

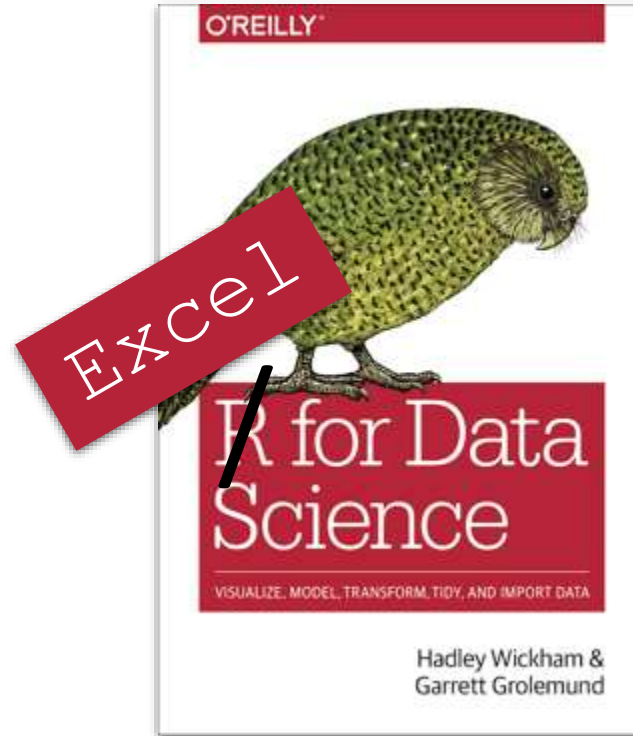
When to stop using Excel and start using R

Nathan Stephens

February 2022

Why do we use Excel?

- Wrangle
- Visualize
- Transform
- Analyze
- Communicate insights about your data

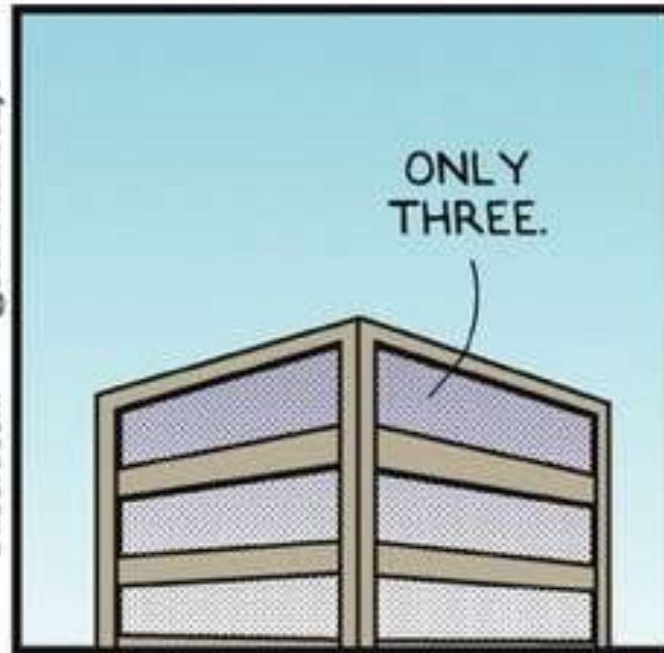


[A few] Problems using Excel

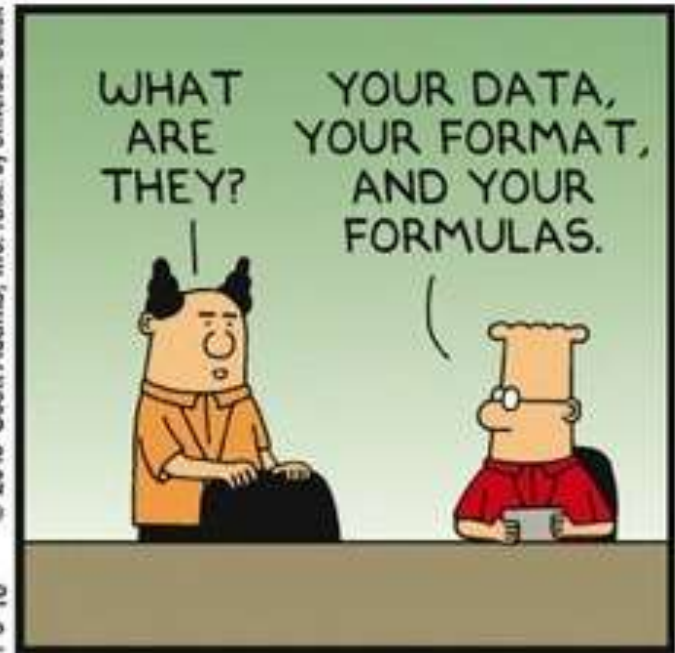
Wednesday January 06, 2016 *Three Problems With Spreadsheet*



Dilbert.com @ScottAdamsSays

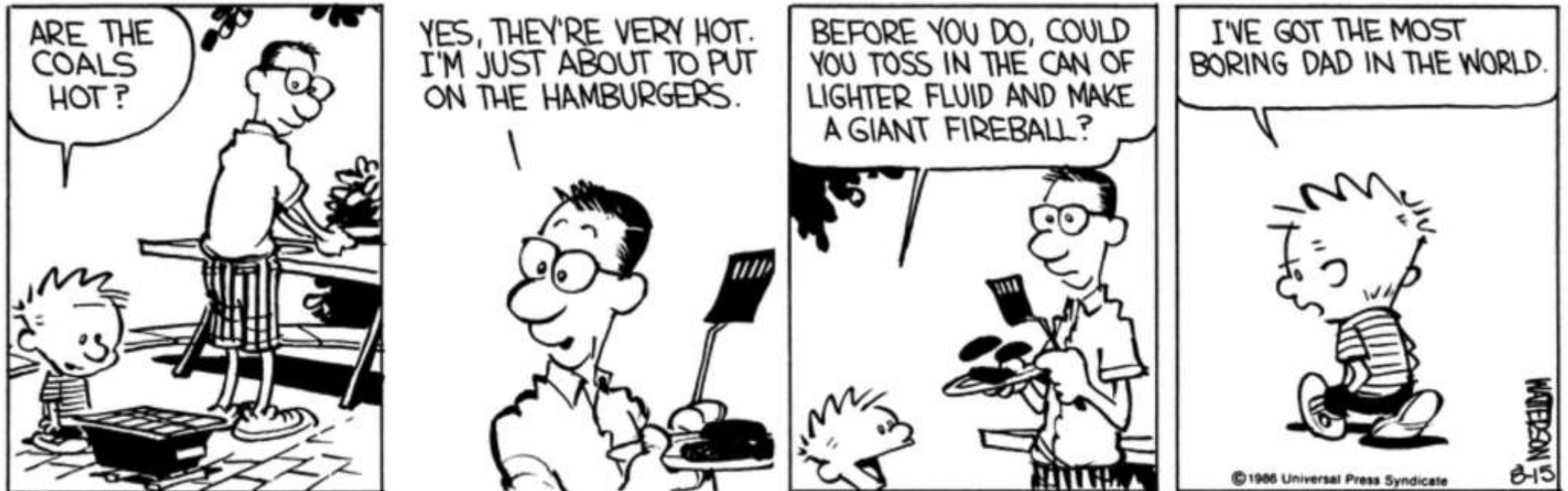


© 2016 Scott Adams, Inc. /Dist. by Universal Uclick 1-6-16



Problem 1: Scaling your data

Can be like playing with fire



Hard limits in Excel

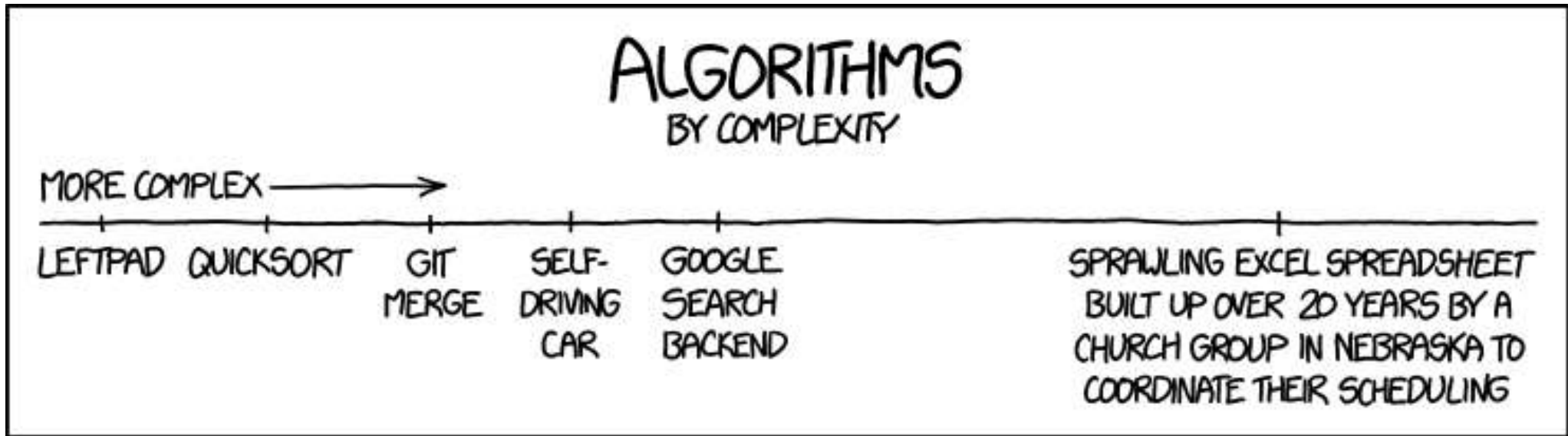
- Maximum rows: 1,048,576
- Maximum columns: 16,384



Memory limitations


1's Mb	10's Mb	100's Mb	1's Gb	10's Gb	100's Gb
--------	---------	----------	--------	---------	----------

Problem 2: Scaling your code




Excel is Turing Complete



- A system that is able to recognize or decide other data-manipulation rule sets.
- In principle -- **although often not in practice** -- it can be used to solve any computation problem.




A screenshot of a tweet from Satya Nadella (@satyanadella) dated February 9, 2021. The tweet states that Excel formulas are now Turing-complete. It includes a retweet from Microsoft Excel (@msexcel) which explains the LAMBDA function. The retweet features a diagram showing a lambda symbol in a green arrow pointing to the text 'custom formulas without code'. Below this, it shows the formula '(x, y, x+y)' with 'x' and 'y' underlined and labeled 'Input names', and 'x+y' underlined and labeled 'Calculation'. An arrow points from the formula to a green box containing the number '4'. Below the formula, it shows '=myLambda (1, 3)'.


Satya Nadella  @satyanadella

Excel formulas, the world's most popular programming language, is now Turing-complete. Go check it out!

 **Microsoft Excel**  @msexcel

Excel keeps evolving to give users even more. Now, with the power of LAMBDA, you can write your own reusable functions with the Excel formula language. See how we're transforming Excel: msft.it/6014pF2Oa

 **custom formulas without code**

(x, y, x+y)  **4**

Input names Calculation =myLambda (1, 3)

4:54 PM · Feb 9, 2021

7.6K Reply Share

[Read 191 replies](#)

Code complexity

Complexity



Example	Simple Workbook	Advanced Workbook
Data sources	Single	Multiple
Analytics	Summary statistics	Advanced algorithms
Data updates	Static or none	Dynamic
Formulas and Functions	One or two arguments	Multi-line and nested
Advanced tools	Single drop-down list or none	Scripts
Pivot Tables and Charts	One or two	Several
Spreadsheets	Few (less than a dozen)	Many (more than a dozen)

Use R to scale your spreadsheets

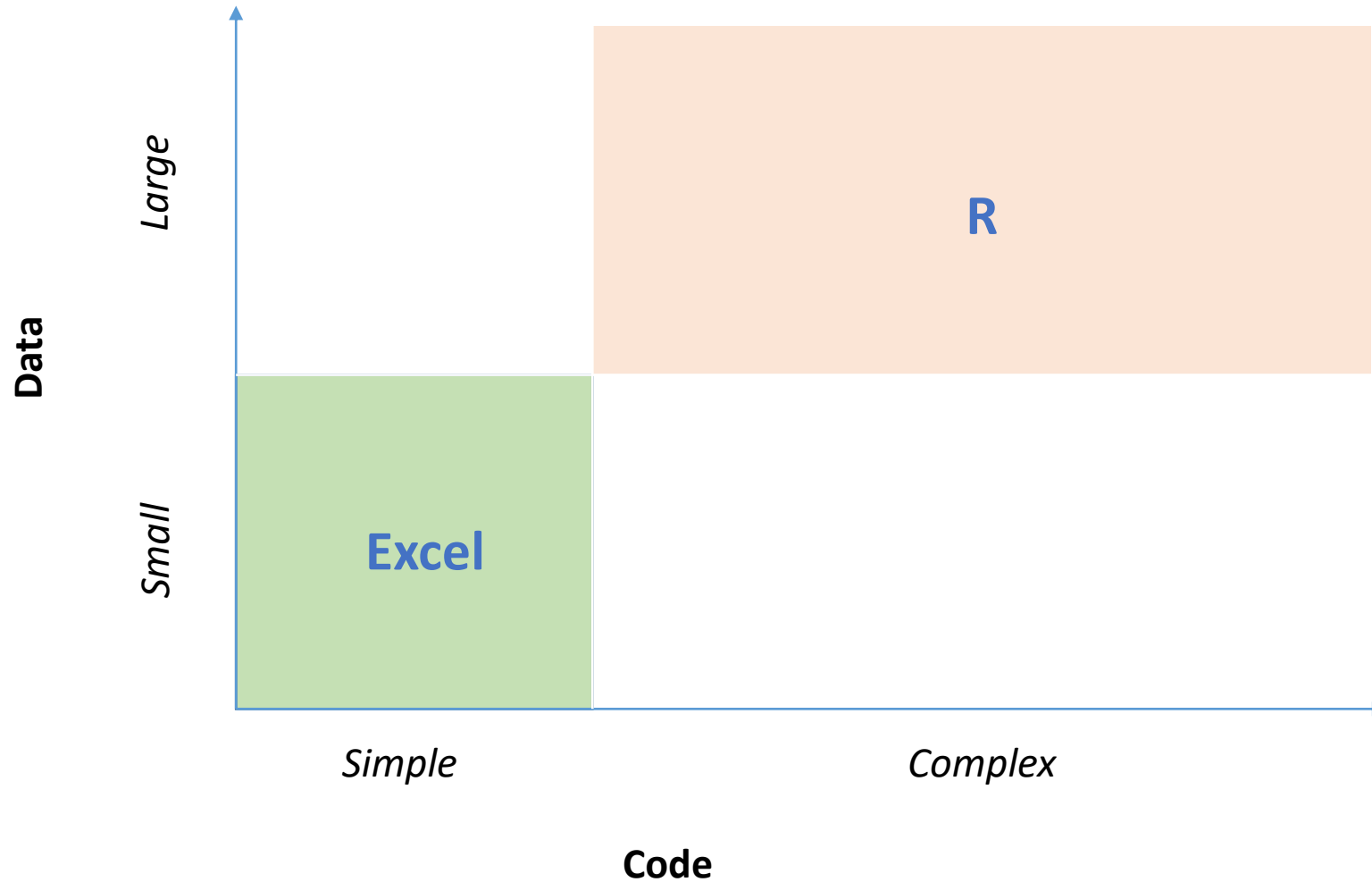
- R handles more data

	1's Mb	10's Mb	100's Mb	1's Gb	10's Gb	100's Gb
Excel						
R						

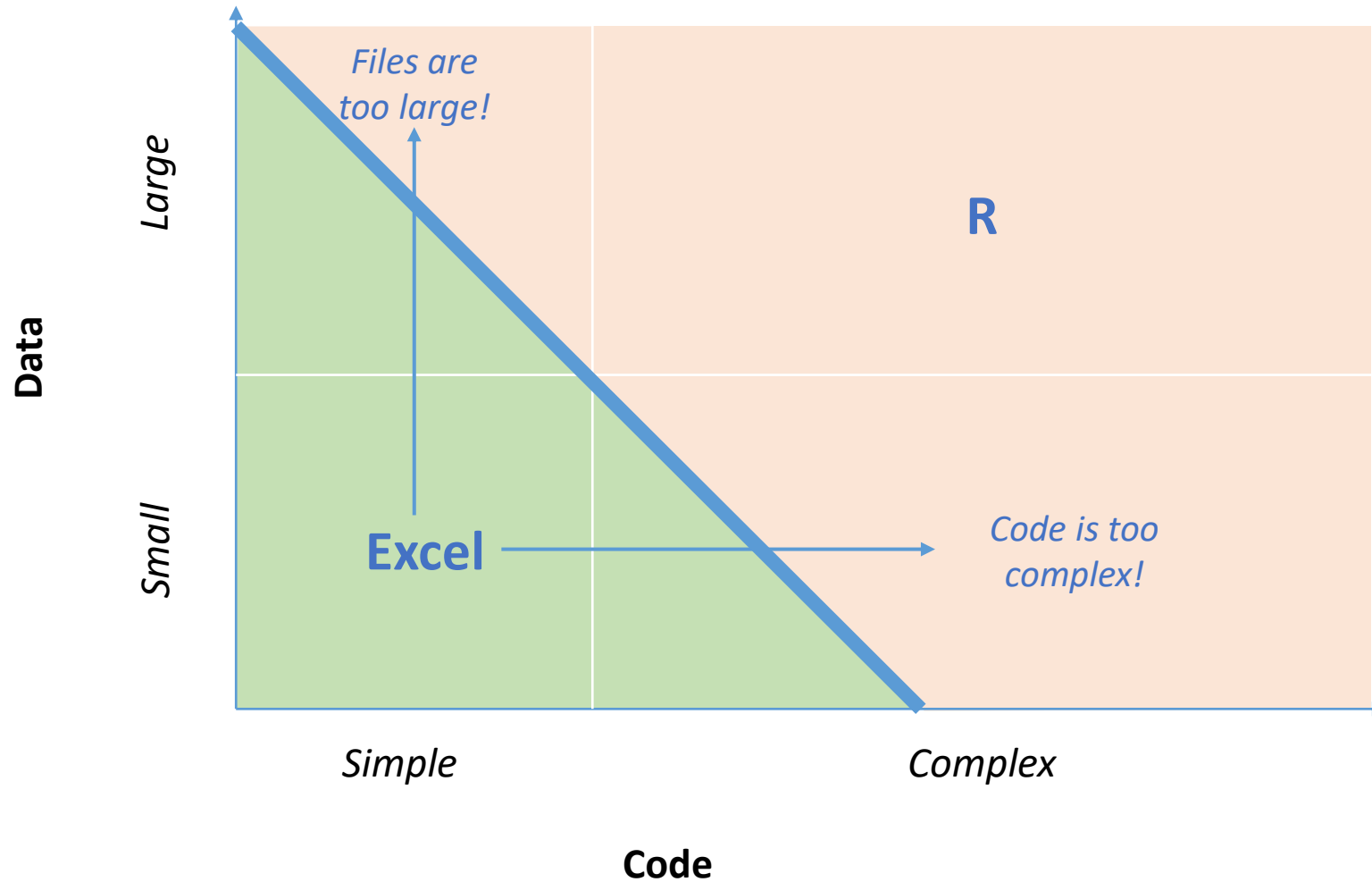
- R code simplifies complex tasks

Simple in R		
Example	Simple Workbook	Advanced Workbook
Data sources	Single	Multiple
Analytics	Summary statistics	Advanced algorithms
Data updates	Static or none	Dynamic
Formulas and Functions	One or two arguments	Multi-line and nested
Advanced tools	Single drop-down list or none	Scripts
Pivot Tables and Charts	One or two	Several
Spreadsheets	Few (less than a dozen)	Many (more than a dozen)

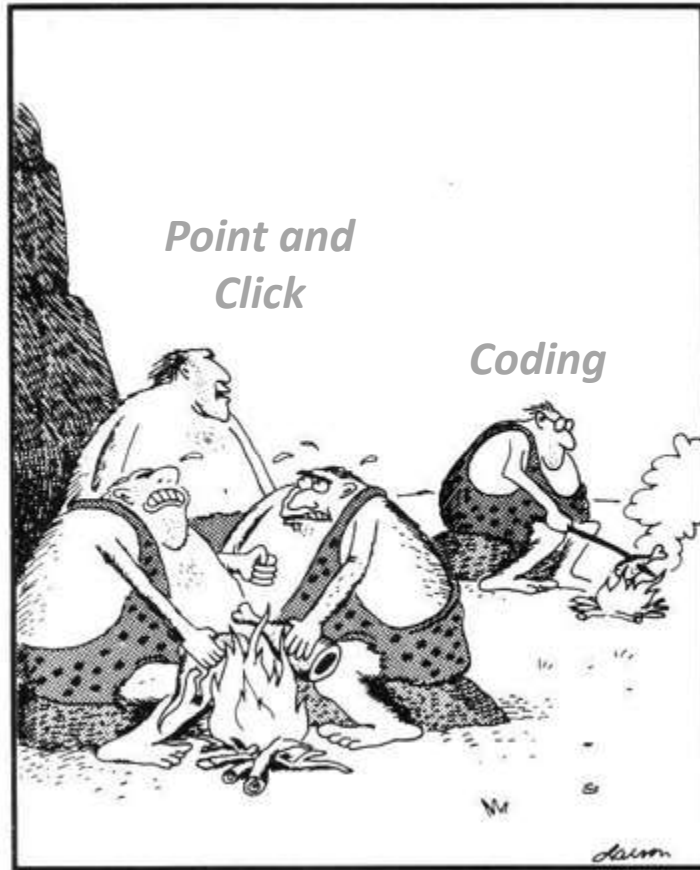
When to stop using Excel and start using R



When to stop using Excel and start using R



For complex code there is a better way



"Hey! Look what Zog do!"

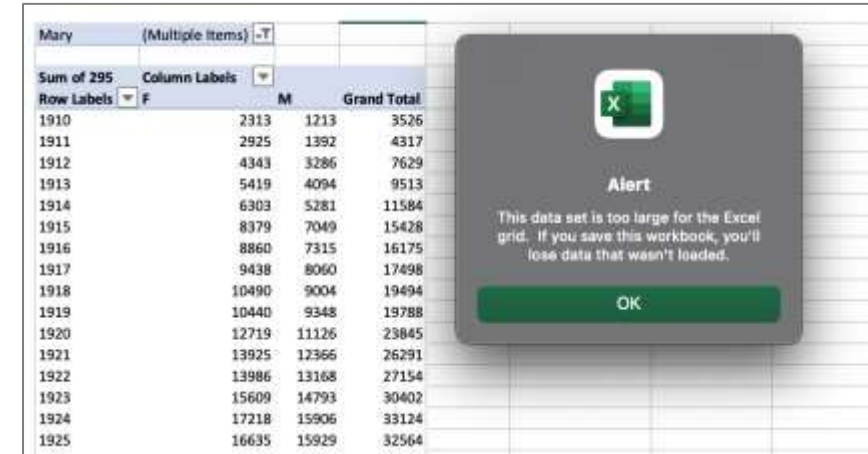
Benefits of using reproducible code

- Wider variety of tools
- More powerful tools
- Reproducible by others
- Easier to inspect for errors

Demos

1) File is too large

- [US State Specific Baby Names](#)



2) Code is too complex

- [Customer Tracker](#)

