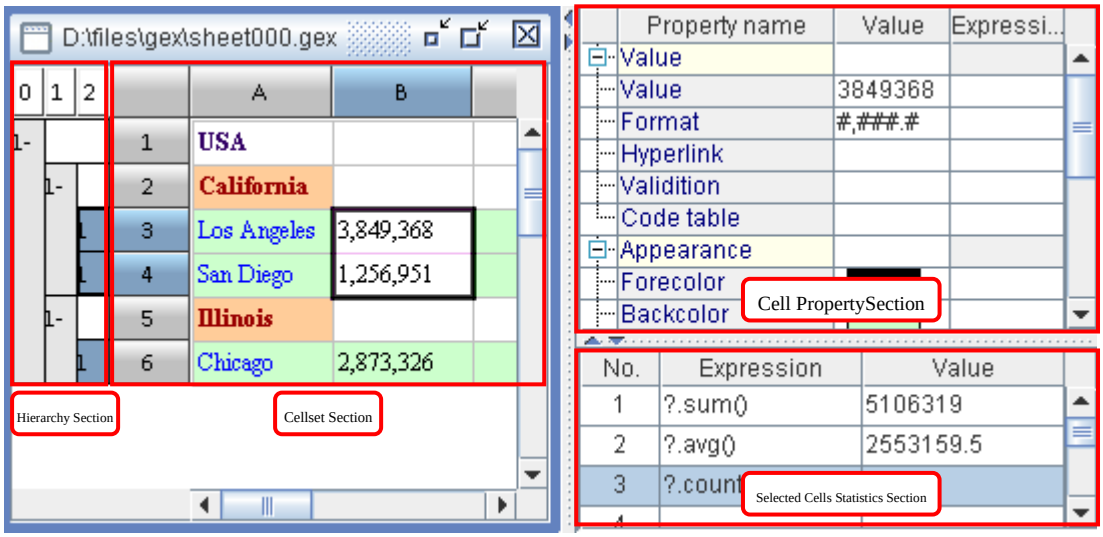
Getting started

1. A Spreadsheet with Hierarchy

esCalc is a powerful spreadsheet preparation tool highly capable of calculation. esCalc has a similar structure to Excel that stores all raw data and calculation expression in the cellset for the user to design and manage cellset or calculate the data conveniently.

However, the spreadsheet of esCalc is unique because it is not a simple 2-D table. Each row in esCalc cellset has an assigned level and a sequence number. The data in the cellset can be interrelated throughout the hierarchy. Such spreadsheet is **Calculation Cellset** by definition.



As shown in the above figure, in the left window, the right section is the **Cellset Section** in which the current calculation cellset file is opened. The left section is the **Hierarchy Section** to display the hierarchy structure of calculation cellset. In the right window, the upper part is the **Cell Property Section** for you to view and set various properties of the cell, and the lower part is the **Selected Cells Statistics Section** for making automatic statistics on selected cells, including sum up, average, and counting. You can add the customized statistics expression in the **Selected Cells Statistics Section**.

In the calculation cellset of esCalc, instead of a simple 2-D table, the data is stored in groups according to levels.

For each row in the calculation cellset, there is a related number in Hierarchy Section on the left. This is the **Sequence Number** of the current row. The column in which the sequence number is located is the **Level** of this row. For example, for the 3rd row, the sequence number is 1 and the level is 2; for the second row, the sequence number is 1 and the level is 1; for the first row, the sequence number is 1 and the level is 0.

Please note that there is a minus sign "-" behind some sequence numbers in the Hierarchy Section. This indicates that this row is the header row of a group (Grouping Row by definition), and all rows of the group are expanded to display (for example, the second row in the above cellset). You can collapse all rows in the group by clicking the sequence number of the grouping row. After clicking the sequence number of the second row, the cellset will be as given below:

0	1	2		A	B	C	D	E	F
1-			1	USA					
	1+		2	California					
	1-		5	Illinois					
		1	6	Chicago	2,873,326				

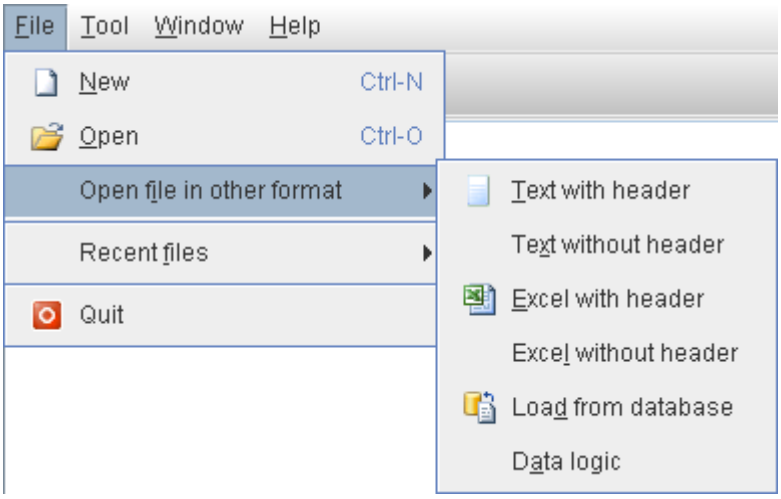
As you can see, the cities below California are collapsed and hide away. At the same time, the symbol behind the sequence number of the second row changes to the plus sign. By clicking the sequence number of the second row again, you can revert to expand the rows in the group, and they will appear again.

The multi-level table can better present the structure of data, and facilitate various complex computations, for example, to calculate the period-on-period comparison and comparison with the same period of any previous year.

2. Reference to Existing Data

In esCalc, you can reference to text file or Excel sheet from which the data will be converted to the esCalc.

Open esCalc, then click **File** and **Open file in other format** from the menu bar. The below menu will prompt.



As shown above, with these options, you can select whether the header row is required in the data when using the data from Txt file or Excel file.

Take the below text file states.txt for example:

```
StateId  Name       Population Abbr  Area    Capital
5        California 37253956   CA    163700  Sacramento
9        Florida    18801310   FL    65755   Tallahassee
13       Illinois   12830632   IL    54826   Springfield
32       New York   19378102   NY    54556   Albany
43       Texas      25145561   TX    268820  Austin
```

In the text file, each line will be separated with the carriage return, and each column is separated with Tab. If chosing to open the text without header, then the resulting calculation cellset is as shown below:

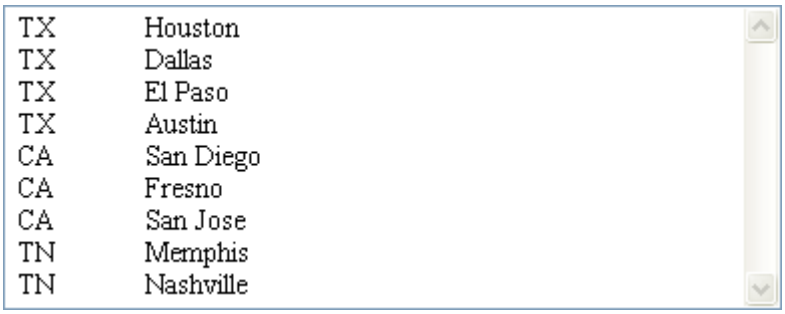
0	1		A	B	C	D	E	F
1-		1						
	1	2	StateId	Name	Population	Abbr	Area	Capital
	1	3	5	California	37253956	CA	163700	Sacramento
	1	4	9	Florida	18801310	FL	65755	Tallahassee
	1	5	13	Illinois	12830632	IL	54826	Springfield
	1	6	32	New York	19378102	NY	54556	Albany
	1	7	43	Texas	25145561	TX	268820	Austin

Because it is the text without header is selected, all rows of text file will be taken as the rows of data when generating the calculation cellset. Import to the cellset one by one, and a blank row for header will be there leading all rows of data. If chosing to open the text with header, then the first row of text file will be imported as the title of each column, as shown below:

0	1		A	B	C	D	E	F
1-		1	StateId	Name	Population	Abbr	Area	Capital
	1	2	5	California	37253956	CA	163700	Sacramento
	1	3	9	Florida	18801310	FL	65755	Tallahassee
	1	4	13	Illinois	12830632	IL	54826	Springfield
	1	5	32	New York	19378102	NY	54556	Albany
	1	6	43	Texas	25145561	TX	268820	Austin

In addition, esCalc allows you to get data directly from the database. Simply click **Load from database** to select the table and field from the specified database, or use the SQL statement directly to import the record from result set to the calculation cellset. The name of each column is the same to that of the result set.

You can also copy the external data and paste it into the calculation cellset. Take the data from the text file record.txt for example:



Copy all text on this text file and create a new gex-file in esCalc as follows:

0	1		A	B	C
1-		1			
	1	2			

Select B2 cell, and click **Homocell Paste** on the right-click menu, and the result is as given below:

0	2		A	B	C
1-		1			
	1	2		TX	Houston
	1	3		TX	Dallas
	1	4		TX	El Paso
	1	5		TX	Austin
	1	6		CA	San Diego
	1	7		CA	Fresno
	1	8		CA	San Jose
	1	9		TN	Memphis
	1	10		TN	Nashville

When pasting the homocells, the row will be appended automatically to the group if the existing rows are not enough to hold the data for pasting in the cellset. To copy the data from outside, you can copy from the files of other formats or the database browser.

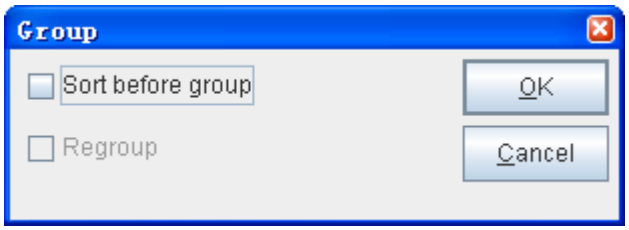
3. Group and Align

Sometimes, you will need to follow a certain rule to group the data, and carry out the data processing for each group when making the statistics. In the calculation cellset, you can perform the grouping action to group the rows.

Choose the **Text with header** to open the GymScores.txt, or open the esCalc_q03_1.gex directly to get the calculation cellset as given below:

0	1		A	B	C
1-		1	Name	Event	Score
	1	2	Grace Miller	Vault	14.0
	1	3	Kayla Rodriguez	Vault	14.575
	1	4	Kayla Rodriguez	Uneven bars	14.675
	1	5	Kayla Rodriguez	Balance beam	13.975
	1	6	Kayla Rodriguez	Floor	14.4
	1	7	Lauren Davis	Vault	14.275
	1	8	Grace Miller	Floor	13.975
	1	9	Grace Miller	Uneven bars	12.875
	1	10	Lauren Davis	Floor	13.0
	1	11	Grace Miller	Balance beam	13.1
	1	12	Daniel Smith	Parallel bars	14.375
	1	13	Lauren Davis	Balance beam	15.225
	1	14	Lauren Davis	Uneven bars	14.275

The above is a player score sheet for which the row in the cellset will be grouped by the name of athlete. To do so, select the A2 cell, and click **Operation** -> **Group** on the right-click menu. After the grouping action, the group action panel will appear, as shown below:



When grouping, you can choose not to sort first by not checking the **Sort before group** option. The resulting esCalc after grouping is as follows:

0	1	2		A	B	C
1-			1	**Name**	**Event**	**Score**
	1-		2	**Grace Miller**		
		1	3	**Grace Miller**	**Vault**	**14.0**
	1-		4	**Kayla Rodriguez**		
		1	5	**Kayla Rodriguez**	**Vault**	**14.575**
		1	6	**Kayla Rodriguez**	**Uneven bars**	**14.675**
		1	7	**Kayla Rodriguez**	**Balance beam**	**13.975**
		1	8	**Kayla Rodriguez**	**Floor**	**14.4**
	1-		9	**Lauren Davis**		
		1	10	**Lauren Davis**	**Vault**	**14.275**
	1-		11	**Grace Miller**		
		1	12	**Grace Miller**	**Floor**	**13.975**
		1	13	**Grace Miller**	**Uneven bars**	**12.875**
	1-		14	**Lauren Davis**		
		1	15	**Lauren Davis**	**Floor**	**13.0**
	1-		16	**Grace Miller**		
		1	17	**Grace Miller**	**Balance beam**	**13.1**
	1-		18	**Daniel Smith**		
		1	19	**Daniel Smith**	**Parallel bars**	**14.375**
	1-		20	**Lauren Davis**		
		1	21	**Lauren Davis**	**Balance beam**	**15.225**
		1	22	**Lauren Davis**	**Uneven bars**	**14.275**

Save this cellset as esCalc_q03_2.gex for future use. In the result, the data of each row is grouped by the name of athlete, and a new row is added on the topmost of each group with the name of each group in it. In the upper left corner of the cell where the name of the current row is located, a small red triangle appears. This is the mark of master row, indicating the cell holding the name is the **Master Cell** of the grouping row. In addition, you will find in the Hierarchy Section that the level of each row of data decreased by one level. The grouping row will be inserted before the row of data as a new level.

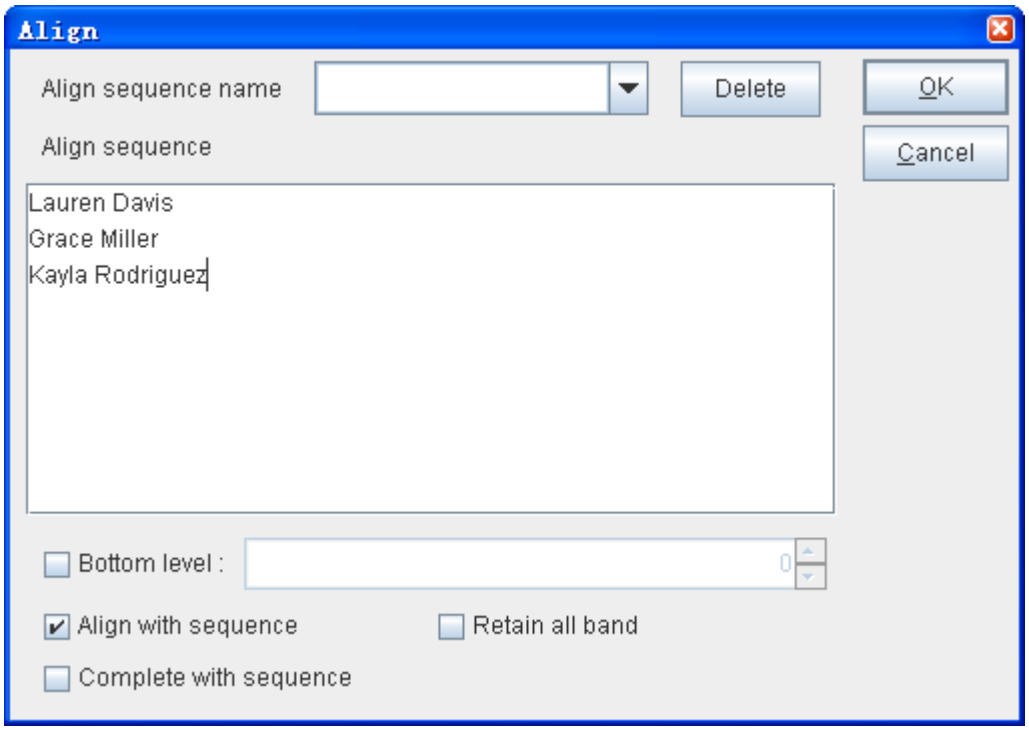
It is obvious that the records are not sorted before grouping. Only the neighboring rows with the same name will be allocated to a same group. In this way, the duplicate grouping row may appear. For example, the Row 2, 11, and 16 are all the grouping rows of Grace Miller.

Reopen the cellset esCalc_q03_1.gex or press Ctrl+Z to undo the last grouping action. Select A2 and group in the original calculation cellset once again. This time, check the **Sort Before Group** option and the calculation cellset after grouping is as given below:

0	1	2		A	B	C
1-			1	Name	Event	Score
	1-		2	Daniel Smith		
		1	3	Daniel Smith	Parallel bars	14.375
	1-		4	Grace Miller		
		1	5	Grace Miller	Vault	14.0
		1	6	Grace Miller	Floor	13.975
		1	7	Grace Miller	Uneven bars	12.875
		1	8	Grace Miller	Balance beam	13.1
	1-		9	Kayla Rodriguez		
		1	10	Kayla Rodriguez	Vault	14.575
		1	11	Kayla Rodriguez	Uneven bars	14.675
		1	12	Kayla Rodriguez	Balance beam	13.975
		1	13	Kayla Rodriguez	Floor	14.4
	1-		14	Lauren Davis		
		1	15	Lauren Davis	Vault	14.275
		1	16	Lauren Davis	Floor	13.0
		1	17	Lauren Davis	Balance beam	15.225
		1	18	Lauren Davis	Uneven bars	14.275

To reach the above result, sort each row ascendingly by the name of athlete, and then perform the grouping action.

The data sorting is often required before grouping. However, the available orders are just either ascending or descending. Such sorting is not always the desired one. For example, suppose that we are planning to group the scores of players following the order of Lauren Davis, Grace Miller, and Kayla Rodriguez, and delete the scores of other players. Then, in this case, the align operation can meet the need. To do so, select the name at grouping level, for example A2, and click **Operation** -> **Align** on the right-click menu to perform the align operation. The setup on the Align panel is as follows:



In the alignment sequence, input the desired name of althelet one by one. Please remember to leave the **Retain all band** option unchecked so as to delete any data no in the align sequence. The result after sorting is as follows:

0	1	2		A	B	C
1-			1	**Name**	**Event**	**Score**
	1-		2	**Lauren Davis**		
		1	3	**Lauren Davis**	**Vault**	**14.275**
		1	4	**Lauren Davis**	**Floor**	**13.0**
		1	5	**Lauren Davis**	**Balance beam**	**15.225**
		1	6	**Lauren Davis**	**Uneven bars**	**14.275**
	1-		7	**Grace Miller**		
		1	8	**Grace Miller**	**Vault**	**14.0**
		1	9	**Grace Miller**	**Floor**	**13.975**
		1	10	**Grace Miller**	**Uneven bars**	**12.875**
		1	11	**Grace Miller**	**Balance beam**	**13.1**
	1-		12	**Kayla Rodriguez**		
		1	13	**Kayla Rodriguez**	**Vault**	**14.575**
		1	14	**Kayla Rodriguez**	**Uneven bars**	**14.675**
		1	15	**Kayla Rodriguez**	**Balance beam**	**13.975**
		1	16	**Kayla Rodriguez**	**Floor**	**14.4**

By aligning the grouping row, every group is sorted again by the name of players, and the groups not in the align sequence are deleted. Although the align operation is usually used to group in the specified order, it can just be performed to sort the row of data.

In the grouping row of this calculation cellset, select a cell, such as A2, and click **Structure**

-> **Dismantle master row** on the right-click menu to dismantle the master row. The result is as given below:

0	1		A	B	C
1-		1	**Name**	**Event**	**Score**
	1	2	**Lauren Davis**	**Vault**	**14.275**
	1	3	**Lauren Davis**	**Floor**	**13.0**
	1	4	**Lauren Davis**	**Balance beam**	**15.225**
	1	5	**Lauren Davis**	**Uneven bars**	**14.275**
	1	6	**Grace Miller**	**Vault**	**14.0**
	1	7	**Grace Miller**	**Floor**	**13.975**
	1	8	**Grace Miller**	**Uneven bars**	**12.875**
	1	9	**Grace Miller**	**Balance beam**	**13.1**
	1	10	**Kayla Rodriguez**	**Vault**	**14.575**
	1	11	**Kayla Rodriguez**	**Uneven bars**	**14.675**
	1	12	**Kayla Rodriguez**	**Balance beam**	**13.975**
	1	13	**Kayla Rodriguez**	**Floor**	**14.4**

As you can see, dismantle the master row is the reverse action of grouping. The group level can be deleted, and the data row from each group can be set in its original order in the group.

4. Sort and Distinct

When populating the cellset with data, the data is not sorted. To facilitate the statistics, sorting the data is necessary.

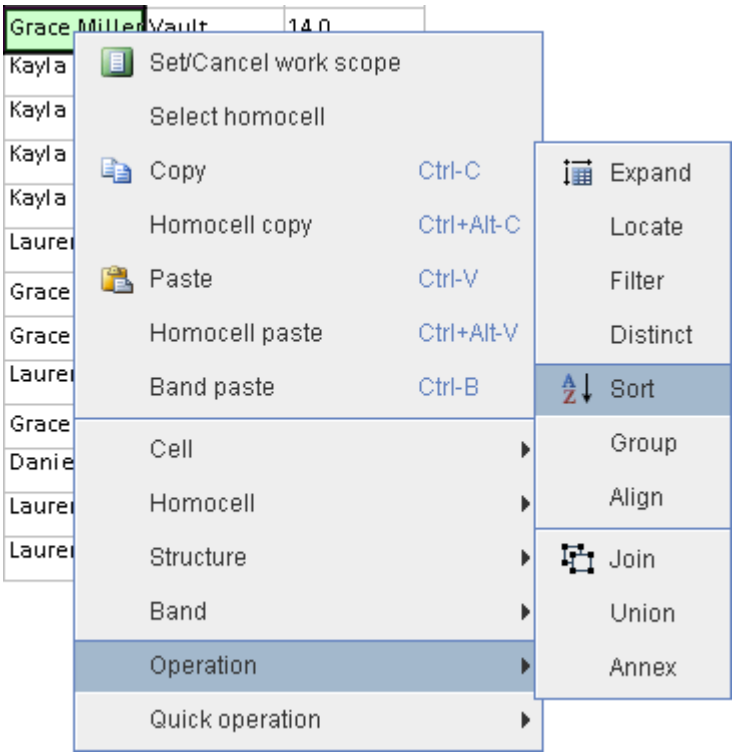
How to sort the data of calculation cellset? Let’s demonstrate the process with the gymnastics score sheet discussed in the previous chapter. Open the GymScores.txt with **Text with header** or open the calculation cellset esCalc_q03_1.gex directly. The cellset is as given below:

0	1		A	B	C
1-		1	**Name**	**Event**	**Score**
	1	2	**Grace Miller**	**Vault**	**14.0**
	1	3	**Kayla Rodriguez**	**Vault**	**14.575**
	1	4	**Kayla Rodriguez**	**Uneven bars**	**14.675**
	1	5	**Kayla Rodriguez**	**Balance beam**	**13.975**
	1	6	**Kayla Rodriguez**	**Floor**	**14.4**
	1	7	**Lauren Davis**	**Vault**	**14.275**
	1	8	**Grace Miller**	**Floor**	**13.975**
	1	9	**Grace Miller**	**Uneven bars**	**12.875**
	1	10	**Lauren Davis**	**Floor**	**13.0**
	1	11	**Grace Miller**	**Balance beam**	**13.1**
	1	12	**Daniel Smith**	**Parallel bars**	**14.375**
	1	13	**Lauren Davis**	**Balance beam**	**15.225**
	1	14	**Lauren Davis**	**Uneven bars**	**14.275**

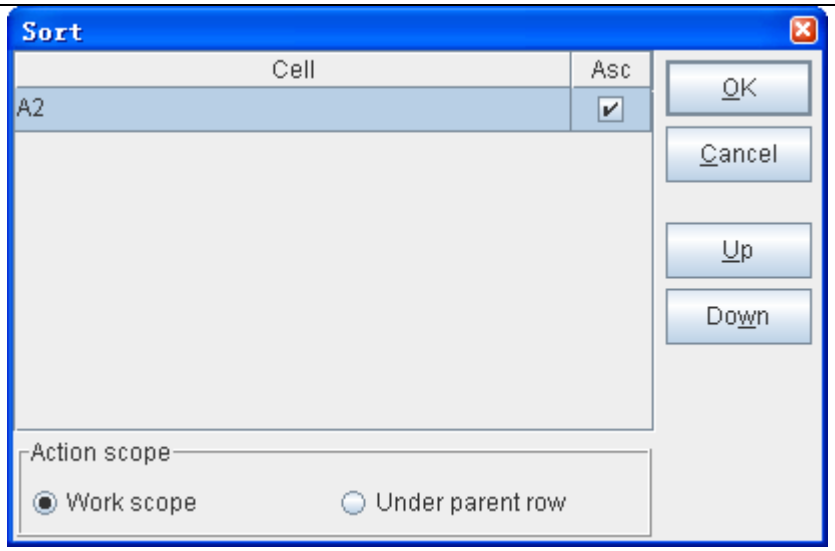
As given in the above table, in the calculation cellset, you may find that the headers of data in each column are given in the first row. Since the second row, each row gives a score of an athlete in a certain event. These rows are similar to a record in the database table.

To facilitate the view and making statistics, you can sort on these rows as desired.

For example, you can sort the scores by name. Select A2 cell, and click **Operation** -> **Sort** to sort the data:



On the above right-click menu, you can choose to perform multiple operations under the Operation option. Once confirmed the sort operation, the action panel as given below will prompt:



On the Sort panel, you can review the selected cell for sorting, and check the **Asc** box for ascending order. If **Asc** box is checked and the action is confirmed, then the result will be as given below:

0	1		A	B	C
1-		1	**Name**	**Event**	**Score**
	1	2	**Daniel Smith**	**Parallel bars**	**14.375**
	1	3	**Grace Miller**	**Vault**	**14.0**
	1	4	**Grace Miller**	**Floor**	**13.975**
	1	5	**Grace Miller**	**Uneven bars**	**12.875**
	1	6	**Grace Miller**	**Balance beam**	**13.1**
	1	7	**Kayla Rodriguez**	**Vault**	**14.575**
	1	8	**Kayla Rodriguez**	**Uneven bars**	**14.675**
	1	9	**Kayla Rodriguez**	**Balance beam**	**13.975**
	1	10	**Kayla Rodriguez**	**Floor**	**14.4**
	1	11	**Lauren Davis**	**Vault**	**14.275**
	1	12	**Lauren Davis**	**Floor**	**13.0**
	1	13	**Lauren Davis**	**Balance beam**	**15.225**
	1	14	**Lauren Davis**	**Uneven bars**	**14.275**

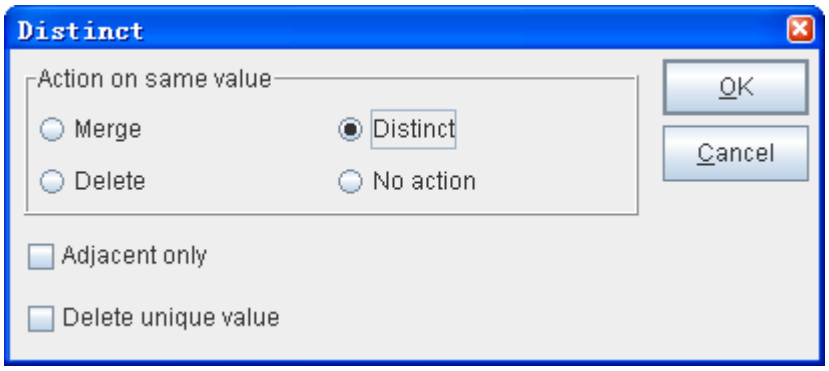
Judging from the result, you can see that all rows in the cellset is sorted by the name alphabetically in ascending order. Scores belong to a same name will remain in the original order.

Then, sort the rows in the cellset by the score in descending order. To do so, select C2 for sorting. On the operation panel, set the cell for sorting as C2, and clear the **Asc** option. The result is as given below:

0	1		A	B	C
1-		1	Name	Event	Score
	1	2	Lauren Davis	Balance beam	15.225
	1	3	Kayla Rodriguez	Uneven bars	14.675
	1	4	Kayla Rodriguez	Vault	14.575
	1	5	Kayla Rodriguez	Floor	14.4
	1	6	Daniel Smith	Parallel bars	14.375
	1	7	Lauren Davis	Vault	14.275
	1	8	Lauren Davis	Uneven bars	14.275
	1	9	Grace Miller	Vault	14.0
	1	10	Grace Miller	Floor	13.975
	1	11	Kayla Rodriguez	Balance beam	13.975
	1	12	Grace Miller	Balance beam	13.1
	1	13	Lauren Davis	Floor	13.0
	1	14	Grace Miller	Uneven bars	12.875

To check the participating athlete, the distinct operation is required.

On the calculation cellset that has just undergone the sorting, select A2 cell and click **Operation** -> **Distinct** to perform the Distinct action. On executing, the Distinct panel will appear. You can choose the **Distinct** option in the **Action on same value** section to remove the same values:



After the action, the result is as follows:

0	1		A	B	C
1-		1	Name	Event	Score
	1	2	Lauren Davis	Balance beam	15.225
	1	3	Kayla Rodriguez	Uneven bars	14.675
	1	4	Daniel Smith	Parallel bars	14.375
	1	5	Grace Miller	Vault	14.0

Compared with the original cellset, the current cellset has been deduplicated to remain distinct by name. Only one row is kept for each name and other rows are all removed.

The sort and below distinct actions can also be performed on grouping rows. Open the

esCalc_q03_2.gex, and the calculation cellset is as follows:

0	1	2		A	B	C
1-			1	Name	Event	Score
	1-		2	Grace Miller		
		1	3	Grace Miller	Vault	14.0
	1-		4	Kayla Rodriguez		
		1	5	Kayla Rodriguez	Vault	14.575
		1	6	Kayla Rodriguez	Uneven bars	14.675
		1	7	Kayla Rodriguez	Balance beam	13.975
		1	8	Kayla Rodriguez	Floor	14.4
	1-		9	Lauren Davis		
		1	10	Lauren Davis	Vault	14.275
	1-		11	Grace Miller		
		1	12	Grace Miller	Floor	13.975
		1	13	Grace Miller	Uneven bars	12.875
	1-		14	Lauren Davis		
		1	15	Lauren Davis	Floor	13.0
	1-		16	Grace Miller		
		1	17	Grace Miller	Balance beam	13.1
	1-		18	Daniel Smith		
		1	19	Daniel Smith	Parallel bars	14.375
	1-		20	Lauren Davis		
		1	21	Lauren Davis	Balance beam	15.225
		1	22	Lauren Davis	Uneven bars	14.275

For this calculation cellset, the record is not sorted before grouping. In this case, if selecting the master cell of a grouping row such as A4 to perform the sorting action, and the **Asc** option is cleared, then the result of sorting is as follows:

0	1	2		A	B	C
1-			1	Name	Event	Score
	1-		2	Lauren Davis		
		1	3	Lauren Davis	Vault	14.275
	1-		4	Lauren Davis		
		1	5	Lauren Davis	Floor	13.0
	1-		6	Lauren Davis		
		1	7	Lauren Davis	Balance beam	15.225
		1	8	Lauren Davis	Uneven bars	14.275
	1-		9	Kayla Rodriguez		
		1	10	Kayla Rodriguez	Vault	14.575
		1	11	Kayla Rodriguez	Uneven bars	14.675
		1	12	Kayla Rodriguez	Balance beam	13.975
		1	13	Kayla Rodriguez	Floor	14.4
	1-		14	Grace Miller		
		1	15	Grace Miller	Vault	14.0
	1-		16	Grace Miller		
		1	17	Grace Miller	Floor	13.975
		1	18	Grace Miller	Uneven bars	12.875
	1-		19	Grace Miller		
		1	20	Grace Miller	Balance beam	13.1
	1-		21	Daniel Smith		
		1	22	Daniel Smith	Parallel bars	14.375

From the above table, you can find that order remains the same for the record row in the group when sorting the grouping row in descending/ascending order. Acutually the record row will move along with its grouping row.

Again, in the grouping row, select the master cell holding the name, for example A21, to perform the Distinct action. On the pop-out Distinct panel, select the **Merge** option for the action on same value. The result is as follows:

0	1	2		A	B	C
1-			1	Name	Event	Score
	1-		2	Lauren Davis		
		1	3	Lauren Davis	Vault	14.275
		1	4	Lauren Davis	Floor	13.0
		1	5	Lauren Davis	Balance beam	15.225
		1	6	Lauren Davis	Uneven bars	14.275
	1-		7	Kayla Rodriguez		
		1	8	Kayla Rodriguez	Vault	14.575
		1	9	Kayla Rodriguez	Uneven bars	14.675
		1	10	Kayla Rodriguez	Balance beam	13.975
		1	11	Kayla Rodriguez	Floor	14.4
	1-		12	Grace Miller		
		1	13	Grace Miller	Vault	14.0
		1	14	Grace Miller	Floor	13.975
		1	15	Grace Miller	Uneven bars	12.875
		1	16	Grace Miller	Balance beam	13.1
	1-		17	Daniel Smith		
		1	18	Daniel Smith	Parallel bars	14.375

Perform the distinct operation on the grouping row, and records from respective groups with the same name will be merged into a same group as a result of selecting the Merge options.

5. Use Expressions in Calculation Cellset

When analyzing the data, it is often required to use the raw data to get the desired results. Open esCalc_q05_1.gex, and the calculation cellset as given below will display:

0	1	2		A	B	C	D	E	F
1-			1	Name	Type	UnitPrice	Unit	Quantity	
	1-		2		Fruit				
		1	3	Apples	Fruit	1.69	LB	1.2	
		1	4	Bananas	Fruit	0.69	LB	3.23	
		1	5	Oranges	Fruit	4.99	BAG	1	
		1	6	Red Grapes	Fruit	0.99	LB	2.87	
		1	7	Peach	Fruit	0.88	LB	1.6	
		1	8	Strawberry	Fruit	1.97	LB	1.25	
	1-		9		Other				
		1	10	Milk	Other	3.99	GAL	2	
	1-		11		Vegetable				
		1	12	Cucumbers	Vegetable	0.77	EACH	3	
		1	13	Onions	Vegetable	0.99	LB	1.33	

This cellset records the shopping list of a certain food supermarket, and groups the records by the type of foods. You can also choose to open reciept.txt in **Text with header**, and then try to produce this calculation cellset by grouping.

There is a number in the Hierarchy Section on the left for each row in the calculation cellset. It is the S**equence Number** of this row. In the above calculation cellset, every row has a sequence number of 1. The column number of the hierarchy column in which the sequence number is located is the **Level** of this row, for example, the level for the 3rd row is 2; the level for the 2nd row is 1; and the level for the 1st row is 0. In this calculation cellset with simple structure, similar data are stored in the rows at the same level, for example, the row 2, 9, and 11 are all the grouping row of goods category, and their levels are all 1; the row 3,4,5,6,7,8,10,12, and 13 are all records of the purchase data of certain food, and their levels are all 2.

In esCalc, if the cell string is started with "=", then the string following the equal mark will be resolved to an e**xpression** whose result is the cell value. Similar to Excel, for any cell with value, esCalc allows for the expression using the **cell name** directly as the variable name to invoke. In addition, esCalc provides another expression started with two equal mark "==".

In the data cellset of shopping list, the unit price and purchased quatity of each item of food is recorded. The price of each category of food can be calculated based on these records. To calculate, multiply the unit price with quantity, and keep two decimal digits. In F3, type in the expression ==round(C3*E3,2) and the result is as follows:

0	1	2		A	B	C	D	E	F
1-			1	Name	Type	UnitPrice	Unit	Quantity	
	1-		2		Fruit				
		1	3	Apples	Fruit	1.69	LB	1.2	2.03
		1	4	Bananas	Fruit	0.69	LB	3.23	2.23
		1	5	Oranges	Fruit	4.99	BAG	1	4.99
		1	6	Red Grapes	Fruit	0.99	LB	2.87	2.84
		1	7	Peach	Fruit	0.88	LB	1.6	1.41
		1	8	Strawberry	Fruit	1.97	LB	1.25	2.46
	1-		9		Other				
		1	10	Milk	Other	3.99	GAL	2	7.98
	1-		11		Vegetable				
		1	12	Cucumbers	Vegetable	0.77	EACH	3	2.31
		1	13	Onions	Vegetable	0.99	LB	1.33	1.32

As given in the above table, the price of apples is calculated in F3. One thing to note is that the calculation results are all generated in respective row at the same level as that of the 3rd row. To select the cells in column F to review their expressions, you will find that the expression in F7 is ==round(C7*E7,2), the expression in F10 is ==round(C10*E10,2), and so on and so forth. In the simple structured calculation cellset, the expressions in the cells in the same colume and the rows at the same level will be set and adjusted automatically. For example, in F7 and F10, the total price of goods in respecieve row is calculated respectively.

By definition, the **Related Calculation Cell** is a cell in which the expression is started with two equal marks. A blue triangle will appear on the lower right corner as an indicator for the related calculation cell. In the simple-structured calculation cellset of esCalc, the cells in the same column and the rows at the same level are refered to as **Homocell**. The expression in the homocell will automatically adjust while setting and no manual intervention is required.

The price of each item is calculated in the cellset. Then, how to calculate the total price of each category of goods? To obtain the total price, we can use an expression with one equal sign in this case. Populate F1 and F2 with a same expression ={F3}.sum(), and the result is as given below:

0	1	2		A	B	C	D	E	F
1-			1	Name	Type	UnitPrice	Unit	Quantity	27.57
	1-		2		Fruit				15.96
		1	3	Apples	Fruit	1.69	LB	1.2	2.03
		1	4	Bananas	Fruit	0.69	LB	3.23	2.23
		1	5	Oranges	Fruit	4.99	BAG	1	4.99
		1	6	Red Grapes	Fruit	0.99	LB	2.87	2.84
		1	7	Peach	Fruit	0.88	LB	1.6	1.41
		1	8	Strawberry	Fruit	1.97	LB	1.25	2.46
	1-		9		Other				7.98
		1	10	Milk	Other	3.99	GAL	2	7.98
	1-		11		Vegetable				3.63
		1	12	Cucumbers	Vegetable	0.77	EACH	3	2.31
		1	13	Onions	Vegetable	0.99	LB	1.33	1.32

The total price of fruits is calculated in F2, and the expression is also set automatically in the homocell F9 and F11, representing the total prices of food categories of Other and Vegetable respectively. In F1, the total price of all foods is calculated. In the expression, {F3} represents all homocells of F3 in the current group. Various aggregation calculations can be conducted on the resulting sets, such as {F3}.sum(), {F3}.count(), {F3}.avg(),{F3}.max(), and others. To make statistics on the total of each group, you can just compose the expression in various grouping rows, for example, compose the expression in the 1st row at level 0 to calculate the total price of all foods.

You can notice that both F1 and F2 are using an expression started with one equal mark. In this case, the lower right corners of cell F1, F2, F9, and F10 are in cyan. Such cells are the **Calculation Cell**.

In the below table, the unit price of apple is adjusted from 1.69 USD to 1 USD. Change the value in C3 to 1, and the result is as follows:

0	1	2		A	B	C	D	E	F
1-			1	Name	Type	UnitPrice	Unit	Quantity	27.57
	1-		2		Fruit				15.96
		1	3	Apples	Fruit	1	LB	1.2	1.2
		1	4	Bananas	Fruit	0.69	LB	3.23	2.23
		1	5	Oranges	Fruit	4.99	BAG	1	4.99
		1	6	Red Grapes	Fruit	0.99	LB	2.87	2.84
		1	7	Peach	Fruit	0.88	LB	1.6	1.41
		1	8	Strawberry	Fruit	1.97	LB	1.25	2.46
	1-		9		Other				7.98
		1	10	Milk	Other	3.99	GAL	2	7.98
	1-		11		Vegetable				3.63
		1	12	Cucumbers	Vegetable	0.77	EACH	3	2.31
		1	13	Onions	Vegetable	0.99	LB	1.33	1.32

As you can see, the total price of apples calculating in F3 is changed accordingly once the C3 cell is modified. However, the total prices of fruits in F2 and that of the all foods in F1 have not been updated. Just like the calculation cell of Excel, the related calculation cell of esCalc will change instantly and recalculate once the invoked cell value is modified. By comparison, the normal calculation cell only calculates once when entering the expression.

In esCalc, since the related calculation cell will use relatively more resources, the calculation cell is more often used. If the data for computation may be changing in the cell, then related calculation cell may have to be used to avoid the computation error.

Open esCalc_q05_5.gex to display the cellset below:

0	1	2		A	B	C
1-			1	State	Year	Population
	1-		2	California		
		1	3	California	1990	29760021
		1	4	California	2000	33871648
		1	5	California	2010	37253956
	1-		6	Texas		
		1	7	Texas	1990	16986510
		1	8	Texas	2000	20851820
		1	9	Texas	2010	25145561
	1-		10	New York		
		1	11	New York	1990	17990455
		1	12	New York	2000	18976457
		1	13	New York	2010	19378102

The data in the calculation cellset is the population statistics of some states in recent years. You can choose to open the population.txt in **Text with header**, and then try to generate the

calculation cellset through the grouping operation.

Add a column on the right part of this table. To do so, select any cell in the 3rd column. For example the C1 cell, and press Ctrl+I at the same time. In the D13 cell of the new column, set the expression =C13-C12 to calculate the population increment and the result is as follows:

0	1	2		A	B	C	D
1-			1	State	Year	Population	
	1-		2	California			
		1	3	California	1990	29760021	29760021
		1	4	California	2000	33871648	4111627
		1	5	California	2010	37253956	3382308
	1-		6	Texas			
		1	7	Texas	1990	16986510	-20267446
		1	8	Texas	2000	20851820	3865310
		1	9	Texas	2010	25145561	4293741
	1-		10	New York			
		1	11	New York	1990	17990455	-7155106
		1	12	New York	2000	18976457	986002
		1	13	New York	2010	19378102	401645

From the above table, you can find that the population increment of homocells of D12 is also calculated by migrating expression. When migrating the expression, the population increment will be calculated by subtracting the population in the previous homocell from the the current population.

With more attention, you will find that there is an error in D7 and D11 because the previous homocell is not in a same group. In this case, you can avoid the error by restricting the homocells with level. To do so, modify the expression in D13 to =C13-C12[A10] so that the previous record of population data in C12 will be restricted to the level below the current grouping level A10. After modification, the calculation result is as follows:

0	1	2		A	B	C	D
1-			1	State	Year	Population	
	1-		2	California			
		1	3	California	1990	29760021	29760021
		1	4	California	2000	33871648	4111627
		1	5	California	2010	37253956	3382308
	1-		6	Texas			
		1	7	Texas	1990	16986510	16986510
		1	8	Texas	2000	20851820	3865310
		1	9	Texas	2010	25145561	4293741
	1-		10	New York			
		1	11	New York	1990	17990455	17990455
		1	12	New York	2000	18976457	986002
		1	13	New York	2010	19378102	401645

Restricting the homocell with levels can steer clear of the potential errors when auto-adjust the expression, and ensure a correct calculation of period-on-period comparison.

Benefited from its simple structure, the calculation cellset provides rather simple relations between rows and an intuitive presentation to recognize the homocell. For the information on homocell in the complicated calculation cellset, please refer to the related topics in the tutorial.

6. Locate and Filter

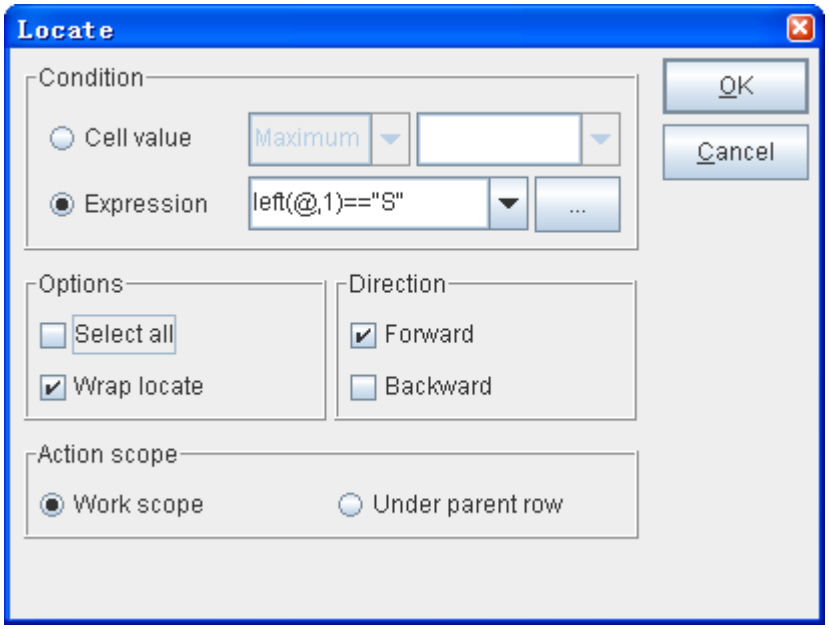
When analyzing data, you often need to locate the record meeting the conditions according to requirement.Then, how to locate in the esCalc?

Open esCalc_q06_1.gex and the below cellset will appear:

0	1	2		A	B	C
1-			1	State	City	Population
	1-		2	TX		
		1	3	TX	Houston	2144491
		1	4	TX	Dallas	1232940
		1	5	TX	El Paso	609415
		1	6	TX	Austin	709893
	1-		7	CA		
		1	8	CA	San Diego	1256951
		1	9	CA	Fresno	466714
		1	10	CA	San Jose	929936
	1-		11	TN		
		1	12	TN	Memphis	670902
		1	13	TN	Nashville	552120

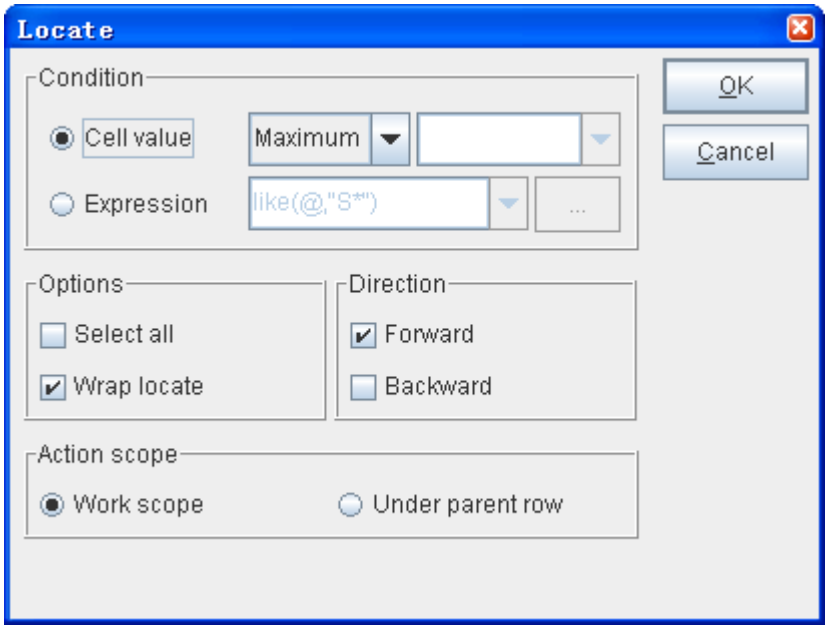
You may also open the cities.txt in **Text with header**, and then try to produce this calculation

cellset through the grouping operation. Let's try to find the first city whose initial is S. To do so, select the B3 cell, and click **Operation** -> **Locate** on the right-click menu to perform the locating operation. The locating operation will judge on the data in all homocells. Move the cursor to the first condition-satisfied homocell and the below Locate panel will prompt as follows:



On the Locate panel, set the conditions for locating to an expression left(@,1)=="S". Of which, "@" represents the value of the current cell, that is, the selected B3 and the value of its homocell. The setup of other actions in the cell is as illustrated above. Once acted, the cell will move to B8. The expression for locating can also be set to like(@,"S*"), and the execution result is the same.

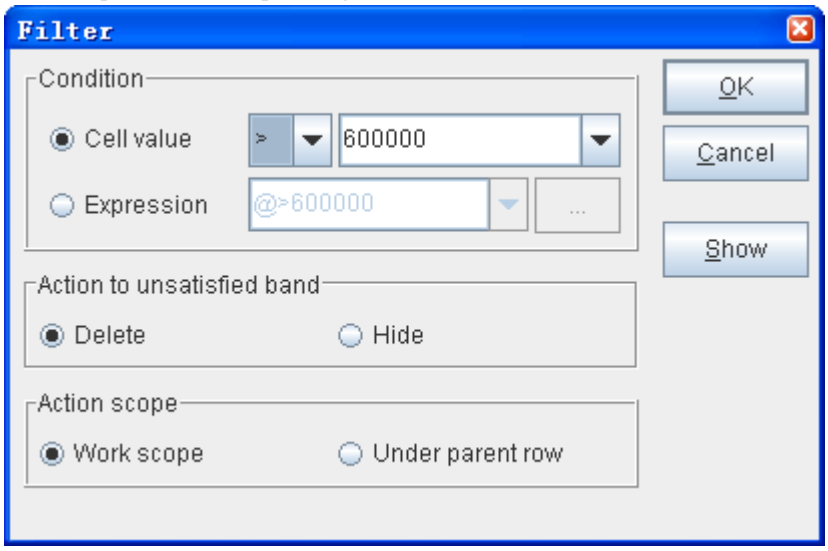
To find the city with the largest population in the cellset, you can pick any cell with population data, for example, C13 cell, and perform the locating. In the action panel, the setup is as follows:



Select the locating condition of **Maximum** cell value. After locating, the censor will move to the cell C3.

When analyzing data, sometimes you need to filter the data and delete the unsatisfied data. What actions shall we take in esCalc?

For example, in the above cellset, we need to filter on the data from respective city to only keep the data from the city whose population is greater than 600,000. In esCalc, the filter action can be taken to judge all data in the homocells. Remove or hide the row to which the homocell unsatisfying the conditions belongs. To do so, select any homocell to hold the population data, for example C6. On the right-click menu, click **Operation** –> **Filter** to take the filter action. On the popout Filter panel, the setup is as given below:



The result of filtering is as given below:

0	1	2		A	B	C
1-			1	State	City	Population
	1-		2	TX		
		1	3	TX	Houston	2144491
		1	4	TX	Dallas	1232940
		1	5	TX	El Paso	609415
		1	6	TX	Austin	709893
	1-		7	CA		
		1	8	CA	San Diego	1256951
		1	9	CA	San Jose	929936
	1-		10	TN		
		1	11	TN	Memphis	670902

In the cellset, records not meeting the requirements will be deleted.

7. Join and Union

In the data processing transaction, the data from multiple database tables are sometimes required to be joined up according to the sequence number and other information, and render in

one table. In esCalc, the data from multiple calculation cellsets can be joined up accordingly through the join action.

Open esCalc_q7_1.gex or open states.txt in **Text with header**, and you will get the below calculation cellset:

0	1		A	B	C	D	E	F
1-		1	StateId	Name	Population	Abbr	Area	Capital
	1	2	5	California	37253956	CA	163700	Sacramento
	1	3	9	Florida	18801310	FL	65755	Tallahassee
	1	4	13	Illinois	12830632	IL	54826	Springfield
	1	5	32	New York	19378102	NY	54556	Albany
	1	6	43	Texas	25145561	TX	268820	Austin

In the calculation cellset, select D2. Under the **cell** option on the right-click menu, select **Set/Cancel master cell** to set the abbreviation of state as the master cell for joining. When setting the master cell, all homocells will be set at the same time.

Select both B2 and C2 at the same time, and press Ctrl+C to prepare to copy the full name and population of the state to other cellsets.

Then reopen the esCalc_q06_1.gex, and the city information cellset is as follows:

0	1	2		A	B	C
1-			1	State	City	Population
	1-		2	TX		
		1	3	TX	Houston	2144491
		1	4	TX	Dallas	1232940
		1	5	TX	El Paso	609415
		1	6	TX	Austin	709893
	1-		7	CA		
		1	8	CA	San Diego	1256951
		1	9	CA	Fresno	466714
		1	10	CA	San Jose	929936
	1-		11	TN		
		1	12	TN	Memphis	670902
		1	13	TN	Nashville	552120

Select B2. On the right-click menu, click **Operation** -> **Join**. On the prompt panel, select the default Left Join, and take the Join action. The result is as given below:

0	1	2		A	B	C
1-			1	State	City	Population
	1-		2	TX	Texas	25145561
		1	3	TX	Houston	2144491
		1	4	TX	Dallas	1232940
		1	5	TX	El Paso	609415
		1	6	TX	Austin	709893
	1-		7	CA	California	37253956
		1	8	CA	San Diego	1256951
		1	9	CA	Fresno	466714
		1	10	CA	San Jose	929936
	1-		11	TN		
		1	12	TN	Memphis	670902
		1	13	TN	Nashville	552120

From the above table, you can find that the name and population data of States have been copied to the cellset of City Information through the align-join action on the State Abbreviation.

When processing the data, some data are not from an individual table. Then, you will need to gather the data from multiple tables for presentatation and computation.

Open esCalc_q07_3.gex and esCalc_q07_4.gex, or open the students1.txt and students2.txt in **Text with header**. Let's check the below two calculation cellsets:

0	1		A	B	C	D
1-		1	ID	Name	Gender	Age
	1	2	1	Emily	F	17
	1	3	2	Elizabeth	F	16
	1	4	3	Sean	M	17
	1	5	4	Lauren	F	15

0	1		A	B	C	D
1-		1	ID	Name	Gender	Age
	1	2	5	Michael	M	16
	1	3	6	John	M	13
	1	4	3	Nicholas	M	16

In both cellsets, there are some student data to be unioned according to their student numbers.

In esCalc, you can union the data from multiple calculation cellsts through the Union operation. Firstly, in both cellsets, set the A2 holding the Sudent ID as the Master Cell by which the union action will be performed. Then, in the second calculation cellset, select any cell (for example, A2) in the second row. By pressing Ctrl+C, copy all student data from the student data cellset 2, and prepare to union with the student data from the student data cellset 1. Pick any cell (for example, C4) of student data row from the first cellset. On the right-click menu, click

Operation -> **Union** to perform the union operation. On the popout control panel, select the way of **Union**. The result is as follows after union:

0	1		A	B	C	D
1-		1	ID	Name	Gender	Age
	1	2	1	Emily	F	17
	1	3	2	Elizabeth	F	16
	1	4	3	Sean	M	17
	1	5	4	Lauren	F	15
	1	6	5	Michael	M	16
	1	7	6	John	M	13

After union, the student data from both calculation cellsets are unioned. In the selected mode of u**nion**, the student data with a duplicate student number will not appear repeatedly.