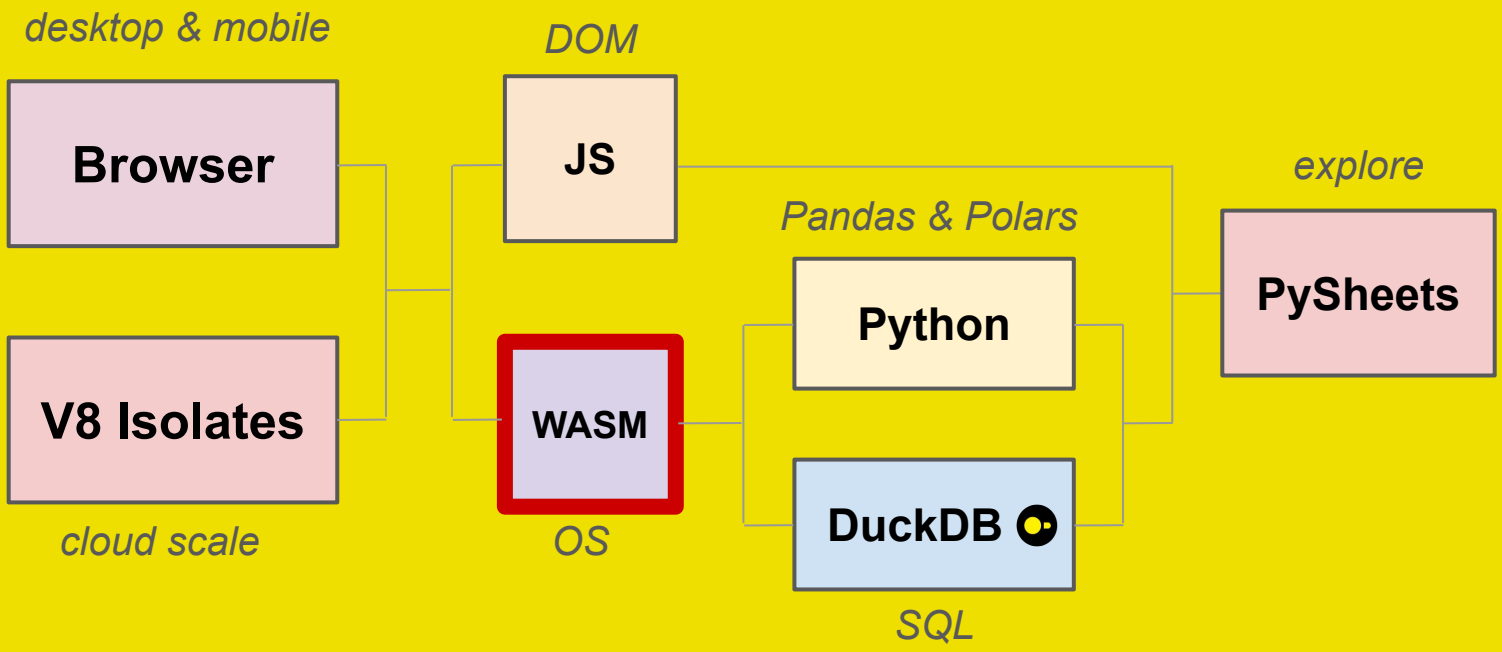
A2 - DataFrame with 21 rows

All	name	elevation_ft	coordinates
0	Damxung Air Base	14105	30.486273, 91.082561
1	Rikaze Dingri Airport	14108	28.604567, 86.798
2	Rutog Heliport	14859	33.65626, 80.45084
3	Gêrzê Heliport	14534	32.30204, 84.02739
4	Nyima Heliport	14997	31.78739, 87.29839
5	Seni Heliport	1	
6	Aksai Chin Heliport	1	
7	Daklam Heliport	1	
8	Daulat Beg Oldi Advanced Landing Ground	1	
9	Chushul Airstrip	1	

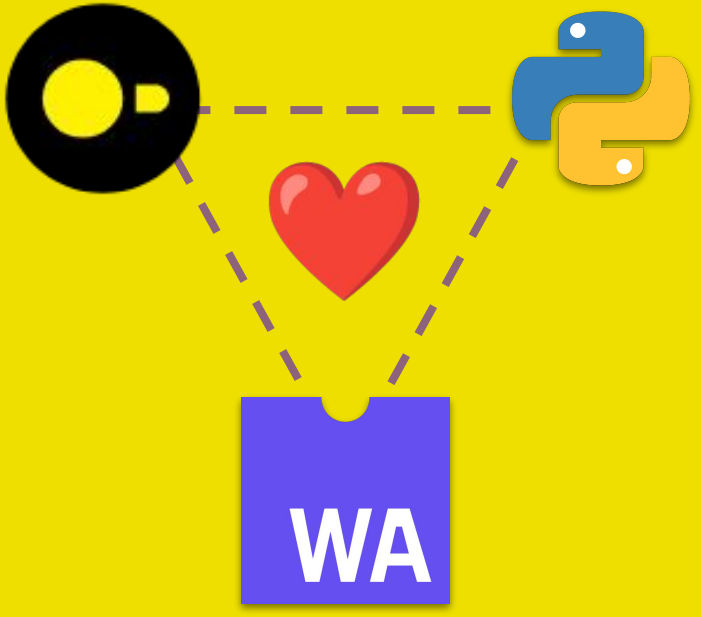
E2 - Map

2025

WASM = The Future



Summary



Questions?

PySheets

C4E

Lifesaver