LUMIVERO

BREAKTHROUGHS 2025

Business Insights Through Data, Using Excel

Guest
Dr. James Abdey
Associate Professor, London School of Economics

Host
Efthalia Anagnastou
Manager, Marketing Partnerships & Strategy, Lumivero

Moderator

Thalia Anagnostou

Marketing Partnerships & Strategy Manager

Lumivero

Thalia Anagnostou
eanagnostou@lumivero.com

Our Statistical Solution for Microsoft Excel ®

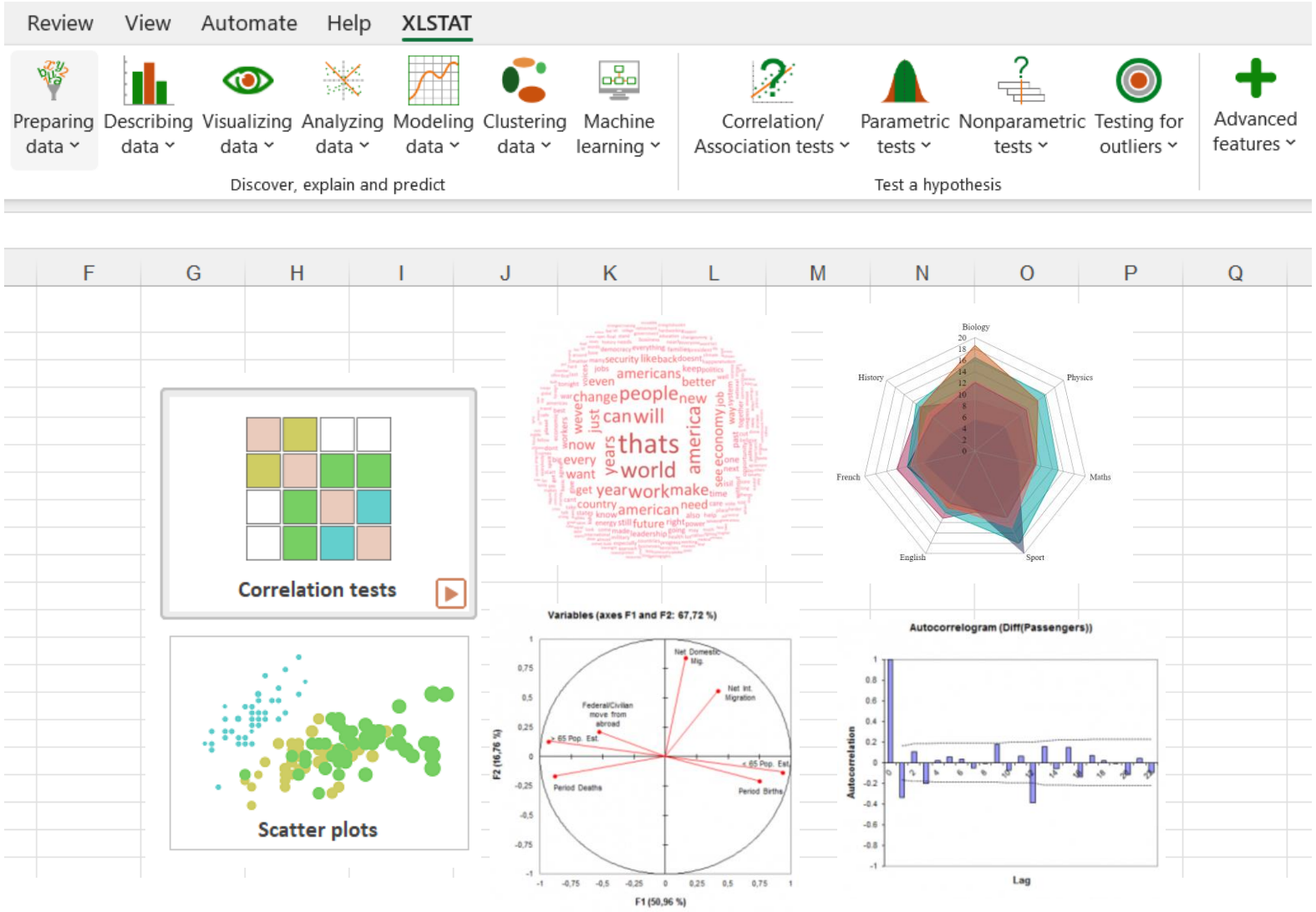
Data Visualization – Exploratory Analysis – Hypothesis Testing – Machine Learning – Time Series and more

XLSTAT

- Integration with Microsoft Excel
- +270 features
- No coding required
- In App customer support access
- 400+ tutorials and videos
- Free version with 14 essential features
- Discounts for students and academics

Learn more about XLSTAT

Request a demo



Lumivero Community

Join Now

Share and Learn

- Research Webinars
- Podcast: Between the Data

Connect

- Product User Group Meetings
- Research Online Groups

Presenter

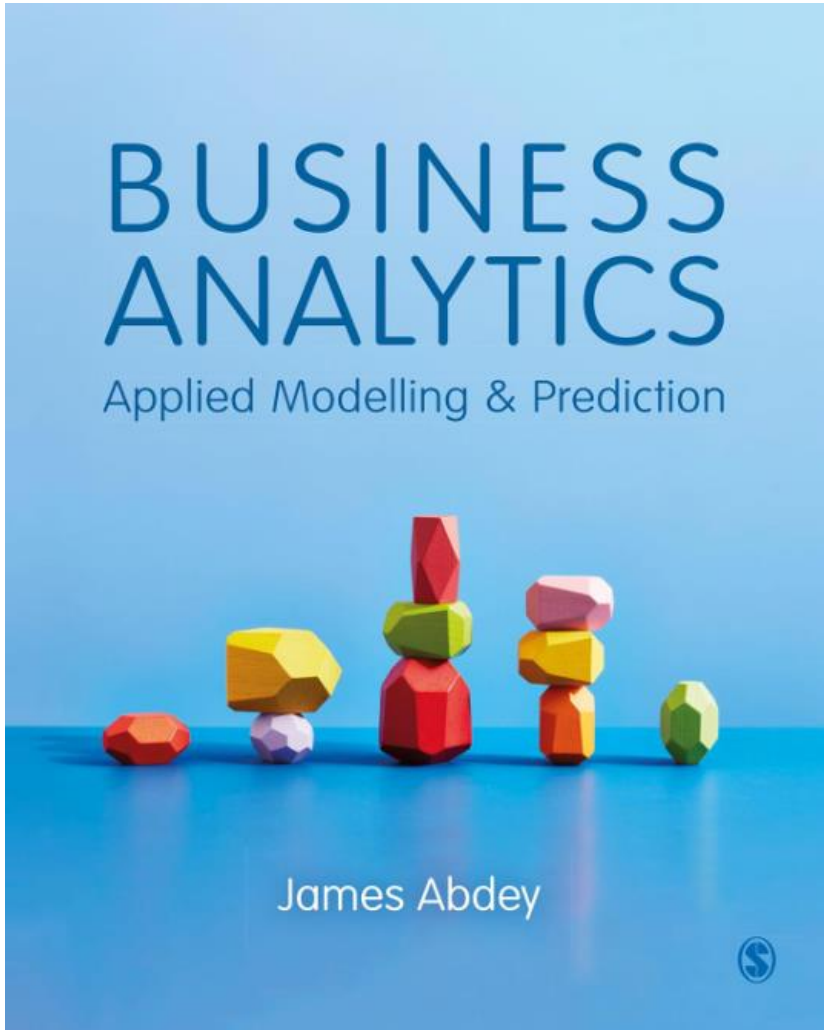
Dr. James Abdey

Dr. James Abdey is an Associate Professor (Education) in Statistics at the London School of Economics, having gained his PhD in 2010 from LSE, asking “To p, or not to p?”! He teaches the Department’s large service-level undergraduate courses in mathematical statistics and quantitative methods, as well as elective courses in market research. His research interests include market research techniques and forensic statistics - the interplay of statistics and the law.

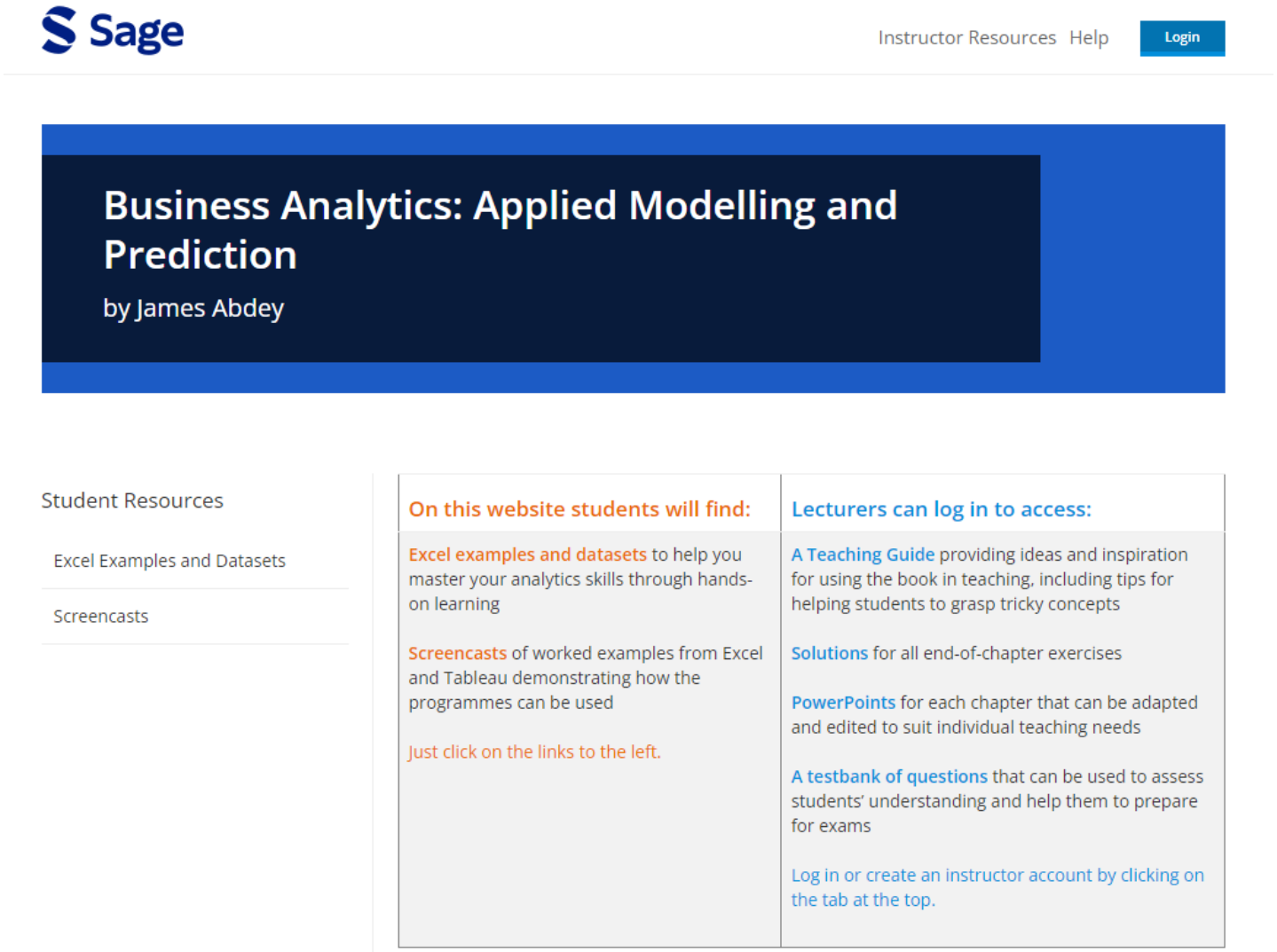
Dr. Abdey has been closely involved with LSE’s Summer School and the University of London International Programmes for a number of years. He helped launch the BSc in Data Science and Business Analytics, with students currently registered in a number of countries. Outside of academia, he has also worked on various quantitative-based consultancy projects in areas including the art market and the World Gold Council.

Dr. James Abdey, Associate Professor
London School of Economics

The future is bright, the future is data



Digital resources



Data: the new oil?

Data: the new oil?

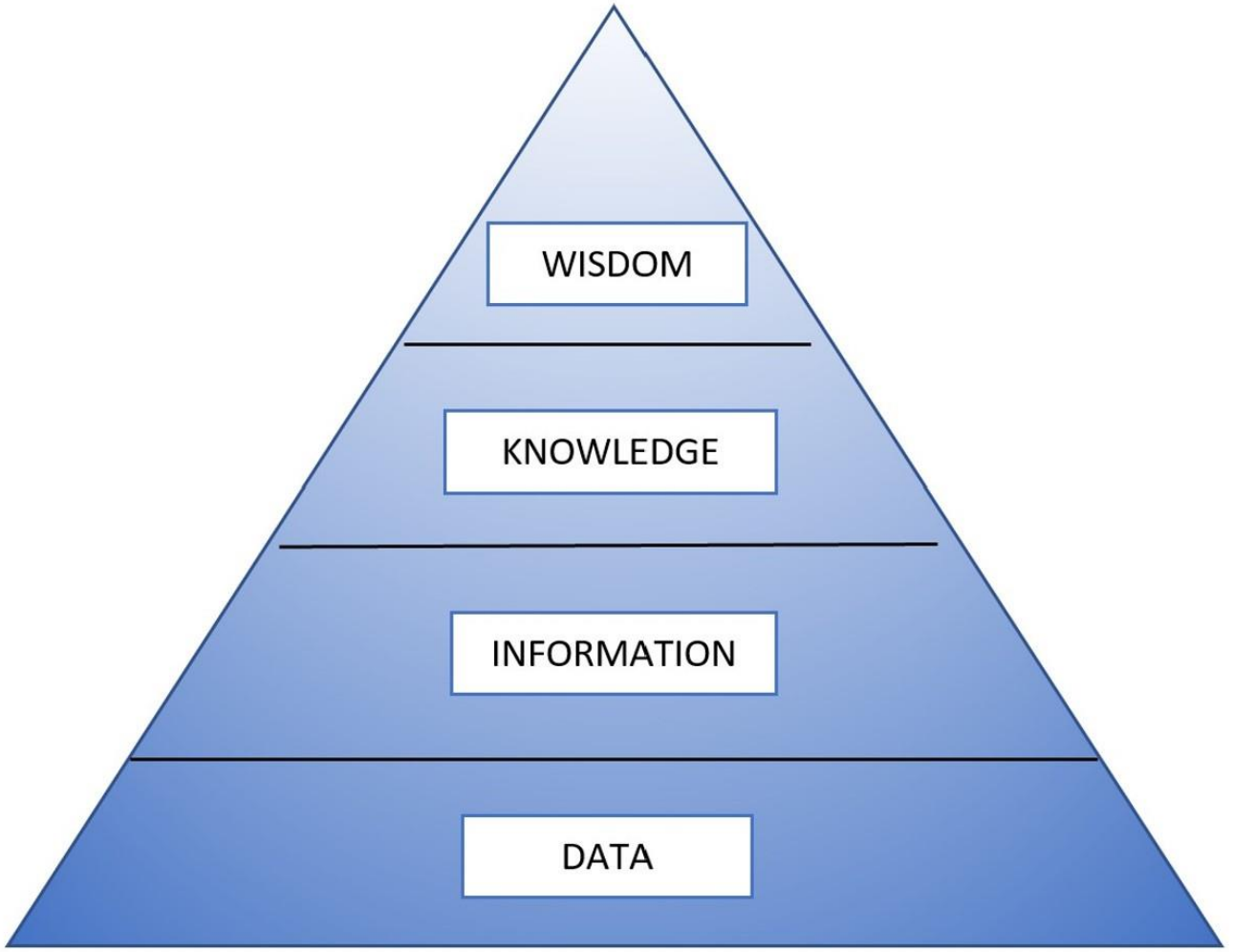
"The world's most valuable resource is no longer oil, but data."

The Economist

Problem: oil is finite, (big) data is near infinite.

But: Data, like oil, derives value after refinement.

DIKW pyramid



In-demand skills

Top Specialized Skills

1. Data Analysis
2. SQL
3. Teamwork / Collaboration
4. Microsoft Power BI
5. Python
6. Stakeholder Management
7. Project Management
8. Key Performance Indicators (KPIs)
9. Tableau
10. Business Intelligence

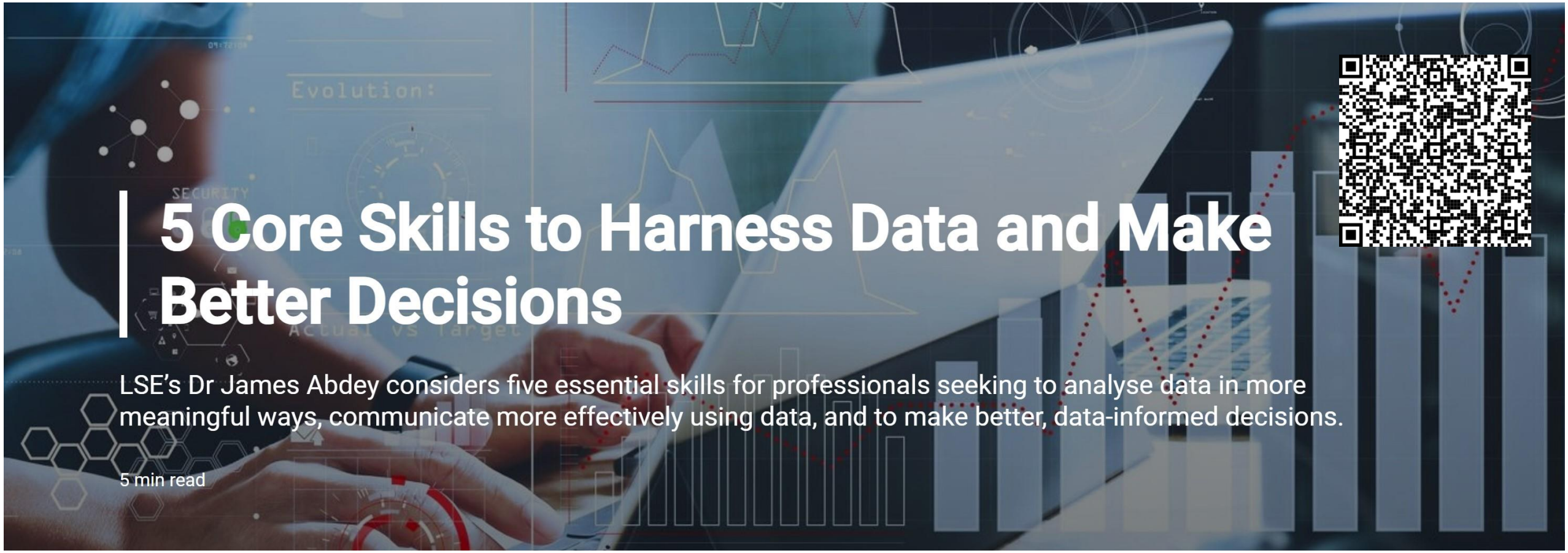
Top Baseline Skills

1. Communication Skills
2. Microsoft Excel
3. Detail-Orientated
4. Problem Solving
5. Research
6. Planning
7. Writing
8. Organisational Skills
9. Presentation Skills
10. Analytical Skills

Top Credentials

1. Microsoft Excel
2. SQL
3. Python
4. Tableau
5. Microsoft Office
6. Microsoft Powerpoint
7. Data Visualisation
8. Salesforce
9. SQL Server
10. Visual Basic for Applications (VBA)

Future-proof yourselves!



5 Core Skills to Harness Data and Make Better Decisions

LSE's Dr James Abdey considers five essential skills for professionals seeking to analyse data in more meaningful ways, communicate more effectively using data, and to make better, data-informed decisions.

5 min read

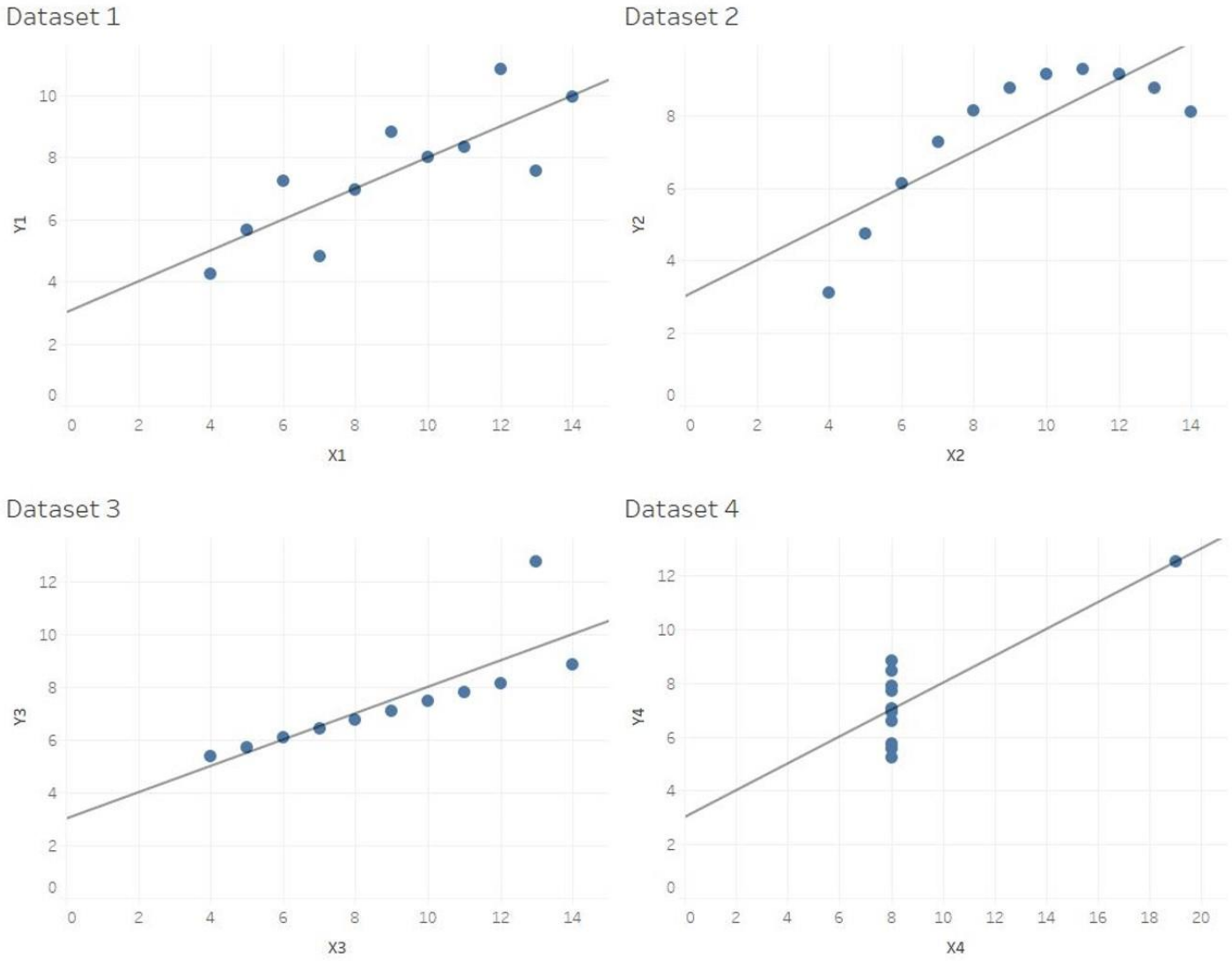
Anscombe's quartet

Dataset 1		Dataset 2		Dataset 3		Dataset 4	
X_1	Y_1	X_2	Y_2	X_3	Y_3	X_4	Y_4
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

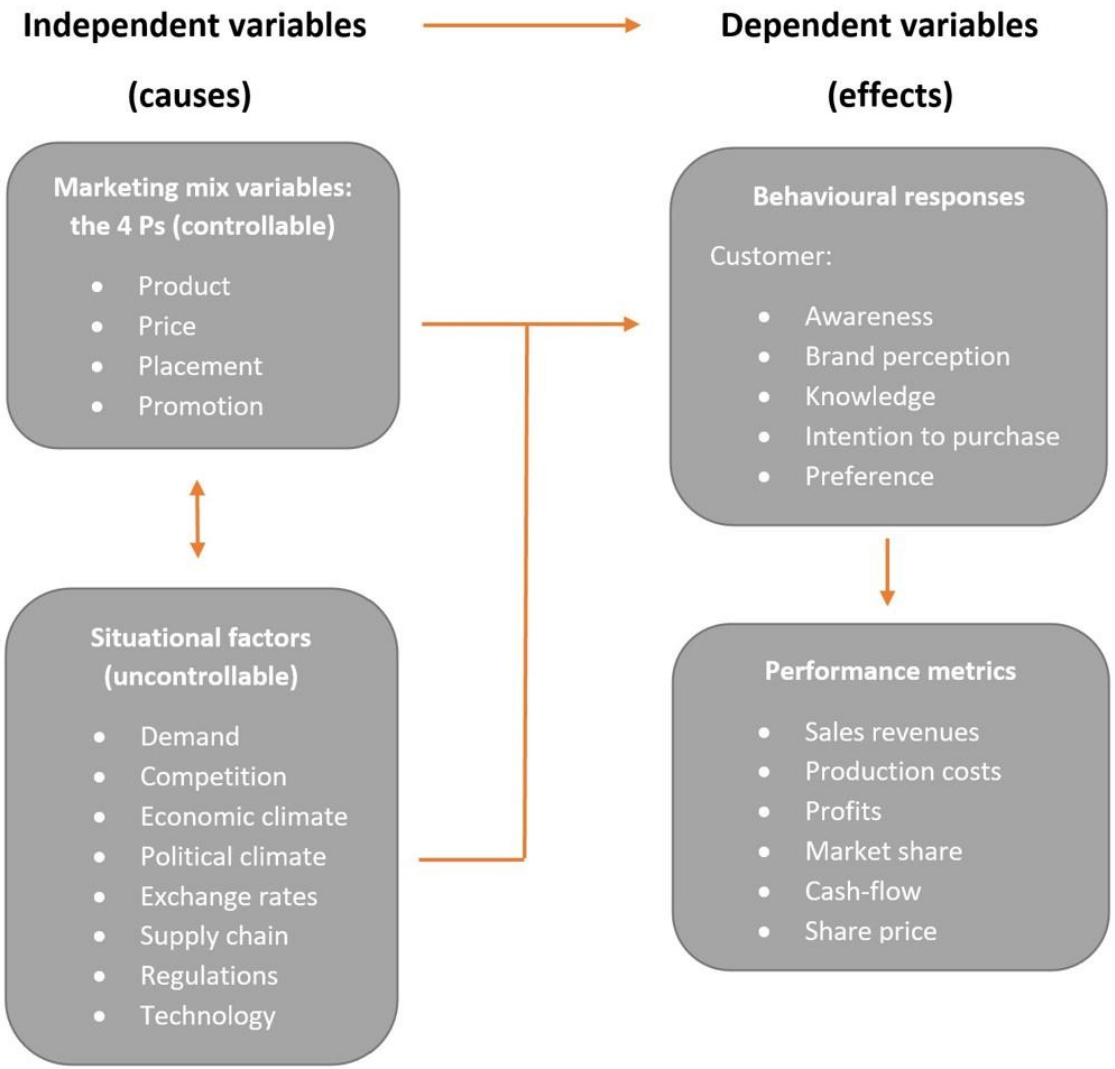
Anscombe’s quartet

	A	B	C	D	E	F	G	H
1	*X1*		*Y1*		*X2*		*Y2*	
2								
3	Mean	9	Mean	7.500909	Mean	9	Mean	7.500909
4	Standard Error	1	Standard Error	0.612541	Standard Error	1	Standard Error	0.612568
5	Standard Deviation	3.316625	Standard Deviation	2.031568	Standard Deviation	3.316625	Standard Deviation	2.031657
6	Sample Variance	11	Sample Variance	4.127269	Sample Variance	11	Sample Variance	4.127629
7	Sum	99	Sum	82.51	Sum	99	Sum	82.51
8	Count	11	Count	11	Count	11	Count	11
9								
10	*X3*		*Y3*		*X4*		*Y4*	
11								
12	Mean	9	Mean	7.5	Mean	9	Mean	7.500909
13	Standard Error	1	Standard Error	0.612196	Standard Error	1	Standard Error	0.612242
14	Standard Deviation	3.316625	Standard Deviation	2.030424	Standard Deviation	3.316625	Standard Deviation	2.030579
15	Sample Variance	11	Sample Variance	4.12262	Sample Variance	11	Sample Variance	4.123249
16	Sum	99	Sum	82.5	Sum	99	Sum	82.51
17	Count	11	Count	11	Count	11	Count	11

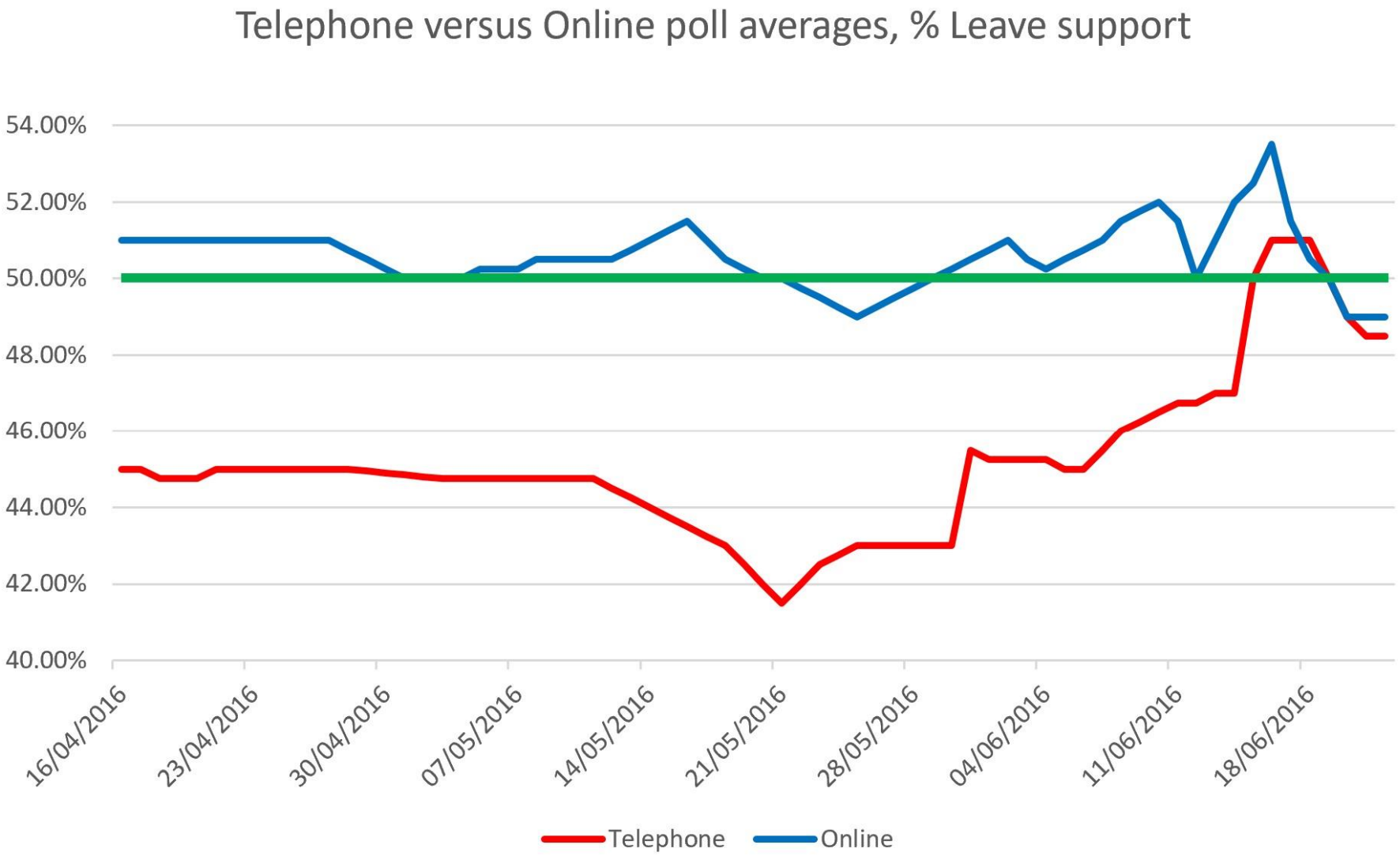
Anscombe’s quartet



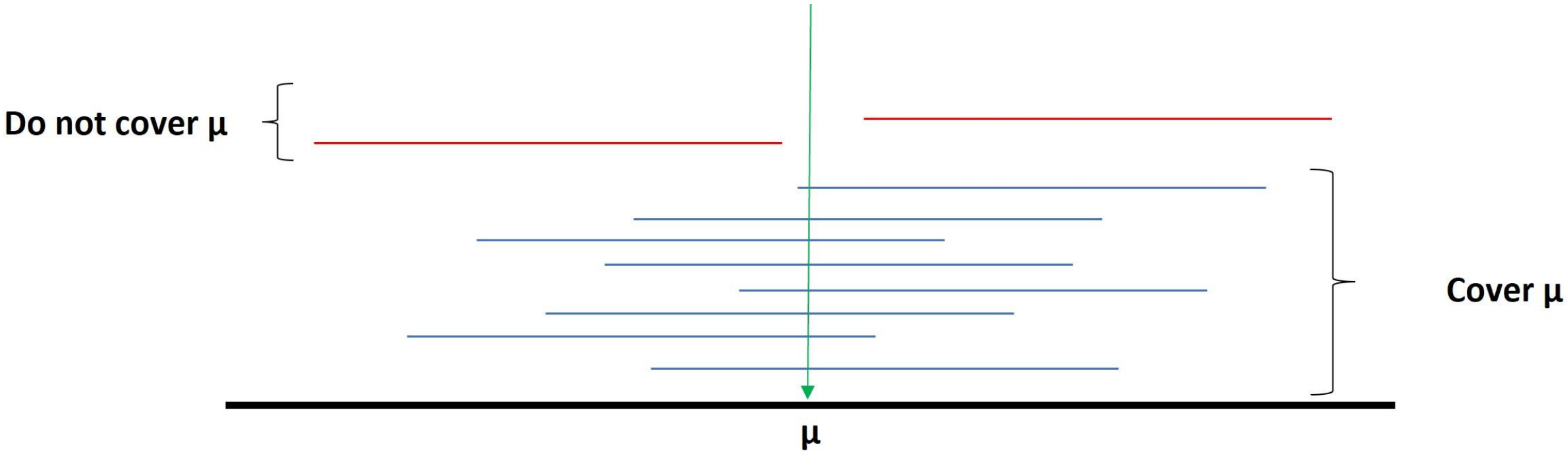
Relationships between variables



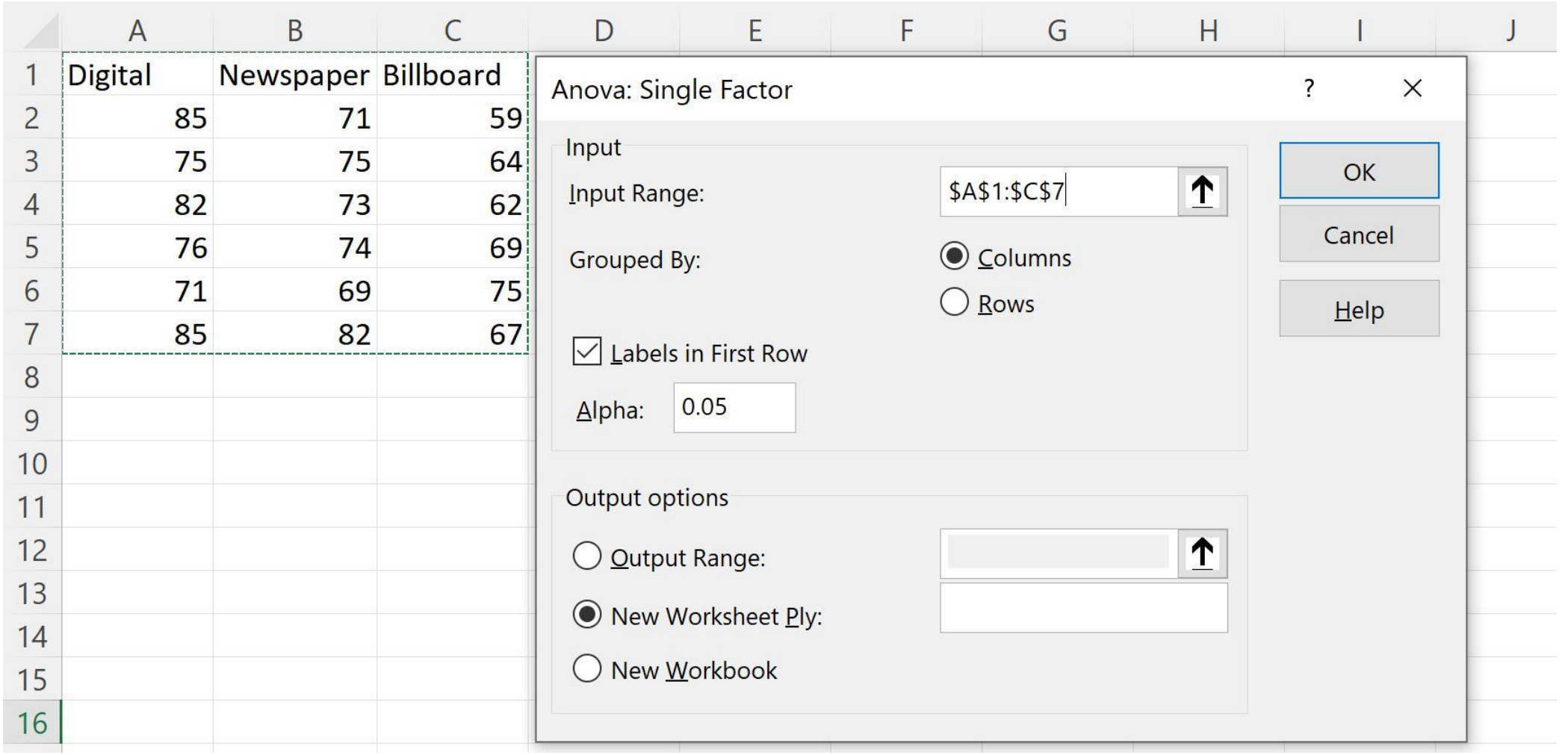
UK’s EU membership referendum poll averages



Coverage probability of confidence intervals



Excel Anova: Single Factor

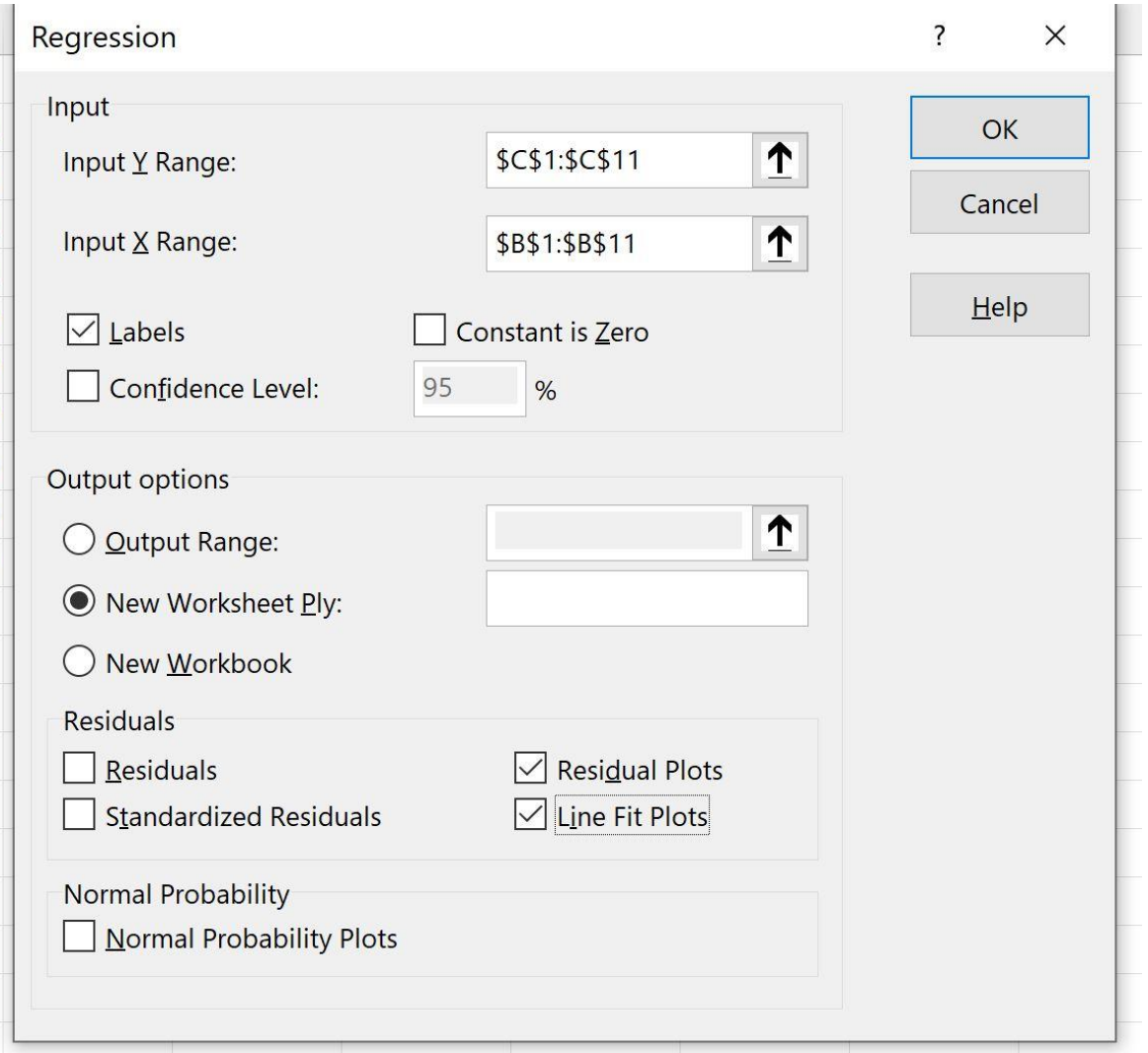


Excel results of a one-way ANOVA

	A	B	C	D	E	F	G
1	Anova: Single Factor						
2							
3	SUMMARY						
4	*Groups*	*Count*	*Sum*	*Average*	*Variance*		
5	Digital	6	474	79	34		
6	Newspaper	6	444	74	20		
7	Billboard	6	396	66	32		
8							
9							
10	ANOVA						
11	*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
12	Between Groups	516	2	258	9	0.002703	3.68232
13	Within Groups	430	15	28.66667			
14							
15	Total	946	17				

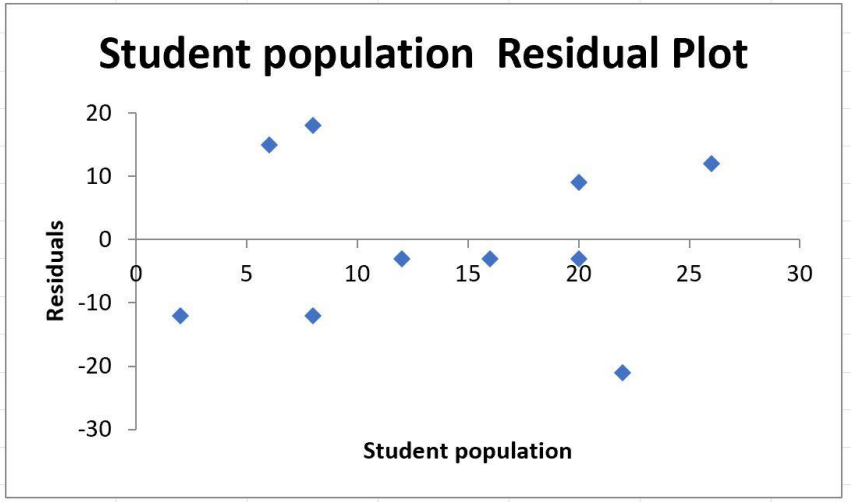
Specifying simple linear regression in Excel

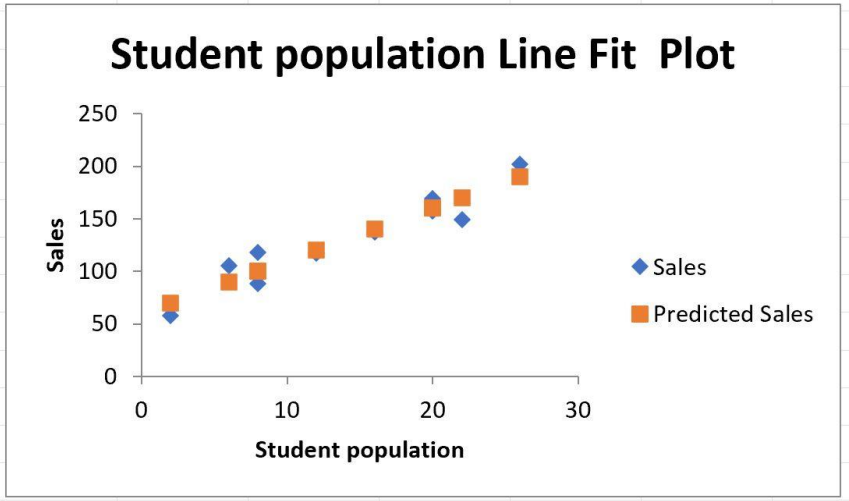
	A	B	C
1	Shop	Student population	Sales
2	1	2	58
3	2	6	105
4	3	8	88
5	4	8	118
6	5	12	117
7	6	16	137
8	7	20	157
9	8	20	169
10	9	22	149
11	10	26	202
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			



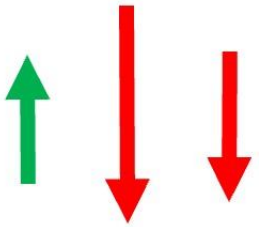
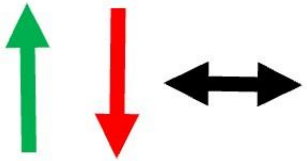
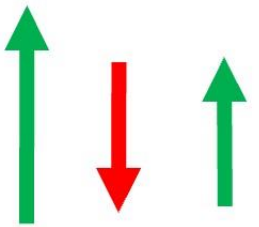
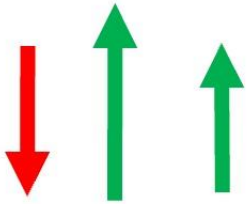
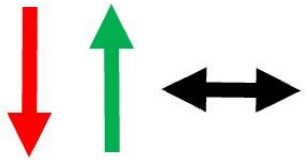
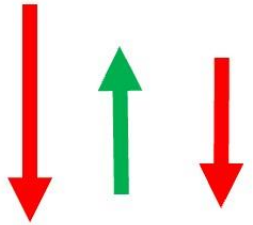
Excel outut for a simple linear regression

	A	B	C	D	E	F	G
1	SUMMARY OUTPUT						
2							
3	*Regression Statistics*						
4	Multiple R	0.950122955					
5	R Square	0.90273363					
6	Adjusted R Square	0.890575334					
7	Standard Error	13.82931669					
8	Observations	10					
9							
10	ANOVA						
11		*df*	*SS*	*MS*	*F*	*Significance F*	
12	Regression	1	14200	14200	74.24837	2.54887E-05	
13	Residual	8	1530	191.25			
14	Total	9	15730				
15							
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
17	Intercept	60	9.22603481	6.503336	0.000187	38.72472558	81.27527442
18	Student population	5	0.580265238	8.616749	2.55E-05	3.661905962	6.338094038
19							
20							
21							
22	RESIDUAL OUTPUT						
23							
24	*Observation*	*Predicted Sales*	*Residuals*				
25	1	70	-12				
26	2	90	15				
27	3	100	-12				





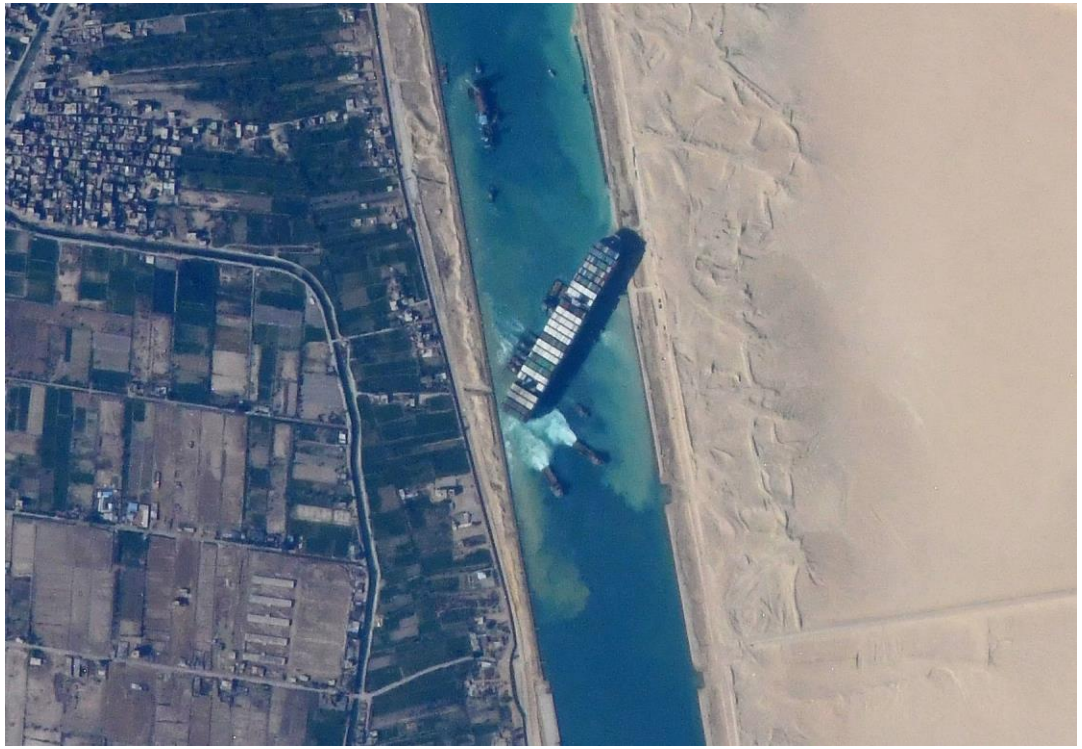
Effect of price elasticity of demand

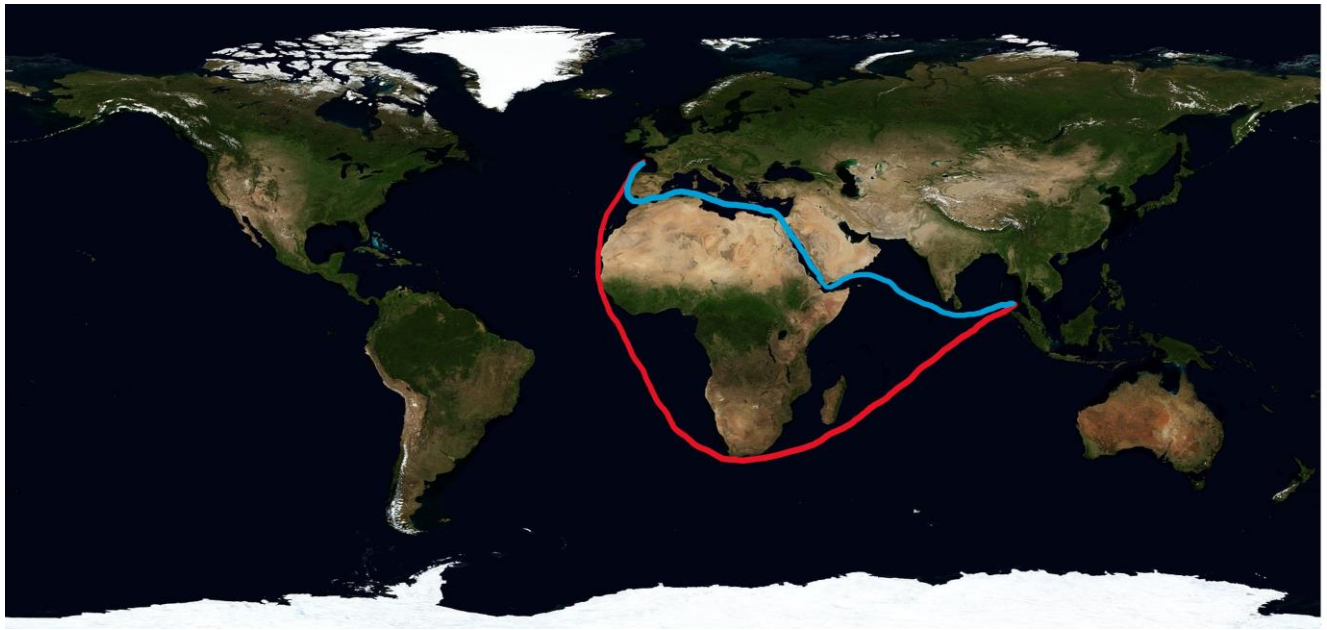
	Price-elastic demand	Unitary-elastic demand	Price-inelastic demand
Price increase	 *P* x *Q* = *TR* Total revenue down	 *P* x *Q* = *TR* Total revenue constant	 *P* x *Q* = *TR* Total revenue up
Price decrease	 *P* x *Q* = *TR* Total revenue up	 *P* x *Q* = *TR* Total revenue constant	 *P* x *Q* = *TR* Total revenue down

Consequences of variable misspecification

Estimated model	True relationship $y = \beta_0 + \beta_1 x_1 + \varepsilon$	$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$
$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1$	Specification is correct, so no issues arise	Omitted variable bias risk with invalid standard errors (Occam's razor-induced)
$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \widehat{\beta}_2 x_2$	Unbiased coefficients (good), but standard errors are inefficient (i.e. large) lowering precision (can occur when fear of omitted variable bias)	Specification is correct so no issues arise

To wait, or not to wait?





Conjoint analysis: Solver solution

=SUM(E13:E15)

	A	B	C	D	E	F	G	H	I
1	**Estimates of regression model coefficients**								
2									
3	^beta_0 =	86.926		^alpha_12 - ^alpha_11 =	-14.111	<- =E14-E13			
4	^beta_1 =	-14.111		^alpha_13 - ^alpha_11 =	-27.778	<- =E15-E13			
5	^beta_2 =	-27.778							
6	^beta_3 =	3		^alpha_22 - ^alpha_21 =	3	<- =E18-E17			
7	^beta_4 =	11.444		^alpha_23 - ^alpha_21 =	11.444	<- =E19-E17			
8	^beta_5 =	-11							
9	^beta_6 =	-21.222		^alpha_32 - ^alpha_31 =	-11	<- =E22-E21			
10				^alpha_33 - ^alpha_31 =	-21.222	<- =E23-E21			
11									
12									
13				^alpha_11 =	13.963				
14				^alpha_12 =	-0.148				
15				^alpha_13 =	-13.815	Sum =	0	<- =SUM(E13:E15)	
16									
17				^alpha_21 =	-4.81467				
18				^alpha_22 =	-1.81467				
19				^alpha_23 =	6.629334	Sum =	0	<- =SUM(E17:E19)	
20									
21				^alpha_31 =	10.74067				
22				^alpha_32 =	-0.25933				
23				^alpha_33 =	-10.4813	Sum =	0	<- =SUM(E21:E23)	
24									
25									
26									
27									

Solver Parameters

Set Objective: $G$15

To: Max Min Value Of: 0

By Changing Variable Cells:

$E$13:$E$15

Subject to the Constraints:

$E$3 = $B$4
$E$4 = $B$5

Add
Change
Delete
Reset All
Load/Save

Make Unconstrained Variables Non-Negative

Select a Solving Method: Simplex LP

Options

Solving Method

Select the GRG Nonlinear engine for Solver Problems that are smooth nonlinear. Select the LP Simplex engine for linear Solver Problems, and select the Evolutionary engine for Solver problems that are non-smooth.

Help
Solve
Close

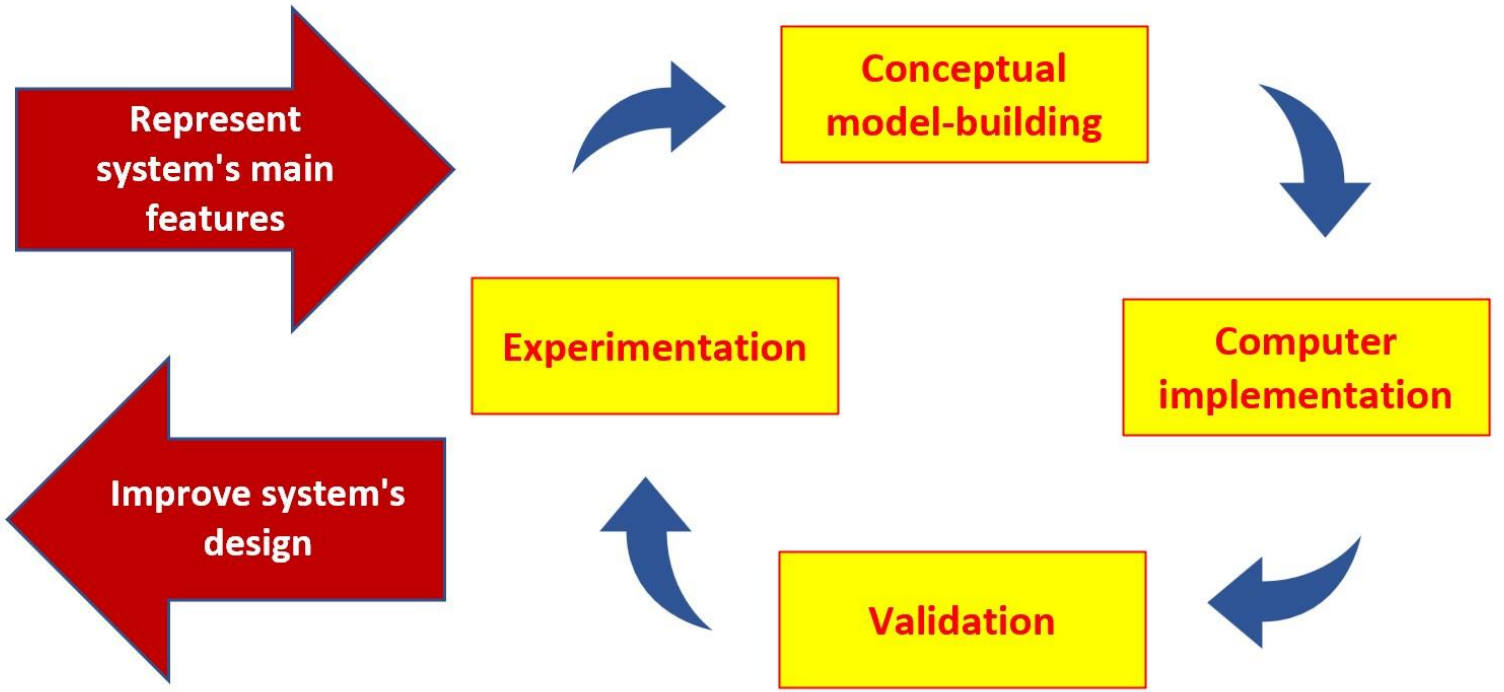
Linking the real world and the abstract world

The real world

Design the system

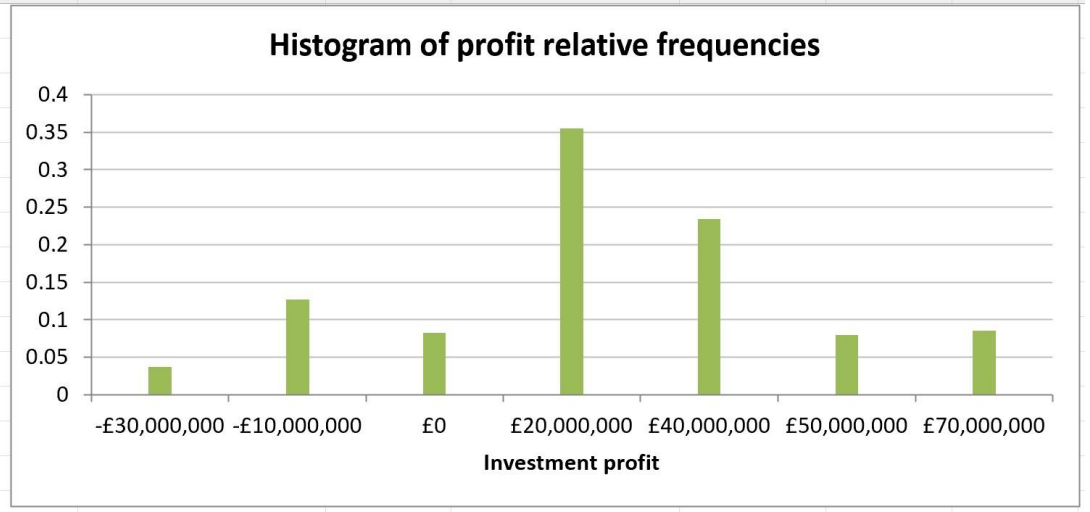
The abstract world

Simulate the system



Output distribution: corporate investment

M28 =SUMPRODUCT(H28:H34,J28:J34)



Distribution of profit

Value	Frequency	Relative frequency
-£30,000,000	185	0.037
-£10,000,000	637	0.1274
£0	411	0.0822
£20,000,000	1775	0.355
£40,000,000	1170	0.234
£50,000,000	397	0.0794
£70,000,000	425	0.085

Expected profit =	£23,996,000
P(loss) =	0.1644

Results of a Monte Carlo simulation

G15 =IF(D15=0,0,MAX(NORM.INV(RAND(),D15,$B$7*D15/$B$6),0))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	**Simulating quarterly profits**															
2																
3	**Assume quarterly revenues are normally distributed**			**Assume quarterly costs are exponentially distributed**												
4																
5	Revenues in Q1			Costs each quarter												
6	Mean	5,000		Lambda	0.00025											
7	Standard deviation	1,800														
8				**Q1 revenue**	**Q1 costs**	**Q1 profit**	**Q2 revenue**	**Q2 costs**	**Q2 profit**	**Q3 revenue**	**Q3 costs**	**Q3 profit**	**Q4 revenue**	**Q4 costs**	**Q4 profit**	**Total profit**
9	Revenues in Q2 through Q4		**Average**	4,992	4,472	520	4,987	4,475	512	5,016	4,442	575	5,056	4,463	593	2,201
10	Mean	The previous quarter	**Standard deviation**	1796	3766	4169	2618	3740	4553	3211	3647	4832	3726	3666	5210	12333
11	Standard deviation	Update to reflect	**Minimum**	0	2,000	-39,670	0	2,000	-33,097	0	2,000	-29,730	0	2,000	-42,399	-59,130
12		realised revenues	**Maximum**	12,059	43,757	9,054	19,283	37,818	15,922	22,299	35,058	20,299	71,018	46,038	69,018	61,511
13																
14			Simulation run	Q1 revenue	Q1 costs	Q1 profit	Q2 revenue	Q2 costs	Q2 profit	Q3 revenue	Q3 costs	Q3 profit	Q4 revenue	Q4 costs	Q4 profit	Total profit
15			1	3,235	2,000	1,235	4,555	2,669	1,886	4,370	2,123	2,247	5,065	2,984	2,081	7,449
16			2	5,759	2,215	3,544	5,392	4,938	454	7,034	2,000	5,034	4,820	3,539	1,281	10,312
17			3	4,462	7,141	-2,679	4,435	5,365	-930	3,933	6,434	-2,502	3,434	2,000	1,434	-4,678
18			4	1,077	5,416	-4,339	836	2,000	-1,164	0	2,000	-2,000	0	2,000	-2,000	-9,502
19			5	6,699	8,452	-1,753	9,648	3,095	6,553	7,370	2,752	4,618	10,473	2,592	7,881	17,299

Questions?

Free Learning Resources

XLSTAT Free 14-Day Trial

XLSTAT Blog

Customer Support
https://www.xlstat.com/en/contact/support

Tutorials & Videos
https://help.xlstat.com/

Our Youtube Channel
https://www.youtube.com/@XLSTAT

The Data Landscape

LUMIVERO

The most complete collection of research, data and decision-making tools available anywhere

NVIVO citavi XLSTAT Sonia @RISK DecisionTools Tevera