Lab 2 Example – Stat 252 – One way ANOVA

NHL Individual Player **Plus/Minus ratings**, recorded at the three-quarter mark (about 60 games into) the 2008-2009 NHL Regular Season. Random players are selected from five randomly chosen teams. Clearly, not all players from every team have been chosen, although the random selections seem to be made only from a list of players who have played a significant number of games with their team (wouldn't made sense to choose a player who only played 5 games).

Categorical variables (defining groups) are teams:

1 - Anaheim Ducks
2 - Columbus Blue Jackets
3 - New York Rangers
4 - Phoenix Coyotes
5 - Toronto Maple Leafs

GENERAL DESCRIPTIVES (get the output for yourself)

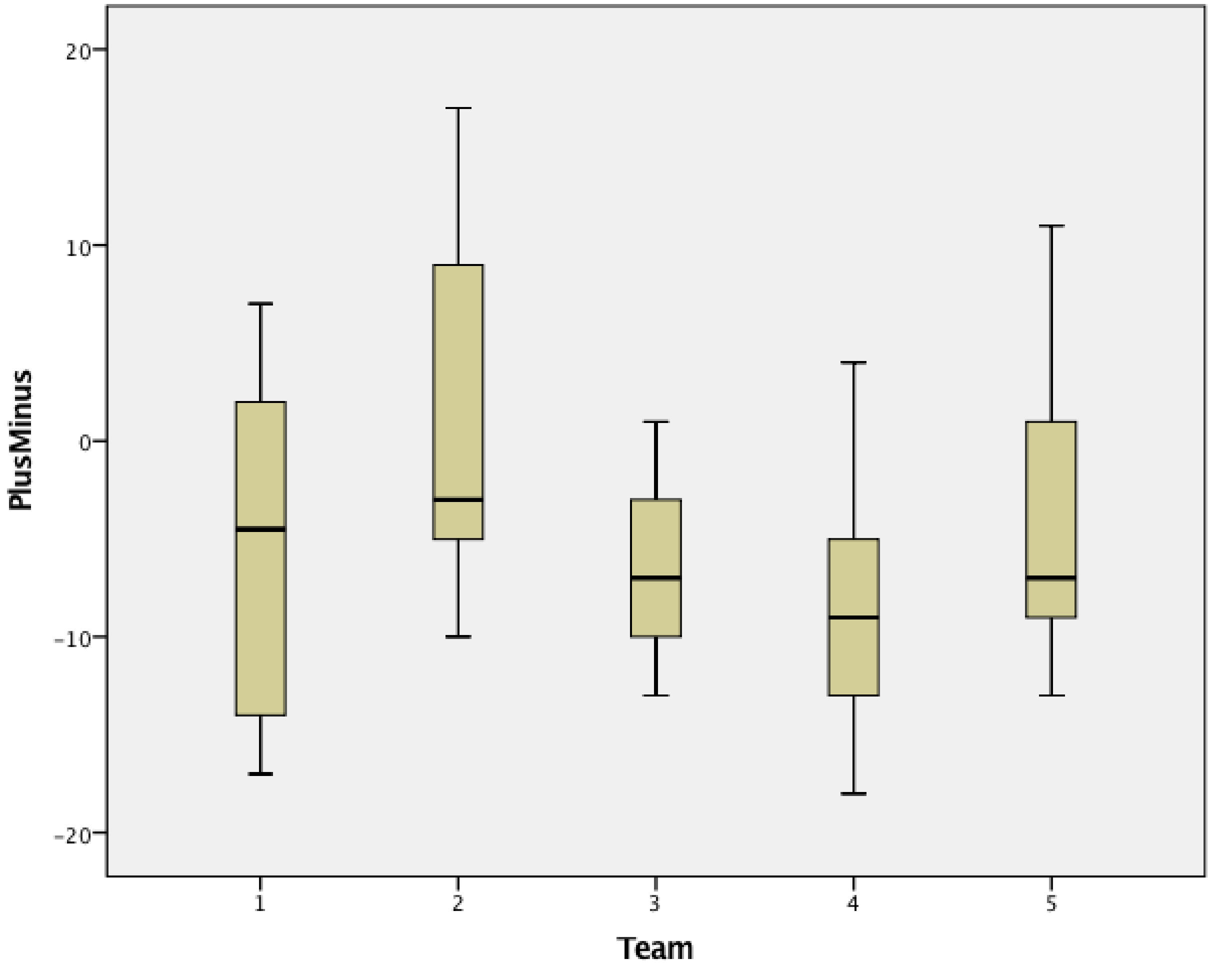
The average plus/minus varies from the lowest of -7.85 to the highest of 1.62. The mean plus/minus for the Columbus Blue Jackets is the highest, while the Phoenix Coyotes have the lowest mean plus/minus, with the New York Rangers mean being nearly as bad (low). The Anaheim Ducks and Toronto Maple Leafs' mean plus/minus values lie between the two "extremes" described above. From this point on, I will refer to teams as "team 1" and not by the city.

The standard deviations vary from the lowest of 4.64 to the highest of 8.75. Given the moderately low sample sizes, these are relatively small differences in the magnitude of standard deviations. The ratio above is 8.75/4.64 which is approximately 1.9, and is relatively high but obviously less than 2.

Note regarding ANOVA and variances

Because ANOVA procedures are not extremely sensitive to unequal standard deviations, it is not necessary to carry out a formal test of equality of standard deviations as a preliminary to the ANOVA. Instead, the spread of observations in the boxplots can be examined and the following rule of thumb may be used: **If the ratio of the largest sample standard deviation to the smallest sample standard deviation is less than 2, the assumption of equal standard deviations is plausible.**

BOXPLOTS



The above side-by-side boxplot confirms and builds on our knowledge obtained from the descriptive stats. Teams 1 and 2 appear to have higher plus/minus values while teams 3, 4, and 5 appear to have considerably lower plus/minus ratings. All five distributions exhibit approximately similar spread, although you can say that teams 1 and 2 exhibit greater spread in their distributions. The distributions are also approximately symmetric, with team 2 having somewhat of a right skewed distribution. There are no outliers. It is important to realize that small sample sizes do not allow us to make strong claims about (lack of) normality for the data.

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
PlusMinus	Based on Mean	2.413	4	59	.059
	Based on Median	1.229	4	59	.309
	Based on Median and with adjusted df	1.229	4	45.797	.312
	Based on trimmed mean	2.357	4	59	.064

For this data, Levene's test is not significant (the p-value in each of the four versions of the test is much larger than the threshold value of 0.05), indicating the null hypothesis cannot be rejected, so the data supports the assumption of equal variances.

QQ PLOTS (obtain them yourself)

The normality plots of the six distributions displayed indicate some slight departures from the normality assumption in some of the six distributions. There is no evidence of any serious departure from the normality assumption. Again, it is crucial to emphasize that small sample sizes do not allow us to make strong claims about normality for the data.

ANOVA TEST

In order to see whether there is any evidence that some teams have players with greater plus/minus, we apply *One-Way ANOVA* feature in SPSS. Denote by Ui the mean plus minus strength of the i-th team, where i=1,2, 3,...5.

Define the null and alternative hypotheses as follows: H_0 : U1=U2=……..=U5 (there is no difference in the mean plus/minus among the five teams)

H_a : there are differences in the plus/minus between the five teams (at least two means differ).

ANOVA

PlusMinus

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	704.752	4	176.188	3.375	.015
Within Groups	3079.686	59	52.198		
Total	3784.438	63			

The sum of squared residuals fitting from the reduced model (equal means) is 3784.438. The sum of square residuals from fitting the full (separate means) model is 3079.686.

The pooled estimate of the variance is 52.198. The F test stat is 3.375 with a P-value of 0.015< alpha=0.05. We conclude at the 5% significance level that some teams have better plus/minus ratings than others (the conclusion on your assignment may or may not be more exciting). Note that the

ANOVA F-test alone doesn't tell us exactly which teams have better/worse mean plus/minus ratings, specifically.

Calculation example of F-test stat (refer to output)

Extra sum of squares = residual sum of squares (reduced) – residual sum of squares (full)

= 704.75

Formula for F-statistic = $\dfrac{\text{(Extra sum of squares) / (extra degrees of freedom)}}{\text{(Full variance)}}$

$$= \frac{(704.75) / (4)}{52.198} = \mathbf{3.375}$$

TUKEY/SCHEFFE (see the long output)

Out of 10 confidence intervals, 8 contain zero and 2 do not contain zero. Equivalently, in 8 cases the test failed to establish a difference between the means. More precisely, there are significant differences between the teams 2 and 3, and teams 2 and 4. (note the low p-value and the confidence interval NOT including zero)

CONTRASTS

For example, we want to compare the **teams 1 and 2**, with **teams 3, 4, and 5** the contrast of interest is

$(U1+U2)/2 - (U3+U4+U5)/3 = 0$

The LCM of 2 and 3 is 6, so multiply the equation by 6 to get contrast coefficients 3, 3, -2, -2, -2.

Contrast Coefficients

	Team				
Contrast	1	2	3	4	5
1	3	3	-2	-2	-2

Contrast Tests							
		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
PlusMinus	Assume equal variances	1	27.02	11.112	2.432	59	.018
	Does not assume equal variances	1	27.02	11.997	2.252	36.700	.030

The p-value of the contrast (two-sided alternative) is 0.018 (equal variances assumed) or 0.030 (equal variances not assumed). At a=0.05, there is sufficient evidence to claim that teams 1 and 2 differ from teams 3, 4, and 5 in terms of plus/minus rating.

LOG TRANSFORM, THEN ANOVA

Since we have negative values for plus/minus, we cannot apply the log-transform to this variable. For your assignment, comment on the spread of the distributions after the log-transform (from boxplots) and comment on how the F-stat changed in ANOVA, as well as the P-value. Of course, do the ANOVA test in the same manner but in terms of mean log values (median values on original scale). See if the conclusion is the same.