EXCEL WORKSHOP I: INTRODUCTION TO EXCEL

With Applications from Ratemaking

BRUIN ACTUARIAL SOCIETY

IF(logical_test, [value_if_true], [value_if_false])

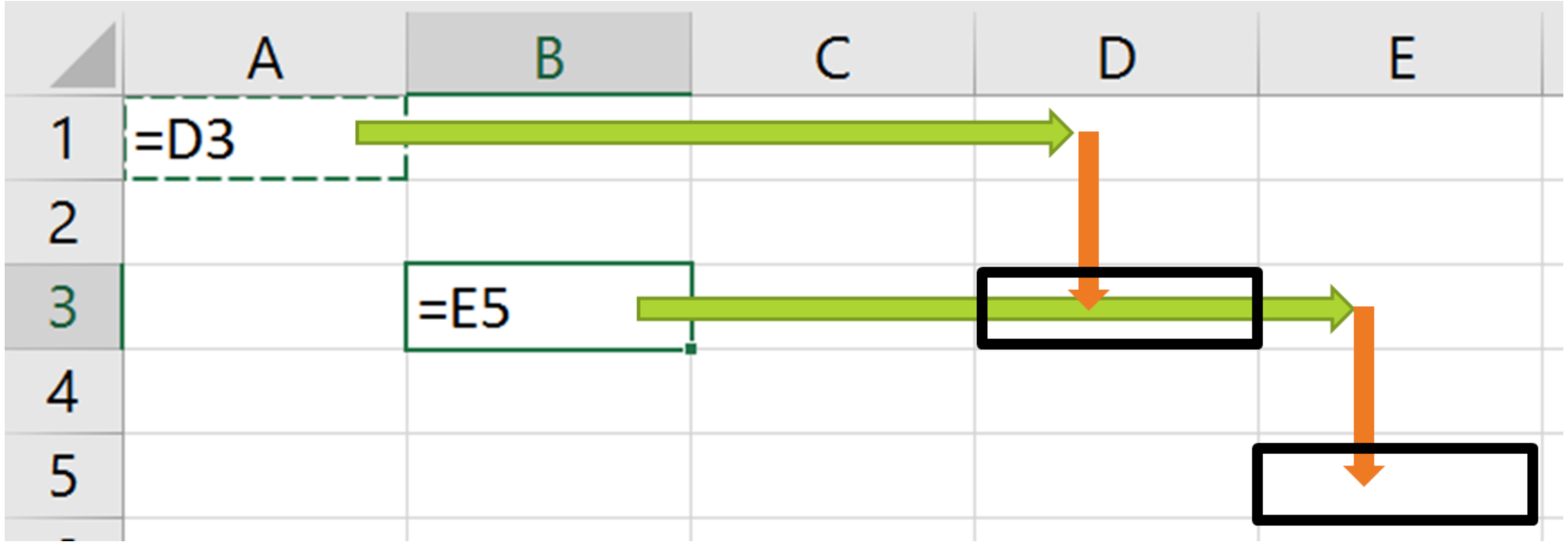
	A	B	C
1		**Formula**	**Output**
2	100	=IF(A2>73,"Pass", "Fail")	Pass
3	80	=IF(A3>73,"Pass", "Fail")	Pass
4	60	=IF(A4>73,"Pass", "Fail")	Fail
5	40	=IF(A5>73,"Pass", "Fail")	Fail
6	20	=IF(A6>73,"Pass", "Fail")	Fail

- If **[value_if_true]** or **[value_if_false]** are not given, they default to TRUE and FALSE
- If **logical_test** is true, output is **[value_if_true]**
- If **logical_test** is false, output is **[value_if_false]**
- IF statements can be nested

BAS.00
BRUIN ACTUARIAL SOCIETY

Relative Cell References:

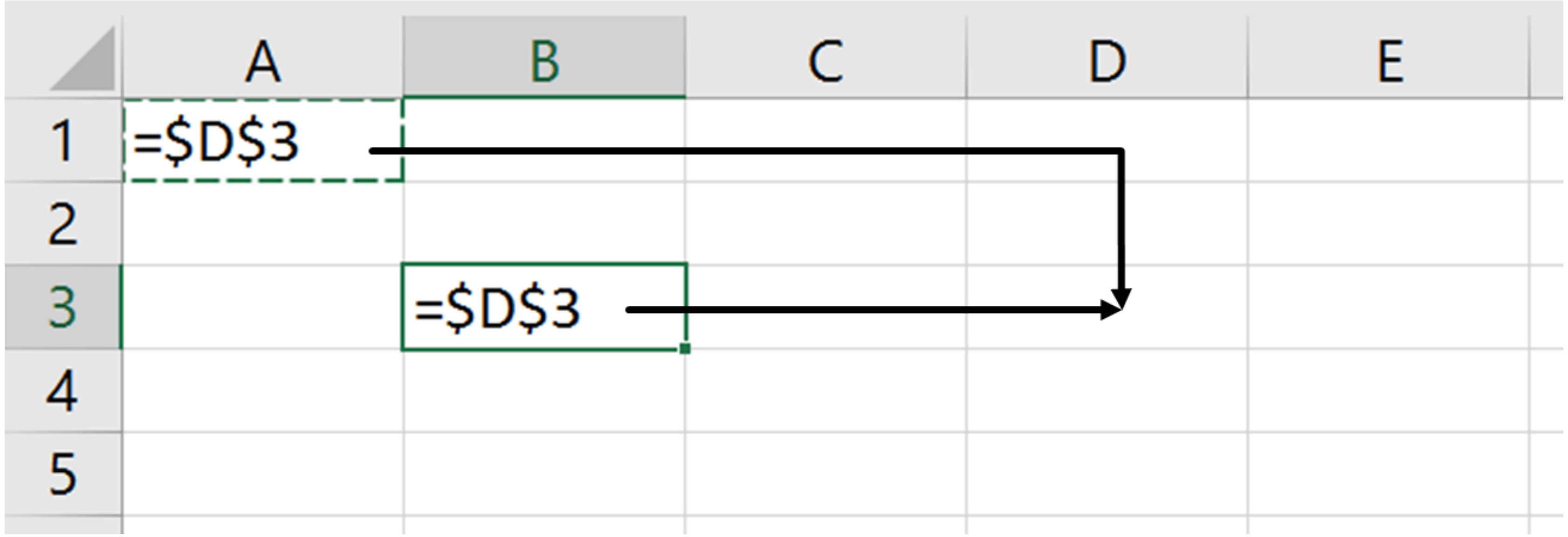
Suppose we copy the formula in cell A1 and paste it in cell B3.



BAS.OO
BRUIN ACTUARIAL SOCIETY

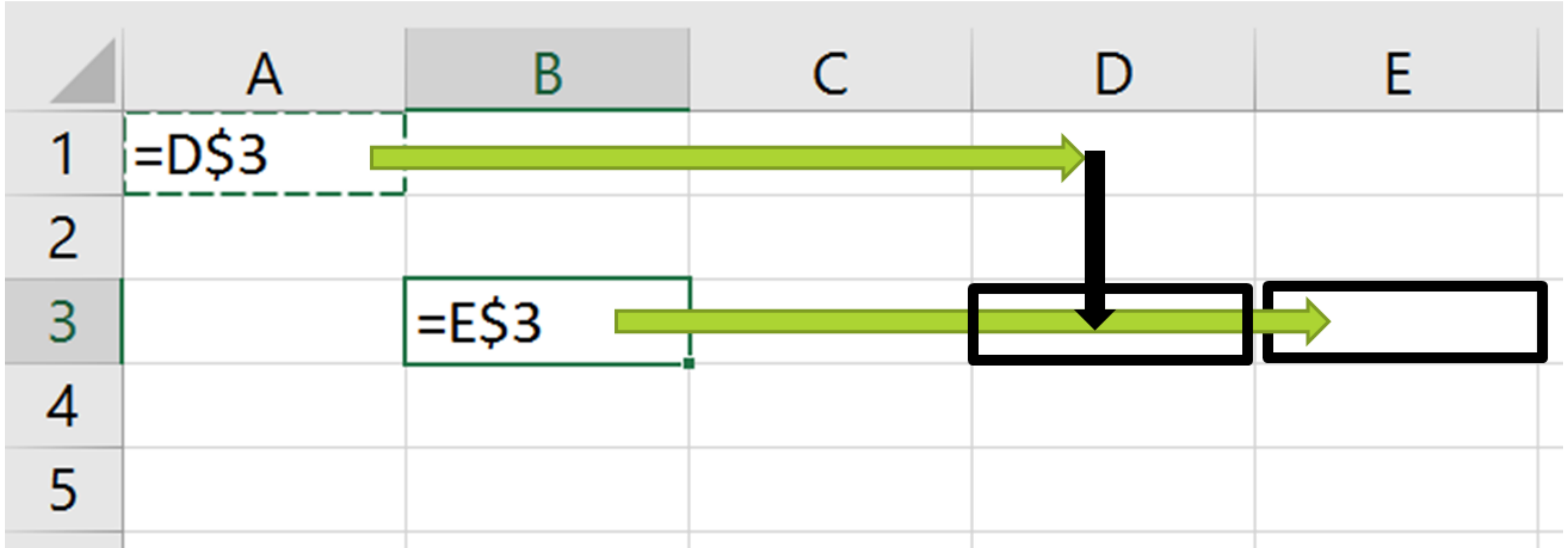
Absolute Cell References:

Suppose we copy the formula in cell A1 and paste it in cell B3.



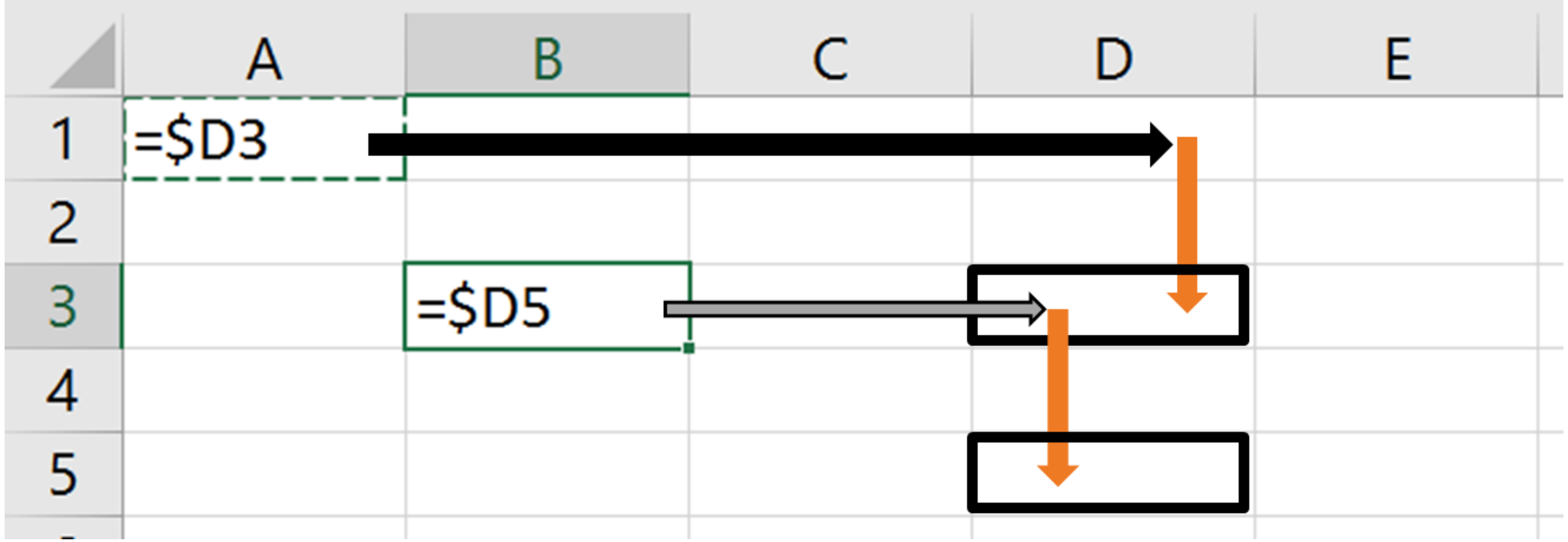
BRUIN ACTUARIAL SOCIETY

Mixed Cell References:
Suppose we copy the formula in cell A1 and paste it in cell B3.



Here, the black arrow does not indicate a relative shift, but an **absolute** shift to row 3.

Mixed Cell References:
Suppose we copy the formula in cell A1 and paste it in cell B3.



Here, the black/gray arrows do not indicate relative shifts, but **absolute** shifts to column D.

BRUIN ACTUARIAL SOCIETY

SUM(number1, [number2], …)
PRODUCT(number1, [number2], …)

	A	B	C	D
1	1		**Formula**	**Output**
2	2		=SUM(A1,A2,A3,A4,A5)	15
3	3		=SUM(A1:A5)	15
4	4		=PRODUCT(A1:A3, A5)	30
5	5			

- Each of the **number**s can be either a reference to a single cell or to a range

BRUIN ACTUARIAL SOCIETY

VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

- Looks for **lookup_value** in the first column of **table_array**
- Returns the value in the column number indicated by **col_index_num**
- **[range_lookup]**: Default is TRUE (approximate match). FALSE means you want an exact match.
  - For approximate match, data must be sorted in ascending order.

VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

	A	B	C	D	E
1	*Name*	*Student ID*		**Formula**	**Output**
2	Joe Bruin	272577814		=VLOOKUP("Joe Bruin",$A$2:$B$4,2,FALSE)	272577814
3	John Doe	644907243		=VLOOKUP("Jane Doe",$A$2:$B$4,2,FALSE)	534763964
4	Jane Doe	534763964		=VLOOKUP("Johnny",$A$2:$B$4,2,FALSE)	#N/A

BRUIN ACTUARIAL SOCIETY

VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

	A	B	C	D	E	F
1	*Percent*	*Grade*			**Formula**	**Output**
2	0	F		52	=VLOOKUP(D2,$A$2:$B$6,2,TRUE)	F
3	60	D		85	=VLOOKUP(D3,$A$2:$B$6,2,TRUE)	B
4	70	C		92	=VLOOKUP(D4,$A$2:$B$6,2,TRUE)	A
5	80	B		63	=VLOOKUP(D5,$A$2:$B$6,2,TRUE)	D
6	90	A		75	=VLOOKUP(D6,$A$2:$B$6,2,TRUE)	C

COUNTIF(range, criteria)
SUMIF(range, criteria, [sum_range])

- COUNTIF counts all cells in **range** that satisfy **criteria**.
  - The **criteria** can be given as a value, in which case the values in **range** must be equal to that value.
  - It can also be given as a boolean expression in quotes.
- SUMIF is similar, but it sums the values in **range** instead of counting them.
  - Or, it sums across a different **[sum_range]**, using the values in corresponding **range**-**criteria** pairs.

BRUIN ACTUARIAL SOCIETY

COUNTIF(range, criteria)
SUMIF(range, criteria, [sum_range])

	A	B	C	D
1	1		**Formula**	**Output**
2	2		=COUNTIF($A$1:$A$7,3)	0
3	4		=COUNTIF($A$1:$A$7,2)	1
4	8		=COUNTIF($A$1:$A$7,">4")	4
5	16		=COUNTIF($A$1:$A$7,"<>4")	6
6	32			
7	64			

COUNTIF(range, criteria)
SUMIF(range, criteria, [sum_range])

	A	B	C	D	E
1	A	1		**Formula**	**Output**
2	B	2		=SUMIF($B$1:$B$7,"<10")	15
3	C	4		=SUMIF($A$1:$A$7,"A",$B$1:$B$7)	73
4	A	8		=SUMIF($A$1:$A$7,"<>B",$B$1:$B$7)	109
5	B	16			
6	C	32			
7	A	64			

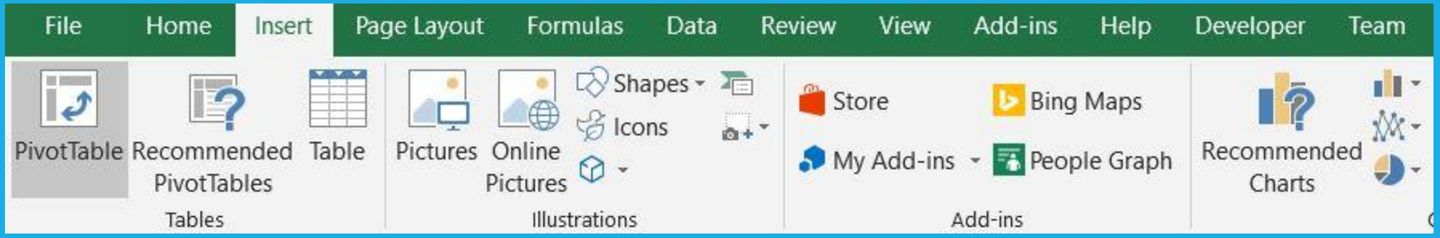
# COUNTIFS(criteria_range1, criteria1,…)
SUMIFS(sum_range, criteria_range1, criteria1,…)

- These do the same thing as COUNTIF and SUMIF, but you can specify multiple criteria.
- Note that for SUMIFS, **sum_range** comes first.

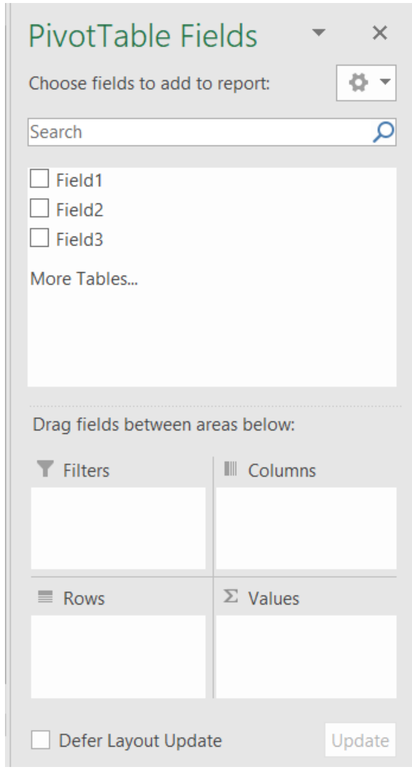
COUNTIFS(criteria_range1, criteria1,…)
SUMIFS(sum_range, criteria_range1, criteria1,…)

	A	B	C	D	E	F
1	Gender	Age	Number of Claims		**Formula**	**Output**
2	M	24	0		=SUMIFS(C2:C14,A2:A14,"F",B2:B14,">30")	1
3	M	37	1			
4	M	38	0			
5	M	31	2			
6	M	28	2			
7	F	36	0			
8	F	25	1			
9	M	60	0			
10	F	27	1			
11	F	61	0			
12	F	35	1			
13	M	32	1			
14	M	54	2			

PivotTables



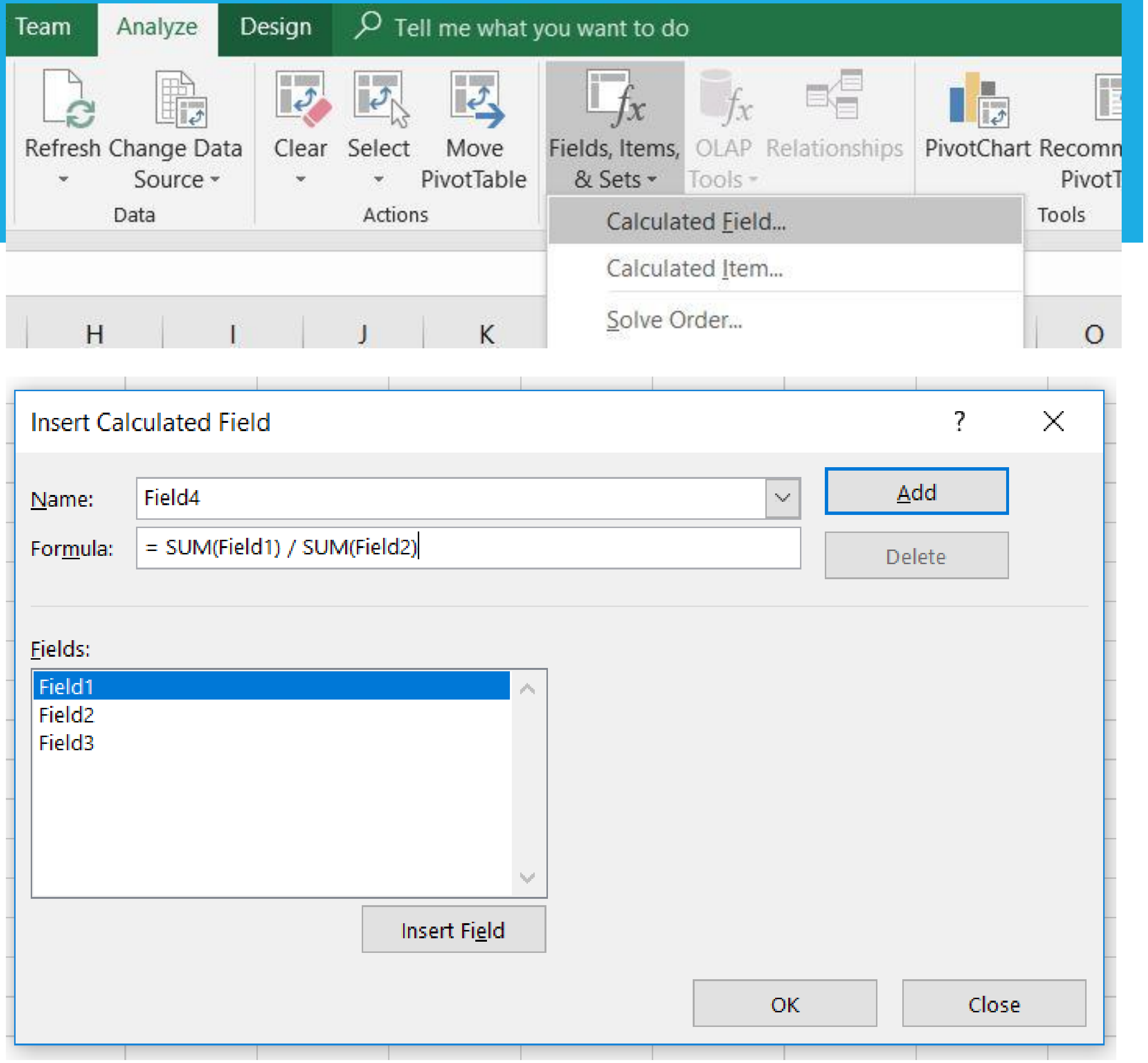
- PivotTables can be used to quickly summarize data
- Click and drag different fields to “columns”, “rows”, or “values”, depending on what you want
- You can also filter data



BRUIN ACTUARIAL SOCIETY

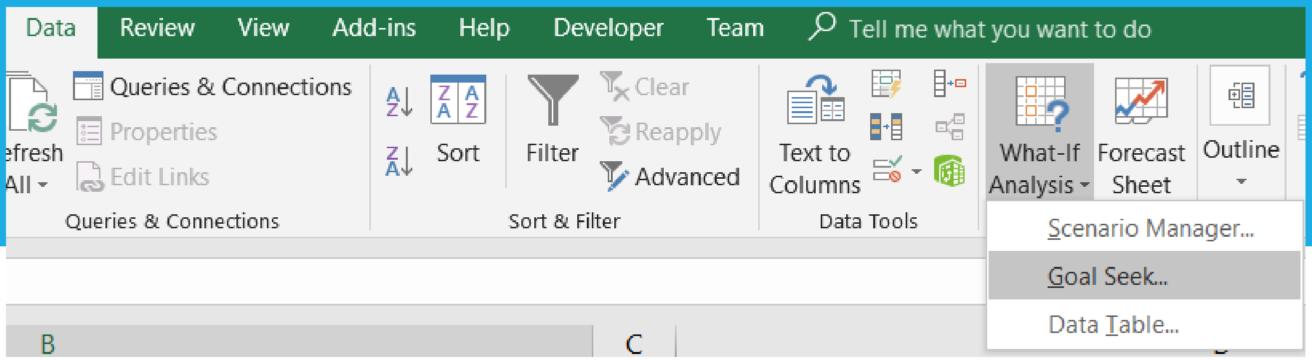
Calculated Fields

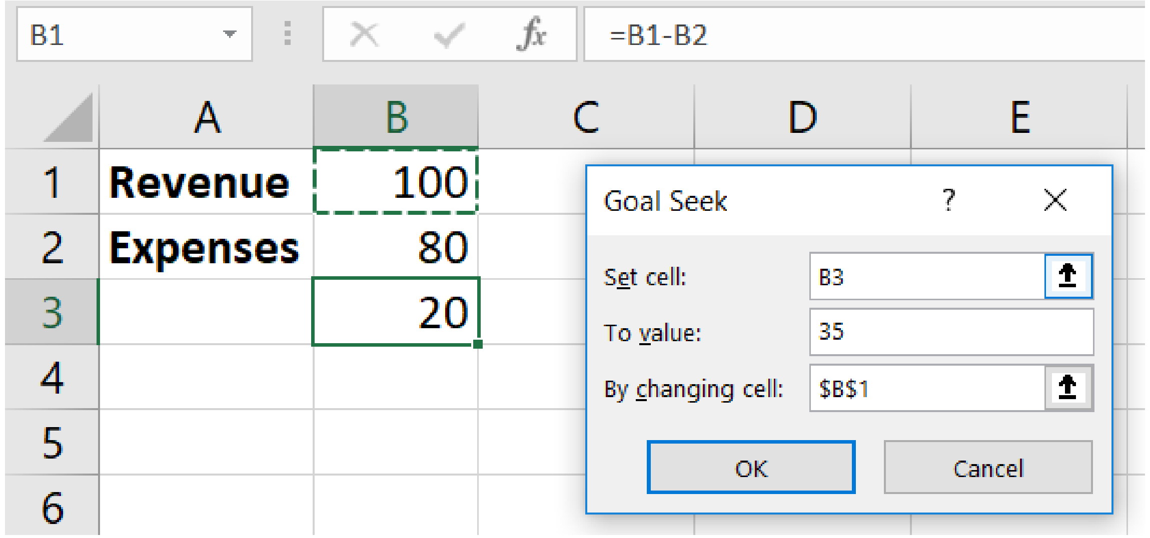
- Calculated fields are useful for aggregate-level calculations
- For example, we could look at the value of Field4 for each Field3

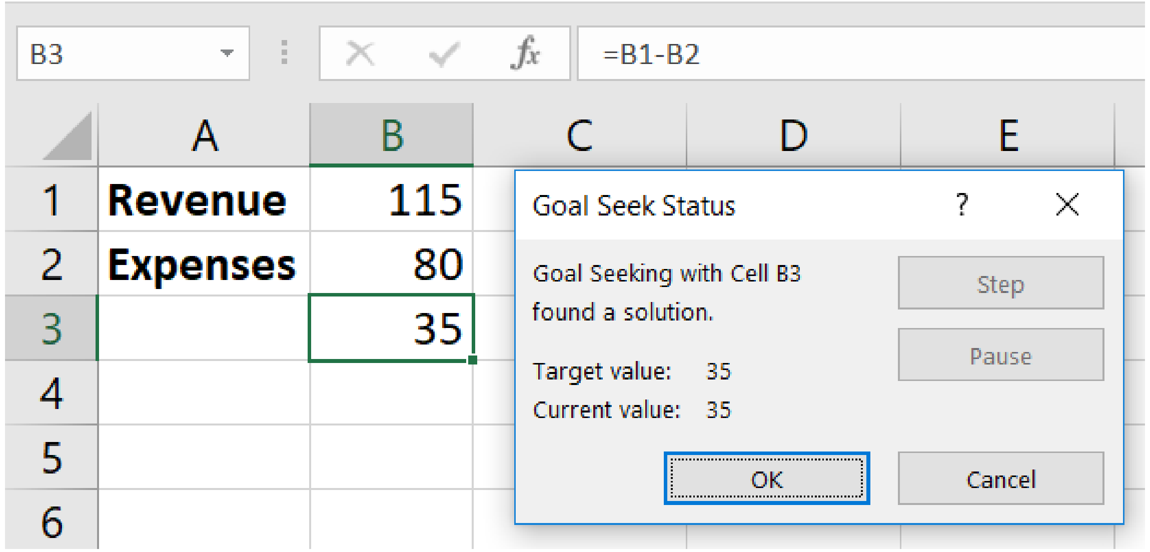


BRUIN ACTUARIAL SOCIETY

GoalSeek







BRUIN ACTUARIAL SOCIETY

APPENDIX

Background on the problem and mathematical justification

BRUIN ACTUARIAL SOCIETY

Background

- You are a pricing actuary for Bruin Auto Liability Insurance, LLC, a Los Angeles based company insuring personal automobile casualty losses in Southern California.
- Due to limited data collected, BAL Insurance only uses two variables in their pricing algorithm: driving frequency and county. Assume these two variables are independent.
- You have been tasked with analyzing the premiums charged and claims filed to determine if the rates are adequate.

Background

- It is January 1st, 2019. Through your discussions, you discover that sufficiently many claims for a given calendar year will be reported in 3 years.
- Therefore, you have enough reliable data to analyze 2014 and 2015. However, volume of claims and policy in 2014 are fairly low.
- Thus, you are to conduct an analysis of premiums and claims for policy year 2015.
- For your convenience, the relevant policy data and claims data have been extracted from the full data tables.

BAS.00
BRUIN ACTUARIAL SOCIETY

Background

- Management's target loss ratio is 85%.
- Conduct an analysis of 2015 data and, based on your results, recommend an alternative rating algorithm to use in the future, if claims in 2016-2018 are similar to those in 2015.
- Backtest your new rating algorithm on data from 2014.
- Determine whether the rates you have set are feasible on the data from policy year 2014, and state the overall loss ratio on that sheet.

Adjusting the loss ratios

- For simplicity, let's only look at the relativities for **county**, though the process for driving frequency is the same.
- Specifically, let's look at Los Angeles. We decompose the premium collected from all LA policyholders into the base rate and the factor:

$$\mathrm{LR}_{\mathrm{LA}} = \frac{\sum \mathrm{Claims}_{\mathrm{LA}}}{\sum \mathrm{Premium}_{\mathrm{LA}}} = \frac{\sum \mathrm{Claims}_{\mathrm{LA}}}{\mathrm{BaseRate} \cdot \sum \mathrm{Factor}_{\mathrm{LA}}}$$

- Note that we cannot simply replace $\mathrm{Factor}_{\mathrm{LA}}$ with 1.20, since Factor refers to the product of county and driving frequency relativities.

BAS.00
BRUIN ACTUARIAL SOCIETY

Adjusting the loss ratios

- Let $\mathrm{LR_T}$ be the target loss ratio. Then, multiplying our previous equation by $\frac{\mathrm{LR_T}}{\mathrm{LR_{LA}}}$,

$$\mathrm{LR_{LA}} \cdot \frac{\mathrm{LR_T}}{\mathrm{LR_{LA}}} = \frac{\sum \mathrm{Claims_{LA}}}{\mathrm{BaseRate} \cdot \sum \mathrm{Factor_{LA}}} \cdot \frac{\mathrm{LR_T}}{\mathrm{LR_{LA}}}$$

- Or, reordering:

$$\mathrm{LR_T} = \frac{\sum \mathrm{Claims_{LA}}}{\mathrm{BaseRate} \cdot \frac{\mathrm{LR_{LA}}}{\mathrm{LR_T}} \cdot \sum \mathrm{Factor_{LA}}}$$

Adjusting the loss ratios

- The key difference between this equation is the additional component of $\frac{\mathrm{LR_{LA}}}{\mathrm{LR_T}}$.

$$\mathrm{LR_{LA}} = \frac{\sum \mathrm{Claims_{LA}}}{\mathrm{BaseRate} \cdot \sum \mathrm{Factor_{LA}}}$$

$$\mathrm{LR_T} = \frac{\sum \mathrm{Claims_{LA}}}{\mathrm{BaseRate} \cdot \frac{\mathrm{LR_{LA}}}{\mathrm{LR_T}} \cdot \sum \mathrm{Factor_{LA}}} = \frac{\sum \mathrm{Claims_{LA}}}{\mathrm{BaseRate} \cdot \sum \left(\frac{\mathrm{LR_{LA}}}{\mathrm{LR_T}} \cdot \mathrm{Factor_{LA}} \right)}$$

- To adjust the loss ratio in LA to be equivalent to the weighted-average loss ratio, we must multiply the current factor by $\frac{\mathrm{LR_{LA}}}{\mathrm{LR_T}}$.

Adjusting the loss ratios

$$LR_{LA} = \frac{\sum Claims_{LA}}{BaseRate \cdot \sum Factor_{LA}}$$

$$LR_T = \frac{\sum Claims_{LA}}{BaseRate \cdot \sum \left(\frac{LR_{LA}}{LR_T} \cdot Factor_{LA} \right)}$$

- We can choose any target loss ratio we want, but we will use the weighted-average loss ratio in this example.
- Regardless of our choice, our final premiums will be the same after base rate modifications.