Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

Data Visualisation Link (Tableau Online):

https://public.tableau.com/profile/denise.chua#!/vizhome/Dataviz_Makeover05/Dashboard1?publish=yes

a. Critiques of Data Visualisation

Clarity	,
S/N	Comments
1	The purpose of this visualisation is not conveyed well in the title of the chart. The title misleads readers into thinking that the data shown corresponds to actual reproducibility of research results, and that the bottom line is at least 70% of research is reportedly reproducible.
	In fact, this visualisation aims to show the percentage of respondents that believe at least 70% of the research in their area of interest is reproducible.
2	The x-axis label, 'At least 70%', is not indicative of what the data values on the bar chart mean. Worse still, the mention of 70% may even mislead readers to think the values are related to the actual reproducibility of published research, as incorrectly insinuated in the title.
3	The sorting of horizontal bars by decreasing proportion of respondents is useful to direct readers' attention on the top-ranked area of interests. However, adding additional information on the absolute number of respondents can provide more insights on the sample size of each sub-group of respondents, which may be important for readers to gauge the statistical significance level of the findings.
4	The static visualisation solely provides the statistics for respondents that selected 70%,80%,90% and 100% (proportion of research believed to be reproducible), and therefore discards statistics on the other responses (0%-60%). It would be more informative if the distribution of responses across the range of available options was provided.

Aesthetic

S/N	Comments
1	The use of color-coding here is presumably used to visually partition out the different
	areas of interest, but is superfluous in this case since the y-axis has labels which
	clearly indicate which bars belong to which category.
	Moreover, the use of colors should help to enhance the conveyance of information,
	but in this case, too many constrasting colors due to the presence of too many
	categories makes it harder to read the data.
2	The legend, which should not appear in the visualisation in the first place (as
	explained in the above comment), is used ineffectively because it is not sorted by
	descending order unlike the bar chart. This makes it harder for readers to map the
	colors to the corresponding category.
3	Units are not indicated on x-axis tick values, and may confuse readers into thinking
	that it is an average value rather than percentage of total.

b. Ways to improve current design

Clarity	
Critiques (from part a)	Suggested Improvement
The purpose of this visualisation is not conveyed	Dashboard title can be modified to
well in the title of the chart. The title misleads	"CONFIDENCE IN REPODUCIBILITY OF

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

readers into thinking that the data shown corresponds to actual reproducibility of research results, and that the bottom line is at least 70% of research is reportedly reproducible. In fact, this visualisation aims to show the percentage of respondents that believe at least 70% of the research in their area of interest is reproducible.	 PUBLISHED RESEARCH" to reflect that the visualization is the confidence level of researchers/ what the researchers believe is the reproducibility of the research in their area of interest. The chart title should incorporate the statement: "How Many Researchers Believe That At Least 70% of Research In Their Field Is Reproducible?"
The x-axis label, 'At least 70%', is not indicative of what the data values on the bar chart mean. Worse still, the mention of 70% may even mislead readers to think the values are related to the actual reproducibility of published research, as incorrectly insinuated in the title.	The x-axis labels will be modified to "%respondents that believe %X of research in their field is reproducible" to reflect the meaning the data values on the bar chart. Note: X% refers to a dynamic value that the user can indicate, with the use of parameter feature in tableau
The sorting of horizontal bars by decreasing proportion of respondents is useful to direct readers' attention on the top-ranked area of interests. However, adding additional information on the absolute number of respondents can provide more insights on the sample size of each sub-group of respondents, which may be important for readers to gauge the statistical significance level of the findings.	To retain the sorting of horizontal bars by decreasing proportion of respondents. A histogram/ count plot of number of responses across the survey response options will be created, and added into the tooltip of the lollipop chart. This allows users to view the distribution and absolute count of responses, based on detail-on- demand.
The static visualisation solely provides the statistics for respondents that selected 70%,80%,90% and 100% (proportion of research believed to be reproducible), and therefore discards statistics on the other responses (0%-60%). It would be more informative if the distribution of responses across the range of available options was provided.	To provide users the ability to select other survey response options (0% - 60%), and also the ability to select "at least", "at most", and "exactly", through the use of parameters. Based on the users' select input, the lollipop chart will vary as the data is computed differently, potentially revealing more insightful information from the dynamic visualisation

Aesthetic

Critiques (from part a)	Suggested Improvement
The use of color-coding here is presumably used to	Instead of using areas of interest to color-
visually partition out the different areas of	code the categories, the % of respondents
interest, but is superfluous in this case since the y-	data values will determine the
axis has labels which clearly indicate which bars	lightness/darkness of a color through the
belong to which category.	use of gradient palette. This puts emphasis
	on the research areas with high confidence
Moreover, the use of colors should help to	levels of results reproducibility.
enhance the conveyance of information, but in this	
case, too many constrasting colors due to the	

presence of too many categories makes it harder	The legend, in this case, would be placed in
to read the data.	the visualization for readers to validate the
The legend, which should not appear in the	visual encoding.
visualisation in the first place (as explained in the	
above comment), is used ineffectively because it is	
not sorted by descending order unlike the bar	
chart. This makes it harder for readers to map the	
colors to the corresponding category.	
Units are not indicated on x-axis tick values, and	Units (%) will be indicated in all the tick
may confuse readers into thinking that it is an	values on the x-axis
average value rather than percentage of total.	

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

Sketch of Proposed Design



- To retain the original intent of the visualisation, lollipop chart (which is a variant of horizontal bar chart) sorted by largest to smallest proportion of respondents was used to put emphasis on the top-ranked area of interest; While the lollipop circle is used as a visual element to highlight data value, which is what the readers are seeking after.
- The use of displaying histogram inside the tooltip box is to provide readers details-ondemand if they are interested to know the distribution of responses within each area of interest. The histogram also provides readers additional information on the actual count of responses.
- Parameter features will be available for users to indicate which survey response options (0% -100%) they would like to see the results of; And users with the ability to select "at least", "at most", and "exactly". Based on the users' select input, the lollipop chart will vary as the data is computed differently, potentially revealing more insightful information from the dynamic visualisation

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

c. Step-by-step description on data visualization preparation

Tab	Step	Action													
Data	Import to	Inspecting the raw excel data file													
Source	tableau	 Column V correspond to the survey responses that are pertinent to the original visualisation intent to show "What proportion of published results 													
oouree	tubicau	- Cont	original visualisation intent to show "What proportion of published results												
		Ung	in your field are reproducible?". Columns F:U and W:CH correspond to												
		in ye	our fie	eld are re	producible	e?". Colu	umns F:U a	and W:CH	correspo	nd to					
		surv	ey re	sponses f	or questio	ns that a	are irrelev	ant and tl	herefore	were					
		dele	eted in	h the exce	el sheet so	as to ob	otain a qu	estion-foc	used data	aset to					
		wor	k with	n in Tahla	au Colum	ns CI·Dk	correspo	nd to cha	ractoristi	rs of					
		0001	respondents/demographic data (including area of interast) are not ontially												
		resp	useful and therefore kept first for further analysis												
		use	useful and therefore kept first for further analysis.												
		$\times \checkmark$	\times \checkmark f_x 'In your opinion, what proportion of published results in your field are reproducible? i.e. the results of a given study could be replicated exactly or reproduced in multiple similar experimental systems with variations of experimental settings such as materials and experimental model)												
		-				o materialo a									
		T	l	,	V		W	X	Y A						
			_					D.							
		is Vhich o	Tow	hai 👻 In your op	pinion, what proj	portion of put	olis * Please *	To wha ▼ den	it di						
								"I think that "I t	hinł						
		There is	a Area	sona <mark>'90%</mark> sona ' 40%			better th	εAgree Str	ong						
		There is	a Note	noug 30%			worse t	Strongly a Str	on						
		There is	s a A rea	sona 50%			worse th	RAgree Ag	ree						
		There is	sa Area sa Area	sona 70% sona 20%			about tr	eDisagree Ag Maree Ag	rec						
		There is a Lam unsur 70% I am unsur Agree Agree													
		There is a Not enoug 30%worse the Strongly a Agree													
		There is	a A rea	sona 80%			better th	eNeither ag Ag	ree						
		 Importing modified excel file (.xlsx) into tableau: It was verified that the headers and data rows were correctly interpretent to the excel file. 													
		*	#	8 8	8	Abc	*	Abe	Abc	Abc					
		Sheet) responseid	Sheet) respid	SheetI interview_start	Sheeti interview_end	Sheet1 status	Sheet] In your opinion, w	Shoet1 How many membe	Sheet] What is your age?	Sheetl Which of the follo					
		24	24	8/10/2015	8/10/2015	complete	0.90000	1-5 members	65 or over	Research Scientist					
		27	27	8/10/2015	8/10/2015	complete	0.40000	1-5 members	25-34	Student					
		107	107	8/10/2015	8/10/2015	complete	0.30000	1-5 members	45-54	· · · · · · · · · · · · · · · · · · ·					
					107 July 107 Bull 107 July 107	complete	0.50000	6-10 members	35 - 44	Associate professor/L Principal Investigator					
		112	112	8/10/2015	8/10/2015	complete	0.50000	6-10 members 6-10 members	35 - 44 18 - 24	Associate professor/L Principal Investigator PhD Student					
		112	112 114	8/10/2015 8/10/2015	8/10/2015 8/10/2015	complete complete complete	0.50000 0.70000 0.20000	6-10 members 6-10 members 1-5 members	35 - 44 18 - 24 25 - 34	Associate professor/L Principal Investigator PhD Student PhD Student					
		112 114 117	112 114 117	8/10/2015 8/10/2015 8/10/2015	8/10/2015 8/10/2015 8/10/2015	complete complete complete	0.50000 0.70000 0.20000 0.70000	6-10 members 6-10 members 1-5 members 1-5 members	35-44 18-24 25-34 25-34	Associate professor/L Principal Investigator PhD Student PhD Student					
		112 114 117 145 179	112 114 117 145 179	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.70000 0.30000 0.80000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 11-25 members	35 - 44 18 - 24 25 - 34 25 - 34 25 - 34 25 - 34 45 - 54	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Laboratory Director/					
		112 114 117 145 179 180	112 114 117 145 179 180	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete	0.5000 0.70000 0.20000 0.70000 0.30000 0.80000 0.80000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 11-25 members 6-10 members	35-44 18-24 25-34 25-34 25-34 45-54 25-34	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student PhD Student Post-doctoral Fellow Laboratory Director/ Post-doctoral Fellow					
		112 114 117 145 179 180 188	112 114 117 145 179 180 188	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.70000 0.30000 0.80000 0.80000 0.80000 0.40000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-25 members 6-10 members 6-10 members	35-44 18-24 25-34 25-34 45-54 25-34 25-34 25-34	Associate professor/L_ Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Laboratory Director/ Post-doctoral Fellow Post-doctoral Fellow					
		112 114 117 145 179 180 188	112 114 117 145 179 180 188	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.80000 0.40000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 11-25 members 6-10 members 6-10 members	35-44 18-24 25-34 25-34 25-34 45-54 25-34 25-34	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow					
	Additional	112 114 117 145 179 180 188	112 114 117 145 179 180 188	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 9/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.80000 0.40000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-25 members 6-10 members 6-10 members 9 rigin. are	35-44 18-24 25-34 25-34 25-34 45-54 25-34 25-34 Separate	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Laboratory Director/ Post-doctoral Fellow Post-doctoral Fellow					
	Additional	112 114 117 145 179 180 188 • Som	112 114 117 145 179 180 188	e/10/2015 e/10/2015	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.80000 0.40000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 6-10 members 6-10 members origin, are	35-44 18-24 25-34 25-34 25-34 25-34 25-34 25-34 Separate	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow Ost-doctoral Fellow					
	Additional data	112 114 117 145 179 180 188 • Som few	112 114 117 145 179 180 188 ne der	e/10/2015 e/10/2015 e/10/2015 e/10/2015 e/10/2015 e/10/2015 e/10/2015 e/10/2015 mographie nns in the	8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015 8/10/2015	complete complete complete complete complete complete complete complete complete	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.40000 Duntry of c therefore	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 6-10 members 6-10 members origin, are contains in	35-44 18-24 25-34 25-34 25-34 25-34 25-34 25-34 Separate many nul	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow					
	Additional data cleaning	• Som few	112 114 117 145 179 180 188 ne der colur	e/10/2015 e/10/2015	(1) (2) (complete com	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.40000 0.40000 0.40000 0.40000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-25 members 6-10	35-44 18-24 25-34 25-34 25-34 25-34 25-34 25-34 Separate many nul	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student PhD Student Laboratory Director/ Post-doctoral Fellow Post-doctoral Fellow d into a Values					
	Additional data cleaning and	• Som few	112 114 117 145 179 180 188 De der colur	هرامردمان سرالاطن سرالاطن سرالاطن سرالاطن	(1)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)	complete com	0.50000 0.70000 0.20000 0.30000 0.80000 0.80000 0.400000 0.400000000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-5 members 1-3 members 6-10 m	35-44 18-24 25-34 25	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow Associate Fellow Post-doctoral Fellow Associate Fello					
	Additional data cleaning and shaping	112 114 117 145 179 180 189 • Som few Ast Gent In which Europe Europe	112 114 117 145 179 180 188 ne der colur	«/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 mographic mographic sin the which country in mult mult	(1) (2) (complete com	0.50000 0.70000 0.20000 0.30000 0.30000 0.400000 0.40000 0.40000 0.40000 0.40000 0.40000 0.40000 0.40000 0.40000 0.40000 0.40000 0.400000000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 6-10	35-44 18-24 25-34 25-34 45-54 25-34 25	Associate professoryL Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow Associate Fellow Associate Fellow Market Student Market Student Post-doctoral Fellow					
	Additional data cleaning and shaping	Iliz Iliz	112 114 117 145 179 180 188 ne der colur	«/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 mographic mographic sin the which country in	(1)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)(2)	complete com	0.50000 0.70000 0.20000 0.30000 0.80000 0.400000000	6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-15 members 6-10	35-44 18-24 25-34 25	Associate professoryL Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Post-doctoral Fellow Associate Fellow Associate Fellow Market Fellow Market Fellow					
	Additional data cleaning and shaping	Iliz Iliz	112 114 117 149 180 188 ne der colur	«/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 «/10/2015 mographic mogr	(1) (2) (complete com		6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-1-5 members 6-10 m	35-44 18-24 25-34 25	Associate professoryL Principal Investigator PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow Associate of post-doctoral Fellow					
	Additional data cleaning and shaping	Iliz Iliz	112 114 117 149 180 188 ne der colur	«10/2015 «10/2015 «10/2015 «10/2015 «10/2015 «10/2015 mographic mographic "Wech county in "With county in " " " " " " " " " " " " " " " " "	(4) (4) (2) (5) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	complete com		6-10 members -6-10 members -1-5 members -1-	35-44 18-24 25-34	Associate professoryL Principal Investigator PhD Student PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow d into a values					
	Additional data cleaning and shaping	Iliz Iliz	112 114 117 145 179 180 188 ne der colur	(۱۹/2015 (۹/19/2015 (۹/19/2015	(1) (2) (2) (2) (2) (2) (2) (2) (2) (2) (2	complete com		6-10 members -6-10 members -1-5 members -1-5 members -1-5 members -1-5 members -1-5 members -1-2 members -1-	35-44 18-24 25-34	Associate professoryL Principal Investigator PhD Student PhD Student PhD Student PhD Student PhD Student Post-doctoral Fellow Post-doctoral Fellow					
	Additional data cleaning and shaping	Iliz Iliz	112 114 117 145 179 180 188 reder colur	8/10/2015 8/10/2015		complete com		6-10 members 6-10 members 1-5 members 1-5 members 1-5 members 1-5 members 6-10 mem	35.44 18-24 25-34 25	Associate professor/L Principal Investigator PhD Student PhD Student PhD Student Dost-doctoral Fellow Post-doctoral Fellow					

		1	To com	hina t	ho d	lata inta	onosina	a aalu		the role	want cal			
			selected	d, and	ne d 'Me	erge Misi	one sing natch Fie	e colu elds' w	imn, vas s	elected	from the	dropdown		
			Sort fields	Data source or	rder *					□ Sh	ow aliases 🗌 Show hid	den fields 1,000 + rows		
			10-	400		•				6	0			
			Other specialty wi	Sheet1 Sheet1 Other specialty wi In which co null Europe null Europe		y wi In which continent.		theri Which country in	Which country in	Sheet) Which countr	ry in	Sheet) Which country in E	Sheeti Which country in	Rename Copy Values Hide
			null	Europe		null	null	nui	11	Germany	null	Create Calculated Field		
			null	Europe		null	null	nul	11	United Kingdom	null	Pivot Marga Microstobad Ealds		
			null	Europe		null	null	nul	11	Switzerland	null	num		
			nutt	South Ame	rica	null	null	nui	11	Germany	null	Argentina		
			null	Europe		null	null	nul	11	United Kingdom	null	null		
			nutt	Europe		null	null	nul	11	Italy	null	null		
			nutt	Australasi	a	null	Australia	nul	11	null	null	null		
			nuti	Europe		null	null	nul	41	Germany	null	null		
			null	North Ame	rica (includ	null	null	nul	11	null	United States	null		
			nutt	North Ame	rica (includ	null	null	nul	11	null	Canada	null		
			The res	ulting	colu	umn:								
			Abc		۵									
			Sheet1	nent	Project Which	country in								
			Europe	incirc	Sloven	ia								
			Europe		Germa	ny								
			Europe		United	Kingdom								
			Europe		Switze	rland								
			South America	Э	Argentina									
			Europe		Germa	ny								
			Europe		United	Kingdom								
			Europe		Italy									
			Australasia		Austra	lia								
			The sam	ne ste	ps w	/ere perf	ormed to	o merg	ge co	olumns p	pertinent	to "What is		
			your sp	ecialty	y"	and "Oth	ier specia	alty"	into	o a single	e column	each.		
		•	Some co	olumr	ns w	ere renai	med to a	more	intu	itive and	d concise	caption.		
			Specific	ally,								•		
			1. "Which of the following best describes your area of interest?" was											
			renamed to "Field of research"											
			2 "In your opinion what proportion of published results in your field are											
			2. In your opinion, what proportion of published results in your field are											
			reproducible? i.e. the results of a given study could be replicated											
			exactly or reproduced in multiple similar experimental systems with											
			variations of experimental settings such as materials and experimental											
			mo	vdol)"	was	chorton	od to "In	vour	onin	ion what	at propo	tion of		
			inc		vvas			yourt	Jhii		at propoi			
			pu	blishe	a re	suits in y	our field	are re	epro	duciple				
Sheets 1	Creating	•	In order	r to au	utom	natically	compute	propo	ortio	n of res	pondents	s dynamically		
and 2	new		based o	n use	rs' s	election.	the follo	wing r	oara	meters	were cre	ated		
	calculated		1 [In	dicate		up of %V	lnaramo	torwa	nc fu	ct croat	od Thic	aramotor		
	Calculated		<u>т</u> . [Ш	uicate	e vai		j parame			stcreat	eu. mis j	Jarameter		
	fields and		allo	ows u	sers	to view t	the prop	ortion	of r	esponde	ents that	indicated %X		
	parameter	1	of	resea	rch i	s reprod	ucible in	their c	opini	ion, whe	ere X is th	ne range of		
	foruser		0.0	tions	(00/	1000/1:	n the cur		1	,				
			op	10/15	0%	- 100%) I	ii the sur	vey.						
	selection													
	and													
	dynamic	1												
	display													

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

		Edit Parameter [Indic	ate value %X]			×	
			-				
		Name: Indicate value	e %X			Comment >>	
		Data type:	Float				
		Current value:	70%				
		Display format:	70%				
		Allowable values: (Range			
		List of unline) Kunge			
		List of values		Display As		Add from Dovomotor	
		0		0%		Add from Field	
		0.1		10%		Danta from Clinkaard	
		0.2		20%		Paste from Clipboard	
		0.4		40%			
		0.5		50%		Clear All	
		0.6		60%			
		0.7		70%	~		
		L		000/			
					L	OK Cancel	
		-	The l	ist of values w	vas imp	orted by selectin	ıg [In your
			onini	on what pror	nortion	of nublished res	ults in your
			field	ara ranraduci	iblo21 in	the dren down	monu from
			neiu		in[;aid	i the drop-down	menu nom
			"Add	from Field" b	outton.		
		-	Data	type was set	to Float	t	
		-	The c	lisplay format	t was se	et to percentage.	0 decimal
			nlace	····[····)			
						00/	
			met	urrent value	set to 7	0%	
	2. 1	The second	paramet	er [%X	of rese	arch believed to	be
	r	eproducible	e] was cr	eated, to allo	w users	s to view proport	ion of
	r	respondents	s that sel	ected options	s which	are <= %X, >= %	X, or exactly
	c	%X. for whic	h the va	lue of X can b	e speci	fied in the first n	arameter
	Γ						
		cuit Parameter [%X OF res	earch believed to be	reproducib	liej	^
		Name:%	X of research b	elieved to be reproduc	ible	C	Comment >>
		Properties					
		Data type:	String	•			
		Current value:	At Least	•			
			The Loude				
		Display format:					
		Allowable values	s: 🔿 All 🔘	List 🔘 Range			
		List of values					
		Value		Display As		Add from Par	rameter
		>=		At Least			
		<=		At Most		Add from I	Field
		=		Exactly		Paste from	Clipboard
		Add					
						Clear	All
							Connect
						OK	Cancel
	-	-	Thel	ist of values a	re man	ually input, as sh	own in the
			scree	enshot above			
					+	ing	
		-	i i i e c	iala lype is se	ະເລະວເເ	^и '5	

7

- Current value is set at "At Least"
Both parameters are displayed by selecting "Show Parameter Control"
in the right dick drendown field
in the right-click dropdown field.
 A new calculated field [%respondents believe %X of research in their field
is reproducible] was created. The formula includes several conditional
statements, specifically values which users selected from the 2 parameters
to compute the properties of respondents to be displayed in the follings
to compute the proportion of respondents to be displayed in the follop
chart. The exact formula used is given below:
CASE [%X of research believed to be reproducible]
WHEN "<=" then
SUM(
IF [In your opinion, what proportion of published results in your field are
reproducible] <= [Indicate value %X] then 1
else 0
END/
SUM//Evolude [In your opinion, what proportion of published results in your
field are reproducible (dimension)). Clind(Number of Decental))
Jiela are reproducible (almension)]: SUIVI([Number of Records])})
WHEN ">=" then
SUM(
IF [In your opinion, what proportion of published results in your field are
reproducible] >= [Indicate value %X] then 1
else ()
END)/ SUNA/(Evoludo [In your opinion what properties of published results in your
SUM({Exclude [In your opinion, what proportion of published results in your
field are reproducible (dimension)]: SUM([Number of Records])})
WHEN "=" then
SUM(
IF [In your opinion, what proportion of published results in your field are
reproducible] = [Indicate value %X] then 1
plco 0
END/
SUM({Exclude [In your opinion, what proportion of published results in your
field are reproducible (dimension)]: SUM([Number of Records])})
END
The use of {Exclude} LOD expression is to ensure that the computation is
done on a [Field of Research] level, and not for individual survey response
addie of a [field of Research] level, and not for individual survey response
option.
arch in their field is reproducible ×
CASE [% X of research believed to be reproducible] WHEN "<=" then
SUM(IF [In your opinion, what proportion of published results in your field are reproducible] <= [Indicate value %X] then 1
erse o END)/ SUM((Exclude [In your opinion, what proportion of published results in your field are reproducible (dimension)); SUM(Number of Records)))
WHEN ">=" then SUM(
<pre>IF [In your opinion, what proportion of published results in your field are reproducible] >= [Indicate value %X] then 1 else 0 FNUD/</pre>
SUM([Exclude [In your opinion, what proportion of published results in your field are reproducible (dimension)]: SUM([Number of Records])) WHEN "=" then
<pre>SUM(IF [In your opinion, what proportion of published results in your field are reproducible] = [Indicate value %X] then 1</pre>
else 0 END)/ SIM(Exclude [In your opinion, what preparties of published results in your field are repreduible (dimension); emp(dumber of period)))
END
The calculation is valid. 4 Dependencies - Apply OK

Creating • ollipop hart	Field of research] was placed on Rows shelf, and [%respondents believe %X of research in their field is reproducible] was placed twice on Columns shelf. Both fields were automatically aggregated (as a result of the use of table calculation in the formula). If Columns Add/trespondents Add/trespondents Sheet 8 Field of research Attronomy and planetary Biology Generative Biology Field of research Biology Sheet 8 Field of research Attronomy and planetary Generative Biology Generative Store of the context of the c
•	 The second [AGG(%respondents believe %X of research in their field is reproducible)] mark was changed to 'dual axis', by selecting the option in the right-click dropdown menu. This automatically changes both marks to circle chart. The first [AGG(%respondents believe %X of research in their field is reproducible)] mark was changed to bar chart, the color changed to grey, and the size was reduced, so as to create skinny grey bars that resemble lollipop stick. The first [AGG(%respondents believe %X of research in their field is reproducible)] mark remained as circle chart, and the size was enlarged to resemble lollipop. [AGG(%respondents believe %X of research in their field is reproducible)] mark remained as circle chart, and the size was enlarged to resemble lollipop. [AGG(%respondents believe %X of research in their field is reproducible)] field was Ctrl+click and dragged onto color, so that higher data values are colored darker blue and lower data values are colored lighter blue, for visal encoding. [AGG(%respondents believe %X of research in their field is reproducible)] field was also dragged onto label, so that the data values are displayed on the circles. The second axis at the top of the chart was removed by deselecting "Show header" in the right-click dropdown menu.
	AGG(%respondents believe %X of research in their field is reproducible)]. Sort [Field of research] Sort By Field Sort Order Ascending Descending Field Name %respondents believe %X of research Aggregation Custom Clear





			Edit To	oltip												×	
			Table	au Book		~	10	~	B	7		• •	E 3		Īr	nsert 🗸 🗙	
					. 1 .			2 '			3''			4 ' ' '	-	5	
			Eiol	d of rose	arch:	Field	lofros	oarc	h>								
			%re	sponder	nts beli	ieve <	Paran	neter	s.Ind	licate v	alue %	∕₀X> of	f rese	earch in	thei	ir field is	
			repr	oducible	e: <ag< td=""><td>G(%re</td><td>espond</td><td>dents</td><td>beli</td><td>eve %X</td><td>ofres</td><td>search</td><td>n in tł</td><td>neir fiel</td><td>d is</td><th></th><td></td></ag<>	G(%re	espond	dents	beli	eve %X	ofres	search	n in tł	neir fiel	d is		
			repr	oducible	e)>												
			Dist	ribution	ofres	ponse	es for	<fiel< td=""><td>d of r</td><td>researc</td><td>h>:</td><td></td><td></td><td></td><td></td><th></th><td></td></fiel<>	d of r	researc	h>:						
			<she< td=""><td>eet nam</td><td>e="His</td><td>stogra</td><td>am" m</td><td>naxwi</td><td>dth=</td><td>"500"</td><td>maxh</td><td>eight=</td><td>="30</td><td>0" filte</td><td>r="<</td><th>All</th><td></td></she<>	eet nam	e="His	stogra	am" m	naxwi	dth=	"500"	maxh	eight=	="30	0" filte	r="<	All	
			rien	137 7													
			<u> </u>	ow tooltip	s Res	ponsive	e - Shov	<i>w</i> toolt	ips ins	stantly ~	-						
			<u> I</u> nc	clude com	mand bu	uttons											
				✓ Allow selection by category													
					_								_				
			R	eset	Prev	view								OK		Cancel	
			How	to too	ltip lo	ooks	like	upc	n h	overi	ng o	ver c	char	t:			
		5		Columns	AGG(%	respondent	s. AGG	(%respon	lents								
			=	Rows	Field of a	research	F									96Y of research believe	ed to be reprodu
		-		Field of researc	h F											At Least At Most	ed to be reprodu
		5		Astronor	hysics						74	196		Dhusice		C Exactly	
		ll Multiple		planetary s	cience						67%	96respo	ondents b	elieve 70% of	research	h in their field is reproduci	ible: 74%
		ar Size	I Label	Matorials S	cience				_	52%		DISCHOO	In your	opinion, what p	proportio	on of publishe	
		il Tooltip		Ear Environmental S	th and		_	_	-	51%		30	0			30	
		96X Multiple fi	of re	В	iology	_	-	-	46%			20 cesponse	0			23	
		Multiple fi	elds	Engin	eering				45%			count of 1			9		
		GG(%resp.	Baa	Me	dicine			35%				0 1	3 2	4 5 4	6	5	
		a at receipt			Other	10%	20%	35%	40%	50% 6	60% 70)%	0 01 02	0.3 0.4 0.5	0.6 0.7	0.8 0.9 1	
						%responde	ents that beli	ieve %X of	research	in their field is	reproducibl	le					
			Tho ta	ool tir	fort	tha f	irct N	лілі	(NI	mhor	· of R		rdc)	mark		s aditad a	c
		•	show	n in th	non u ne scr	reen	shot	hel	(190 nw	mber	011		lusj	mark	vva	is culted a	3
			5110 44			Edit T	Cooltin	ber									×
			Marks			Luit I	lass			10		-			-	Trant	
			∼ All		O	Tabl	eau Book	: 	· · ·	10	<u>В</u>		3		± : 4 '	Insert •	
				Numb	O	Fiel	ld of res	search		<field o<="" td=""><td>fresea</td><td>rch></td><td></td><td></td><td></td><th></th><td></td></field>	fresea	rch>					
			🕑 Pie		¥	Nur	mber of	Reson	dents	SUM(N	lumber	of Reco	ords)>				
			Colour	6) Size	T												
			Colour														
			Detail	Tooltip	Angle												
			ଚ si	JM(Numbe	er												
			T Fie	ud of resea JM(N <u>umb.</u>	ar Δ	⊠ s	how toolti	ips Re	esponsiv	ve - Show t	tooltips in	stantly	~				
			:: SI	JM(Numbe	er	⊡ Ir	nclude cor	mmand	buttons	3							
				Numbe	A	⊠ A	llow selec	tion by	categor	у							
			✓ IVIIIV(NUMDE	0		Reset	Pre	eview					[0	K Cancel	
								-	-		-	-	-		-		
Dashhoar	Create	•	Craat	0 2 0 0		shh	oard	cho	<u>_</u>	Shoot	- 1	nd Ch	100	2 14/0	re	renamed	
d	Dashhoard		l ollir	e a וופ. החחי בי	w ud חל nd 'ח	00011	uaiu t′r≏o	SILE	eu. tive	alv an	.⊥ai d dra	ום או ממספ	neel nd in	.∠we to th₄	a da	ashhoard s	heet
u	Fdit titles		rout	op a		onu	1 1 63	pec		iy all	u ui	ugge	.u m		_ uc		
	, _ and the of																

and Add	The title of the dashboard was changed to 'CONFIDENCE IN
additio	REPODUCIBILITY OF PUBLISHED RESEARCH', with grey background and
elemen	white text.
	 A textbox element was placed below the dashboard title with the following text to provide context to the purpose of the overall data visualization: How Many Researchers Believe That
	<parameters %x="" of="" research<="" th=""></parameters>
	believed to be reproducible>
	<parameters.indicate %x="" value=""> of</parameters.indicate>
	Research In Their Field Is Reproducible?
	1,576 researchers were surveyed, and the responses varied significantly across different fields of research.
	 A textbox was placed on top of the lollipop chart to provide instructions to
	users to "HOVER over the lollipops to view the distribution of responses"
	• A textbox was added on top of the parameters to direct users to "SELECT A SURVEY RESPONSE".

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

Dashboard Screenshot

CONFIDENCE IN REPODUCIBILITY OF PUBLISHED RESEARCH

How Many Researchers Believe That At Least 70% of Research In Their Field Is Reproducible?

1,576 researchers were surveyed, and the responses varied significantly across different fields of research. Here is the breakdown:



Source

1. Original article: https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970.

The raw survey data: <u>https://www.nature.com/news/polopoly_fs/7.36742!/file/Reproducibility%20Survey%20Raw%20Data.xlsx.</u>
 The survey questions: <u>https://www.nature.com/news/polopoly_fs/7.36741!/file/Reproducibility%20Questionnaire.doc</u>

d. Information Revealed by Proposed Data Visualisation

- 1. The confidence level that respondents show in the reproducibility of published work in their field varies significantly across each field of research.
- 2. Physics and chemistry have significantly higher count of responses in the 70%/80%/90% survey options (right-skewed distribution); This means that physicists and chemists are generally more confident of their published research.

Denise Chua (deniseadele.2019@mitb.smu.edu.sg)



%respondents believe 70% of research in their field is reproducible: 74%

Distribution of responses for Physics:



3. While Astronomy and planetary science appears second-ranked in terms of proportion of respondents that believe at least 70% of the published research in the field is reproducible, a closer look at the count of responses reveal that the statistics is unreliable since the sample size is too small



Denise Chua (deniseadele.2019@mitb.smu.edu.sg)

4. On the reverse scale, 'Other', 'Earth and environmental science' and 'engineering' appear top 3 for proportion of researchers believe that at most 30% of published research in their field is reproducible, indicating low levels of confidence.

