



Getting Started with QlikView Part XIII Addressing the Data model issue (Synthetic Key issue)

Description:

BISP is committed to provide BEST learning material to the beginners and advance learners. In the same series, we have prepared a complete end-to end Hands-on Beginner's Guide for Qlikview Report, Synthetic key issue, Subfield Function. The document focuses on script editor. **Join our professional training program and learn from experts.**

History:

Version	Description Change	Author	Publish Date
0.1	Initial Draft	Surbhi sahu	21 st Aug 2012
0.1	Review#1	Rajkumar N	29 th Aug 2012

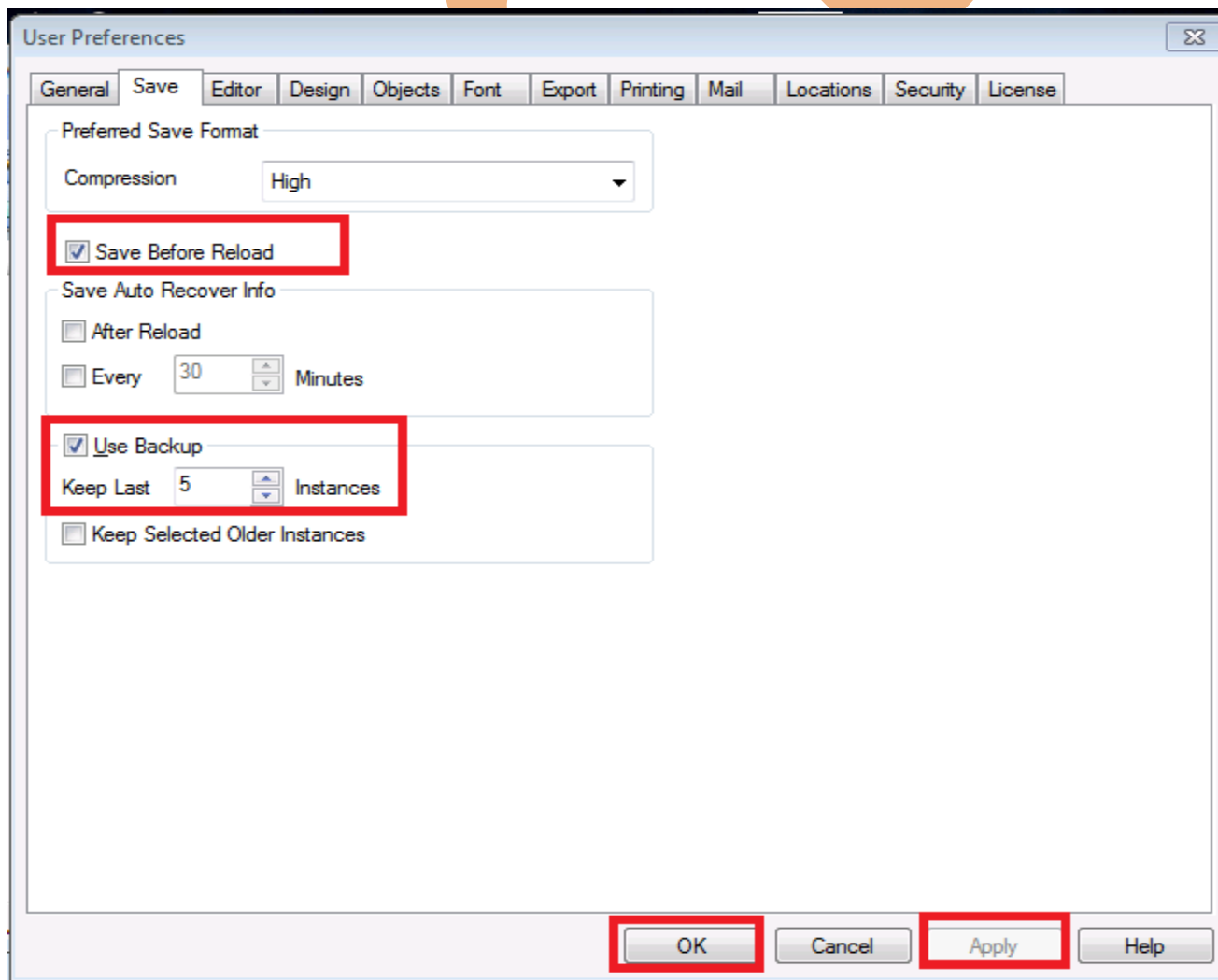
Table of Contents

Table of Contents.....	2
1)Aspects related to reloading the application:.....	3
2)The Script Editor.....	4
2.1)Menu and toolbar.....	4
2.2) Script pane.....	5
2.3) Tool pane.....	5
2.4) Script statement.....	5
3) Loading the aircraft information.....	6
4)Synthetic key issue.....	13
The problem generated due to synthetic key.....	14
.....	14
6)Concatenate statement.....	15
7) Join statement.....	17
8)Subfield Function.....	18
Subfield function are used for splitting the field via subfield function we can split two fields within a single field.....	18

1) **Aspects related to reloading the application:**

As we familiar about the reload function is to load the data from different data sources into the QlikView data model. By default, when QlikView encounters errors during the reload of a document, it automatically closes the document and reloads the last saved version of the file / data. It can be a frustrating experience when you have just written a lot of script, only to see all of it lost because you forgot a semicolon somewhere and didn't save the script which you have written all this time. And you forgot to save the script before reloading the data.

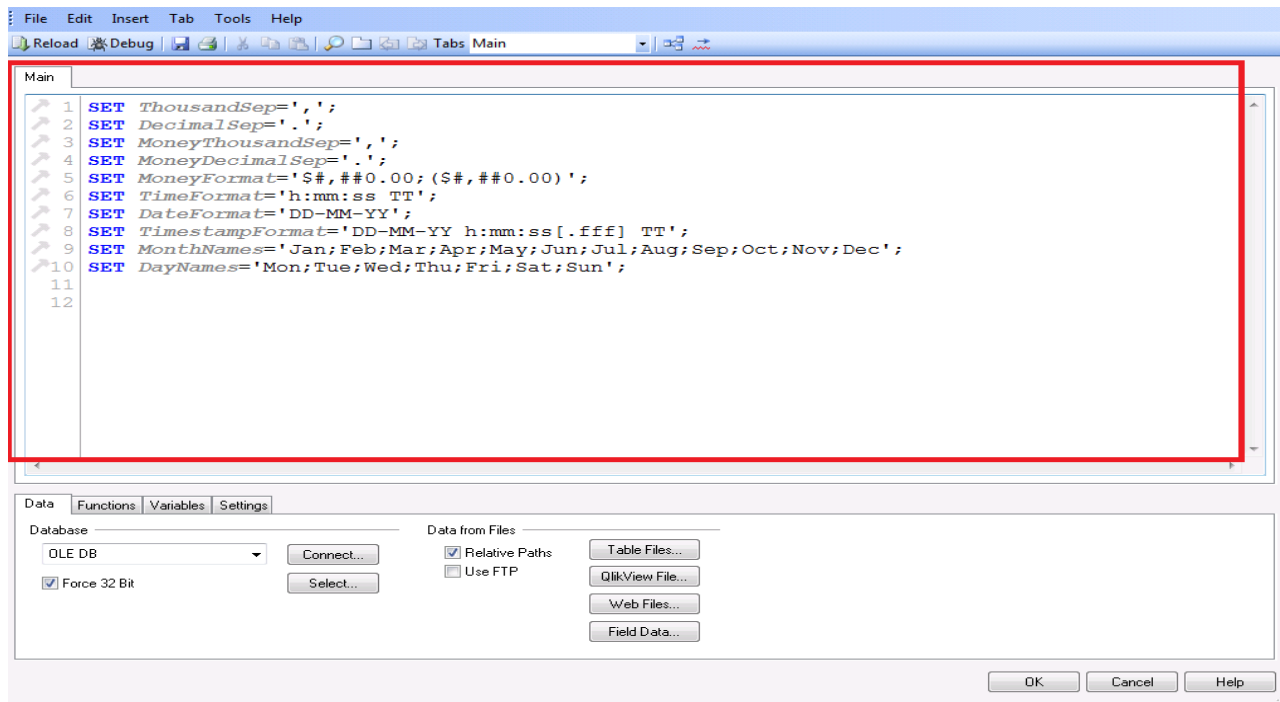
One way to avoid this problem is by always first saving your script before reloading. This can be done by 1) File > Save Entire Document from the menu, 2) Ctrl + S, 3) Clicking on the Save icon in the toolbar, 4) Another more safe way is to set in QlikView to automatically save the file before each reload. To do this, close the script editor and open the User Preferences menu by selecting Settings > User Preferences from the menu, or by pressing Ctrl + Alt + U. In the menu, select the Save tab and tick the checkbox labeled Save Before Reload. It is also advisable to tick the checkbox Use Backup and set the field Keep Last Instances to 5. This last option ensures that the last 5 versions of the QlikView file are kept as backup.



Then->Apply->ok

2) The Script Editor

Script editor is the portion where the script are written and tables are loaded in application. This is mainly termed as the “Heart of QlikView”. All the ETL work is done here.



The script editor consists of the following areas:

- A menu bar
- A toolbar
- A script pane
- A tool pane

2.1) Menu and toolbar

The menu offers a wide range of options

File | Reload= Runs the entire script to reload the data File |

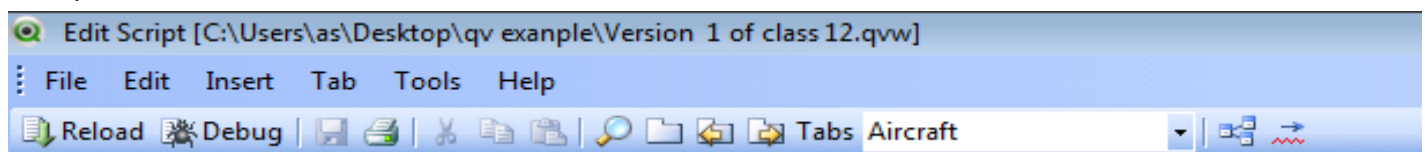
Save Entire Document= Saves the entire document, not just the

Tab | Add Tab = script QlikView scripts can be organized using tabs; this function adds a new tab

Tab | Rename= Renames the currently selected tab

Tab | Promote= Moves the currently selected tab to the left

Tab | Demote= Moves the currently selected tab to the right

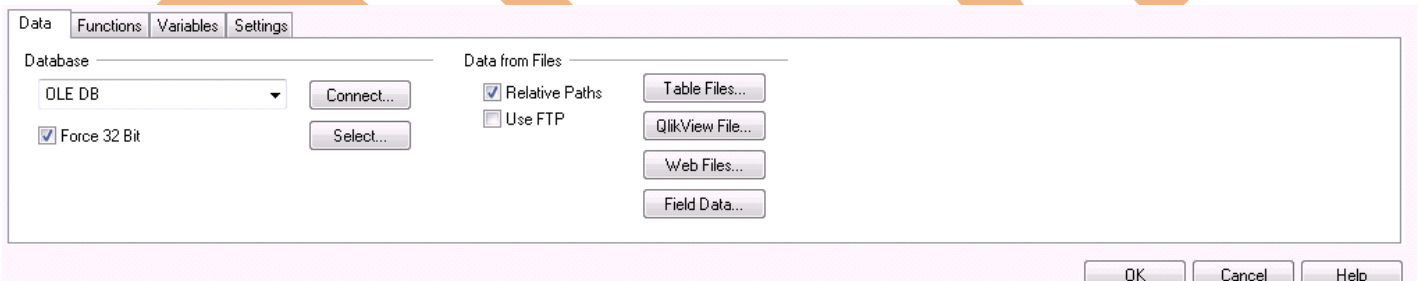


2.2) Script pane

This is the main working area of the script editor. When a new QlikView data model is created, the new script is populated with a Main tab and a few lines where number interpretation variables are added. These lines tell QlikView how to interpret various numbers and are generated automatically based on your operating system settings.

```
Main
1 SET ThousandSep=',';
2 SET DecimalSep='.';
3 SET MoneyThousandSep=',';
4 SET MoneyDecimalSep='.';
5 SET MoneyFormat='$#,##0.00;($#,##0.00)';
6 SET TimeFormat='h:mm:ss TT';
7 SET DateFormat='DD-MM-YY';
8 SET TimestampFormat='DD-MM-YY h:mm:ss[.fff] TT';
9 SET MonthNames='Jan;Feb;Mar;Apr;May;Jun;Jul;Aug;Sep;Oct;Nov;Dec';
10 SET DayNames='Mon;Tue;Wed;Thu;Fri;Sat;Sun';
11
12
```

2.3) Tool pane



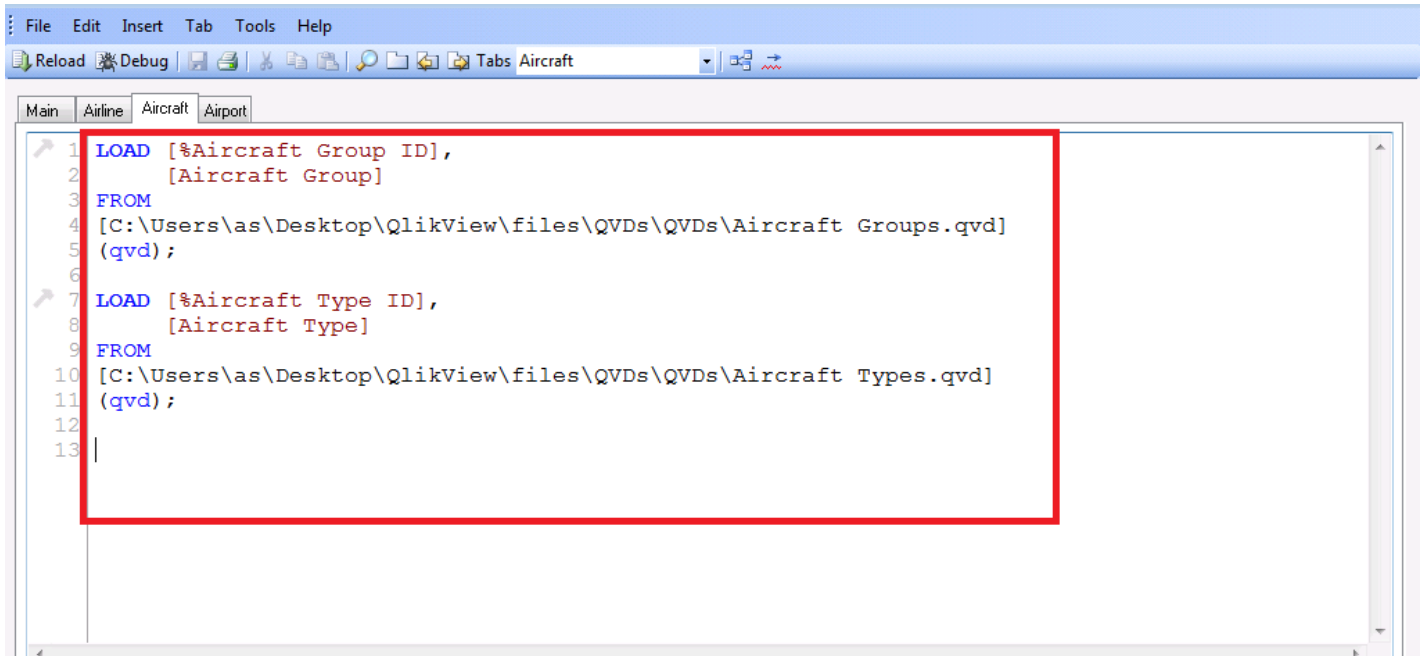
- The Data tab gives QV the flexibility to get / connect to different data sources to get data.
- The Functions tab gives a categorized overview of the script functions available in QlikView.
- The Variables tab shows user and system variables. This tab is populated after each reload, so in a new document it will be blank.
- The Settings tab contains some additional settings with regards to system access and password scrambling.

2.4) Script statement

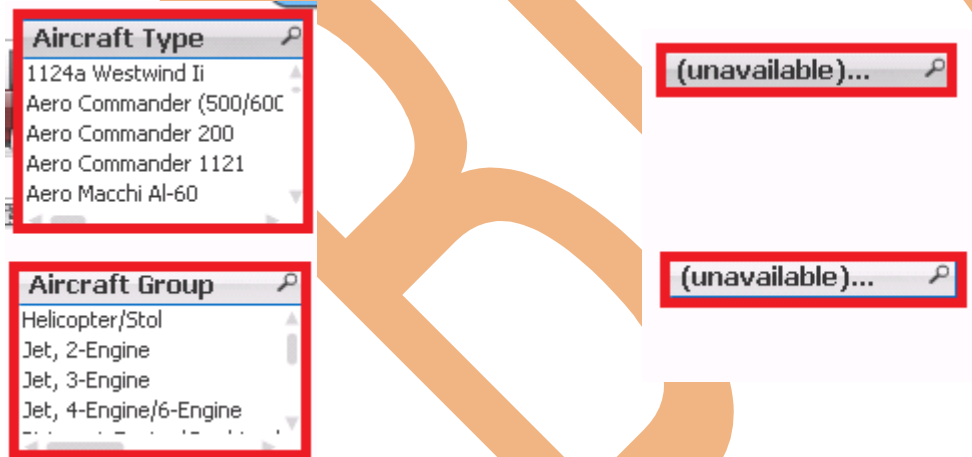
A QlikView script is made up of a sequence of LOAD / SQL statements. These statements are typically used to either manipulate the data, or to conditionally control the way in which the

script is executed. For example, we may want to combine two tables together, or skip over a part of a script if a condition is not met.

We look some other actions which we can perform through the QlikView



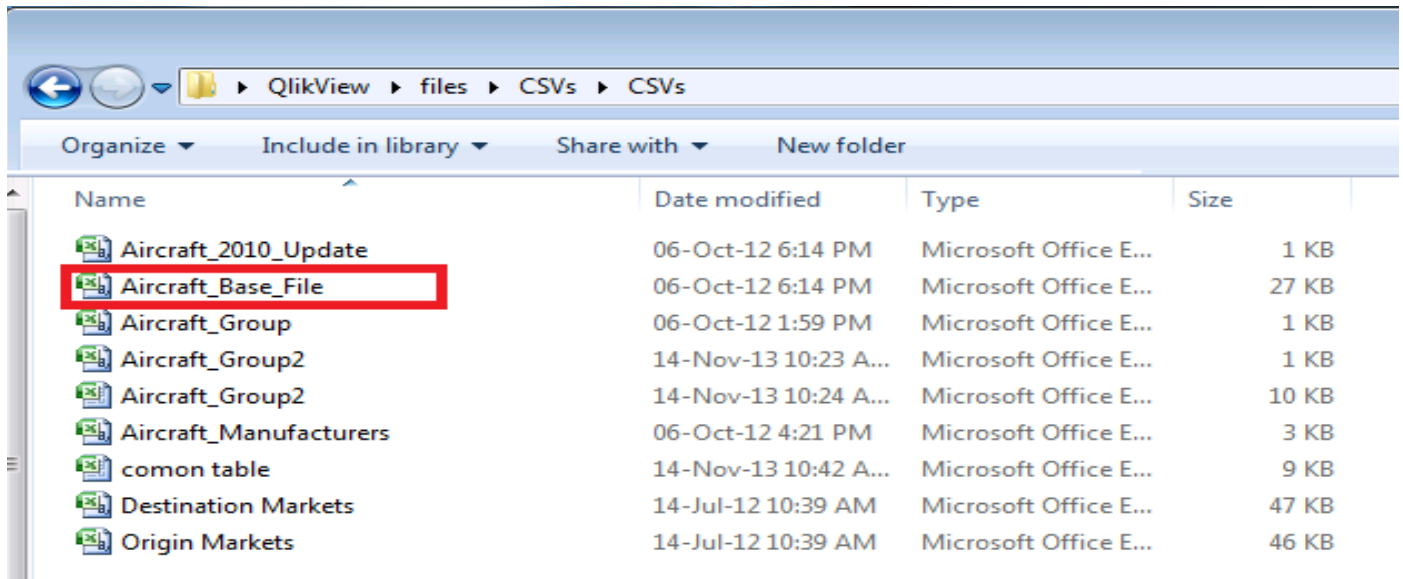
Here you can see in the script editor window the script is here for the table Aircraft Groups and Aircraft Types If we erase the script from here this will disappear the data from both tables in the sheet. Look here the both fields are shown as (unavailable).



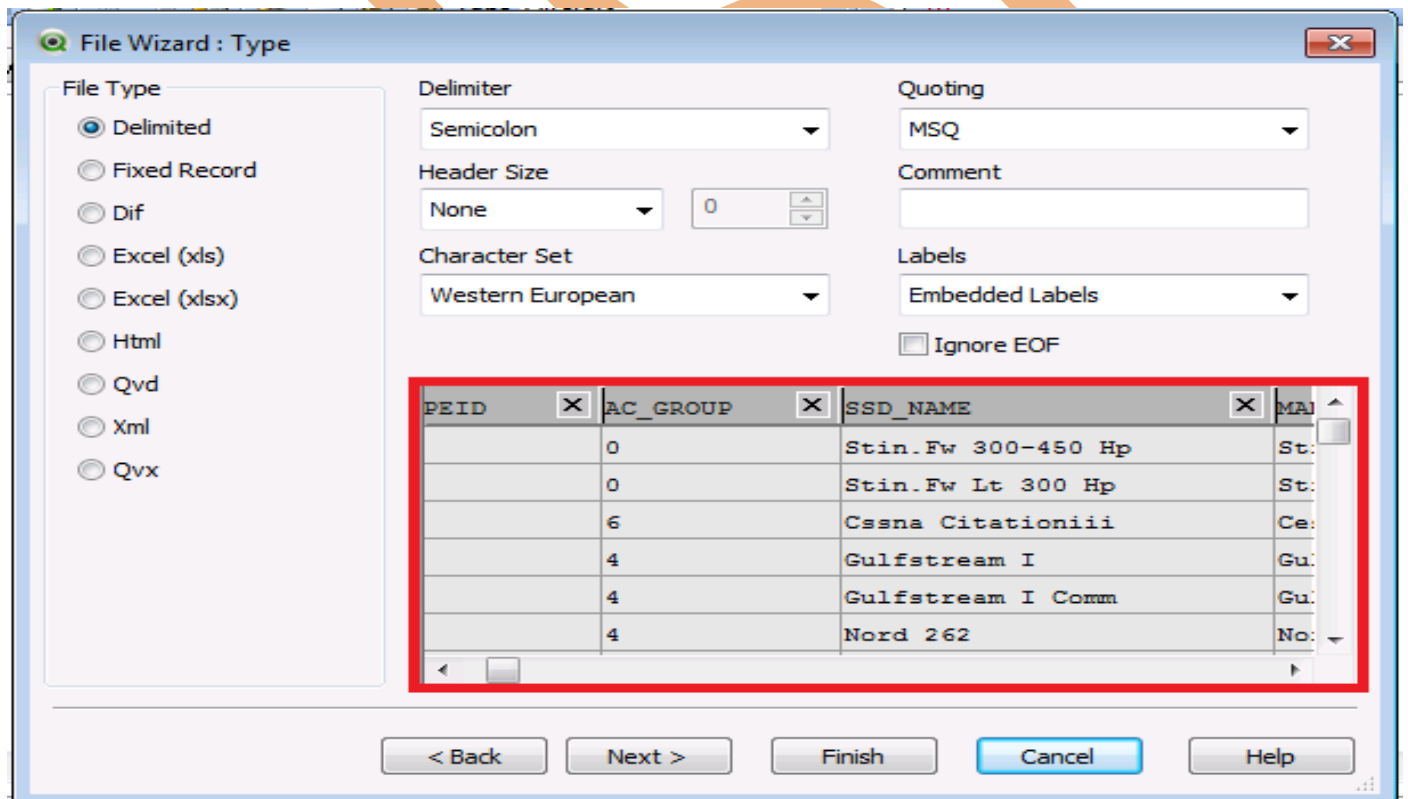
3) Loading the aircraft information

Loading of aircraft information using these steps:

Step1: Choose the file for upload by clicking on -> Table files in the Tool Pane > Browse the location of the particular file which we want to load, select "Aircraft_Base_File.csv" -> OK basically this file consists of all the updates prior to the year 2010. We have loaded it in to the application.



Step 2: Now the table will be loaded in the QlikView



The QV File Wizard has identified the file type as "Delimited", which we are loading > Finish.

Step 4:

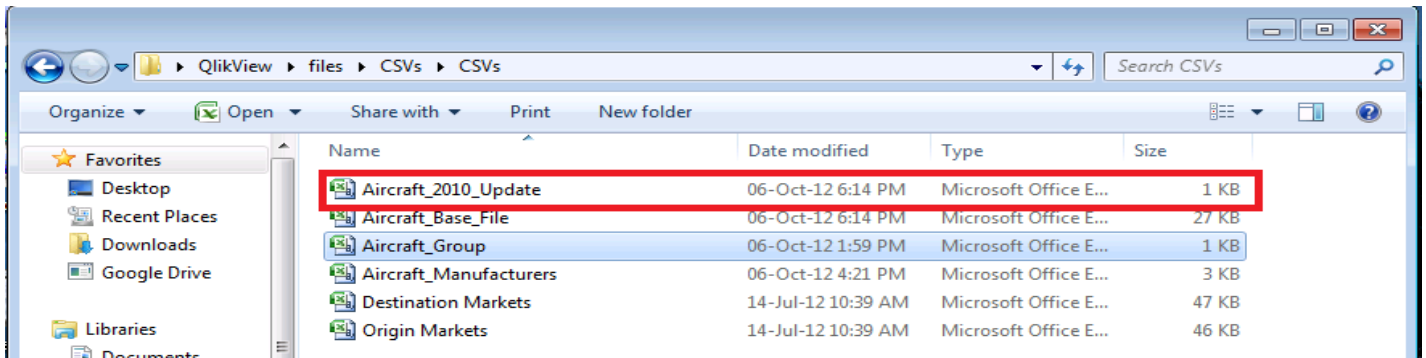
```
Main | Airline | Aircraft | Airport
1 Directory;
2 LOAD AC_TYPEID,
3     AC_GROUP,
4     SSD_NAME,
5     MANUFACTURER,
6     LONG_NAME,
7     SHORT_NAME,
8     BEGIN_DATE,
9     END_DATE
10 FROM
11 [..\QlikView\files\CSVs\CSVs\Aircraft_Base_File.csv]
12 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
13
```

Here the script is loaded in the table but the name of the fields name are not properly maintained (so as to maintain joins between the tables) so we rename the fields by using 'as' keyword.

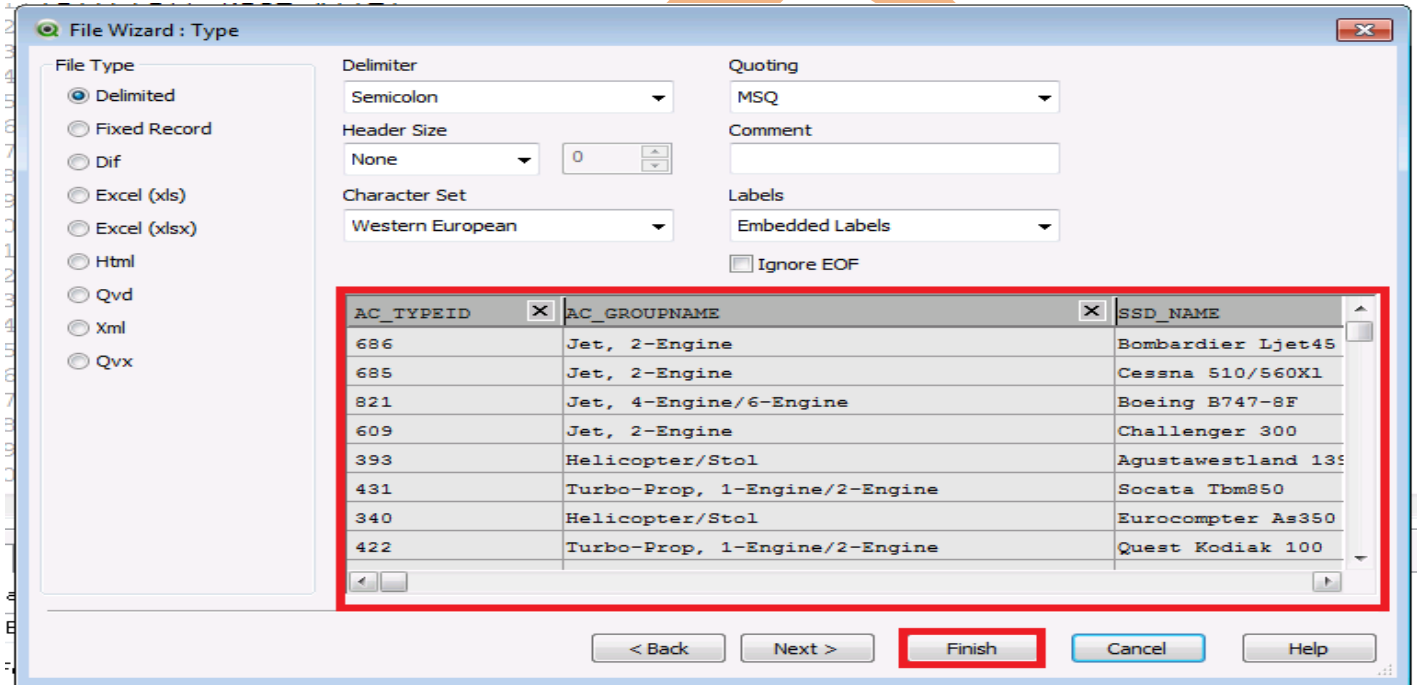
[Aircraft Types]:

```
LOAD AC_TYPEID as [%Aircraft Type ID],
     AC_GROUP as [%Aircraft Group Type],
     SSD_NAME as [Aircraft Name],
     MANUFACTURER as [Aircraft Manufacturer],
     LONG_NAME as [Aircraft Name Full],
     SHORT_NAME as [Aircraft Name Abbreviated],
     BEGIN_DATE as [Aircraft Begin Date],
     END_DATE as [Aircraft End Date]
FROM
[..\Data Files\CSVs\Aircraft_Base_File.csv]
(txt, codepage is 1252, embedded labels, delimiter is ';', msq);
```

Step 5: Now we are moving through our next table which is "Aircraft_2010_Update.csv" which gives the information about the updates of 2010 year and later.



Step 6: Now the table will be loaded > Finish



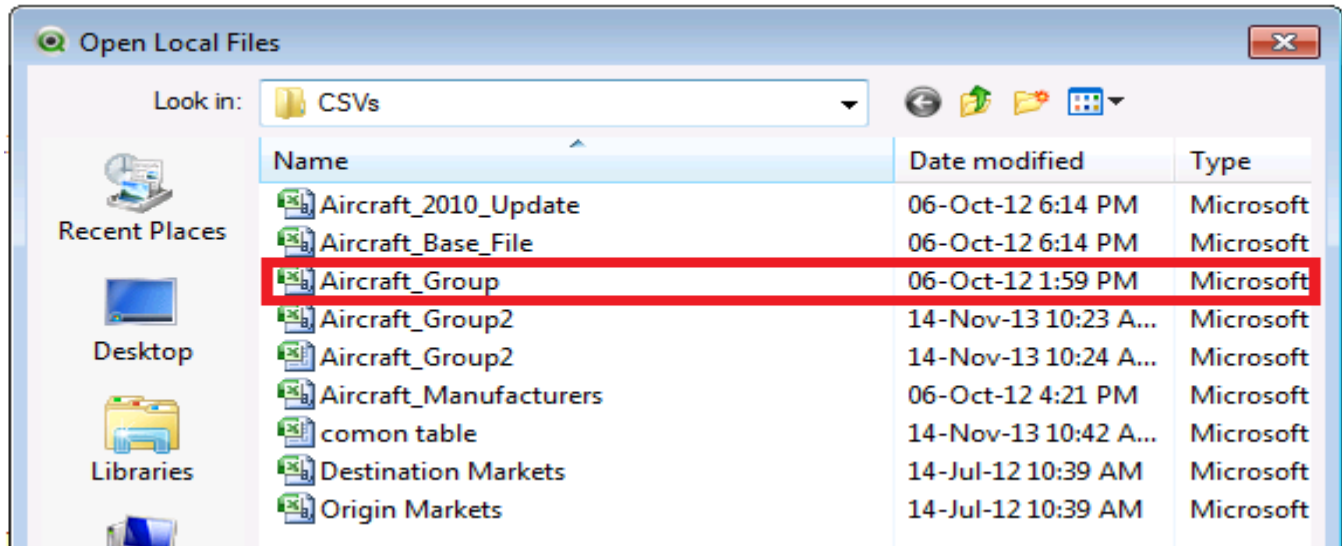
Step 7: Then Rename the columns of the table "Aircraft_2010_Update" using 'as' keyword as shown here.

```
[Aircraft Types 2010]:
LOAD AC_TYPEID as [%Aircraft Type ID],
      AC_GROUPNAME,
      SSD_NAME as [Aircraft Name],
      MANUFACTURER as [Aircraft Manufacturer],
      LONG_NAME as [Aircraft Name Full],
      SHORT_NAME as [Aircraft Name Abbreviated],
      BEGIN_DATE as [Aircraft Begin Date],
      END_DATE as [Aircraft End Date]
FROM
```

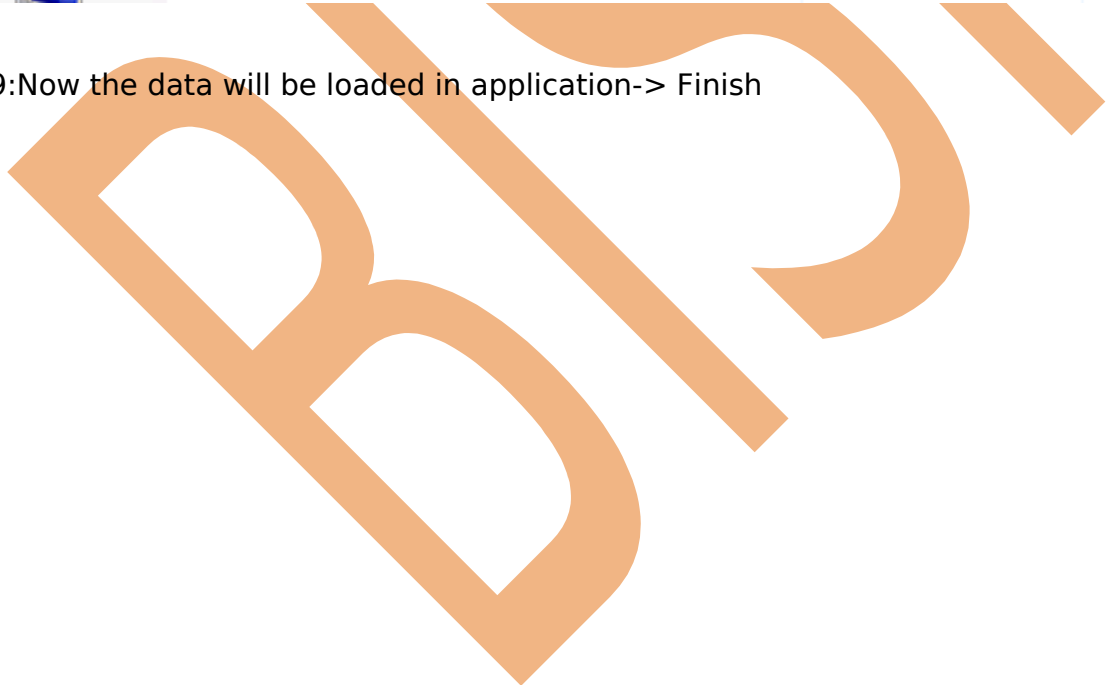
[..\Data Files\CSVs\Aircraft_2010_Update.csv]

(txt, codepage is 1252, embedded labels, delimiter is ';', msq);

Step 8: Then the Table Aircraft Group will be loaded in the QlikView application. This Field contains the information about Aircraft Group ID ,Aircraft Engine type and Aircraft number of Engine. For this browse the location and open the particular file.



Step 9: Now the data will be loaded in application-> Finish



File Wizard : Type

File Type

- Delimited
- Fixed Record
- Dif
- Excel (xls)
- Excel (xlsx)
- Html
- Qvd
- Xml
- Qvx

Delimiter: Semicolon

Quoting: MSQ

Header Size: None, 0

Comment:

Character Set: Western European

Labels: Embedded Labels

Ignore EOF

Aircraft Group ID	Aircraft Engine Type	Aircraft Number Of Engines
0	Piston	1-Engine/Combined Single Engine
1	Piston	2-Engine
2	Piston	3-Engine/4-Engine
3	Helicopter/Stol	1-Engine/Combined Single Engine
4	Turbo-Prop	1-Engine/2-Engine
5	Turbo-Prop	4-Engine
6	Jet	2-Engine
7	Jet	3-Engine

< Back Next > Finish Cancel Help

Aircraft_GroupType

Aircraft_2010_Update

script differ

Delimiter: Semicolon

Quoting: MSQ

Header Size: None, 0

Comment:

Character Set: Western European

Labels: Embedded Labels

Ignore EOF

Aircraft Group ID	Aircraft Engine Type	Aircraft Number Of Engines
0	Piston	1-Engine/Combined Single Engine
1	Piston	2-Engine
2	Piston	3-Engine/4-Engine
3	Helicopter/Stol	1-Engine/Combined Single Engine
4	Turbo-Prop	1-Engine/2-Engine
5	Turbo-Prop	4-Engine
6	Jet	2-Engine
7	Jet	3-Engine

< Back Next > Finish Cancel Help

script differ

Delimiter: Semicolon

Quoting: MSQ

Header Size: None, 0

Comment:

Character Set: Western European

Labels: Embedded Labels

Ignore EOF

%Aircraft Type Id	%Aircraft Group Type	Aircraft Name	Aircr
97	0	Stin.Fw 300-450 Hp	Stins
98	0	Stin.Fw Lt 300 Hp	Stins
646	6	Cesna Citationiii	Cesna
470	4	Gulfstream I	Gulfst
473	4	Gulfstream I Comm	Gulfst
475	4	Nord 262	Nord 2
477	4	Mohawk 298	Nord 2
478	4	Piper T-1040	Piper

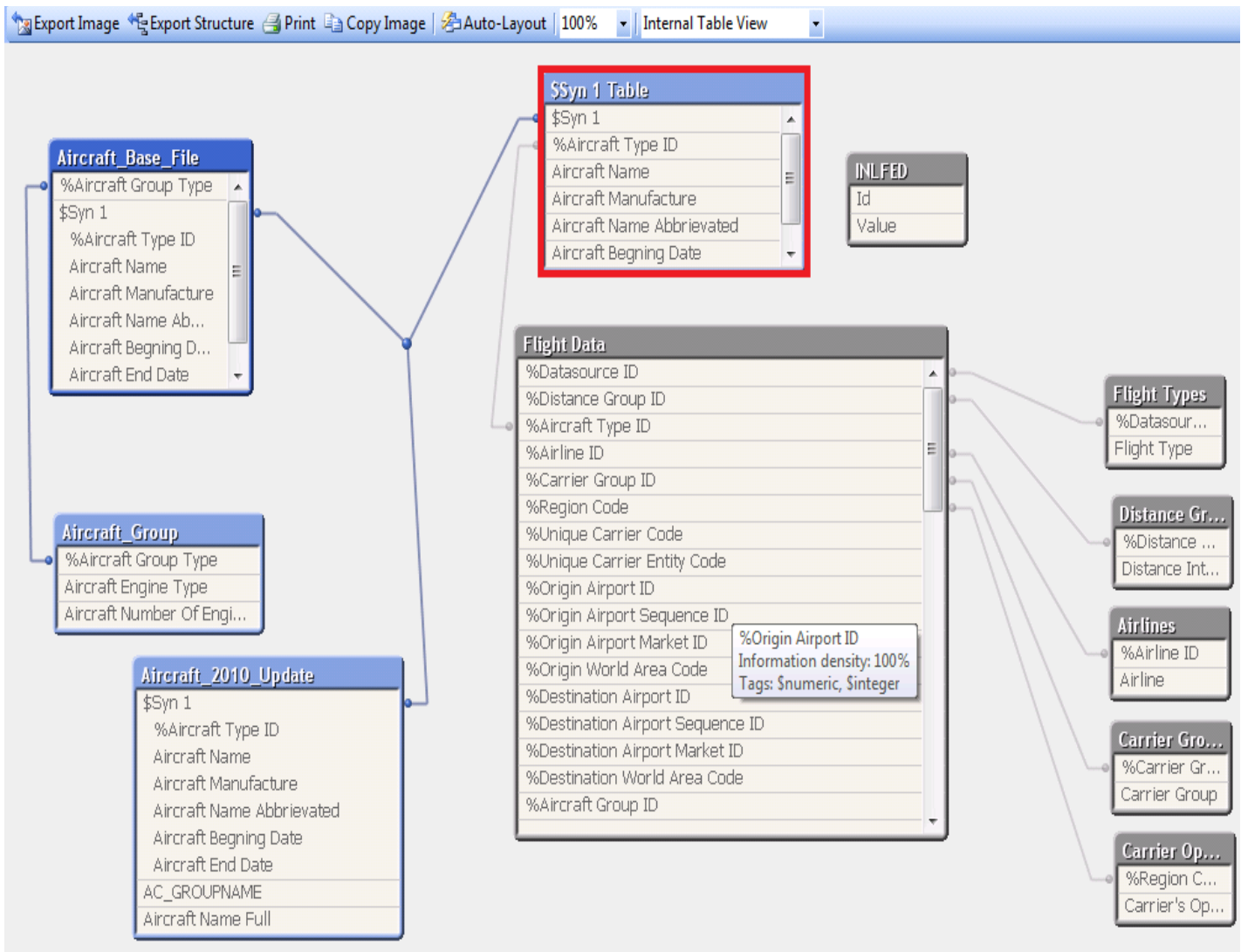
< Back Next > Finish Cancel Help

inigs.

So, it's very illogical to avoid this as it will have large no of records. We should reduce our data as much as possible. So that we can easily manage our data and can reduce the extra statement using 'as' keyword which can be applied in the both table. And joining these 2 tables with this column / field name.

```
Main | Airline | Aircraft | Airport
6 Directory;
7 LOAD [Aircraft Group ID] as [%Aircraft Group Type],
8     [Aircraft Engine Type],
9     [Aircraft Number Of Engines]
10 FROM
11 [..\QlikView\files\CSVs\CSVs\Aircraft_Group.csv]
12 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
13
14
15
16
17
```

we used as Aircraft Group Id in %Aircraft Group Type these both considered as a single column. Now loaded the data. Have a look at the table viewer.



Look here the field \$syn1 Table which we have not import but exist in the table viewer window. This is the Synthetic key which is auto-matically generated.

4) Synthetic key issue

Synthetic table and Synthetic keys gets created when there are more than one common fields between the 2 or more tables. QlikView join tables based on the common fields names in the table and if there is more than one common field then QlikView creates Synthetic tables. Synthetic keys are not beneficial in data model because they are resource intensive and may slow down calculations. They also make data model hard to understand. Synthetic keys should always be resolved.

Look the three tables which we have loaded in the table having many fields common as we see Aircraft_Base_File having almost fields similar to Aircarft_2010_Update since in table field and these two contain one field similar to the table Aircraft Group is Aircraft Group ID.

A	B	C
Aircraft Group	Aircraft_Base_File	Aircraft_2010_Update
Aircraft Group ID	AC_TYPEID	AC_TYPEID
Aircraft Engine Type	AC_GROUP	AC_GROUPNAME
Aircraft Number Of Engines	SSD_NAME	SSD_NAME
	MANUFACTURER	MANUFACTURER
	LONG_NAME	LONG_NAME
	SHORT_NAME	SHORT_NAME
	BEGIN_DATE	BEGIN_DATE
	END_DATE	END_DATE

The problem generated due to synthetic key

For looking the problem create a multibox ->right click on the sheet ->go to new sheet objects ->multibox.

Multi Box Properties []

General | Sort | Presentation | Number | Font | Layout | Caption

Title: Object ID: MB04

Available Fields:

- Aircraft Name Full
- # Air Time
- # Available Seats
- # Departures Performed
- # Departures Scheduled
- # Payload
- # Ramp-To-Ramp Time
- # Transported Freight
- # Transported Mail
- # Transported Passengers
- %Aircraft Configuration ID
- %Aircraft Group ID
- %Aircraft Group Type
- %Aircraft Type ID
- %Airline ID
- %Carrier Group ID
- %Datasource ID
- %Destination Airport ID

Fields Displayed in Multibox:

- Aircraft Name
- Aircraft Number Of Engines
- Aircraft Engine Type
- AC_GROUPNAME
- Aircraft Beginning Date

Buttons: Add >, Add All >>, < Remove, << Remove All, Expression >, Edit...

Buttons: Count Order, Load Order, Text Order

Promote, Demote

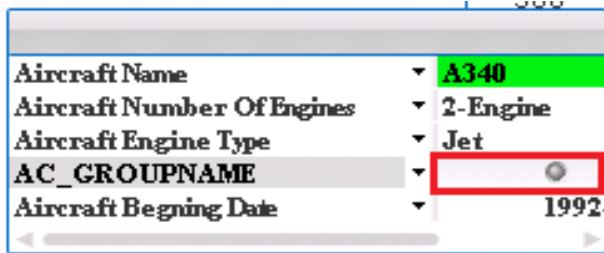
Settings for Selected Field:

Label:

Sort by Applicability

Now five fields are selected those are Aircraft Name,Aircraft number of engine,Aircraft engine type

Aircraft beging date,AC_GROUPNAME.now->Apply->ok.



look here the data on AC_GROUPNAME field disappeared due to generation of synthetic keys here we have two ways to resolve this issue.

- ⇒ Using concatenate statement
- ⇒ Using joins

6)Concatenate statement

The CONCATENATE statement is also a prefix to the LOAD statement, but instead of matching and merging rows between tables, this statement appends the rows of one table to another table.

Example Look here we have table Table1 and Table 2

Table1			Table2		
Key	A	B	Key	C	D
1	A1	B1	1	C1	D1
2	A2	B2	2	C2	D2
3	A3	B3			

Concatenate appends the rows of one table to another. Concatenate never merges any rows. The number of rows in a concatenated table is always the sum of the rows from the two input tables.

CONCATENATE (Table1) LOAD * RESIDENT Table2;

CONCATENATE Result					
Row	Key	A	B	C	D
1	1	A1	B1	-	-
2	2	A2	B2	-	-
3	3	A3	B3	-	-
4	1	-	-	C1	D1
5	2	-	-	C2	D2

Rows are like Key values and are not merged together. The rows from Table2 are simply appended to Table1. Because the tables have different fields, rows will have null values for the fields from the "other" table.

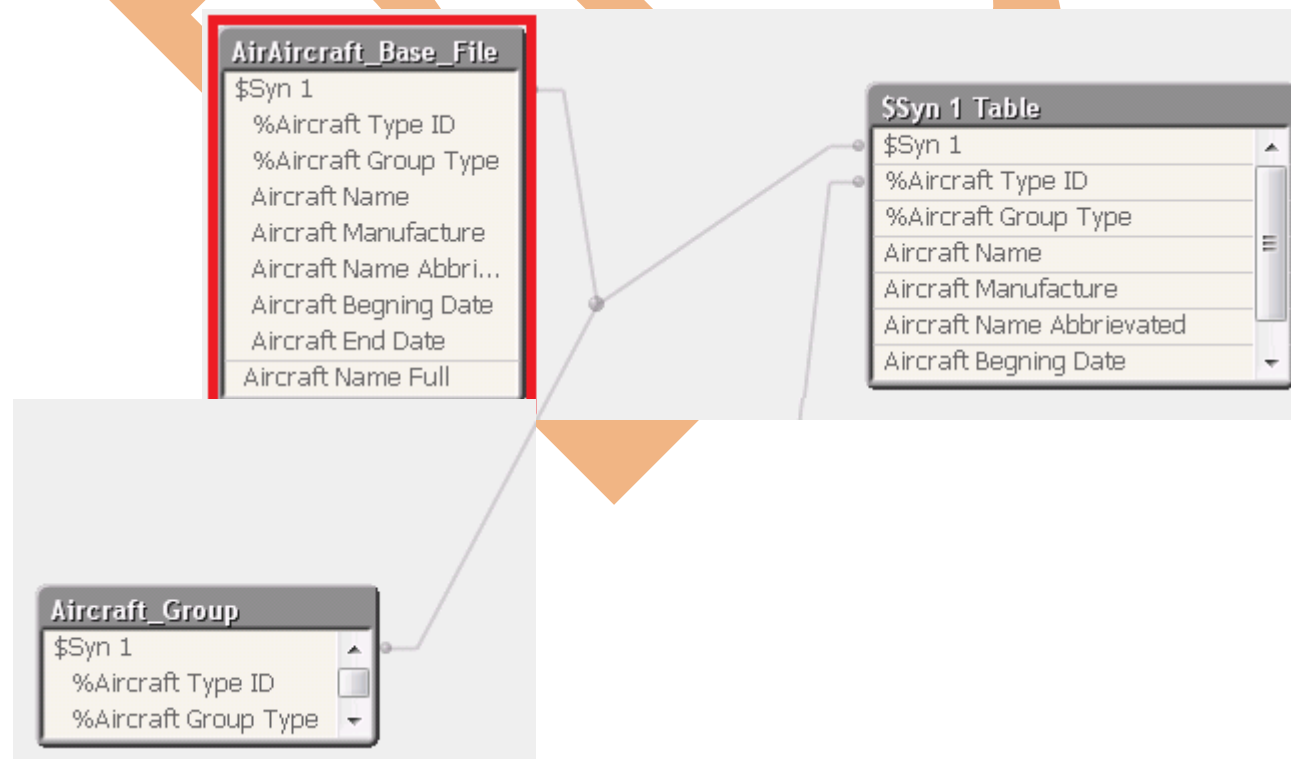
Here we want the whole information in this the table Aircraft_2010_Update contain information of 2010 and Aircraft_Base_File information before 2010 we want to perform

Aircraft_Base_File= Aircraft_Base_File(previous)+ Aircraft_2010_Update

we want that the data of 2010 append to the base file.

```
Main | Airline | Aircraft | Airport
16 Concatenate [AirAircraft_Base_File]:|
17 LOAD AC_TYPEID as [%Aircraft Type ID],
18     AC_GROUPNAME,
19     SSD_NAME as [Aircraft Name],
20     MANUFACTURER as [Aircraft Manufacture],
21     LONG_NAME as [Aircraft Name Full],
22     SHORT_NAME as [Aircraft Name Abbrievated],
23     BEGIN_DATE as [Aircraft Begning Date],
24     END_DATE as [Aircraft End Date]
25 FROM
26 [..\QlikView\files\CSVs\CSVs\Aircraft_2010_Update.csv]
27 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
28
```

->Reload the table->go to the table viewer



From the table viewer window the table Aircraft_2010_Update disappear and concatenates into the table Aircraft_Base_File.

7) Join statement

The JOIN statement is a prefix to the LOAD statement. It is used to join the table that is being loaded to a previously loaded table. The two tables are joined using a natural join, this means that the columns in both tables are compared and the join is made over those columns that have the same column names. This means that if multiple columns are shared between tables, the match will be made over the distinct combinations of those columns.

Example we have 2 table Table1 and Table 2

Table1			Table2		
Key	A	B	Key	C	D
1	A1	B1	1	C1	D1
2	A2	B2	2	C2	D2
3	A3	B3			

Now the join will result

CONCATENATE (Table1) LOAD * RESIDENT Table2;

JOIN Result					
Row	Key	A	B	C	D
1	1	A1	B1	C1	D1
2	2	A2	B2	C2	D2
3	3	A3	B3	-	-

Here we have two tables as Aircraft_group and Aircraft_Base_File in this we have previously seen that the data in the 2 fields as Aircraft Group ID and %Aircraft Group Type is same so we want to merge this row of the table Aircraft_group

```
Main | Airline | Aircraft | Airport
13
14 Left join [Aircraft Base File]:
15 LOAD [Aircraft Group ID] as [%Aircraft Group Type],
16     [Aircraft Engine Type],
17     [Aircraft Number Of Engines]
18 FROM
19 [..\QlikView\files\CSVs\CSVs\Aircraft_Group.csv]
20 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
21
```

This will merge the row in the table Aircraft_base_File.

8) Subfield Function

Subfield function are used for splitting the field via subfield function we can split two fields within a single field.

Look the field AC_GROUPNAME of Aircraft base table

AC_TYPEID	AC_GROUPNAME	SSD_NAME
686	Jet, 2-Engine	Bombardier Ljet45
685	Jet, 2-Engine	Cessna 510/560X1
821	Jet, 4-Engine/6-Engine	Boeing B747-8F
609	Jet, 2-Engine	Challenger 300
393	Helicopter/Stol	Agustawestland 139
431	Turbo-Prop, 1-Engine/2-Engine	Socata Tbm850
340	Helicopter/Stol	Eurocompter As350
422	Turbo-Prop, 1-Engine/2-Engine	Quest Kodiak 100

And these two fields of table Aircraft_Group

Aircraft Engine Type	Aircraft Number Of Engines
Piston	1-Engine/Combined Single Engine
Piston	2-Engine
Piston	3-Engine/4-Engine
Helicopter/Stol	1-Engine/Combined Single Engine
Turbo-Prop	1-Engine/2-Engine
Turbo-Prop	4-Engine
Jet	2-Engine
Jet	3-Engine

Here the field AC_GROUPNAME from table Aircraft_Base_File covers the information of Aircraft Engine Type and Aircraft Number OF Engines of table Aircraft Group individually so we can split these two columns within a single table.

AC_GROUPNAME = Aircraft Engine Type+Aircraft Number OF Engines

This function is very case sensitive, in field AC_GROUPNAME information separated by comma(,) in syntax as we are loading the .csv file which is a comma separated file.

SUBFIELD(AC_GROUPNAME,',1) as Aircraft Engine Type

SUBFIELD(AC_GROUPNAME,',2) as Aircraft Number OF Engines

are numbered as 1 and 2 to identify the fields which are going to get split.

```

Main | Airline | Aircraft | Airport
12 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
13
14 Left join [Aircraft_Base_File]:
15 LOAD [Aircraft Group ID] as [%Aircraft Group Type],
16     [Aircraft Engine Type],
17     [Aircraft Number Of Engines]
18 FROM
19 [..\QlikView\files\CSVs\CSVs\Aircraft_Group.csv]
20 (txt, codepage is 1252, embedded labels, delimiter is ';', msq);
21
22 Concatenate [Aircraft_Base_File]:
23 LOAD AC TYPEID as [%Aircraft Type ID],
24     SUBFIELD(AC_GROUPNAME,',',1) as [Aircraft Engine Type],
25     SUBFIELD(AC_GROUPNAME,',',2) as [Aircraft Number Of Engines],
26     SSD_NAME as [Aircraft Name],
27     MANUFACTURER as [Aircraft Manufacture],
28     LONG_NAME as [Aircraft Name Full],
29     SHORT_NAME as [Aircraft Name Abbrievated],
30     BEGIN_DATE as [Aircraft Begning Date],
31     END_DATE as [Aircraft End Date]

```

Now reload the table and go to the table viewer window. Look here the synthetic key resolved and data managed in formatted way.

