

MMI 1

Evaluate Results of a User Study

Author: Alexander De Luca + Bettina Conradi (for lecture MMI1 SS11) -
LMU Munich

Study Design I

Compare 2 text input methods on a mobile phone

Type of study: Quantitative evaluation

Method 1: QWERTY

Method 2: Swype

H1: Swype is faster than QWERTY

H2: Swype is less error-prone than QWERTY

Study Design II

Task: Type 10 given sentences with each method.

Participants: at least 10 people

Dependent variables: time (stop watch), error rate (log)

Independent variables: text input method

Within subjects design

Counterbalancing: even number of participants, 50% start with method 1, 50% with method 2

Study Design III

Drawback: Participants might be used to QWERTY and not to Swype text input

-> Training beforehand: before logging time and error rate for each method, participant can train method with writing 2 sentences.

Evaluate Results - Exemplary Study

Task and Results are fictional!!

Task: Compare web vs. mobile app for getting trip information from Munich to Berlin on www.bahn.de

Results in seconds:

Web	Mobile App
50.7	52.6
46.8	50.8
52.3	49.9
49.6	51.9
56.2	56.2
47.6	52.7
52.1	54.8
49.3	56.3
47.5	49.8
51.4	51.6

Average	50.4	52.7
---------	------	------

-> Average values of both do not provide enough insights!

How to evaluate results: statistics

When & why to use which test -> see lecture

Statistical tests can be calculated:

- by hand ;)
- Excel/OpenOffice (statistical test are rather limited)
- SPSS (license required)
- R (for free)
- ...

⇒ This tutorial will focus on Excel and R

Excel – Box-and-Whisker Diagram

Visualize distribution of results

-> Create a Box-and-Whiskers Diagram:

Calculate min/max, median, 1st/3rd quartile (e.g. values in B2:B10):

=MAX(B2:B10)

=PERCENTILE(B2:B10,0.75)

=MEDIAN(B2:B10)

=PERCENTILE(B2:B10,0.25)

=MIN(B2:B10)

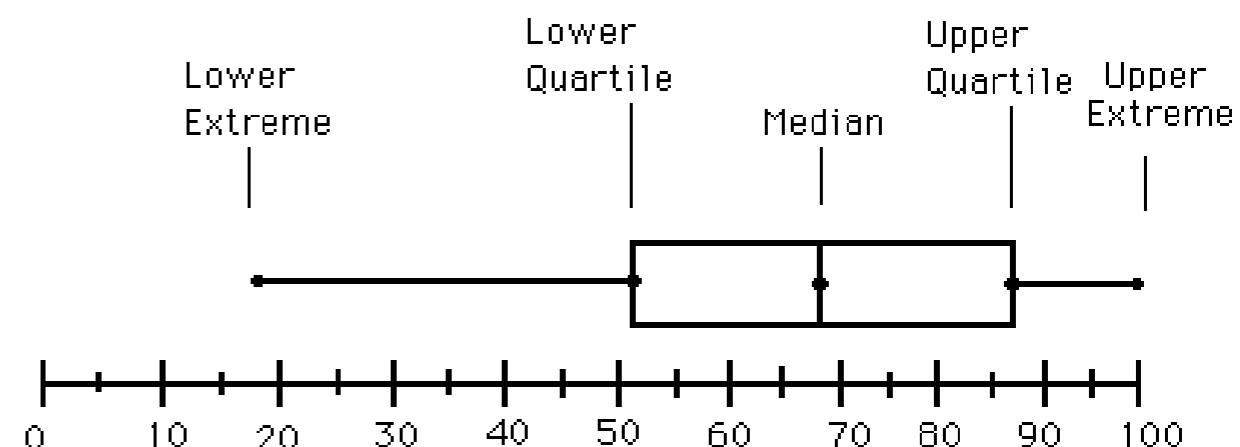


Image from <http://ellerbruch.nmu.edu/cs255/jnord/boxplot.html>

Tutorials: <http://blog.immeria.net/2007/01/box-plot-and-whisker-plots-in-excel.html>

<http://www.bloggpro.com/box-plot-for-excel-2007/>

Excel – Calculate absolute/relative values

Create a Box-and-Whiskers Diagram:

	Web	Mobile
	50.7	52.6
	46.8	50.8
	52.3	49.9
	49.6	51.9
	56.2	56.2
	47.6	52.7
	52.1	54.8
	49.3	56.3
	47.5	49.8
	51.4	51.6
Max	56.2	56.3
3rd Quartile	51.9	54.3
Median	50.2	52.3
1st Quartile	48.0	51.0
Min	46.8	49.8

absolute

In order to create a box plot with Excel we do not need absolute values (left), but relative values (right)

(relative to lowest drawn box -> 1st Quartile)



Max-3rd Quartile
3rd Quartile-Median
Median-1st quartile
1st Quartile-Min

Max	4.3	2.0
3rd Quartile	1.8	2.0
Median	2.1	1.3
1st Quartile	48.0	51.0
Min	1.2	1.2

Relative to 1st Quartile

Excel – Create a stacked column chart I

The screenshot shows a Microsoft Excel interface. The ribbon at the top includes tabs for File, Home, Insert, Page Layout, Formulas, Review, View, and Acrobat. The Insert tab is selected. Below the ribbon, there are icons for PivotTable, Table, Picture, Clip Art, Shapes, SmartArt, and Screenshot. The main area displays a table of data with columns A, B, and C. Row 1 contains headers 'Web' and 'Mobile'. Rows 2 through 12 contain numerical values. Rows 13 through 17 are labeled 'Max', '3rd Quartile', 'Median', '1st Quartile', and 'Min' respectively. Rows 19 through 23 also have these labels. A red box highlights the range from B20 to C23. A callout bubble with the number '1.' points to this highlighted range. The chart gallery is open, showing various chart types under the 'Column' category. The 'Stacked Column' option is selected, with a callout bubble containing the text: 'Compare the contribution of each value to a total across categories by using vertical rectangles.' and 'Use it to emphasize the total across series for one category.' A red box highlights the 'Stacked Column' icon. A red number '2.' is placed over the 'Stacked Column' icon.

	A	B	C
1		Web	Mobile
2		50,7	52,6
3		46,8	50,8
4		52,3	49,9
5		49,6	51,9
6		56,2	56,2
7		47,6	52,7
8		52,1	54,8
9		49,3	56,3
10		47,5	49,8
11		51,4	51,6
12			
13	Max	56,2	56,3
14	3rd Quartile	51,9	54,3
15	Median	50,2	52,3
16	1st Quartile	48,0	51,0
17	Min	46,8	49,8
18			
19	Max	4,3	2,0
20	3rd Quartile	1,8	2,0
21	Median	2,1	1,3
22	1st Quartile	48,0	51,0
23	Min	1,2	1,2

1. select 3rd Quartile, Median and 1st Quartile values

2. create a stacked column chart

-> resulting chart has wrong rows/columns

Excel – Create a stacked column chart II

The screenshot shows a Microsoft Excel spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Web	Mobile											
2		50,7	52,6											
3		46,8	50,8											
4		52,3	49,9											
5		49,6	51,9											
6		56,2	56,2											
7		47,6	52,7											
8		52,1	54,8											
9		49,3	56,3											
10		47,5	49,8											
11		51,4	51,6											
12														
13	Max	56,2	56,3											
14	3rd Quartile	51,9	54,3											
15	Median	50,2	52,3											
16	1st Quartile	48,0	51,0											
17	Min	46,8	49,8											
18														
19	Max	4,3	2,0											
20	3rd Quartile	1,8	2,0											
21	Median	2,1	1,3											
22	1st Quartile	48,0	51,0											
23	Min	1,2	1,2											
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														

The chart area displays two stacked bars. The first bar represents the 'Web' category with values: Series1 (blue) = 5, Series2 (red) = 2, Series3 (green) = 43. The second bar represents the 'Mobile App' category with values: Series1 (blue) = 2, Series2 (red) = 1, Series3 (green) = 51.

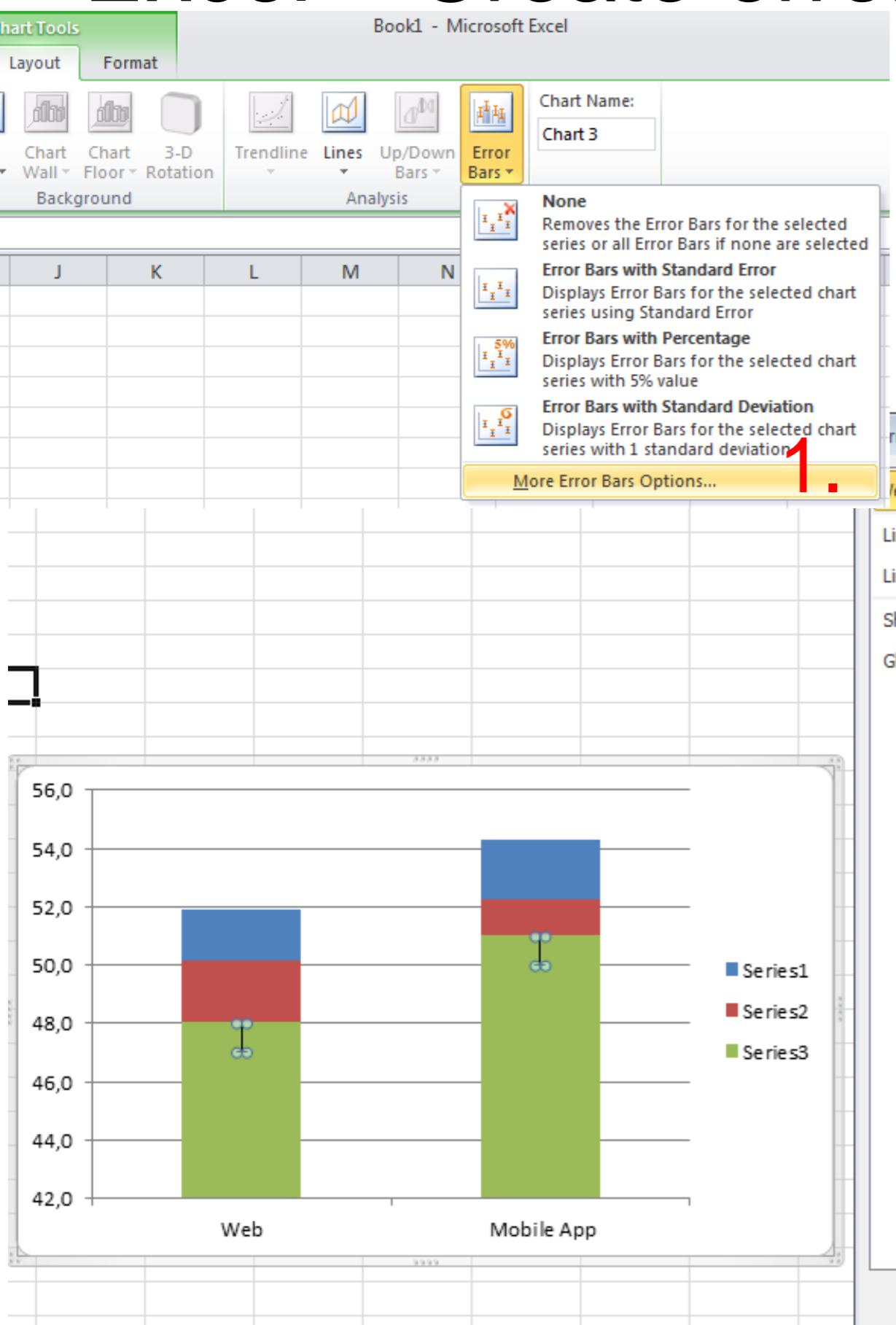
The 'Select Data Source' dialog box is open, showing the following settings:

- Chart data range: (empty)
- Legend Entries (Series): Series3, Series2, Series1
- Horizontal (Category) Axis Labels: Web, Mobile App

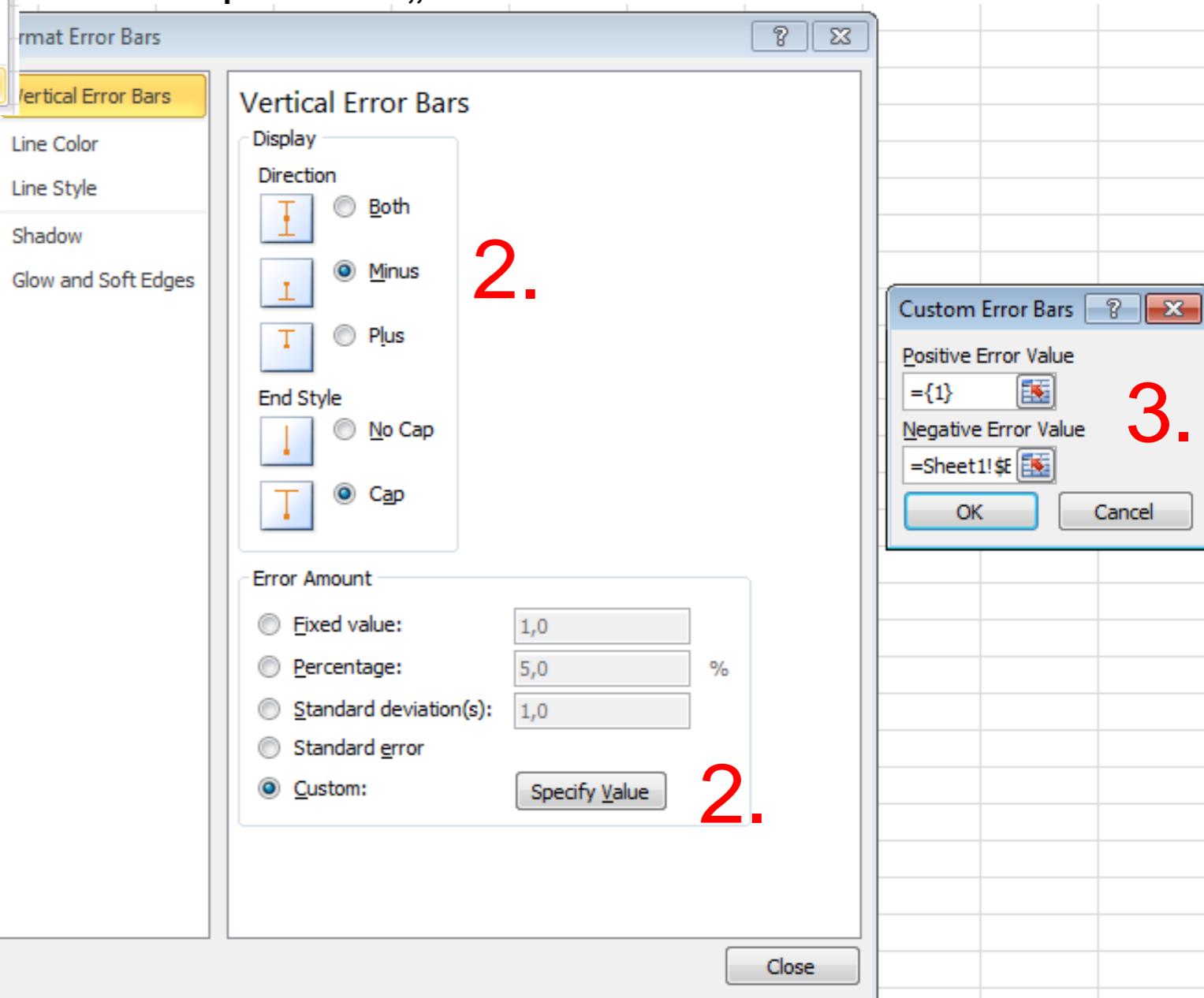
Red numbers 1, 2, 3, and 4 are overlaid on the interface to indicate the steps:

1. Switch rows and columns
2. Select data
3. Swap legend entries (1st quartile is at bottom)
4. Name labels for x-axis

Excel – Create error bars

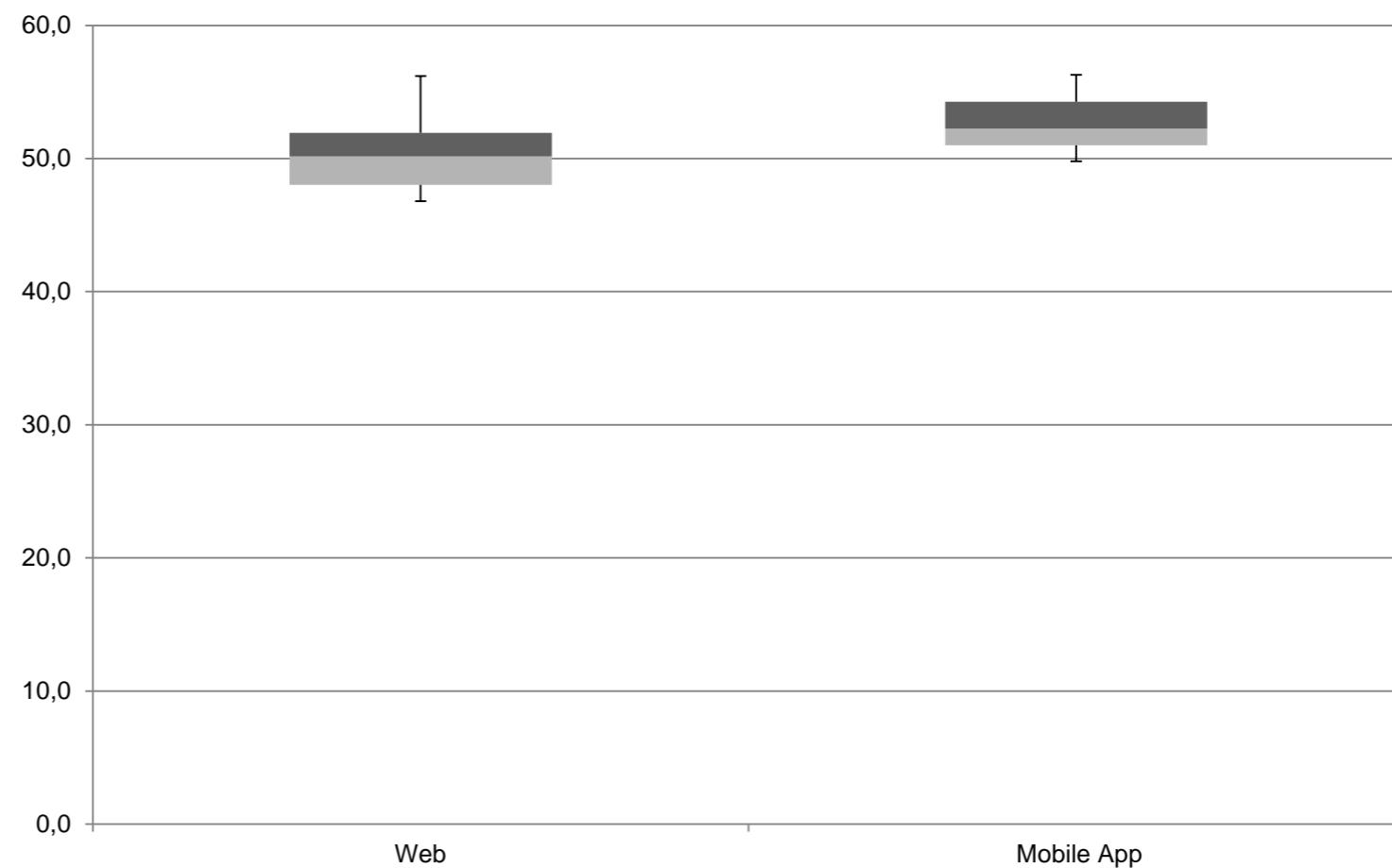


1. Select Chart -> Layout -> Insert error bar with more options
2. Select „Minus“ and „custom“ error amount
3. Set as negative value „Min“ values
4. Repeat for „Max“ error bar



Excel – Tweak visualization

1. Select boxes of first quartile and format (no fill, no border)
2. Optional: Choose different colors
3. Admire result



Excel – t-test

Excel-Funktion: TTEST oder T.Test (2010)

TTEST(array1,array2,tails,type)

Array1 is the first data set.

Array2 is the second data set.

Tails specifies the number of distribution tails.

- 1 one-tailed distribution (testing for a directed hypothesis, e.g. higher, larger, faster)
- 2 two-tailed distribution (difference between array1 and array2)

Type is the kind of t-Test to perform.

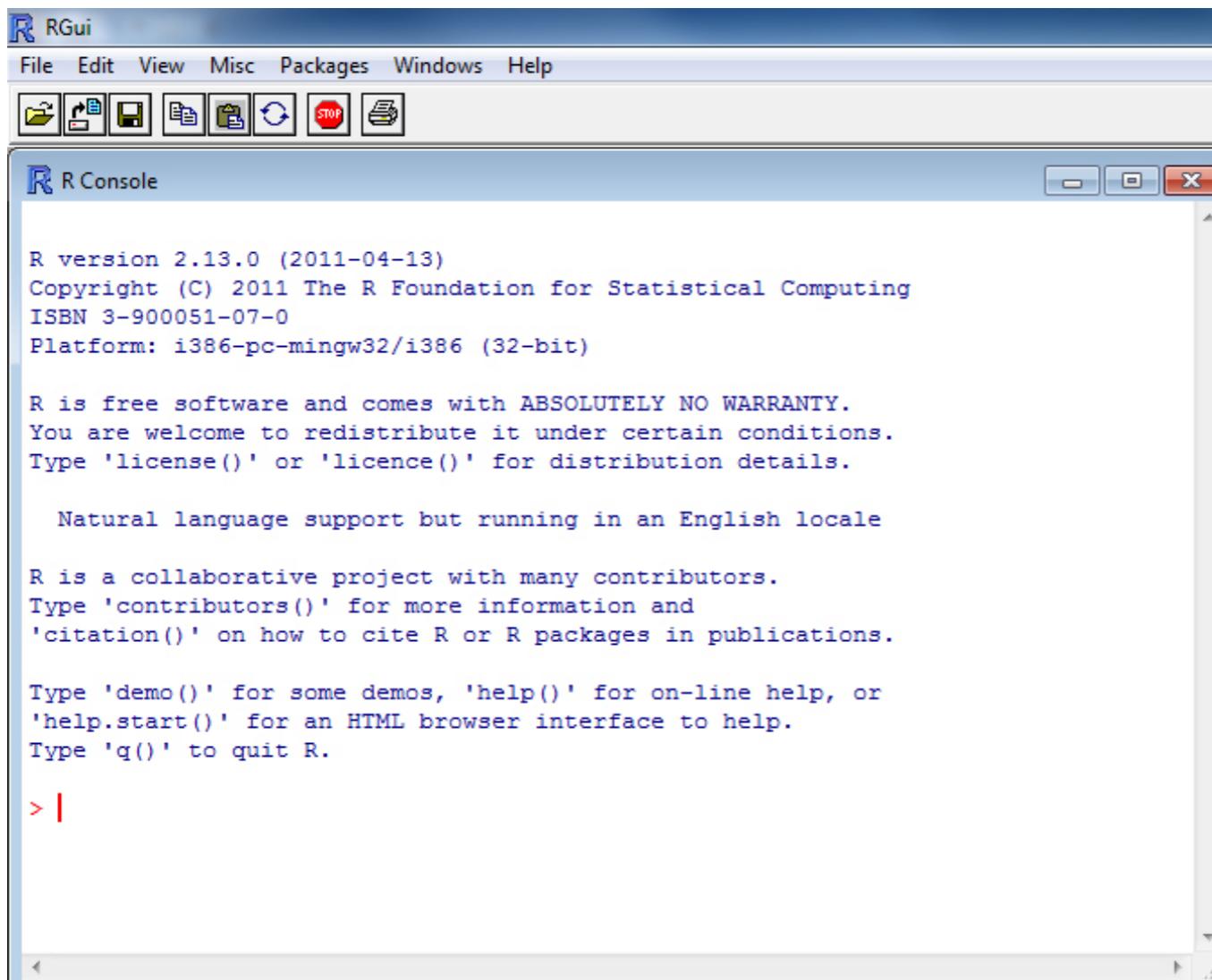
- 1 Paired (within subjects)
- 2 Two-sample equal variance (between subjects)
- 3 Two-sample unequal variance (between subjects)

e.g. =TTEST(B2:B11;C2:C11;2;1)

If result of t-test < 0.05 differences are significant (for 5% significance level)



- Mathematical/statistical computing software
- Free alternative to SPSS
- Also offers a comprehensive programming language



R - Assign values

```
> a = 1 # a stores the value 1
```

or

```
> a <- 1 # a stores the value 1
```

```
> a <- c(1,3,4,5) # c is a method stands for combine
```

```
> a[2] # access an element in the list
```

```
[1] 3
```

R – Read a CSV file

```
> a <- read.csv(file="file.csv", head=TRUE,sep=",")
```

```
# file: location of the file, head: does it have a header or not, sep =  
seperator
```

R – Read a CSV file

```
> a <- read.csv(file="file.csv", head=TRUE,sep=",")  
# file: location of the file, head: does it have a header or not, sep  
= separator
```

Web	Mobile
50,7	52,6
46,8	50,8
52,3	49,9
49,6	51,9
56,2	56,2
47,6	52,7
52,1	54,8
49,3	56,3
47,5	49,8
51,4	51,6

```
> attach(a) # make the columns of file.csv available to R  
> web  
[1] 50.7 46.8 52.3 49.6 56.2 47.6 52.1 49.3 47.5 51.4
```

R - paired sampled T-Test

```
> t.test(web,mobile,paired=T)  
# paired data t-test
```

Paired t-test

```
data: mobile and web  
t = -2.7197, df = 9, p-value = 0.02362  
alternative hypothesis: true difference in means  
is not equal to 0  
95 percent confidence interval:  
-4.2314137 -0.3885863  
sample estimates:  
mean of the differences  
-2.31
```

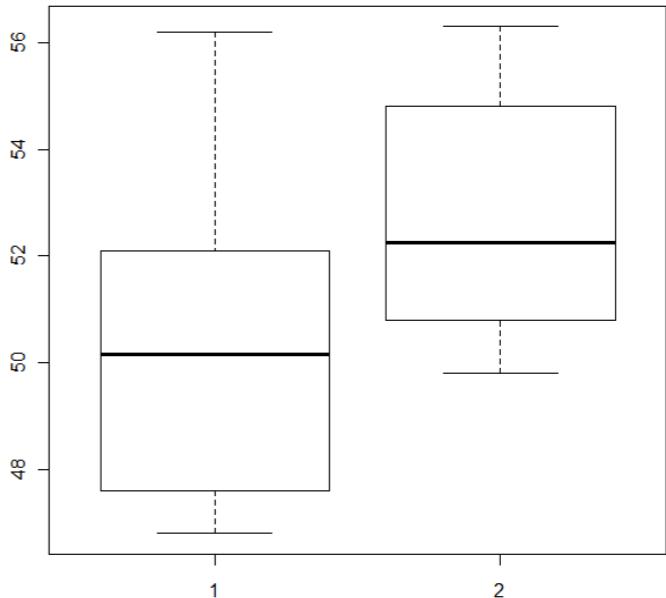
statistically significant

R – independent samples T-Test

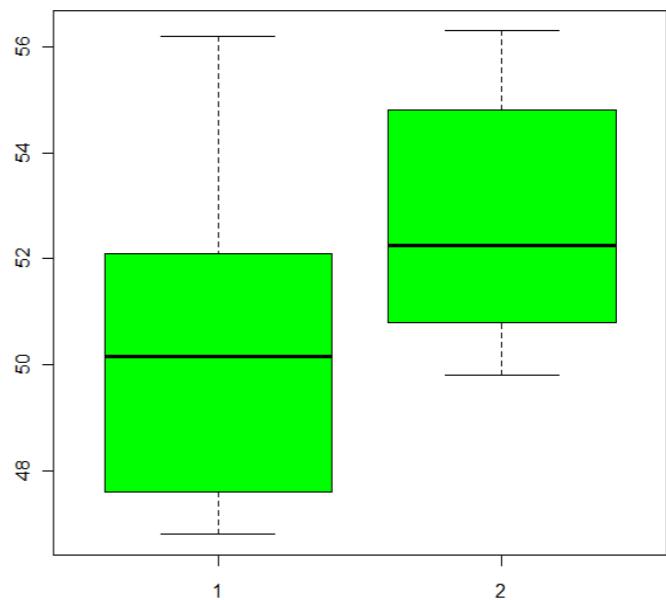
```
> t.test(web,mobile)
# variances unequal Welch procedure t-test
```

R - Boxplot

```
> boxplot(web,mobile)
```



```
> boxplot(web,mobile,col="green")
```



To be honest – this example was boring

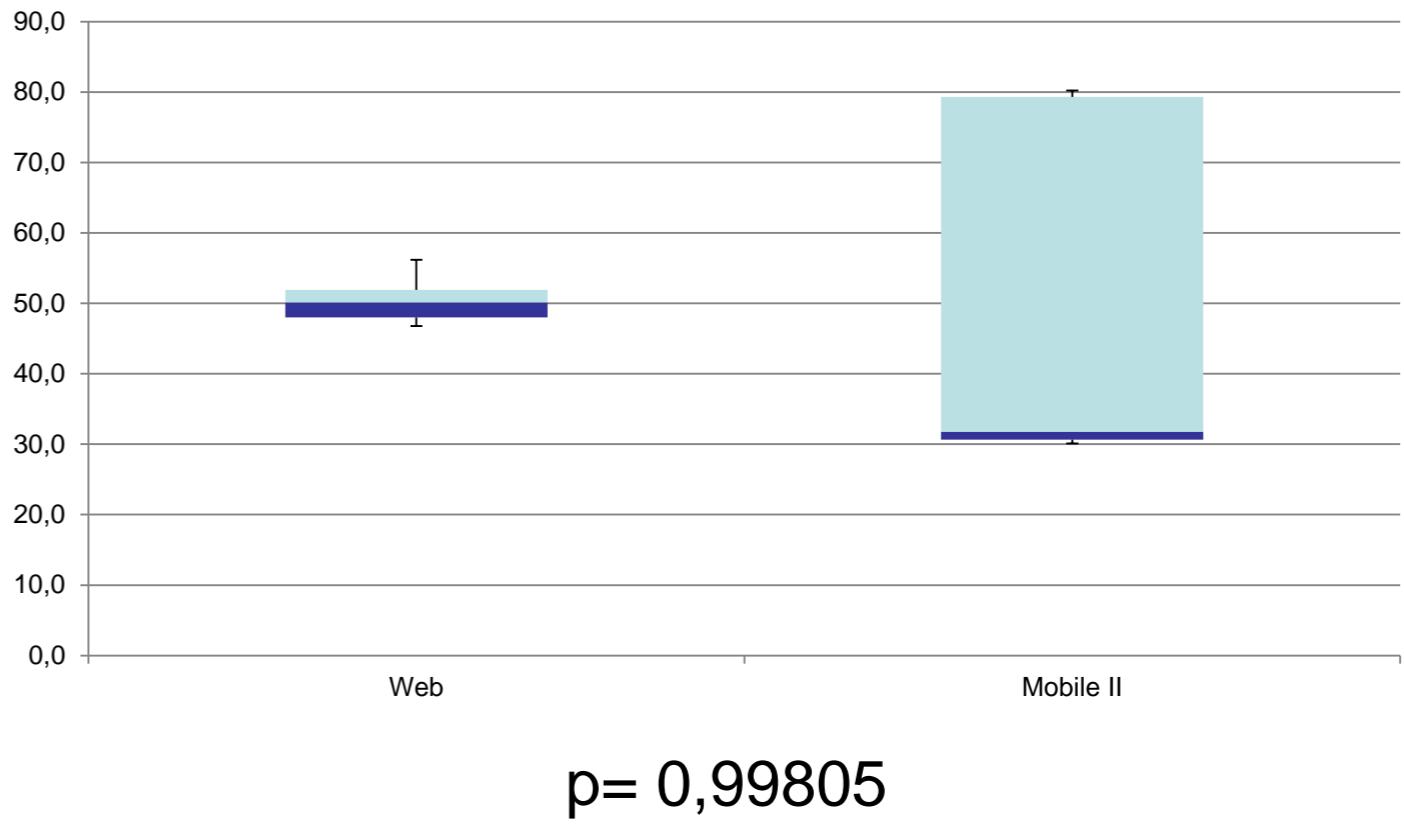
But what if results had been like this:

	Web	Mobile II
	50,7	80,2
	46,8	30,1
	52,3	30,5
	49,6	80,2
	56,2	32,4
	47,6	31,2
	52,1	78,1
	49,3	79,7
	47,5	30,2
	51,4	31,1
Average	50,4	50,4

“On average web and mobile app have the same task completion time” (!!!)

Boxplot and t-test reveal more insights

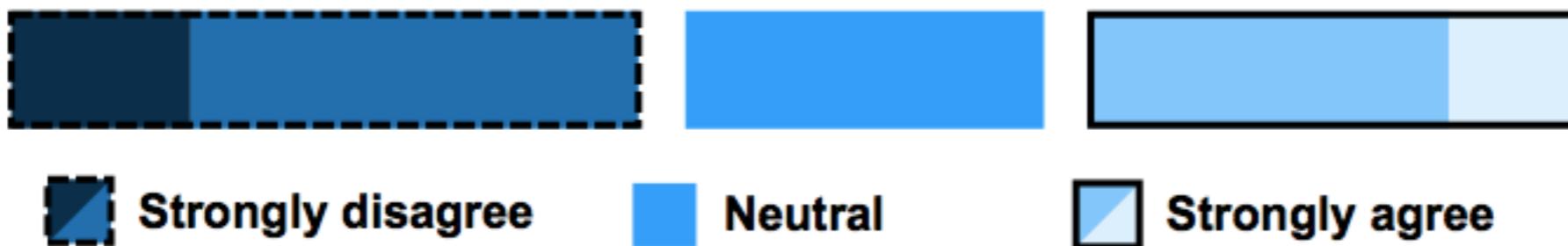
	Web	Mobile II
50,7		80,2
46,8		30,1
52,3		30,5
49,6		80,2
56,2		32,4
47,6		31,2
52,1		78,1
49,3		79,7
47,5		30,2
51,4		31,1
Average	50,4	50,4



- Median of mobile app is much lower (-18.4 seconds)
 - but: also very high task completion times were measured (max = 80,2 seconds)
- => Look into qualitative data for explanation (e.g. participants who found the “retrieve current location” function were much faster

Reporting Qualitative Evaluations

- Let participants rate statements on Likert Scales
- Again: do not report average result, but present frequencies for Likert scales



Visualization by Max Maurer. Script available here <http://www.paje-systems.de/likert/>

- Quotes from participants can also be very interesting