CheckCell

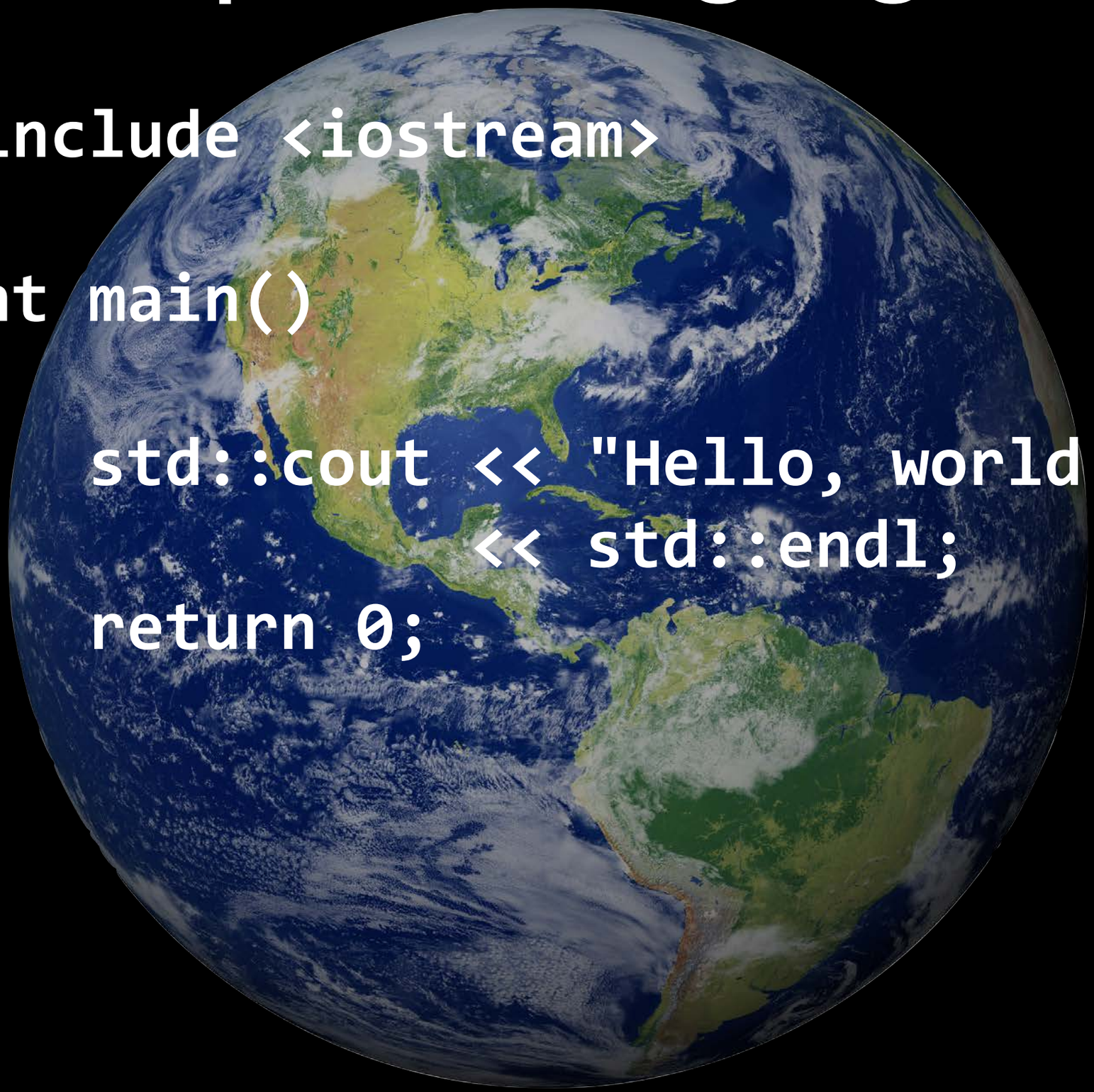
Dan Barowy, Dimitar Gochev, **Emery Berger**

UMASSCS

SCHOOL OF COMPUTER SCIENCE

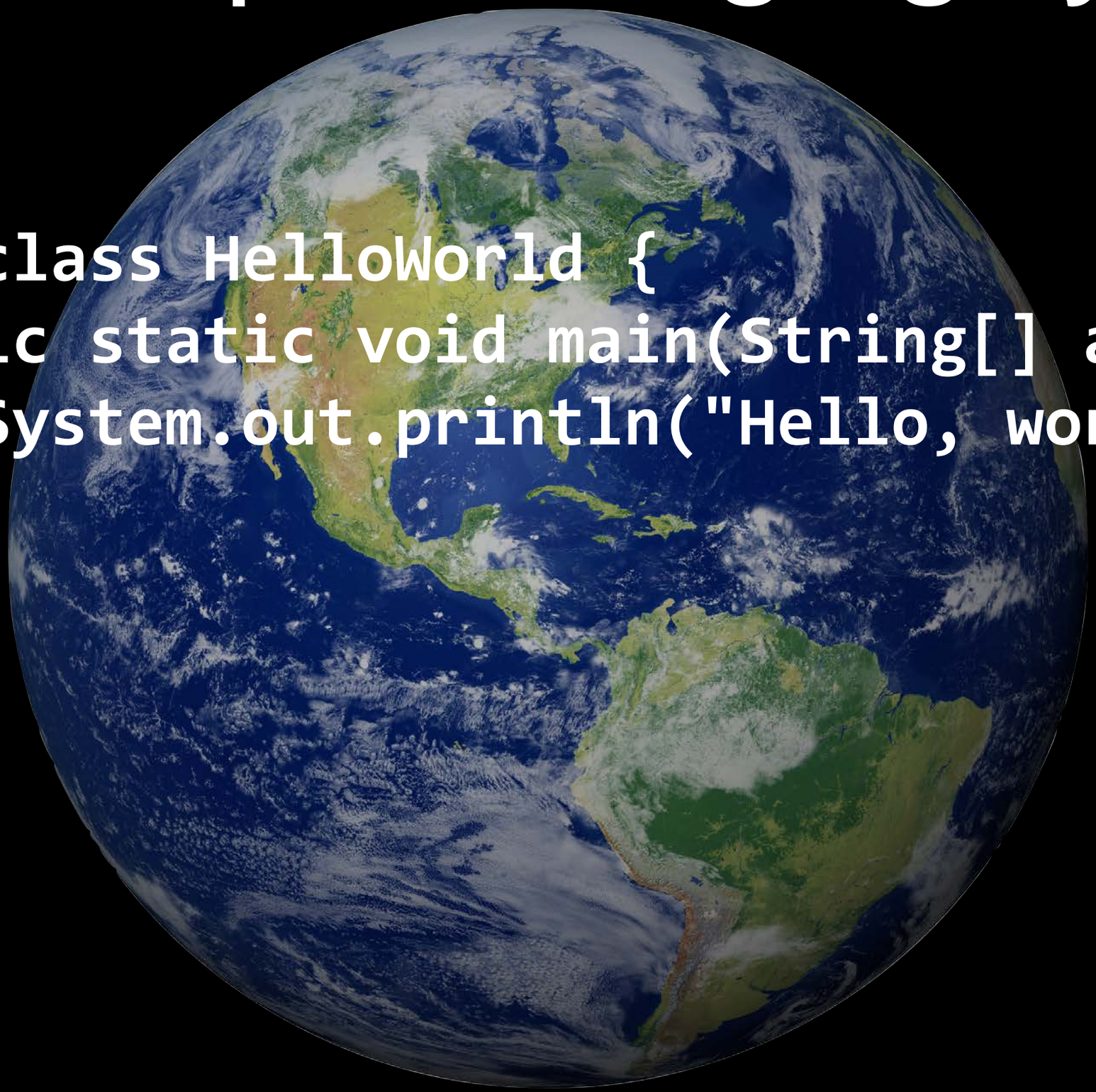
Most Popular Language?

Most Popular Language: C++?



```
#include <iostream>

int main()
{
    std::cout << "Hello, world!"
              << std::endl;
    return 0;
}
```

Most Popular Language: Java?

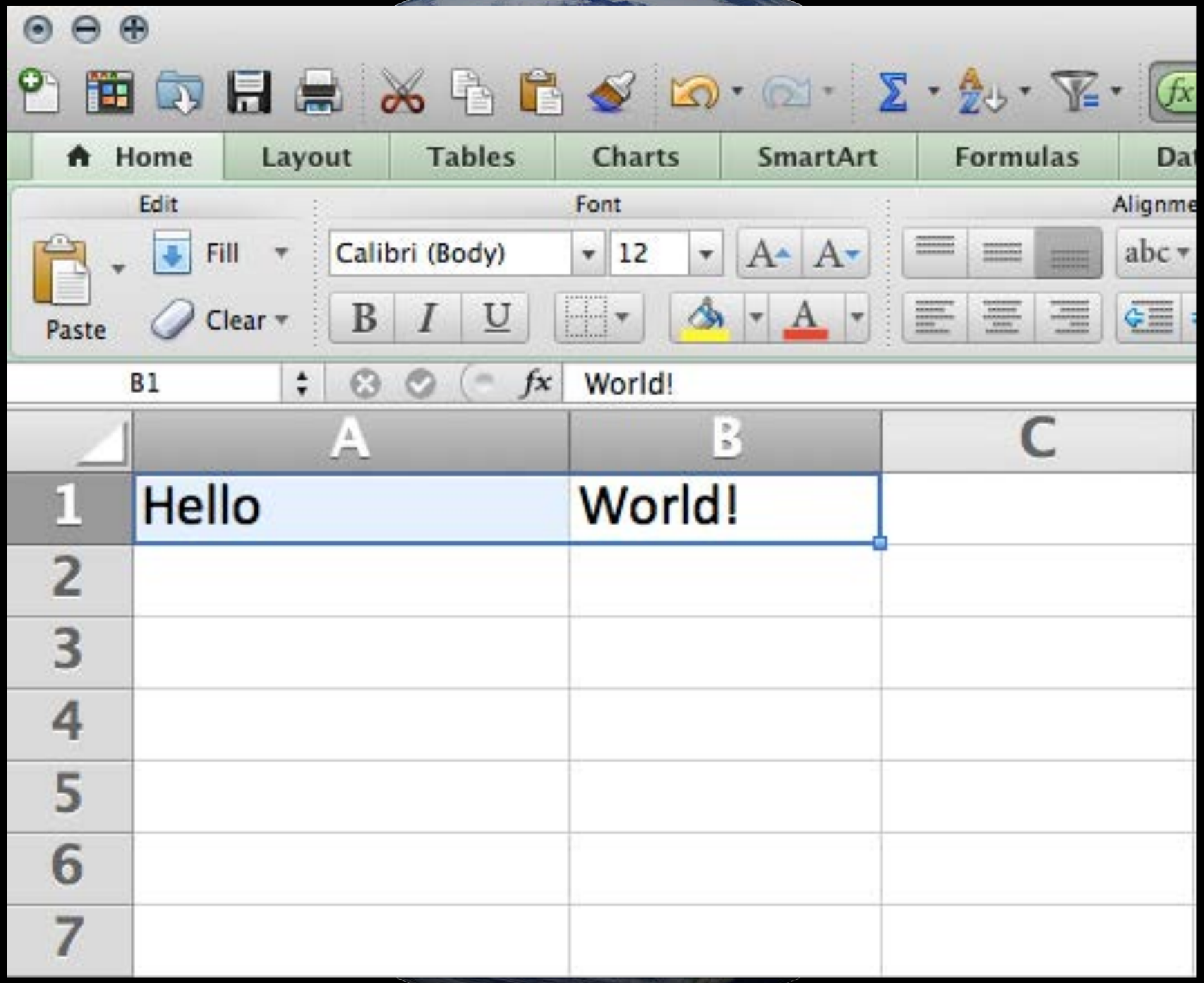


```
public class HelloWorld {
    public static void main(String[] args) {
        System.out.println("Hello, world!");
    }
}
```

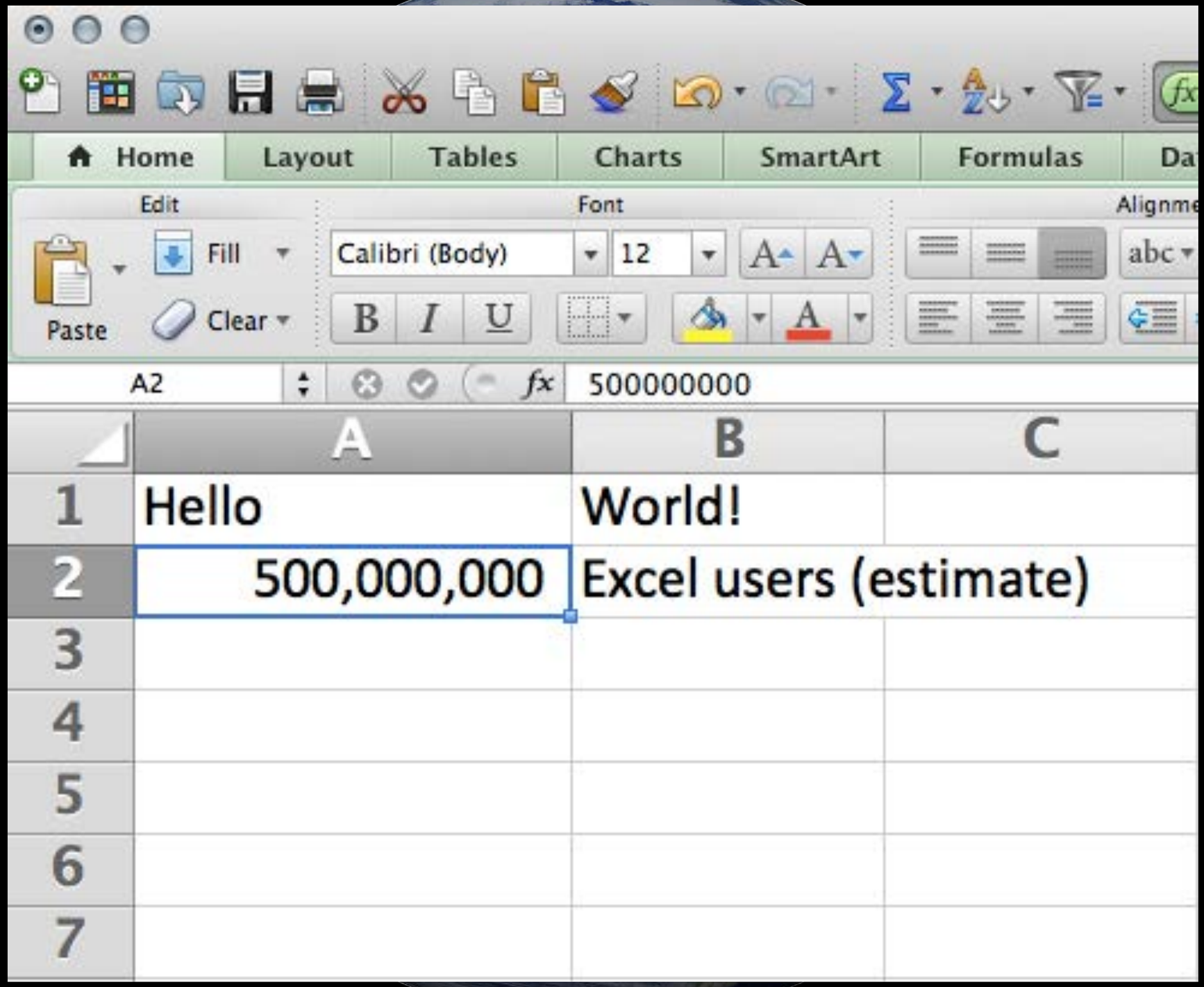
Most Popular Language: Java?

```
import static org.junit.Assert.assertEquals;
import java.io.ByteArrayOutputStream;
import java.io.PrintStream;
import org.junit.Test;

public class HelloWorldTest {
    @Test
    public void sayHelloWorld() {
        ByteArrayOutputStream outContent = captureSystemOut();

        HelloWorld.say();

        assertEquals("Hello, World!", outContent.toString());
    }

    ByteArrayOutputStream captureSystemOut() {
        ByteArrayOutputStream outContent = new ByteArrayOutputStream();
        System.setOut(new PrintStream(outContent));
        return outContent;
    }
}
public class HelloWorld {
    public static void say() {
        System.out.print("Hello, World!");
    }
}
```

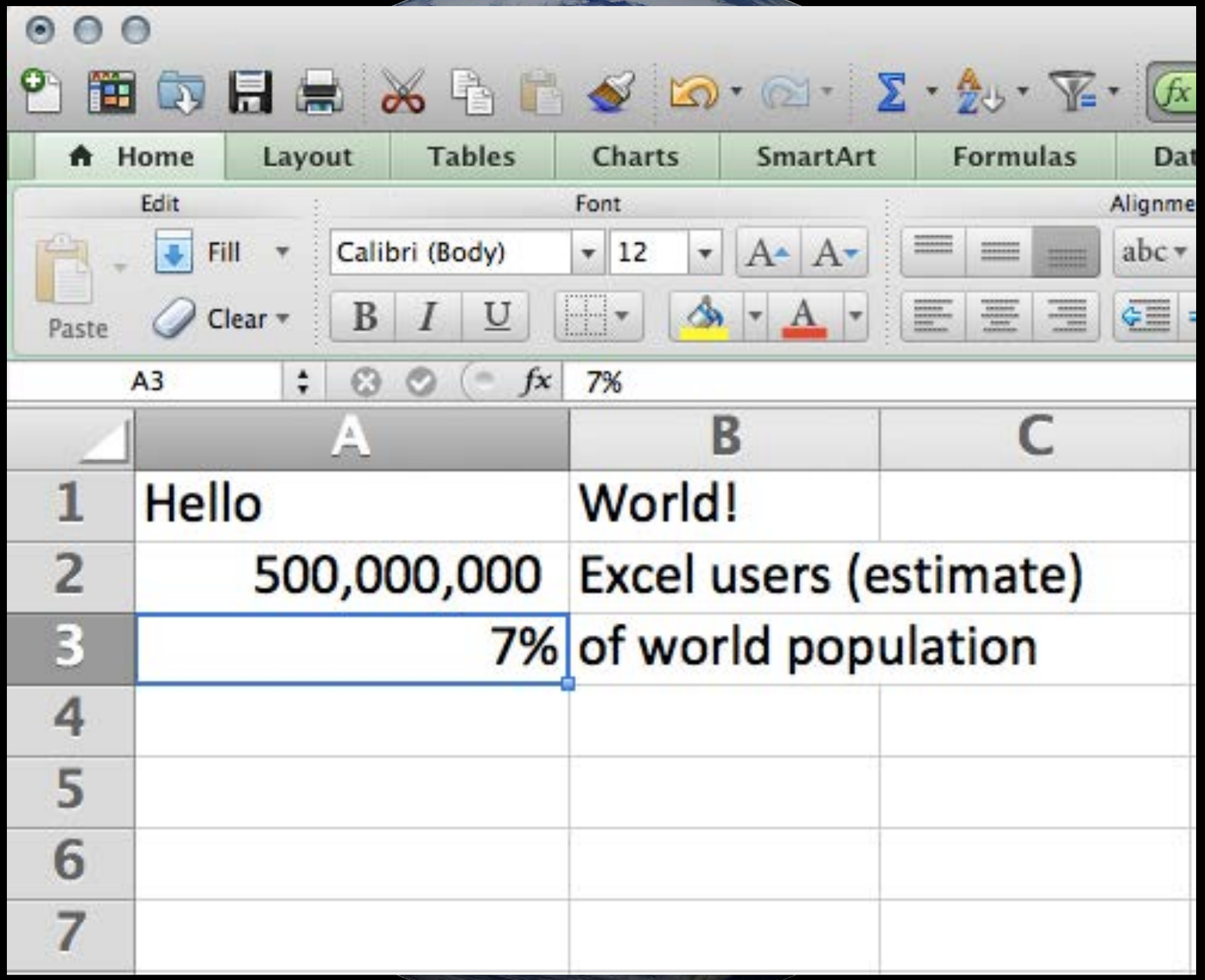
Most Popular Language: Excel



Most Popular Language: Excel



Most Popular Language: Excel



Tim Worstall
Contributor

FOLLOW

I write about business and technology.
full bio →

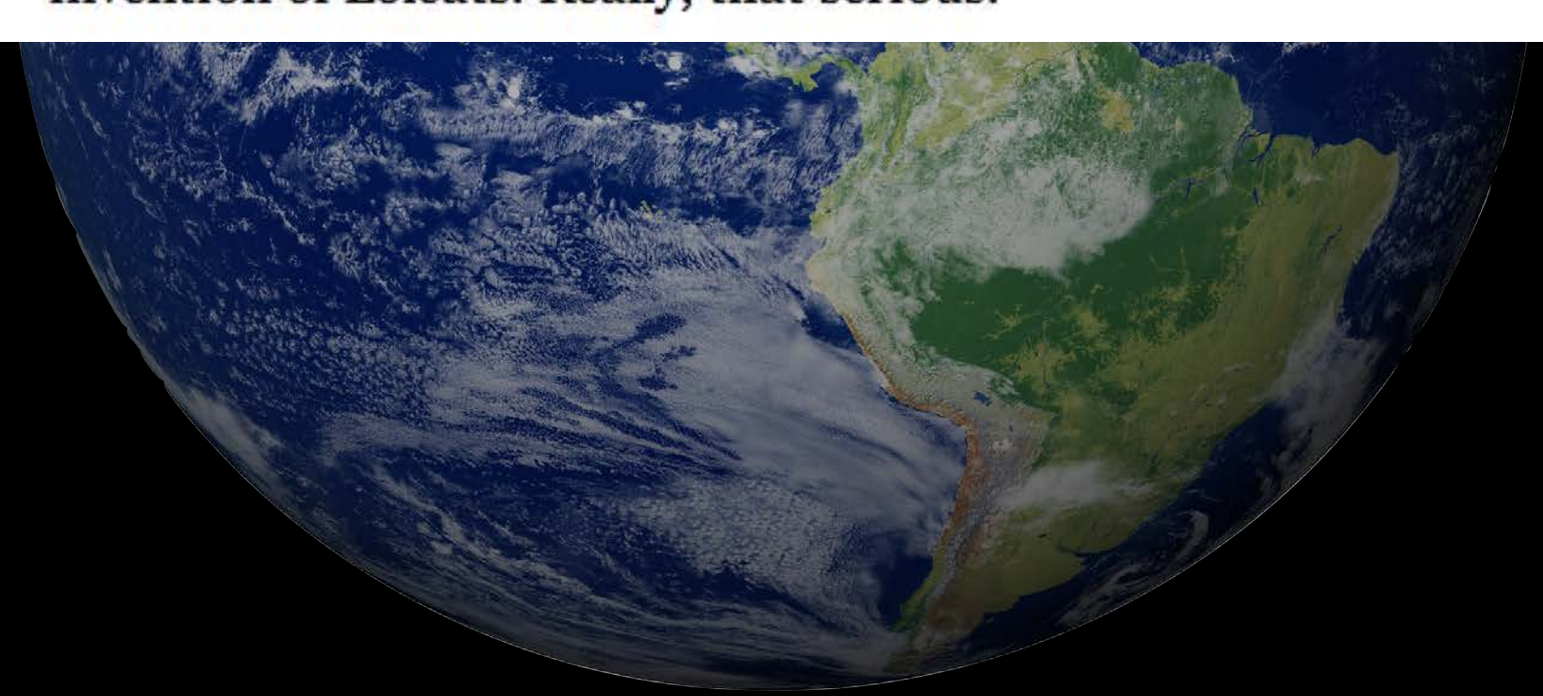
Opinions expressed by Forbes Contributors are their own.

TECH 2/13/2013 @ 9:37AM | 91,335 views

Microsoft's Excel Might Be The Most Dangerous Software On The Planet

+ Comment Now + Follow Comments

No, really, it's possible that Microsoft's Excel is the most dangerous software on the planet. Yes, more dangerous than rogue code running a nuclear power plant, than the Stuxnet that was deliberately sent off to sabotage Iran's nuclear program, worse, even, than whatever rent in the fabric of space time led to the invention of Lolcats. Really, that serious.



Tim Worstall
Contributor

FOLLOW

I write about business and technology.
full bio →

Opinions expressed by Forbes Contributors are their own.

Microsoft's Excel Might Be The Most Dangerous Software On The Planet

+ Comment Now + Follow Comments

No, really, it's possible that Microsoft's Excel is the most dangerous software on the planet. Yes, more dangerous than rogue code running a nuclear power plant, than the Stuxnet that was deliberately sent off to sabotage Iran's nuclear program, worse, even, than whatever rent in the fabric of space time led to the invention of Lolcats. Really, that serious.

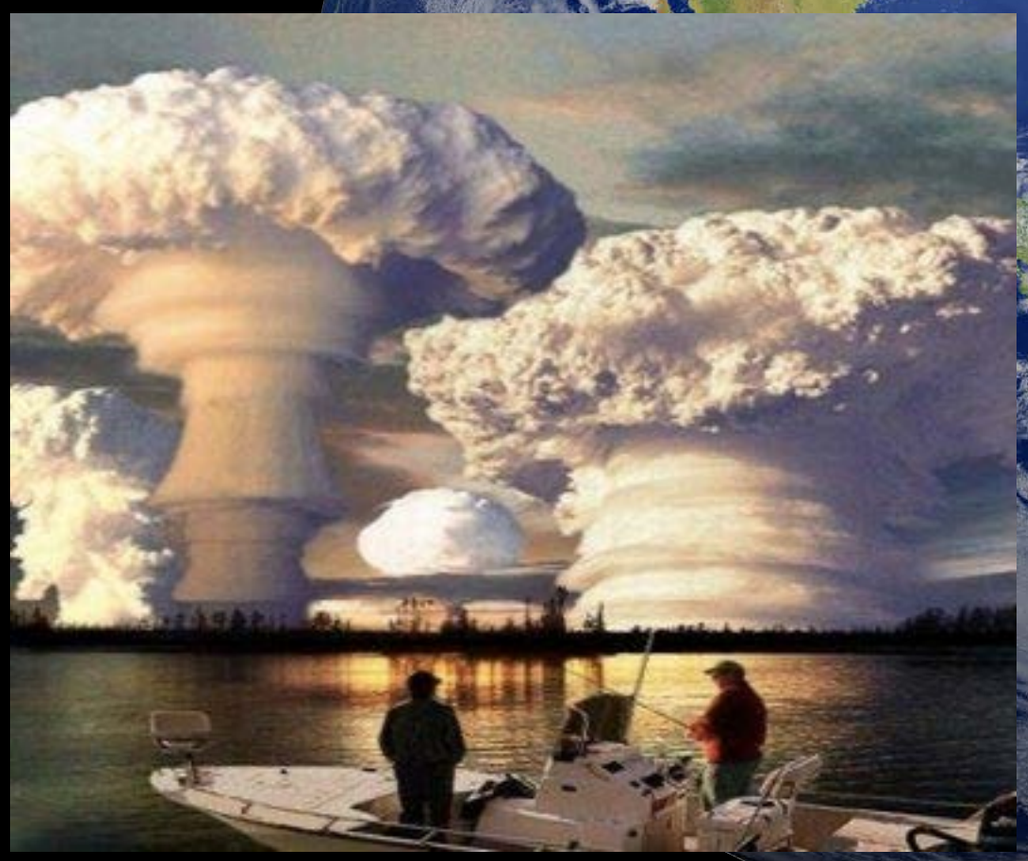
Tim Worstall
Contributor

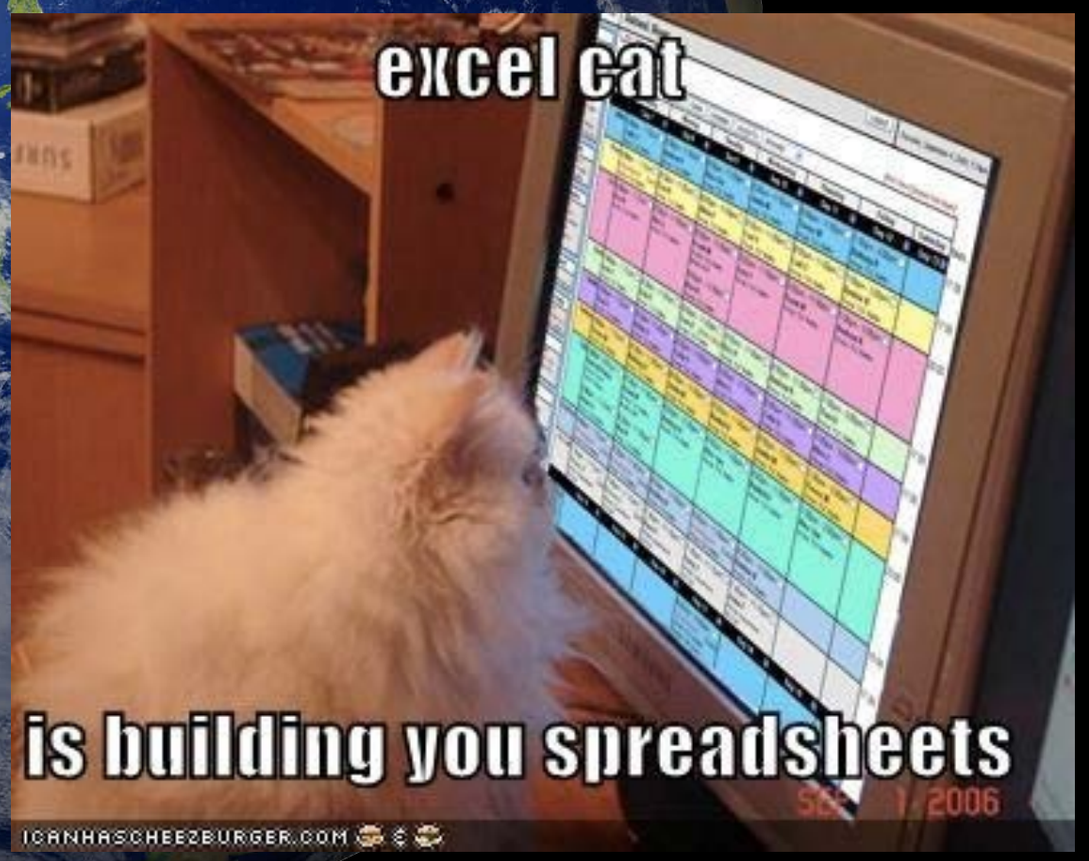
FOLLOW

I write about business and technology.
full bio →

Microsoft's Excel Might Be The Most Dangerous Software On The Planet

No, really, it's possible that Microsoft's Excel is the most dangerous software on the planet. Yes, more dangerous than rogue code running a nuclear power plant, than the Stuxnet that was deliberately sent off to sabotage Iran's nuclear program, worse, even, than whatever rent in the fabric of space time led to the invention of Lolcats. Really, that serious.





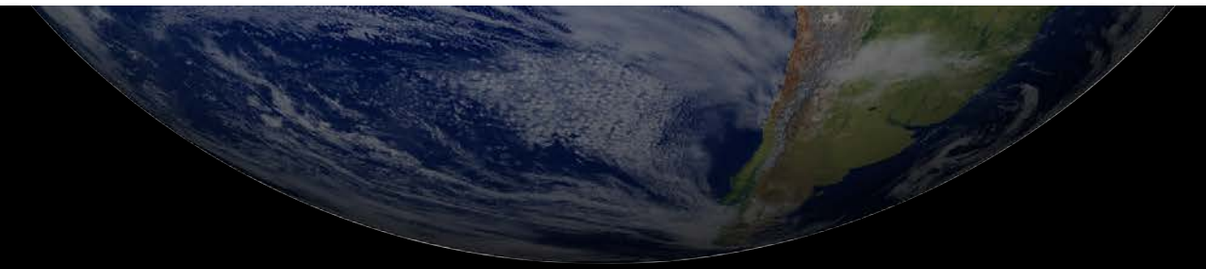
Excel snafu costs firm $24m

Some cleric, some error

By Drew Cullen, 19th June 2003

A simple spreadsheet error cost a firm a whopping US$24m.

The mistake led to TransAlta, a big Canadian power generator, buying more US power transmission hedging contracts in May at higher prices than it should have.



Flintshire County Council school cash blunder 'down to spreadsheet error'

A £1M council cash 'cock-up' which prompted calls for the resignation of a top Tory councillor was down to a spreadsheet error, it has been claimed.

A £1M council cash 'cock-up' which prompted calls for the resignation of a top Tory councillor was down to a spreadsheet error, it has been claimed.

County Hall officials have not published details of an internal investigation carried out in the wake of the blunder, which left Flintshire Council £1m out of pocket after they over-allocated money cash to the county's 12 high schools last year.

But the Chronicle understands the mix-up occurred because someone 'put money in the wrong column' – and no one spotted it before the money was handed out.

Computing error at Fidelity's Magellan fund

Kathy Godfrey <kgodfrey@BBN.COM>
Wed, 4 Jan 95 17:44:41 EST

There was a big flap recently over Fidelity's Magellan fund estimating in November that they would make a $4.32/share distribution at the end of year, and then not doing so. A letter of explanation was sent to the shareholders (including me) from J. Gary Burkhead, the President of Fidelity, including the following pertinent items:

"During the estimating process, a tax accountant is required to transcribe the net realized gain or loss from the fund's financial records (which were correct at all times) to a separate spreadsheet, where additional calculations are performed. The error occurred when the accountant omitted the minus sign on a net capital loss of $1.3 billion and incorrectly treated it as a net capital gain on this separate spreadsheet. This meant that the dividend estimate spreadsheet was off by $2.6 billion....

"Some people have asked how, in this age of technology, such a mistake could be made. While many of our processes are computerized, the requirements of

Computing error at Fidelity's Magellan fund

Kathy Godfrey <kgodfrey@BBN.COM>
Wed, 4 Jan 95 17:44:41 EST

There was a big flap recently over Fidelity's Magellan fund estimating in November that they would make a $4.32/share distribution at the end of year, and then not doing so. A letter of explanation was sent to the shareholders (including me) from J. Gary Burkhead, the President of Fidelity, including the following pertinent items:

"During the estimating process, a tax accountant is required to transcribe the net realized gain or loss from the fund's financial records (which were correct at all times) to a separate spreadsheet, where additional calculations are performed. The error occurred when the accountant omitted the minus sign on a net capital loss of $1.3 billion and incorrectly treated it as a net capital gain on this separate spreadsheet. This meant that the dividend estimate spreadsheet was off by $2.6 billion....

"Some people have asked how, in this age of technology, such a mistake could be made. While many of our processes are computerized, the requirements of

FAQ: Reinhart, Rogoff, and the Excel Error That Changed History

By Peter Coy | April 18, 2013

SEND TO kindle

How big is this mistake?

Reinhart and Rogoff wrote in their 2010 paper that average annual growth was negative 0.1 percent in countries with episodes of gross government debt equal to 90 percent or more of GDP between 1945 and 2009. Liberal economists have been critical of their work for years (just economists being their usual cranky selves), but now three economists at UMass say Reinhart and Rogoff made several mistakes and omissions. According to the UMass scholars, the "corrected" number is positive 2.2 percent—which means GDP still grows, even when debt levels are very high.

Photograph by Gregor Schuster

The Excel Depression

By PAUL KRUGMAN
Published: April 18, 2013 | 470 Comments

In this age of information, math errors can lead to disaster. NASA's Mars Orbiter crashed because engineers forgot to convert to metric measurements; JPMorgan Chase's "London Whale" venture went bad in part because modelers divided by a sum instead of an average. So, did an Excel coding error destroy the economies of the Western world?

Enlarge This Image

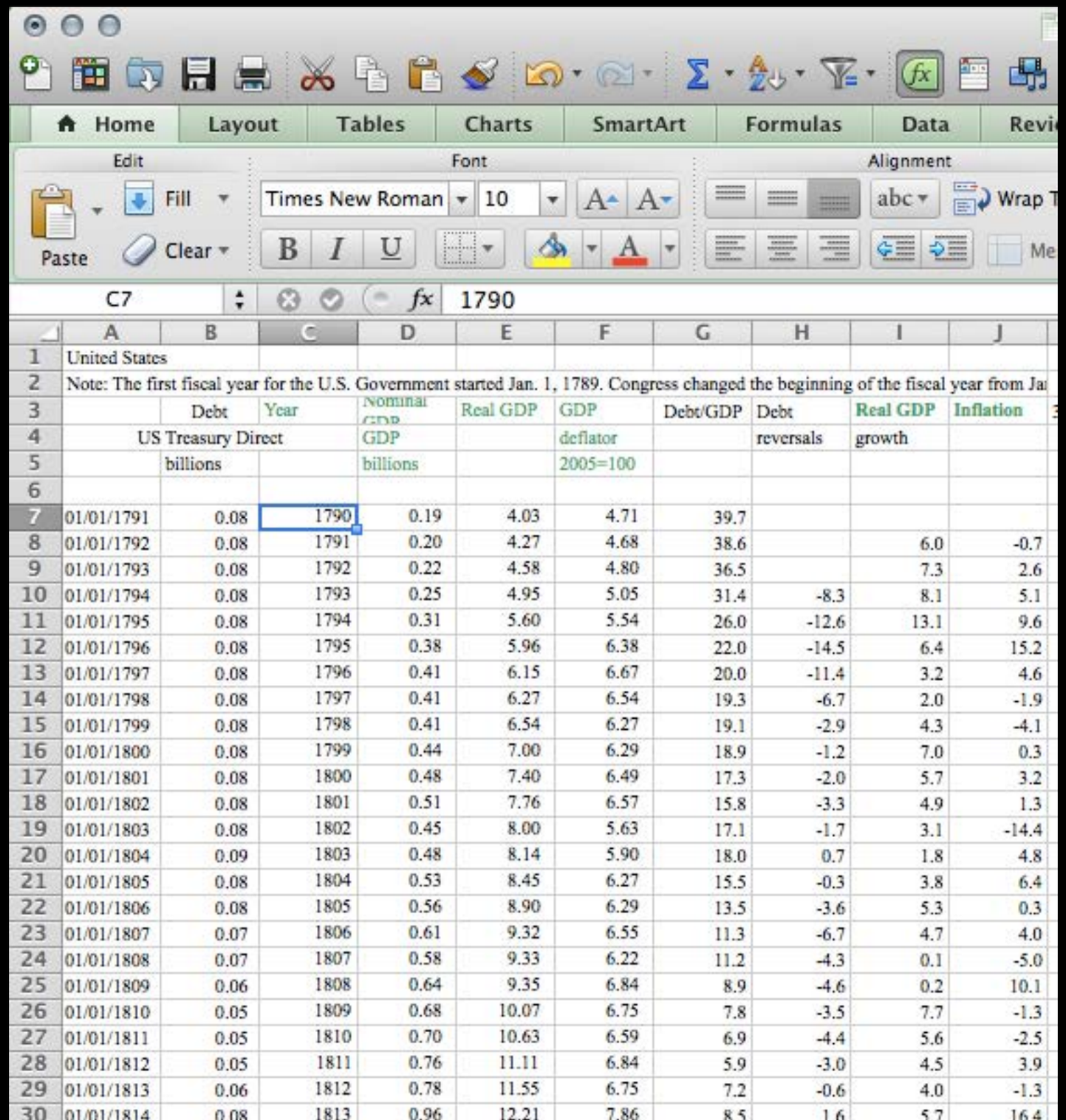
Fred R. Conrad/The New York Times
Paul Krugman

Go to Columnist Page »

The story so far: At the beginning of 2010, two Harvard economists, Carmen Reinhart and Kenneth Rogoff, circulated a paper, "Growth in a Time of Debt," that purported to identify a critical "threshold," a tipping point, for government indebtedness. Once debt exceeds 90 percent of gross domestic product, they claimed, economic growth drops off sharply.

Ms. Reinhart and Mr. Rogoff had credibility thanks to a widely admired earlier book on the history of financial crises, and their timing was impeccable. The paper came out just after Greece went into crisis and played right into the desire of many officials to "pivot" from stimulus to austerity. As a result, the paper instantly became famous; it was, and is, surely the most influential economic analysis of recent years.

What's the Problem?



Home | Layout | Tables | Charts | SmartArt | Formulas | Data | Revi

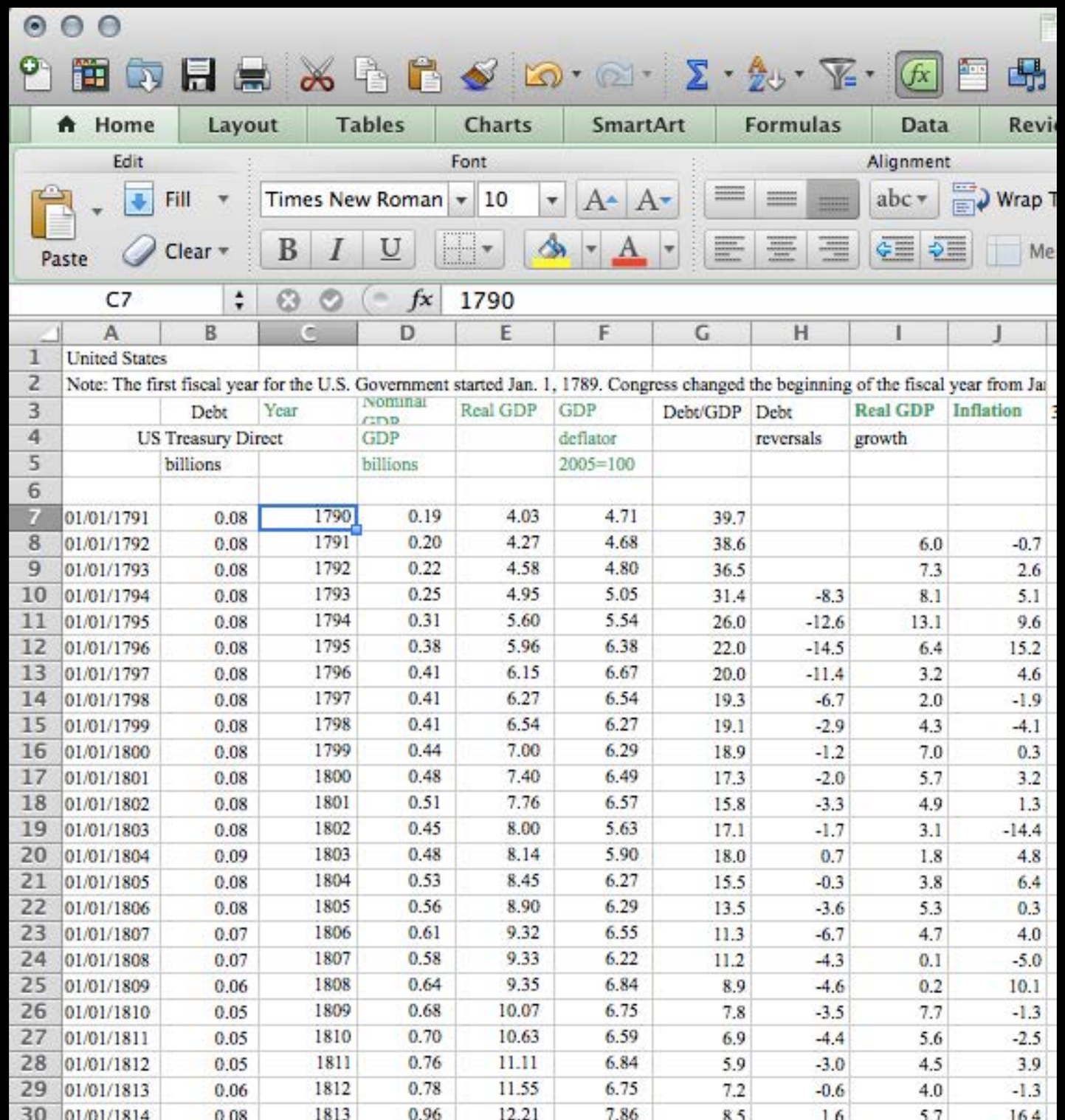
Edit | Font | Alignment

Paste | Fill | Clear | Times New Roman | 10 | B I U | abc | Wrap T

C7 | fx | 1790

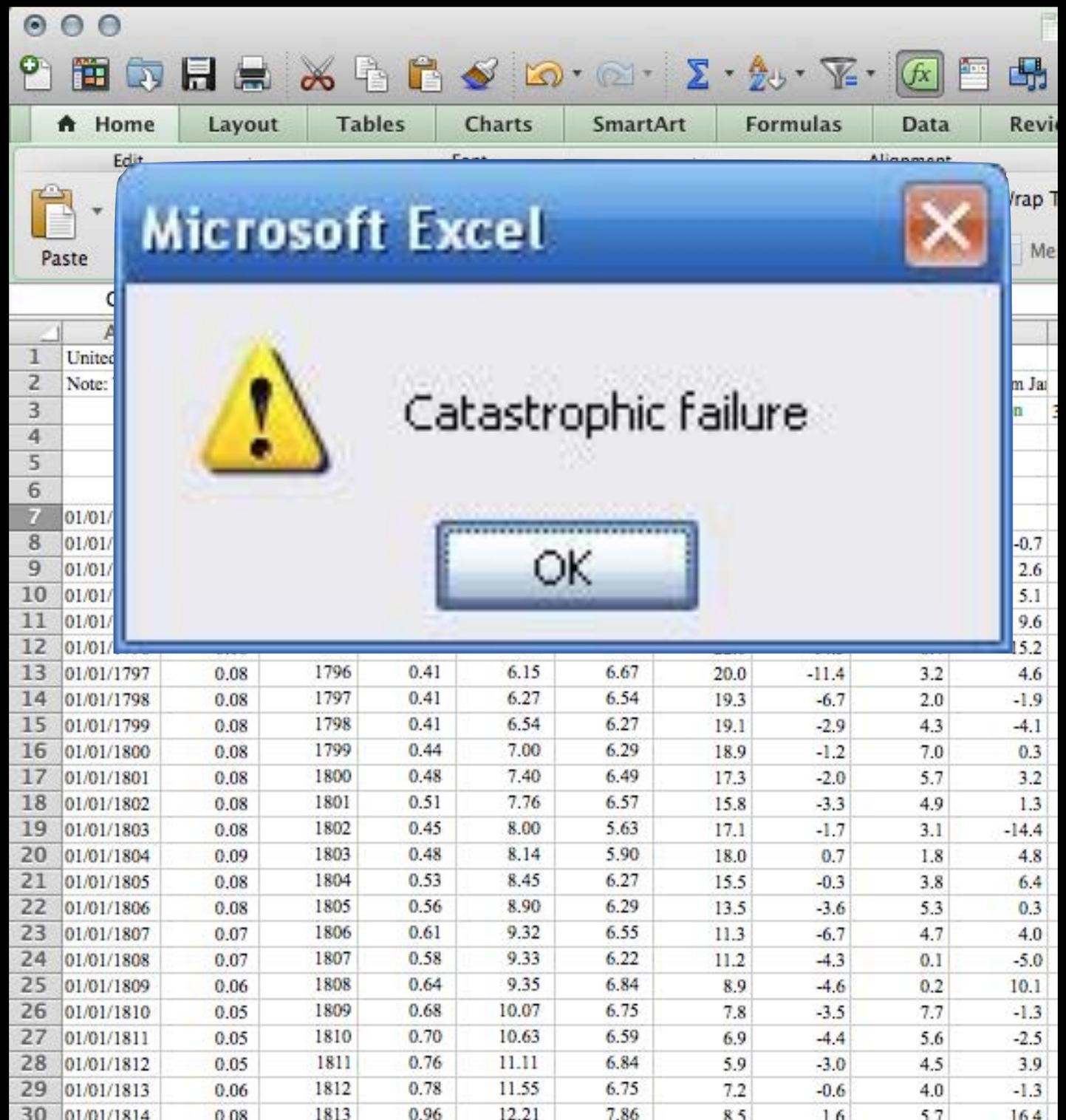
	A	B	C	D	E	F	G	H	I	J
1	United States									
2	Note: The first fiscal year for the U.S. Government started Jan. 1, 1789. Congress changed the beginning of the fiscal year from Ja									
3		Debt	Year	Nominal GDP	Real GDP	GDP	Debt/GDP	Debt	Real GDP	Inflation
4		US Treasury Direct		GDP		deflator		reversals	growth	
5		billions		billions		2005=100				
6										
7	01/01/1791	0.08	1790	0.19	4.03	4.71	39.7			
8	01/01/1792	0.08	1791	0.20	4.27	4.68	38.6		6.0	-0.7
9	01/01/1793	0.08	1792	0.22	4.58	4.80	36.5		7.3	2.6
10	01/01/1794	0.08	1793	0.25	4.95	5.05	31.4	-8.3	8.1	5.1
11	01/01/1795	0.08	1794	0.31	5.60	5.54	26.0	-12.6	13.1	9.6
12	01/01/1796	0.08	1795	0.38	5.96	6.38	22.0	-14.5	6.4	15.2
13	01/01/1797	0.08	1796	0.41	6.15	6.67	20.0	-11.4	3.2	4.6
14	01/01/1798	0.08	1797	0.41	6.27	6.54	19.3	-6.7	2.0	-1.9
15	01/01/1799	0.08	1798	0.41	6.54	6.27	19.1	-2.9	4.3	-4.1
16	01/01/1800	0.08	1799	0.44	7.00	6.29	18.9	-1.2	7.0	0.3
17	01/01/1801	0.08	1800	0.48	7.40	6.49	17.3	-2.0	5.7	3.2
18	01/01/1802	0.08	1801	0.51	7.76	6.57	15.8	-3.3	4.9	1.3
19	01/01/1803	0.08	1802	0.45	8.00	5.63	17.1	-1.7	3.1	-14.4
20	01/01/1804	0.09	1803	0.48	8.14	5.90	18.0	0.7	1.8	4.8
21	01/01/1805	0.08	1804	0.53	8.45	6.27	15.5	-0.3	3.8	6.4
22	01/01/1806	0.08	1805	0.56	8.90	6.29	13.5	-3.6	5.3	0.3
23	01/01/1807	0.07	1806	0.61	9.32	6.55	11.3	-6.7	4.7	4.0
24	01/01/1808	0.07	1807	0.58	9.33	6.22	11.2	-4.3	0.1	-5.0
25	01/01/1809	0.06	1808	0.64	9.35	6.84	8.9	-4.6	0.2	10.1
26	01/01/1810	0.05	1809	0.68	10.07	6.75	7.8	-3.5	7.7	-1.3
27	01/01/1811	0.05	1810	0.70	10.63	6.59	6.9	-4.4	5.6	-2.5
28	01/01/1812	0.05	1811	0.76	11.11	6.84	5.9	-3.0	4.5	3.9
29	01/01/1813	0.06	1812	0.78	11.55	6.75	7.2	-0.6	4.0	-1.3
30	01/01/1814	0.08	1813	0.96	12.21	7.86	8.5	1.6	5.7	16.4

Data >> Formulas



	A	B	C	D	E	F	G	H	I	J
1	United States									
2	Note: The first fiscal year for the U.S. Government started Jan. 1, 1789. Congress changed the beginning of the fiscal year from Ja									
3		Debt	Year	Nominal GDP	Real GDP	GDP	Debt/GDP	Debt	Real GDP	Inflation
4		US Treasury Direct		GDP		deflator		reversals	growth	
5		billions		billions		2005=100				
6										
7	01/01/1791	0.08	1790	0.19	4.03	4.71	39.7			
8	01/01/1792	0.08	1791	0.20	4.27	4.68	38.6		6.0	-0.7
9	01/01/1793	0.08	1792	0.22	4.58	4.80	36.5		7.3	2.6
10	01/01/1794	0.08	1793	0.25	4.95	5.05	31.4	-8.3	8.1	5.1
11	01/01/1795	0.08	1794	0.31	5.60	5.54	26.0	-12.6	13.1	9.6
12	01/01/1796	0.08	1795	0.38	5.96	6.38	22.0	-14.5	6.4	15.2
13	01/01/1797	0.08	1796	0.41	6.15	6.67	20.0	-11.4	3.2	4.6
14	01/01/1798	0.08	1797	0.41	6.27	6.54	19.3	-6.7	2.0	-1.9
15	01/01/1799	0.08	1798	0.41	6.54	6.27	19.1	-2.9	4.3	-4.1
16	01/01/1800	0.08	1799	0.44	7.00	6.29	18.9	-1.2	7.0	0.3
17	01/01/1801	0.08	1800	0.48	7.40	6.49	17.3	-2.0	5.7	3.2
18	01/01/1802	0.08	1801	0.51	7.76	6.57	15.8	-3.3	4.9	1.3
19	01/01/1803	0.08	1802	0.45	8.00	5.63	17.1	-1.7	3.1	-14.4
20	01/01/1804	0.09	1803	0.48	8.14	5.90	18.0	0.7	1.8	4.8
21	01/01/1805	0.08	1804	0.53	8.45	6.27	15.5	-0.3	3.8	6.4
22	01/01/1806	0.08	1805	0.56	8.90	6.29	13.5	-3.6	5.3	0.3
23	01/01/1807	0.07	1806	0.61	9.32	6.55	11.3	-6.7	4.7	4.0
24	01/01/1808	0.07	1807	0.58	9.33	6.22	11.2	-4.3	0.1	-5.0
25	01/01/1809	0.06	1808	0.64	9.35	6.84	8.9	-4.6	0.2	10.1
26	01/01/1810	0.05	1809	0.68	10.07	6.75	7.8	-3.5	7.7	-1.3
27	01/01/1811	0.05	1810	0.70	10.63	6.59	6.9	-4.4	5.6	-2.5
28	01/01/1812	0.05	1811	0.76	11.11	6.84	5.9	-3.0	4.5	3.9
29	01/01/1813	0.06	1812	0.78	11.55	6.75	7.2	-0.6	4.0	-1.3
30	01/01/1814	0.08	1813	0.96	12.21	7.86	8.5	1.6	5.7	16.4

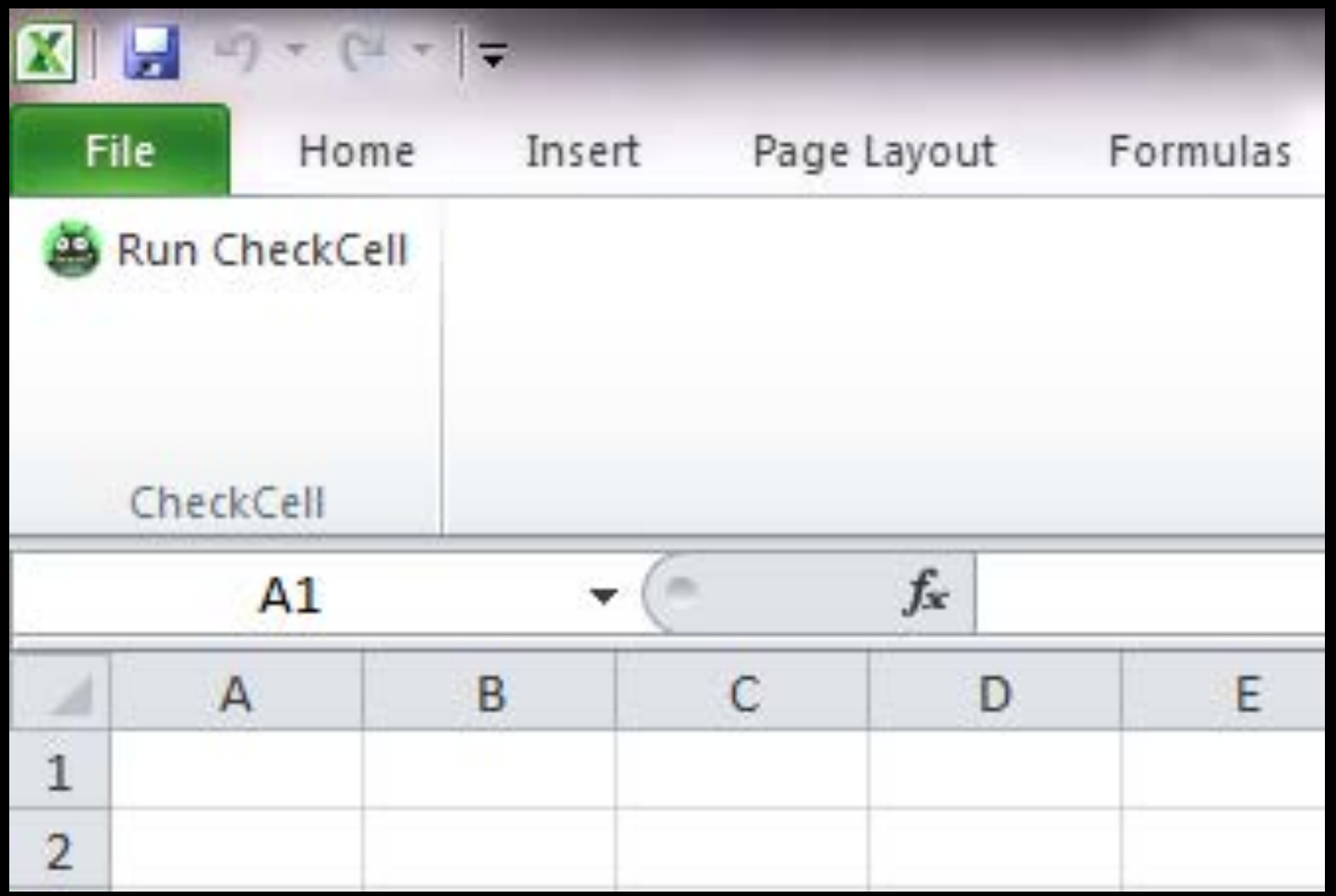
Data >> Formulas

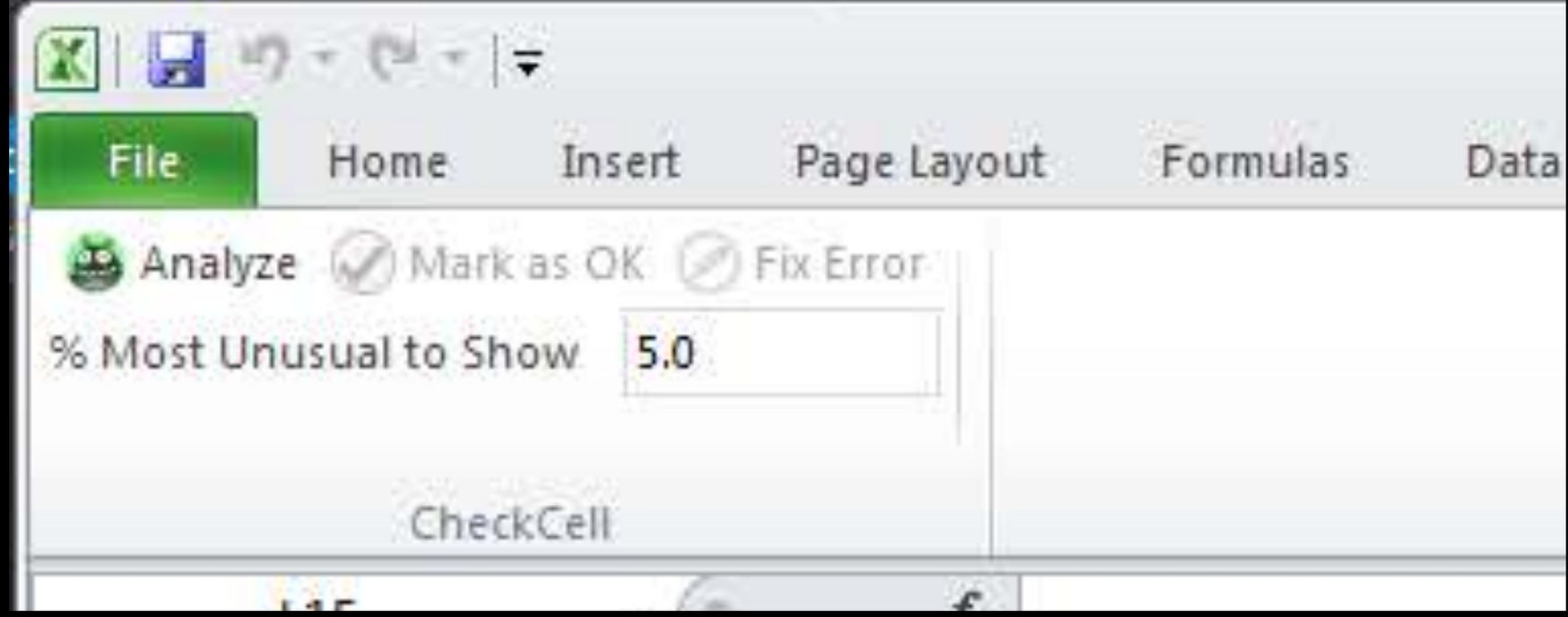


CheckCell

leverages interaction of **data &**
formulas to find ***potential*** **data bugs**

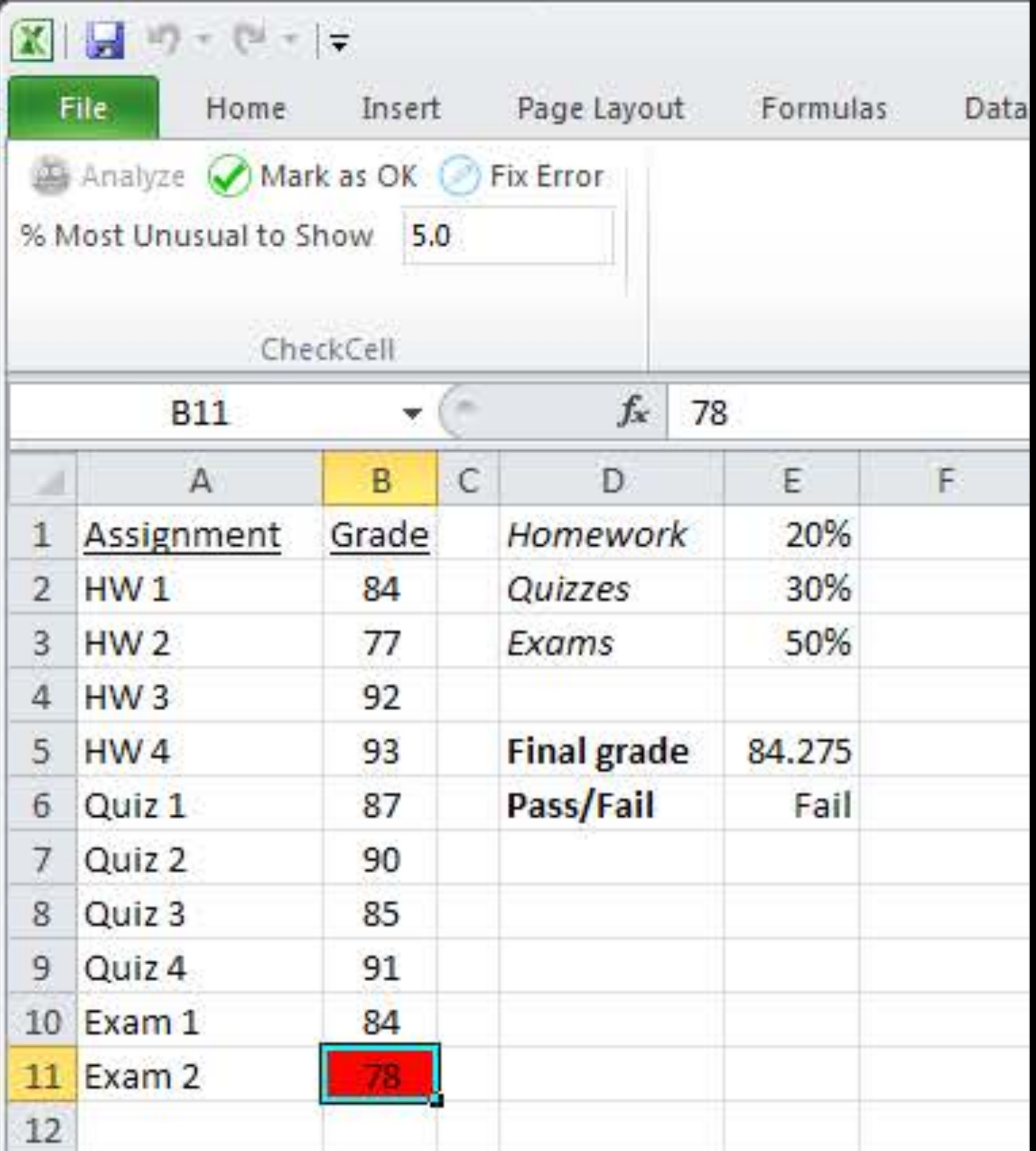
CheckCell


File
Home
Insert
Page Layout
Formulas
Run CheckCell
CheckCell
A1
A
B
C
D
E
1
2



Values with **unusually high impact:** really **important** to be **correct!**

5	HW 4	93		Final grade	84.275
6	Quiz 1	87		Pass/Fail	Fail
7	Quiz 2	90			
8	Quiz 3	85			
9	Quiz 4	91			
10	Exam 1	84			
11	Exam 2	78			
12					



	A	B	C	D	E	F
1	Assignment	Grade		*Homework*	20%	
2	HW 1	84		*Quizzes*	30%	
3	HW 2	77		*Exams*	50%	
4	HW 3	92				
5	HW 4	93		**Final grade**	84.275	
6	Quiz 1	87		**Pass/Fail**	Fail	
7	Quiz 2	90				
8	Quiz 3	85				
9	Quiz 4	91				
10	Exam 1	84				
11	Exam 2	78				
12						

File | Home | Insert | Page Layout | Formulas | Data | Review | Vi

Analyze | Mark as OK | Fix Error

% Most Unusual to Show 5.0

CheckCell

B11 | 78

	A	B	C	D	E	F	G	H
1	Assignment	Grade		*Homework*	20%			
2	HW 1							
3	HW 2							
4	HW 3							
5	HW 4							
6	Quiz 1							
7	Quiz 2							
8	Quiz 3							
9	Quiz 4							
10	Exam 1	84						
11	Exam 2	78						
12								

CellFixForm

Enter the corrected value:

87

Cancel | Fix

File | Home | Insert | Page Layout | Formulas | Data

Analyze | Mark as OK | Fix Error

% Most Unusual to Show 5.0

CheckCell

B11 | 87

	A	B	C	D	E	F
1	Assignment	Grade		*Homework*	20%	
2	HW 1	84		*Quizzes*	30%	
3	HW 2	77			50%	
4	HW 3	92				
5	HW 4	93			86.525	
6	Quiz 1	87			Pass	
7	Quiz 2	90				
8	Quiz 3	85				
9	Quiz 4	91				
10	Exam 1	84				
11	Exam 2	87				
12						

No bugs remain.

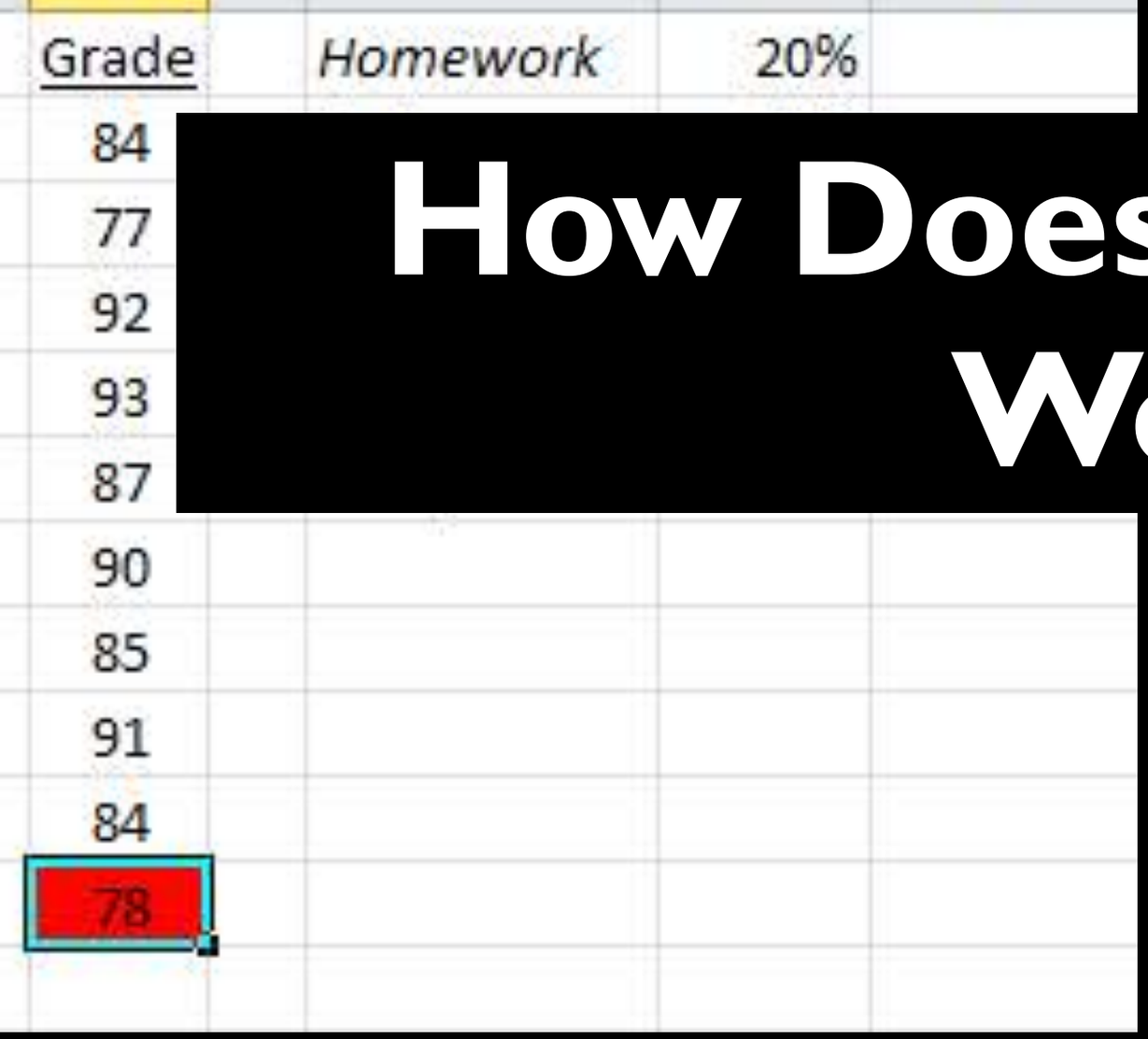
OK

How Does CheckCell Work?

CheckCell

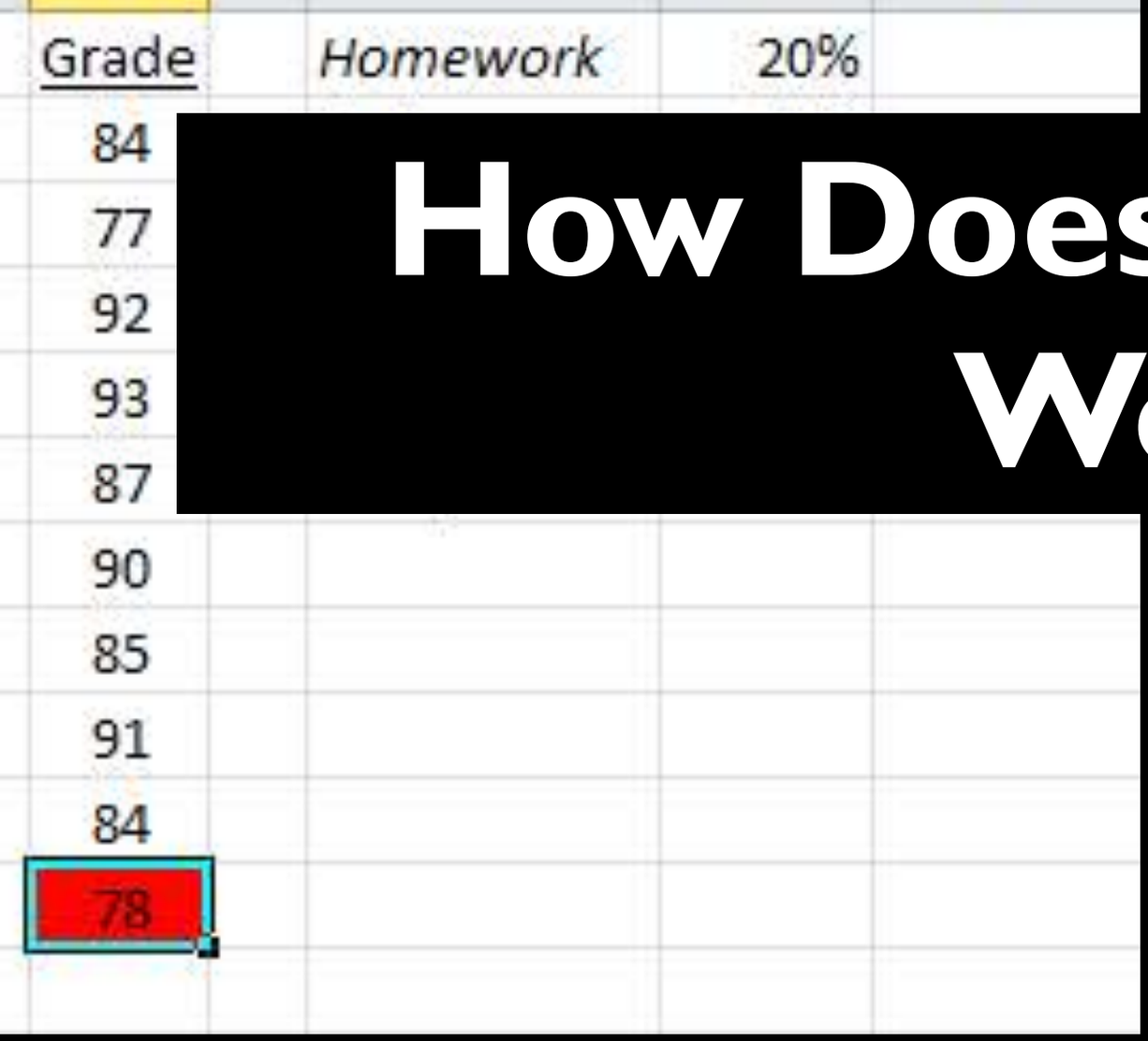
B11 | fx | 78

	A	B	C	D	E	F
1	Assignment	Grade		*Homework*	20%	
2	HW 1	84		*Quizzes*	30%	
3	HW 2	77		*Exams*	50%	
4	HW 3	92				
5	HW 4	93		**Final grade**	84.275	
6	Quiz 1	87		**Pass/Fail**	Fail	
7	Quiz 2	90				
8	Quiz 3	85				
9	Quiz 4	91				
10	Exam 1	84				
11	Exam 2	78				
12						



How Does CheckCell Work?

Combination of **dependence analysis** and **non-parametric statistical analysis**



How Does CheckCell Work?

Combination of **dependence analysis** and **non-parametric statistical analysis**

(assumes nothing about data distribution)

How Does CheckCell Work?

Assignment	Grade		*Homework*	20%
HW 1	84		*Quizzes*	30%
HW 2	77		*Exams*	50%
HW 3	92			
HW 4	93		**Final grade**	84.275
Quiz 1	87		**Pass/Fail**	Fail
Quiz 2	90			
Quiz 3	85			
Quiz 4	91			
Exam 1	84			
Exam 2	78			

Dependence Analysis

Assignment	Grade		*Homework*	20%
HW 1	84		*Quizzes*	30%
HW 2	77		*Exams*	50%
HW 3	92			
HW 4	93		**Final grade**	84.275
Quiz 1	87		**Pass/Fail**	Fail
Quiz 2	90			
Quiz 3	85			
Quiz 4	91			
Exam 1	84			
Exam 2	78			

Dependence Analysis

Assignment	Grade		*Homework*	20%
HW 1	84		*Quizzes*	30%
HW 2	77		*Exams*	50%
HW 3	92			
HW 4	93		**Final grade**	84.275
Quiz 1	87		**Pass/Fail**	Fail
Quiz 2	90			
Quiz 3	85			
Quiz 4	91			
Exam 1	84			
Exam 2	78			

Dependence Analysis

Assignment	Grade		*Homework*	20%
HW 1	84		*Quizzes*	30%
HW 2	77		*Exams*	50%
HW 3	92			
HW 4	93		**Final grade**	84.275
Quiz 1	87		**Pass/Fail**	Fail
Quiz 2	90			
Quiz 3	85			
Quiz 4	91			
Exam 1	84			
Exam 2	78			

Dependence Analysis

Assignment	Grade		*Homework*	20%
HW 1	84		*Quizzes*	30%
HW 2	77		*Exams*	50%
HW 3	92			
HW 4	93		**Final grade**	84.275
Quiz 1	87		**Pass/Fail**	Fail
Quiz 2	90			
Quiz 3	85			
Quiz 4	91			
Exam 1	84			
Exam 2	78			

Statistical Analysis

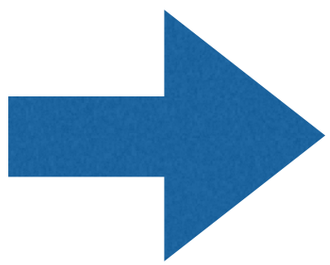
Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

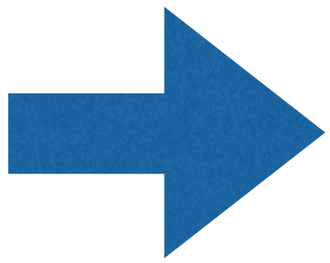


93
92
84
84

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

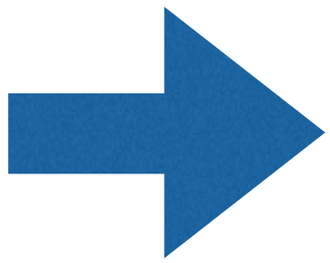


93	84
92	93
84	93
84	92

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

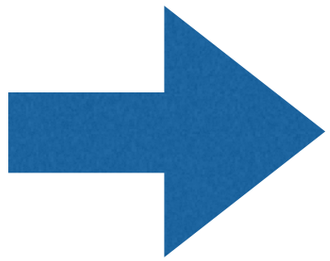


93	84	93
92	93	93
84	93	92
84	92	93

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

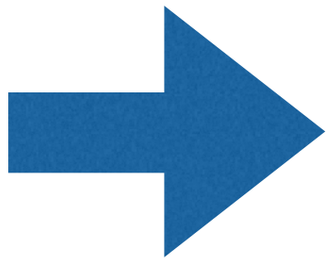


93	84	93	84
92	93	93	93
84	93	92	84
84	92	93	92

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93

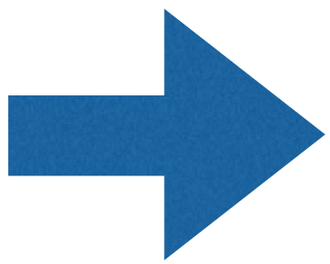


93	84	93	84
92	93	93	93
84	93	92	84
84	92	93	92

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

84
77
92
93



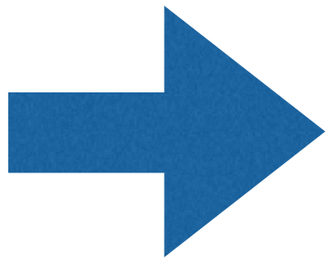
93	84	93	84
92	93	93	93
84	93	92	84
84	92	93	92

Can do statistical tests over single sample

Statistical Analysis

Adaptation of Efron's **bootstrap method** (resampling)

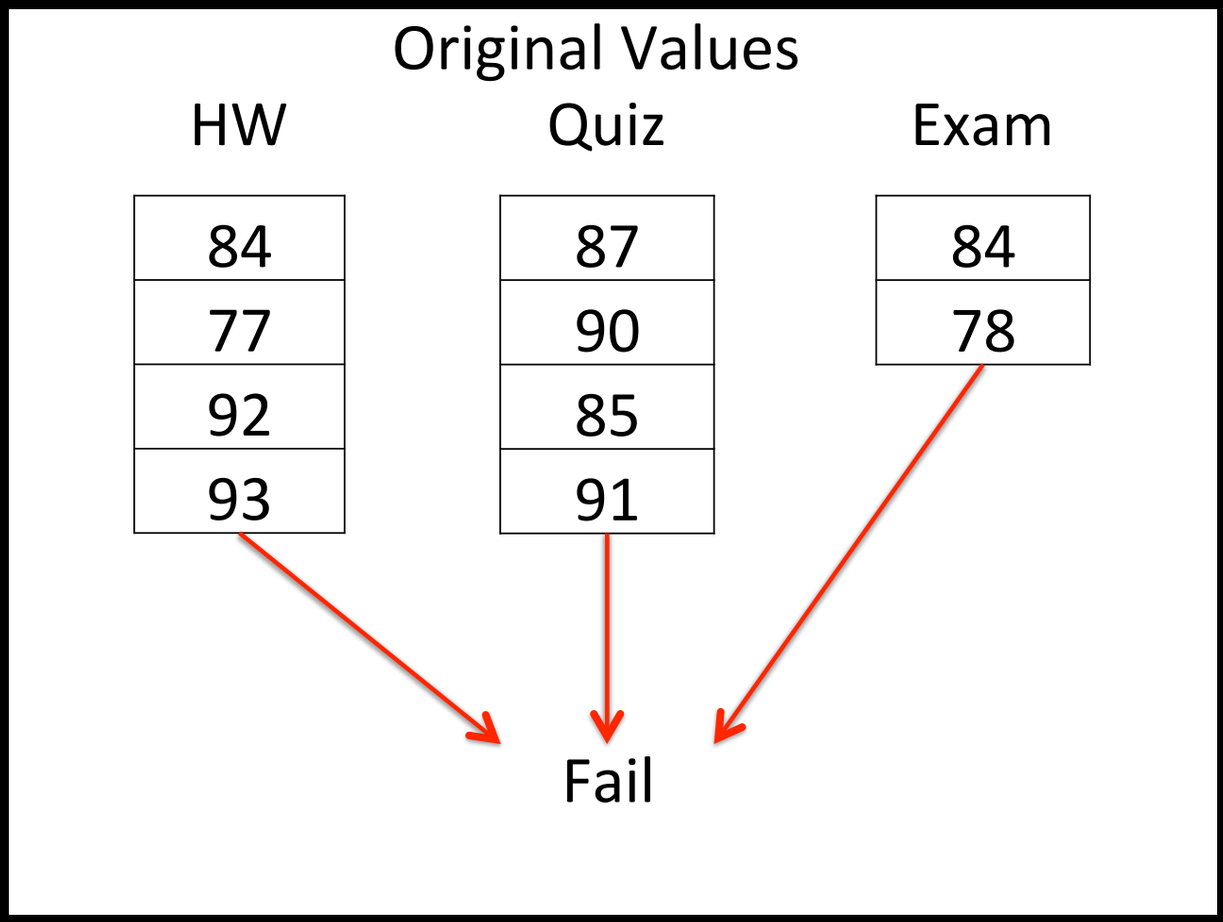
84
77
92
93



93	84	93	84
92	93	93	93
84	93	92	84
84	92	93	92

Can do statistical tests over single sample
Extends to non-numerical data

Guided Bootstrap



Guided Bootstrap

Original Values

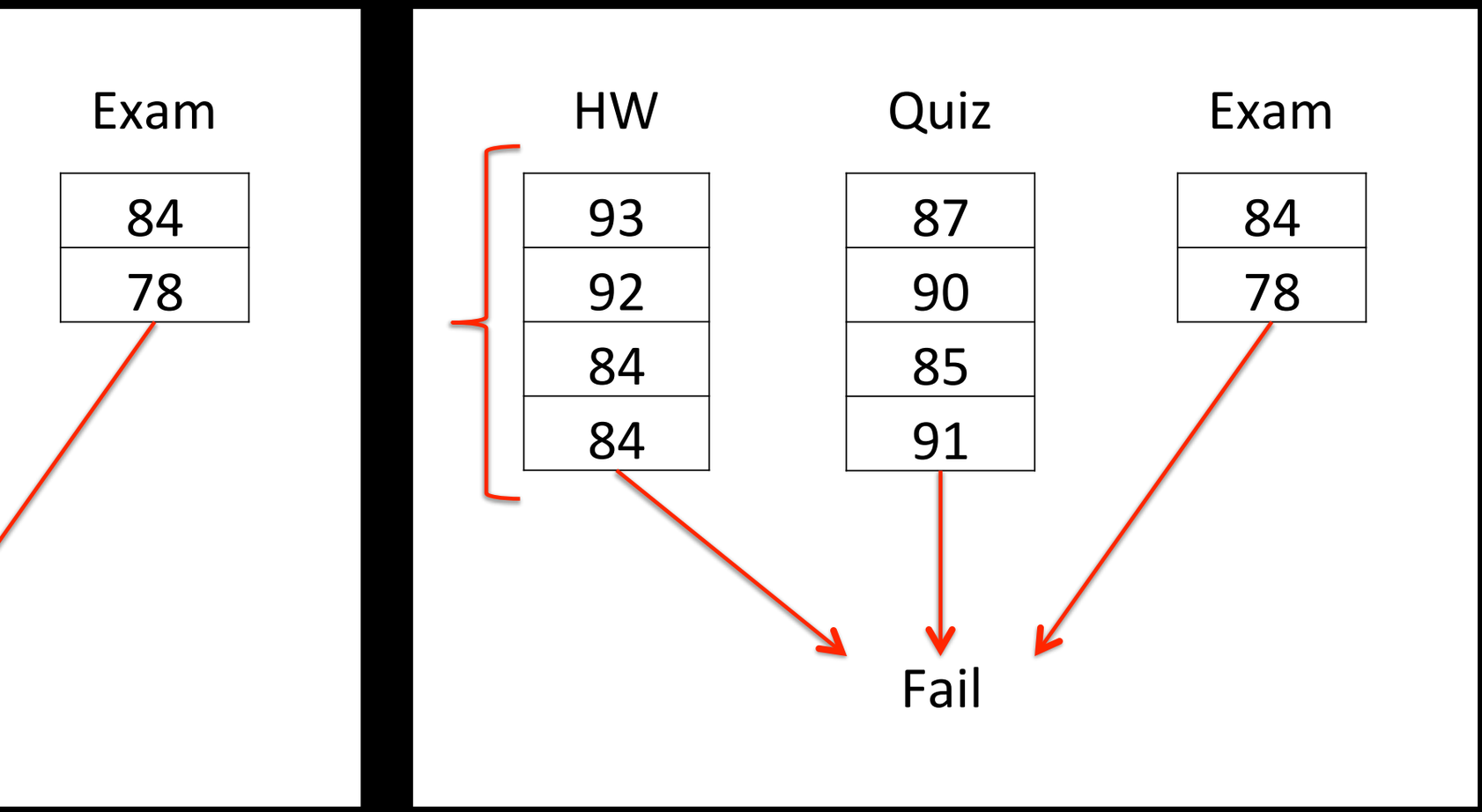
HW	Quiz	Exam
84	87	84
77	90	78
92	85	
93	91	

Fail

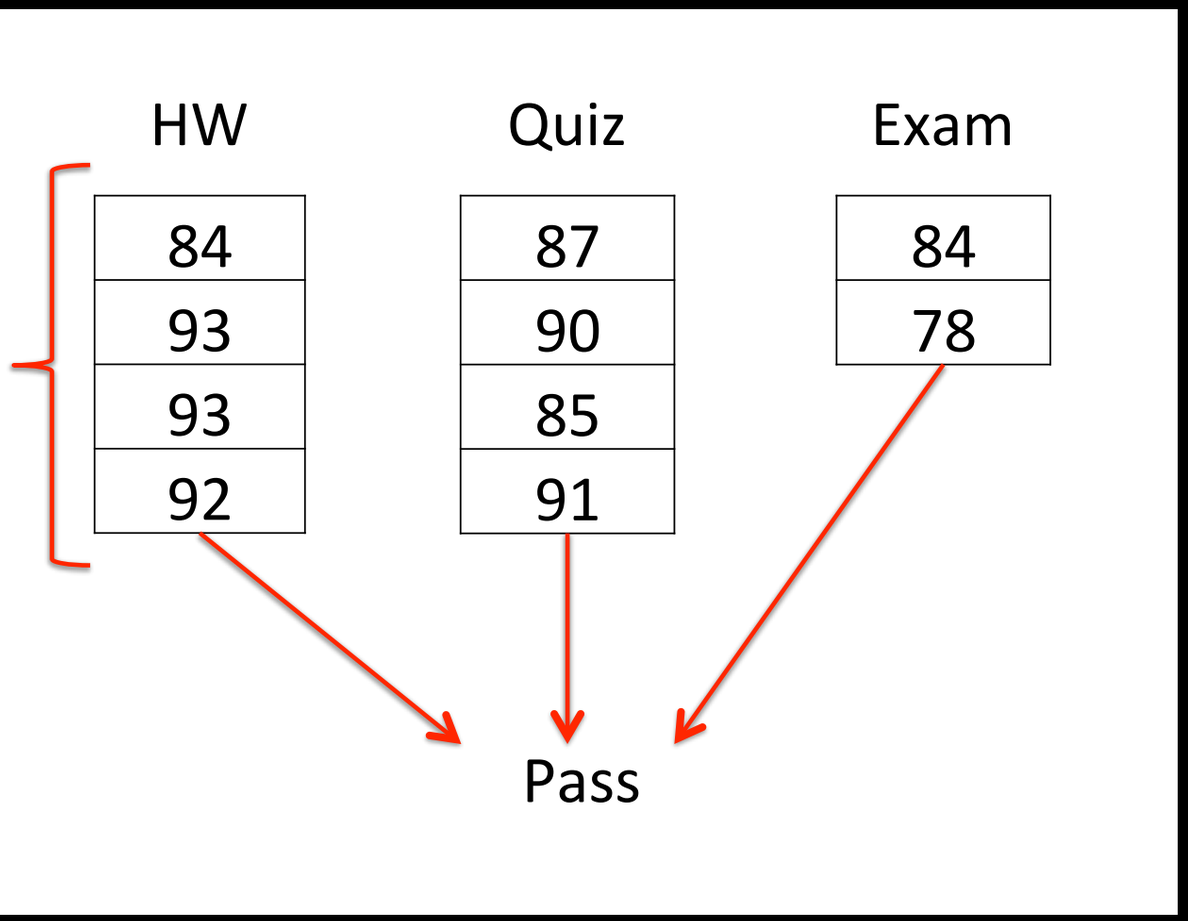
HW	Quiz	Exam
93	87	84
92	90	78
84	85	
84	91	

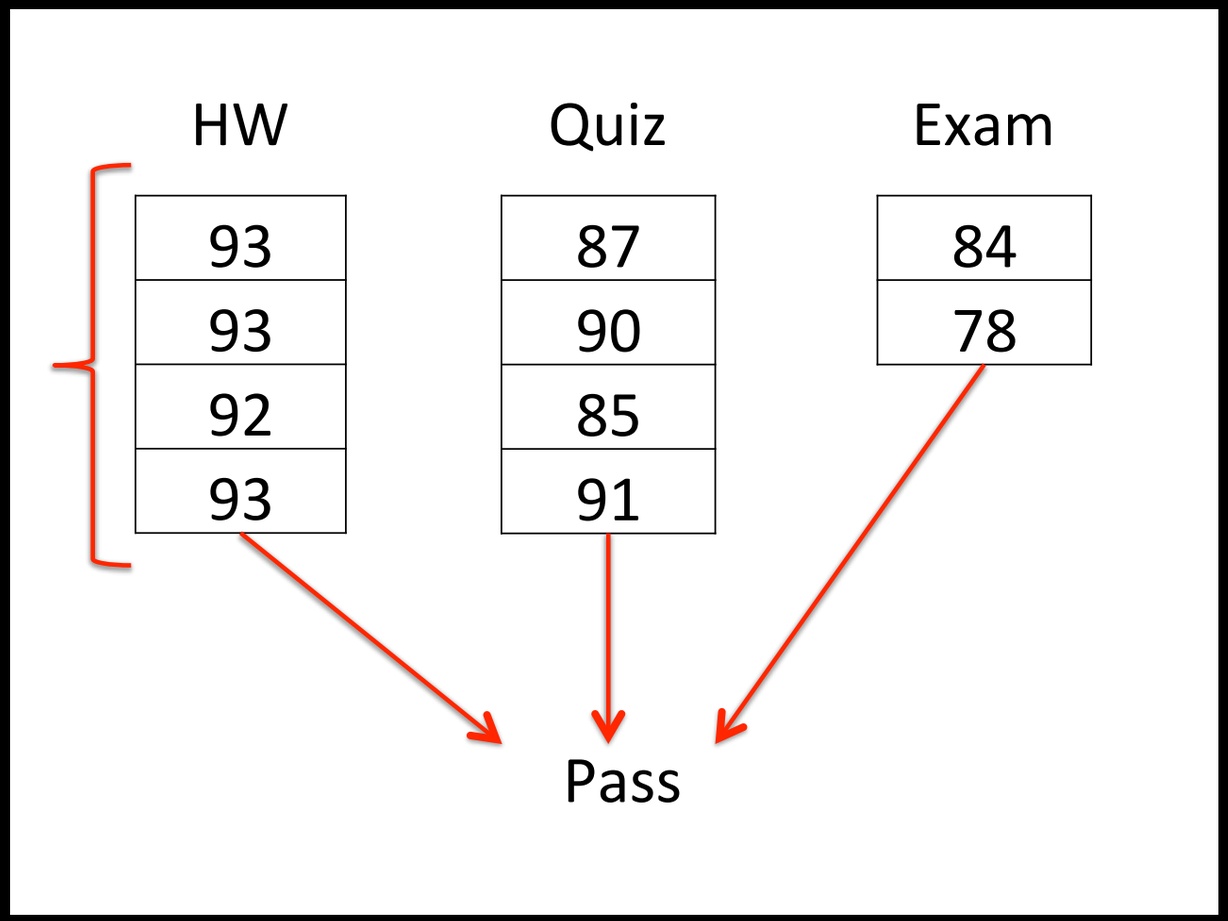
Fail

Guided Bootstrap

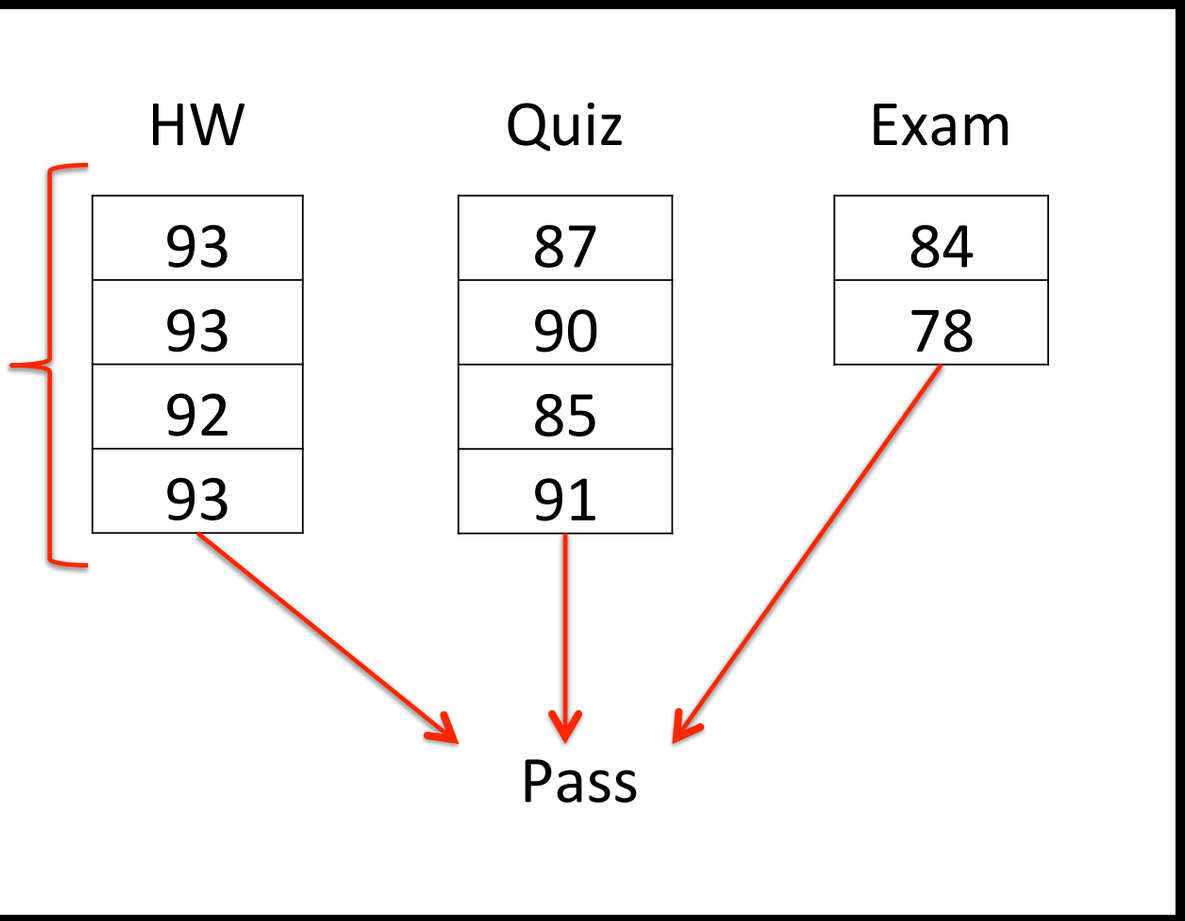


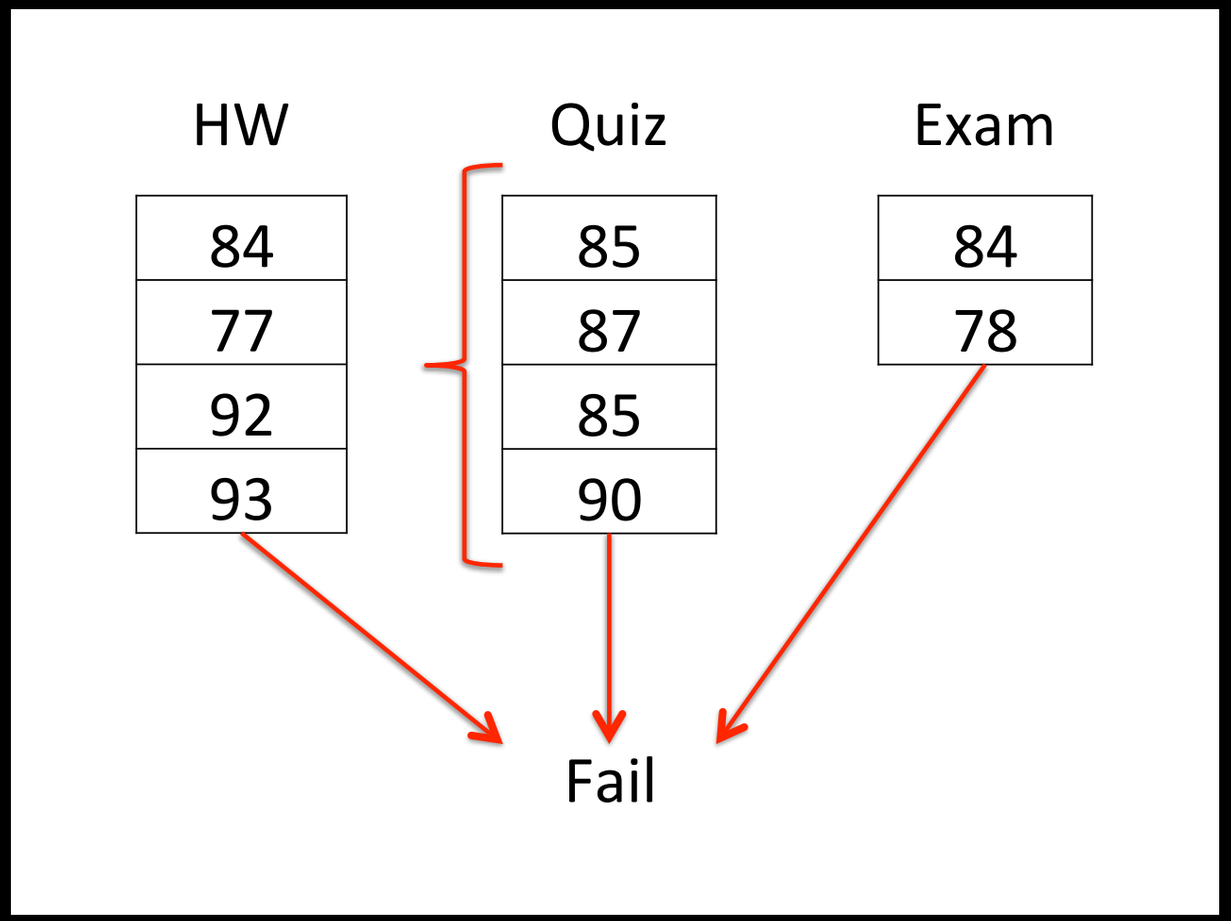
Guided Bootstrap



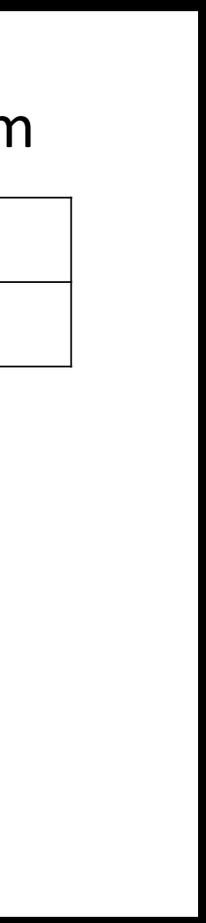


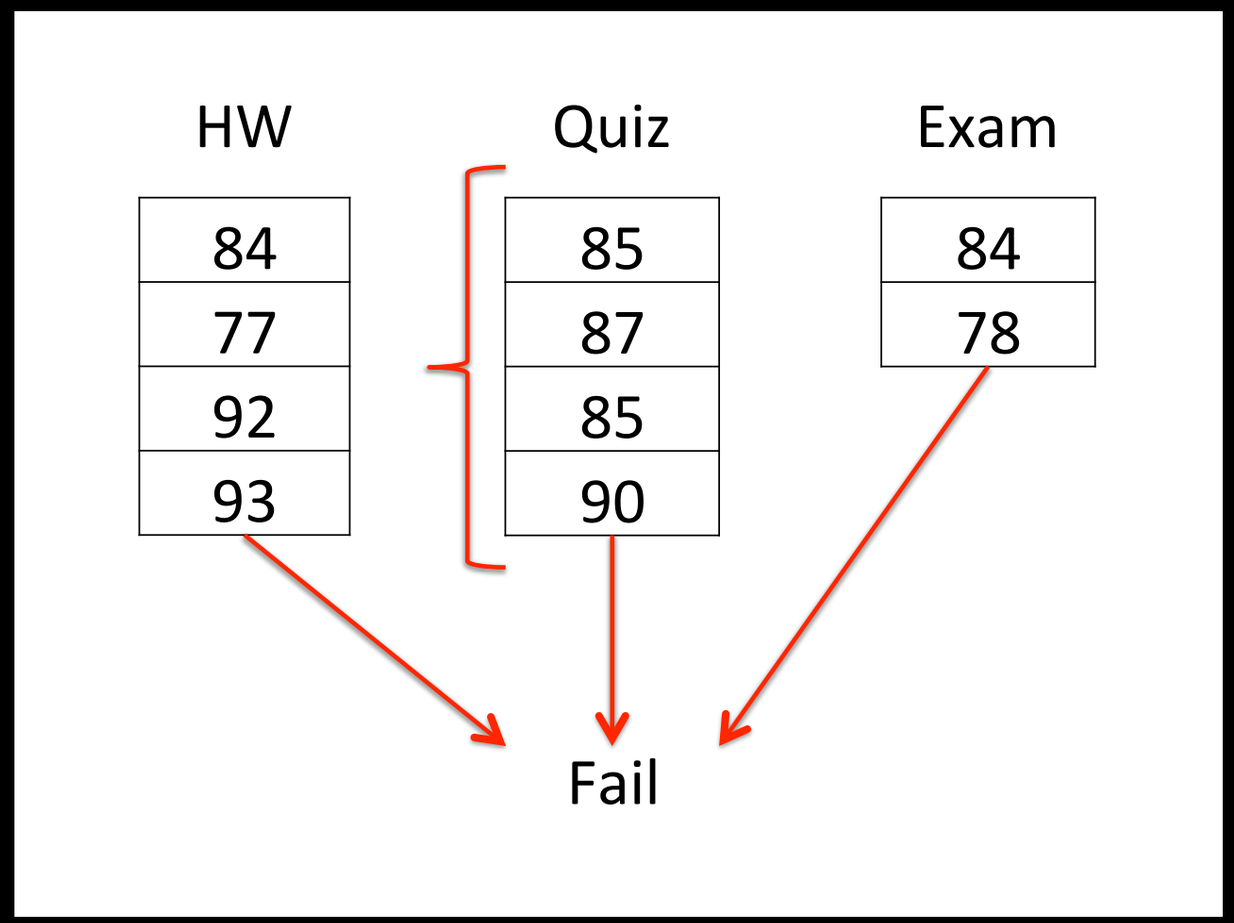
Guided Bootstrap

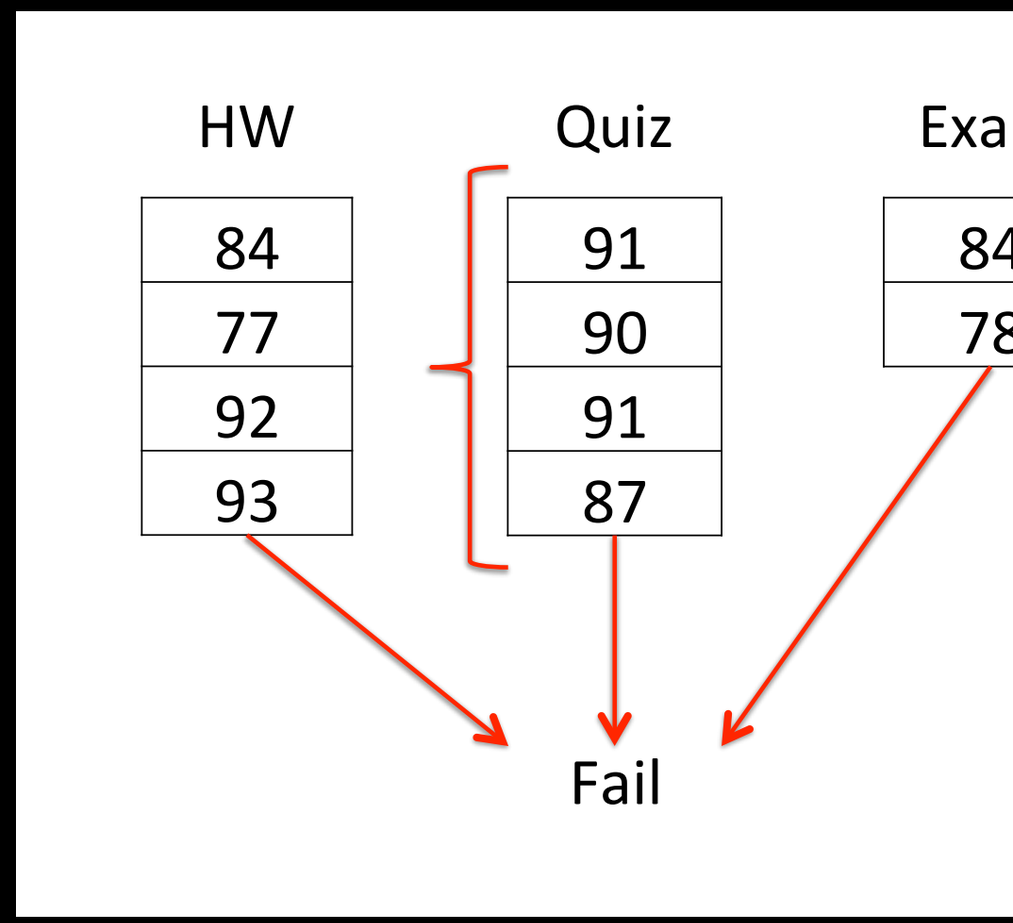




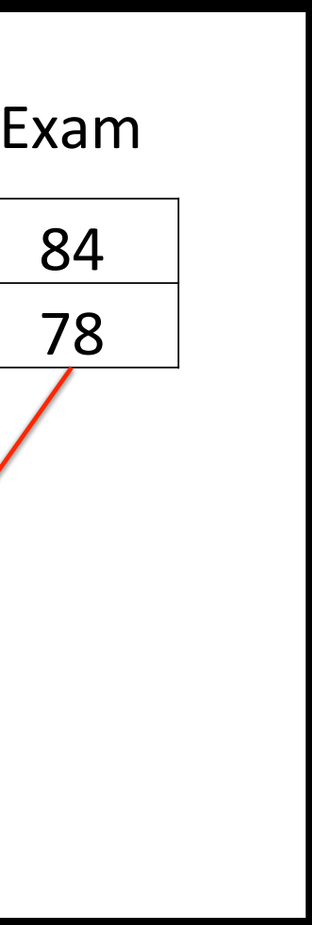
Guided Bootstrap

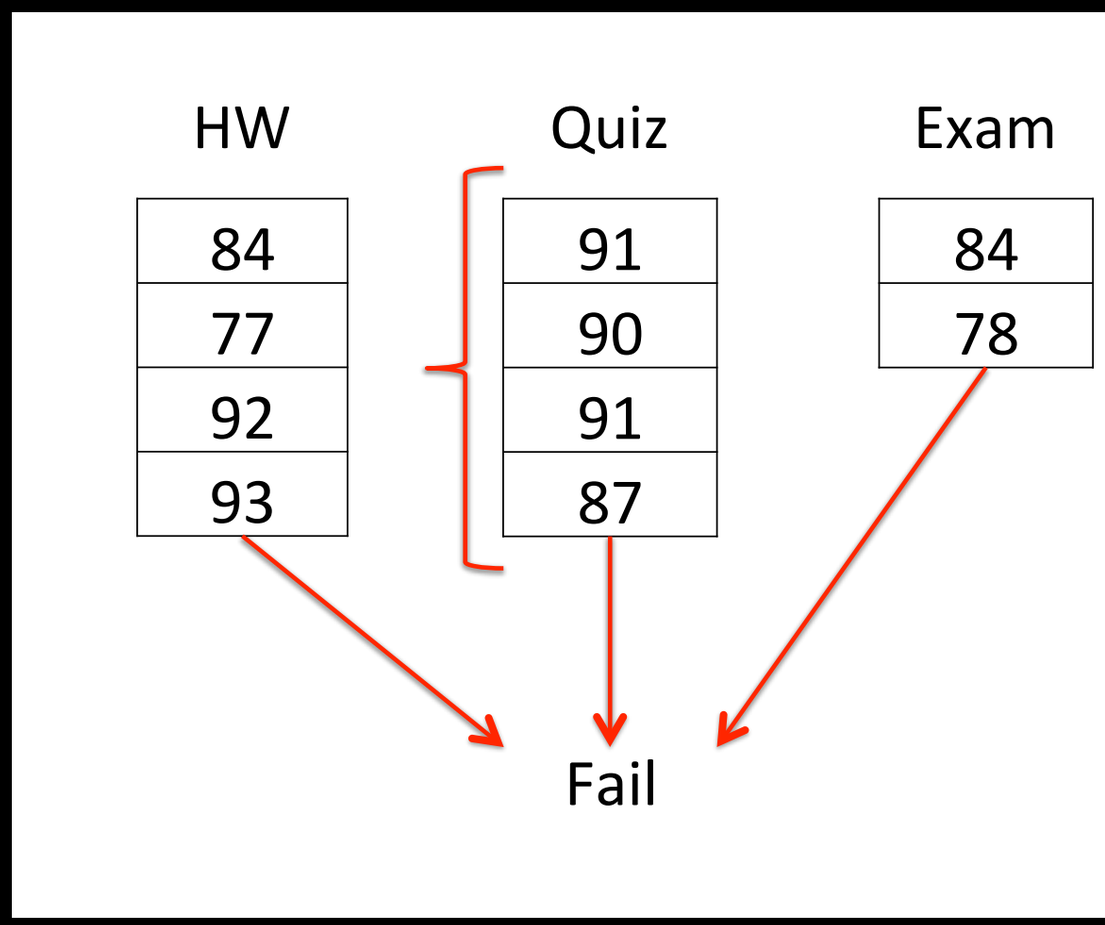


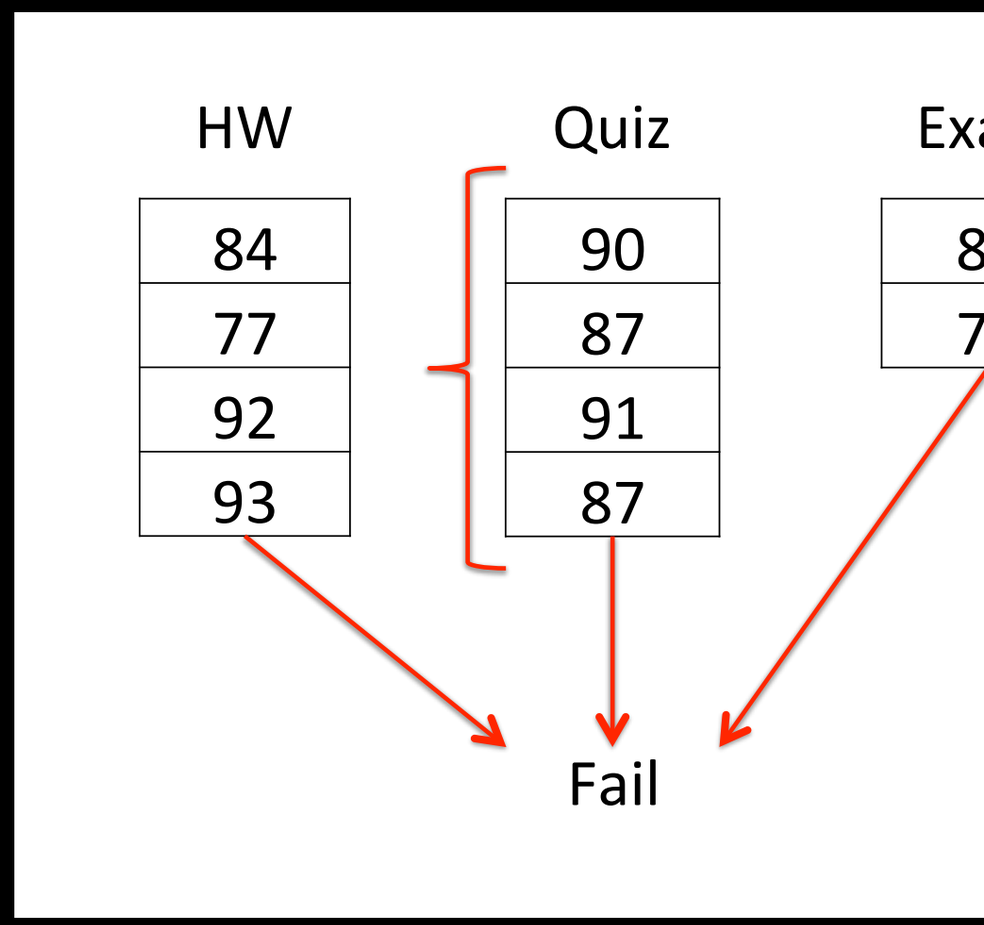




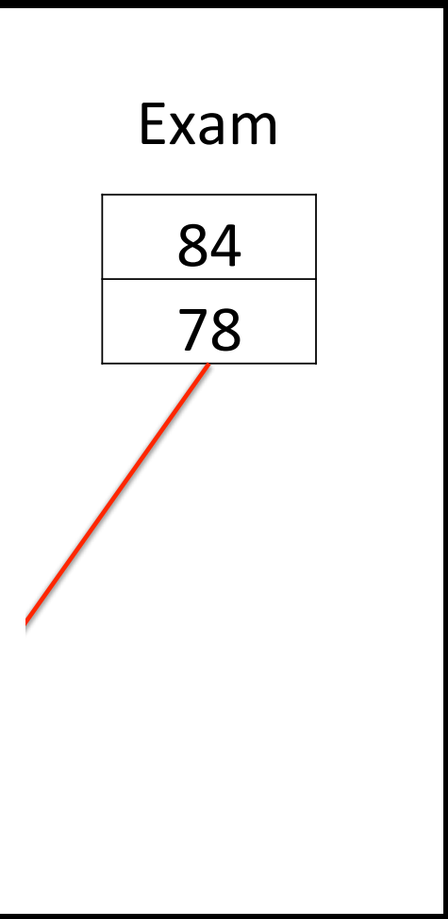
Guided Bootstrap

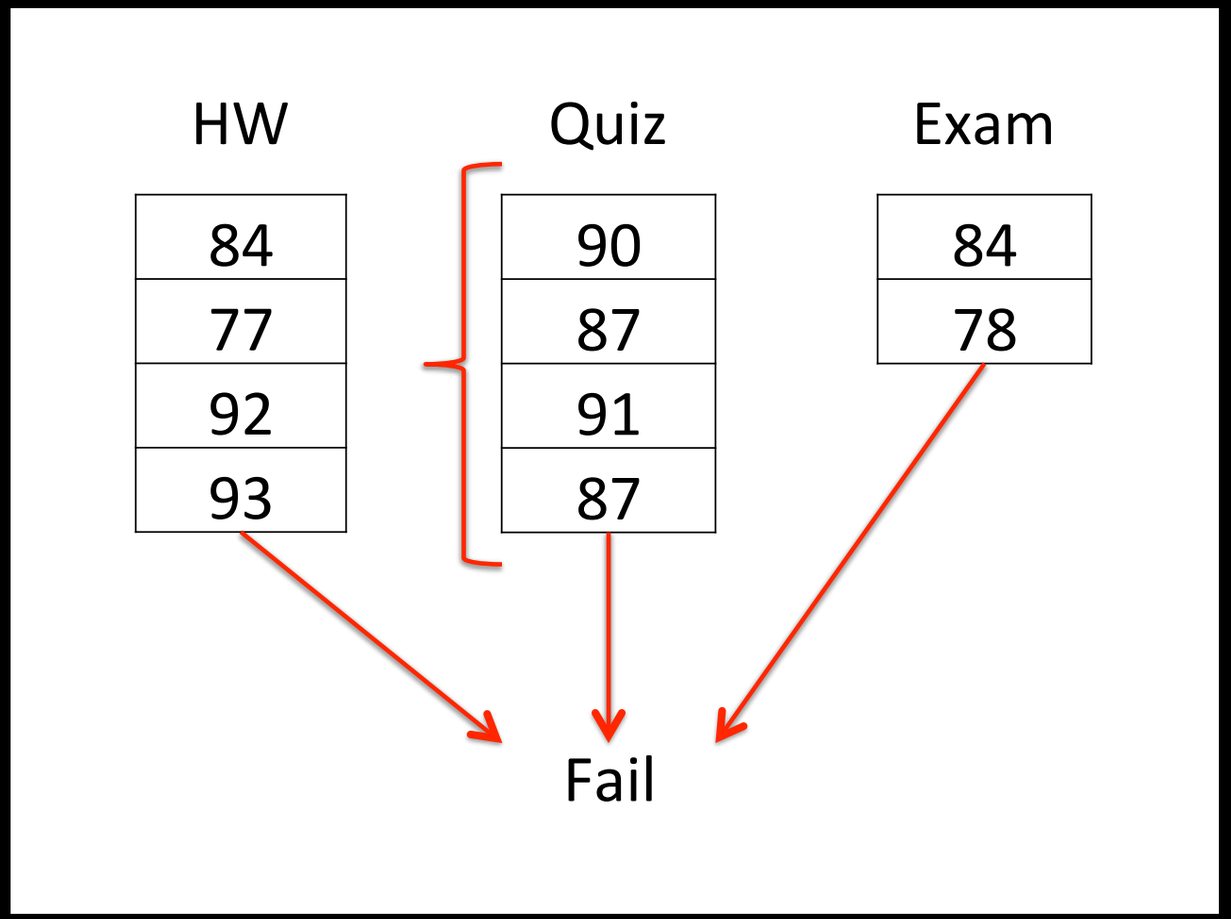


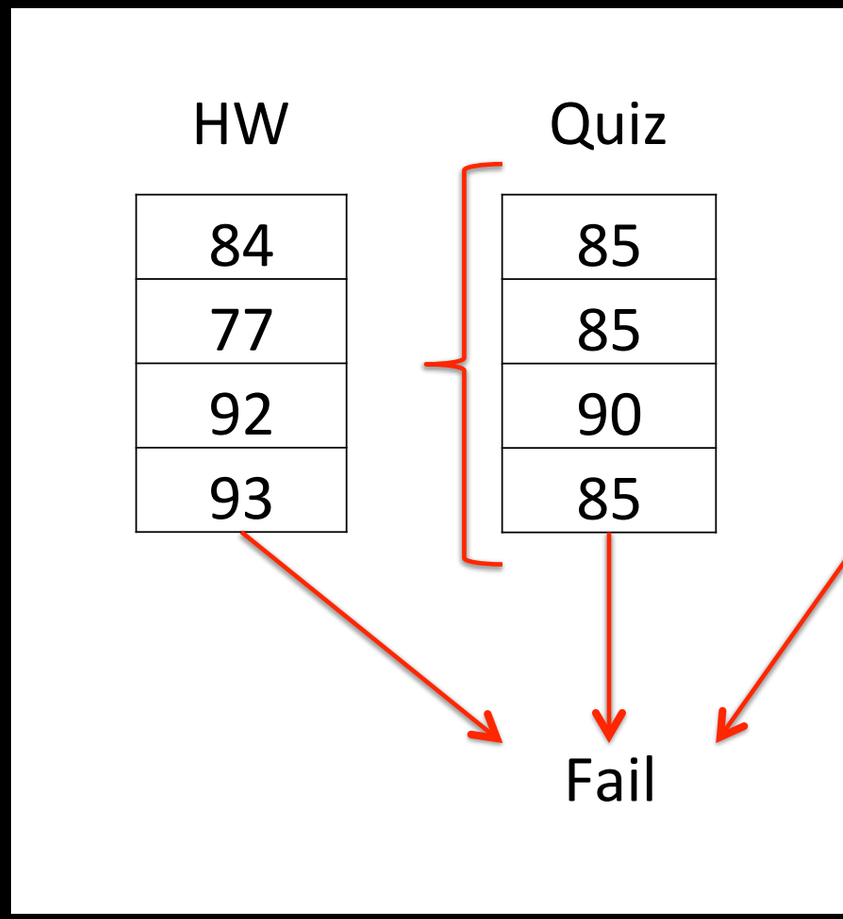




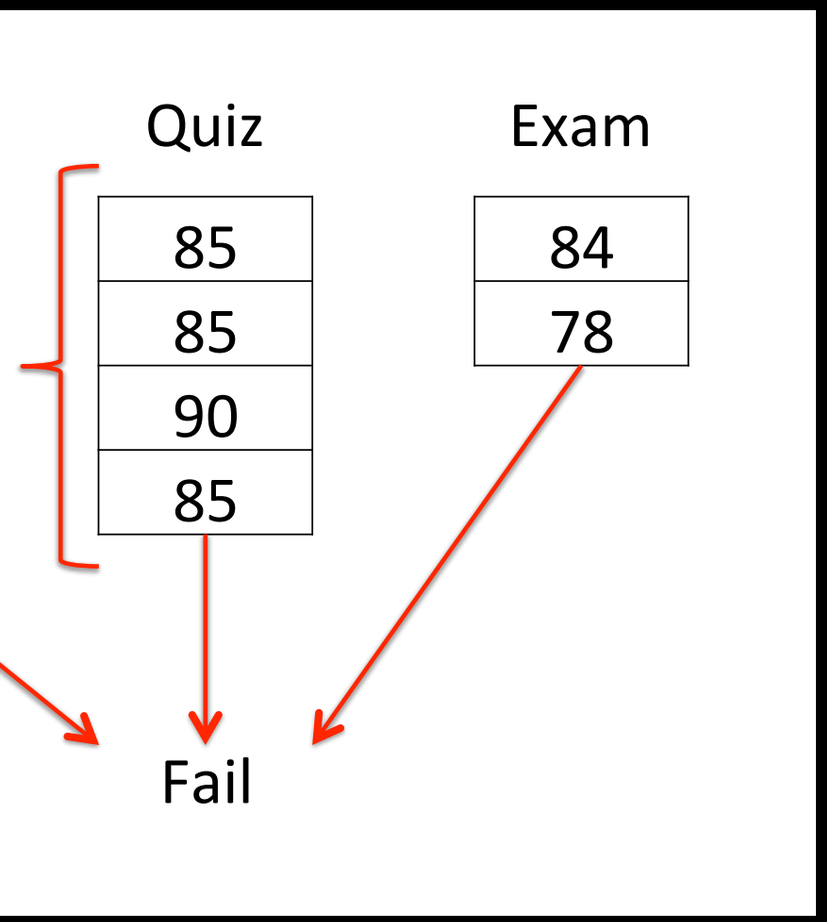
Guided Bootstrap

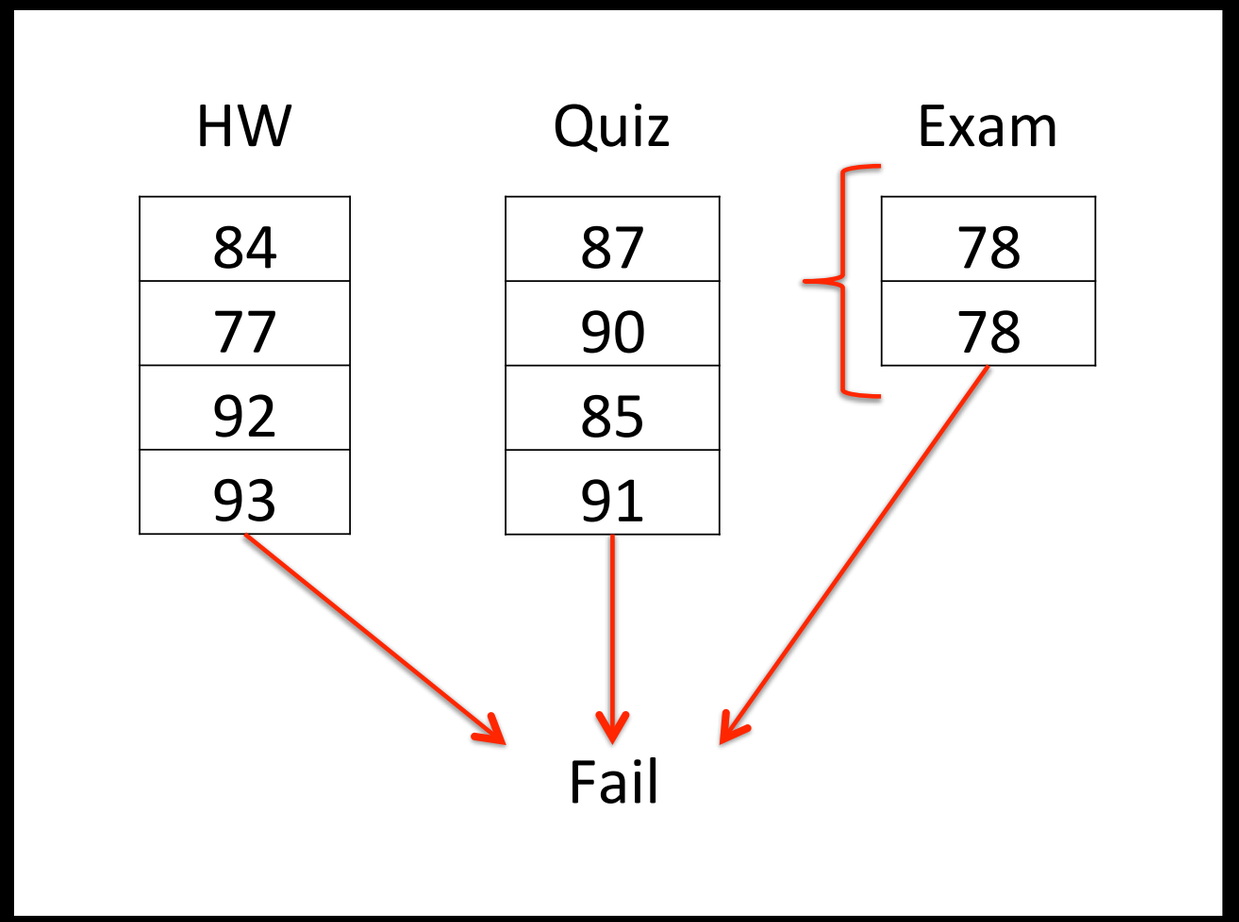




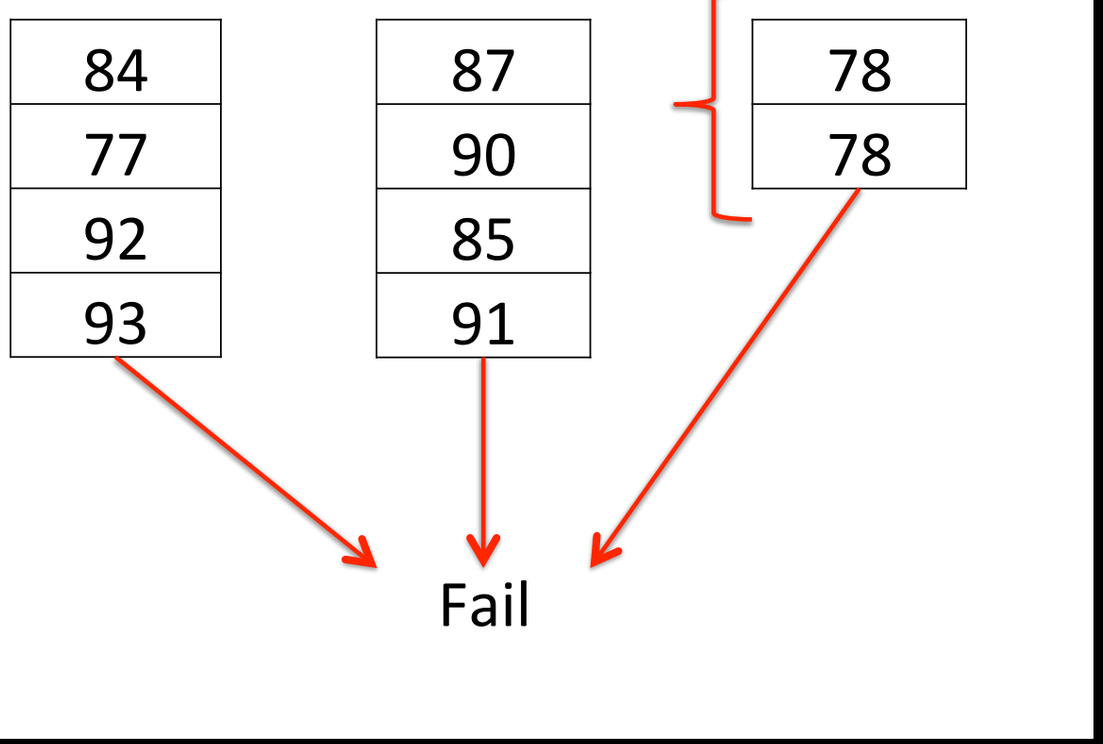


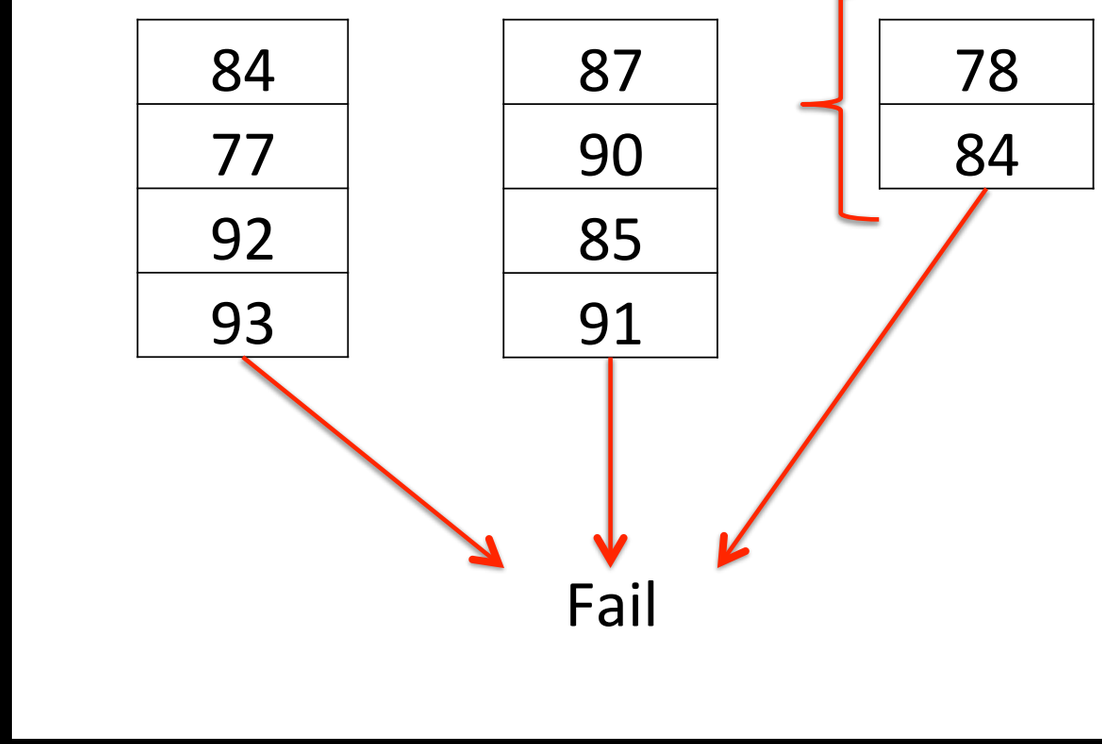
Guided Bootstrap



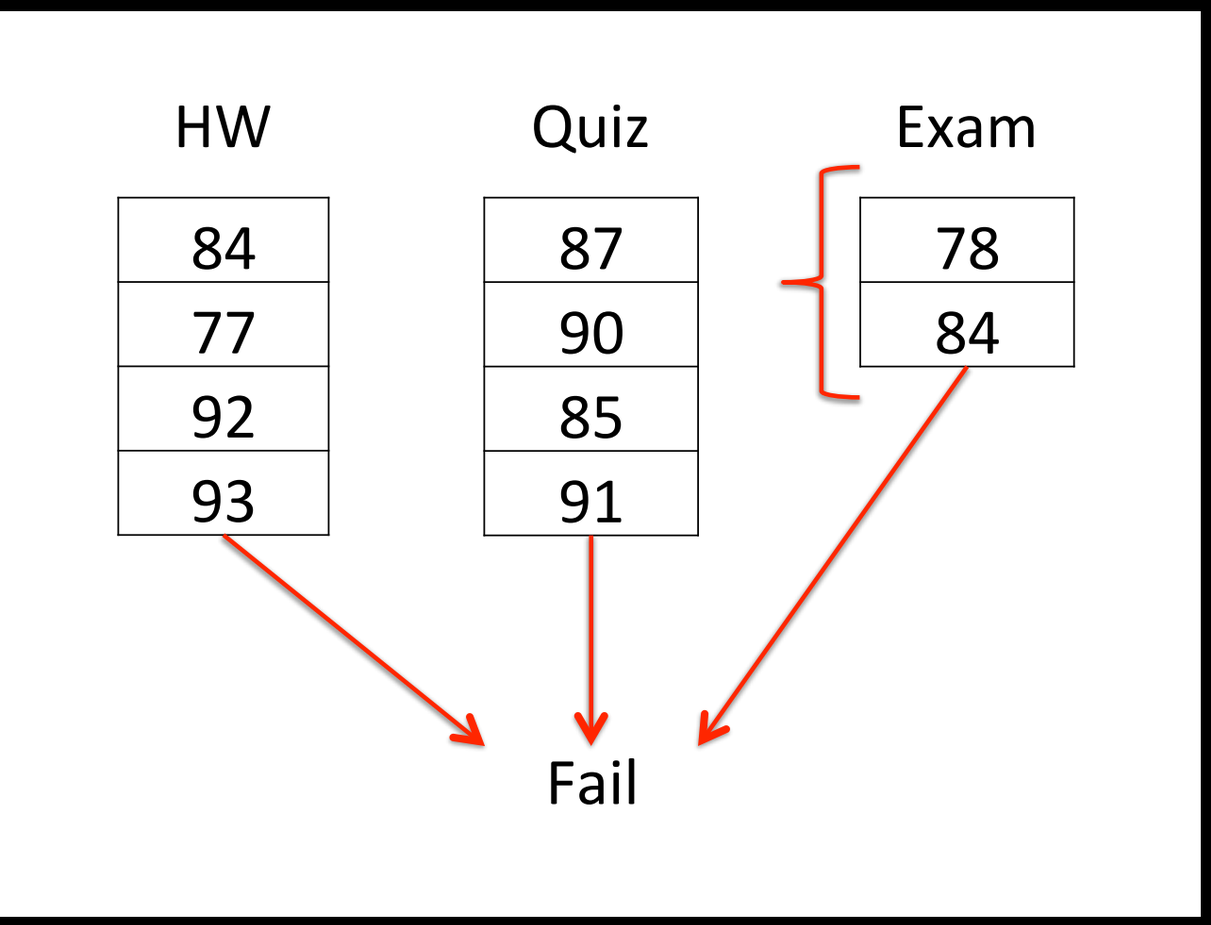


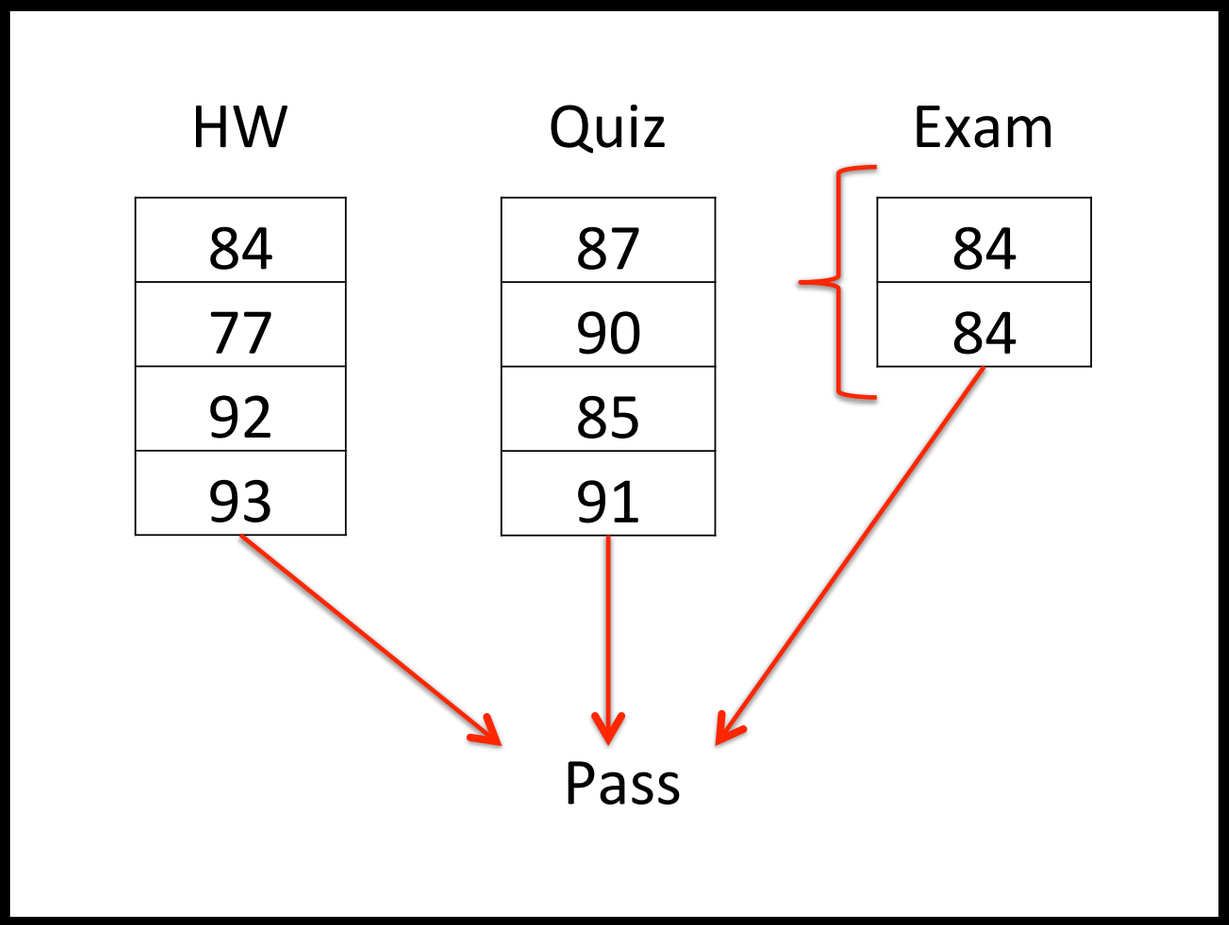
Guided Bootstrap



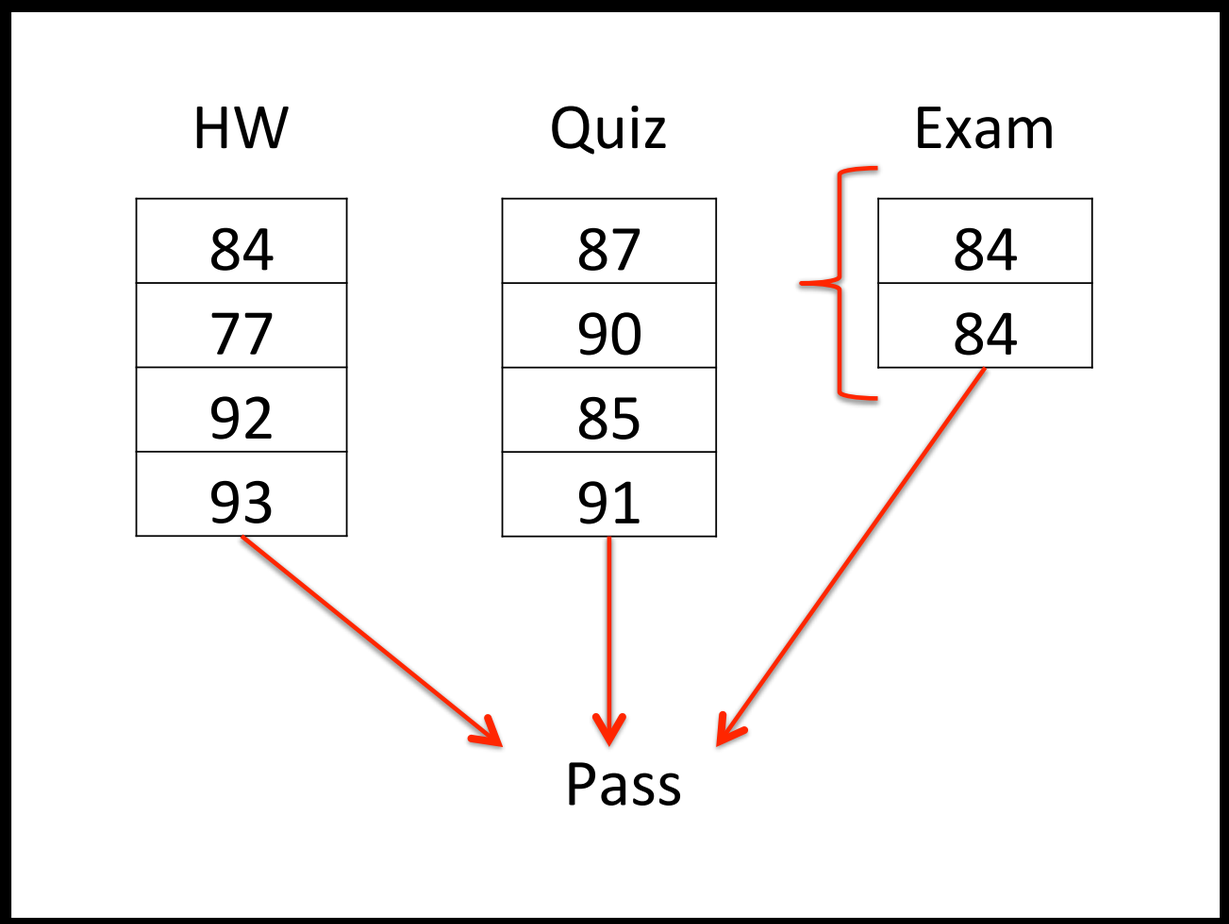


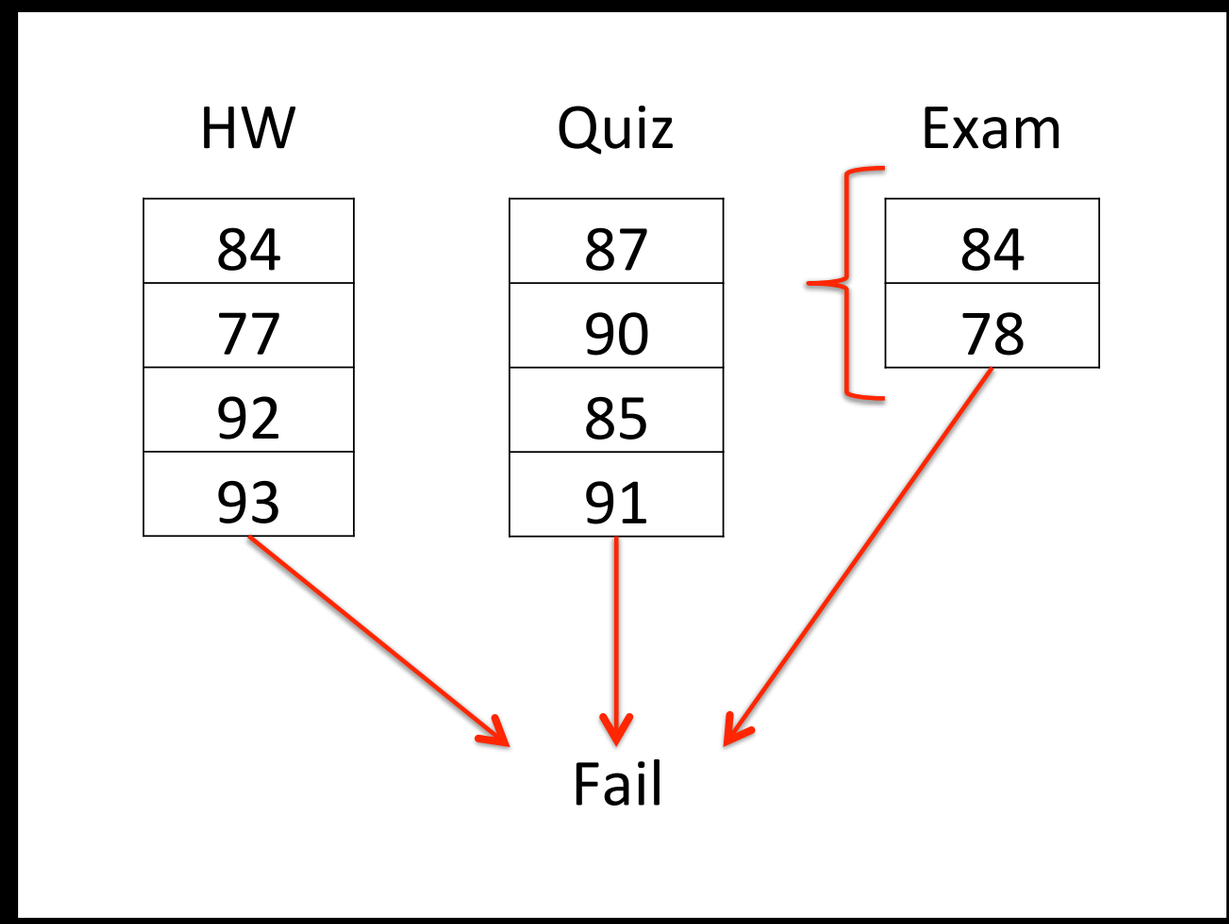
Guided Bootstrap



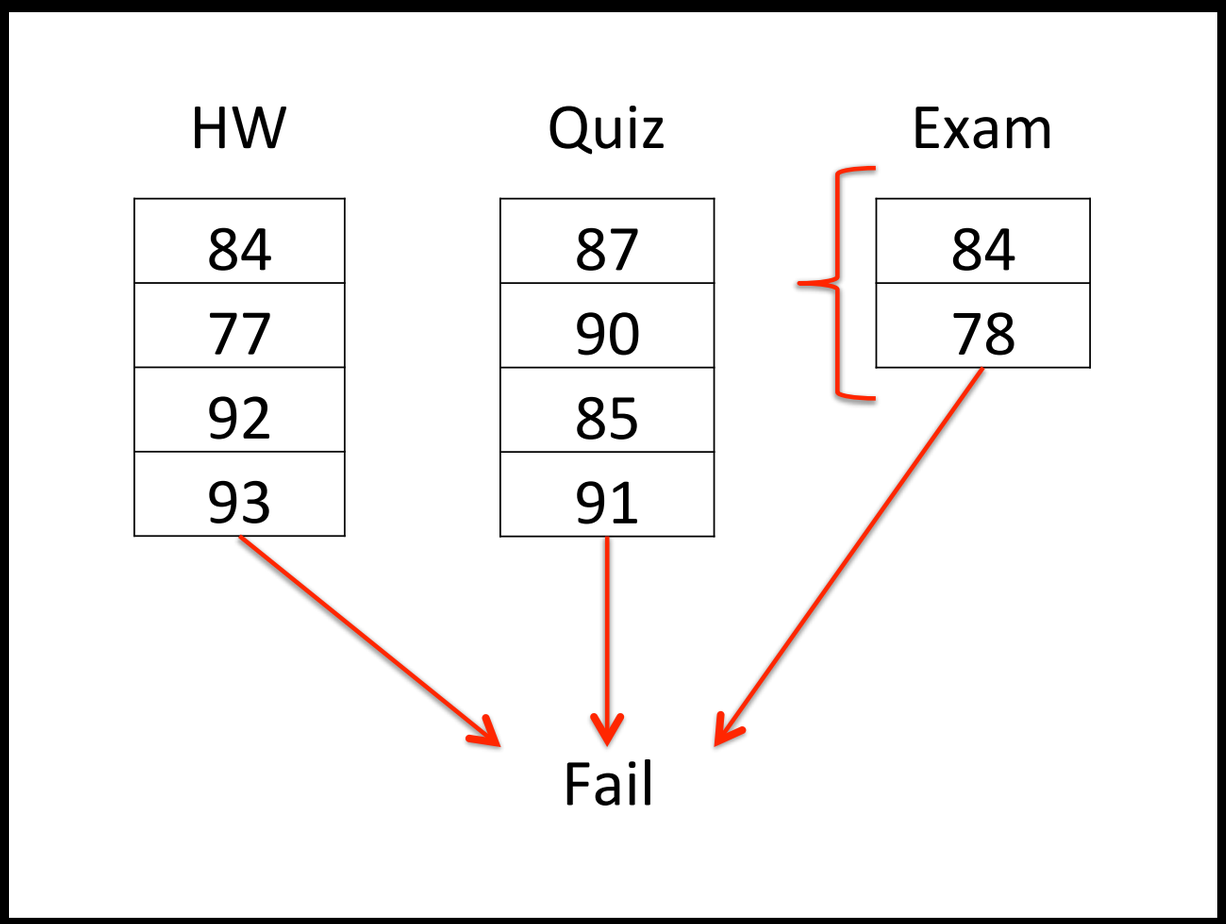


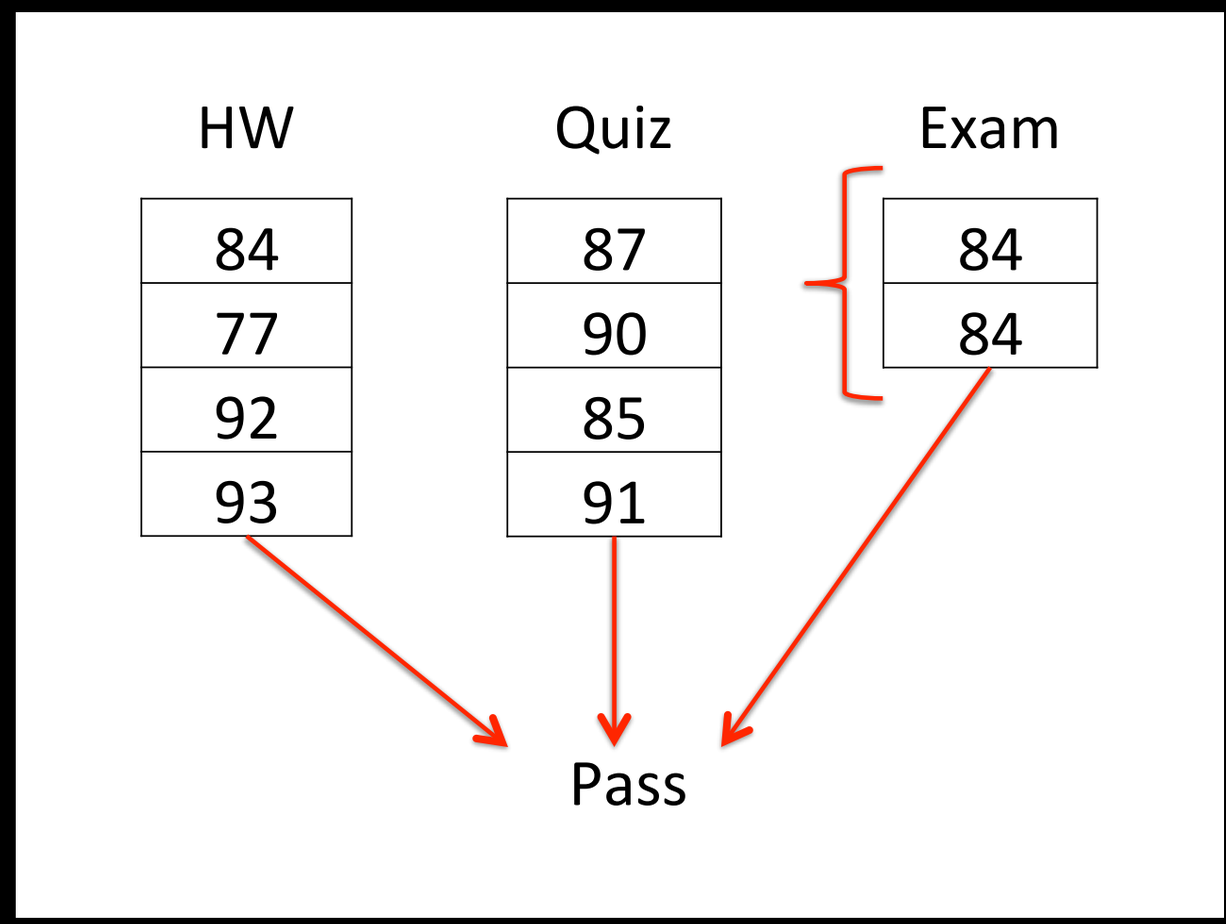
Guided Bootstrap



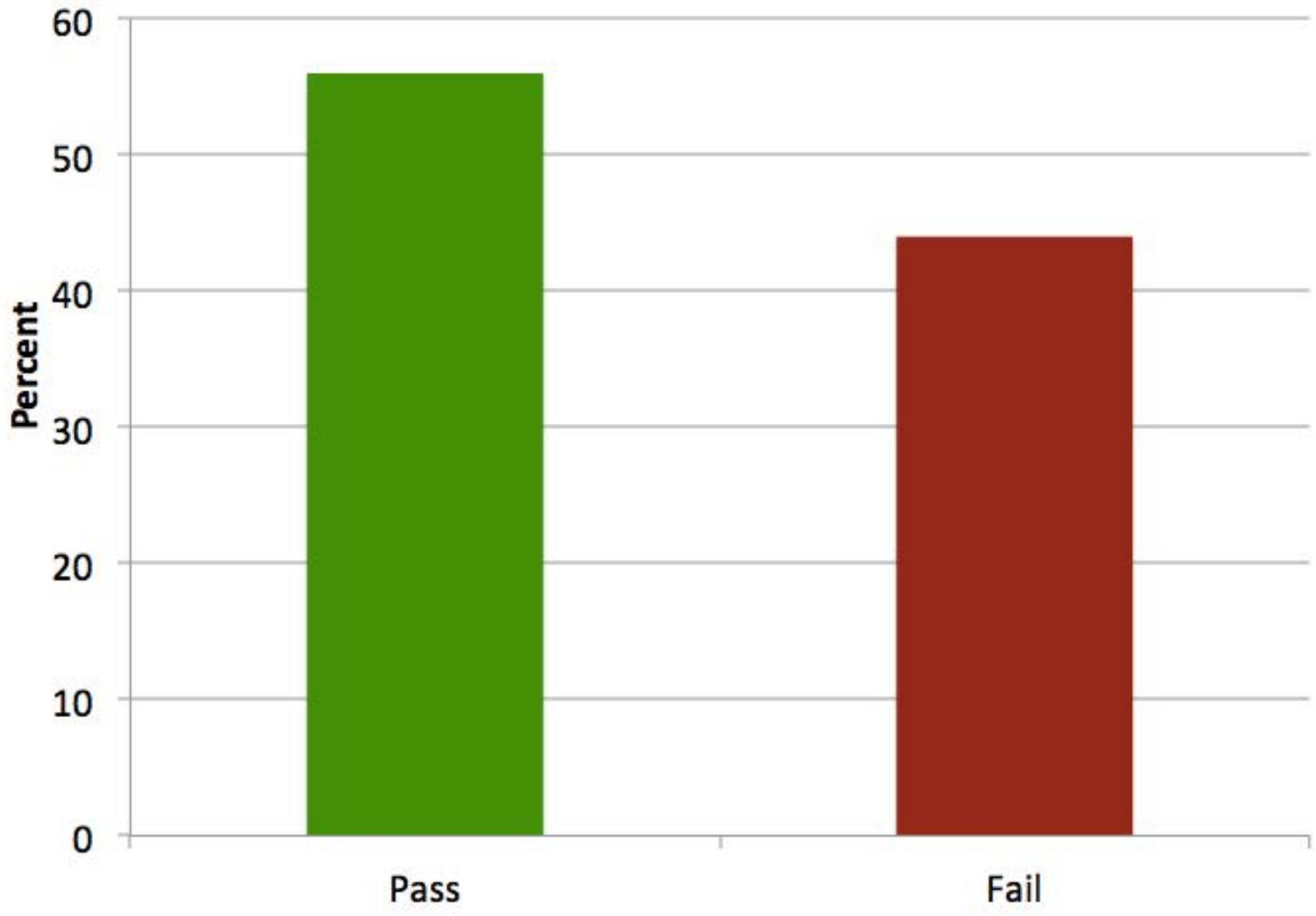


Guided Bootstrap

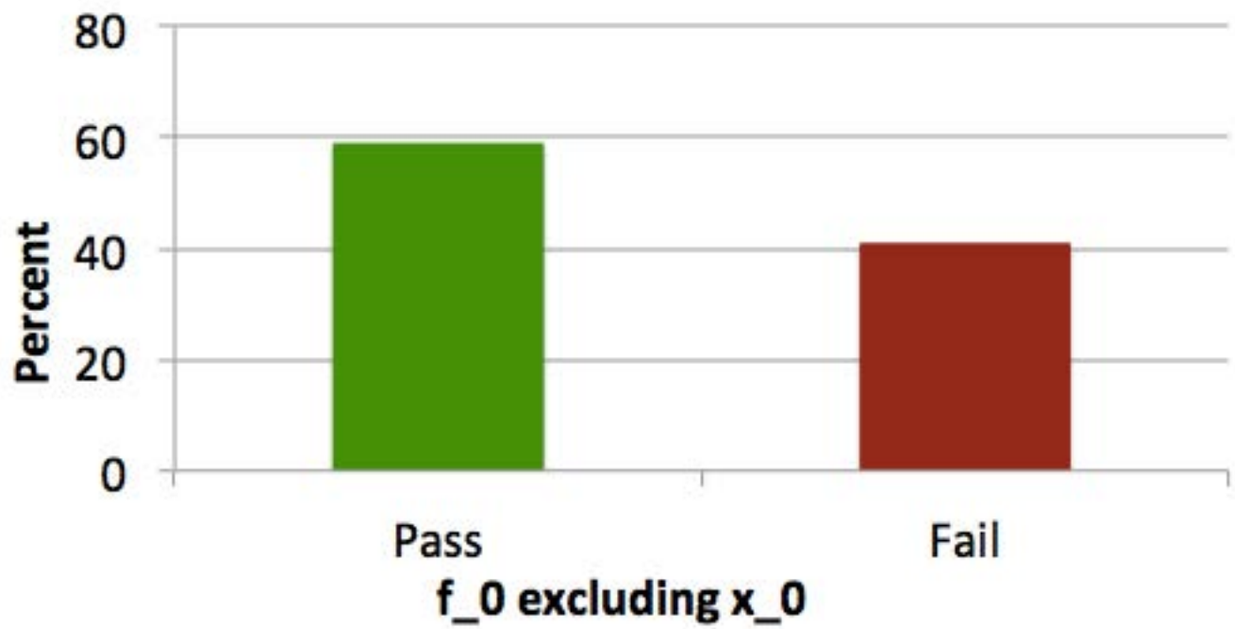




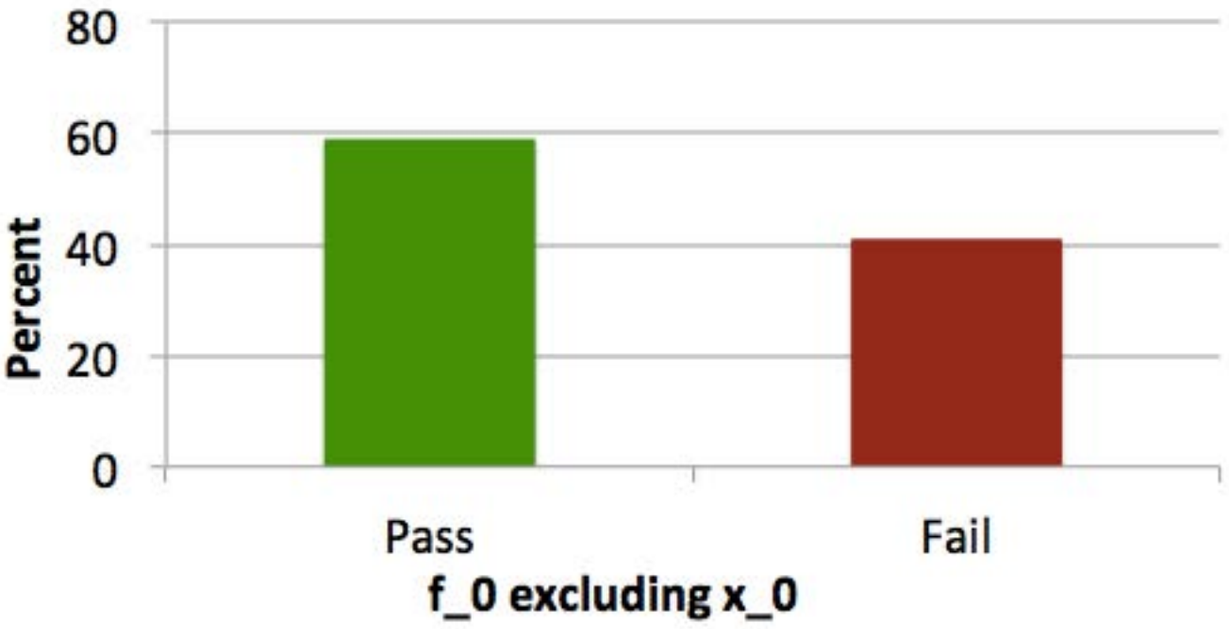
Outputs:
Empirical Distribution

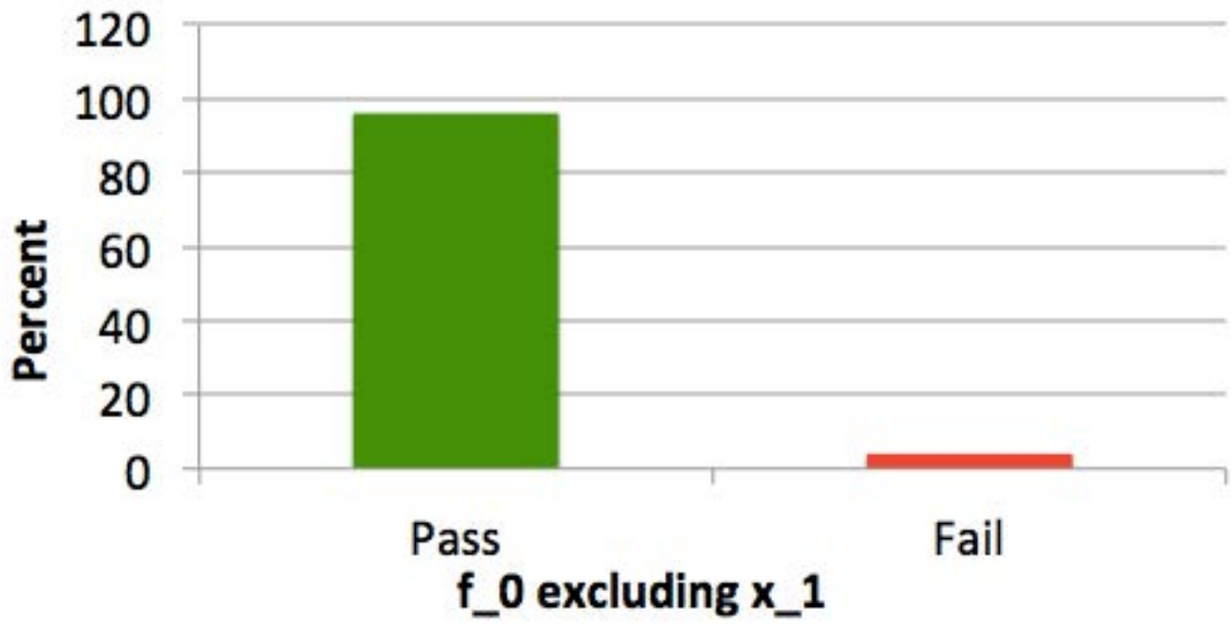


Condition on Presence/ Absence of Values

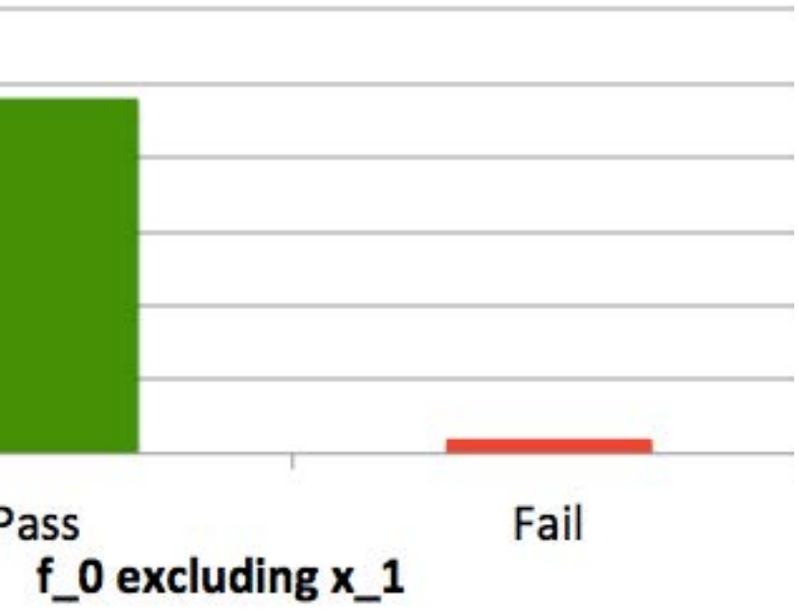


Condition on Presence/ Absence of Values





Input Unusual in Conditional Distribution?



input	score
0	0
1	1
2	0
3	0

Marginalize Over Functions

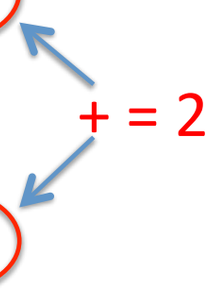
	score
0	0
1	1
2	0
3	0

function	input	score
0	0	0
0	1	1
0	2	0
0	3	0
1	0	0
1	1	1
1	2	0
1	3	0

+ = 2

Rank Results

input	score
0	0
1	1
2	0
3	0
0	0
1	1
2	0
3	0

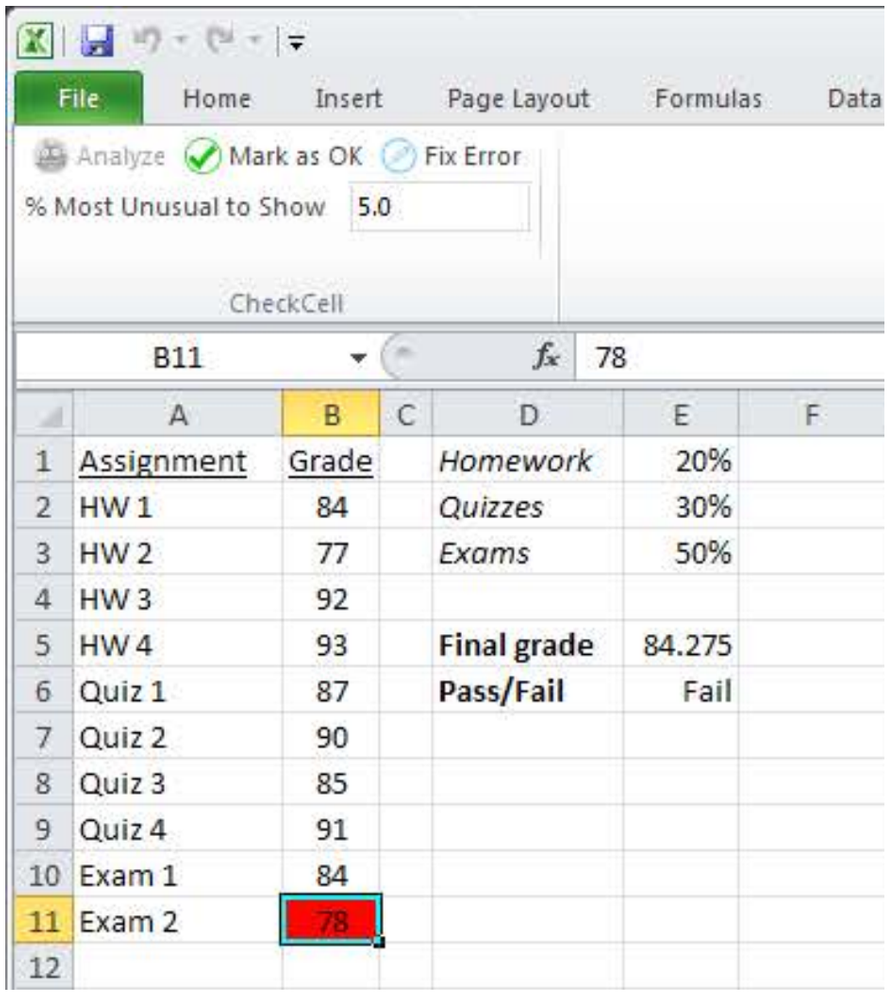


input	score
1	2
0	0
2	0
3	0

Rank Results

ut	score
1	2
0	0
2	0
3	0

most unusual



Empirical Evaluation

Challenges

Where do we find buggy spreadsheets?

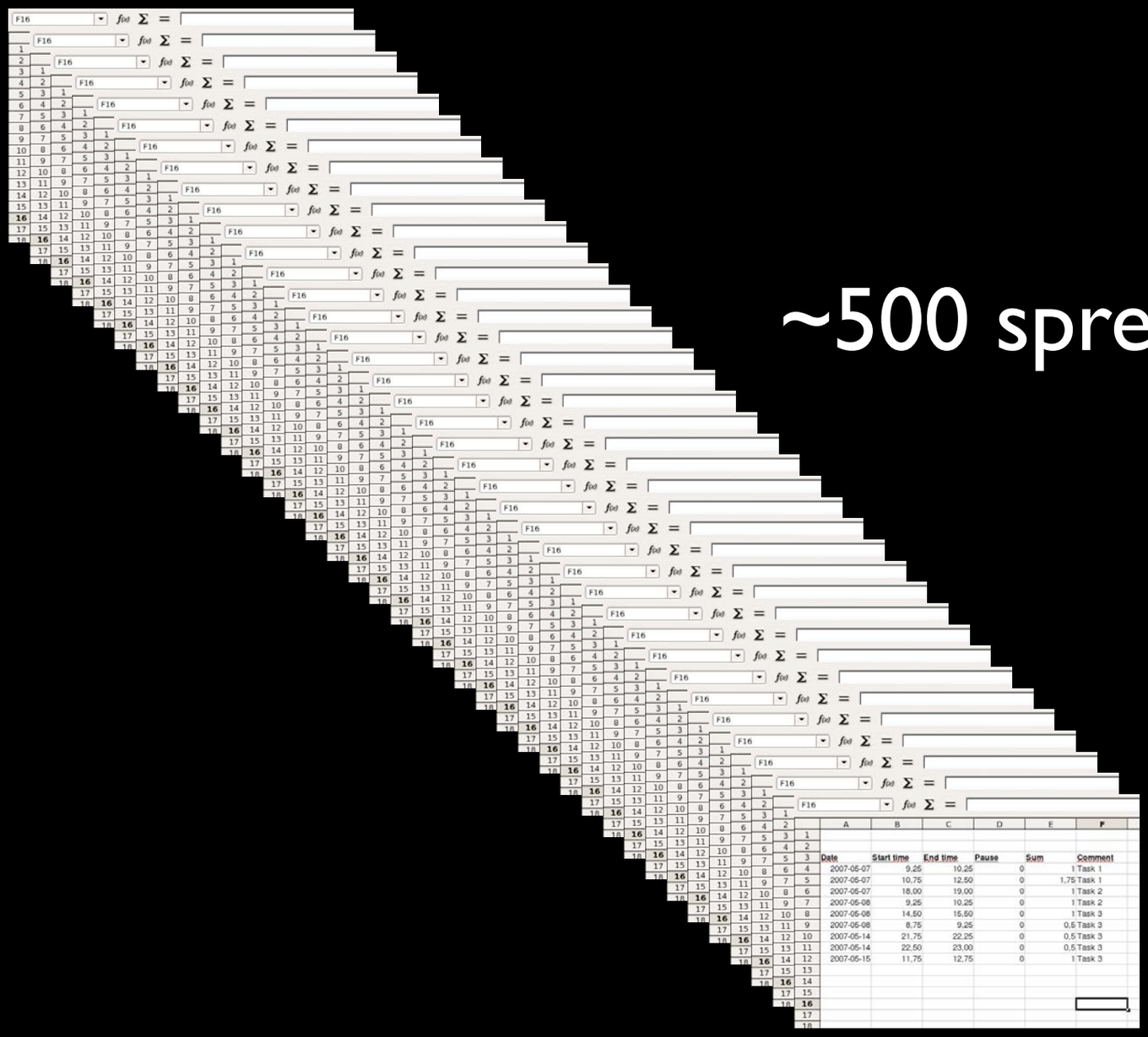
Challenges

Where do we find buggy spreadsheets?

How do we know what's **correct**?

Experimental Methodology

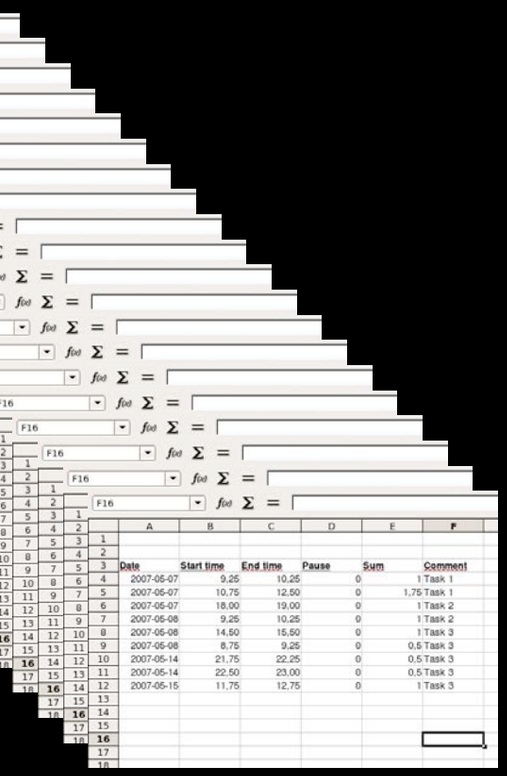
Experimental Methodology



~500 spreadsheets (EUSES)

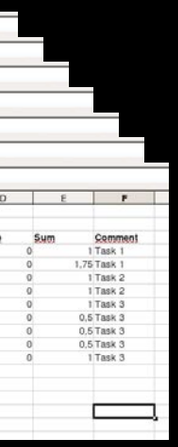
Experimental Methodology

~500 spreadsheets (EUSES)



~ 900 Mechanical Turk workers retype data (presented as images)

Experimental Methodology



~ 900 Mechanical Turk workers
retype data (presented as images)

Fill in this spreadsheet

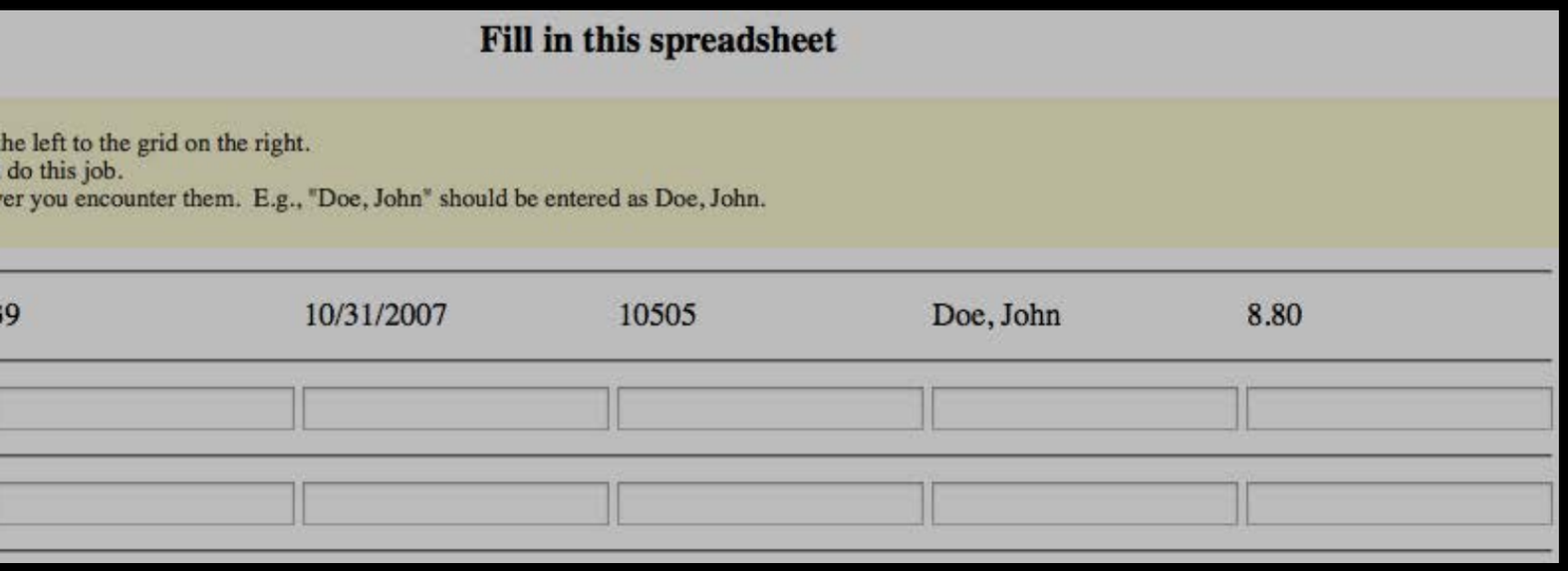
- Copy the values from the image on the left to the grid on the right.
- The first row shows how you should do this job.
- Remove quotes from strings whenever you encounter them. E.g., "Doe, John" should be entered as Doe, John.

39,10/31/2007,10505,"Doe, John",8.80	39	10/31/2007	10505	Doe, John	8.80
2779,5799,5726,1152,4132					
6641,840,3643,137,198					

Experimental Methodology

~ 900 Mechanical Turk workers

retype data (presented as images)



Generative model
of typographical errors

Experimental Methodology

Generative model
of typographical errors

Injected into ~70 randomly-chosen spreadsheets (5% error rate)

Experimental Methodology

Generative model
of typographical errors

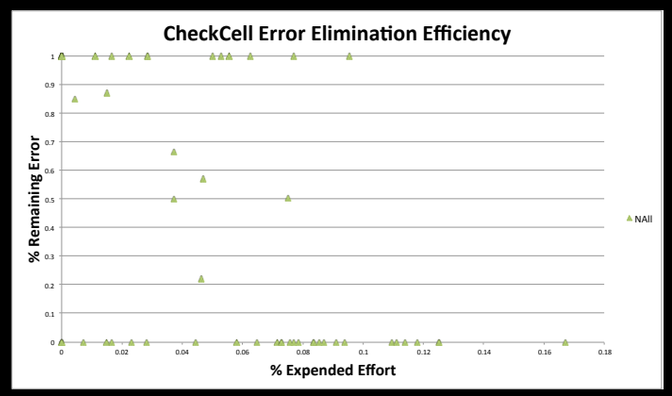
Injected into ~70 randomly-chosen spreadsheets (5% error rate)

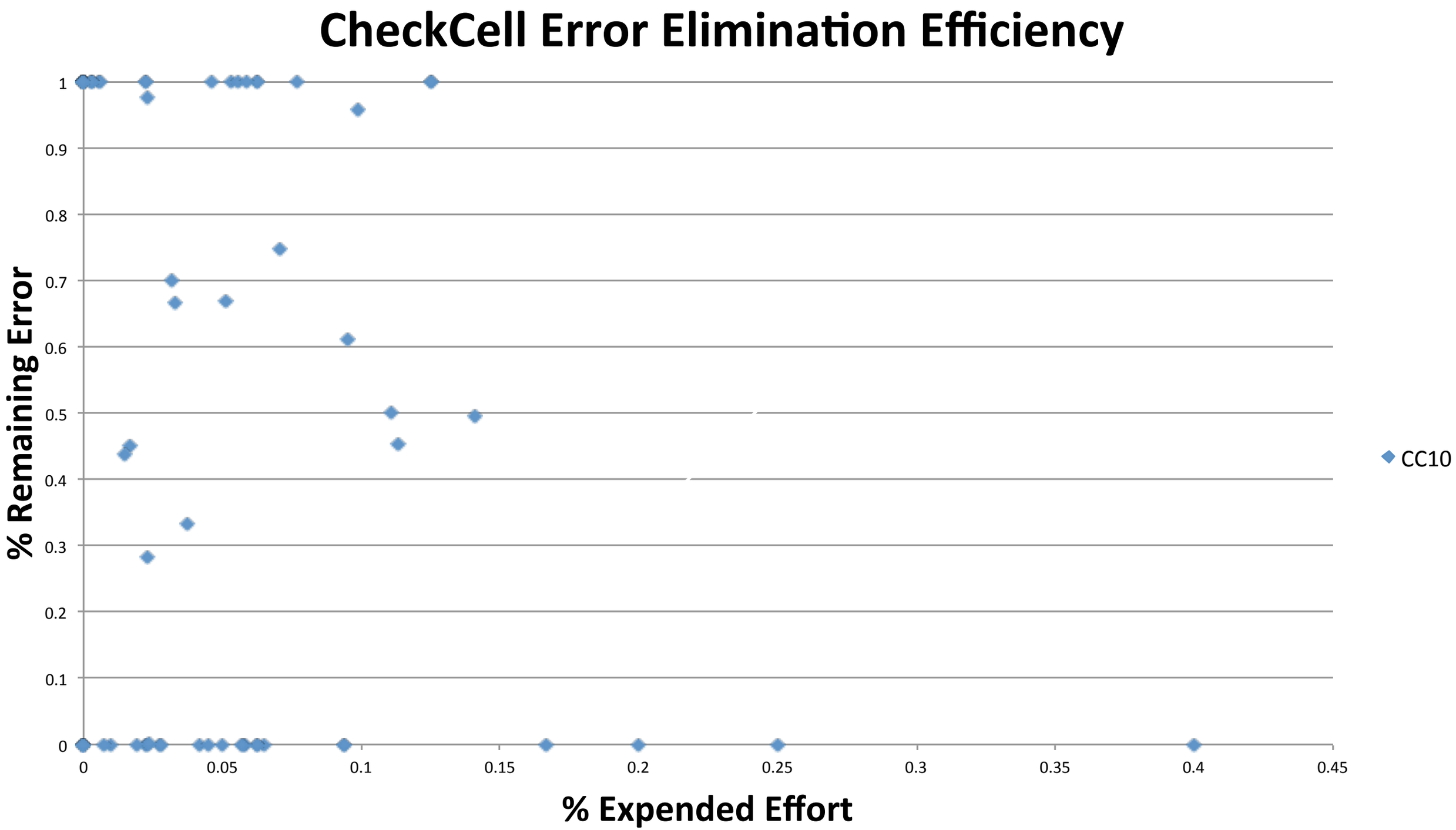
Originals = oracle (“correct”)

Experimental Methodology

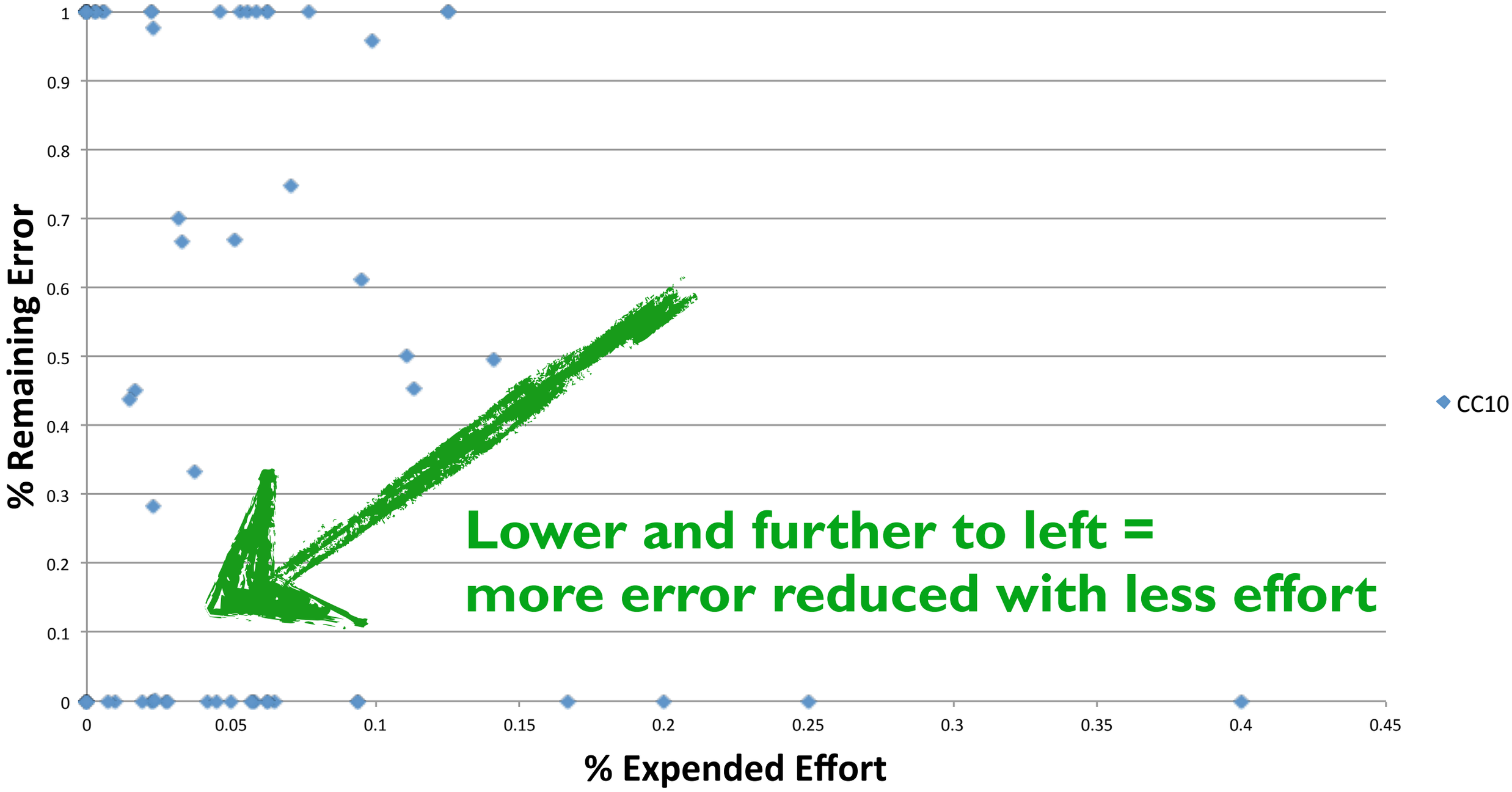
Injected into ~70 randomly-chosen spreadsheets (5% error rate)

Measure **error elimination efficiency**: Reduction in errors versus % cells inspected

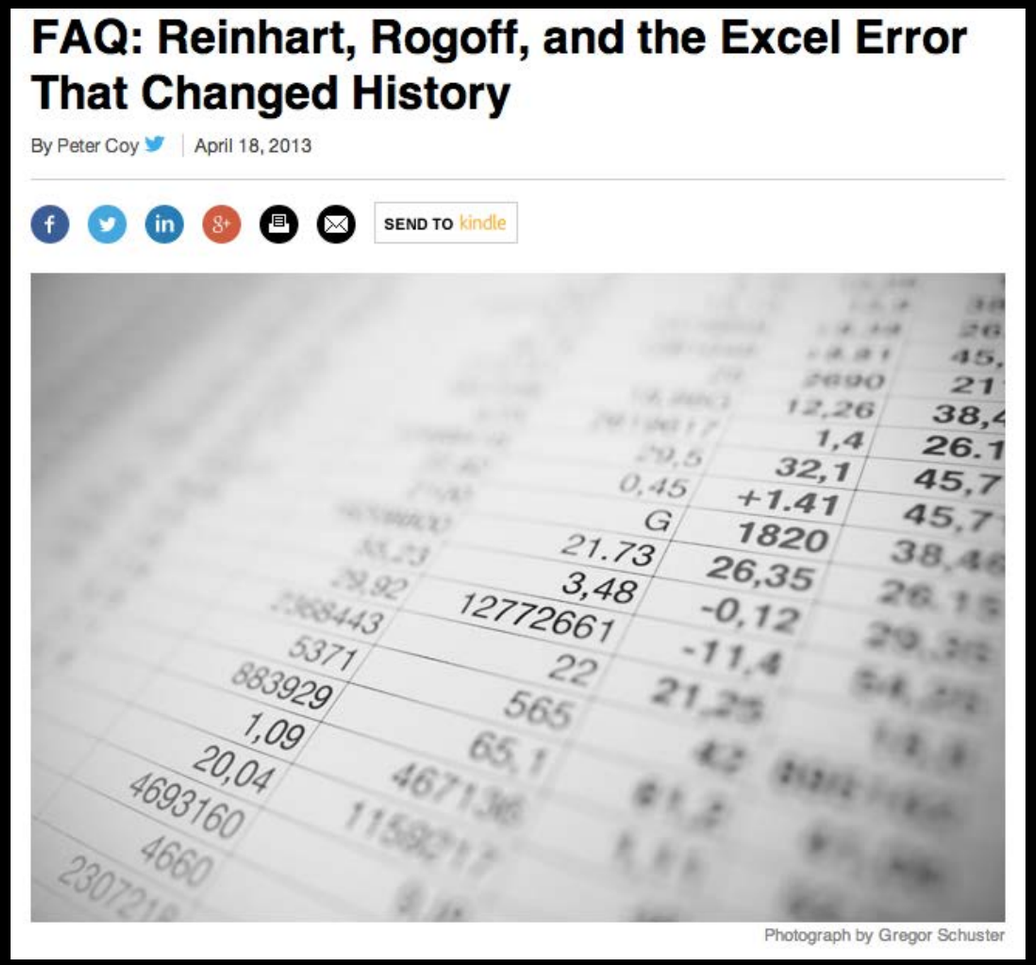



CheckCell Error Elimination Efficiency
% Remaining Error
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
% Expended Effort
CC10

CheckCell Error Elimination Efficiency



Case Study



FAQ: Reinhart, Rogoff, and the Excel Error That Changed History

By Peter Coy | April 18, 2013

SEND TO kindle

Photograph by Gregor Schuster

Reinhart-Rogoff spreadsheet

Case Study

FAQ: Reinhart, Rogoff, and the Excel Error That Changed History

By Peter Coy | April 18, 2013

Photograph by Gregor Schuster

e: December 5, 2009

			Number of observations				Average							
			Debt/GDP				Real GDP growth Debt/GDP				Inflation Debt/GDP			
Country	Coverage	Total	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or ab
US	1791-2009		129	59	23	5	4.0	3.4	3.3	-1.8	1.1	1.8	2.3	6
UK	1830-2009		3	68	27	82	2.5	2.2	2.1	1.8	0.8	4.2	1.4	2
Sweden	1880-2009		79	40	11	0	2.9	2.9	2.7	n.a.	2.8	4.6	4.2	n.
Spain	1850-2009		26	53	47	30	1.6	3.2	1.3	2.8	9.9	5.5	2.3	0
Portugal	1880-2009		42	10	39	0	4.8	2.5	1.4	n.a.	8.8	3.3	0.9	n.
Norway	1880-2009		98	25	1	0	2.9	4.4	10.2	n.a.	4.4	-0.1	0.0	n.
New Zealand	1932-2009		9	33	17	19	2.5	2.9	3.9	3.6	2.6	7.4	5.0	2
Netherlands	1880-2009		17	50	32	8	4.1	2.8	2.4	2.0	6.4	1.5	0.0	-2
Japan	1885-2009		47	42	11	11	4.9	3.7	3.9	0.7	6.0	2.1	3.2	-1
Italy	1880-2009		26	12	39	49	5.4	4.9	1.9	0.7	5.6	11.1	10.6	13
Ireland	1949-2009		8	14	32	7	4.4	4.5	4.0	2.4	2.9	4.8	7.3	5
Greece	1884-2009		13	5	11	55	4.0	0.3	4.8	2.5	13.3	19.4	12.3	2
Germany	1880-2009		96	11	0	0	3.6	0.9	n.a.	n.a.	1.8	1.5	n.a.	n.
France	1880-2009		26	21	19	37	4.9	2.7	2.8	2.3	5.2	5.0	1.5	1
Finland	1914-2009		69	18	6	3	3.2	3.0	4.3	1.9	10.3	5.4	13.2	32
Denmark	1880-2009		57	16	17	0	3.1	1.7	2.4	n.a.	2.5	4.7	3.3	n.
Canada	1925-2009		3	52	23	7	1.9	4.5	3.0	2.2	2.2	4.1	0.6	6
Belgium	1835-2009		37	60	32	31	3.0	2.6	2.1	3.3	1.0	2.0	3.0	3
Austria	1880-2009		43	32	35	0	4.3	3.0	2.3	n.a.	5.3	2.4	0.7	n.
Australia	1902-2009		38	33	23	8	3.1	4.1	2.3	4.6	5.9	2.9	5.2	3
		2317	866	654	445	352	3.7	3.0	3.5	1.7	5.5	5.2	4.6	5
Minimum							1.6	0.3	1.3	-1.8	0.8	-0.1	0.0	-2
Maximum							5.4	4.9	10.2	3.6	13.3	19.4	13.2	32

Reinhart-Rogoff spreadsheet

Case Study

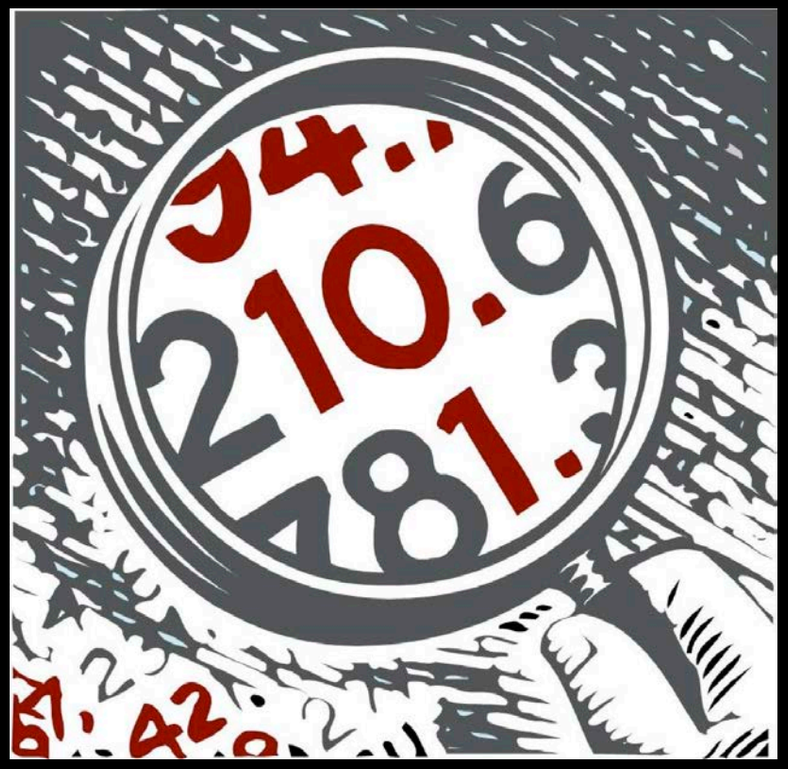
e: December 5, 2009			Number of observations				Average											
							Real GDP growth Debt/GDP				Inflation Debt/GDP				Real GDP growth Debt/GDP			
			Debt/GDP															
Country	Coverage	Total	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abo
US	1791-2009		129	59	23	5	4.0	3.4	3.3	-1.8	1.1	1.8	2.3	6.1	4.0	3.7	3.4	-0.
UK	1830-2009		3	68	27	82	2.5	2.2	2.1	1.8	0.8	4.2	1.4	2.0	2.0	2.6	1.8	2.
Sweden	1880-2009		79	40	11	0	2.9	2.9	2.7	n.a.	2.8	4.6	4.2	n.a.	3.3	3.0	2.9	n.a
Spain	1850-2009		26	53	47	30	1.6	3.2	1.3	2.8	9.9	5.5	2.3	0.5	1.7	3.3	0.8	2.
Portugal	1880-2009		42	10	39	0	4.8	2.5	1.4	n.a.	8.8	3.3	0.9	n.a.	5.4	2.4	1.4	n.a
Norway	1880-2009		98	25	1	0	2.9	4.4	10.2	n.a.	4.4	-0.1	0.0	n.a.	3.0	4.4	10.2	n.a
New Zealand	1932-2009		9	33	17	19	2.5	2.9	3.9	3.6	2.6	7.4	5.0	2.8	2.8	3.0	2.9	4.
Netherlands	1880-2009		17	50	32	8	4.1	2.8	2.4	2.0	6.4	1.5	0.0	-2.2	4.2	3.1	2.0	1.
Japan	1885-2009		47	42	11	11	4.9	3.7	3.9	0.7	6.0	2.1	3.2	-1.1	6.2	3.5	1.9	1.
Italy	1880-2009		26	12	39	49	5.4	4.9	1.9	0.7	5.6	11.1	10.6	13.1	5.8	3.1	1.6	1.
Ireland	1949-2009		8	14	32	7	4.4	4.5	4.0	2.4	2.9	4.8	7.3	5.3	5.3	4.1	3.7	3.
Greece	1884-2009		13	5	11	55	4.0	0.3	4.8	2.5	13.3	19.4	12.3	2.8	3.9	0.5	3.8	3.
Germany	1880-2009		96	11	0	0	3.6	0.9	n.a.	n.a.	1.8	1.5	n.a.	n.a.	3.6	1.2	n.a.	n.a
France	1880-2009		26	21	19	37	4.9	2.7	2.8	2.3	5.2	5.0	1.5	1.2	5.4	2.7	2.8	1.
Finland	1914-2009		69	18	6	3	3.2	3.0	4.3	1.9	10.3	5.4	13.2	32.7	3.3	3.2	3.8	0.
Denmark	1880-2009		57	16	17	0	3.1	1.7	2.4	n.a.	2.5	4.7	3.3	n.a.	2.8	0.8	2.6	n.a
Canada	1925-2009		3	52	23	7	1.9	4.5	3.0	2.2	2.2	4.1	0.6	6.0	2.5	4.2	4.1	2.
Belgium	1835-2009		37	60	32	31	3.0	2.6	2.1	3.3	1.0	2.0	3.0	3.2	2.8	2.8	2.6	2.
Austria	1880-2009		43	32	35	0	4.3	3.0	2.3	n.a.	5.3	2.4	0.7	n.a.	4.6	2.3	2.1	n.a
Australia	1902-2009		38	33	23	8	3.1	4.1	2.3	4.6	5.9	2.9	5.2	3.7	3.5	4.7	3.4	6.
		2317	866	654	445	352	3.7	3.0	3.5	1.7	5.5	5.2	4.6	5.7	3.9	3.1	2.8	1.
Minimum							1.6	0.3	1.3	-1.8	0.8	-0.1	0.0	-2.2	1.7	0.5	0.8	-0.
Maximum							5.4	4.9	10.2	3.6	13.3	19.4	13.2	32.7	6.2	4.4	10.2	4.

Case Study

e: December 5, 2009		Number of observations					Average											
			Debt/GDP				Real GDP growth Debt/GDP				Inflation Debt/GDP				Real GDP growth Debt/GDP			
Country	Coverage	Total	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abov	30 or less	30 to 60	60 to 90	90 or abo
US	1791-2															.7	3.4	-0.
UK	1830-2															.6	1.8	2.
Sweden	1880-2															.0	2.9	n.a
Spain	1850-2															.3	0.8	2.
Portugal	1880-2															.4	1.4	n.a
Norway	1880-2															.4	10.2	n.a
New Zealand	1932-2															.0	2.9	4.
Netherlands	1880-2															.1	2.0	1.
Japan	1885-2															.5	1.9	1.
Italy	1880-2															.1	1.6	1.
Ireland	1949-2															.1	3.7	3.
Greece	1884-2															.5	3.8	3.
Germany	1880-2															.2	n.a.	n.a
France	1880-2															.7	2.8	1.
Finland	1914-2															.2	3.8	0.
Denmark	1880-2															.8	2.6	n.a
Canada	1925-2															.2	4.1	2.
Belgium	1835-2															.8	2.6	2.
Austria	1880-2															.3	2.1	n.a
Australia	1902-2															.7	3.4	6.
		2317	866	654	445	352	3.7	3.0	3.5	1.7	5.5	5.2	4.6	5.7	3.9	3.1	2.8	1.
Minimum							1.6	0.3	1.3	-1.8	0.8	-0.1	0.0	-2.2	1.7	0.5	0.8	-0.
Maximum							5.4	4.9	10.2	3.6	13.3	19.4	13.2	32.7	6.2	4.4	10.2	4.

"This one extraordinary growth experience (Norway) contributes fully 5.3 percent (1/19) of the weight [...] even though it constitutes only 0.2 percent (1/445) of the country-years in this category."

— *Herndon et al. (2013)*



CheckCell:

Data Debugging for Spreadsheets

Dan Barowy, Dimitar Gochev, **Emery Berger**

[OOPSLA 2014]

checkcell.org

UMASSCS

SCHOOL OF COMPUTER SCIENCE