Unit 3 – Introduction to Nonparametrics
Practice Problems
Solutions – *Online Apps*

1. (*Source: Moore, D and McCabe, GP. Introduction to the Practice of Statistics, Third Edition)*
Food products are often enriched with vitamins and other supplements. Does the level of a supplement decline over time so that the user receives less than the manufacturer intended? The following are n=9 observations of vitamin C levels (milligrams per 100 grams) in a wheat soy blend, a flour-like product supplied by international aid programs mainly for feeding children. Each bag was measured twice, first at the factory and a second time five months later in Haiti. Researchers suspect that vitamin C levels are generally higher at the factory than they were five months later. We would like to test the hypotheses:

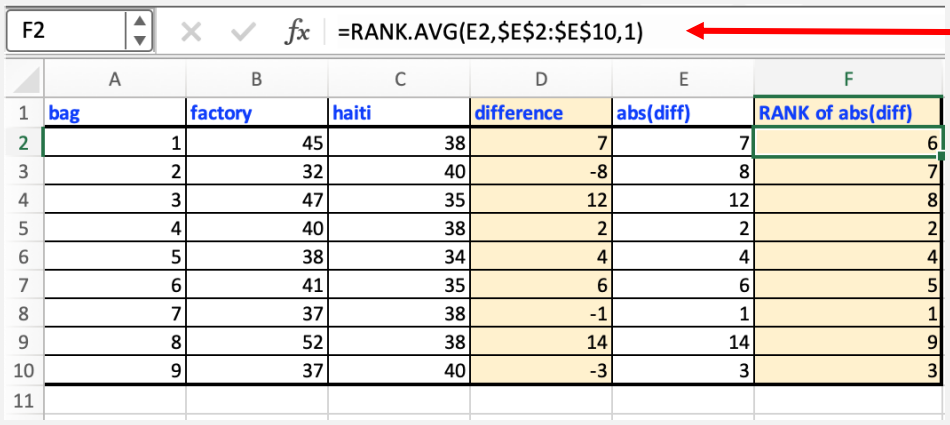
Null, H_O: vitamin C has the same distribution at both times
Alternative, H_A: vitamin C is systematically higher at the factory (one sided)

Bag	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Factory	45	32	47	40	38	41	37	52	37
Haiti	38	40	35	38	34	35	38	38	40

(a) What is the correct nonparametric test here?

Wilcoxon Signed Rank Test for single sample of paired data.

(b) Produce a copy of the table above that shows the differences and the ranks of the *absolute* differences.
Tip. Consider putting your data into Excel first. For future questions, you can then do EDIT/COPY/PASTE.



	A	B	C	D	E	F
1	bag	factory	haiti	difference	abs(diff)	RANK of abs(diff)
2	1	45	38	7	7	6
3	2	32	40	-8	8	7
4	3	47	35	12	12	8
5	4	40	38	2	2	2
6	5	38	34	4	4	4
7	6	41	35	6	6	5
8	7	37	38	-1	1	1
9	8	52	38	14	14	9
10	9	37	40	-3	3	3
11						

Key: `RANK.AVG(cell of interest, full list of cells to rank, 1 for ascending ranking)`
I used this to populate cell F2. From there I did a COPY of F2 followed by PASTE to populate cells F3 ... F10
In order for the EDIT/COPY/PASTE to work properly the function in cell F2 MUST have the dollar signs ($) that you see.

(c) Corresponding to your answer to "a", what is the corresponding normal theory test that would have been performed if the assumptions were met?

Paired t-test for single sample of paired data.

(d) By any means you like, perform the nonparametric test you gave in "a". In 1-2 sentences at most, report. What do you conclude?

I used the app suggested on page 13 of the Unit 3 notes:
https://astatsa.com/WilcoxonTest/

Mann Whitney Test calculator (for unpaired data), Wilcoxon Signed Rank Test calculator (for paired data)

- **Mann Whitney** test, also known as **Mann Whitney U** test, **Mann Whitney Wilcoxon** test, and as **Wilcoxon rank sum test**, when applied to *unpaired* two sample data
- **Wilcoxon signed rank test**, when applied to *paired* one sample data

Select:

○ **Mann Whitney test** for *unpaired* data, two sample (default)
◉ **Wilcoxon signed rank test** for *paired* data, one sample

Enter your two columns of *paired* numerical data below:

Paired [A, B] across rows,
comma or space separated

```
45 38
32 40
47 35
40 38
38 34
41 35
37 38
52 38
37 40
```

Clear

Additional parameters and options:

Alternative hypothesis	Null hypothesis	Confidence interval of $\hat{\mu}$
○ two sided, estimated location shift $\hat{\mu} \neq \mu_0$ ◉ greater, $\hat{\mu} > \mu_0$ ○ lesser, $\hat{\mu} < \mu_0$	location shift (mean) μ_0 0 Replace/ edit default $\mu_0 = 0$	◉ 95% ○ 99%
Calculate exact p-value ◉ exact ○ approximate, for large samples	Continuity correction ◉ True, Yes ○ False, No	

	A	B
Obs.#		
1	45	38
2	32	40
3	47	35
4	40	38
5	38	34
6	41	35
7	37	38
8	52	38
9	37	40

Wilcoxon signed rank test

Test statistic V : 34

p-value : 0.101562

null hypothesis $\mu_0 = 0.0$

alternative hypothesis: greater, $\hat{\mu} > \mu_0$

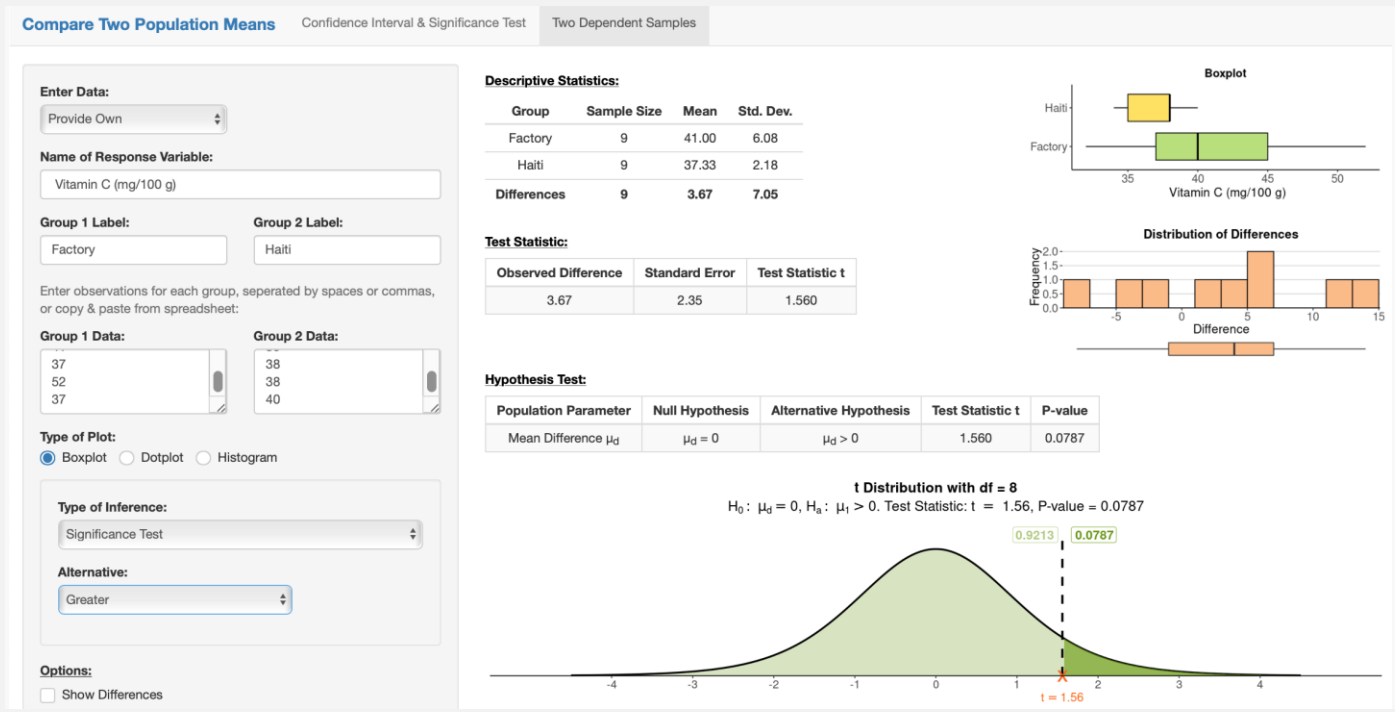
R code to reproduce these results:

```
# Copy-paste these lines into the R command prompt.
# Lines that begin with the # character are taken as comment lines by R.

A <- c(45, 32, 47, 40, 38, 41, 37, 52, 37)

B <- c(38, 40, 35, 38, 34, 35, 38, 38, 40)

wilcox.test(A, B, paired = TRUE, alternative = "greater", mu = 0.0,
      exact = TRUE, correct = TRUE, conf.int = TRUE, conf.level = 0.95)
```

The null hypothesis is ***not*** rejected (p-value = .10). This sample does ***not*** provide statistically significant evidence that the level of vitamin C supplement declines with time after it leaves the factory.

(e) By any means you like (and just for comparison), perform the normal theory test you gave in "c", even knowing that it is not appropriate here.
I used art of stat https://artofstat.com

Online Web Apps > Compare Two Means > at top, tab**: Two Dependent Samples**



The p-value from the (*admittedly not appropriate*) normal theory paired t-test is not so different (p-value = .08) from what was obtained by the Wilcoxon Signed Rank test (p-value = .10).

2. (*Source: Moore, D and McCabe, GP. Introduction to the Practice of Statistics, Third Edition).*
The most used measure of economic growth is the rate of growth in a country's total output of goods and services gauged by the gross domestic product (GDP) adjusted for inflation. The level of a country's GDP growth reflects the growth of businesses, jobs, and personal income. The following are World Bank data on the average growth of GDP (percent per year) for the period 2010 to 2013 in developing countries of Europe:

Developing Countries: Europe

Country	Growth	Country	Growth
Albania	2.3	Macedonia, FYR	2.1
Armenia	4.4	Moldova	5.5
Azerbaijan	3.2	Montenegro	1.7
Belarus	4.0	Romania	1.3
Bosnia and Herzegovina	0.4	Serbia	0.9
Bulgaria	0.9	Turkey	6.0
Georgia	5.6	Ukraine	2.9
Kosovo	3.4		

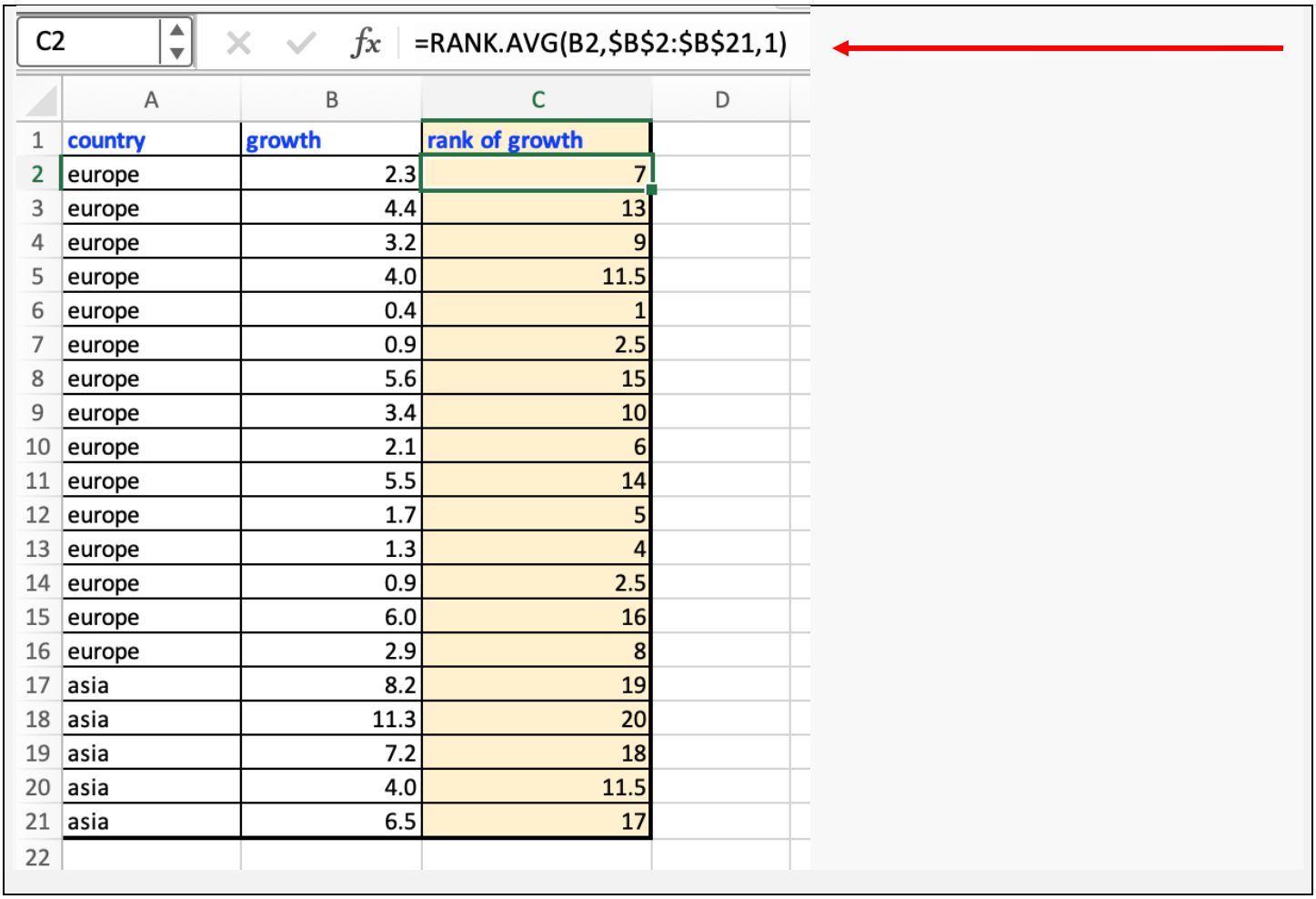
Developing Countries: Central Asia

Country	Growth	Country	Growth
Uzbekistan	8.2	Kyrgyz Republic	4.0
Turkmenistan	11.3	Kazakhstan	6.5
Tajikistan	7.2		

a). Suppose we are interested in researching the similarity of average growth of GDP in the two groups of developing countries: Europe versus Central Asia. State the null and alternative hypotheses.

H_O: Distribution of growth in Europe = distribution of growth in Central Asia.
H_A: Distribution of growth in Europe ≠ distribution of growth in Central Asia (two-sided).

(b). Produce a copy of the table above that shows the ranks of the 20 observations. Take care in your ranking to handle ties.



	A	B	C	D
1	country	growth	rank of growth	
2	europe	2.3	7	
3	europe	4.4	13	
4	europe	3.2	9	
5	europe	4.0	11.5	
6	europe	0.4	1	
7	europe	0.9	2.5	
8	europe	5.6	15	
9	europe	3.4	10	
10	europe	2.1	6	
11	europe	5.5	14	
12	europe	1.7	5	
13	europe	1.3	4	
14	europe	0.9	2.5	
15	europe	6.0	16	
16	europe	2.9	8	
17	asia	8.2	19	
18	asia	11.3	20	
19	asia	7.2	18	
20	asia	4.0	11.5	
21	asia	6.5	17	
22				

(c). What is the correct nonparametric test here?

Wilcoxon Rank Sum Test for Two Independent Groups

(d) Consider your answer to "c". What is the corresponding normal theory test that would have been performed if the assumptions were met?

Two Sample t-Test for Two Independent Groups

(e) By any means you like, perform the nonparametric test you gave in "c". In 1-2 sentences at most, report. What do you conclude?

Again, I used the app suggested on page 13 of the Unit 3 notes:
https://astatsa.com/WilcoxonTest/

Mann Whitney Test calculator (for unpaired data), Wilcoxon Signed Rank Test calculator (for paired data)

- **Mann Whitney** test, also known as **Mann Whitney U** test, **Mann Whitney Wilcoxon** test, and as **Wilcoxon rank sum test**, when applied to *unpaired* two sample data
- **Wilcoxon signed rank test**, when applied to *paired* one sample data

Select:

◉ **Mann Whitney test** for *unpaired* data, two sample (default)
○ **Wilcoxon signed rank test** for *paired* data, one sample

Enter your two columns of *unpaired* numerical data below:

A	B
2.3 4.4 3.2 4.0 0.4 0.9 5.6 3.4 2.1 5.5 1.7 1.3	8.2 11.3 7.2 4.0 6.5

Additional parameters and options:
Defaults are pre-selected, change as necessary.

Alternative hypothesis	Null hypothesis location shift (mean) μ_0	Confidence interval of $\hat{\mu}$
◉ two sided, estimated location shift $\hat{\mu} \neq \mu_0$ ○ greater, $\hat{\mu} > \mu_0$ ○ lesser, $\hat{\mu} < \mu_0$	0 Replace/ edit default $\mu_0 = 0$	◉ 95% ○ 99%
Calculate exact p-value	Continuity correction	
◉ exact ○ approximate, for large samples	◉ True, Yes ○ False, No	

Mann Whitney (Wilcoxon) test for unpaired data, i.e. two sample

Input Data (scroll as required):

Obs.#	A	B
1	2.3	8.2
2	4.4	11.3
3	3.2	7.2
4	4.0	4.0
5	0.4	6.5
6	0.9	
7	5.6	
8	3.4	
9	2.1	

Results:

Mann Whitney test, also known as Wilcoxon rank sum test with continuity correction

Test statistic W : 4

p-value : 0.004526

There are ties in the ranks, so the exact p-value cannot be computed. The p-value shown is inexact but fairly robust when the number of ties is small.

null hypothesis $\mu_0 = 0.0$

alternative hypothesis: two sided, $\hat{\mu} \neq \mu_0$

R code to reproduce these results:

```
# Copy-paste these lines into the R command prompt.
# Lines that begin with the # character are taken as comment lines by R.

A <- c(2.3, 4.4, 3.2, 4.0, 0.4, 0.9, 5.6, 3.4, 2.1, 5.5, 1.7, 1.3, 0.9, 6.0, 2.9)

B <- c(8.2, 11.3, 7.2, 4.0, 6.5)

wilcox.test(A, B, paired = FALSE, alternative = "two.sided", mu = 0.0,
      exact = TRUE, correct = TRUE, conf.int = TRUE, conf.level = 0.95)
```

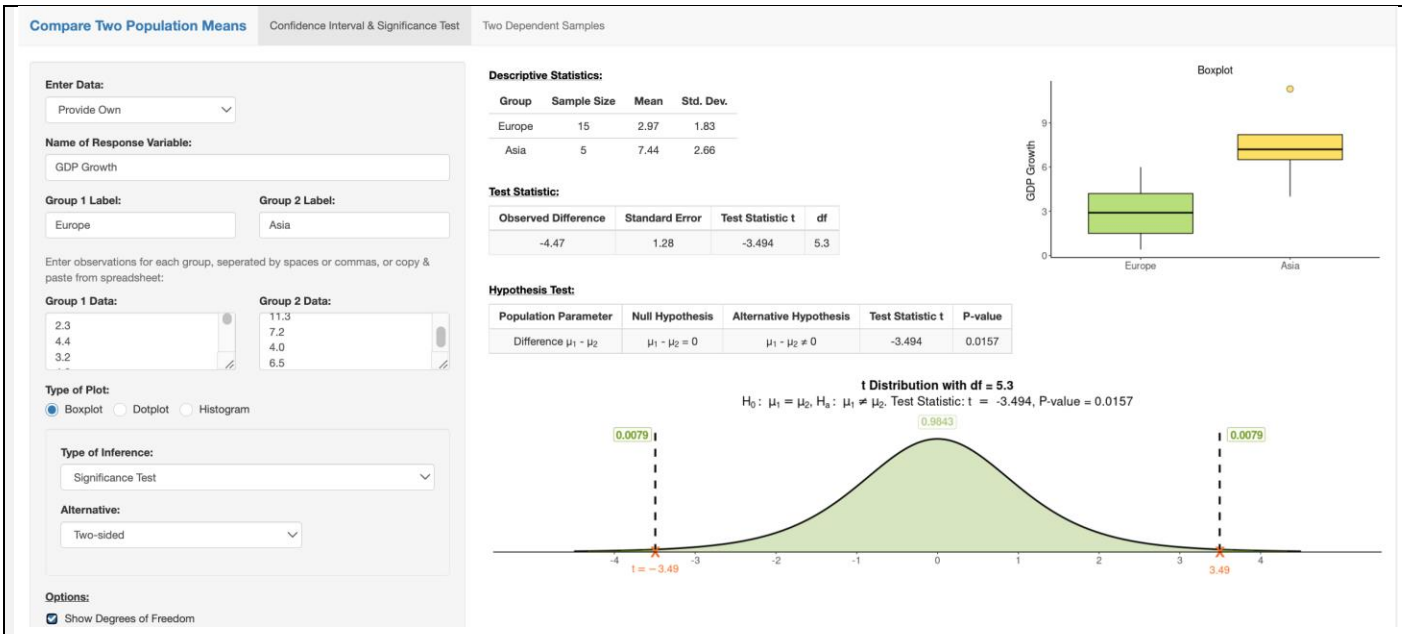
The null hypothesis *is* rejected (p-value = .005). Assumption of the null hypothesis and its application to the data has led to a highly unlikely result. This sample ***does*** provide statistically significant evidence that average growth of GDP (percent per year) for the period 2010 – 2013 was different in developing countries of Central Asia than it was in developing countries of Europe.

NOTE: The R output here doesn't come with much by way of clues regarding the nature of this two-sided difference however. A cursory look at the data, however, suggests that the average growth of GDP was higher in the developing countries of Central Asia.

(f) By any means you like (and just for comparison), perform the normal theory test you gave in "d", even knowing that it is not appropriate here.

I used art of stat https://artofstat.com

Online Web Apps > Compare Two Means > at top, tab**: Confidence Interval & Significance Test**



Interesting. Here, the p-value from the (*admittedly not appropriate*) normal theory two independent samples t-test (p-value = .0157) is quite a bit different from what was obtained by the Wilcoxon Rank Sum test (p-value = .005). Note – the discrepancy could go either way.