Method comparison and Bland-Altman plots

Robert Haase, Myers lab, MPI CBG

Based on material by Martin J. Bland
and Douglas G. Altman

June 2021

@haesleinhuepf

Method comparison studies

- Scenario
  - You work in a lab and try to improve procedures
    - Chemical protocols
    - Sample preparation
    - Analysis protocols
      - Physical measurements
      - Image analysis

Unpaired data
- Analyze independent sample sets
- Conclude about their similarity or relationship

Inferential statistics

Paired data
- The same dataset analyzed twice with different methods
- The same dataset analyzed twice with the same method

Direct method comparison; descriptive statistics

Method comparison studies

Martin Bland and Douglas Altman work on Method Comparison (excerpt)

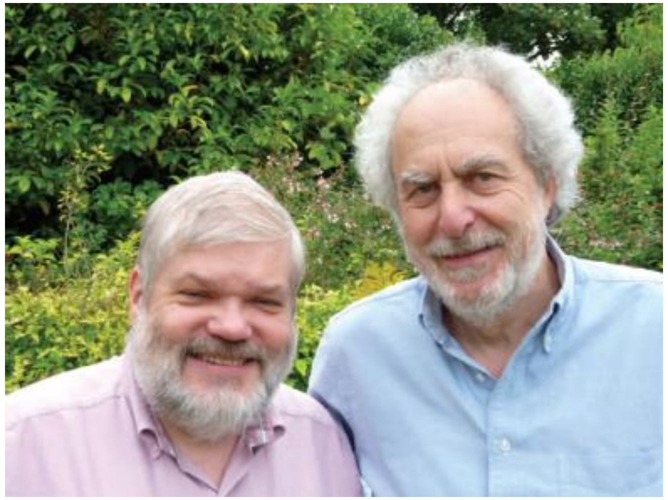
- https://www-users.york.ac.uk/~mb55/meas/ab83.pdf (Open Access)
- https://www.sciencedirect.com/science/article/pii/S0140673686908378

The Statistician 32 (1
© 1983 Institute of St

Measurem
Compariso

D. G. ALTMAN

Division of Com
Research Centre
‡ Department of
St George's Hospital Medical School, Cranmer Terrace, London SW17

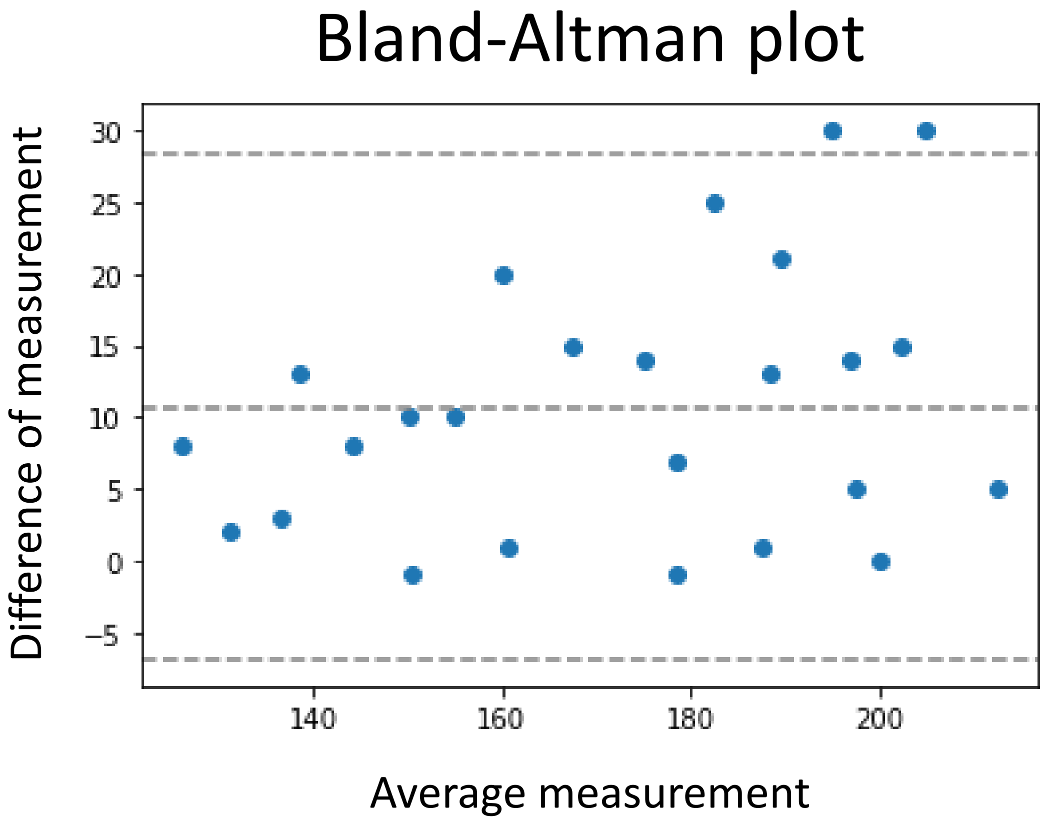


Copyright J. Martin Bland and Douglas G. Altman.

Search ScienceDirect
Advanced
Recommended articles
Comparison of refractive assessment by ...
Journal of Optometry, Volume 8, Issue 2, 2015,...
Download PDF
View details
Estimation of the concordance correlatio...
Computer Methods and Programs in Biomedi...
Purchase PDF
View details
Retinal Degeneration in Choroideremia f...
Ophthalmology, Volume 125, Issue 7, 2018, pp...
Download PDF
View details
1 2 Next
Citing articles (34736)
Feedback

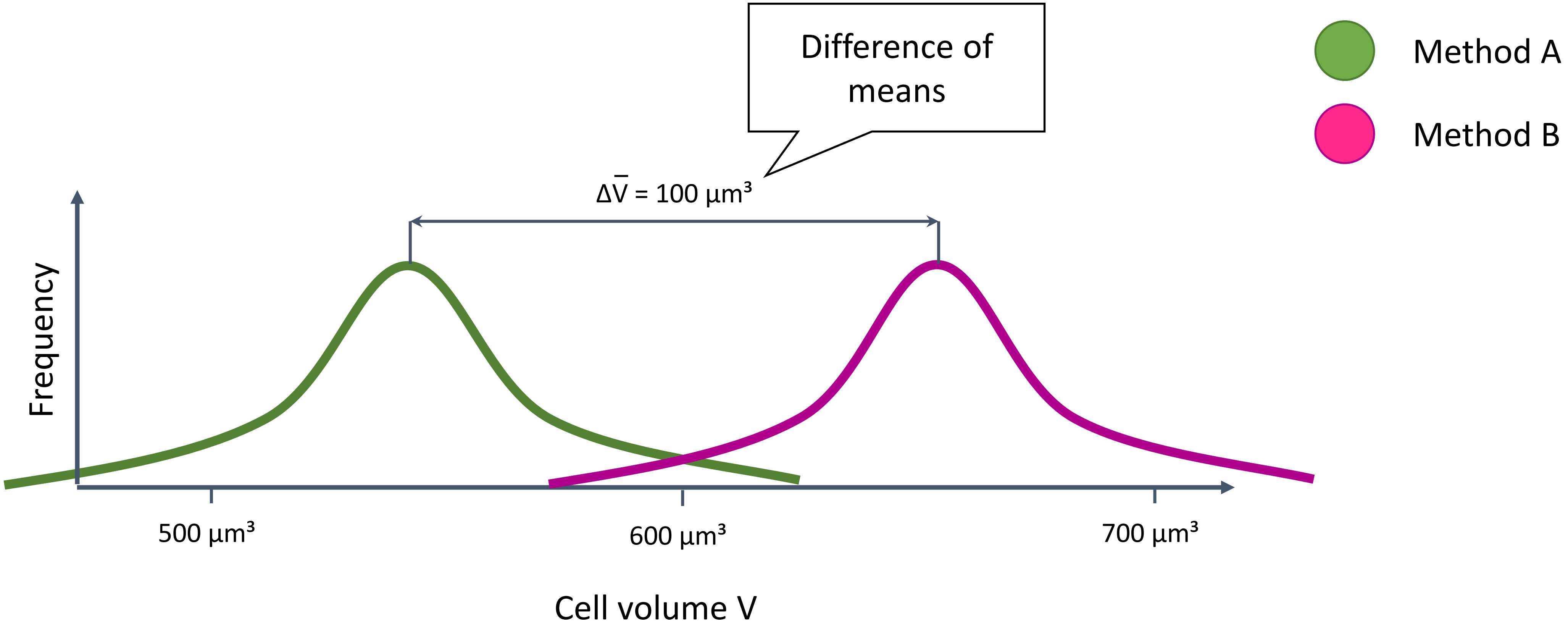
Bland-Altman plots

- Visualized method comparison



Comparison of means

- Comparing mean measurements appears reasonable on the first view.



Comparison of means

- Are two methods doing the same if their mean measurement is similar?

A	B
1	4
9	5
7	5
1	7
2	4
8	5
9	4
2	6
1	6
7	5
8	4

```
Mean(A) = 5.0
Mean(B) = 5.0
```

What if means were "very" different?

Yes

Method B cannot replace method A

No

Similar means is a necessary condition, but is it sufficient?

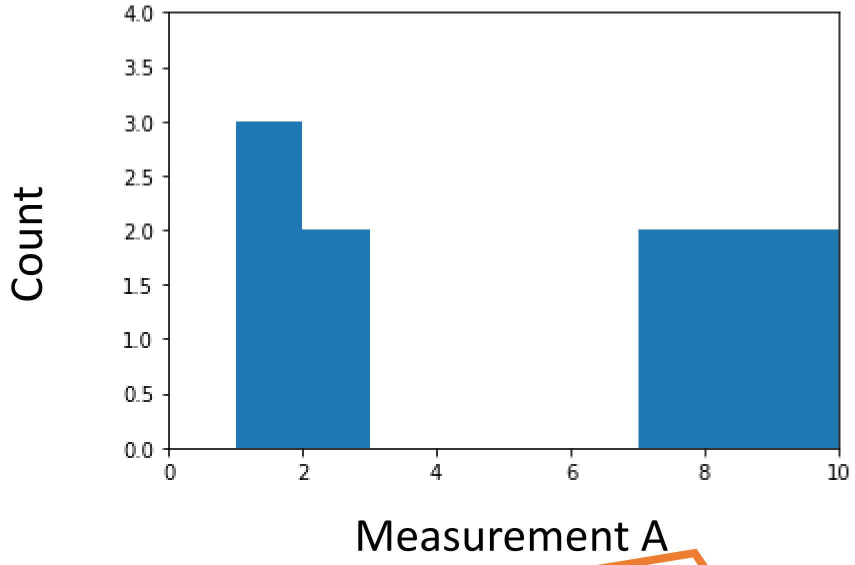
Comparison of means

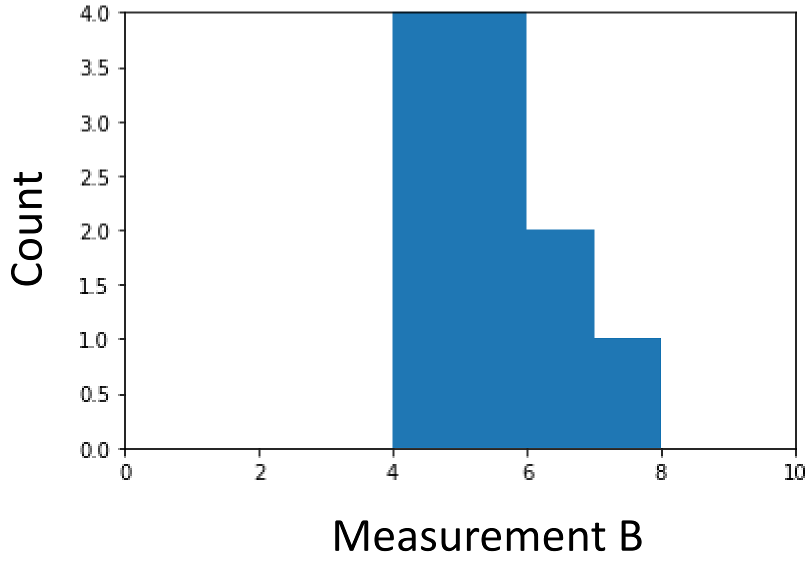
- Are two methods doing the same if their mean measurement is similar?

A	B
1	4
9	5
7	5
1	7
2	4
8	5
9	4
2	6
1	6
7	5
8	4

```
Mean(A) = 5.0
Mean(B) = 5.0
```

- Draw histograms! How can two methods do the same if histograms from their measurements are different?





Similar means is a necessary condition, but it is NOT sufficient!

The scientific method: Show that a method doesn't work with *just one* example. And you have *proven* that the method doesn't work in general.

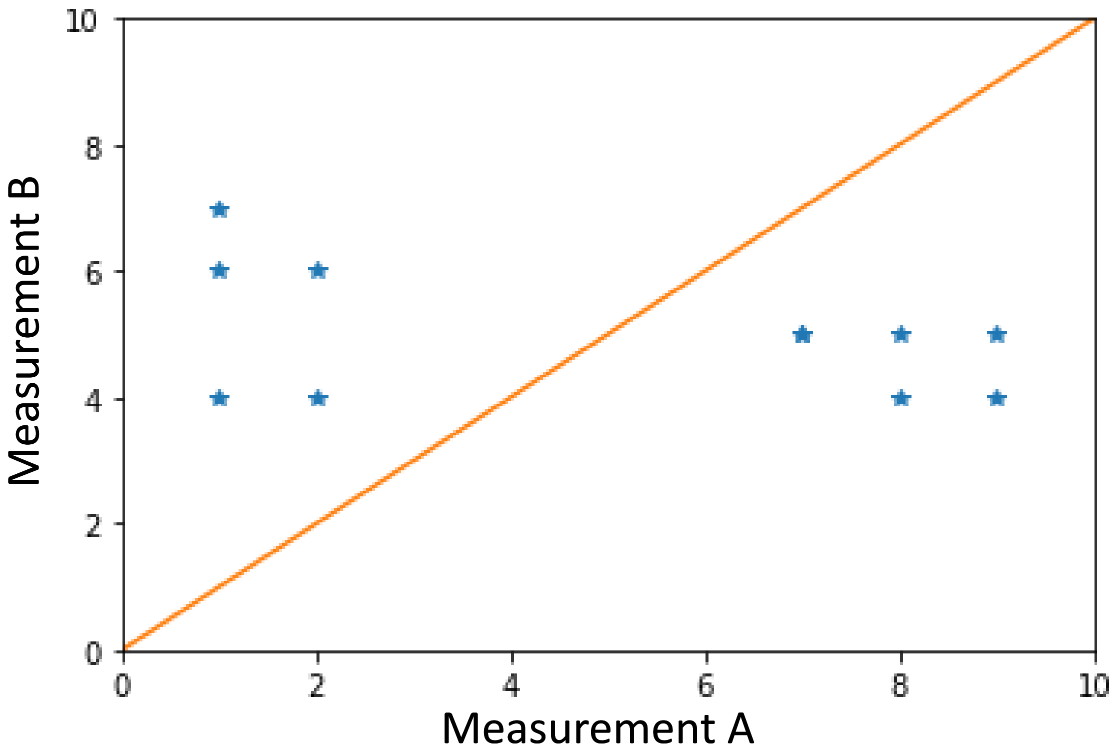
Comparison of means

- Are two methods doing the same if their mean measurement is similar?

A	B
1	4
9	5
7	5
1	7
2	4
8	5
9	4
2	6
1	6
7	5
8	4

```
Mean(A) = 5.0
Mean(B) = 5.0
```

- Plot the measurements against each other. What does it mean if they lie on a straight line? What if not?



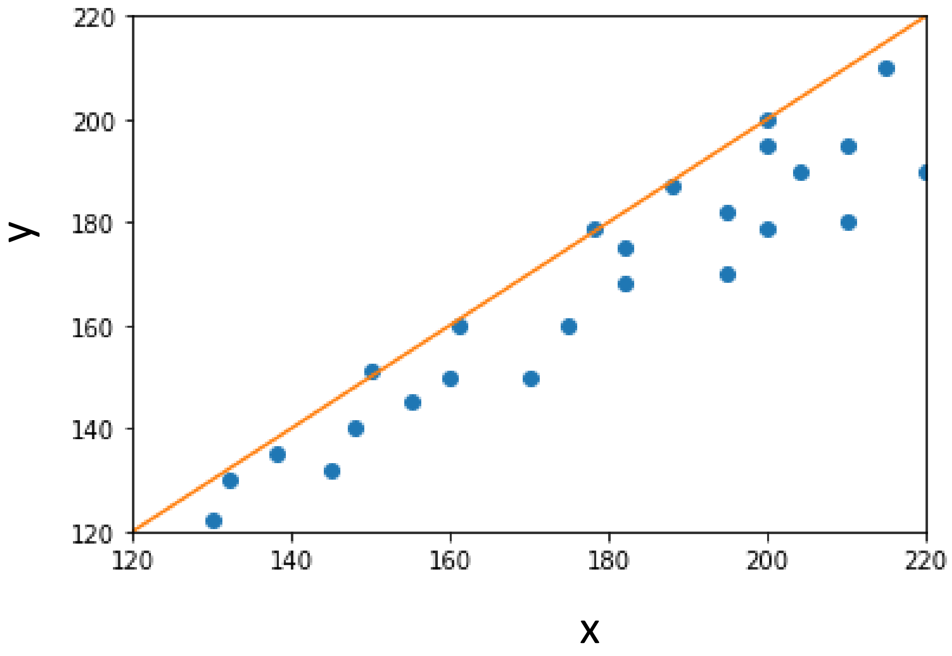
The "criterion of agreement was that the two methods gave the same mean measurement; 'the same' appears to stand for 'not significantly different'. Clearly, this approach tells us very little about the accuracy of the methods."[1]

[1] Altman & Bland, The Statistician 32, 1983

Correlation

- Are two methods doing the same if they correlate?
  - Correlation: Any kind of relationship.
  - Measurable; e.g. using Pearson's Correlation Coefficient *r* enumerated linear correlation.

Comparison of two methods of measuring systolic blood pressure (Data take from [1])



Expectation E

Mean average μ

$$r(X,Y) = \frac{E(X - \mu_X)(Y - \mu_y)}{\sigma_X \sigma_Y}$$

Standard deviation σ

Disclosure: Mean and standard deviation must be obtained from the whole population or from a sample set which is sufficiently large.

In practice *E* is the weighted sum:

$$r(X,Y) = \frac{\sum_{x \in X, y \in Y} \frac{(x - \mu_X)(y - \mu_Y)}{n}}{\sigma_X \sigma_Y}$$

Number of measurements *n*

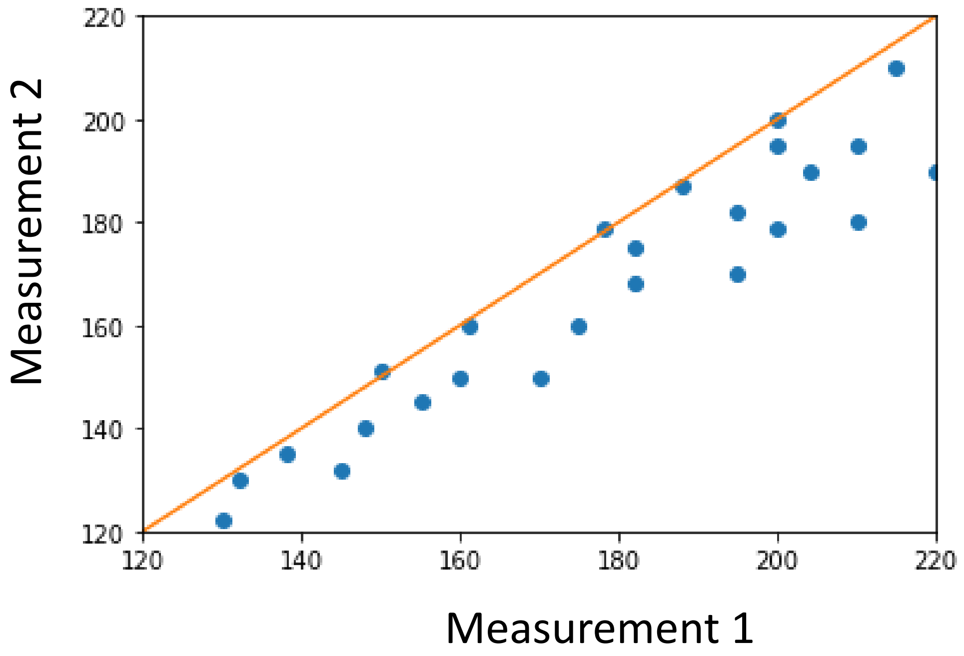
[1] Altman & Bland, The Statistician 32, 1983

Correlation

Physics of Life
TU Dresden

- Are two methods doing the same if they correlate?
  - Correlation: Any kind of relationship.
  - Measurable; e.g. using Pearson's Correlation Coefficient *r* enumerated linear correlation.

Comparison of two methods of measuring systolic blood pressure (Data take from [1])



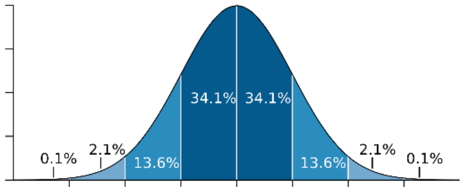
$$r(X,Y) = \frac{\sum_{x \in X, y \in Y} \frac{(x-\mu_X)(y-\mu_Y)}{n}}{\sigma_X \sigma_Y} = 0.94$$

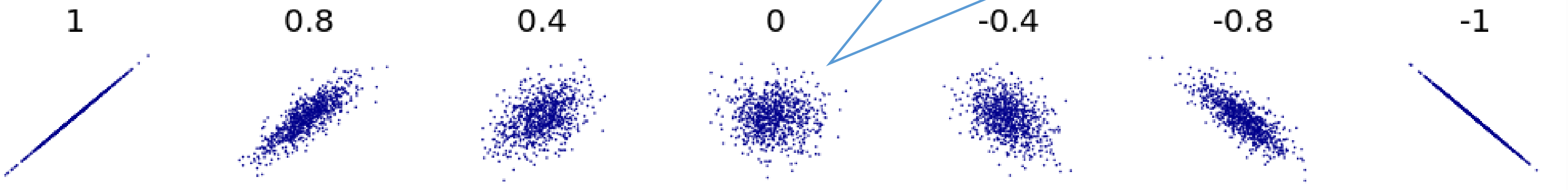
[1] Altman & Bland, The Statistician 32, 1983

Correlation: Pearson's *r*

- Pearson's *r* lies between -1 and 1
  - 1: Positive linear correlation
  - 0: No linear correlation
  - -1: Negative linear correlation

2-dimensional normal distribution



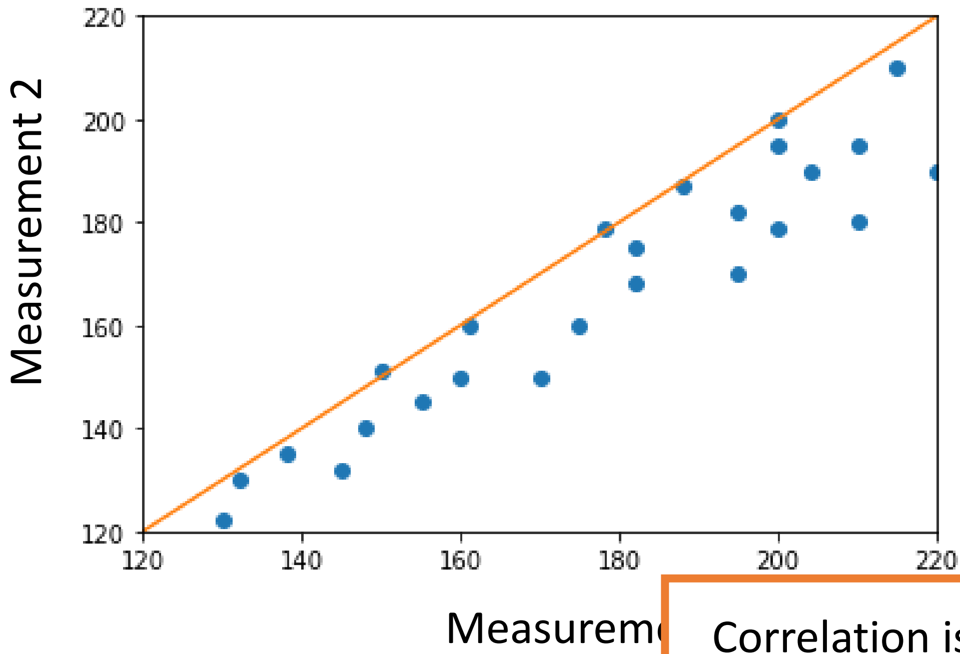


https://en.wikipedia.org/wiki/Pearson_correlation_coefficient#/media/File:Correlation_examples2.svg

Correlation

- Are two methods doing the same if they correlate?
  - Correlation: Any kind of relationship.
  - Measurable; e.g. using Pearson's Correlation Coefficient *r* enumerated linear correlation.

Comparison of two methods of measuring systolic blood pressure (Data take from [1])



$$r(X,Y) = \frac{\sum_{x\in X, y\in Y} \frac{(x-\mu_X)(y-\mu_Y)}{n}}{\sigma_X \sigma_Y} = 0.94$$

"Positive linear correlation"

Measurement 1 is almost always larger than measurement 2

Correlation is necessary, but it is NOT sufficient!

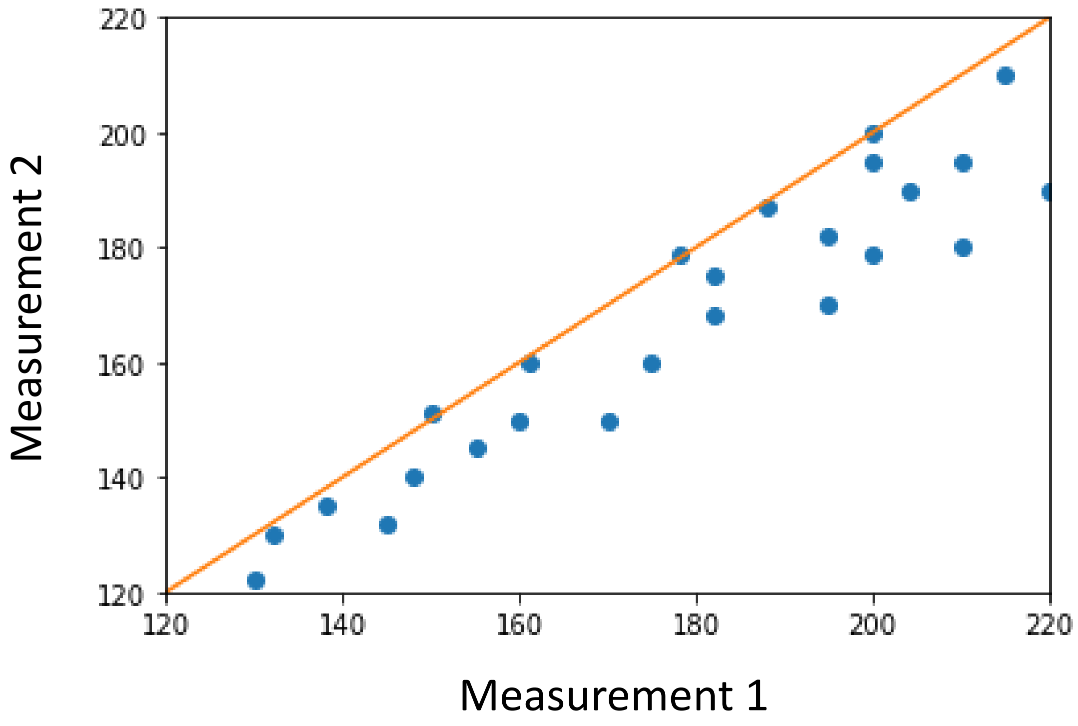
The scientific method: Show that a method doesn't work with *just one* example. And you have *proven* that the method doesn't work in general.

[1] Altman & Bland, The Statistician 32, 1983

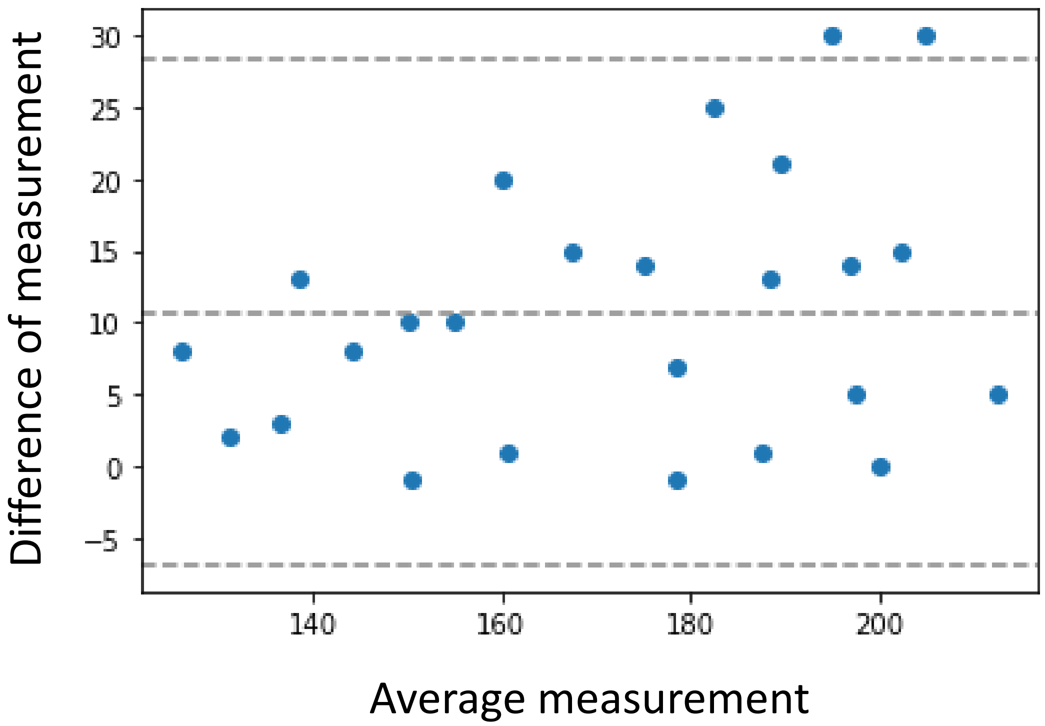
Bland-Altman plots

- In order to evaluate the difference between two methods, you should visualize them first.
- *“The purpose of computing is insight, not numbers.”*, Richard Hamming

Scatter plot



Bland-Altman plot



The confidence interval

- “The British Standards Institution (1979) define a coefficient of repeatability as ‘the value below which the difference between two single test results ... may be expected to lie with a specified probability; in the absence of other indications, the probability is 95 per cent’.”[1]

The confidence interval *CI* of agreement

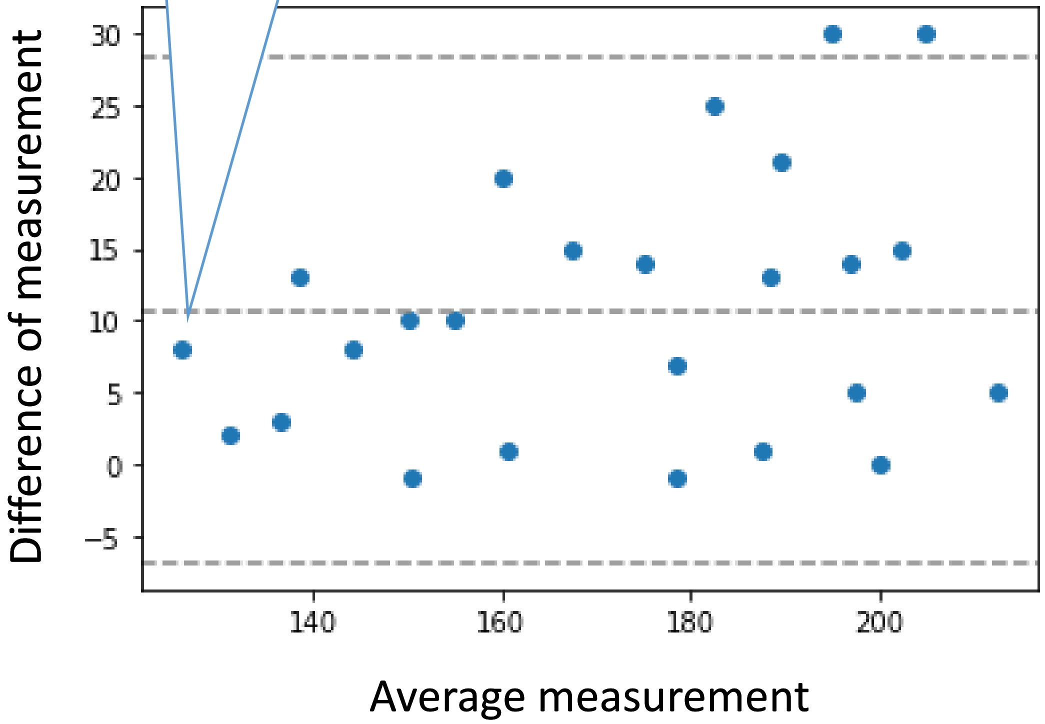
$$CI = (\mu - 2\sigma, \mu + 2\sigma)$$

Mean difference

Standard deviation of differences

The mean difference between the methods

Bland-Altman plot

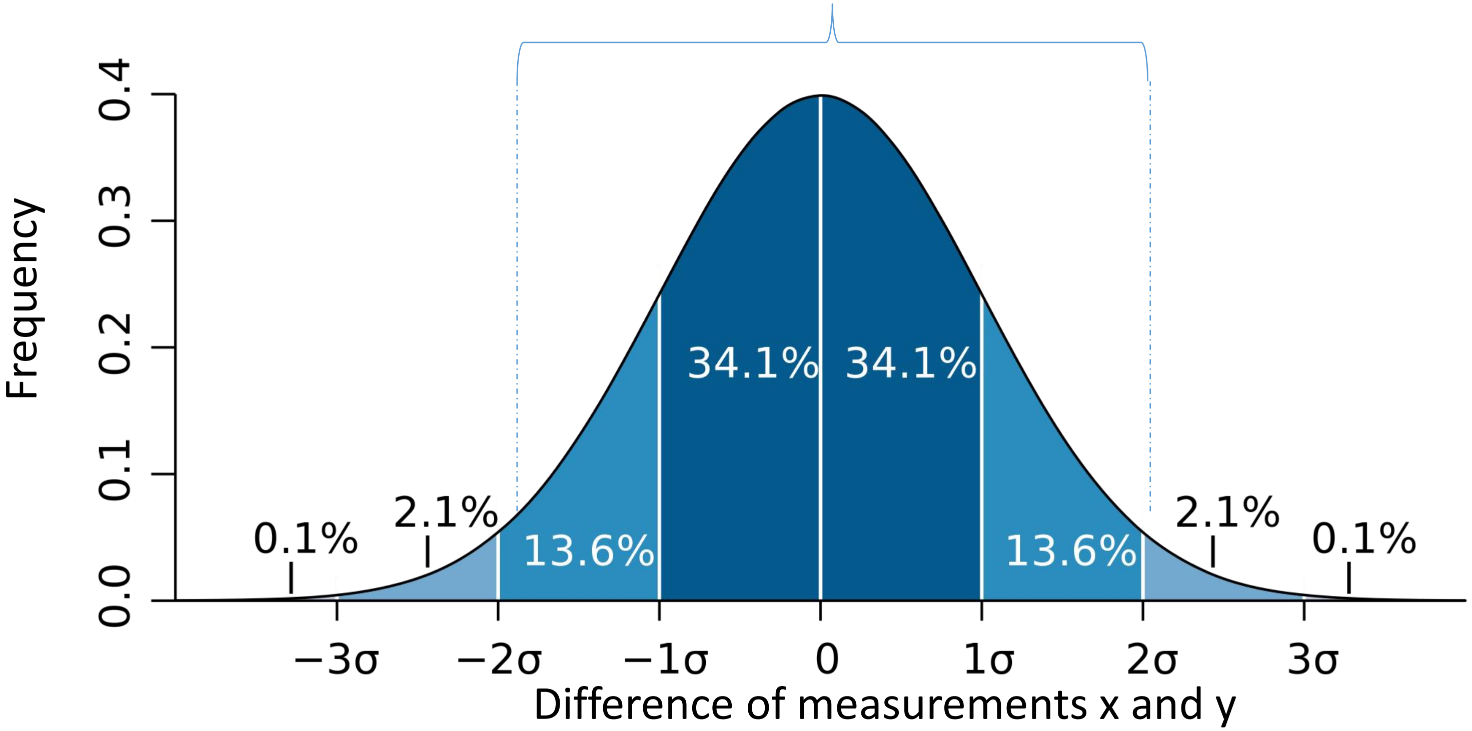


[1] Altman & Bland, The Statistician 32, 1983

The confidence interval & the coefficient of repeatability.

- "The British Standards Institution (1979) define a coefficient of repeatability as 'the value below which the difference between two single test results ... may be expected to lie with a specified probability; in the absence of other indications, the probability is 95 per cent'."[1]

In a normal distribution 95% of measurements lie in this area around the mean

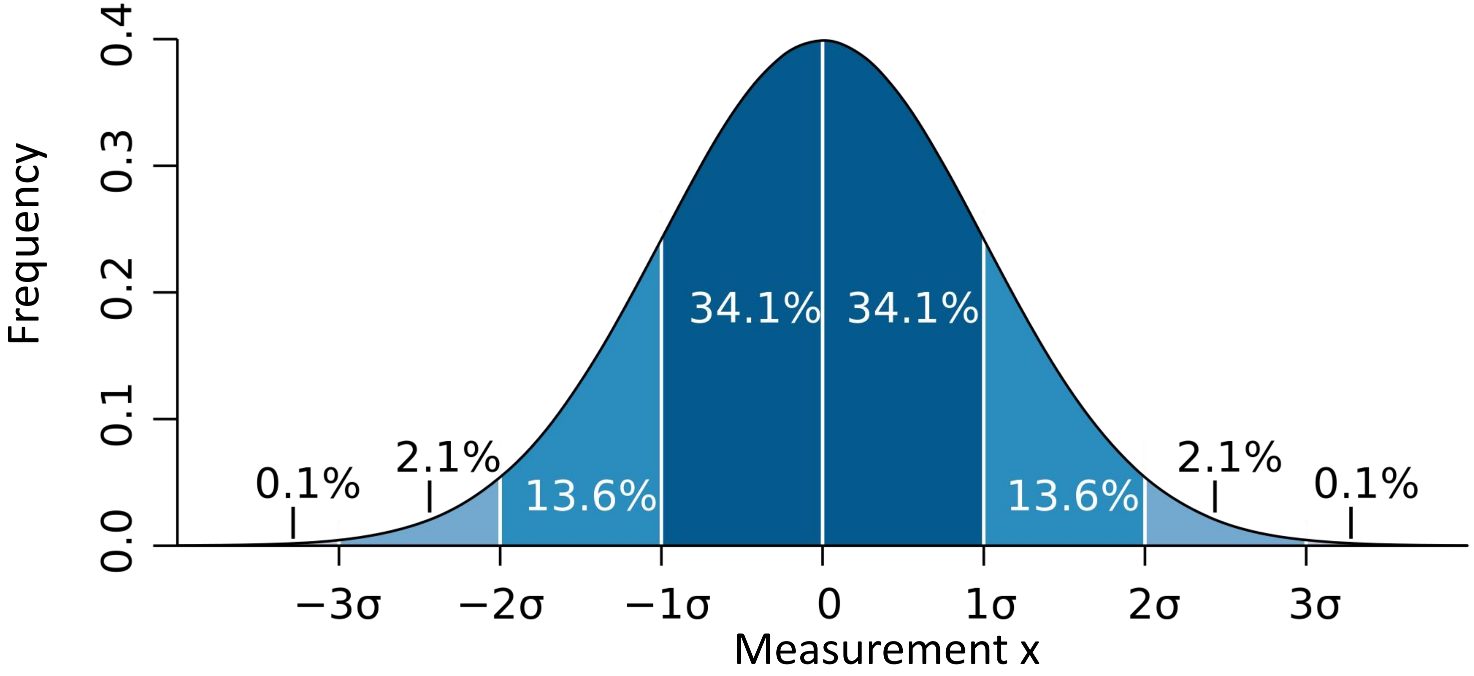


Graph adapted from: M. W. Toews - Own work, based (in concept) on figure by Jeremy Kemp, on 2005-02-09, CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=1903871

[1] Altman & Bland, The Statistician 32, 1983

The confidence interval & the coefficient of repeatability.

- "The British Standards Institution (1979) define a coefficient of repeatability as 'the value below which the difference between two single test results ... may be expected to lie with a specified probability; in the absence of other indications, the probability is 95 per cent'."[1]
- If the two measurements come from the same method which just repeated twice, we can assume that the main difference is zero. The coefficient of repeatability *CR* can then be estimated: It's the standard deviation of differences.[2]



$$\mathrm{CR}(X,Y) = \sqrt{\sum_{x\in X, y\in Y} \frac{(x-y)^2}{n}}$$

Graph adapted from: M. W. Toews - Own work, based (in concept) on figure by Jeremy Kemp, on 2005-02-09, CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=1903871

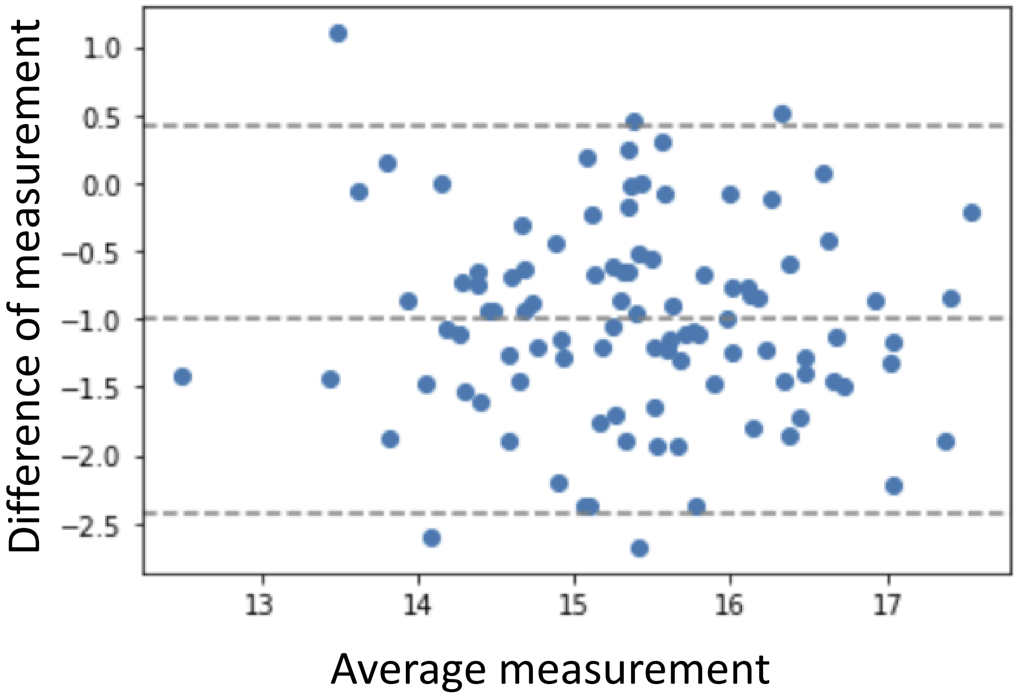
[1] Altman & Bland, The Statistician 32, 1983

2 Bland & Altman, Lancet , 1986

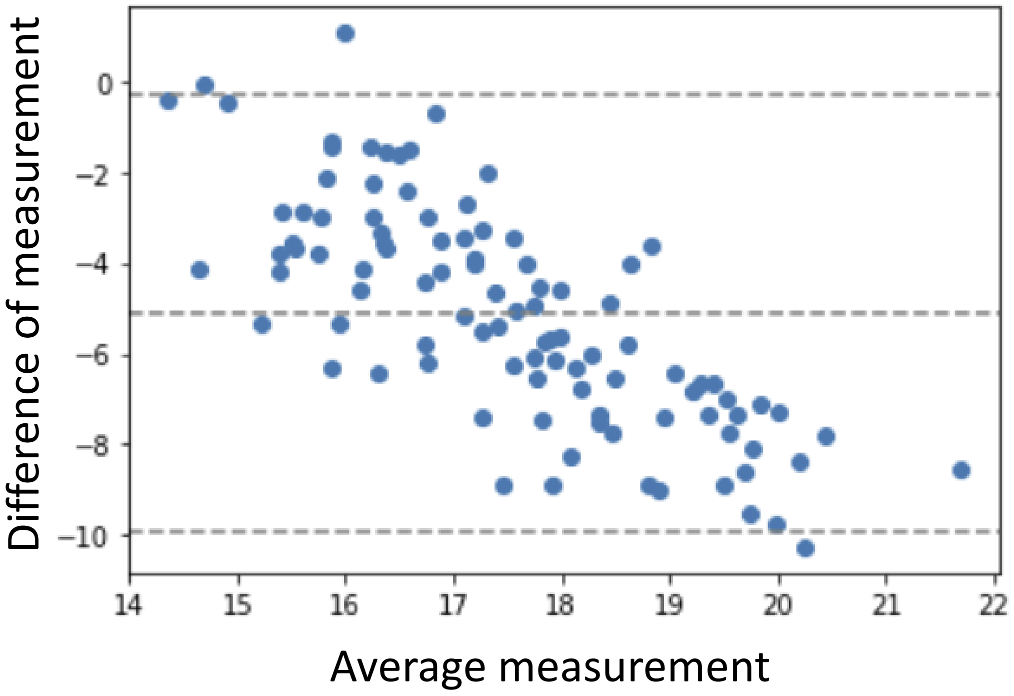
In practice:

- Bland-Altman plots allow us to differentiate various kinds of bias.

Agreement with additive error (offset)



Multiplicative bias



Processing tables with Python

Robert Haase

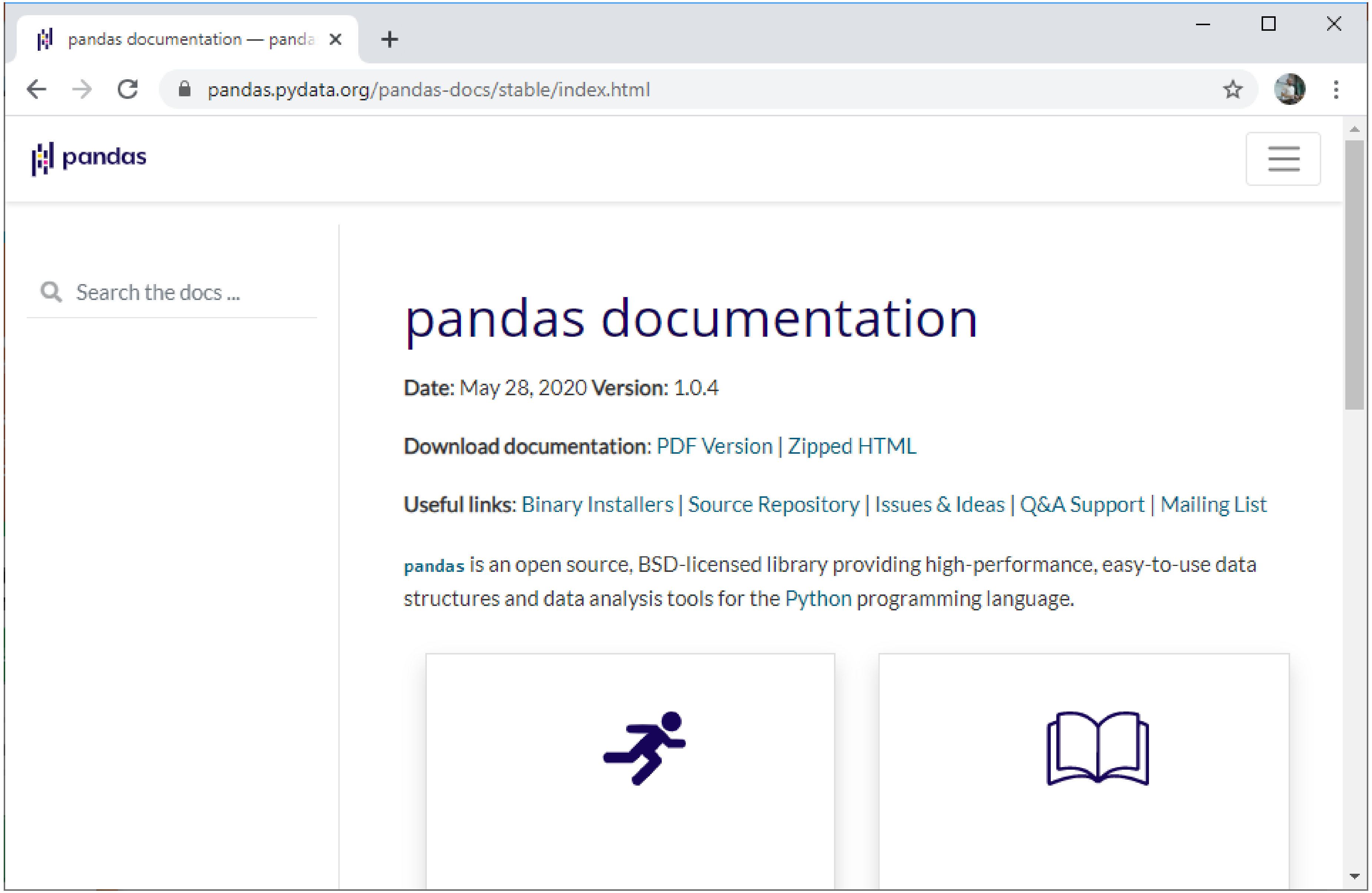
June 2021

@haesleinhuepf

Pandas

- pandas is a library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

```
conda install pandas
```



Processing tables with pandas

- Typical use-case:
  - You get data from a colleague in form of a table.
  - Using pandas, you can analyze it in python.

Results

File Edit Font Results

	Area	Mean	Circ.	AR	Round	Solidity
1	2610	96.920	0.773	1.289	0.776	1.000
2	2100	90.114	0.660	2.333	0.429	1.000
3	27	110.222	0.108	27.000	0.037	1.000

- Loading a table in python using pandas:

```python
import pandas as pd

data_frame = pd.read_csv("Measurements_ImageJ.csv", delimiter=',')
data_frame
```

		Area	Mean	Circ.	AR	Round	Solidity
0	1	2610	96.920	0.773	1.289	0.776	1.0
1	2	2100	90.114	0.660	2.333	0.429	1.0
2	3	27	110.222	0.108	27.000	0.037	1.0

Processing tables with pandas

- Typical use-case:
  - You get data from a colleague in form of a table.
  - Using pandas, you can analyze it in python.

Results

File Edit Font Results

	Area	Mean	Circ.	AR	Round	Solidity
1	2610	96.920	0.773	1.289	0.776	1.000
2	2100	90.114	0.660	2.333	0.429	1.000
3	27	110.222	0.108	27.000	0.037	1.000

- Accessing a column

```
data_frame["Mean"]
```

```
0     96.920
1     90.114
2    110.222
Name: Mean, dtype: float64
```

- Determining mean of a column

```
import numpy as np
np.mean(data_frame["Mean"])
```

```
99.08533333333332
```

- Accessing an individual cell

```
data_frame["Mean"][0]
```

```
1.2890000000000001
```

Processing tables with pandas

- Creating tables with pandas

- Creating a new table

```
header = ['A', 'B', 'C']

data = [
    [1, 2, 3],  # this will later be colum A
    [4, 5, 6],  #                          B
    [7, 8, 9]   #                          C
]

# convert the data and header arrays in a pandas data frame
data_frame = pd.DataFrame(data, header)

# show it
data_frame
```

	0	1	2
A	1	2	3
B	4	5	6
C	7	8	9

- Rotate a table

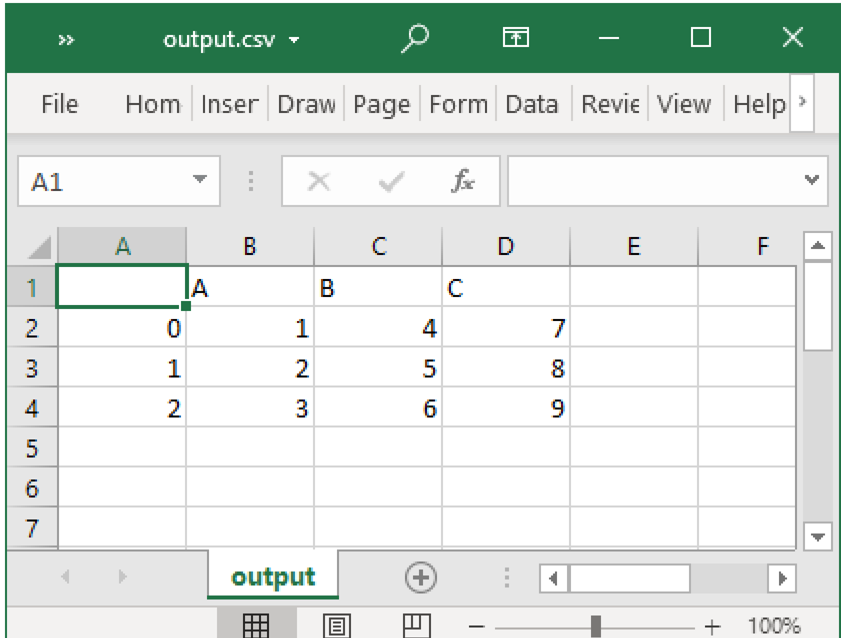
```
# rotate/flip it
data_frame = data_frame.transpose()

# show it
data_frame
```

	A	B	C
0	1	4	7
1	2	5	8
2	3	6	9

- Save it to disc

```
# save a dataframe to a CSV
data_frame.to_csv("output.csv")
```



Selecting rows and columns

	City	Country	Population	Area_km2
0	Tokyo	Japan	13515271	2191
1	Delhi	India	16753235	1484
2	Shanghai	China	24183000	6341
3	Sao Paulo	Brazil	12252023	1521
4	Mexico City	Mexico	9209944	1485

- Selecting columns

```
cities[['City', 'Country']]
```

	City	Country
0	Tokyo	Japan
1	Delhi	India
2	Shanghai	China
3	Sao Paulo	Brazil
4	Mexico City	Mexico

- Selecting rows

```
cities[cities['Area_km2'] > 2000]
```

	City	Country	Population	Area_km2
0	Tokyo	Japan	13515271	2191
2	Shanghai	China	24183000	6341

Combining tables

- The big art in data science is the ability of combining information from multiple sources to gain new insights.

	Country	Population
0	Japan	127202192
1	India	1352642280
2	China	1427647786
3	Brazil	209469323
4	Mexico	126190788

	City	Country	Population	Area_km2
0	Tokyo	Japan	13515271	2191
1	Delhi	India	16753235	1484
2	Shanghai	China	24183000	6341
3	Sao Paulo	Brazil	12252023	1521
4	Mexico City	Mexico	9209944	1485

```
combined = countries.merge(cities, on='Country', suffixes=['_country', '_city'])
combined
```

	Country	Population_country	City	Population_city	Area_km2
0	Japan	127202192	Tokyo	13515271	2191
1	India	1352642280	Delhi	16753235	1484
2	China	1427647786	Shanghai	24183000	6341
3	Brazil	209469323	Sao Paulo	12252023	1521
4	Mexico	126190788	Mexico City	9209944	1485

```
# compute ratio
combined['City_Country_population_ratio'] = combined['Population_city'] / combined['Population_country']

# only show selected columns
combined[['City', 'City_Country_population_ratio']]
```

	City	City_Country_population_ratio
0	Tokyo	0.106250
1	Delhi	0.012386
2	Shanghai	0.016939
3	Sao Paulo	0.058491
4	Mexico City	0.072984

PoL
Physics of Life
TU Dresden

CENTER FOR SYSTEMS BIOLOGY DRESDEN

Functional parameters in Python

Robert Haase

June 2021

@haesleinhuepf

Functional parameters

- So far, we know variables can contain numbers, list, strings, images, …
- We also know functions.
- Variables can also be functions.

```python
def double_number(x):
    return x * 2
```

```python
double_number(values)
```

```
array([ 2,  4,  6,  8, 20])
```

```python
import numpy as np

values = np.asarray([1, 2, 3, 4, 10])
```

```python
my_function = double_number

my_function(values)
```

```
array([ 2,  4,  6,  8, 20])
```

Functional parameters

- Using the same concept, you can pass over a function to a function.
- Function will then execute function.

```python
import matplotlib.pyplot as plt
from skimage.measure import label

def count_blobs(image, threshold_algorithm):
    # binarize the image using a given
    # threshold-algorithm
    threshold = threshold_algorithm(image)
    binary = image > threshold

    # show intermediate result
    # plt.imshow(binary)

    # return count blobs
    labels = label(binary)
    return labels.max()
```

```python
from skimage.filters import threshold_otsu

count_blobs(blobs_image, threshold_otsu)
```

```
64
```

```python
from skimage.filters import threshold_yen

count_blobs(blobs_image, threshold_yen)
```

```
67
```

Useful for method comparison ;-)

Summary

- Descriptive statistics
    - Method comparison
        - Comparison of means (not sufficient!)
        - Pearson's correlation coefficient
        - Scatter plots
        - Bland-Altman plots
        - Confidence interval
        - Coefficient of repeatability.
- Coming up next
    - Hypothesis testing