

## SESSION 4: Conditional Statements

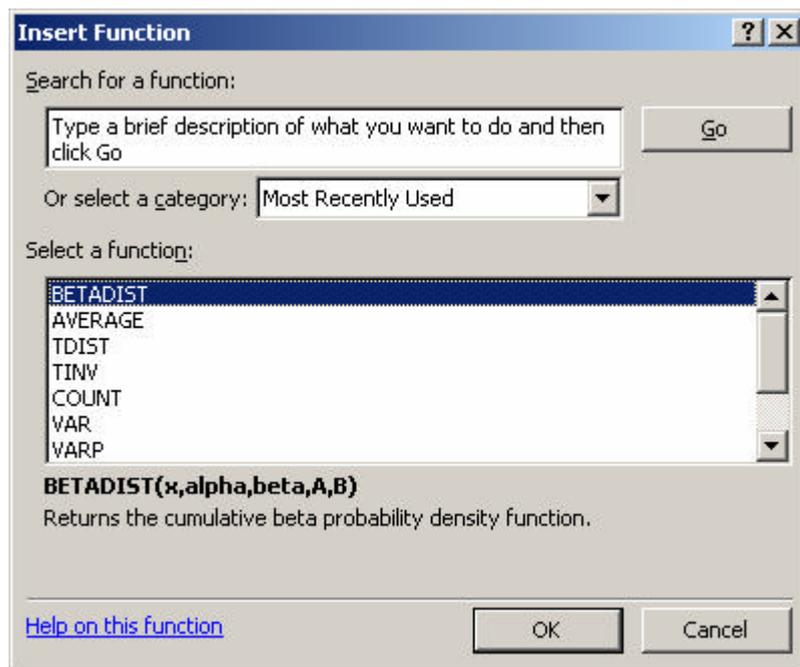
### Using the Function Wizard

**Excel** provides us with a function wizard that guides us through entering complex functions.

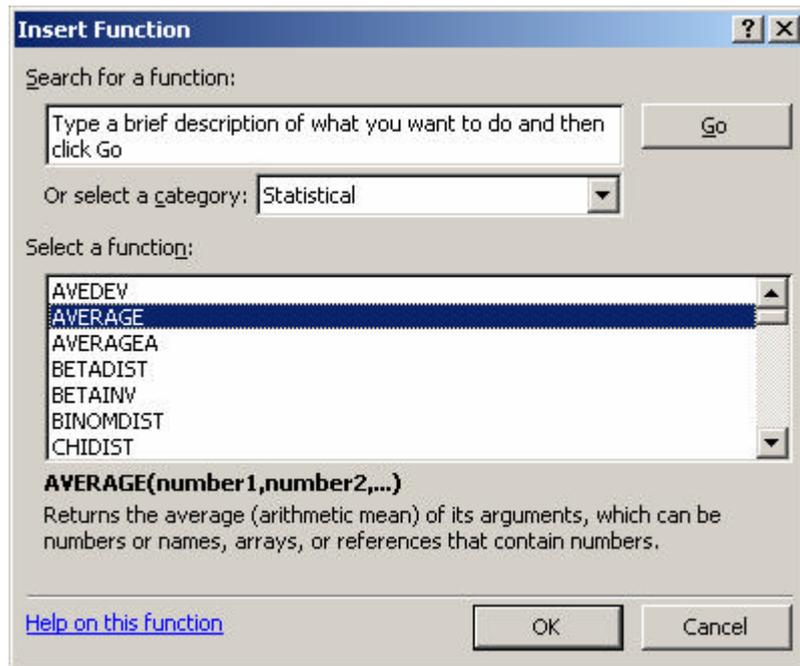
To demonstrate this, we will make use of the file **sample1.xls**. The function wizard is opened by clicking on the paste function, .

You have to remember to click on the destination cell before opening the function wizard. Once you have chosen this cell, click on the paste function. The layout of the form has changed a bit from the 2000 version, however the content of the form is still the same.

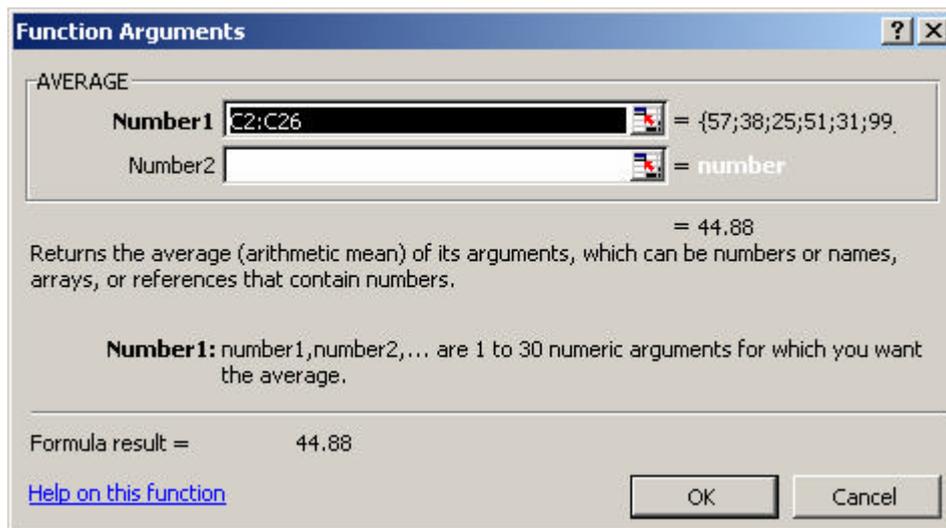
Suppose we wanted to find the average age of the sample. Click on **C27** and then press the paste function. The following appears.



You can search for a function by typing what you want to achieve at the end, or if you know which function to use, you can select the category, and then choose the function. Note that when you change the category, the list of functions changes as well. Click on **Statistical** > **Average**. You can see that **Excel** provides a quick help of what the arguments should be, and of what the function returns.



Press **OK**.



**Excel** tries to be smart and gives you a list of what it thinks your chosen data is. If this is correct, then press **OK**, if not then you have to choose the correct range of data. In this case, the data chosen is correct, so press **OK**. The average is 45.

If you look again at the range chosen, you will see that the value of 99 has been picked inside the range. Be careful as this is missing data. **Excel** has no way of determining that this number is in fact a missing number, so we have to define it ourselves. Click on the cell, and press the **Delete** key. You can see that **Excel** has changed the calculated average to 43 without the need of refreshing the formula.

24	1	44
25	1	45
Average		43

The paste function allows us to use more complex statistical functions, such as finding the t-values, correlations, f-values, etc.

## Using the 'IF' Function

This command is essentially a computational command with a conditional statement in front of it.

The **If** function returns one value if a condition you specify evaluates to TRUE and another value if it evaluates to FALSE. Use IF to conduct conditional tests on values and formulas.

### IF(conditional expression, value if true, value if false)

You can use **conditional expressions** to apply transformations to a selected subset of cases.

Let's suppose we want to check who is older than 50, and who is not. Those older than 50 will be issued with an advantage card, while those younger will not obtain one.

C
RAGE
57
38
25
51
31
71
38
37
23
35
71
33
44
39
39
31
51
51
51
68
29
21
44
45

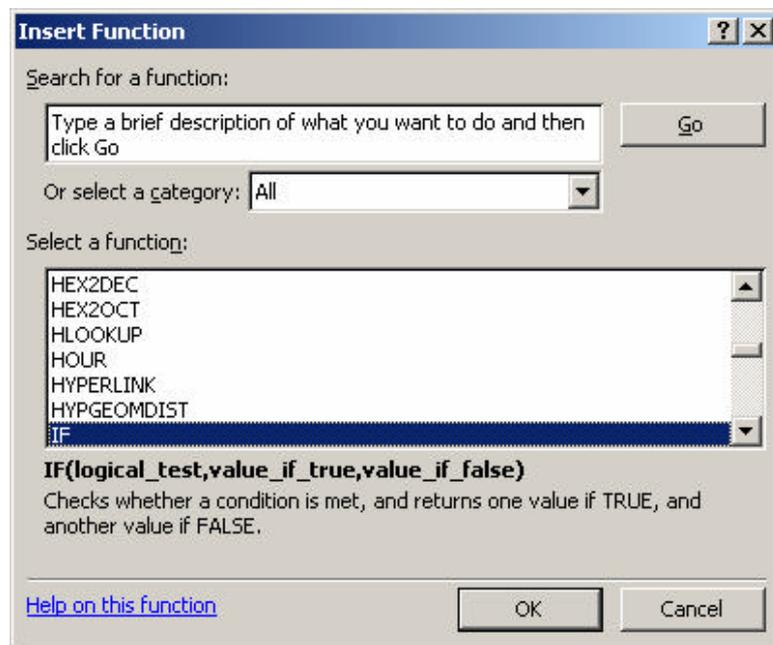
We would need to create a new column and put a **Yes** if the person is going to get an advantage card and **No** otherwise.

This can be re-expressed in terms of an **IF** statement as follows:

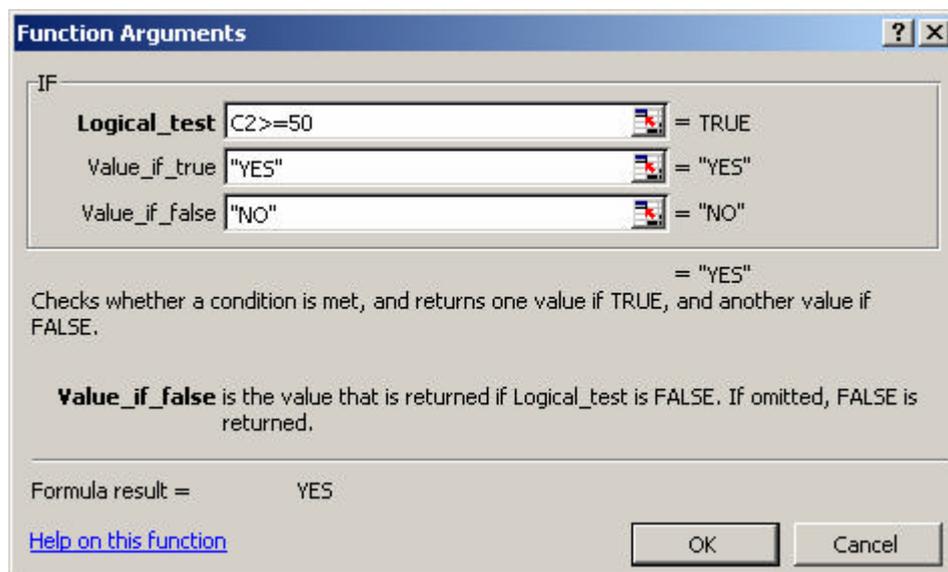
**IF(Age>=50,"Yes","No")**

Note that since we want **Excel** to output text, we have to type this inside double quotes.

We can use the paste function to write down the conditional statement. Click on the paste function, choose **All** and use the scroll buttons to find **IF**.



Press **OK**.



Fill the 1<sup>st</sup> edit box with the condition that you want to check on. Remember that you can only check one case at a time, so insert the 1<sup>st</sup> case, which is C2. In the 2<sup>nd</sup> edit box, write down the text that you would obtain if the condition is accepted, and in the last box write down the text that you would obtain if the condition was false. **Excel** will give you a preview of the result.

Press **OK**, and **Excel** will print **YES** on cell **G2**. Drag the pointer to copy this formula for all the cases. You will then see that where we had deleted the 99, then **Excel** just outputted a **NO**.

## Relational and Logical Operators and Functions

In the example above I used the relational operator **>=**; there are several other operators that we need to use when we are dealing with data.

### 1. Relational operators

=	Equal to
<	less than
>	greater than
<>	not equal to
<=	less than or equal to
>=	greater than or equal to

In addition you can join conditional expressions using logical operators.

### 2. Logical operators

AND
OR
NOT

### 3. Example using both relational and logical operators

Let us take an example from **bsas91.xls**. Suppose we have the condition

**IF (RAGE>=65 AND RSEX = 1) OR (RAGE>=60 AND RSEX =2)  
PENSION=1**

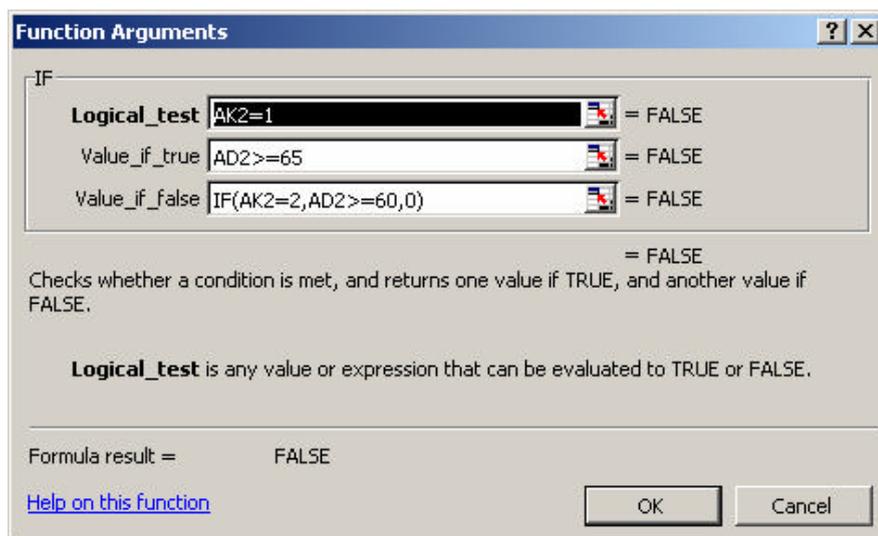
i.e. if the respondent is aged 65 or over and is male, or if the respondent is aged 60 or over and is female, set the variable **PENSION** to 1. If the respondent is not of pension able age, **PENSION** is set to 0.

The above example is NOT as easy to do in **Excel** as in other packages, because **Excel** does not allow the concatenation of expression by **AND** or by **OR** directly. We therefore have to make use of nested **IF** statements.

We have to check first on whether the age is greater or equal to 65. If it is, we have to check the 2<sup>nd</sup> condition, i.e. if RSEX =1. It would be easier to check on the gender first, and then depending on the gender check on the age. So if we look at the following table:

RSEX=1	RAGE>=65	1
RSEX=1	RAGE<65	0
RSEX=2	RAGE>=60	1
RSEX=2	RAGE<60	0

Therefore in **Excel**, create a new column called pension and click on the paste function. Choose the **IF** function.



This form needs some explanation. Basically, we are first checking to see whether the gender is male. If this is male, then we would then check on the age if it is greater than 65 or not. If it is, then **Excel** will output **TRUE**, else it will output **FALSE**. (These could be modified to 1 and 0 respectively).

If the gender is not male, then **Excel** will try to evaluate the 2<sup>nd</sup> expression. We have created a nested **IF** condition here. If the gender is not 1 but it is 2, then check if the age is greater than 60, and this will output a **TRUE** if both are satisfied. If not, it will output **FALSE**.

The total **IF** condition is

```
=IF(AK2=1,AD2>=65,IF(AK2=2,AD2>=60,0))
```

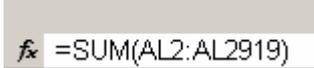
You can change the **TRUE** and **FALSE** to **1** and **0** although the **IF** command would be more complicated.

=IF(AK2=1,IF(AD2>=65,1,0),IF(AK2=2,IF(AD2>=60,1,0),0))



*f<sub>x</sub>* =IF(AK2=1,IF(AD2>=65,1,0),IF(AK2=2,IF(AD2>=60,1,0),0))

If we wanted a count of all the people eligible for pension, then click on the last cell between the new column, and sum all the values (if you have changed the cells to contain a **1** instead of **TRUE** and a **0** instead of **FALSE**).



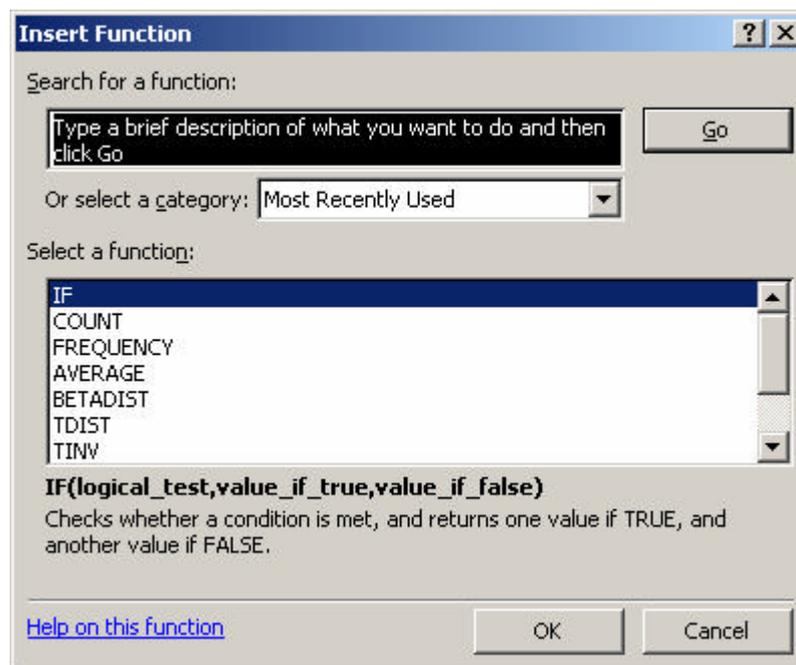
*f<sub>x</sub>* =SUM(AL2:AL2919)

The total number is **792**.

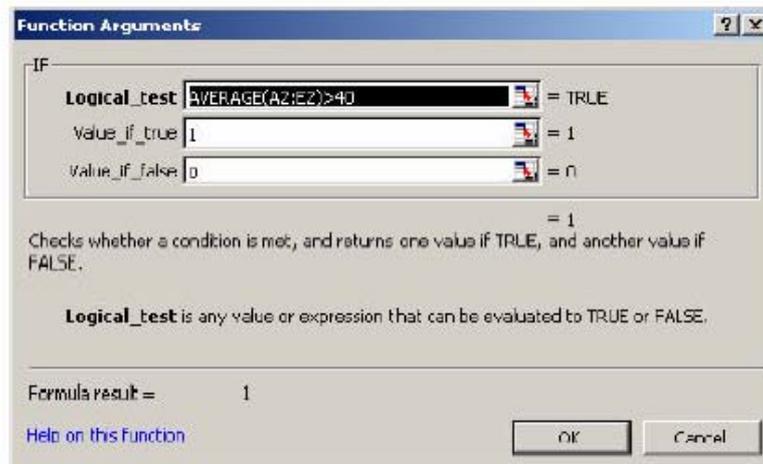
#### 4. Example using functions

Open the file **marks.xls**. This data set contains marks of 100 students for 5 examinations. A student needs to have a mean mark of 40% over a set of five exam papers to pass.

We are again going to use the **IF** command, combined with the **average** function. Click on the paste function. If you click on Most Recently Used, you will find that Excel automatically brings up the **IF** function, and all the last commands that you have used.



Press **OK**.



Type the logical statement in the 1<sup>st</sup> edit box. So we are asking **Excel** to find the average and to compare this result with 40. If this average is greater than 40, then the student will pass, else the student fails. Press **OK** and copy the formula down to the last case.