



Week 1: Introduction to Statistics

Unit 4: Bias

Introduction

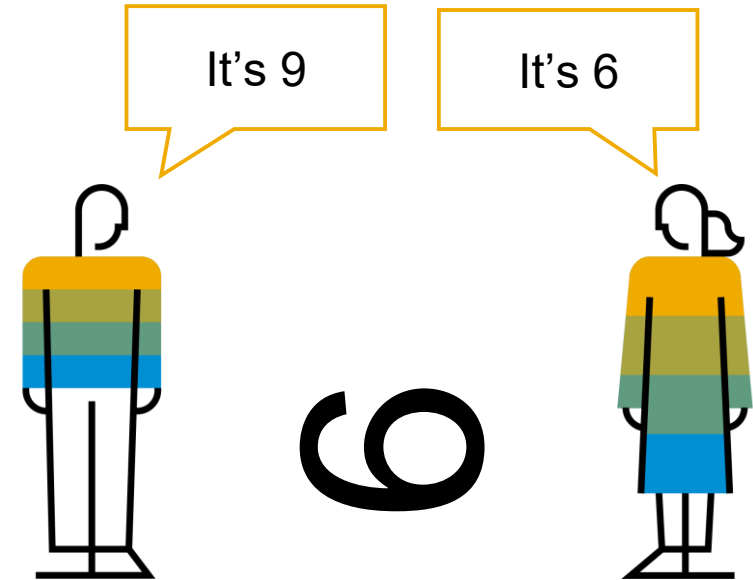
- **Bias** refers to the tendency of a measurement process to over or under-estimate the value of a population parameter.
- There are many different types of bias.



It is inevitable!

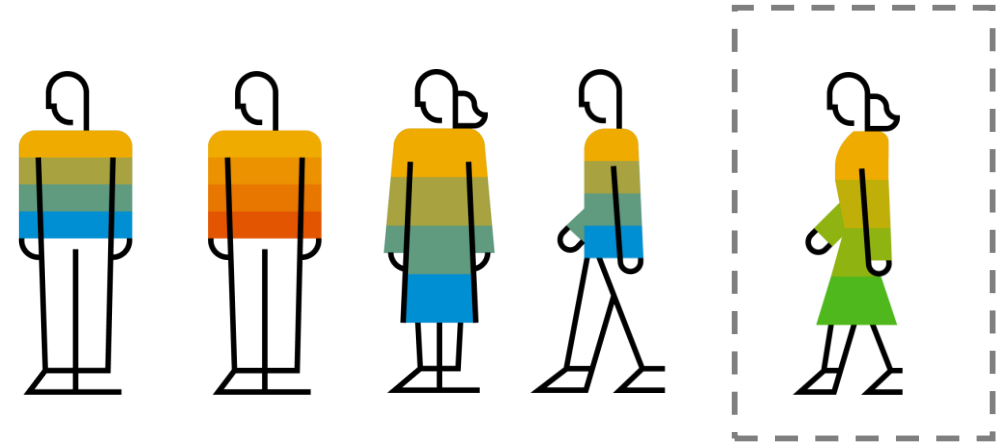
“With careful and prolonged planning, we may reduce or eliminate many potential sources of bias, but seldom will we be able to eliminate all of them. Accept bias as inevitable and then endeavor to recognize and report all exceptions that do slip through the cracks.”

Good and Hardin (2006) Common Errors in Statistics (and How to Avoid Them), p. 113



Different types of bias

- We will look at some of the different types of bias:
 - Selection/sampling bias
 - Self-selection bias
 - Confirmation bias
 - Overfitting
- “Unlike error related to random variability, bias cannot be assessed without external knowledge of the world”



Herbert I. Weisberg (2010), *Bias and Causation: Models and Judgment for Valid Comparisons*, p. 26

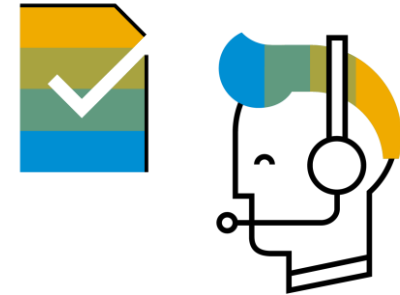
Sampling bias



Surveying of attitudes to Brexit
on Facebook
(favors younger audiences)




Face-to-face surveying of
attitudes to equality at a
football match (self-
selection)



Telephone surveying of
holiday preferences

Self-selection bias

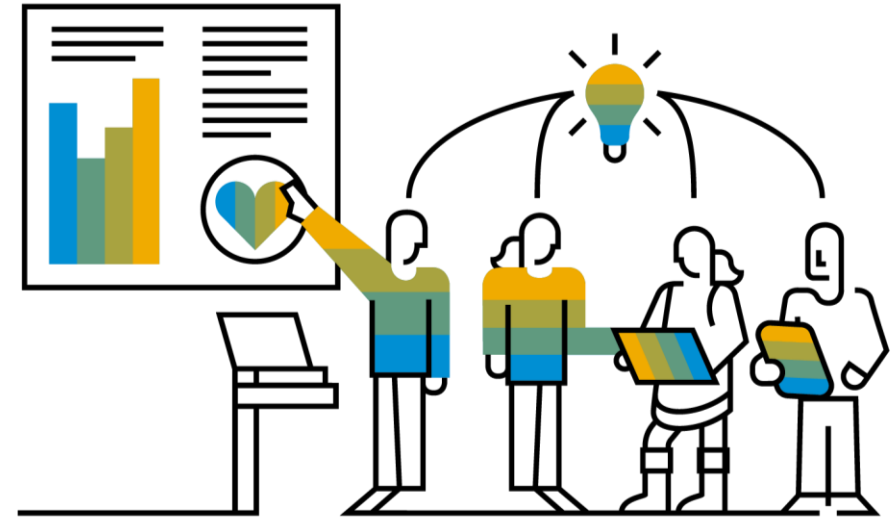
- If you use data taken from a *voluntary response sample*, i.e. the participants volunteered to take part, it becomes very difficult to avoid bias.
- The self-selected group will contain more participants with a particular set of beliefs about your study.



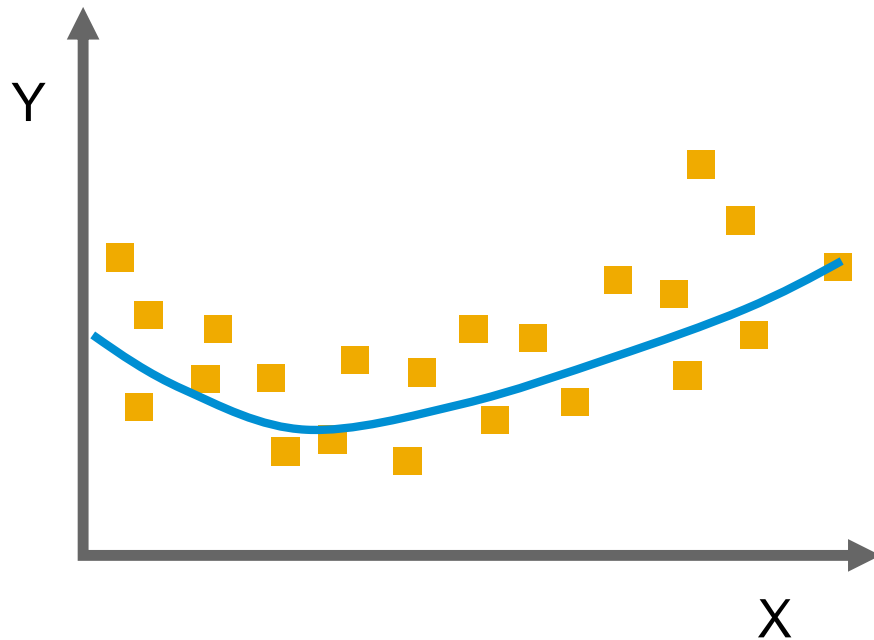
<input checked="" type="checkbox"/>	EXCELLENT
<input type="checkbox"/>	GOOD
<input type="checkbox"/>	AVERAGE
<input type="checkbox"/>	POOR

Confirmation bias

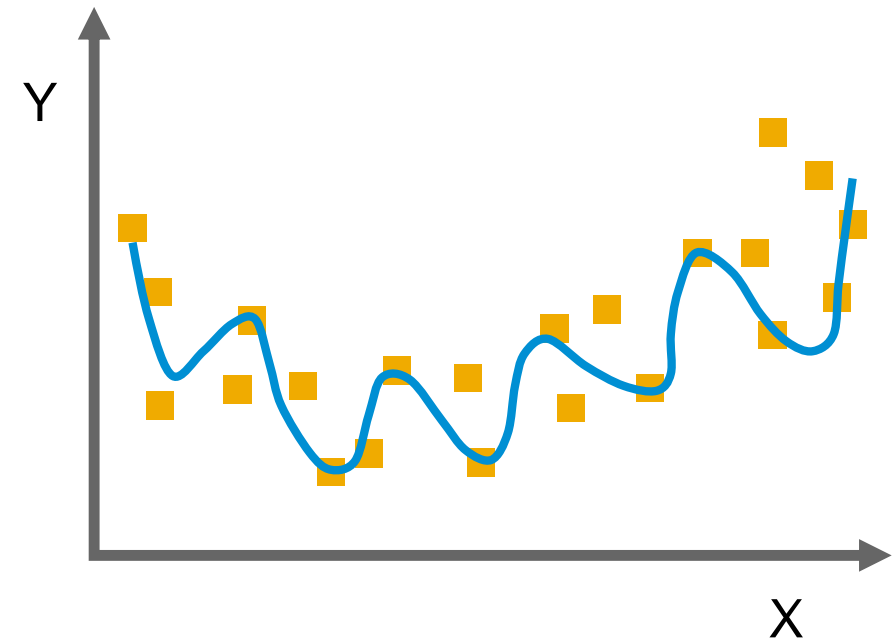
- “Confirmation bias” is one of a range of “cognitive biases” that affect how we read and interpret the insights we think we have found.
- “Cognitive bias” means that it is inbuilt into us, as humans, and how we think.
- “Confirmation bias” reflects our tendency to pick out those parts of the data and information in a way to support our previously held beliefs.



Overfitting and extrapolation



Just right!



overfitting

Bias

Summary

- It is almost impossible to avoid bias in its various forms, but an awareness of bias can help mitigate its worst effects.
- There are various forms of bias, e.g. technical, cognitive, and others, which impact what data to collect and how it should be interpreted.
- In this unit we have examined 4 key examples of bias – sampling, self-selection, confirmation, and overfitting. There are many other kinds of bias.

See: [https://en.wikipedia.org/wiki/Bias \(statistics\)](https://en.wikipedia.org/wiki/Bias_(statistics))



Thank you.

Contact information:

open@sap.com

Follow all of SAP



www.sap.com/contactsap

© 2019 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platforms, directions, and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies.

See www.sap.com/copyright for additional trademark information and notices.