

M 365 Excel Class Video 16: Power Query Group By in Excel & Power BI



p

1. What Does Power Query Group By feature do?	2
2. Concept of “Group By” is to group records together based on a condition or criteria:	3
This example groups records by Product:	3
3. Excel or Power BI: Where is Group By feature located? :	3
4. You Already Have Used Group By Concepts Before with These 5 Tools:	4
Table.Group M Code Function (First Three Arguments)	4
Video Topics:	5
How to Hack Group By dialog box (Standard Deviation calculation)	5
Underscore gets Table of Grouped Records	5
Type M Code for Group By Calculation (Unique List of SalesReps who Sold Product)	5
Power Query Rank with Grouping and Table.AddRankColumn	6
Group By: Two Conditions	7
Group By Back-To-Back: Frequency Distribution For Failed Student Class Combination	7
Group By: Add Blank Row After Each Group (Bill Szysz)	8
Group By 4th Argument: Consecutive Occurrences	9

1. What Does Power Query Group By feature do?

- It creates tabular results with a unique list of items and specified calculations for each item in the unique list.
- The Group By feature will:
 1. Create a unique list of items from one or more columns and
 2. Make an aggregate calculation for each item in the unique list.
- Group By Power Query Feature is similar to:
 - i. PivotTables
 1. Where you get a Unique List of items in the Rows Area and make an Aggregate Calculation for each row in that Unique List.
 - ii. SUMIFS and COUNTIFS (and other similar functions)
 1. Where the function makes an aggregate calculation for each row in a column filled with a Unique List.
 - iii. SQL Group By
 1. Where the SQL code allows you to create a unique list of items and aggerate for each row.
- Example of Group By in Power Query:

What "Group By" Does:

Sales Table:

Date	Product	SalesRep	Sales
2/29/2020	Quad	Chauntel	1379.01
2/25/2020	Carlota	Tyrone	1018.27
2/2/2020	Bellen	Chauntel	2162.58
2/21/2020	Quad	Gigi	73.39
2/19/2020	Sunshine	Pham	2205.1
2/23/2020	Sunshine	Sioux	112.79
2/27/2020	Carlota	Tyrone	1171.23
2/12/2020	Carlota	Chauntel	617.22
2/12/2020	Bellen	Gigi	937.34
2/15/2020	Carlota	Chauntel	1562.88
2/12/2020	Bellen	Sioux	496.82
2/11/2020	Quad	Tyrone	2177.72
2/14/2020	Sunshine	Chauntel	2394.35
2/21/2020	Sunshine	Gigi	1097.95
2/20/2020	Quad	Pham	1920.33
2/17/2020	Quad	Sioux	1108.04
2/2/2020	Carlota	Tyrone	1170.44
2/4/2020	Carlota	Chauntel	2418.86

Group By Feature to get Report:

Product	Total Sales	Deviation of Sales	Who Sold Product
Bellen	\$3,597	\$863	Chauntel, Gigi, Sioux
Carlota	\$7,959	\$616	Tyrone, Chauntel
Quad	\$6,658	\$821	Chauntel, Gigi, Tyrone, Pham, Sioux
Sunshine	\$5,810	\$1,061	Pham, Sioux, Chauntel, Gigi

1) Extract Unique List from Column/Columns

2) Make Aggregate Calculations on Specified Number Column (Sales)

OR

2) Make Aggregate Calculations on Specified Text Column (SalesRep)

2. Concept of “Group By” is to group records together based on a condition or criteria:

This example groups records by Product:

Date	Product	SalesRep	Sales
2/2/2020	Bellen	Chauntel	2162.58
2/12/2020	Bellen	Gigi	937.34
2/12/2020	Bellen	Sioux	496.82

Date	Product	SalesRep	Sales
2/25/2020	Carlota	Tyrone	1018.27
2/27/2020	Carlota	Tyrone	1171.23
2/12/2020	Carlota	Chauntel	617.22
2/15/2020	Carlota	Chauntel	1562.88
2/2/2020	Carlota	Tyrone	1170.44
2/4/2020	Carlota	Chauntel	2418.86

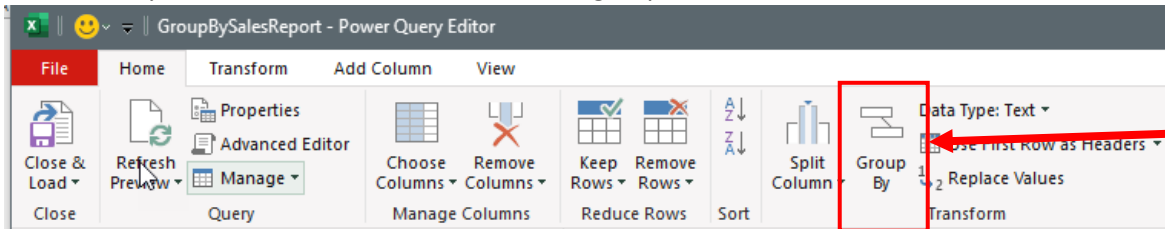
Date	Product	SalesRep	Sales
2/29/2020	Quad	Chauntel	1379.01
2/21/2020	Quad	Gigi	73.39
2/11/2020	Quad	Tyrone	2177.72
2/20/2020	Quad	Pham	1920.33
2/17/2020	Quad	Sioux	1108.04

Date	Product	SalesRep	Sales
2/19/2020	Sunshine	Pham	2205.1
2/23/2020	Sunshine	Sioux	112.79
2/14/2020	Sunshine	Chauntel	2394.35
2/21/2020	Sunshine	Gigi	1097.95

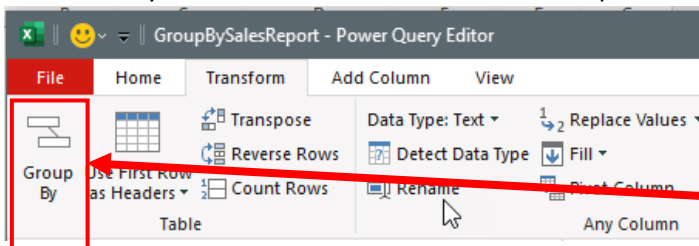
The Concept of Group By = Group Records together based a Condition or Criteria.

3. Excel or Power BI: Where is Group By feature located? :

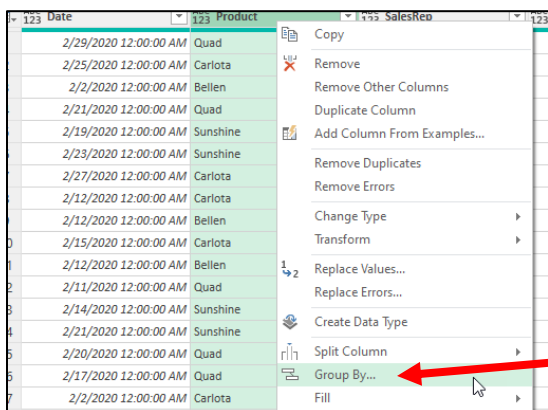
Power Query Home Ribbon Tab in the Transform group:



Power Query Transform Ribbon Tab in Table Group:



Right-Click Column/Columns you want to Group By:



4. You Already Have Used Group By Concepts Before with These 5 Tools:

Sales Table: "SalesT"

Date	Product	SalesRep	Sales
2/29/2020	Quad	Chauntel	1379.01
2/25/2020	Carlota	Tyrone	1018.27
2/2/2020	Bellen	Chauntel	2162.58
2/21/2020	Quad	Gigi	73.39
2/19/2020	Sunshine	Pham	2205.1
2/23/2020	Sunshine	Sioux	112.79
2/27/2020	Carlota	Tyrone	1171.23
2/12/2020	Carlota	Chauntel	617.22
2/12/2020	Bellen	Gigi	937.34
2/15/2020	Carlota	Chauntel	1562.88
2/12/2020	Bellen	Sioux	496.82
2/11/2020	Quad	Tyrone	2177.72
2/14/2020	Sunshine	Chauntel	3242.65
2/21/2020	Sunshine	Gigi	1097.95
2/20/2020	Quad	Pham	1920.33
2/17/2020	Quad	Sioux	1108.04
2/2/2020	Carlota	Tyrone	1170.44
2/4/2020	Carlota	Chauntel	2418.86

Group By With 5 Other Tools:

1) Standard PivotTable:

Product	Sum of Sales	SStdDev Sales
Bellen	3,596.74	863.14
Carlota	7,958.90	615.74
Quad	6,658.49	821.26
Sunshine	6,658.49	1,355.43
Grand Total	24,872.62	841.79

2) Power BI Matrix Visual:

Product	Total Sales (\$)	SStdDev (\$)
Bellen	3596.74	863.14
Carlota	7958.90	615.74
Quad	6658.49	821.26
Sunshine	6658.49	1,355.43
Total	24872.62	841.79

3) Worksheet Formulas:

Product	Total Sales	SStdDev Sales	Who Sold This Product?
Bellen	\$3,596.74	\$863.14	Chauntel, Gigi, Sioux
Carlota	\$7,958.90	\$615.74	Chauntel, Tyrone
Quad	\$6,658.49	\$821.26	Chauntel, Gigi, Pham, Sioux, Tyrone
Sunshine	\$6,658.49	\$1,355.43	Chauntel, Gigi, Pham, Sioux

G16: =SORT(UNIQUE(C5:C22))
H16: =SUMIFS(SalesT[Sales],SalesT[Product],G16:G19)
I16: =STDEV.S(IF(SalesT[Product]=G16,SalesT[Sales]))
J16: =TEXTJOIN(", ",SORT(UNIQUE(FILTER(SalesT[SalesRep],SalesT[Product]=G16))))

4) Data Model DAX PivotTable:

Product	Total Sales (\$)	StandDev (\$)	Who Sold Product?
Bellen	3,596.74	863.14	Chauntel, Gigi, Sioux
Carlota	7,958.90	615.74	Chauntel, Tyrone
Quad	6,658.49	821.26	Chauntel, Gigi, Pham, Sioux, Tyrone
Sunshine	6,658.49	1,355.43	Chauntel, Gigi, Pham, Sioux
Grand Total	24,872.62	841.79	Chauntel, Gigi, Pham, Sioux, Tyrone

Total Sales (\$):=SUM(SalesT[Sales])
StandDev (\$):=ROUND(STDEV.S(SalesT[Sales]),2)
Who Sold Product?:=CONCATENATEX(VALUES(SalesT[SalesRep]),SalesT[SalesRep],", ",SalesT[SalesRep],ASC)

5) SQL Code:

```
SELECT Product, SUM(Quantity) AS TotalUnits
FROM fTransactions
GROUP BY Product
```

Product	TotalUnits
Alpine	5,286,164
Aspen	6,319,712
Bellen	11,572,168
Bower Aussie	6,299,017
Carlota	12,601,534
Carlota Doubl	387,333
Crested Beaut	11,522,799
Darnell Tri Fly	4,171,419
Eagle	4,184,394
Sunset	5,252,626
Sunshine	6,300,895
Sunspot	3,145,017
Yanaki	6,290,484

Server: pond.highline.edu
Database: boomerang
User: excelisfun
PW: ExcellisFun!

Table.Group M Code Function (First Three Arguments)

`= Table.Group(Source, {"Product"}, {"Total Sales", each List.Sum([Sales]), type number})`

Product	Total Sales
1 Quad	6658.49
2 Carlota	7958.9
3 Bellen	3596.74
4 Sunshine	6658.49

3) Previous step (table)

2) List of fields that determine unique list/table

1) List in List with:

- 1) Field name
- 2) Formula for each row
- 3) Data type

Video Topics:

How to Hack Group By dialog box (Standard Deviation calculation)

1. Step 1: Using the Group By dialog box, add a place holder function like Sum
2. Step 2: Edit M Code in formula bar, like in the below picture where we changed the List.Sum function to List.StandardDeviation function

The formula bar shows the following M code:

```
= Table.Group(Source, {"Product"}, {"Total Sales ($)"} each List.Sum([Sales]), type number}, {"Standard Deviation ($)"} each List.StandardDeviation([Sales]), type number}})
```

Arrows point to the `List.Sum` and `List.StandardDeviation` functions in the code.

Product	Total Sales (\$)	Standard Deviation (\$)
1 Quad	6658.49	821.26
2 Carlota	7958.9	615.74
3 Bellen	3596.74	863.14
4 Sunshine	6658.49	1355.43

Underscore gets Table of Grouped Records

The formula bar shows the following M code:

```
= Table.Group(Source, {"Product"}, {"Total Sales ($)"} each List.Sum([Sales]), type number}, {"Standard Deviation ($)"} each Number.Round(List.StandardDeviation([Sales]),2), type number}, {"Sales Reps"} each _))
```

The table below shows the 'Sales Reps' column containing table objects:

Product	Total Sales (\$)	Standard Deviation (\$)	Sales Reps
1 Quad	6658.49	821.26	Table
2 Carlota	7958.9	615.74	Table
3 Bellen	3596.74	863.14	Table
4 Sunshine	6658.49	1355.43	Table

An arrow points from the 'Sales Reps' column header to the expanded table below.

Date	Product	SalesRep	Sales
2/29/2020 12:00:00 AM	Quad	Chauntel	1279.01
2/21/2020 12:00:00 AM	Quad	Gigi	73.39
2/11/2020 12:00:00 AM	Quad	Tyrone	2177.72
2/20/2020 12:00:00 AM	Quad	Pham	1920.33
2/17/2020 12:00:00 AM	Quad	Sioux	1108.04

Type M Code for Group By Calculation (Unique List of SalesReps who Sold Product)

The formula bar shows the following M code:

```
= Table.Group(Source, {"Product"}, {"Total Sales ($)"} each List.Sum([Sales]), type number}, {"Standard Deviation ($)"} each Number.Round(List.StandardDeviation([Sales]),2), type number}, {"Sales Reps"} each Text.Combine(List.Sort(List.Distinct([SalesRep]),0),", "), type text}})
```

Arrows point to the `List.Distinct` and `Text.Combine` functions in the code.

Product	Total Sales (\$)	Standard Deviation (\$)	Sales Reps
1 Quad	6658.49	821.26	Chauntel, Gigi, Pham, Sioux, Tyrone
2 Carlota	7958.9	615.74	Chauntel, Tyrone
3 Bellen	3596.74	863.14	Chauntel, Gigi, Sioux
4 Sunshine	6658.49	1355.43	Chauntel, Gigi, Pham, Sioux

Power Query Rank with Grouping and Table.AddRankColumn

Table.AddRankColumn(Table , "NewColumnName" , {"RankColumnName",Order.Descending or Order.Ascending}, [RankKind = RankKind.XXX])			
RankKind:	Value	Result (tie 2)	Description
RankKind.Competition	0	1,2,2,4	Items which compare as equal receive the same ranking number and then a gap is left before the next ranking.
RankKind.Dense	1	1,2,2,3	Items which compare as equal receive the same ranking number and the next item is numbered consecutively with no gap.
RankKind.Ordinal	2	1,2,3,4	All items are given a unique ranking number even if they compare as equal.

The below M Code shows how we accomplished these types of sort in the video:

```

let
    Source = Excel.CurrentWorkbook(){[Name="fSalesVideo"]}[Content],

    GroupProductSales = Table.Group(Source, {"Product"},
    {"Sales($)", each List.Sum([Sales]), type number})),

    Rank1224 = Table.AddRankColumn(GroupProductSales,"Rank1224","Sales($)",Order.Descending),
    [RankKind=RankKind.Competition]),
    Rank1223 = Table.AddRankColumn(Rank1224,"Rank1223","Sales($)",Order.Descending),
    [RankKind=RankKind.Dense]),

    AddIndex = Table.AddIndexColumn(Rank1223, "Index", 1, 1, Int64.Type),
    GroupSales = Table.Group(AddIndex, {"Sales($)", {"T", each _},
    {"RankAve", each List.Average([Index]), type number})),
    Expand = Table.ExpandTableColumn(GroupSales, "T",
    {"Product", "Rank1224", "Rank1223"}, {"Product", "Rank1224", "Rank1223"}),

    MoveFields = Table.ReorderColumns(Expand,
    {"Product", "Sales($)", "Rank1224", "Rank1223", "RankAve"})
in
    MoveFields
    
```

Result of M Code:

Product	Sales(\$)	Rank1224	Rank1223	RankAve
Carlota	7958.9	1	1	1
Quad	6658.49	2	2	2.5
Sunshine	6658.49	2	2	2.5
Bellen	3596.74	4	3	4

Group By: Two Conditions

✕ ✓ fx

```
= Table.Group(AddDataTypes,
{"Product", "SalesRep"}, ← Group By 2 columns
{"Total Sales ($)", each List.Sum([Sales]), type nullable number}})
```

	A ^B C Product	A ^B C SalesRep	1.2 Total Sales (\$)
1	Quad	Chauntel	1379.01
2	Carlota	Tyrone	3359.94
3	Bellen	Chauntel	2162.58
4	Quad	Gigi	73.39

Group By Back-To-Back: Frequency Distribution For Failed Student Class Combination

```
let
    Source = CsvFileDynamicFilePath,

    GetFailedClasses = Table.SelectRows(Source, each [Grade] < 0.7),

    GroupIDandQuarter = Table.Group(GetFailedClasses, {"SID", "Quarter"},
    {"Count", each Table.RowCount(_), Int64.Type},
    {"Classes", each Text.Combine([Class], ", "), type text}},

    RemoveOneFailedClass = Table.SelectRows(GroupIDandQuarter, each ([Count] <> 1)),

    GroupFailClassComboCount = Table.Group(RemoveOneFailedClass, {"Classes"},
    {"Count", each Table.RowCount(_), Int64.Type}},

    SortCount = Table.Sort(GroupFailClassComboCount, {"Count", Order.Descending})
in
    SortCount
```

Result of M Code:

Classes	Count
English Composition I, Interm Alg For Calculus	124
College Writ Strategies, English Composition I	53
English Composition I, General Psychology	46
English Composition I, Essentials Of Interm Alg	35
Essentials Of Interm Alg, General Psychology	19
Ell Ex.C.E.L. English 3, College Success Seminar	19
Intro To College Writing, Reading Skills Iii, Intro To Theatre	17
Intro To Chemistry, Introduction To Stats	17
Public Speaking, English Composition I	15
Hs Cont. World Issues, Hs English 2, College Success Seminar	15
S.C.O.P.E. English 3, S.C.O.P.E. English 3, College Success Seminar	15
College Success Seminar, English Composition I	15
Introduction To Comm, English Composition I	14
English Composition I, College Algebra	14
English Composition I, Us History Iii	13


Group By: Add Blank Row After Each Group (Bill Szysz)

The goal was to add a blank row after each set of grouped records. To accomplish this, we used the Grouped By formula:

`_ɾtable({"Product"},{null})`

Where the underscore retrieved the table of grouped records in each row, the table literal created a one column table with a single null value, and the Join Operator (& Ampersand) allows us to append two tables.

Date	Product	Sales
2/14/2023	Quad	305
2/14/2023	Carlota	1021
2/14/2023	Aspen	854
2/14/2023	Yanaki	560
2/14/2023	Quad	1039
2/14/2023	Carlota	888
2/14/2023	Aspen	1263
2/14/2023	Yanaki	139
2/14/2023	Quad	1183
2/14/2023	Carlota	138
2/14/2023	Aspen	117



Date	Product	Sales
2/14/2023	Aspen	1007
2/14/2023	Aspen	117
2/14/2023	Aspen	854
2/14/2023	Aspen	1263
2/14/2023	Carlota	138
2/14/2023	Carlota	888
2/14/2023	Carlota	1021
2/14/2023	Carlota	175
2/14/2023	Quad	1255

```
let
    Source = Excel.CurrentWorkbook(){[Name="AddBlanks"]}[Content],
    AddDataTypes = Table.TransformColumnTypes(Source,{{"Date", type date},
{"Product", type text}, {"Sales", Int64.Type}}),
    SortProductAtoZ = Table.Sort(AddDataTypes,{{"Product", Order.Ascending}}),
    AddBlankRows = Table.Combine(Table.Group(SortProductAtoZ, {"Product"},
{"Temp", each _&#x27Etable({"Product"},{null}),type table}))[Temp]
in
    AddBlankRows
```


Group By 4th Argument: Consecutive Occurrences

- Table.Group M Code function arguments:
 - Table.Group(Table, GroupByColumns, ListAsListWithAggregations(NameAggregationType), GroupKind)
 - ** The is a 5th argument, not discussed here.
- The fourth argument is GroupKind:
 - GroupKind.Global = 1 = Default
 - A Unique List is determined from the Group By Column and then the Group By Aggregation is performed.
 - GroupKind.Local = 0
 - The Group By Aggregation is performed based on consecutive occurrences of items in the Group By Column.

Baseball Example from video to create a Win/Loss Consecutive Occurrence Report, as show here:

✕ ✓ fx

```
= Table.Group(SortDates, {"Result"},  
{{"Count", each Table.RowCount(_), Int64.Type},  
{"Dates", each Text.From(List.Min([Date]))&" to "&Text.From(List.Max([Date]))}, type text})  
,0)
```

← **0 to group by Consecutive Occurrences**

	A ^B _C Result	1 ² ₃ Count	A ^B _C Dates
1	Loss		2 4/8/2022 to 4/9/2022
2	Win		2 4/10/2022 to 4/11/2022
3	Loss		1 4/12/2022 to 4/12/2022
4	Win		2 4/13/2022 to 4/14/2022
5	Loss		1 4/15/2022 to 4/15/2022