

QLIKVIEW - DIRECT DISCOVERY

Contents

1. Introduction	2
1.1 Data Loading	2
1.1.1 Dimension	3
1.1.2 Measure	3
1.1.3 Detail	3
1.2 Qlikview Object	4
1.3 Data Sources	4
1.4 Cancelling a Direct Discovery Query	4
1.5 Unsupported Qlikview Functionality	4
1.6 Features	4
1.7 Drawback	5
1.8 Security	5
1.9 In Memory.....	5
1.10 Direct Discovery	6
1.11 Reference	6

1. Introduction

Qlikview Direct Discovery capability expands the potential use cases for Business Discovery, enabling business users to conduct associative analysis on larger data sources.

The data for Direct Discovery Fields are retrieved from the database everytime and In-Memory Fields are already fetched and stored in RAM as qlikview normally works.

Direct Discovery feature is aimed for big data collections that do not fit into memory during load.

Direct Discovery is used for real time analysis in Dashboards. For example in share market, every minute the share values are keep on changing to see the changes in Dashboard as and when it changes Direct Discovery is very useful.

1.1 Data Loading

To apply the direct discovery method, the keyword “SQL SELECT” is replaced with “DIRECT QUERY” in order to get the data from the corresponding database and additional keywords – DIMENSION, MEASURE, DETAIL and NATIVE are introduced.

ODBC CONNECT TO Database;

DIRECT QUERY

DIMENSION

InvoiceId,
InvoiceDate,

NATIVE('month([InvoiceDate])') as InvoiceMonth,

NATIVE('Year([InvoiceDate])') as InvoiceYear

MEASURE

InvoiceAmount,
Paid

DETAIL

ShippedDate,
InvoiceNumber

DETACH

SalesOrderID,
CustomerID

FROM Invoice;

In the example above, with the use of “DIRECT QUERY” and “DIMENSION” keywords, only columns like InvoiceId and InvoiceDate are loaded into memory as symbol tables. Other columns following the “MEASURE” and “DETAIL” keywords are exist in the source data table within the database and they are not part of “in memory” data model.

Preceding load cannot be used with Direct Discovery as it only relates to the data that is loaded in memory.

It is also possible to execute source Database SQL functions with the direct discovery table on the load script by using the keyword NATIVE which should be used within single quotation marks.

NATIVE('month([InvoiceDate])') as InvoiceMonth,

NATIVE('Year([InvoiceDate])') as InvoiceYear

The following table describes the possible usage of Direct Query Keywords.

Field Type	In Memory?	Forms Association?	Used in List Box?	Used in Chart expressions?	Used in Table Box?
DIMENSION	Yes	Yes	Yes	Yes	Yes
MEASURE	No	No	In aggr() expressions only	Yes	Yes
DETAIL	No	No	No	No	Yes
DETACH	Yes	No	Yes	Yes	Yes

1.1.1 Dimension

These are the fields that are listed after DIMENSION keywords in the load script. The unique data values of these fields are loaded into the “in memory” symbol tables.

The main use for dimension fields is:

- To define dimension chart values
- To create the association SQL between the in-memory data and the direct discovery data
- To define list boxes with direct discovery data

1.1.2 Measure

A measure field is a field that Qlikview is aware of on a “meta level”. The measure fields are required to have an aggregation function when used on the user interface.

For example, if the expression Sum(InvoiceAmount) is used in a Qlikview chart and InvoiceAmount is a measure field will do the sum in database not in qlikview.

1.1.3 Detail

These are the fields that the user may want to display in a drill-to-details table box but that should *not be* involved in any chart expressions. The data will be displayed at the lowest level without aggregation within a table box.

The Detach keyword has been introduced to flag certain dimension fields NOT to be part of the associative experience BUT to be part of the query generated passed to the data source.

1.2 Qlikview Object

- Detail records can only be displayed in a Table box.
- Statistics boxes are not supported with measure direct discovery fields.
- The `aggr()` function should be used to show the aggregated value of the measure field for any dimension field.
- Only measure direct discovery fields can be used in expression and dimension direct discovery fields can be used as dimensions.
- It is also possible to use the in memory fields and direct discovery fields on the same chart. In these cases, Qlikview associative automatically handle the associations between in memory fields and direct discovery fields.

1.3 Data Sources

Qlikview Direct Discovery can only be used against SQL compliant data sources. The following data sources are supported

- ODBC/OLEDB data sources - All ODBC/OLEDB sources are supported, including SQLServer, Teradata and Oracle.
- Custom connectors which support SQL – SAP SQL Connector, Custom QVX connectors for SQL compliant data stores.
- SAP HANA, Cloudera/MapR/Horton Works Hadoop, HP Vertica and Parstream via ODBC.

Both the 32-bit and 64-bit connections are supported.

1.4 Cancelling a Direct Discovery Query

The cancel icon on Qlikview charts can be used to abort the direct discovery query. The direct discovery query running for a chart will be automatically aborted when the Qlikview application is closed or a new tab is selected.

1.5 Unsupported Qlikview Functionality

The following Qlikview features are not supported

- Set Analysis
- Direct Discovery Measure and Detail fields are not supported on Global Search
- Desktop Section access and data reduction
- Loop and Reduce
- Synthetic keys on the Direct Discovery table
- Pivot table and Mini charts are not supported

1.6 Features

- Possible to apply the section access on Qlikview server
- Single Calculated dimensions are supported
- Possible to rename the Direct Discovery table
- Possible to Global search on the dimension fields
- Can reflect updated records without reloads (not new records)
- Can support more than one Direct Discovery table in certain scenarios

1.7 Drawback

- Performance slow when compare to memory apps
- Direct Discovery will always be slower that evaluating data that has been loaded into memory during load.
- It is not a real time solution
- Direct Discovery can be used with more than one table **only where the cardinality of the key field in the join is low. Cardinality denotes the uniqueness of data in a column.**

1.8 Security

Some security best practices should be taken into considerations when using Direct Discovery feature.

- All of the users using the same Qlikview application with the Direct Discovery capability will be using the same connection. Authentication pass through or credentials per user are not supported.
- There is no logging capability
- It is possible to flood the database with requests from the client.
- It is possible to get detailed error messages from the Qlikview Server log files.

1.9 In Memory

Taking the following script as an example:

```
LOAD CustomerID, SalesPersonID, SalesOrderID, SubTotal;  
SQL SELECT CustomerID, SalesPersonID, SalesOrderID, SubTotalFROM SalesOrderHeader;
```

During the reload process the data is transformed into two table type one associative data table and multiple symbol tables (one per field). The symbol tables contain one row per distinct value of the field. Each row contains a pointer and the value of the field, both the numeric value and the textual component.

Unique values are loaded into the structures (symbol tables)

<u>CustomerID</u>		<u>SalesPersonID</u>		<u>SalesOrderID</u>		<u>SubTotal</u>	
Pointer	Value	Pointer	Value	Pointer	Value	Pointer	Value
10	Acme Ltd	100	Paul S.	1101	AB1234	10001	2345.57
11	Smith Co.	101	Mary D.	1110	AC1234	10010	45.78

The associative data table is simultaneously populated

<u>CustomerID</u>	<u>SalesPersonID</u>	<u>SalesOrderID</u>	<u>SubTotal</u>
10	100	1101	10001
11	100	1110	10010
10	101	1110	10001
10	100	1101	10010
...

1.10 Direct Discovery

Modifying the script to enable Direct Discovery:

DIRECT QUERY DIMENSION CustomerID, SalesPersonID, SalesOrderID

MEASURE SubTotal

FROM SalesOrderHeader;

During the reload process the only the DIMENSION data is transformed into multiple symbol tables (DETAIL and MEASURE column data remains in the source) and the associative data table remains empty with only the metadata being loaded (column names)

CustomerID		SalesPersonID		SalesOrderID		SubTotal
Pointer	Value	Pointer	Value	Pointer	Value	
10	Acme Ltd	100	Paul S.	1101	AB1234	
11	Smith Co.	101	Mary D.	1110	AC1234	

The associative data table is not populated on reload and created by selections (with caching)

CustomerID	SalesPersonID	SalesOrderID	SubTotal



The association table is populated on the fly per user click with the result set being cached. The exception to this rule would be dimension fields flagged as “DETACH” which would not populate the association data table at all resulting in a faster query execution BUT without any associations.

1.11 Reference

- Qlikview Direct Discovery FAQ 11.2_SR7
- DS-Qlikview -Direct-Discovery-EN