Data Schema

1. Q0

For the ride prediction task, we collect open NYC transportation data ranging from 2016-01-01 to 2016-06-30. The table schema is shown below.

id,string
vendor_id,int
pickup_datetime,timestamp
dropoff_datetime,timestamp
passenger_count,int
pickup_longitude,double
pickup_latitude,double
dropoff_longitude,double
dropoff_latitude,double
store_and_fwd_flag,string
trip_duration,int

Train

The data schema consists of a single table with 11 columns, including two timestamp columns, two string columns, and seven numeric columns.

2. Q1

For the flu forecast task, we have collected covid19 data ranging from 2020-01-22 to 2020-04-07. The table schema is shown below.

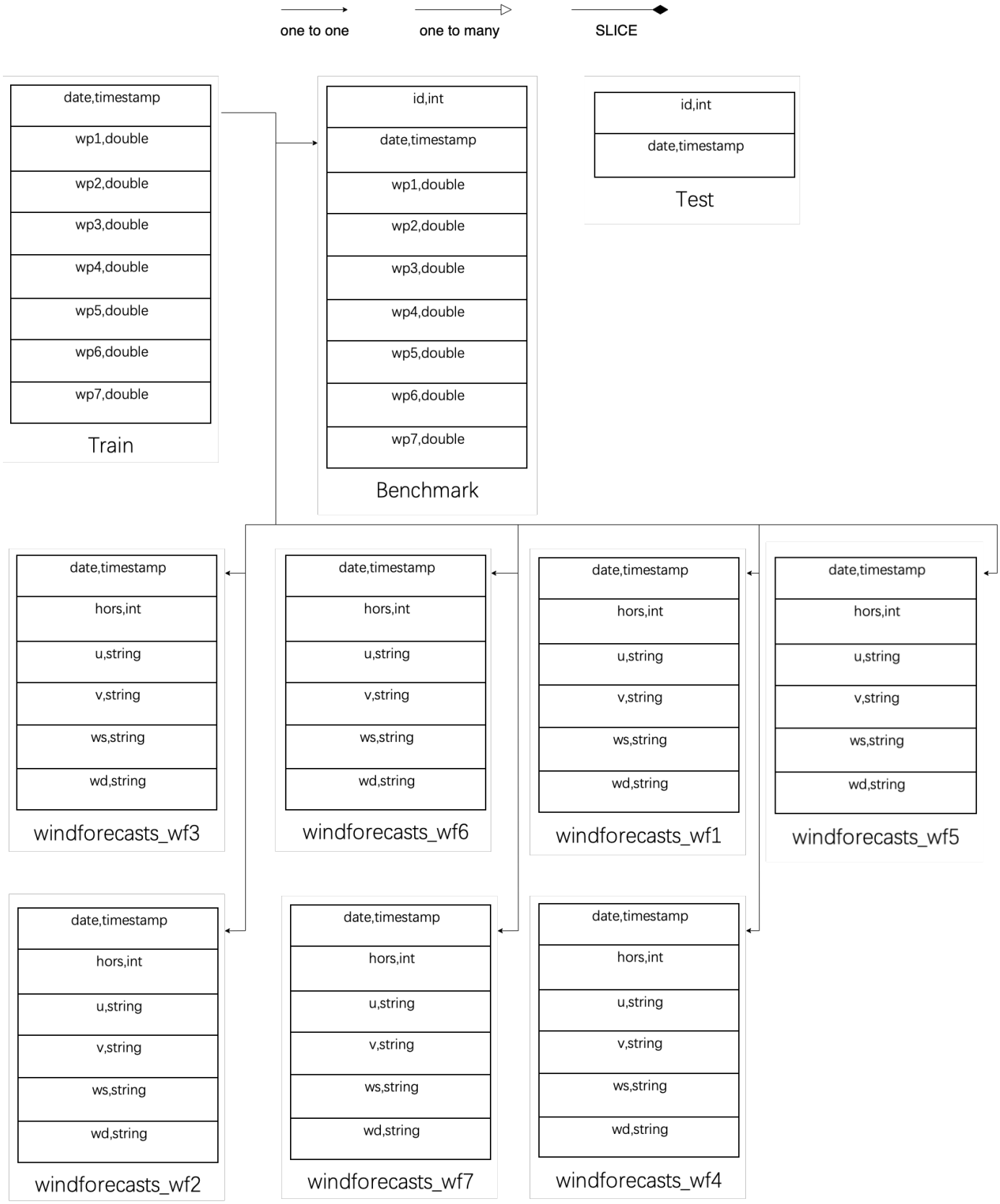
Id,int
Province_State,string
Country_Region,string
Date,timestamp
ConfirmedCases,double
Fatalities,double

Train

The data schema contains one base table with only 6 columns, including one timestamp columns, two string columns, three numeric columns.

3. Q2

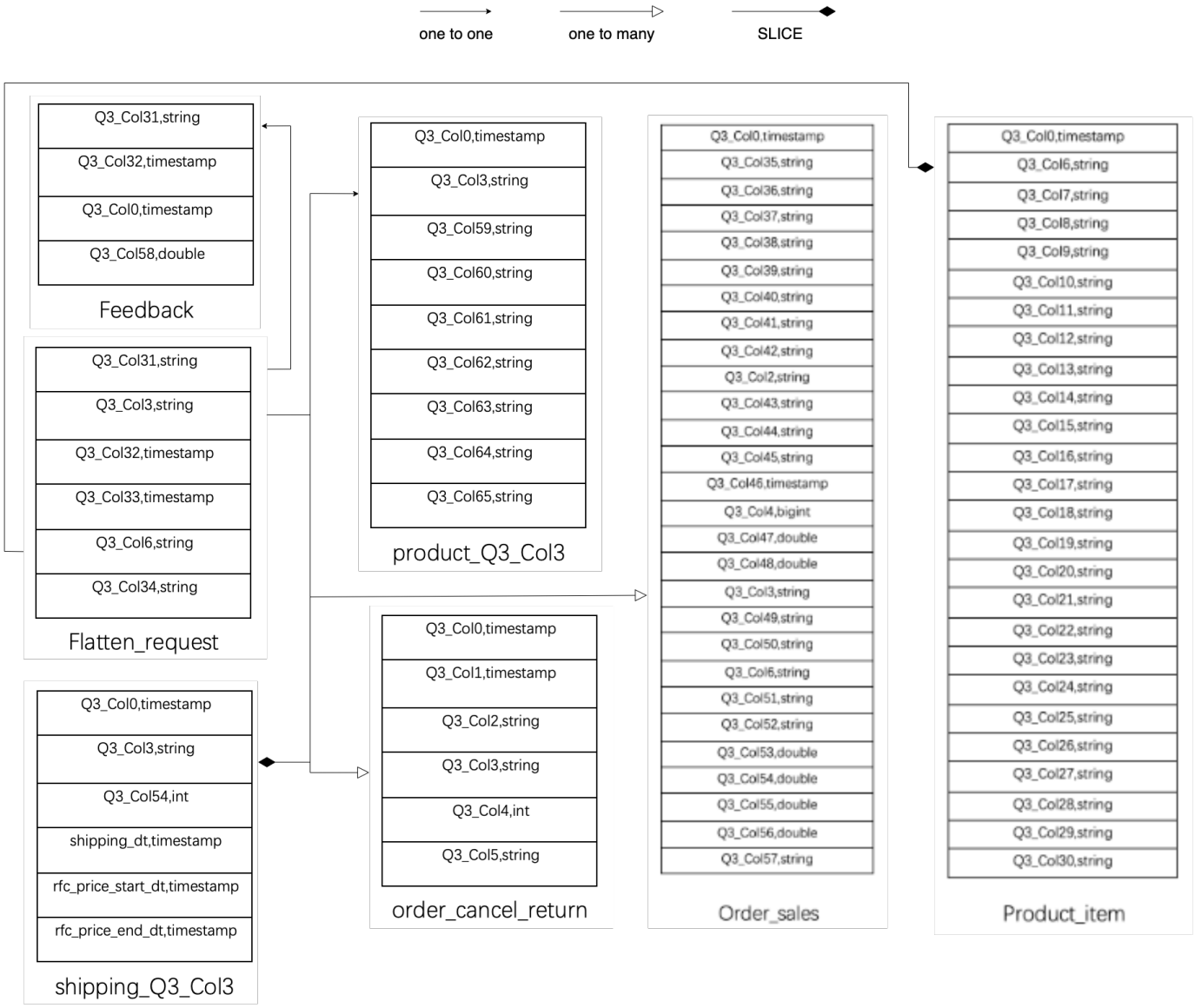
For the power forecast task, we have collected the open wind-power data ranging from 2009-07-01 to 2012-06-26. The table schema is shown below.



The data schema contains 10 tables with 61 columns, in which seven tables separately record the power changes in each farm (farm tables) and the rest three tables are used to train the prediction model (training tables).

4. Q3

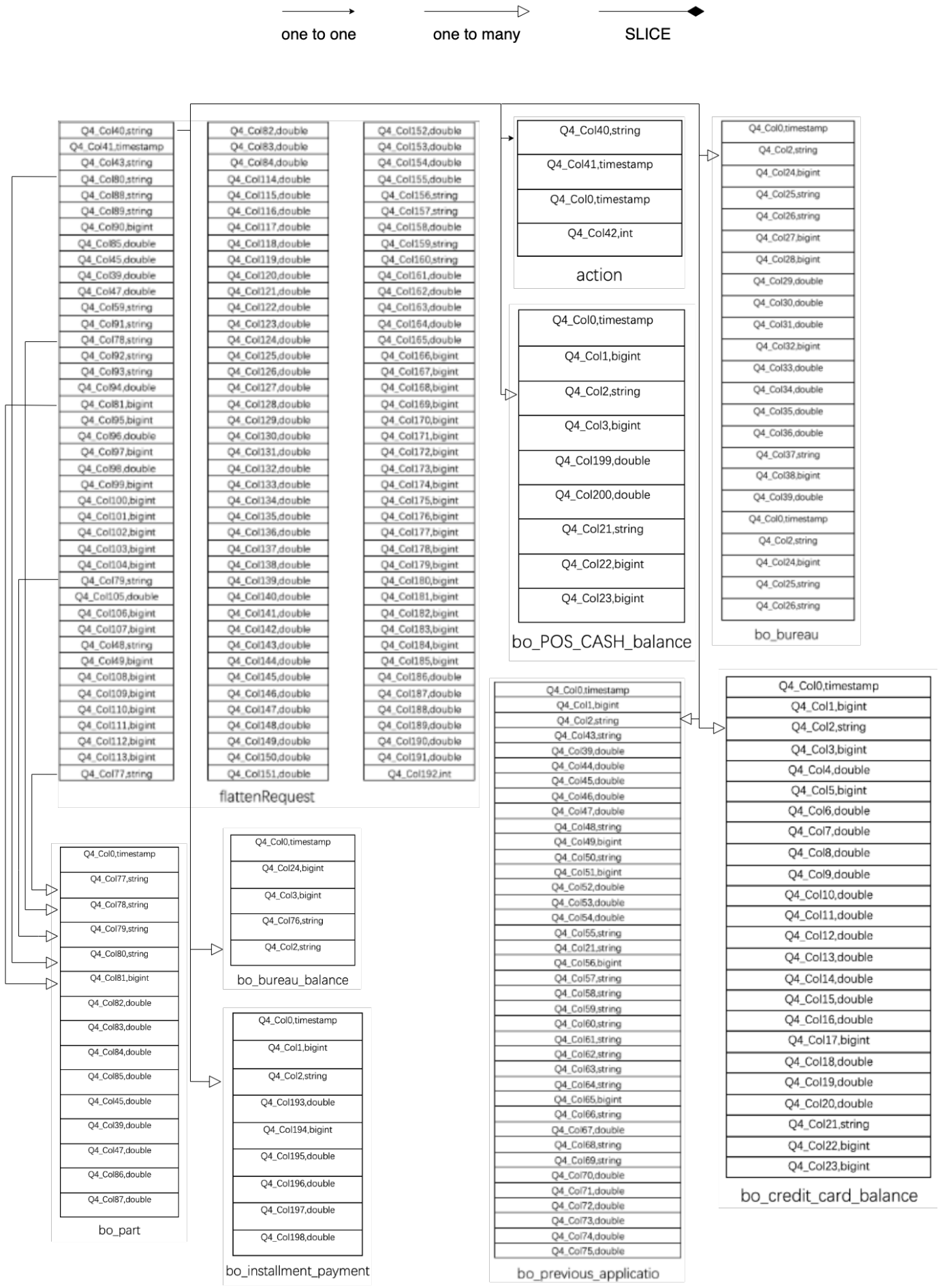
For the sales-prediction task, we have collected real data in Uniqlo ranging from 2017-12-31 to 2021-05-30. The table schema is shown below.



The data schema owns 7 tables together with 85 columns, in which there are both steam tables, such as storing (canceled) orders, and attribute tables like product information.

5. Q4

The task aims to predict whether customers will pay back their loans on time for a credit card company. Q4 owns 9 tables together with 1GB data and 245 columns.


one to one
one to many
SLICE
Q4_Col40,string
Q4_Col41,timestamp
Q4_Col43,string
Q4_Col80,string
Q4_Col88,string
Q4_Col89,string
Q4_Col90,bigint
Q4_Col85,double
Q4_Col45,double
Q4_Col39,double
Q4_Col47,double
Q4_Col59,string
Q4_Col91,string
Q4_Col78,string
Q4_Col92,string
Q4_Col93,string
Q4_Col94,double
Q4_Col81,bigint
Q4_Col95,bigint
Q4_Col96,double
Q4_Col97,bigint
Q4_Col98,double
Q4_Col99,bigint
Q4_Col100,bigint
Q4_Col101,bigint
Q4_Col102,bigint
Q4_Col103,bigint
Q4_Col104,bigint
Q4_Col79,string
Q4_Col105,double
Q4_Col106,bigint
Q4_Col107,bigint
Q4_Col48,string
Q4_Col49,bigint
Q4_Col108,bigint
Q4_Col109,bigint
Q4_Col110,bigint
Q4_Col111,bigint
Q4_Col112,bigint
Q4_Col113,bigint
Q4_Col77,string
Q4_Col82,double
Q4_Col83,double
Q4_Col84,double
Q4_Col114,double
Q4_Col115,double
Q4_Col116,double
Q4_Col117,double
Q4_Col118,double
Q4_Col119,double
Q4_Col120,double
Q4_Col121,double
Q4_Col122,double
Q4_Col123,double
Q4_Col124,double
Q4_Col125,double
Q4_Col126,double
Q4_Col127,double
Q4_Col128,double
Q4_Col129,double
Q4_Col130,double
Q4_Col131,double
Q4_Col132,double
Q4_Col133,double
Q4_Col134,double
Q4_Col135,double
Q4_Col136,double
Q4_Col137,double
Q4_Col138,double
Q4_Col139,double
Q4_Col140,double
Q4_Col141,double
Q4_Col142,double
Q4_Col143,double
Q4_Col144,double
Q4_Col145,double
Q4_Col146,double
Q4_Col147,double
Q4_Col148,double
Q4_Col149,double
Q4_Col150,double
Q4_Col151,double
Q4_Col152,double
Q4_Col153,double
Q4_Col154,double
Q4_Col155,double
Q4_Col156,string
Q4_Col157,string
Q4_Col158,double
Q4_Col159,string
Q4_Col160,string
Q4_Col161,double
Q4_Col162,double
Q4_Col163,double
Q4_Col164,double
Q4_Col165,double
Q4_Col166,bigint
Q4_Col167,bigint
Q4_Col168,bigint
Q4_Col169,bigint
Q4_Col170,bigint
Q4_Col171,bigint
Q4_Col172,bigint
Q4_Col173,bigint
Q4_Col174,bigint
Q4_Col175,bigint
Q4_Col176,bigint
Q4_Col177,bigint
Q4_Col178,bigint
Q4_Col179,bigint
Q4_Col180,bigint
Q4_Col181,bigint
Q4_Col182,bigint
Q4_Col183,bigint
Q4_Col184,bigint
Q4_Col185,bigint
Q4_Col186,double
Q4_Col187,double
Q4_Col188,double
Q4_Col189,double
Q4_Col190,double
Q4_Col191,double
Q4_Col192,int
flattenRequest
Q4_Col40,string
Q4_Col41,timestamp
Q4_Col0,timestamp
Q4_Col42,int
action
Q4_Col0,timestamp
Q4_Col1,bigint
Q4_Col2,string
Q4_Col3,bigint
Q4_Col199,double
Q4_Col200,double
Q4_Col21,string
Q4_Col22,bigint
Q4_Col23,bigint
bo_POS_CASH_balance
Q4_Col0,timestamp
Q4_Col2,string
Q4_Col24,bigint
Q4_Col25,string
Q4_Col26,string
Q4_Col27,bigint
Q4_Col28,bigint
Q4_Col29,double
Q4_Col30,double
Q4_Col31,double
Q4_Col32,bigint
Q4_Col33,double
Q4_Col34,double
Q4_Col35,double
Q4_Col36,double
Q4_Col37,string
Q4_Col38,bigint
Q4_Col39,double
Q4_Col0,timestamp
Q4_Col2,string
Q4_Col24,bigint
Q4_Col25,string
Q4_Col26,string
bo_bureau
Q4_Col0,timestamp
Q4_Col77,string
Q4_Col78,string
Q4_Col79,string
Q4_Col80,string
Q4_Col81,bigint
Q4_Col82,double
Q4_Col83,double
Q4_Col84,double
Q4_Col85,double
Q4_Col45,double
Q4_Col39,double
Q4_Col47,double
Q4_Col86,double
Q4_Col87,double
bo_part
Q4_Col0,timestamp
Q4_Col24,bigint
Q4_Col3,bigint
Q4_Col76,string
Q4_Col2,string
bo_bureau_balance
Q4_Col0,timestamp
Q4_Col1,bigint
Q4_Col2,string
Q4_Col193,double
Q4_Col194,bigint
Q4_Col195,double
Q4_Col196,double
Q4_Col197,double
Q4_Col198,double
bo_installment_payment
Q4_Col0,timestamp
Q4_Col1,bigint
Q4_Col2,string
Q4_Col43,string
Q4_Col39,double
Q4_Col44,double
Q4_Col45,double
Q4_Col46,double
Q4_Col47,double
Q4_Col48,string
Q4_Col49,bigint
Q4_Col50,string
Q4_Col51,bigint
Q4_Col52,double
Q4_Col53,double
Q4_Col54,double
Q4_Col55,string
Q4_Col21,string
Q4_Col56,bigint
Q4_Col57,string
Q4_Col58,string
Q4_Col59,string
Q4_Col60,string
Q4_Col61,string
Q4_Col62,string
Q4_Col63,string
Q4_Col64,string
Q4_Col65,bigint
Q4_Col66,string
Q4_Col67,double
Q4_Col68,string
Q4_Col69,string
Q4_Col70,double
Q4_Col71,double
Q4_Col72,double
Q4_Col73,double
Q4_Col74,double
Q4_Col75,double
bo_previous_applicatio
Q4_Col0,timestamp
Q4_Col1,bigint
Q4_Col2,string
Q4_Col3,bigint
Q4_Col4,double
Q4_Col5,bigint
Q4_Col6,double
Q4_Col7,double
Q4_Col8,double
Q4_Col9,double
Q4_Col10,double
Q4_Col11,double
Q4_Col12,double
Q4_Col13,double
Q4_Col14,double
Q4_Col15,double
Q4_Col16,double
Q4_Col17,bigint
Q4_Col18,double
Q4_Col19,double
Q4_Col20,double
Q4_Col21,string
Q4_Col22,bigint
Q4_Col23,bigint
bo_credit_card_balance

6. Q5

Q5 owns 10 tables with over 13GB origin data and 773 columns. The schema is shown below.