

# 1 ( '□' ) ^ \_ \_

## 1.1 创建数据库并读入

在excel中打开数据，从其他源里面选取 Microsoft query。在Query里面选择新的数据源可以从创建新的数据源，并在创建的过程中注意驱动程序的选择和数据库文件。这样就可以创建一个新的数据库。同时也可以从query里面选择需要读入的数据。在具体读入的时候，可以选择具体的要读入的表。

## 1.2 条件查询

不同列之间的条件做并，不同行之间的条件作或。针对日期字段的查询要写为 #1996/7/1#。其他的数字值可以直接用>,<之类的符号进行筛选。

## 1.3 多表查询

### 1.3.1 内连接

内连接：是将多个表中符合条件的记录挑出来组成一个结果。在内连接查询中，几个表之间一定要有关系。如果没有表表和之间没有关系，需要使用中间过渡的表来使得我们需要读取数据的表相互连接。也可以手动将两个字段关联起来，但是要保证字段中的类型和长度要一致，字段名字可以不同，相当于同种数据直接根据自身的值的到其他表中进行对应以查到进一步的信息。

连接的条件和查询之间的关系：每一个连接会有一个等式表示要对应的数据关系，这个等式的意思是：其连接字段值同时出现在两个表中的（比如「2」这个数字都在两个表中的相连接字段中出现了，那么在查询1表的某些值时会出现连接字段值为2的数据）每一个对象会作为查询的总范围。

书上的两个例子：不同表中表示同一个东西的不同名字的「运货商」，「运货商id」；也可以时表示层级关系的上下级关系对应（具体是其上级栏位中对应了上级的雇员id，这样就可以和另一个复制表中的雇员id建立关系）。但是如果没有寻找到对应关系的字段将不会显示出来。

### 1.3.2 外连接

外连接适用于将不符合连接条件的记录一并查询出来。在连接的选项中，可以选择将表中没有列入查询范围的对象可以在此表中查询到（所有的此表对象均显示）。

## 1.4 计算字段

Query中也可以使用类似excel的方式进行简单的计算。在计算的时候在选择变量的时候需要具体指定在哪个表的字段。例如：数量 \* 订单明细.单价。在excel中计算的时候注意公式需要从字段中拖过来保证能够用( '□' ) ^ \_ \_。

# 2 数据分类汇总

这个是在excel里面进行操作。我们只用D函数和透视表(0.0)

## 2.1 一些汇总操作之前的操作

存储在excel表中的数据也可能作为一种数据源的方式，然后里面里面的主表可能是有ID构成然后有字表来解释ID的意义。这时候我们也可以用query然后在新的excel文档中打开。注意选取正确的驱动程序来打开新的数据源。在query里面打开的excel作为的数据源时要添加表需要在「选项」里面打开「系统表」。excel之间的表需要手动建立关系。

## 2.2 数据透视表

在「插入」选项中可以插入数据透视表(图)。数据区域不包含字段名，会自动忽略然后将他们加入到字段列表里面。由「报表筛选」「列标签」「行标签」「数值」来管理各种的标签和最终汇总表样式。透视表完成之后在光标放在数据透视表中时可以在「数据透视工具」里面可以做存储数据透视图。也可以直接按Alt + D + P启动数据透视表创建向导然后直接从外部数据源导入数据建立数据透视图。

将字段放在「报表筛选」之中表示下面表格中出现的项都属于报表筛选中规定的属性。在「行标签」中，最直观显示出来的就是最上面的分类项，其下面的则是对于其父项的更细的分类。

### 2.2.1 按月汇总

随便点击一个在数据透视表其中的一个单元格，右键「创建组」就可以对于不同的条件进行汇总。得到数据透视图的折线图时，点击图右键可以使用「趋势线」进行预测。在「预测线」界面可以选择显示方程和R值。R值处于[0, 1)，越靠近1越贴合。由于结束的月份只包含很少的天数，所以我们不包括最后结束的月份以保证预测的正确性。在标签里面可以选择不显示隐藏的月份。对于月份的缺失（当月数据为0会不出现在透视表中），可能需要插入这些月份。在字段列表里面点击下面区域中的相应字段右键「字段设置」，在布局 and 打印中恢复无数据项。对于无数据项进行值修改，右键此字段点击「数据透视表选项.布局」，对于空单元格显示0，注意设置为数字0而非字符0。在数据透视表透视里面可以点击设计可以更高改显示放法（比如表格形式）。在将需要的数据提取出来的时候，可以在第一个位置指定对应的位置，然后将框拖下去就可以得到下面的对应的数据的复制。复制的时候注意单元格的格式设置。

在query里面的合并：在query里面的字段名双击可以「求和」。有两种用法：通过一个没啥用的字段来合并一二关键字相同的项，用于计数。第二种是合并的按照一二关键字的某个有意义数值，比如销售额求和。

### 2.2.2 对于销售次数的频率分布的汇总

首先得到一个含有订单时间，销量和产品名称的汇总表。将数量分别拖到行标签和数量中表示「销售数量为x的销售有多少次」。注意到在最开始的默认值字段设置里面指汇总方式的求和表示的是「销售量的求和」，是对于字段的数值的求和。那么对于销售量为2的产品进行统计的时候就会得到要求答案的2倍。将值字段设置为「计数」就是数值出现的次数的求和，满足要求。要得到每一种销量占有所有销量的比，在值字段中「值显示方式」修改为列汇总百分比。这里可以指导一个大包装里面装入的最优的数量。

### 2.2.3 计算百分比

目标；统计某一个范围内数据的数目占比。在query里面直接对于第一关键字（客户）和第二关键字（订单id）相同的订单直接进行合并（对于合并字段右键求和）。excel对于列进行排序（右键或者上面的「栏位」之中）。对于所选的一块创建组就是直接拉一个组（区别与点一个然后创建组）。

在得出了柱形图之后，在同时有两种不同单位的数据的时候，可以使用副轴来作为一种单位的轴。可以通过鼠标来电或者用键盘上下去选择比教不容易用鼠标选中的图。右键或者在上面数据透视工具「数据系列格式」「系列选项」中可以设置在副轴上。通过样在系列选项里面可以设计柱形图的间隔。对于坐标轴的显示范围和格式也可以右键「设置坐标轴格式」「坐标轴选项」里面进行设置

用散点图可以使选中的第一列为x坐标，其余的列都是y轴。

## 2.3 D函数 模拟运算表 控件

### 2.3.1 模拟运算表

模拟运算表在 数据-数据工具-模拟分析。拉出一块有一列为自变量，另一列为应变量（最上面额外以个值用于指定在旁边的函数定义），然后将应用列的单元格为函数定义中选择自变量对应的符号定义的例值（也是通过在excel表中的位置来关联的）。模拟运算表中的单格内容无法被编辑修改，只能被引用到其他的位置。如果自变量和应变量放在两行，则是应用列的单元格函数定义。如果是二元函数，行列分别设置为自变量和应变量，然后函数值放在行列相交的地方。

### 2.3.2 D函数

D函数用于数据库操作。D函数的一般形式：

D函数名称(数据列表,汇总字段, 条件区域)

在罗列数据列表的时候，字段名要注意在里面。数据区域可以自定义命名：左上角的编辑栏。

D函数结合模拟运算表：函数在excel里面被定义在每一个单元格上。应用的时候就要指定是引用带有函数表达式的单元格作为函数定义的参考。在D函数中，变量为条件区域的时候，那么行和列的标签就应该指向这些条件自变量。在构建模拟数据表的时候，可以使用 数据高级筛选 利用空的条件加上不显示相同的值就可以贴出来列自变量。行自变量可以通过提取到列后通过复制时 开始复制选项中 选择转置。

D函数给出汇总的具体内容，然后通过模拟数据表罗列出所有的汇总的范围的变量然后将它们作为D函数条件区域的变量得到一个汇总。如果列出的变量没有出现在原表中的时候，会默认认为是现在计算的销售额的那个变量。通过对于模拟数据表进行值复制就可以对它进行操作如排序。对于汇总值进行季度分类，要用到一下函数

left函数：从左往右截取x个函数

left(文本，字符数目)

mid函数:从左往右数到x位置截取y个字符

mid(文本, 开始位置， 截取字符数)

这两个函数可以从字符串中提取信息，如从「1996年3季度」中提取出1996和3。

date函数:将字符转为日期格式

date(年, 月, 日) 对于每一个月的最后一天，由于最后一天会随着月份不同，可以用下个月的第一天-1来得到。

通过以上函数可以设置函数条件：

"<=" & date(x, y, z)

">" & date(x, y, z)

x, y, z 就是通过给出的文字提取的年月来计算具体的季度上下限。

### 2.3.3 控件

控件就是下拉按钮(☐) 在文件-选项-自定义功能区=主选项卡-勾选开发工具。对于需要的单元格做控件，用表单控件中的「列表框控件」。在随便一个位置画好框框以后，右键可以设置控件格式设置数据源区域和单元格连接。点击控件里面的选项会在单元格链接中显示序号。

Index 函数：在矩形的数据区域中，根据行号和列号取出行列交叉点数据。

Index(数据区域, 行号, 列号)

通过将行号或者列号指向控件的单元格链接，就可以通过控件修改对应单元格链接的值来修改单元格的显示内容。对于修改条件区域的控件，可能会出现将「全部」转为条件区域中的空来代表全部的集合。这里要用到IF 函数

if(条件, 条件T的表达式, 条件F的表达式)

对应例11, 12; 自己回家做.cpp (☐) 在文件-选项-自定义功能区=主选项卡-勾选开发工具。

### 2.3.4 3-例13的一些简单说明

目标：根据在控件之中所选择的总销售额前10的公司来展示一张该公司不同季度的图

第一步：调入数据

第二步：构造d函数，通过模拟运算表做出所有公司的销售额，并进行排序，选出前10。

第三步：在某个公司为和某个季度间隔中为条件下的销售额汇总。通过d函数的条件区域来设定以上条件，最终通过汇总的模拟运算表得到折线图。注意，季度的条件是由有一个由函数得出的单元格得出的，而这个单元格的函数又上溯于一个原始条件。在使用模拟运算表汇总的时候注意要追溯到原始条件来设定变量。

第四步：通过控件来控制能够得出季度汇总d函数的公司条件。如上面所讲通过链接单元格能够加上index函数来满足以上要求。

一些扯白：更加复杂的汇总都可以像第三步一样明确条件之后再考虑具体的d函数设计，特别是条件本身的设计方式决定了模拟运算表选择列/行的单元格的例子条件。控件可以用于条件的控制。