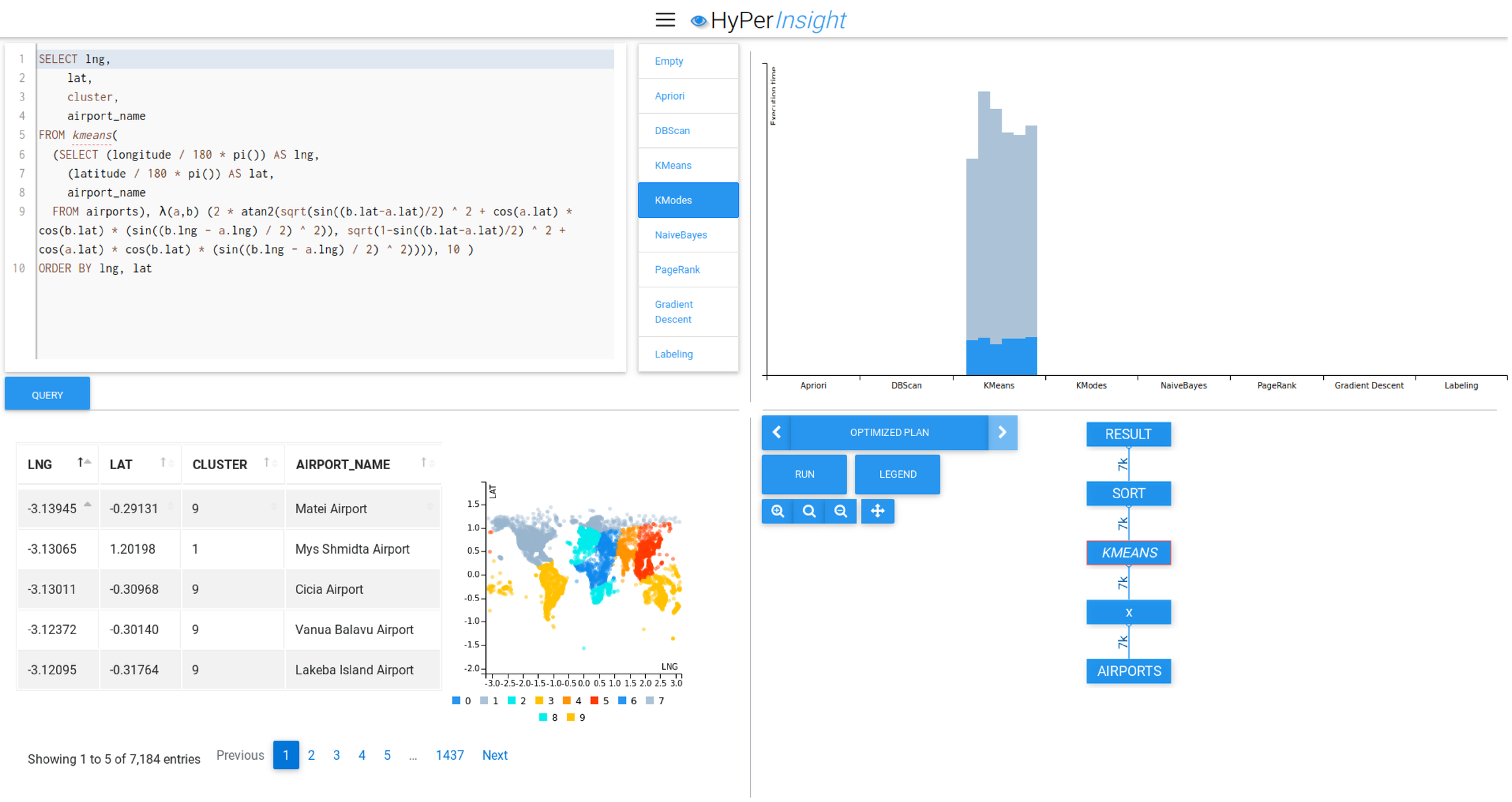
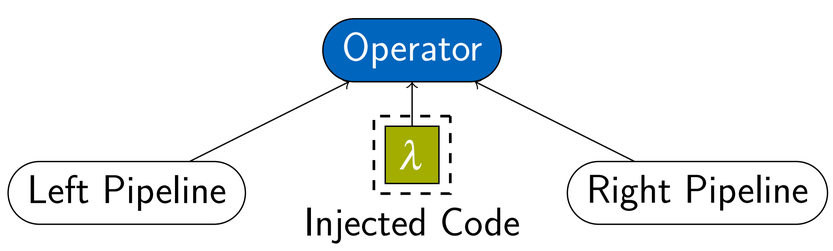
Chair III: Database Systems, Professorship of Data Mining and Analytics
Chair XXV: Data Science and Engineering
Department of Informatics
Technical University of Munich

The Power of SQL Lambda Functions

Maximilian E. Schüle, Dimitri Vorona, Linnea Passing, Harald Lang, Alfons Kemper, Stephan Günnemann, Thomas Neumann
{schuele,vorona,passing,langh,kemper,guennemann,neumann}@in.tum.de

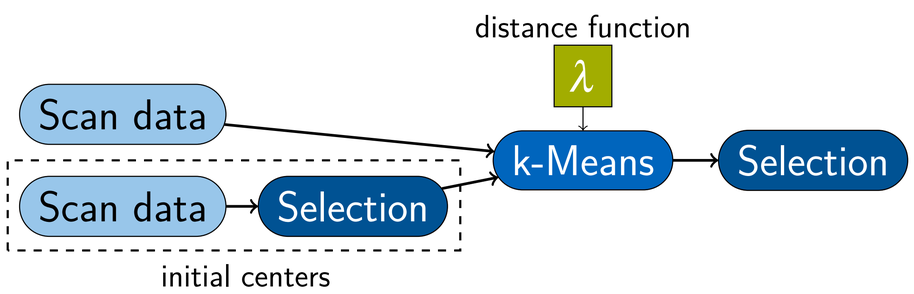


Lambda Functions for SQL



+ Allow users to inject expressions in operators
+ Applicable on unary and binary operators as well

Example: k-Means



+ Distance function as lambda expression
+ User-specified distance functions are allowed

```
CREATE TABLE data(x FLOAT, y INT);
CREATE TABLE center(x FLOAT, y INT);
INSERT INTO ...

SELECT * FROM kmeans(
  (SELECT x,y FROM data),
  (SELECT x,y FROM center),
  -- the distance function
  λ(a,b) (a.x-b.x)^2+(a.y-b.y)^2,
  3 -- max. number of iterations
);
```

Listing 1: k-Means.

```
CREATE TABLE edges (a BIGINT,b BIGINT);
INSERT INTO ...



SELECT * FROM pagerank(
   (SELECT * FROM edges),
   λ(src) src.a, -- source
   λ(dst) dst.b, -- destination
   0.85,     -- damping factor
   0.00001, -- threshold
   100       -- iterations
);
```

Listing 2: PageRank.

```
CREATE TABLE data (x FLOAT, y FLOAT);
CREATE TABLE weights(a FLOAT, b FLOAT);
INSERT INTO ...

SELECT * FROM gradientdescent(
  -- the loss function
  λ(d,w) (w.a*d.x+w.b-d.y)^2,
  -- training data set
  (SELECT x,y FROM data d),
  -- initial weights
  (SELECT a,b FROM weights w),
  0.05, -- learning rate
  100    -- max. number of iteration
);
```

Listing 3: Gradient descent.

```
CREATE TABLE testdata (x FLOAT);
CREATE TABLE weights(a FLOAT, b FLOAT);
INSERT INTO ...



SELECT * FROM labeling(
  -- the model function
  λ(d,w) (w.a*d.x+w.b),
  -- test data set
  (SELECT x,y FROM testdata d),
  -- optimized weights
  (SELECT a,b FROM weights w)
);
```

Listing 4: Labeling.

hyper-db.com