In [17]:

```
import sys
import importlib as imp
if ('Jupytils' in sys.modules):
    reloaded = imp.reload(Jupytils)
else:
    import Jupytils
```

Mann Whitney test (also called the Mann-Whitney-Wilcoxon (MWW))

is a nonparametric test of the null hypothesis that two samples come from the same population against an alternative hypothesis, especially that a particular population tends to have larger values than the other.

- Non parametric test used when data comes from non-normal distribution
- Can be used with small samples

There are some situations when it is clear that the outcome does not follow a normal distribution. These include situations:

- when the outcome is an ordinal variable or a rank,
- when there are definite outliers or
- when the outcome has clear limits of detection

> **Use** : To compare a continuous outcome in two independent samples $group_1$ and $group_2$.

> **Null Hypothesis** : H0: Two populations are equal

> **Test Statistic** : The test statistic is U, the smaller of n_1, n_2 is the number of entries in group1 and group2

$$U = min(U_1, U_2)$$

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

where R_1 and R_2 are the sums of the ranks in $group_1$ and $group_2$, respectively.

Decision Rule: Reject H_0 if $U <$ critical value from table in favor of H_a the research hypothesis

References:

- http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Nonparametric/BS704_Nonparametric4.html (http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_Nonparametric/BS704_Nonparametric4.html)

-

####NOTE: If the sample size is at least 20, then one could use Z-values to test

(Reminder z for two tailed z test is 1.96 and for one tailed it is 1.65 I think)

If $Z_{calculated}$ is less than -1.96 or greater than +1.96, we reject the Null hypothesis"

Example

```
Following example is taken from you tube video
https://www.youtube.com/watch?v=hw3z49QoB1s

Data:
Data is a list of "Scores" obtained in an exam by two groups who were sressed and not-stre
ssed.
Question, is these a difference between these groups?

Y Stressed Y:    44    50    68    70    72    75    76    81    83    88    92    94
No Stress  N:    74    78    79    82    87    90    91    92    92    93

H0: There is no difference in scores between Stress and no-stress Groups
Ha: There is a difference

Test: 2 sided (because of some difference)
n1 = 12
n2 = 10
U critical (from table): 29
Result: if U is less than 29 we reject H0 in favor of Ha (i.e there is difference)

PDF: http://ocw.umb.edu/psychology/psych-270/other-materials/RelativeResourceManager.pdf

From the calculations below, we find:
Z = 1.84626532551 p = 0.129 (Note we multiply by 2 for two sided p-value)
U = 32.0 p = 0.0694  (Note we multiply by 2 for two sided p-value)

32 is not less than 29, therefore we fail to reject the H0.
(Also the p > 0.05, Z value is within -1.96 and +1.96 - all unabling to reject H_0)
i.e. There is no difference in scores between Stress and No-Stress groups
```

-

In [16]:

```
fileName="data/mann-whitney-test1.csv"

dfL  = LoadDataSet(fileName, columns=None);
displayDFs([dfL])
d1 = dfL.loc[dfL['Stress'] == 'N']['Score']
d2 = dfL.loc[dfL['Stress'] == 'Y']['Score']

z_stat1, p_val1 = stats.ranksums(d1, d2)
u, p_val2 = stats.mannwhitneyu(d1, d2,1)
print ('''
Mann Whitney fails with large values of P
''',d1.shape, d2.shape, "\nU-statistic: ",z_stat1, " P value: ", p_val1 , "\nMWW U stat: ", u, " P
```

22x2 var: DFF_PY_VAR_tableID_1579399727219

	Score	Stress
0	44	Y
1	50	Y
2	68	Y
3	70	Y
4	72	Y
5	74	N

<< >> Save ↻ Search>

```
Mann Whitney fails with large values of P
 (10,) (12,)
U-statistic:  1.8462653255082035  P value:  0.06485368983936193
MWW U stat:  32.0  P:  0.06947051710021464
```

```
##Another example
A physician is interested in the effect of an anaesthetic on reaction times. Two groups are
compared,
    * Group A taking anaesthetic
    * Group B without taking the anaesthetic.

Subjects had to react on a simple visual stimulus. Reaction times are not normally distributed
in this experiment, so data is analysed with the Mann-Whitney U-Test for ordinal scaled
measurements. The table below shows the rank-ordered data:

####Example taken from:
* Example From https://secure.brightstat.com/index.php?p=c&d=1&c=2&i=5
```

-

* Look at the results:
https://secure.brightstat.com/img/content/npartests/UTest/ex/Example_MWU.pdf

H0: There is no difference in reaction times in groups taking anaesthetic or not
Ha: There is a difference

Test: 1 sided (We want Anaesthetic group to be slower) at 5% confidence
n1 = 14
n2 = 12
U critical (from table): 51 (Fro two tailed it is 45)
Result: if U is less than 51 (less than 45) we reject H0 in favor of Ha (i.e there is difference)

PDF: http://ocw.umb.edu/psychology/psych-270/other-materials/RelativeResourceManager.pdf

From the calculations below, we find:
Z = -2.16 p = 0.0307 (Note we multiply by 2 for two sided p-value)
U = 42.0 p = 0.0163 (Note we multiply by 2 for two sided p-value)

42 is less than 51, therefore we reject the H0.
(Also the p < 0.05, Z value is outside of 1.65 (or for 2 tailed -1.96 and +1.96) - all reject H_0)
i.e. There is a difference The anaesthetic group shows significantly slower reaction times than the non-anaesthetic group

-

In [14]:

```
fileName="data/mann-whitney-test2.csv"

dfL  = LoadDataSet(fileName, columns=None, comment='#');
d2 = dfL.loc[dfL['Group'] == 'A']['Mean']
d1 = dfL.loc[dfL['Group'] == 'B']['Mean']


z_stat1, p_val1 = stats.ranksums(d1, d2)
u, p_val2 = stats.mannwhitneyu(d1, d2,1)


print( '''
Mann Whitney fails with large values of P
''',"n1: ", d1.shape, " n2:", d2.shape, "\nRank Sums: z: ",z_stat1, " p: ", p_val1 , "\nMann-Whitr

displayDFs(dfL);
```

```
Mann Whitney fails with large values of P
 n1:  (12,)  n2: (14,)
Rank Sums: z:  -2.1602468994692865  p:  0.03075356125927459
Mann-Whitneyt U:  42.0  p:  0.01632518745228646
```

26x2 var: DFF_PY_VAR_tableID_1579399714034

	Mean	Group
0	131	B
1	135	A
2	138	B
3	138	B
4	139	A
5	141	B

<< >> Save ↻ Search>

-

In [5]:

```
# Example from: https://www.youtube.com/watch?v=nRAAAp1Bgnw
#
#
s1= [28,31,36,35,32,33,12,18,19,14,20,19]
s2= "a,a,a,a,a,a,b,b,b,b,b,b".split(",");
dfL = pd.DataFrame( {"Data":s1, "Group":s2})

displayDFs(dfL)
d2 = dfL.loc[dfL['Group'] == 'a']['Data']
d1 = dfL.loc[dfL['Group'] == 'b']['Data']


z_stat1, p_val1 = stats.ranksums(d1, d2)
u, p_val2 = stats.mannwhitneyu(d1, d2,1)
print ('''
Mann Whitney fails with large values of P
''',d1.shape, d2.shape, "\n",z_stat1, p_val1*2 , "\n", u, p_val2 * 2)
```

12x2 var: DFF_PY_VAR_tableID_1579399452960

	Data	Group
0	28	a
1	31	a
2	36	a
3	35	a
4	32	a
5	33	a

<< >> Save ↻ Search>

```
Mann Whitney fails with large values of P
 (6,) (6,)
 -2.8823067684915684 0.007895503713806915
 0.0 0.004998124765082452
```

In [ ]:

```

```

-