Exercise 1

Three sword fighters are comparing their swords. Obi Wan Kenobi claims that light sabers are the best, Beatrix Kiddo swears by Hattori Hanzo katanas and Frodo Baggins prefers Elvish daggers.



Exercise 1

- Five different people are randomly assigned to each weapon (15 people all together) and sent out to battle.
- The number of kills they achieve are listed in the table to the right.
- State the null and alternative hypotheses and perform an ANOVA on the data.
- What are your conclusions?
- If you reject the null hypothesis, then conduct post-hoc tests (Scheffé) to determine the best sword.

Light Saber	Katana	Dagger
6	6	0
8	5	4
5	9	0
4	4	1
2	6	0
T = 25	T = 30	T = 5

ANOVA

Step 1: State the hypothesis & specify the alpha level

Step 2: Locate the critical region

Step 3: Perform the analysis

1. SS (sum of square)
2. MS (mean of square)
3. F-ratio

Step 1: State the hypotheses, and specify the alpha level

- The null hypothesis states that there is no difference among the swords in terms of number of dismemberments/kills committed:
    - H_0: $\mu_{Light\ Saber} = \mu_{Katana} = \mu_{Dagger}$ (the type of sword used has no effect).

- There are a number of possible statements for the alternative hypothesis. The general alternative hypothesis is:
    - H_1: At least one of the mean numbers of dismemberments/kills is different.

  That is, the type of sword has an effect on the number of dismemberments/kills.

- For this test, we'll use $\alpha = 0.05$.

Step 2: Locate the critical region

- $df_{between} = k - 1 = 3 - 1 = 2$
- $df_{within} = N - k = 15 - 3 = 12$

- $F(2,12) = 3.8853$

- The obtained F ratio must exceed this value to reject H0.

F - Distribution ($\alpha = 0.05$ in the Right Tail)

Denominator Degrees of Freedom (df_2) / Numerator Degrees of Freedom (df_1)

df_2 \ df_1	1	2	3	4	5	6	7	8	9
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371	19.385
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	9.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	6.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876
16	4.4940	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377
17	4.4513	3.5915	3.1968	2.9647	2.8100	2.6987	2.6143	2.5480	2.4943
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.3660
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419
23	4.2793	3.4221	3.0280	2.7955	2.6400	2.5277	2.4422	2.3748	2.3201
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002
25	4.2417	3.3852	2.9912	2.7587	2.6030	2.4904	2.4047	2.3371	2.2821
26	4.2252	3.3690	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655
27	4.2100	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501
28	4.1960	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.2360
29	4.1830	3.3277	2.9340	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588
∞	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799

https://web.stanford.edu/dept/radiology/cgi-bin/classes/stats_data_analysis/lesson_5/234_8_l.html

Step 3: Perform analysis

0. Mean and SS (sum of squares)

Light Saber	Katana	Dagger
6	6	0
8	5	4
5	9	0
4	4	1
2	6	0
T = 25	T = 30	T = 5

- Total number of observations: $N = N_1 + N_2 + N_3 = 5 + 5 + 5 = 15$
- Grand total: $T = T_1 + T_2 + T_3 = 25 + 30 + 5 = 60$
- Number of groups: $k = 3$
- $df_{between} = k - 1 = 3 - 1 = 2$
- $df_{within} = N - k = 15 - 3 = 12$

	Light Saber	Katana	Dagger
Mean	$\frac{6+8+5+4+2}{5} = 5$	$\frac{6+5+9+4+6}{5} = 6$	$\frac{0+4+0+1+0}{5} = 1$
SS	$(6-5)^2 + (8-5)^2 + (5-5)^2 + (4-5)^2 + (2-5)^2 = 20$	$(6-6)^2 + (5-6)^2 + (9-6)^2 + (4-6)^2 + (6-6)^2 = 14$	$(0-1)^2 + (4-1)^2 + (0-1)^2 + (1-1)^2 + (0-1)^2 = 12$

Step 3.1: Mean Squares between swords

$$SS_{between} = \sum_{group} \frac{T_{group}^2}{N_{group}} - \frac{T^2}{N}$$

$$(or) = \sum_{group} N_{group} * \left(\overline{x_{group}} - \bar{x}\right)^2$$

- $SS_{between} = \frac{25^2}{5} + \frac{30^2}{5} + \frac{15^2}{5} - \frac{60^2}{15} = 125 + 180 + 5 - 240 = 70$
- $MS_{between} = \frac{SS_{between}}{df_{between}} = \frac{70}{2} = 35$

Step 3.2: Mean Squares within swords

$$SS_{within} = \sum_{group} SS_{group}$$

$$(or) = \sum_{group\ (g)} \sum_{observation\ (o)} \left(x_{g,o} - \overline{x_g}\right)^2$$

- $SS_{within} = SS_{Light\ Saber} + SS_{Katana} + SS_{Dagger} = 20 + 14 + 12 = 46$
- $MS_{within} = \frac{SS_{within}}{df_{within}} = \frac{46}{12} = 3.8\overline{3}$

Step 4: compute F

$$F = \frac{MS_{between}}{MS_{within}} = \frac{35}{3.8\overline{3}} = 9.1304$$

Step 4: Make a decision about H0, and state a conclusion

The obtained F of 9.1304 exceeds the critical value of 3.8853. Therefore, we can **reject the null hypothesis H0**.

The following table summarizes the results of the analysis:

Source	SS	df	MS	F
Between	70	2	35	**9.13**
Within	46	12	3.8333	
Total	116	14		

Post-hoc tests (Scheffé)

Just need to compute between treatment Mean Square (MS)

1. **$SS_{between}$**
2. **$Ms_{between}$**
3. **F - ratio**

Step 5.1: post-hoc tests (Scheffé)

- Light Saber VS Katana (1 VS 2)

- $T_{1+2} = T_1 + T_2 = 25 + 30 = 55$
- $N_{1+2} = N_1 + N_2 = 5 + 5 = 10$

- $SS_{between1+2} = \frac{25^2}{5} + \frac{30^2}{5} - \frac{55^2}{10} = 2.5$
- $MS_{between1+2} = \frac{2.5}{2} = 1.25$
- $F_{1+2} = \frac{1.25}{3.8\overline{3}} = 0.33$
- No difference between Light Saber and Katana

Step 5.2: post-hoc tests (Scheffé)

- Light Saber VS Dagger (1 VS 3)

- $T_{1+3} = T_1 + T_3 = 25 + 5 = 30$
- $N_{1+3} = N_1 + N_3 = 5 + 5 = 10$

- $SS_{between1+3} = \frac{25^2}{5} + \frac{5^2}{5} - \frac{30^2}{10} = 40$
- $MS_{between1+3} = \frac{40}{2} = 20$
- $F_{1+3} = \frac{20}{3.8\overline{3}} = 5.22$
- Significant difference between Light Saber and Dagger!

Step 5.3: post-hoc tests (Scheffé)

- Katana VS Dagger (2 VS 3)

- $T_{2+3} = T_2 + T_3 = 30 + 5 = 35$
- $N_{2+3} = N_2 + N_3 = 5 + 5 = 10$

- $SS_{between2+3} = \frac{30^2}{5} + \frac{5^2}{5} - \frac{35^2}{10} = 62.5$
- $MS_{between2+3} = \frac{62.5}{2} = 31.25$
- $F_{2+3} = \frac{31.25}{3.8\overline{3}} = 8.15$
- Significant difference between Katana and Dagger!

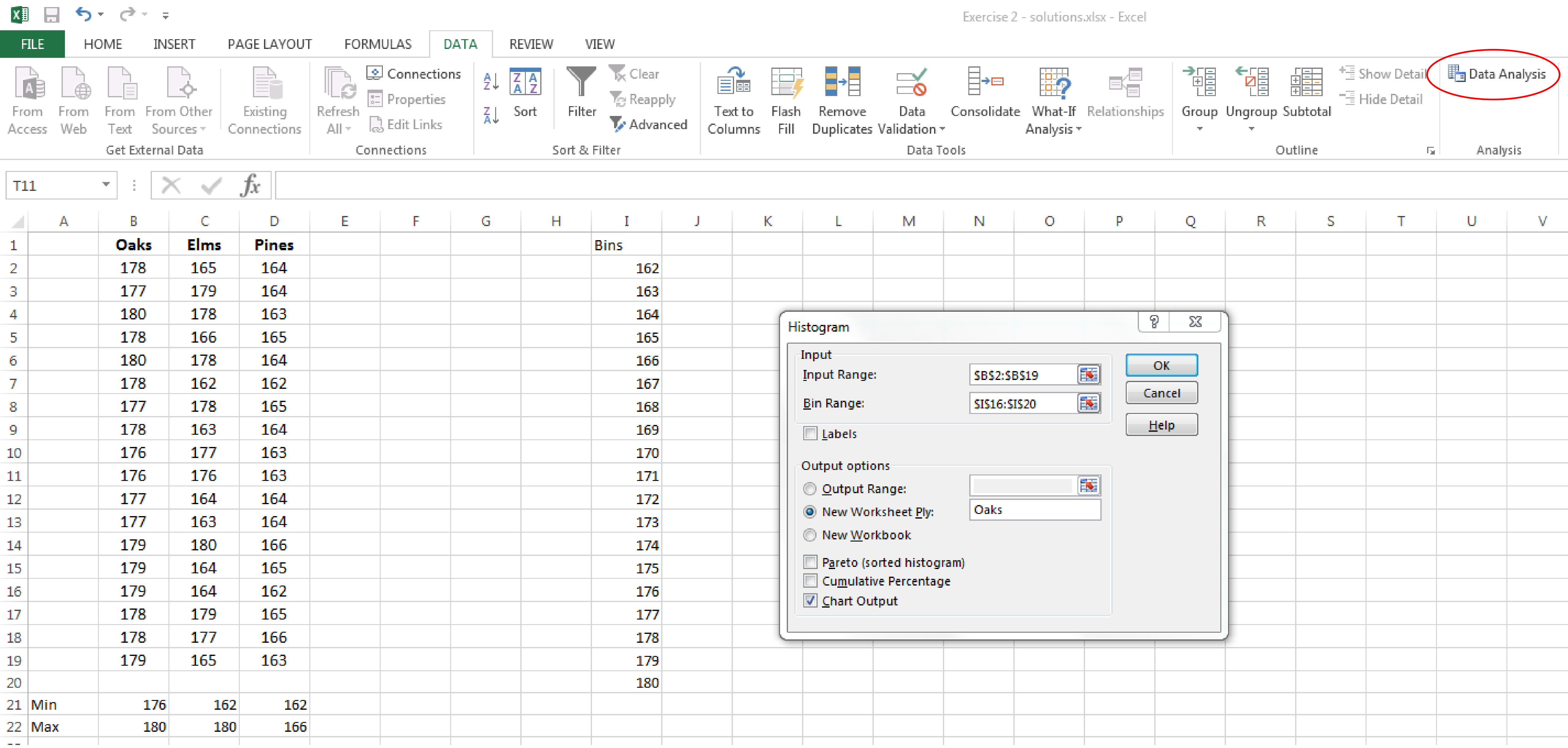
Exercise 2

- The heights of Oaks, Elms and Pines are given in the table.
- Plot histograms of each groups' heights.
- If we wanted to compare these groups what assumption(s) of the ANOVA does/do these data violate?

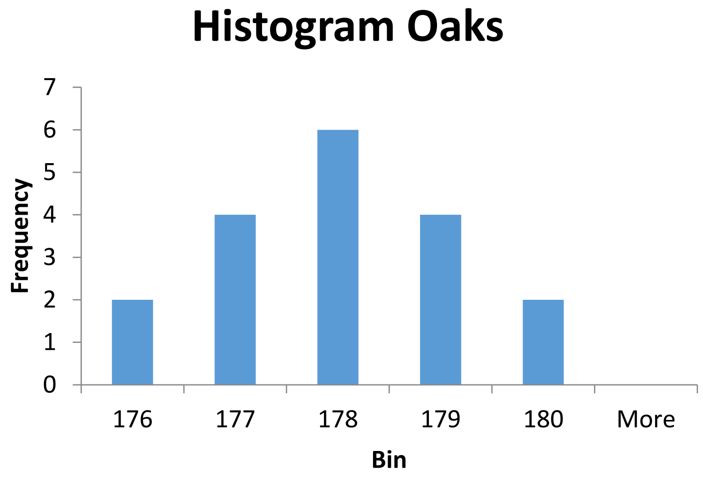
Oaks	Elms	Pines
178	165	164
177	179	164
180	178	163
178	166	165
180	178	164
178	162	162
177	178	165
178	163	164
176	177	163
176	176	163
177	164	164
177	163	164
179	180	166
179	164	165
179	164	162
178	179	165
178	177	166
179	165	163

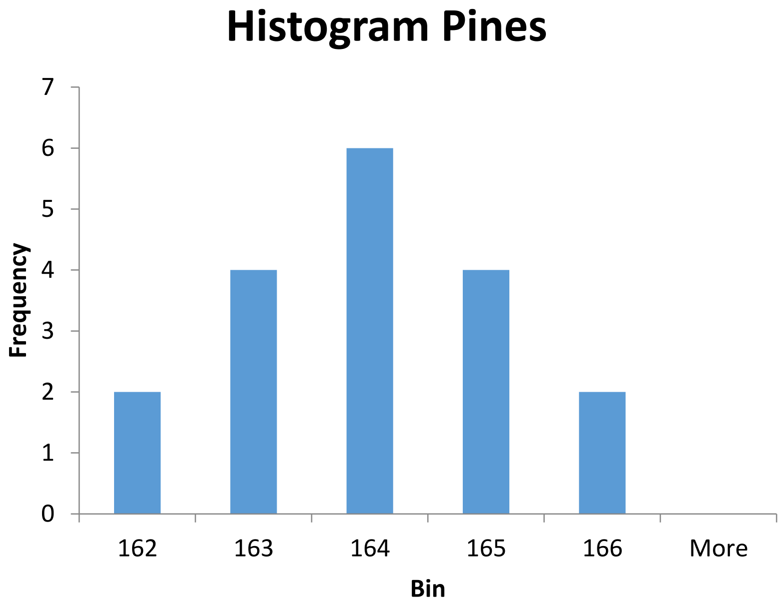
Assumptions of ANOVA

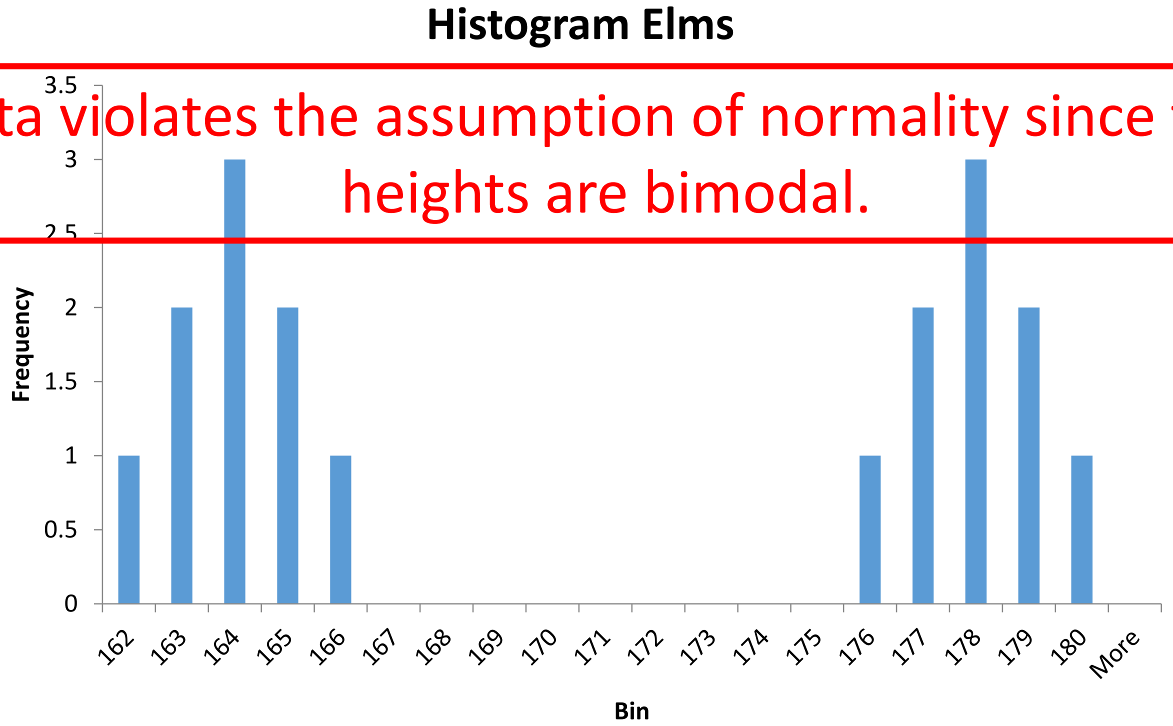
1. Independent samples
   (samples to be obtained independently from all other samples)
2. Homogeneity of variance
3. Normally distributed populations



	A	B	C	D	I
1		Oaks	Elms	Pines	Bins
2		178	165	164	162
3		177	179	164	163
4		180	178	163	164
5		178	166	165	165
6		180	178	164	166
7		178	162	162	167
8		177	178	165	168
9		178	163	164	169
10		176	177	163	170
11		176	176	163	171
12		177	164	164	172
13		177	163	164	173
14		179	180	166	174
15		179	164	165	175
16		179	164	162	176
17		178	179	165	177
18		178	177	166	178
19		179	165	163	179
20					180
21	Min	176	162	162	
22	Max	180	180	166	







The data violates the assumption of normality since the Elms heights are bimodal.

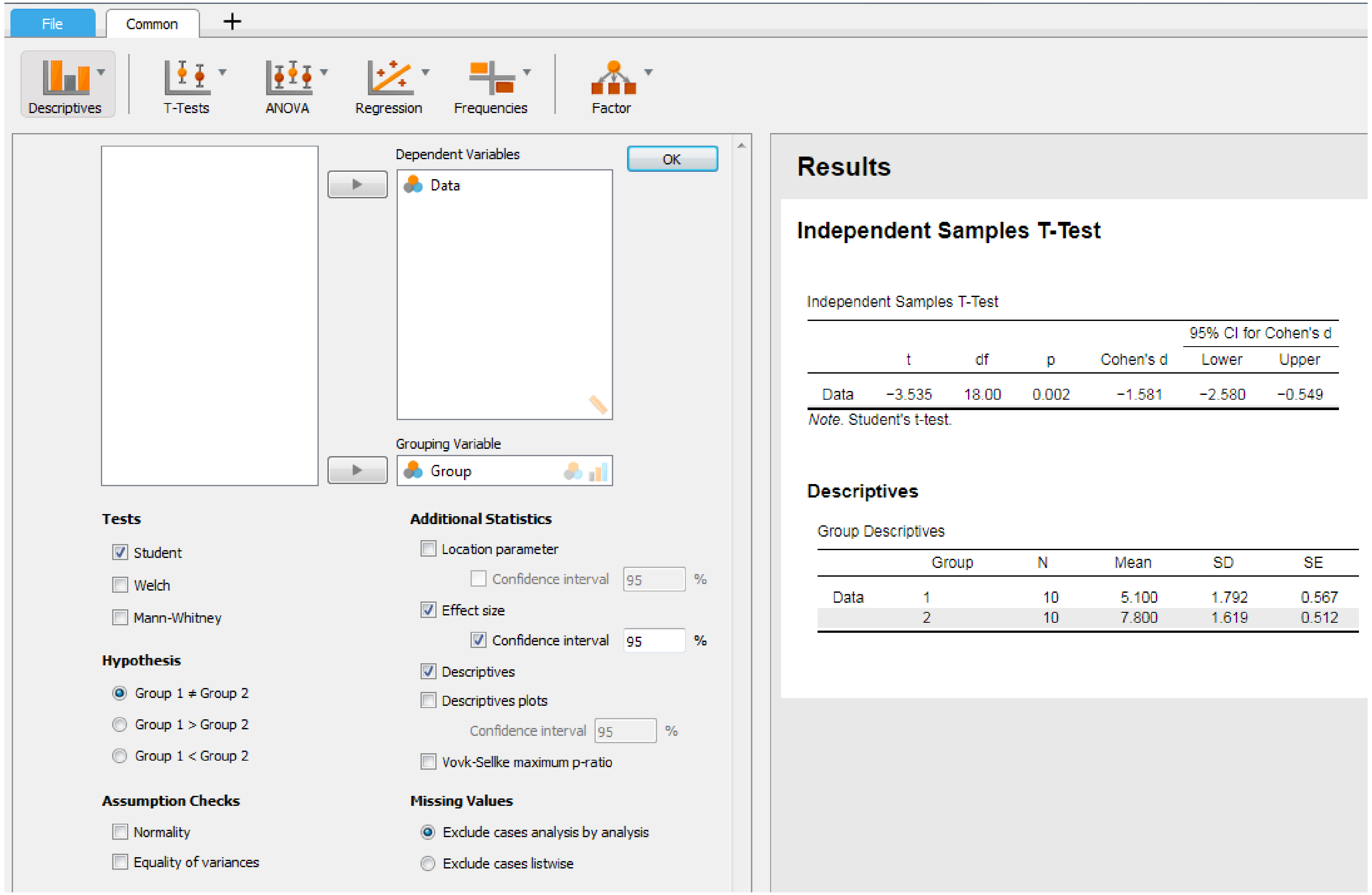
Exercise 3

- Given the data on the right, use your favourite software to compute both an independent measures t-test and an independent measures ANOVA and verify that $F = t^2$.

Group 1	Group 2
5	9
4	7
7	7
1	7
6	5
7	8
6	8
6	9
5	11
4	7

JASP – independent t-test

	Group	Data
1	1	5
2	1	4
3	1	7
4	1	1
5	1	6
6	1	7
7	1	6
8	1	6
9	1	5
10	1	4
11	2	9
12	2	7
13	2	7
14	2	7
15	2	5
16	2	8
17	2	8
18	2	9
19	2	11
20	2	7



JASP – F-test

$$F = t^2$$

$$12.50 = (-3.535)^2$$