

# How to use pandas the wrong way

Pietro Battiston  
University of Milan Bicocca, Italy  
<http://pietrobattiston.it>

Saint Petersburg, November 4th, 2017



# Disclaimer

## Disclaimer



pandas has bugs

## Disclaimer



pandas has bugs

A screenshot of the GitHub repository page for pandas-dev/pandas. The page shows the repository name, a 'Code' button, and an 'Issues' button with a count of 2,055. Below this is a search bar with filters and a dropdown menu. At the bottom of the screenshot, it shows '2,055 Open' and '8,838 Closed' issues.

(quite a bit of them)

## Disclaimer



pandas has bugs

You are *very* welcome to help fixing them

A screenshot of the pandas GitHub repository page. The header shows 'pandas-dev / pandas'. Below the header, there are tabs for 'Code' and 'Issues 2,055'. A search bar contains the filter 'is:issue is:open'. At the bottom of the screenshot, it displays '2,055 Open' with a warning icon and '8,838 Closed' with a checkmark icon.

(quite a bit of them)

## Disclaimer



pandas has bugs

A screenshot of the GitHub repository page for pandas-dev/pandas. The page shows the repository name, a 'Code' button, and an 'Issues' button with a notification icon and the number '2,055'. Below this is a search bar with the filter 'is:issue is:open' and a 'Filters' dropdown. At the bottom of the screenshot, it shows '2,055 Open' and '8,838 Closed' issues.

(quite a bit of them)

You are very welcome to help fixing them  
(Come talk to me if you're interested, or just curious!)

## Disclaimer



pandas has bugs

A screenshot of the GitHub repository page for pandas-dev/pandas. The page shows the repository name, a 'Code' button, and an 'Issues' button with a notification icon and the number '2,055'. Below this is a 'Filters' dropdown and a search bar containing 'is:issue is:open'. At the bottom of the screenshot, it displays '2,055 Open' with a notification icon and '8,838 Closed' with a checkmark icon.

(quite a bit of them)

You are very welcome to help fixing them  
(Come talk to me if you're interested, or just curious!)

... but this talk is *not* about pandas bugs.

## Disclaimer - ctd.

This talk is not *even* about (necessarily) wrong design decisions




## Disclaimer - ctd.

This talk is not *even* about (necessarily) wrong design decisions

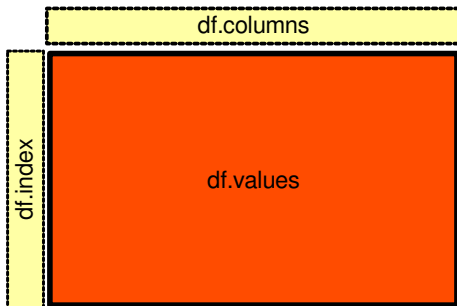
It is rather about a library which seems so intuitive<sup>1</sup> but *cannot always afford* to be intuitive.

---

<sup>1</sup>At least assuming that you find numpy intuitive. 

Let's code

## The structure of a df - outside



# The structure of a df - inside

