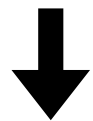


Quizzes

Missing data

by Sina Rüeger

Notes & Material



https://github.com/sinarueeger/teaching/tree/master/missing_data

The following excerpt was part of a larger questionnaire that was sent to a number of individuals.

The goal of the questionnaire was to assess income and number of hours worked.

Gender			
<input type="checkbox"/>	Female	Yearly income in CHF	<input type="text"/>
<input type="checkbox"/>	Male	Average hours worked/week	<input type="text"/>

In several cases, no gender was reported, but a yearly income and the average hours worked per week provided.

Gender			
<input type="checkbox"/>	Female	Yearly income in CHF	<input type="text" value="80'000"/>
<input type="checkbox"/>	Male	Average hours worked/week	<input type="text" value="40"/>

Questions

1. What could be the reason for the missing data in “gender”?
2. For the subsequent analysis of income and hours worked per week, observations with missing gender were removed. What is the danger of doing an analysis with removed observations that had missing values?