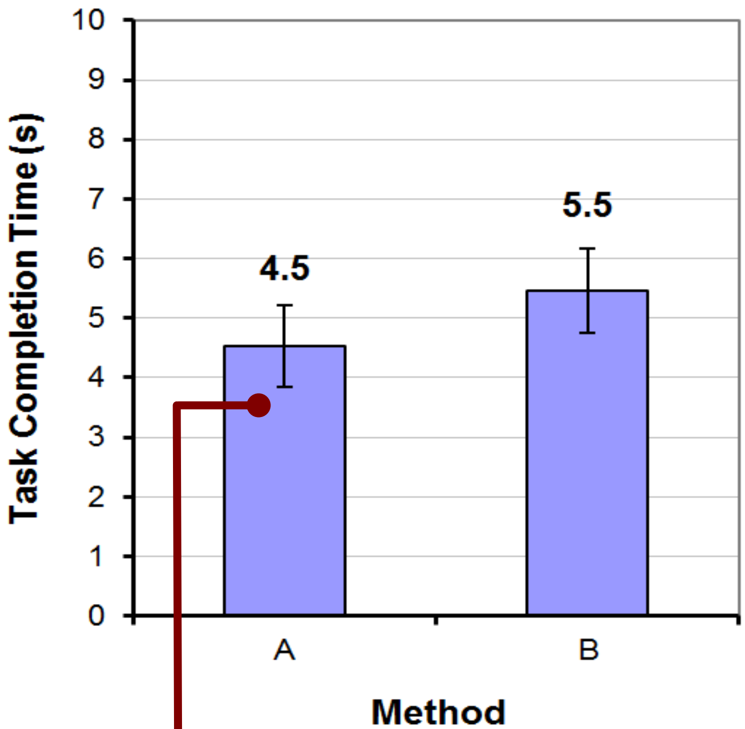
Hypothesis Testing – SPSS Practice

21.10.5
Taejun Kim

Based on Chapter 6 of
Human-Computer Interaction by I. S. MacKenzie

Example #1 - Details



Note: Within-subjects design

Participant	Method	
	A	B
1	5.3	5.7
2	3.6	4.8
3	5.2	5.1
4	3.6	4.5
5	4.6	6.0
6	4.1	6.8
7	4.0	6.0
8	4.8	4.6
9	5.2	5.5
10	5.1	5.6
Mean	4.5	5.5
SD	0.68	0.72

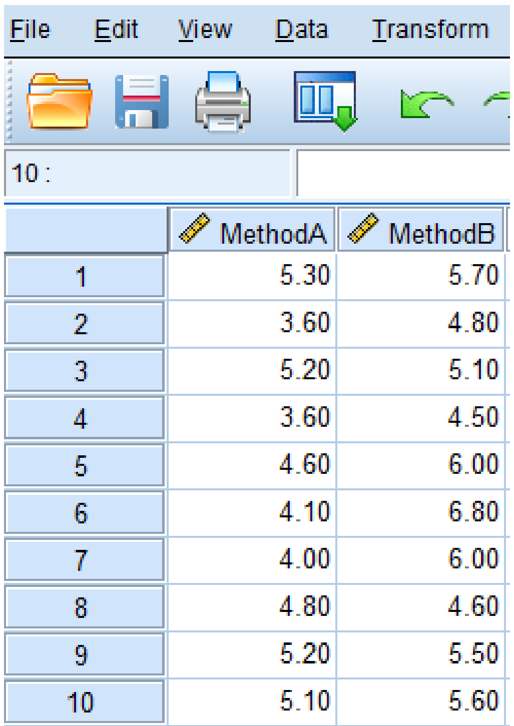
Error bars show ±1 standard deviation

Note: *SD* is the square root of the variance

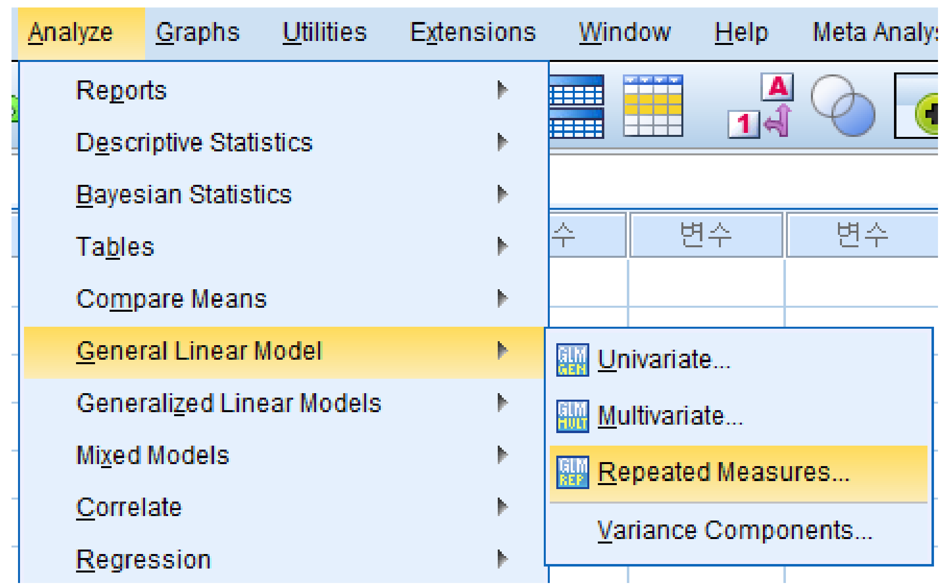
1. One-way repeated measure ANOVA

[TODO] With SPSS (Example #1 – ANOVA)

①



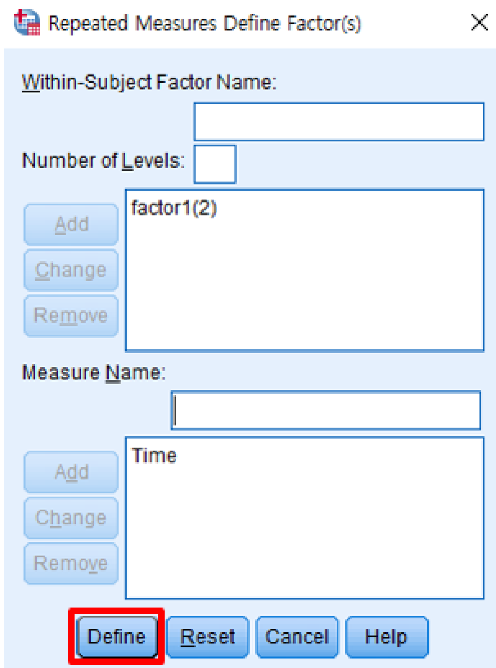
②



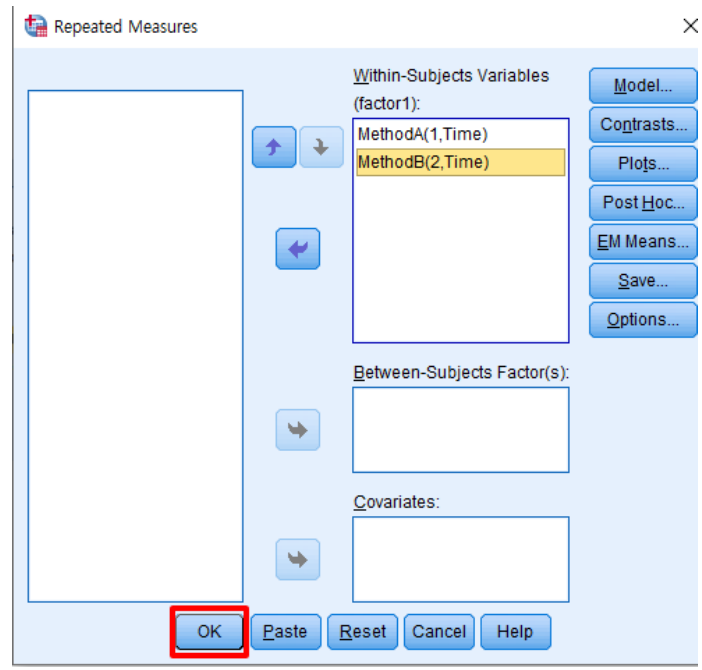
1. One-way repeated measure ANOVA

[TODO] With SPSS (Example #1 – ANOVA)

③



④



[TODO] With SPSS (Example #1 – ANOVA)

```
================================================================================
Effect              df        SS           MS          F          p
--------------------------------------------------------------------------------
Participant          9      4.884        0.543
F1                   1      4.140        4.140       9.593      0.0128
F1_x_Par             9      3.885        0.432
================================================================================
```

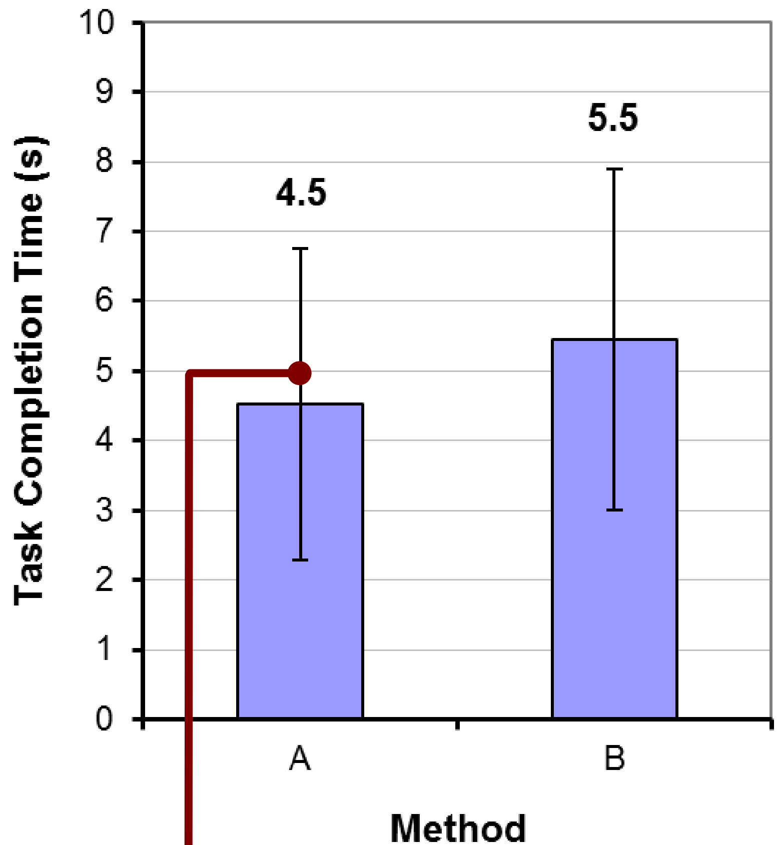
Tests of Within-Subjects Effects

Measure: Time

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	4.141	1	4.141	9.593	.013
	Greenhouse-Geisser	4.141	1.000	4.141	9.593	.013
	Huynh-Feldt	4.141	1.000	4.141	9.593	.013
	Lower-bound	4.141	1.000	4.141	9.593	.013
Error(factor1)	Sphericity Assumed	3.885	9	.432		
	Greenhouse-Geisser	3.885	9.000	.432		
	Huynh-Feldt	3.885	9.000	.432		
	Lower-bound	3.885	9.000	.432		

1. One-way repeated measure ANOVA

Example #2 - Details



Participant	Method	
	A	B
1	2.4	6.9
2	2.7	7.2
3	3.4	2.6
4	6.1	1.8
5	6.4	7.8
6	5.4	9.2
7	7.9	4.4
8	1.2	6.6
9	3.0	4.8
10	6.6	3.1
Mean	4.5	5.5
SD	2.23	2.45

Error bars show
±1 standard deviation

1. One-way repeated measure ANOVA

[TODO] With SPSS (Example #2 – ANOVA)

Tests of Within-Subjects Effects

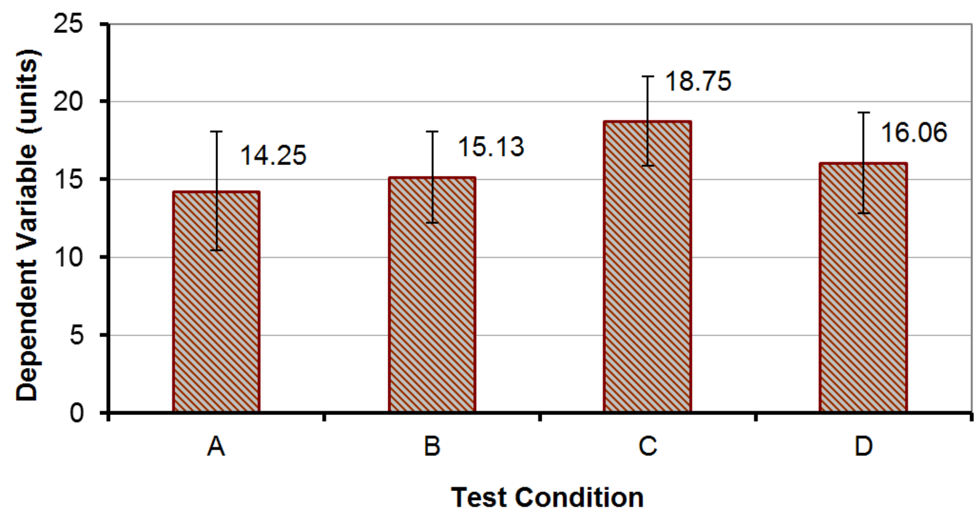
Measure: Time

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	4.324	1	4.324	.626	.449
	Greenhouse-Geisser	4.324	1.000	4.324	.626	.449
	Huynh-Feldt	4.324	1.000	4.324	.626	.449
	Lower-bound	4.324	1.000	4.324	.626	.449
Error(factor1)	Sphericity Assumed	62.140	9	6.904		
	Greenhouse-Geisser	62.140	9.000	6.904		
	Huynh-Feldt	62.140	9.000	6.904		
	Lower-bound	62.140	9.000	6.904		

2. Post hoc comparison of (One-way RM ANOVA)

More Than Two Test Conditions (Levels)

Participant	Test Condition			
	A	B	C	D
1	11	11	21	16
2	18	11	22	15
3	17	10	18	13
4	19	15	21	20
5	13	17	23	10
6	10	15	15	20
7	14	14	15	13
8	13	14	19	18
9	19	18	16	12
10	10	17	21	18
11	10	19	22	13
12	16	14	18	20
13	10	20	17	19
14	10	13	21	18
15	20	17	14	18
16	18	17	17	14
Mean	14.25	15.13	18.75	16.06
SD	3.84	2.94	2.89	3.23

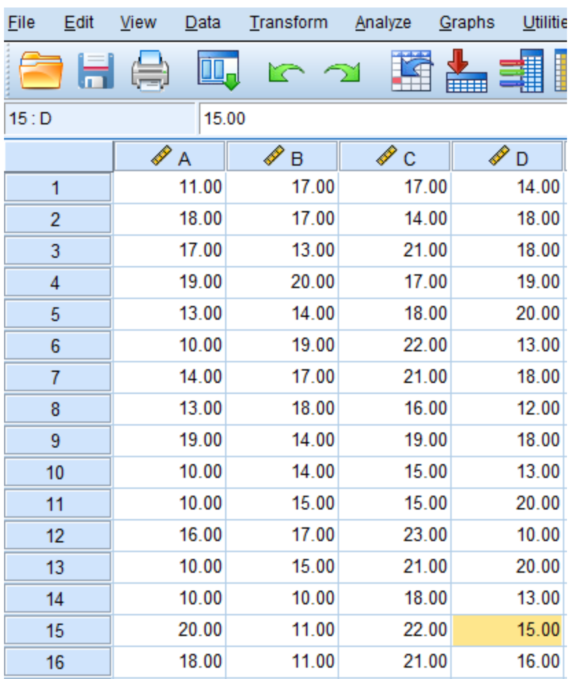


Provided in posthoc-ex1.txt

2. Post hoc comparison of (One-way RM ANOVA)

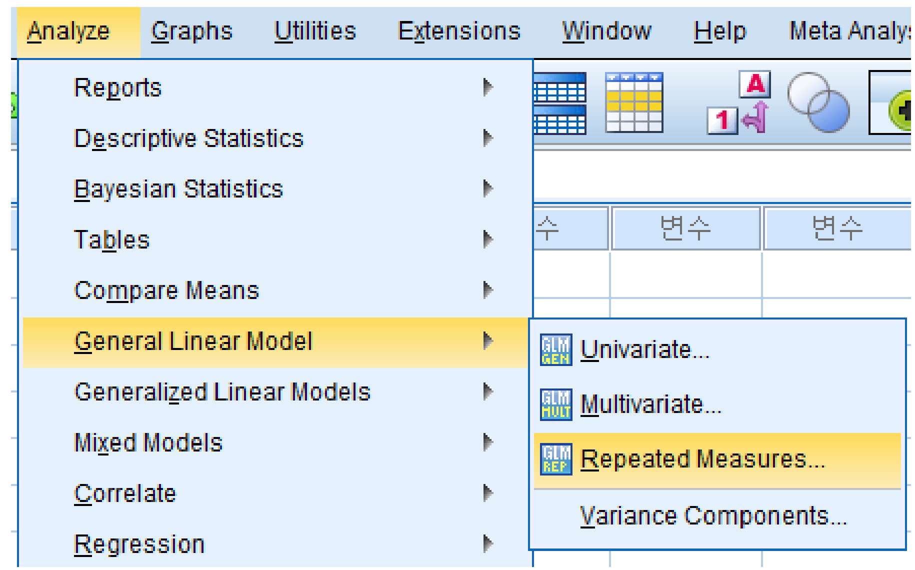
[TODO] With SPSS

①



	A	B	C	D
1	11.00	17.00	17.00	14.00
2	18.00	17.00	14.00	18.00
3	17.00	13.00	21.00	18.00
4	19.00	20.00	17.00	19.00
5	13.00	14.00	18.00	20.00
6	10.00	19.00	22.00	13.00
7	14.00	17.00	21.00	18.00
8	13.00	18.00	16.00	12.00
9	19.00	14.00	19.00	18.00
10	10.00	14.00	15.00	13.00
11	10.00	15.00	15.00	20.00
12	16.00	17.00	23.00	10.00
13	10.00	15.00	21.00	20.00
14	10.00	10.00	18.00	13.00
15	20.00	11.00	22.00	15.00
16	18.00	11.00	21.00	16.00

②

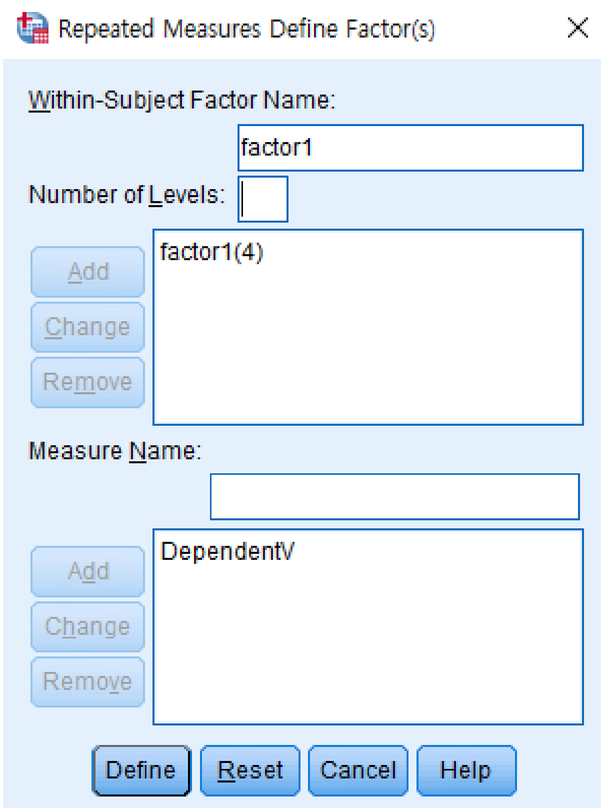


* In SPSS, we'll use paired sample t-test for post hoc comparison

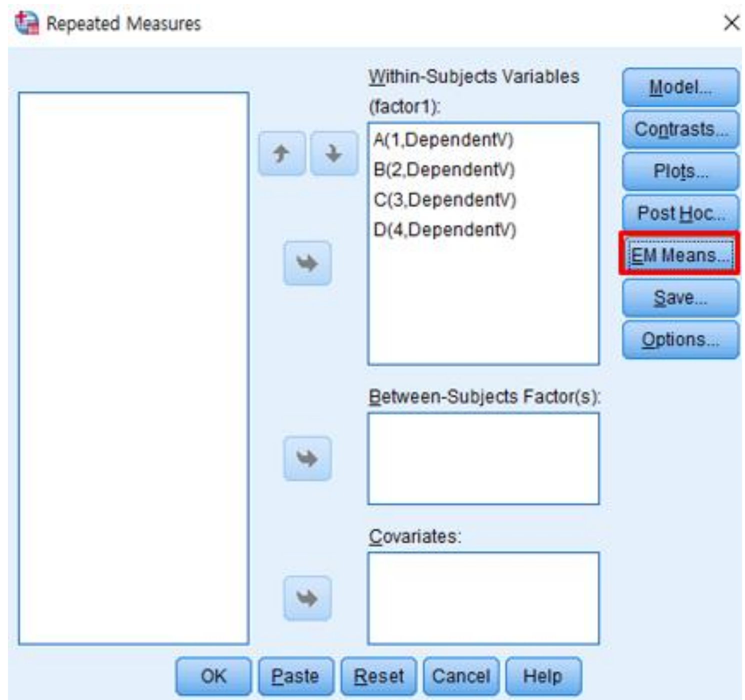
2. Post hoc comparison of (One-way RM ANOVA)

[TODO] With SPSS

③



④



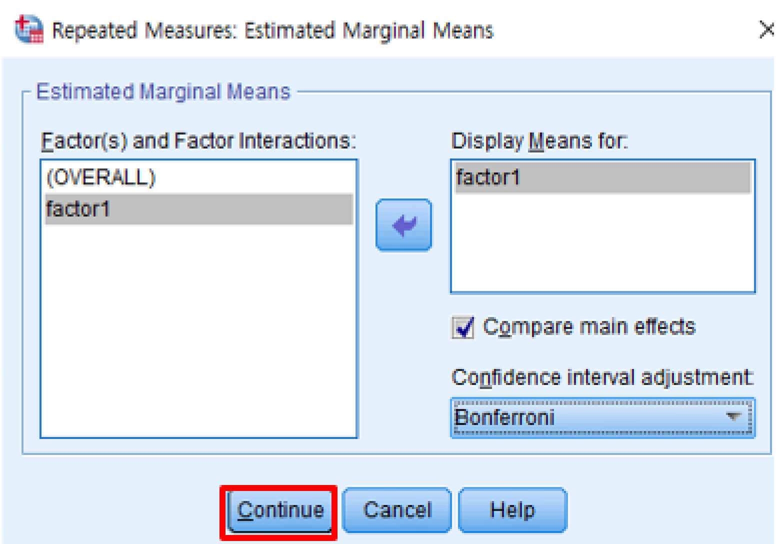
* In SPSS, we'll use paired sample t-test for post hoc comparison

Human-Computer Interaction Lab @kaist

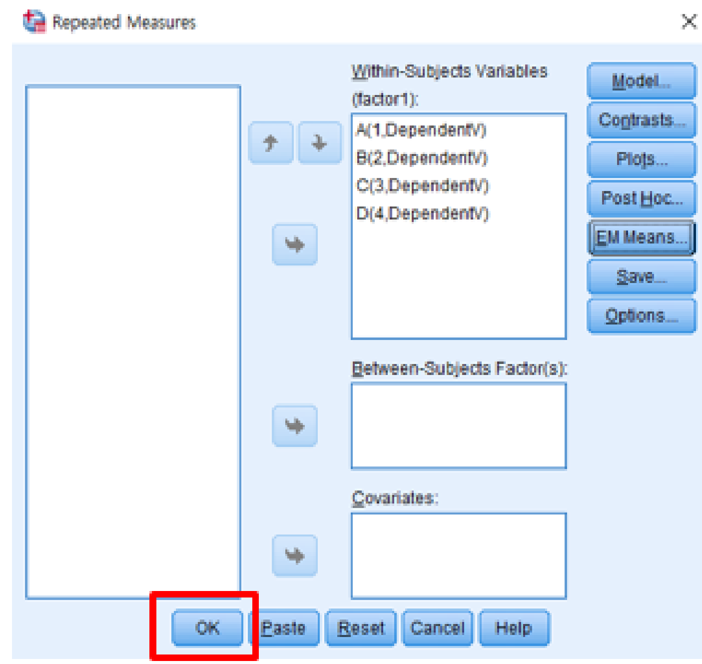
2. Post hoc comparison of (One-way RM ANOVA)

[TODO] With SPSS

⑤



⑥



* In SPSS, we'll use paired sample t-test for post hoc comparison

2. Post hoc comparison of (One-way RM ANOVA)

[TODO] With SPSS

```
------------------------------------------------------------
---------- Pairwise Comparisons (Scheffe) ------------
------------------------------------------------------------
Pair 1:2 -->     0.88   >    3.30   ?     -
Pair 1:3 -->     4.50   >    3.30   ?     * (significant)
Pair 1:4 -->     1.81   >    3.30   ?     -
Pair 2:3 -->     3.63   >    3.30   ?     * (significant)
Pair 2:4 -->     0.94   >    3.30   ?     -
Pair 3:4 -->     2.69   >    3.30   ?     -
------------------------------------------------------------
```

Pairwise Comparisons

Measure: DependentV

(I) factor1	(J) factor1	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	-.875	1.258	1.000	-4.694	2.944
	3	-4.500*	1.065	.004	-7.732	-1.268
	4	-1.812	1.148	.812	-5.299	1.674
2	1	.875	1.258	1.000	-2.944	4.694
	3	-3.625*	1.118	.033	-7.018	-.232
	4	-.937	1.112	1.000	-4.315	2.440
3	1	4.500*	1.065	.004	1.268	7.732
	2	3.625*	1.118	.033	.232	7.018
	4	2.688	1.175	.223	-.881	6.256
4	1	1.813	1.148	.812	-1.674	5.299
	2	.938	1.112	1.000	-2.440	4.315
	3	-2.687	1.175	.223	-6.256	.881

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

* In SPSS, we'll use paired sample t-test for post hoc comparison

Between-subjects Design

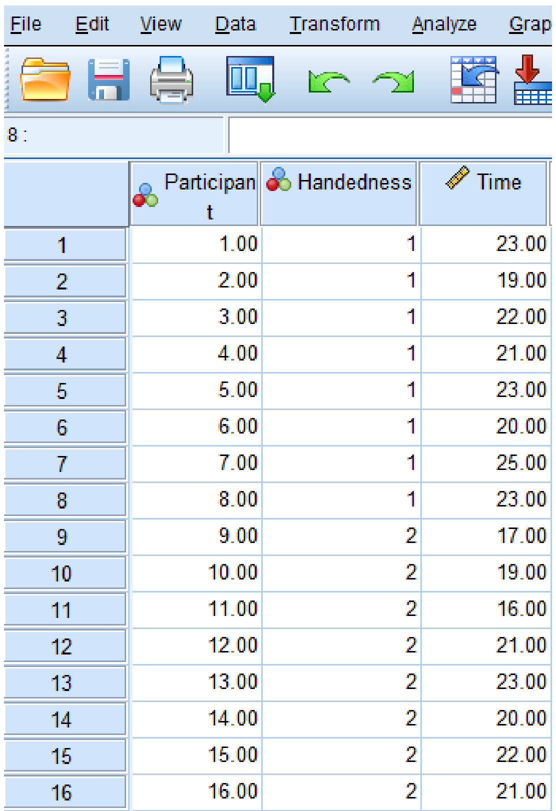
- Research question:
  - Do left-handed users and right-handed users differ in the time to complete an interaction task?
- The independent variable (handedness) must be assigned between-subjects
- Example data set →

Participant	Task Completion Time (s)	Handedness
1	23	L
2	19	L
3	22	L
4	21	L
5	23	L
6	20	L
7	25	L
8	23	L
9	17	R
10	19	R
11	16	R
12	21	R
13	23	R
14	20	R
15	22	R
16	21	R
Mean	20.9	
SD	2.38	

3. One-way ANOVA (Between-subjects)

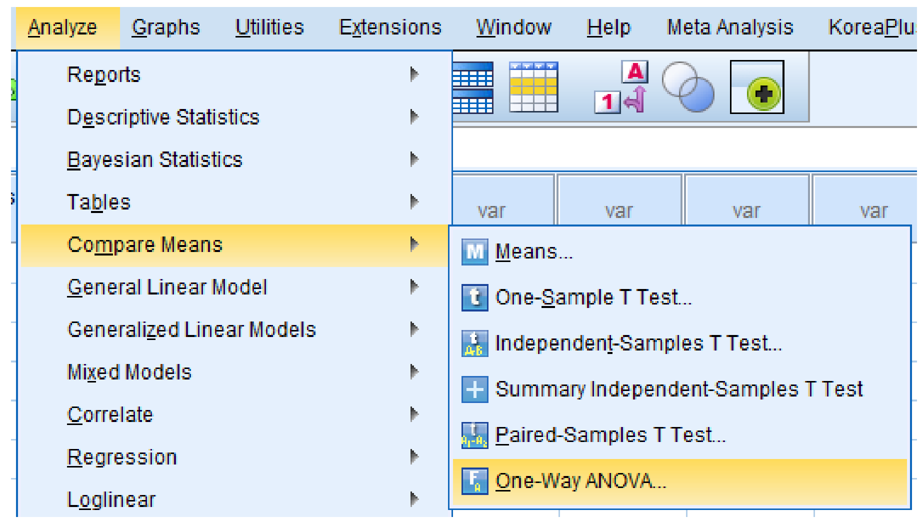
[TODO] With SPSS

①



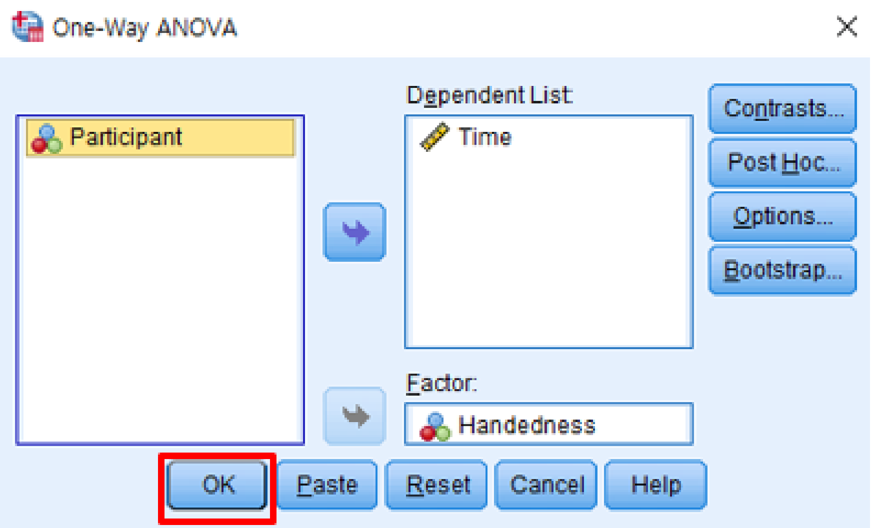
	Participant	Handedness	Time
1	1.00	1	23.00
2	2.00	1	19.00
3	3.00	1	22.00
4	4.00	1	21.00
5	5.00	1	23.00
6	6.00	1	20.00
7	7.00	1	25.00
8	8.00	1	23.00
9	9.00	2	17.00
10	10.00	2	19.00
11	11.00	2	16.00
12	12.00	2	21.00
13	13.00	2	23.00
14	14.00	2	20.00
15	15.00	2	22.00
16	16.00	2	21.00

②



[TODO] With SPSS

③



3. One-way ANOVA (Between-subjects)

[TODO] With SPSS

```
=========================================================================
Effect              df        SS          MS         F        p
-------------------------------------------------------------------------
F3                   1      18.063      18.063     3.781    0.0722
Residual            14      66.875       4.777
=========================================================================
```

Oneway

ANOVA

Time

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	18.063	1	18.063	3.781	.072
Within Groups	66.875	14	4.777		
Total	84.938	15			

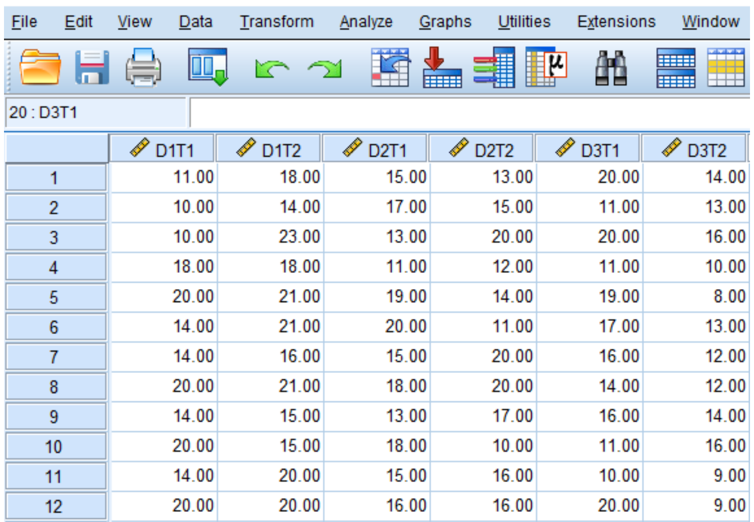
Data Set

Participant	Device 1		Device 2		Device 3	
	Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
1	11	18	15	13	20	14
2	10	14	17	15	11	13
3	10	23	13	20	20	16
4	18	18	11	12	11	10
5	20	21	19	14	19	8
6	14	21	20	11	17	13
7	14	16	15	20	16	12
8	20	21	18	20	14	12
9	14	15	13	17	16	14
10	20	15	18	10	11	16
11	14	20	15	16	10	9
12	20	20	16	16	20	9
Mean	15.4	18.5	15.8	15.3	15.4	12.2
SD	4.01	2.94	2.69	3.50	3.92	2.69

4. Two-way repeated measures ANOVA

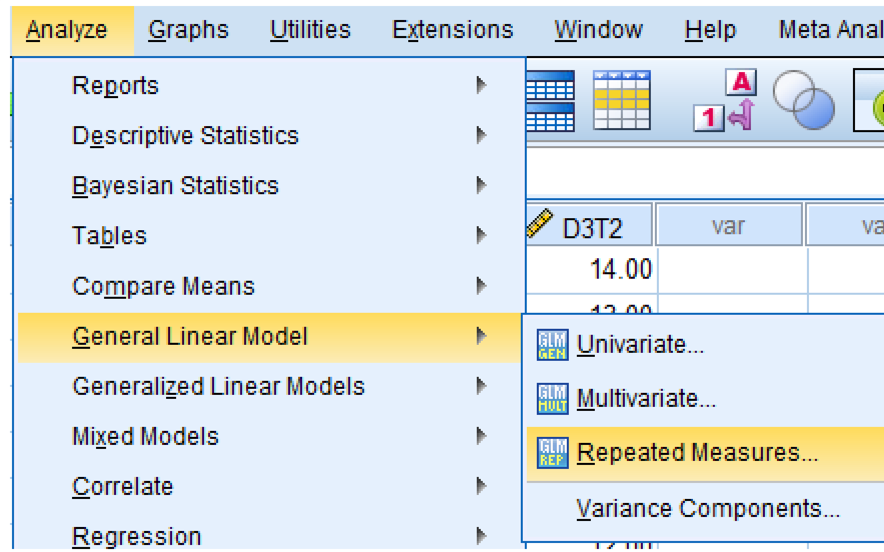
[TODO] With SPSS

①



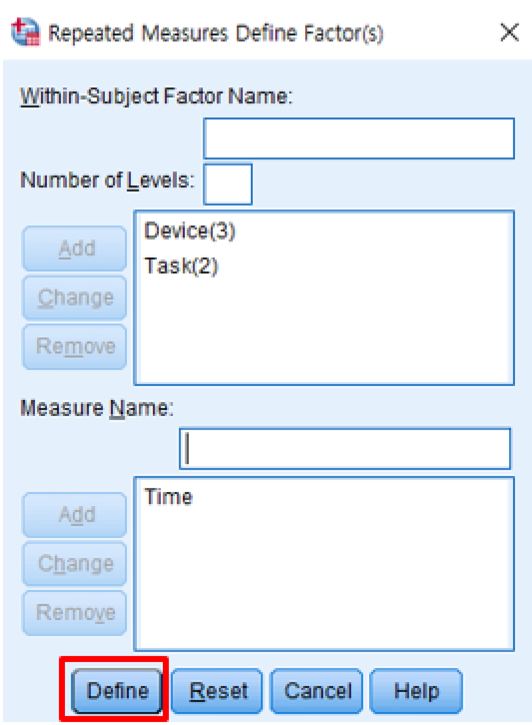
	D1T1	D1T2	D2T1	D2T2	D3T1	D3T2
1	11.00	18.00	15.00	13.00	20.00	14.00
2	10.00	14.00	17.00	15.00	11.00	13.00
3	10.00	23.00	13.00	20.00	20.00	16.00
4	18.00	18.00	11.00	12.00	11.00	10.00
5	20.00	21.00	19.00	14.00	19.00	8.00
6	14.00	21.00	20.00	11.00	17.00	13.00
7	14.00	16.00	15.00	20.00	16.00	12.00
8	20.00	21.00	18.00	20.00	14.00	12.00
9	14.00	15.00	13.00	17.00	16.00	14.00
10	20.00	15.00	18.00	10.00	11.00	16.00
11	14.00	20.00	15.00	16.00	10.00	9.00
12	20.00	20.00	16.00	16.00	20.00	9.00

②

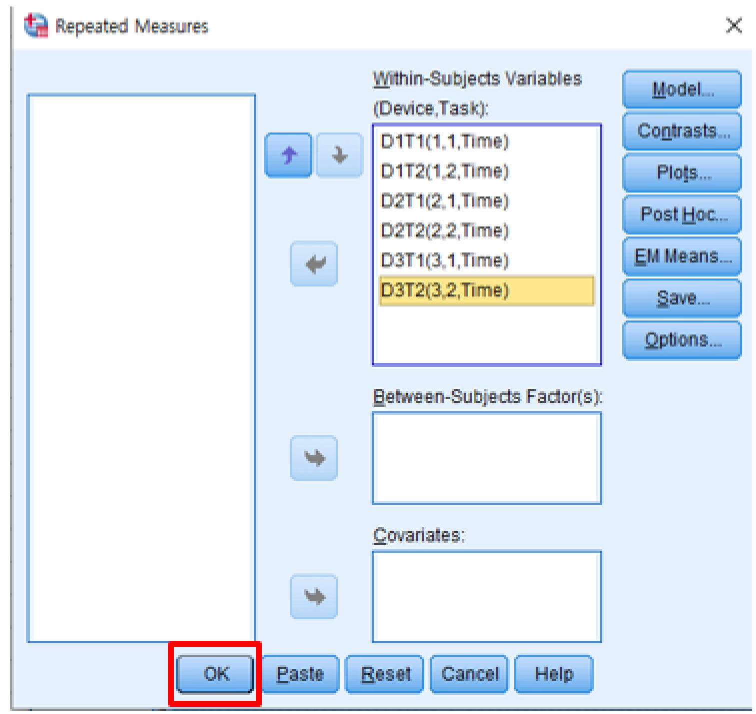


[TODO] With SPSS

③



④



4. Two-way repeated measures ANOVA

[TODO] With SPSS

Effect	df	SS	MS	F	p
Participant	11	134.778	12.253		
F1	2	121.028	60.514	5.865	0.0091
F1_x_Par	22	226.972	10.317		
F2	1	0.889	0.889	0.076	0.7875
F2_x_Par	11	128.111	11.646		
F1_x_F2	2	121.028	60.514	5.435	0.0121
F1_x_F2_x_Par	22	244.972	11.135		

⑤

Tests of Within-Subjects Effects

Measure: Time

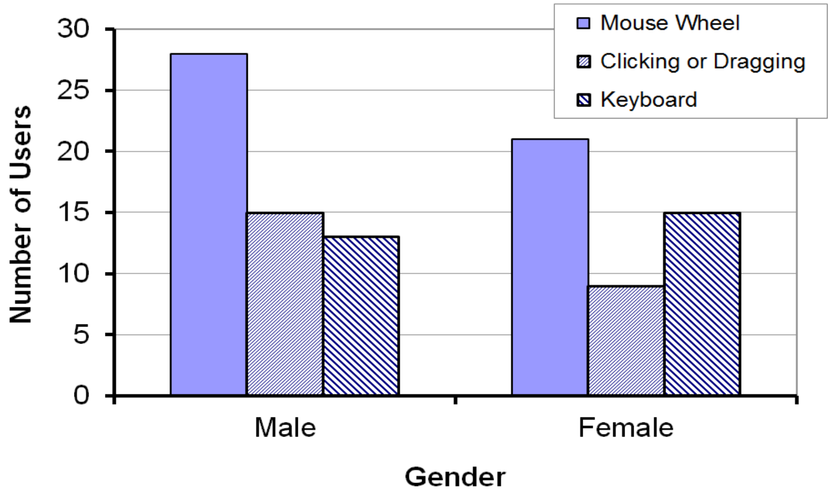
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Device	Sphericity Assumed	121.028	2	60.514	5.865	.009
	Greenhouse-Geisser	121.028	1.713	70.671	5.865	.013
	Huynh-Feldt	121.028	1.997	60.593	5.865	.009
	Lower-bound	121.028	1.000	121.028	5.865	.034
Error(Device)	Sphericity Assumed	226.972	22	10.317		
	Greenhouse-Geisser	226.972	18.838	12.049		
	Huynh-Feldt	226.972	21.971	10.330		
	Lower-bound	226.972	11.000	20.634		
Task	Sphericity Assumed	.889	1	.889	.076	.787
	Greenhouse-Geisser	.889	1.000	.889	.076	.787
	Huynh-Feldt	.889	1.000	.889	.076	.787
	Lower-bound	.889	1.000	.889	.076	.787
Error(Task)	Sphericity Assumed	128.111	11	11.646		
	Greenhouse-Geisser	128.111	11.000	11.646		
	Huynh-Feldt	128.111	11.000	11.646		
	Lower-bound	128.111	11.000	11.646		
Device * Task	Sphericity Assumed	121.028	2	60.514	5.435	.012
	Greenhouse-Geisser	121.028	1.838	65.839	5.435	.015
	Huynh-Feldt	121.028	2.000	60.514	5.435	.012
	Lower-bound	121.028	1.000	121.028	5.435	.040
Error(Device*Task)	Sphericity Assumed	244.972	22	11.135		
	Greenhouse-Geisser	244.972	20.221	12.115		
	Huynh-Feldt	244.972	22.000	11.135		
	Lower-bound	244.972	11.000	22.270		

Chi-square – Example #1

- Research question:
  - Do males and females differ in their method of scrolling on desktop systems?

Observed Number of Users				
Gender	Scrolling Method			Total
	MW	CD	KB	
Male	28	15	13	56
Female	21	9	15	45
Total	49	24	28	101

MW = mouse wheel
CD = clicking, dragging
KB = keyboard



Chi-square – Example #2

- Research question:
  - Do students, professors, and parents differ in their responses to the question: Students should be allowed to use mobile phones during classroom lectures?

- Data:

Observed Number of People				
Opinion	Category			Total
	Student	Professor	Parent	
Agree	10	12	98	120
Disagree	30	48	102	180
Total	40	60	200	300

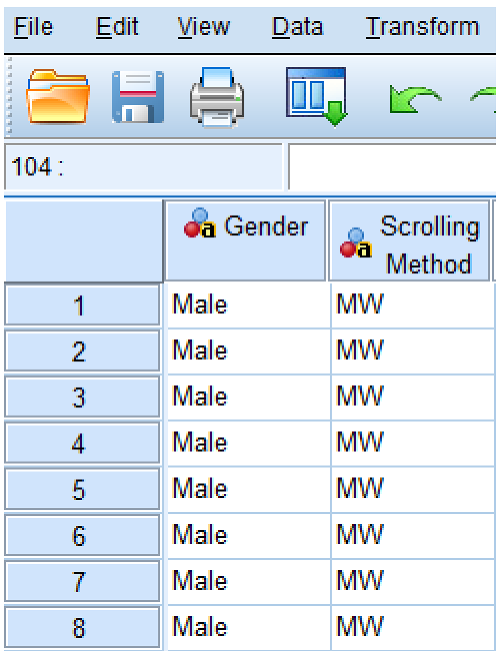
- Result: significant difference in responses ($\chi^2 = 20.5$, $p < .0001$)

- Q: Then, which of the three pairs is different?

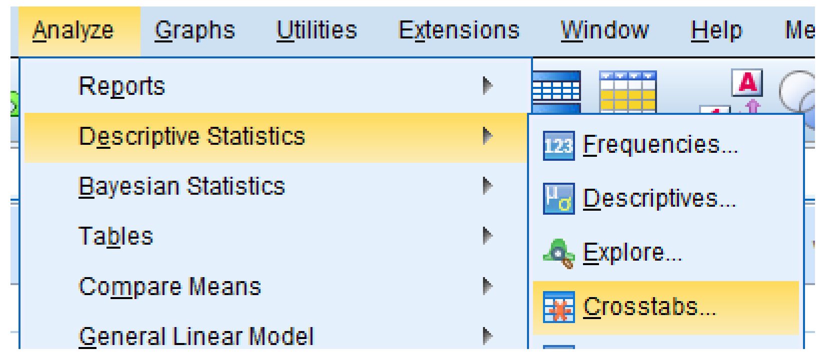
5. Chi-square test

[TODO] With SPSS – Example #1

①



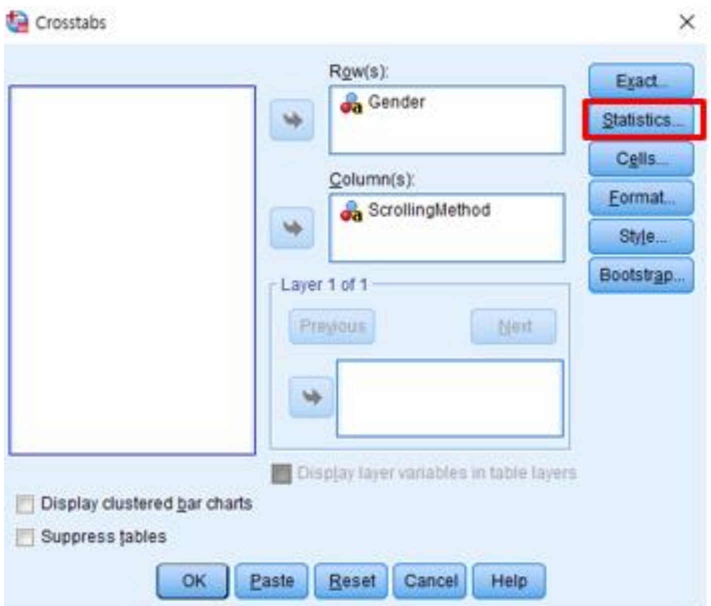
②



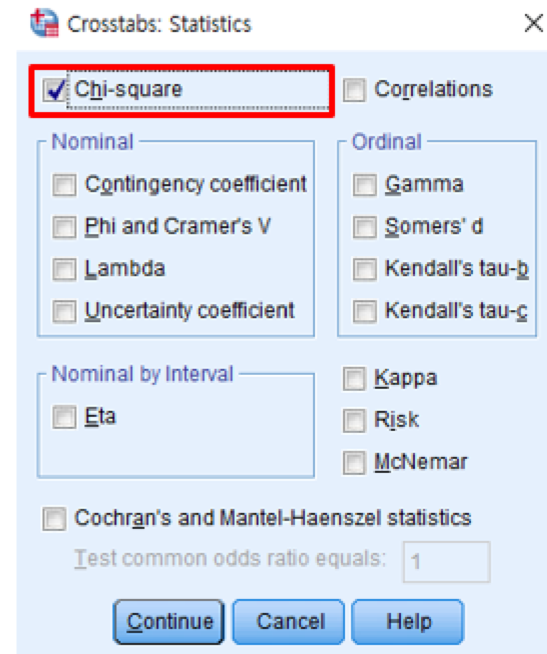
5. Chi-square test

[TODO] With SPSS – Example #1

③

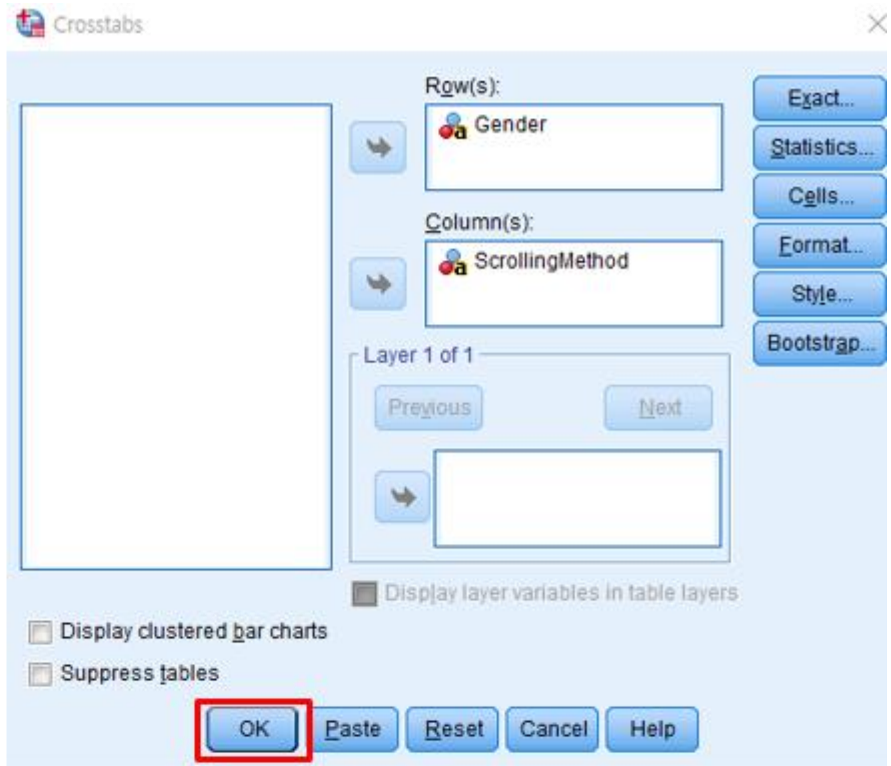


④



[TODO] With SPSS – Example #1

⑤



[TODO] With SPSS – Example #1

```
CMD
text>type chisquare-ex1.txt
28   15   13
21    9   15

text>java ChiSquare chisquare-ex1.txt
Chi-square(2) = 1.462
p = 0.4814

text>
```

Chi-Square Tests

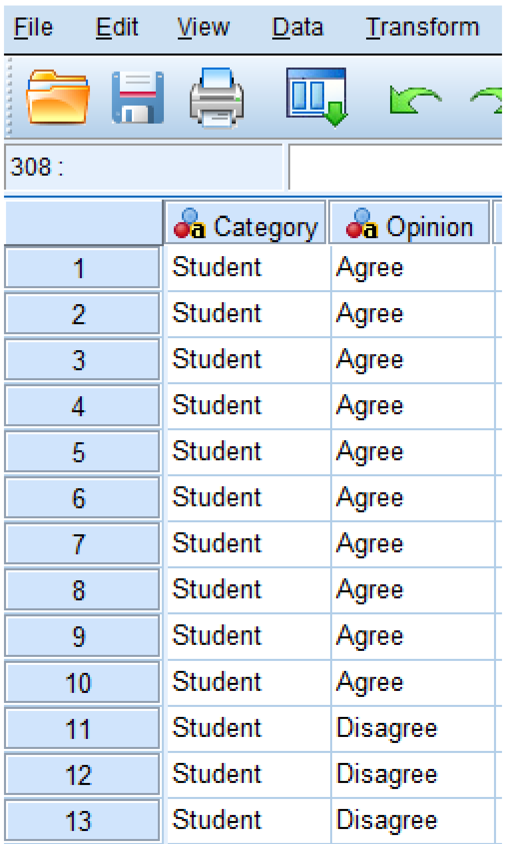
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.462[a]	2	.481
Likelihood Ratio	1.462	2	.481
N of Valid Cases	101		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.69.

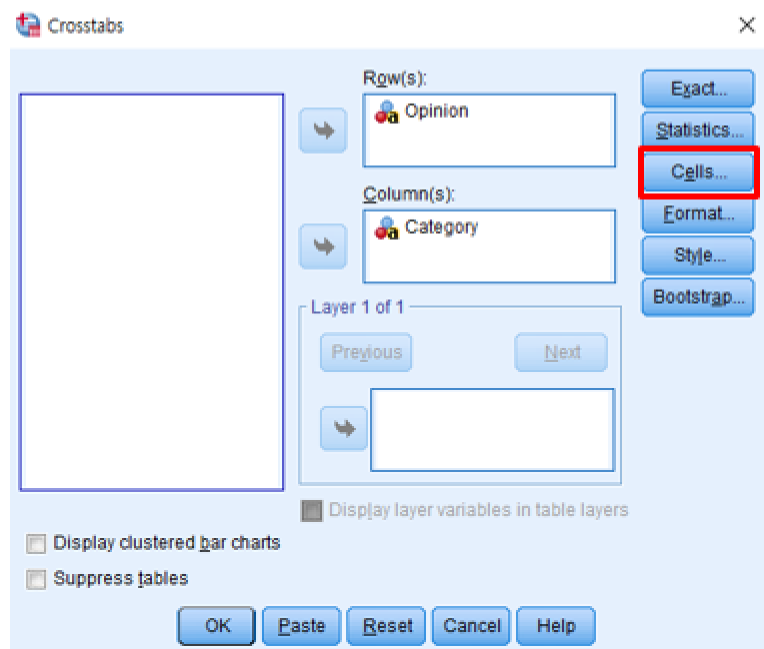
5. Chi-square test

[TODO] With SPSS – Example #2

①

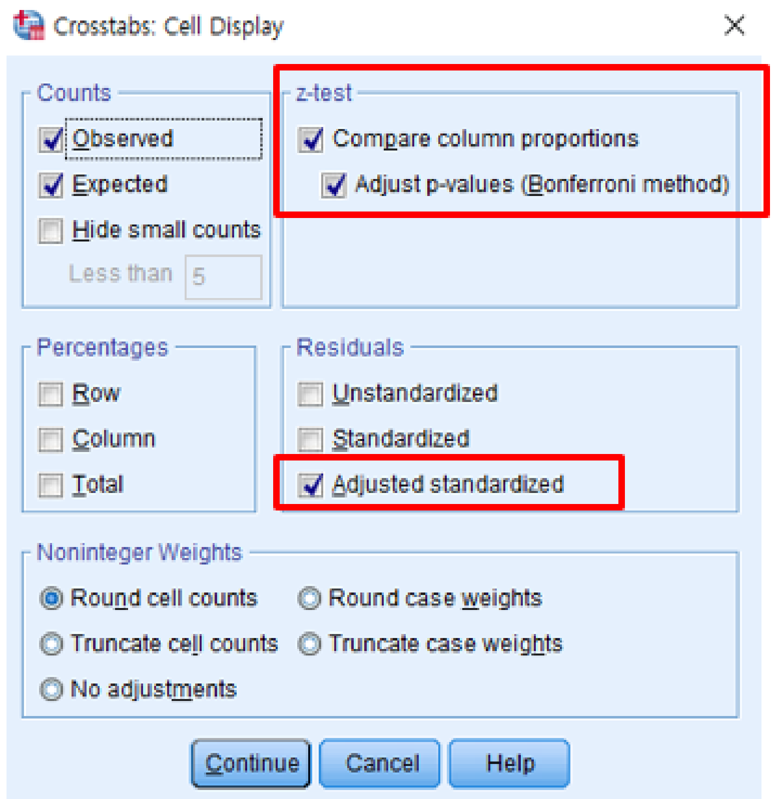


②



[TODO] With SPSS – Example #2

③



[TODO] With SPSS – Example #2

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	20.500[a]	2	.000
Likelihood Ratio	21.593	2	.000
N of Valid Cases	300		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 16.00.

Opinion * Category Crosstabulation

			Category			
			Parent	Prof	Student	Total
Opinion	Agree	Count	98_a	12_b	10_b	120
		Adjusted Residual	4.5	-3.5	-2.1	
	Disagree	Count	102_a	48_b	30_b	180
		Adjusted Residual	-4.5	3.5	2.1	
Total		Count	200	60	40	300

Each subscript letter denotes a subset of Category categories whose column proportions do not differ significantly from each other at the .05 level.

"within each row,
percentages that don't share a subscript are significantly different."
(Parent vs. Student: $p < .05$), (Prof vs. Student: $p < .05$)

Reference: https://www.spss-tutorials.com/spss-chi-square-test-with-pairwise-z-tests/

Data (Example #1)

- Means:
  - 3.7 (Mac users)
  - 4.5 (PC users)
- Data suggest PC users more right-leaning, but is the difference statistically significant?
- Data are ordinal (at least), ∴ a non-parametric test is used
- Which test?

Design	Conditions: 2	Conditions: 3 or more
Between-subjects (independent samples)	Mann-Whitney U	Kruskal-Wallis
Within-subjects (correlated samples)	Wilcoxon Signed-Rank	Friedman

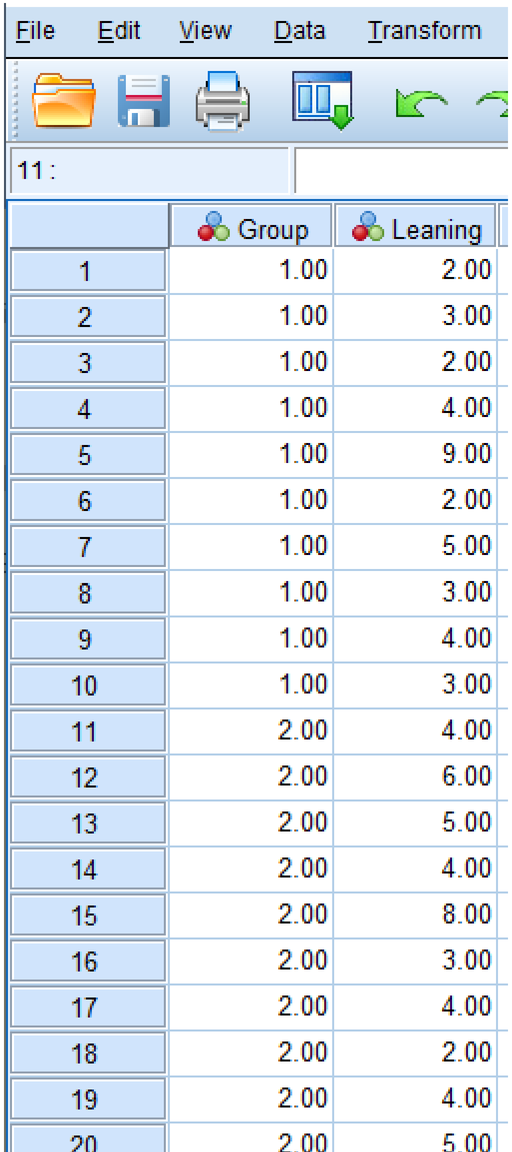
Mac Users	PC Users
2	4
3	6
2	5
4	4
9	8
2	3
5	4
3	2
4	4
3	5
3.7	**4.5**

Means for interval data?

6. Mann Whitney U test

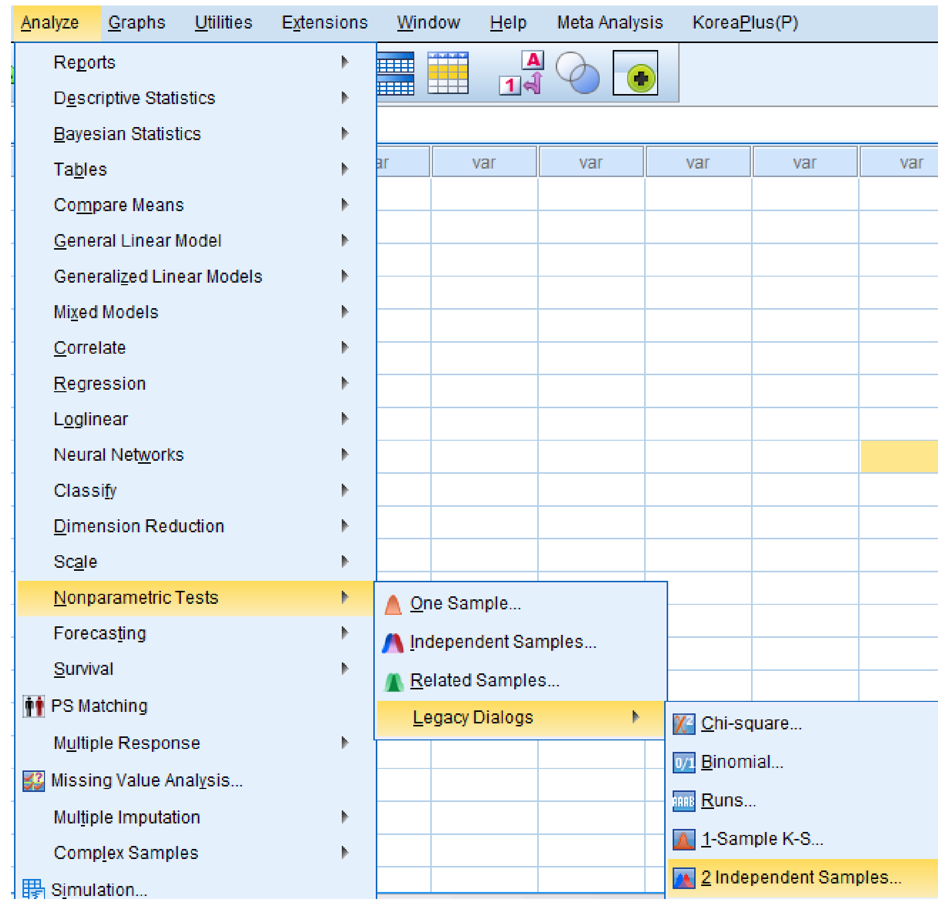
[TODO] With SPSS

①



	Group	Leaning
1	1.00	2.00
2	1.00	3.00
3	1.00	2.00
4	1.00	4.00
5	1.00	9.00
6	1.00	2.00
7	1.00	5.00
8	1.00	3.00
9	1.00	4.00
10	1.00	3.00
11	2.00	4.00
12	2.00	6.00
13	2.00	5.00
14	2.00	4.00
15	2.00	8.00
16	2.00	3.00
17	2.00	4.00
18	2.00	2.00
19	2.00	4.00
20	2.00	5.00

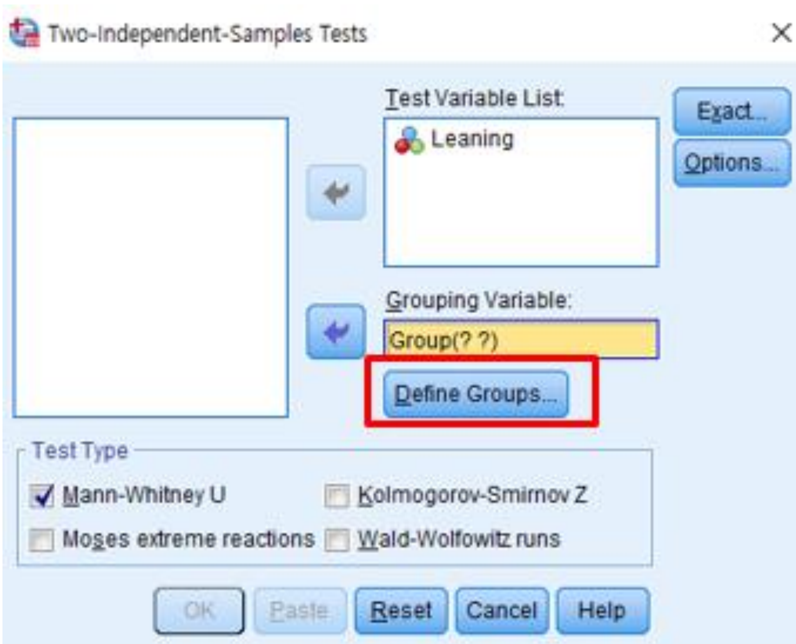
②



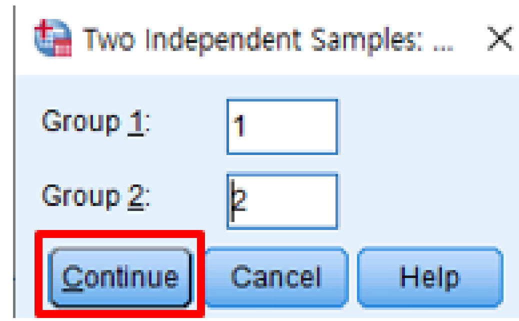
6. Mann Whitney U test

[TODO] With SPSS

③



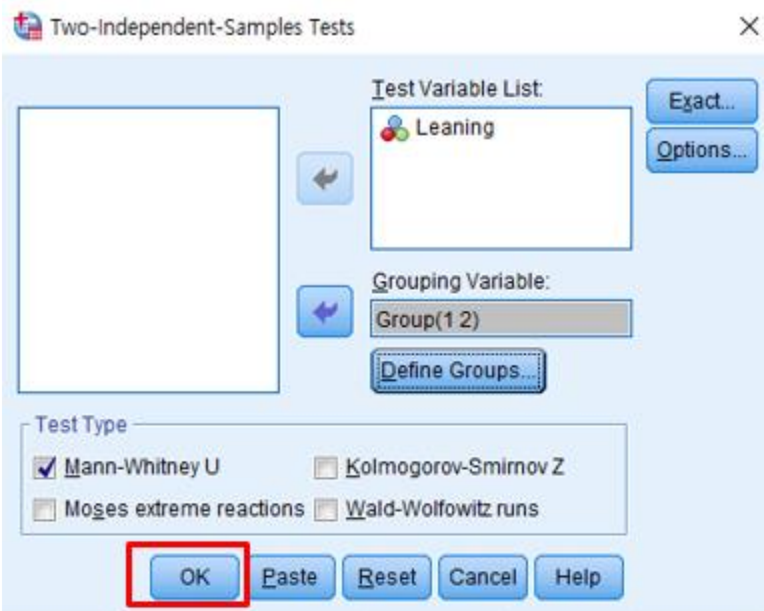
④



6. Mann Whitney U test

[TODO] With SPSS

⑤



6. Mann Whitney U test

[TODO] With SPSS

```
book>type mannwhitneyu-ex1.txt
2       4
3       6
2       5
4       4
9       8
2       3
5       4
3       2
4       4
3       5

book>java MannWhitneyU mannwhitneyu-ex1.txt
U = 31.0
z = -1.436, p = 0.1509
z' = -1.469, p' = 0.1418

book>
```

Mann-Whitney Test

Ranks

	Group	N	Mean Rank	Sum of Ranks
Leaning	1.00	10	8.60	86.00
	2.00	10	12.40	124.00
	Total	20		

Test Statistics[a]

	Leaning
Mann-Whitney U	31.000
Wilcoxon W	86.000
Z	-1.469
Asymp. Sig. (2-tailed)	.142
Exact Sig. [2*(1-tailed Sig.)]	.165[b]

a. Grouping Variable: Group

b. Not corrected for ties.

Data (Example #2)

- Means
  - 6.4 (MPA)
  - 3.7 (MPB)
- Data suggest MPA has more "cool appeal", but is the difference statistically significant?
- Data are ordinal (at least), ∴ a non-parametric test is used
- Which test?

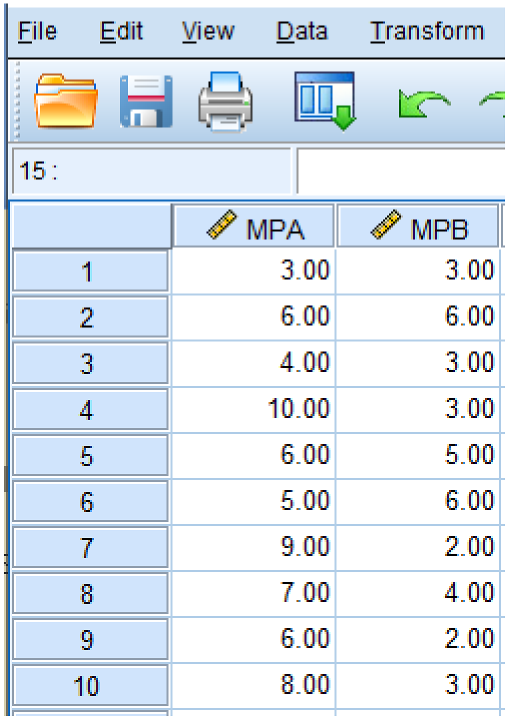
Design	Conditions	
	2	3 or more
Between-subjects (independent samples)	Mann-Whitney U	Kruskal-Wallis
Within-subjects (correlated samples)	Wilcoxon Signed-Rank	Friedman

Participant	MPA	MPB
1	3	3
2	6	6
3	4	3
4	10	3
5	6	5
6	5	6
7	9	2
8	7	4
9	6	2
10	8	3
	6.4	**3.7**

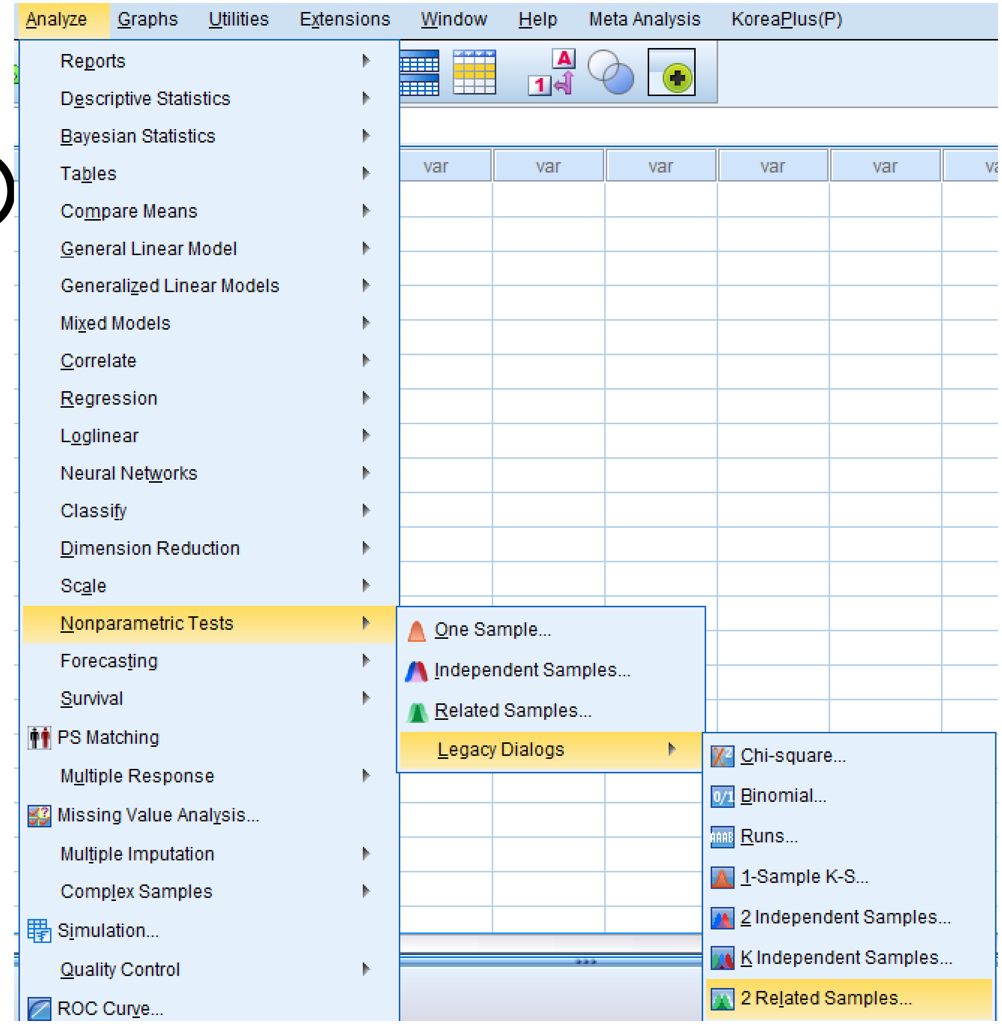
7. Wilcoxon signed-rank test

[TODO] With SPSS

①



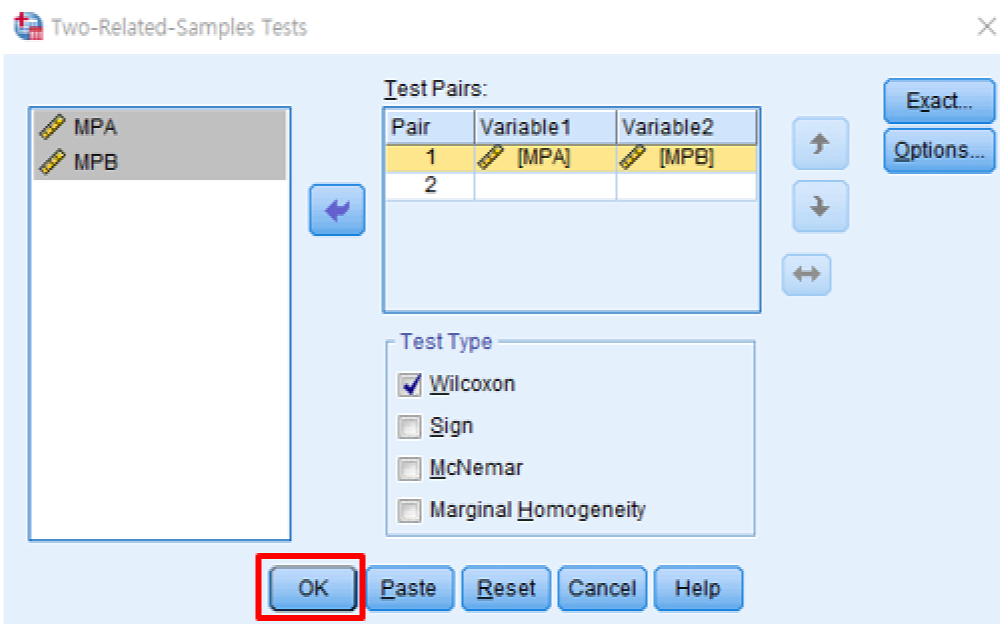
②



7. Wilcoxon signed-rank test

[TODO] With SPSS

③



7. Wilcoxon signed-rank test

[TODO] With SPSS

Test Statistics[a]

	MPB - MPA
Z	-2.254[b]
Asymp. Sig. (2-tailed)	.024

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Participant	MPA	MPB
1	3	3
2	6	6
3	4	3
4	10	3
5	6	5
6	5	6
7	9	2
8	7	4
9	6	2
10	8	3
	6.4	**3.7**

Data (Example #3)

- Means
  - 7.1 (20-29)
  - 4.0 (30-39)
  - 2.9 (40-49)
- Data suggest differences by age, but are differences statistically significant?
- Data are ordinal (at least), ∴ a non-parametric is used
- Which test?

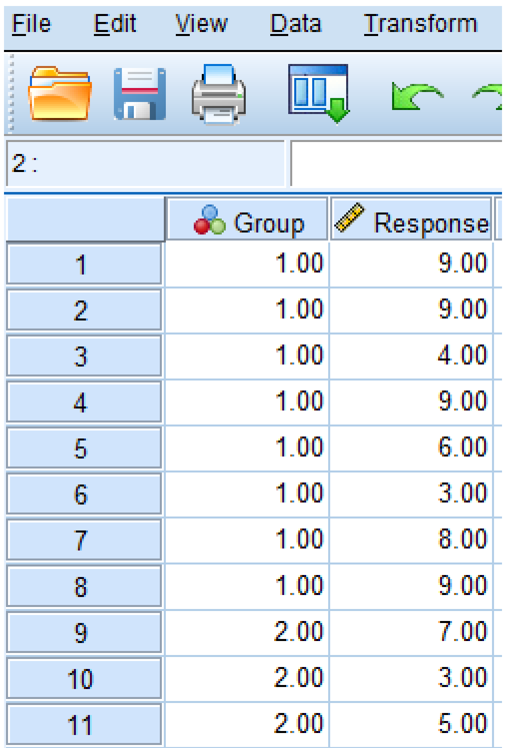
Design	Conditions	
	2	3 or more
Between-subjects (independent samples)	Mann-Whitney U	Kruskal-Wallis
Within-subjects (correlated samples)	Wilcoxon Signed-Rank	Friedman

A20-29	A30-39	A40-49
9	7	4
9	3	5
4	5	5
9	3	2
6	2	2
3	1	1
8	4	2
9	7	2
7.1	**4.0**	**2.9**

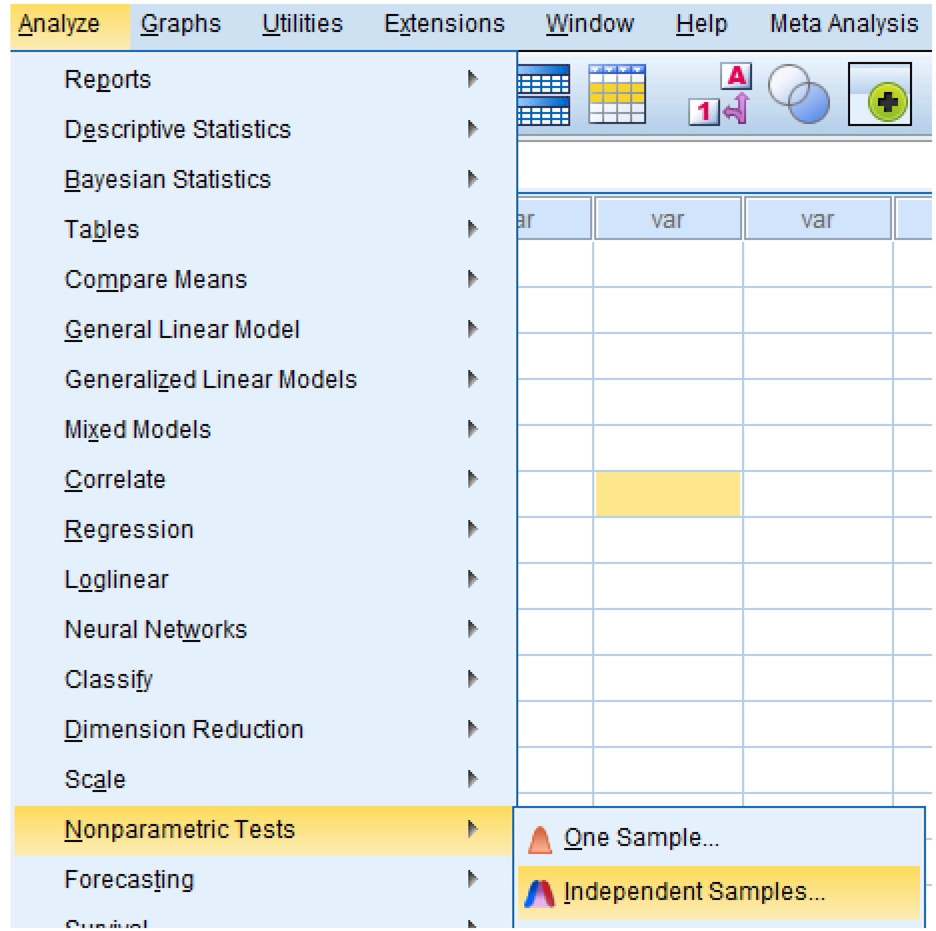
8. Kruskal-Wallis Test

[TODO] With SPSS

①

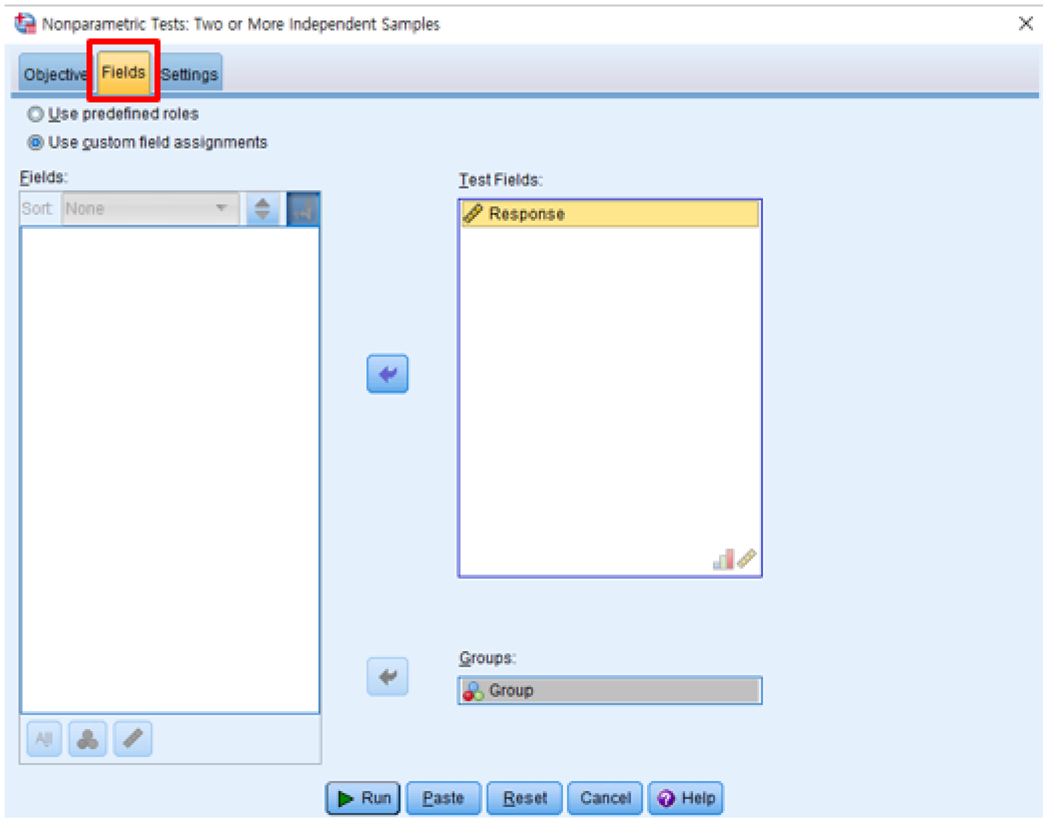


②



[TODO] With SPSS

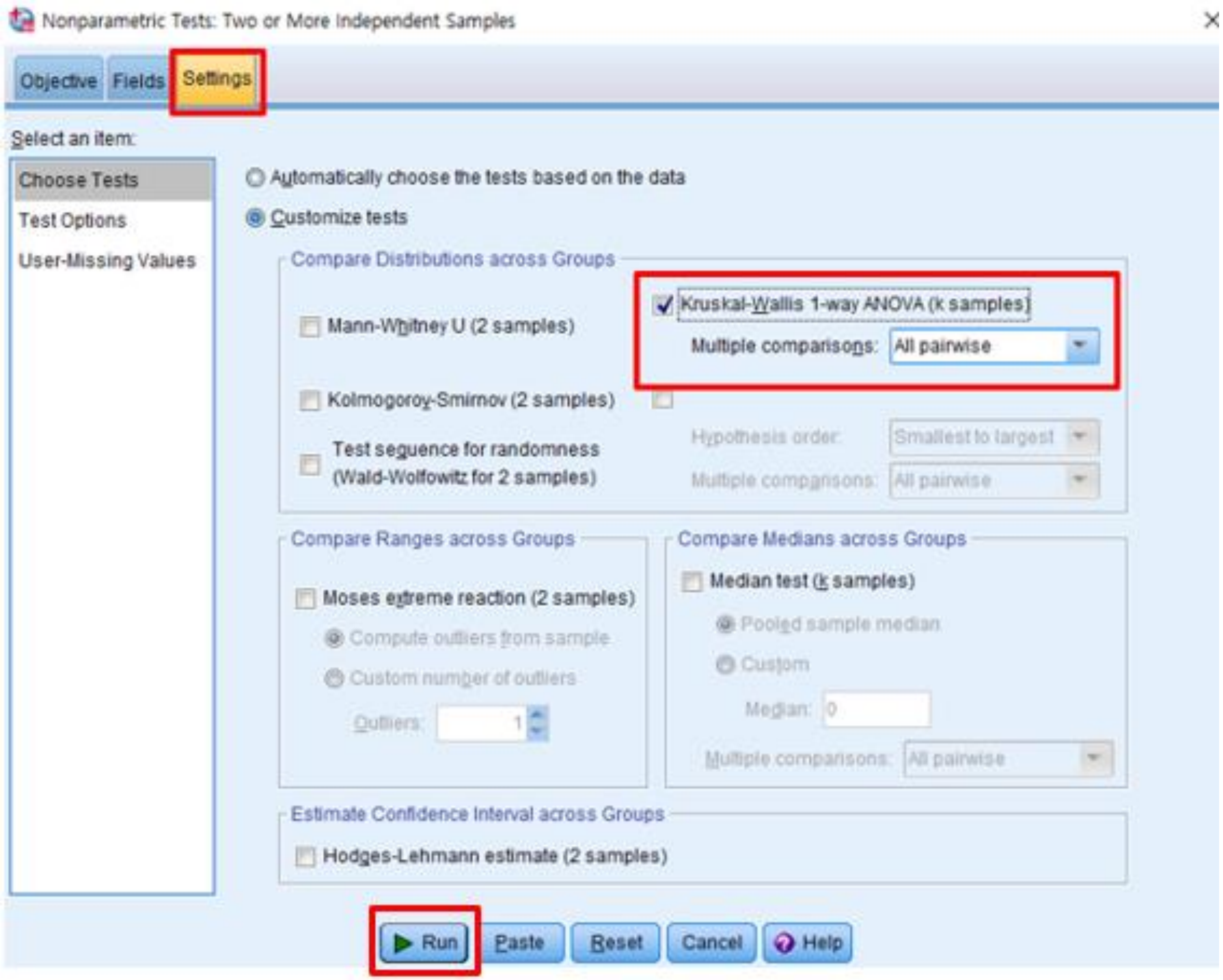
③



8. Kruskal-Wallis Test

[TODO] With SPSS

④



8. Kruskal-Wallis Test

[TODO] With SPSS

CMD

```
book>java KruskalWallis kruskalwallis-ex1.txt -ph
H = 9.421, p = 0.0090
H' = 9.605, p' = 0.0082

-----------------------------------------------
----- Multiple Comparisons Test (alpha = .05) ------
-----------------------------------------------
Pair 1:2 -->  7.4375 >=  7.6103  ?   -
Pair 1:3 --> 10.5625 >=  7.6103  ?   * (significant)
Pair 2:3 -->  3.1250 >=  7.6103  ?   -
-------------------------------------------------

book>_
```

Independent-Samples Kruskal-Wallis Test Summary

Total N	24
Test Statistic	9.605[a]
Degree Of Freedom	2
Asymptotic Sig.(2-sided test)	.008

a. The test statistic is adjusted for ties.

A20-29	A30-39	A40-49
9	7	4
9	3	5
4	5	5
9	3	2
6	2	2
3	1	1
8	4	2
9	7	2
7.1	**4.0**	**2.9**

Pairwise Comparisons of Group

Sample 1-Sample 2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.[a]
3.00-2.00	3.125	3.502	.892	.372	1.000
3.00-1.00	10.563	3.502	3.017	.003	.008
2.00-1.00	7.438	3.502	2.124	.034	.101

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

a. Significance values have been adjusted by the Bonferroni correction for multiple tests.

Data (Example #4)

- Means
  - 71.0 (A), 68.1 (B), 60.9 (C), 69.8 (D)
- Data suggest a difference in quality of results, but are the differences statistically significant?
- Data are ordinal (at least), ∴ a non-parametric test is used
- Which test?

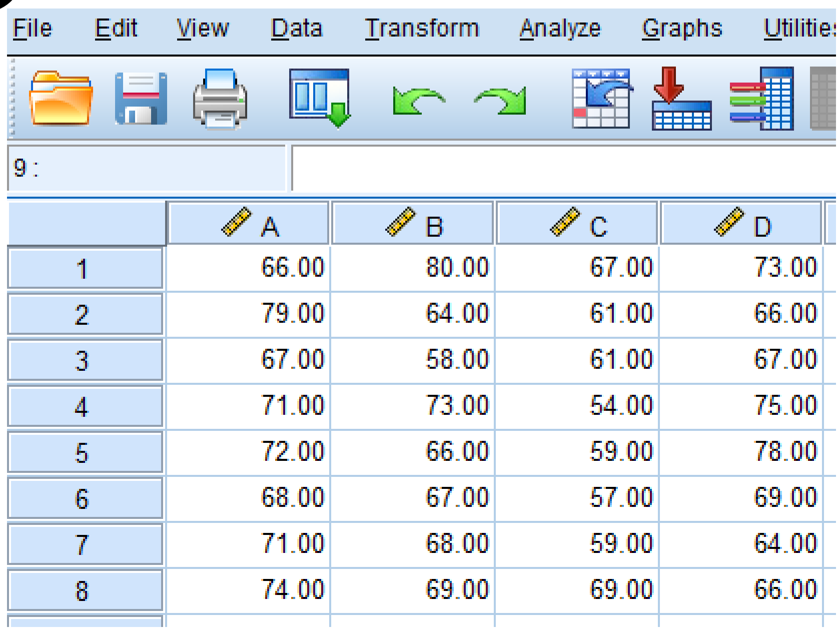
Design	Conditions	
	2	3 or more
Between-subjects (independent samples)	Mann-Whitney U	Kruskal-Wallis
Within-subjects (correlated samples)	Wilcoxon Signed-Rank	Friedman

Participant	A	B	C	D
1	66	80	67	73
2	79	64	61	66
3	67	58	61	67
4	71	73	54	75
5	72	66	59	78
6	68	67	57	69
7	71	68	59	64
8	74	69	69	66
	71.0	**68.1**	**60.9**	**69.8**

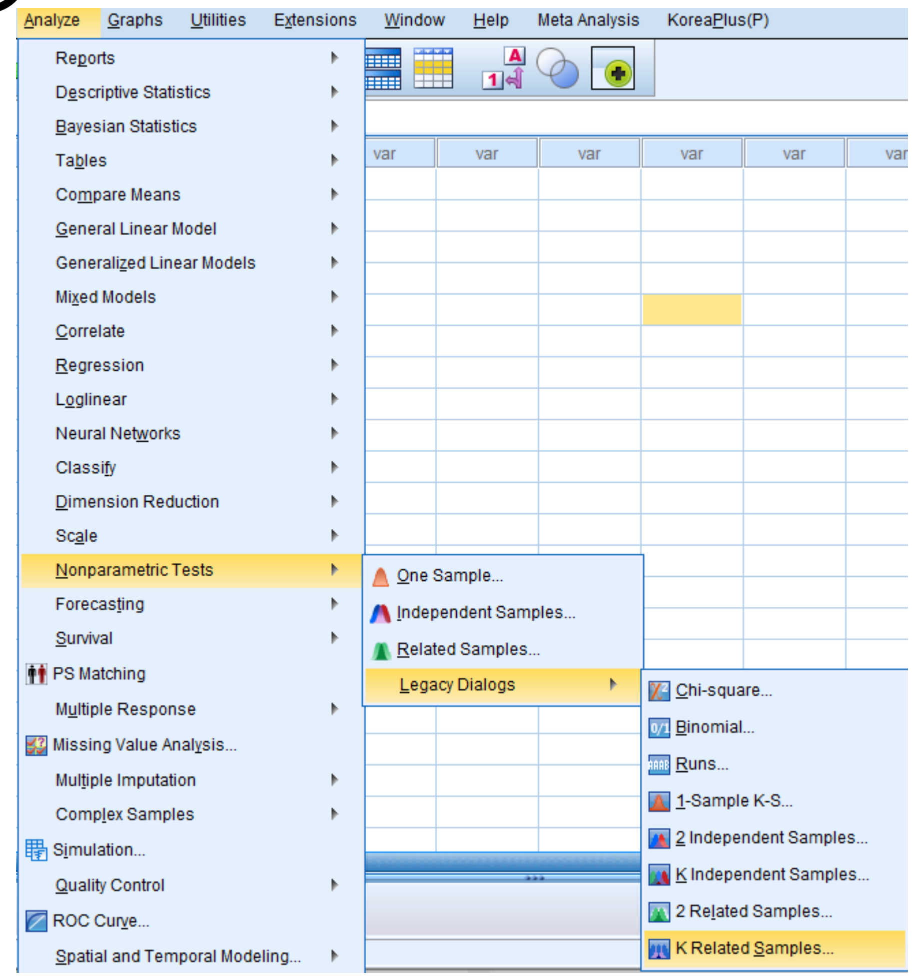
9. Friedman test

[TODO] With SPSS

①



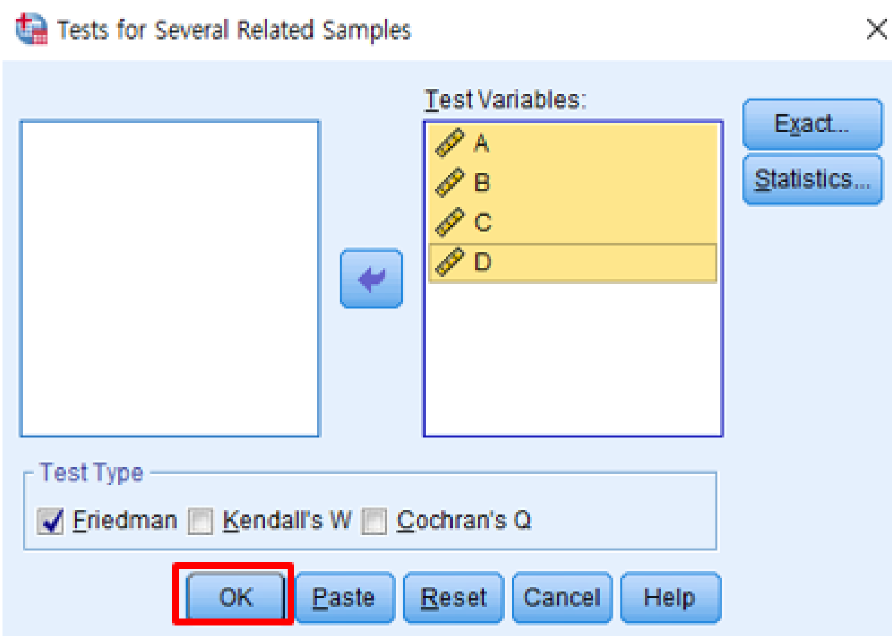
②



9. Friedman test

[TODO] With SPSS

③



[TODO] With SPSS

Test Statistics[a]

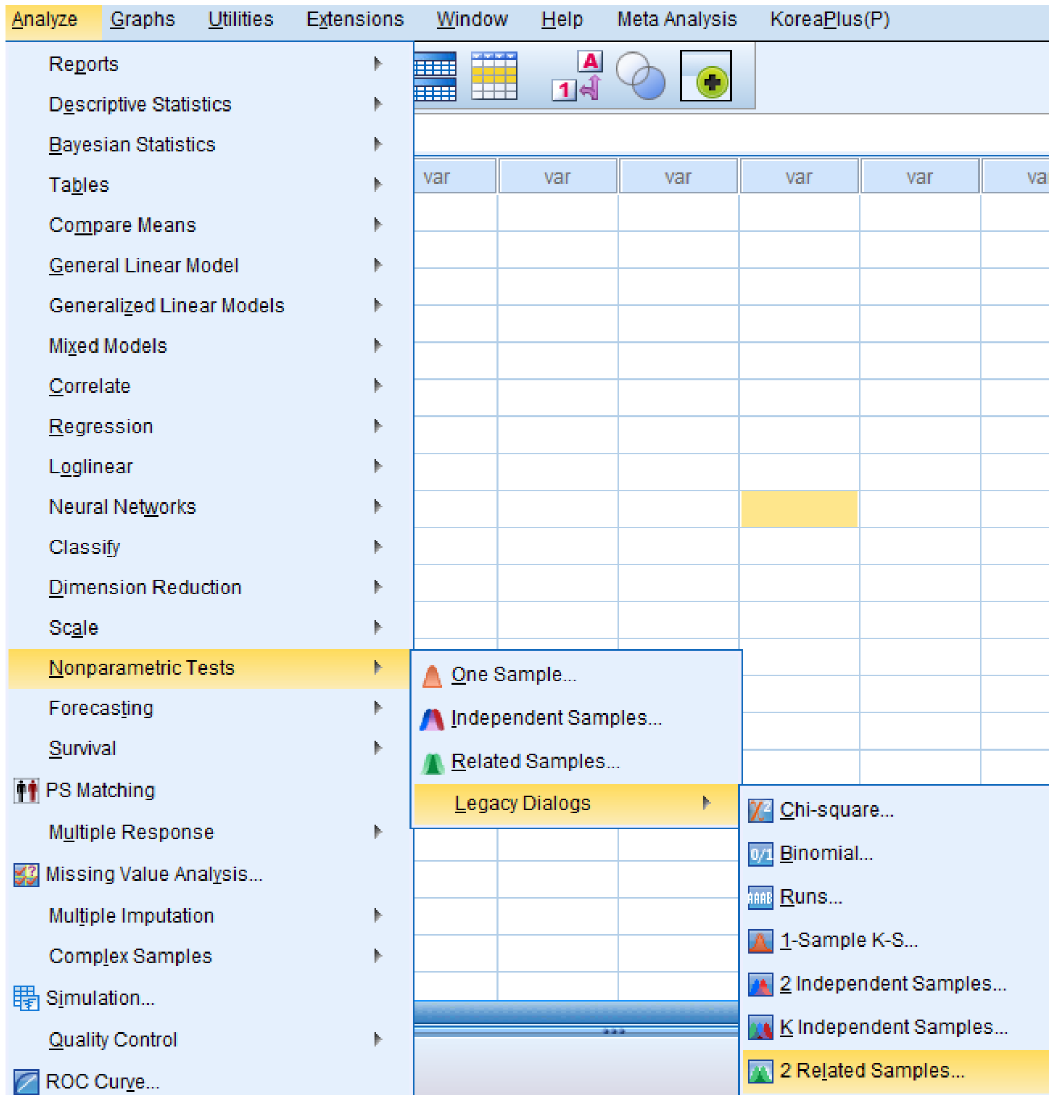
N	8
Chi-Square	8.692
df	3
Asymp. Sig.	.034

a. Friedman Test

9. Friedman test

[TODO] With SPSS – Post hoc comparison

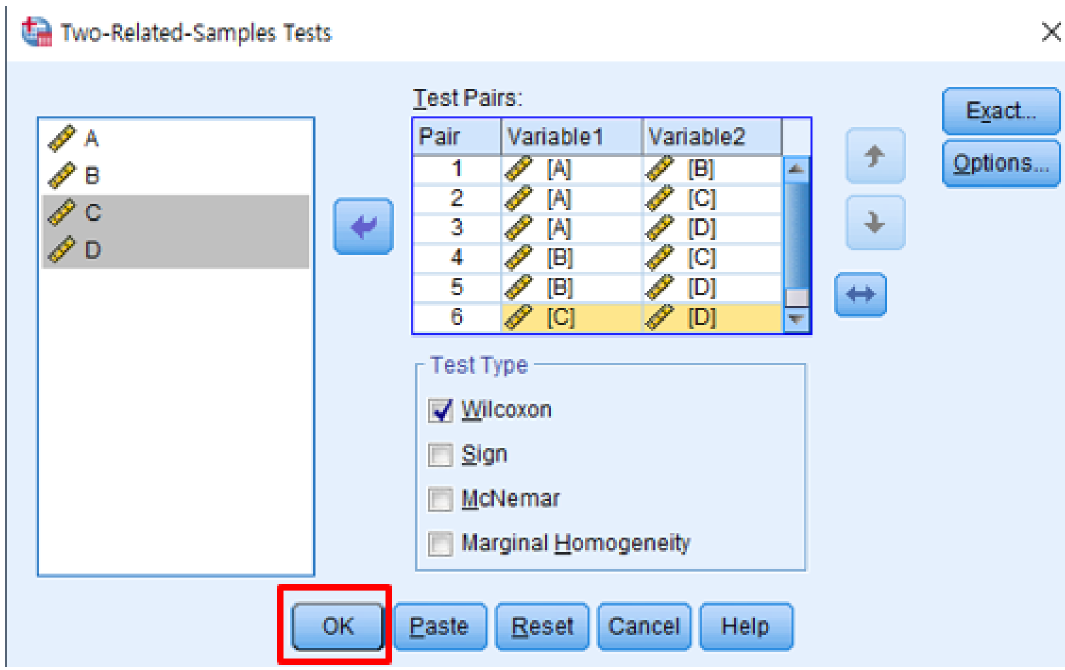
①



* We'll use Wilcoxon signed-rank test for post hoc comparison

[TODO] With SPSS – Post hoc comparison

②



* We'll use Wilcoxon signed-rank test for post hoc comparison

[TODO] With SPSS – Post hoc comparison

Test Statistics[a]

	B - A	C - A	D - A	C - B	D - B	D - C
Z	-1.260[b]	-2.380[b]	-.593[b]	-2.117[b]	-.422[c]	-2.386[c]
Asymp. Sig. (2-tailed)	.208	.017	.553	.034	.673	.017

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

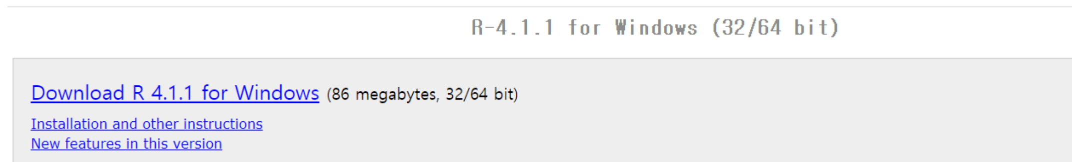
c. Based on negative ranks.

* We'll use Wilcoxon signed-rank test for post hoc comparison

Hypothesis Testing – R Practice (ART)

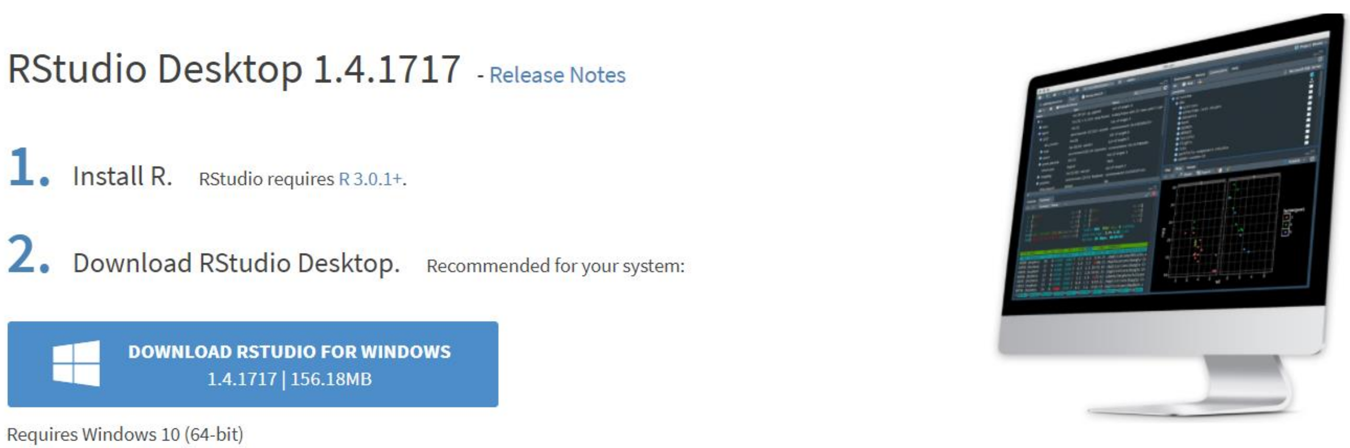
Based on Chapter 6 of
Human-Computer Interaction by I. S. MacKenzie

IDE and R download

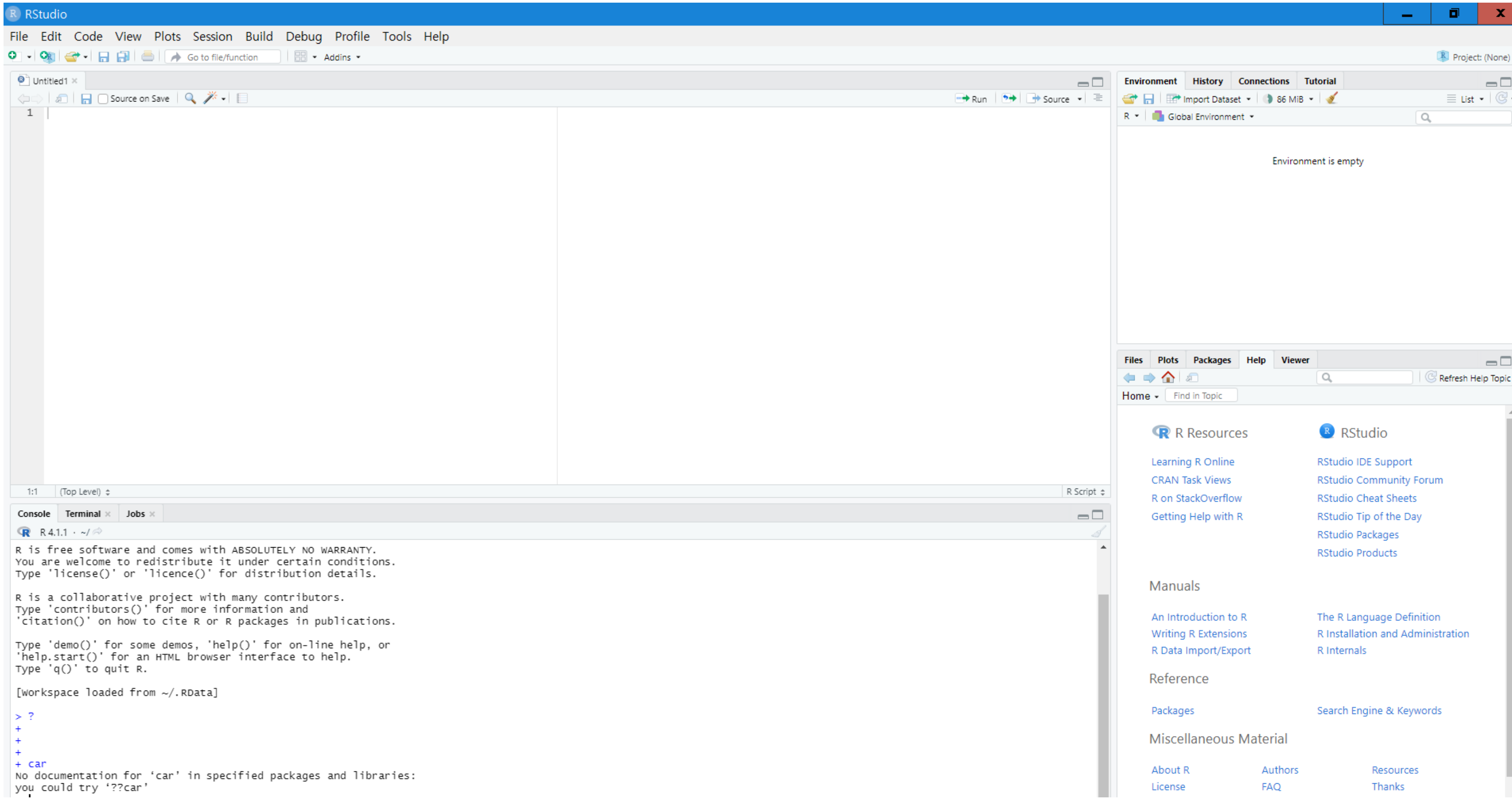


- R version – 4.1.1
  - Download: https://cran.rstudio.com/

- Rstudio version – 1.4.1717 (latest)
  - Download: https://www.rstudio.com/products/rstudio/download/#download



Rstudio


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Console Terminal Jobs
R 4.1.1 · ~/
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
[Workspace loaded from ~/.RData]
> ?
+
+
+
+ car
No documentation for 'car' in specified packages and libraries:
you could try '??car'
Environment History Connections Tutorial
Import Dataset
86 MiB
List
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Refresh Help Topic
Home
Find in Topic
R Resources
RStudio
Learning R Online
CRAN Task Views
R on StackOverflow
Getting Help with R
RStudio IDE Support
RStudio Community Forum
RStudio Cheat Sheets
RStudio Tip of the Day
RStudio Packages
RStudio Products
Manuals
An Introduction to R
Writing R Extensions
R Data Import/Export
The R Language Definition
R Installation and Administration
R Internals
Reference
Packages
Search Engine & Keywords
Miscellaneous Material
About R
Authors
Resources
License
FAQ
Thanks

Working directory

```
2:1    (Top Level) ÷

Console   Terminal ×   Jobs ×

R 4.1.1 · C:/Users/HCIL/Desktop/2021_10_15_R/

R version 4.1.1 (2021-08-10) -- "Kick Things"
Copyright (C) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

> setwd("C:/Users/HCIL/Desktop/2021_10_15_R")
> 
```

Basic R grammar for this class

Useful document

- https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf

The on-line help of R

- If you want to find the lm() (linear model)
  - ?lm
  - help(lm)
  - Help("lm")

  "help" only searches in the packages which are loaded in memory
- If no documentation for 'lm' in specified packages and libraries: you could try
  - 'help.search("lm")
  - ??lm
-

aov {formula, data, …}

- Fit an Analysis of Variance Model by a call to lm (linear model) for each stratum
- In "stats" package which is default package if you install R

Formula

- Y ~ model
  - Y: the analyzed response (dependent variable)
  - Model: a set of terms for which some parameters are to be estimated.
  - For this class..
    - A : estimate the effect of A, for one-way ANOVA (one independent variable)
    - A + B : estimate the effects of A and B, but not consider the interaction effect between A, B (two independent variables: A, B)
    - A:B : interactive effect between A and B
    - A*B : estimate the effects of A and B, and interactive effect between A and B
      - It is same to A + B + A:B

aov {formula, data, ...}

aov() accepts a particular syntax to define random effects (within-subjects design)

For repeated measure (within)

- Y ~ A + Error(Subjects/A)
-

ANOVA Example #1

```
data = read.table("anova-ex1a.txt", header = FALSE,
col.names = c("DV", "IV", "Participant"))
```

```
str(data) # check a data mode
```

The result:

```
> str(data)
'data.frame':	20 obs. of  3 variables:
 $ DV         : num  5.3 3.6 5.2 3.6 4.6 4.1 4 4.8 5.2 5.1 ...
 $ IV         : int  1 1 1 1 1 1 1 1 1 1 ...
 $ Participant: int  1 2 3 4 5 6 7 8 9 10 ...
```

ANOVA Example #1

Because the data mode is numeric or integer, we change to the mode of independent variable to factor.

```
# ANOVA (within)
fit <- aov(DV ~ factor(IV) +
Error(factor(Participant)/factor(IV)), data)

# result
summary(fit)
```

```
Error: factor(Participant)
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  9  4.884  0.5427

Error: factor(Participant):factor(IV)
           Df Sum Sq Mean Sq F value Pr(>F)
factor(IV)  1  4.141   4.141   9.593 0.0128 *
Residuals   9  3.884   0.432
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Example #1 – ANOVA

$ java Anova2 anova-ex1.txt 10 2 . . -a

```
=================================================================
Effect            df        SS          MS          F        p
-----------------------------------------------------------------
Participant        9     4.884       0.543
F1                 1     4.140       4.140      9.593    0.0128
F1_x_Par           9     3.885       0.432
=================================================================
```

Probability of obtaining the observed data if the null hypothesis is true.

Thresholds for “p”

- .05, .01, .005, .001, .0005, .0001

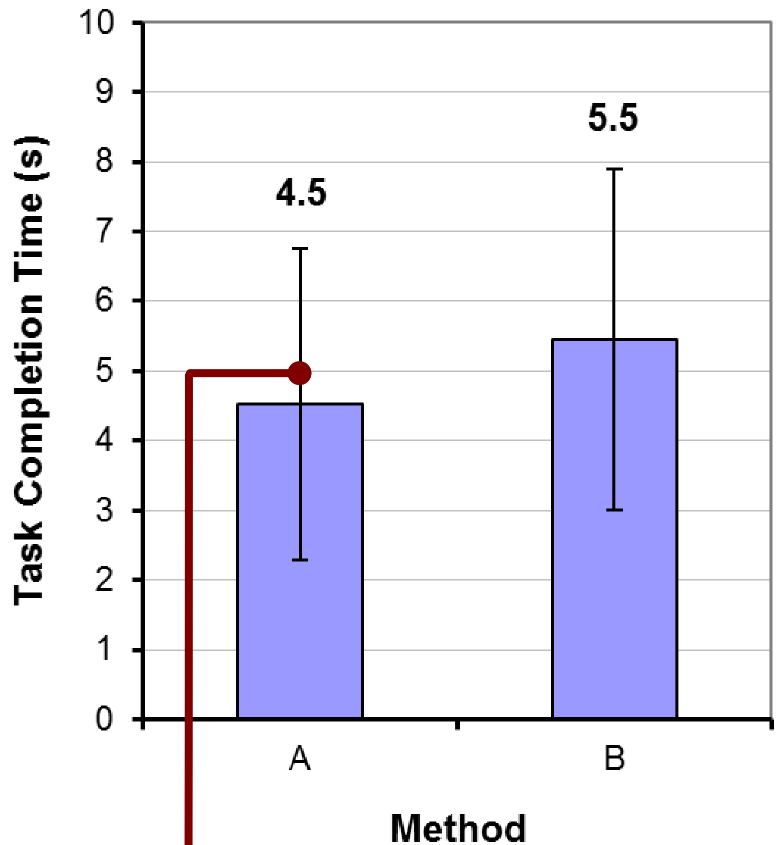
The mean task completion time for Method A was 4.5 s. This was 20.1% less than the mean of 5.5 s observed for Method B. The difference was statistically significant ($F_{1,9} = 9.593$, $p < 0.05$).

ANOVA Example #1 (cont.)

```
# In order to change the mode, you can use as.factor()
 data$Participant <- as.factor(data$Participant)
 data$IV <- as.factor(data$IV)
 str(data)

# same function
fit <- aov(DV ~ IV + Error(Participant/IV), data)
summary(fit)
```

Example #2 – Do It your self



Participant	Method	
	A	B
1	2.4	6.9
2	2.7	7.2
3	3.4	2.6
4	6.1	1.8
5	6.4	7.8
6	5.4	9.2
7	7.9	4.4
8	1.2	6.6
9	3.0	4.8
10	6.6	3.1
Mean	4.5	5.5
SD	2.23	2.45

Error bars show
±1 standard deviation

ANOVA Example #2

```
data = read.table("anova-ex2a.txt", header = FALSE,
col.names = c("DV", "IV", "Participant"))
```

Estimate the effect of IV on the DV.
All participants performed two methods, A and B (within-subject design)

```
fit <- aov(DV ~ factor(IV) +
Error(factor(Participant)/factor(IV)), data)
summary(fit)
```

```
Error: factor(Participant)
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals  9  37.37   4.152

Error: factor(Participant):factor(IV)
           Df Sum Sq Mean Sq F value Pr(>F)
factor(IV)  1   4.32   4.325   0.626  0.449
Residuals   9  62.14   6.904
>
```

Example #2 – ANOVA (Compare to the result)

```
$ java Anova2 anova-ex2.txt 10 2 . . -a
```

Effect	df	SS	MS	F	p
Participant	9	37.373	4.153		
F1	1	4.324	4.324	0.626	0.4491
F1_x_Par	9	62.140	6.904		

Probability of obtaining the observed data if the null hypothesis is true.

Note: For non-significant effects, use "ns" if $F < 1.0$, or "$p > .05$" if $F > 1.0$.

The mean task completion time was 4.5 s for Method A and 5.5 s for Method B. As there was substantial variation in the observations across participants, the difference was not statistically significant as revealed in an analysis of variance ($F_{1,9} = 0.626$, ns).

Between-subjects Design

```
data = read.table("anova-ex3.txt", header = FALSE,
col.names = c("DV", "IV"))
fit <- aov(DV ~ factor(IV), data)
summary(fit)
```

```
==========================================================================
Effect                df        SS             MS          F        p
--------------------------------------------------------------------------
F3                     1        18.063         18.063      3.781    0.0722
Residual              14        66.875          4.777
==========================================================================
```

```
            Df Sum Sq Mean Sq F value Pr(>F)
factor(IV)   1  18.06  18.063   3.781 0.0722 .
Residuals   14  66.88   4.777
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

What is difference of the command from within-subject design?

Two-way ANOVA (within)

```
data <- read.table("anova-ex4a.txt", header = FALSE)
colnames(data) <- c("Participant", "Device", "Task", "DV")
head(data)
str(data)
```

The formula for two-way anova

- Y ~ A * B
  - Identical to Y ~ A + B + A:B
- + Error (subjects / (A*B))
  - For within-subject design
  - Because A, B is nested condition under the subjects

Two-way ANOVA (within)

```
# check the error term
fit <- aov(DV ~ factor(Device)*factor(Task) +
Error(factor(Participant)/(factor(Device)*factor(Task))),
data)
summary(fit)
```

Two-way ANOVA (within)

Result:

```
Error: factor(Participant)
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 11  134.8   12.25

Error: factor(Participant):factor(Device)
               Df Sum Sq Mean Sq F value  Pr(>F)
factor(Device)  2    121   60.51   5.865 0.00909 **
Residuals      22    227   10.32
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: factor(Participant):factor(Task)
             Df Sum Sq Mean Sq F value Pr(>F)
factor(Task)  1   0.89   0.889   0.076  0.787
Residuals    11 128.11  11.646

Error: factor(Participant):factor(Device):factor(Task)
                           Df Sum Sq Mean Sq F value
factor(Device):factor(Task)  2    121   60.51   5.435
Residuals                   22    245   11.14
                           Pr(>F)
factor(Device):factor(Task) 0.0121 *
Residuals
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA

$ java Anova2 anova-ex4.txt 12 3 2 . -a

Effect	df	SS	MS	F	p
Participant	11	134.778	12.253		
F1	2	121.028	60.514	5.865	0.0091
F1_x_Par	22	226.972	10.317		
F2	1	0.889	0.889	0.076	0.7875
F2_x_Par	11	128.111	11.646		
F1_x_F2	2	121.028	60.514	5.435	0.0121
F1_x_F2_x_Par	22	244.972	11.135		

ART : Aligned ranked transform in R

- For Multi-factor Non-parametric Data
  - Preprocess data before ANOVA
  - Aligning step
    - Remove the effects of all factors and interactions except for one.
  - Rank-transform
    - Remove the skewness of the distribution.
  - ANOVA
    - Calculate the effect of the factor in focus.
  - (Repeat for other factors and interactions)

- The "ARTool" package allows very easy way to use ART
  - Install.packages("ARTool")
  - Library(ARTool)
  - reference: http://depts.washington.edu/acelab/proj/art/index.html

ART example

Data: Higgins1990Table5

```
library(ARTool)
data(Higgins1990Table5, package = "ARTool")

str(Higgins1990Table5)
head(Higgins1990Table5, n=8)
```

ART example

Step 1: Transform the data (Between)

```
m <- art(DryMatter ~ Moisture * Fertilizer , data =
Higgins1990Table5)

m$residuals
m$estimated.effects
m$aligned
m$aligned.ranks
```

In this process,

1) Step 1: Compute residuals: Y – cell mean
2) Compute estimated effects for all main and interaction effects
3) Compute aligned response Y′
4) Assign averaged ranks Y′′

ART example

Step 2: Verify appropriateness of ART

To verify that the ART procedure was correctly applied and is appropriate for this dataset, we can look at the output of summary

`Summary(m)`

```
> summary(m)
Aligned Rank Transform of Factorial Model

Call:
art(formula = DryMatter ~ Moisture * Fertilizer, data = Higgins1990Table5)

Column sums of aligned responses (should all be ~0):
           Moisture            Fertilizer Moisture:Fertilizer
                  0                     0                   0

F values of ANOVAs on aligned responses not of interest (should all be ~0):
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      0       0       0       0       0       0
> 
```

ART example

Step 3: Run the ANOVA

`anova(m)`

```
> anova(m)
Analysis of Variance of Aligned Rank Transformed Data

Table Type: Anova Table (Type III tests)
Model: No Repeated Measures (lm)
Response: art(DryMatter)

                     Df Df.res F value     Pr(>F)    
1 Moisture            3     32 41.5199 3.8513e-11 ***
2 Fertilizer          3     32 69.4604 4.2015e-14 ***
3 Moisture:Fertilizer 9     32  2.9388   0.011685   *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

If you want to use different model like repeated measures ANOVA, you add Error() term to the art() formula as using aov().

ART homework

CS584 Homework 6

Problem 7. (ART)

- Data: Dataset(p1-p12).csv
- Within-subjects (SubjectName)
- Independent variables: Radius (level: 6, 8, 10), and itemNum (level: 4, 6)
- Dependent variable: time

Using ARTool, 1) Estimate the effects of Radius and itemNum on time, and 2) report the p-value and F-value of the estimations. (You can use read.csv() instead of read.table())