SELECT count(\*) FROM `my-data-403501.DA04\_K296.pizza`; --300 dòng

SELECT \* FROM `my-data-403501.DA04\_K296.pizza`

LIMIT 1000;

-- Build model

-- Scale by min-max scaler

-- th1: k=2

CREATE OR REPLACE MODEL DA04\_K296.pizza\_clustering\_mm\_2

OPTIONS(model\_type='kmeans',num\_clusters=2)

AS

SELECT ML.MIN\_MAX\_SCALER(fat) over() as fat\_mm,

ML.MIN\_MAX\_SCALER(cal) over() as cal\_mm

FROM `my-data-403501.DA04\_K296.pizza`;

-- th1: k=3

CREATE OR REPLACE MODEL DA04\_K296.pizza\_clustering\_mm\_3

OPTIONS(model\_type='kmeans',num\_clusters=3)

AS

SELECT ML.MIN\_MAX\_SCALER(fat) over() as fat\_mm,

ML.MIN\_MAX\_SCALER(cal) over() as cal\_mm

FROM `my-data-403501.DA04\_K296.pizza`;

--chọn k=3

-- clustering

SELECT CENTROID\_ID, fat\_mm, cal\_mm, fat, cal

FROM

ML.PREDICT( MODEL `DA04\_K296.pizza\_clustering\_mm\_3`,(

SELECT ML.MIN\_MAX\_SCALER(fat) over() as fat\_mm,

ML.MIN\_MAX\_SCALER(cal) over() as cal\_mm, fat, cal

FROM `my-data-403501.DA04\_K296.pizza`

));