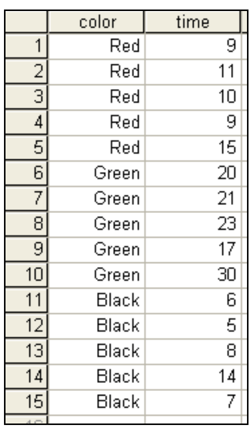
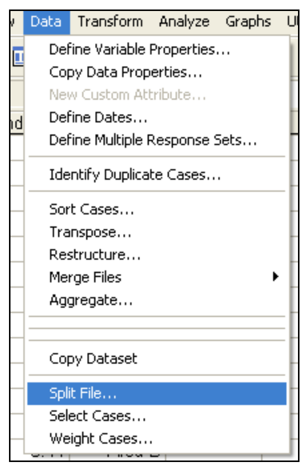
ANOVA and Multiple Comparisons in SPSS **STAT 314**

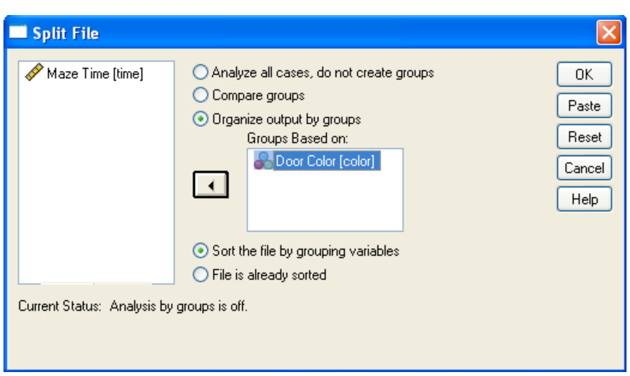
Three sets of five mice were randomly selected to be placed in a standard maze but with different color doors. The response is the time required to complete the maze as seen below. Perform the appropriate analysis to test if there is an effect due to door color. (Use $\alpha = 0.01$)

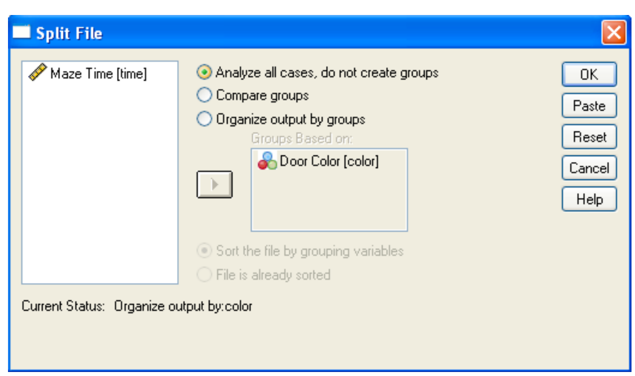
Color	**Time**				
Red	9	11	10	9	15
Green	20	21	23	17	30
Black	6	5	8	14	7

1. Enter the group number (1 for Red, 2 for Green, 3 for Black) into one variable and the corresponding time values into another variable *(see upper-right figure, below)*. Be sure to code your variables appropriately. Now it is time to check the normality assumption. Select "Split File" from the "Data" menu so that we can tell SPSS that we want separate Q–Q Plots for each group *(see upper-right figure, below)*. Select "Organize output by groups" and enter "color" as the variable that groups are based upon *(see lower-left figure, below)*. Now create Normal Q–Q Plots to assess the normality of each group *(see separate handout on Normal Q–Q Plots)*. Once you've created your Q–Q Plots and determined that your groups are approximately normally distributed, select "Split File" from the "Data" menu and then select "Analyze all cases, do not create groups" in order to return SPSS to its normal data analysis mode *(see lower-right figure, below)*.

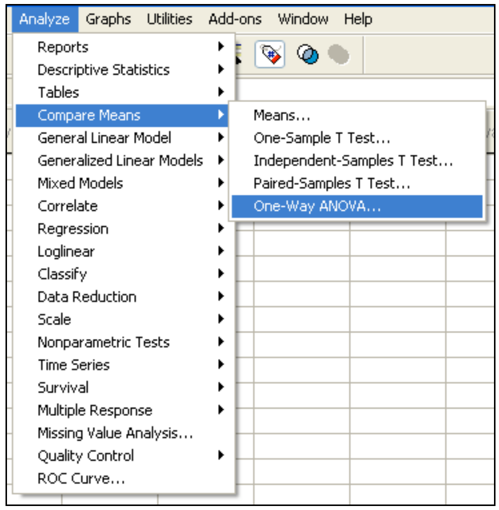




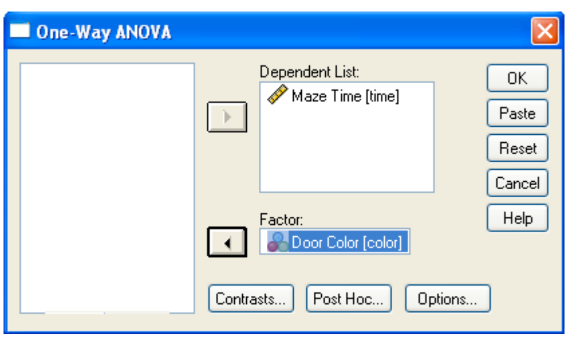


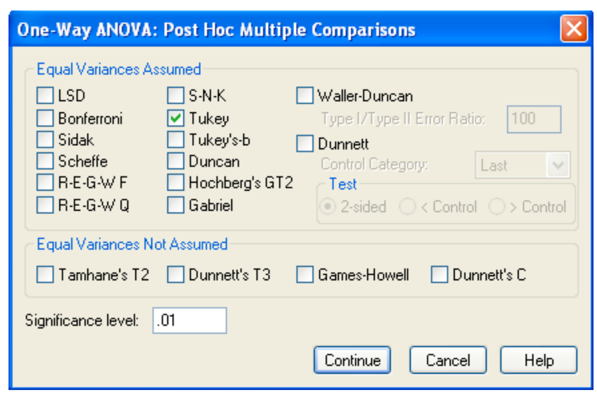


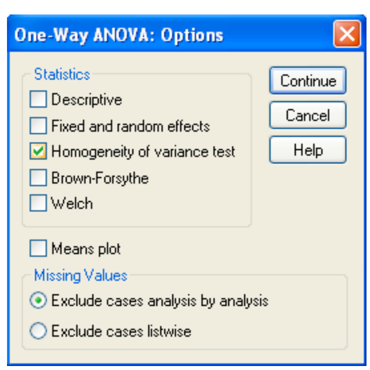
2. Select Analyze → Compare Means → One-Way ANOVA… *(see figure, below)*.



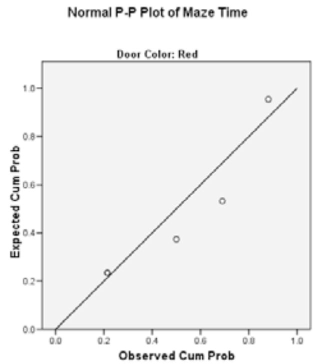
3. Select "Maze Time" as the dependent test variable, select "Door Color" as the grouping factor, and click "Post Hoc…". Select the "Tukey" procedure, enter 0.01 for the significance level (99% CI corresponds to a 1% (0.01) significance level) to obtain the 99% Tukey-Kramer multiple comparison confidence intervals and grouping diagram, and click "Continue". Click the "Options…" button and select "Homogeneity-of-Variance" (Levene's Test checks the assumption of equal variances). Click "Continue" to close the options and then click "OK" (*see the 3 figures, below*).

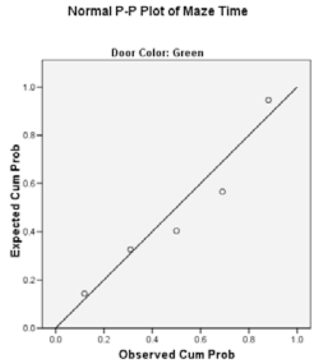


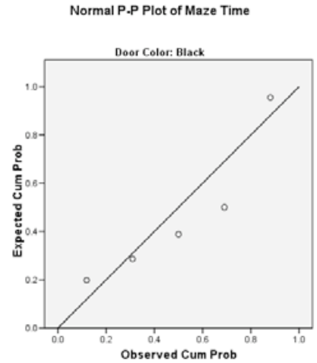




4. Your output should look like this.







Oneway

Test of Homogeneity of Variances

Maze Time

Levene Statistic	df1	df2	Sig.
.6522	2	12	.5384

ANOVA

Maze Time

	df	Sum of Squares	Mean Square	F	Sig.
Between Groups	2	565.7333	282.8667	20.0142	.0002
Within Groups	12	169.6000	14.1333		
Total	14	735.3333			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: Maze Time
Tukey HSD

(I) Door Color	(J) Door Color	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval	
					Lower Bound	Upper Bound
Red	Green	-11.4000*	2.3777	.0012	-19.8836	-2.9164
	Black	2.8000	2.3777	.4879	-5.6836	11.2836
Green	Red	11.4000*	2.3777	.0012	2.9164	19.8836
	Black	14.2000*	2.3777	.0002	5.7164	22.6836
Black	Red	-2.8000	2.3777	.4879	-11.2836	5.6836
	Green	-14.2000*	2.3777	.0002	-22.6836	-5.7164

*. The mean difference is significant at the .01 level.

Homogeneous Subsets

Maze Time

Tukey HSD[a]

Door Color	N	Subset for alpha = .01	
		1	2
Black	5	8.0000	
Red	5	10.8000	
Green	5		22.2000
Sig.		.4879	1.0000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 5.000.

5. You should use the output information in the following manner to answer the question.

Step 0: **Check Assumptions**

Since the points of each Q-Q Plot lie close to their respective diagonal lines, we conclude that each of the data groups is from an approximately normally distributed population.
Also, the Levene Statistic *p*-value = Sig. = 0.5384 is greater than $\alpha = 0.01$ (from Step 2), so we fail to reject the null hypothesis that the variances are all equal. Since the variances appear to be equal (and we have random/independent samples), we may continue with ANOVA.

Test of Homogeneity of Variances

Maze Time

Levene Statistic	df1	df2	Sig.
.6522	2	12	.5384

Step 1: **Hypotheses**

$H_0: \mu_{Red} = \mu_{Green} = \mu_{Black}$

$H_a:$ *at least one* μ_i *is different*

Step 2: **Significance Level**

$\alpha = 0.01$

Step 3: **Rejection Region**

Reject the null hypothesis if *p*-value ≤ 0.01.

Step 4: **Construct the One-way ANOVA Table**

ANOVA

Maze Time

	df	Sum of Squares	Mean Square	F	Sig.
Between Groups	2	565.7333	282.8667	20.0142	.0002
Within Groups	12	169.6000	14.1333		
Total	14	735.3333			

From the output, $F = 20.0142$ with 2 and 12 degrees of freedom.
p-value = Sig. = 0.0002

Step 5: **Conclusion**

Since *p*-value = 0.0002 $\le$ 0.01 = α, we shall reject the null hypothesis.

Step 6: **State conclusion in words**

At the $\alpha = 0.01$ level of significance, there exists enough evidence to conclude that there is a difference in the mean times to complete the maze based on door color (i.e., there is an effect due to door color).

6. Since we rejected the null hypothesis (we found differences in the means), we should perform a Tukey-Kramer (Tukey's W) multiple comparison analysis to determine which means are similar and which means are different. Here is how such an analysis might appear.

Multiple Comparisons

Dependent Variable: Maze Time
Tukey HSD

(I) Door Color	(J) Door Color	Mean Difference (I-J)	Std. Error	Sig.	99% Confidence Interval Lower Bound	99% Confidence Interval Upper Bound	
Red	Green	-11.4000*	2.3777	.0012	-19.8836	-2.9164	Different (interval does not contain zero)
	Black	2.8000	2.3777	.4879	-5.6836	11.2836	Similar (interval contains zero)
Green	Red	11.4000*	2.3777	.0012	2.9164	19.8836	
	Black	14.2000*	2.3777	.0002	5.7164	22.6836	Different (interval does not contain zero)
Black	Red	-2.8000	2.3777	.4879	-11.2836	5.6836	
	Green	-14.2000*	2.3777	.0002	-22.6836	-5.7164	

*. The mean difference is significant at the .01 level.

Thus, we are 99% confident that mazes with green doors seem to take longer on average to complete than do mazes with red or black doors (which have similar population mean times to completion).

Maze Time

Tukey HSD[a]

Door Color	N	Subset for alpha = .01: 1	2
Black	5	8.0000	
Red	5	10.8000	
Green	5		22.2000
Sig.		.4879	1.0000

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 5.000.

This table corresponds to our diagram. Note that the black and red sample means (8.0000 & 10.8000) are grouped together (separately from the differing green sample mean (22.2000)). This shows that we are 99% confident that the black and red population means are similar, yet both differ from the green population mean (which agrees with the conclusion based on the simultaneous confidence intervals).