Scientific Letter

Suggested Statistical Methods to Analyze Air Power Operations Course Surveys

Background

The RCAF has conducted two serials of the Air Power Operations Course (APOC). Students are administered surveys at the beginning of the course and marks are obtained during the course. The Lessons Learned Branch of CFAWC was asked to conduct analysis of the results; they requested the Operational Research and Analysis Branch's assistance. The purpose of this letter is to describe the statistical method(s) applied to each question so that they can be repeated as further data are collected, as well as provide the results from the analysis of the first two serials of APOC.

Statement of results

The questions and the statistical tests applied/suggested were (see Annex A for results):

- **Did the completion of Air Force Officer Development (AFOD) 5 result in higher marks on APOC?**
    - A box plot can be used to visibly represent the assessed average versus the completion (Yes, No, N/A of AFOD 5)
    - An Independent two-sample t-test for the equality of means was applied.
- **Did the completion of AFOD 5 result in higher marks on the Service Paper?**
    - The Service Paper was assessed as Below Standard (BS), Met Standard With Difficulty (MSWD), Met Standard (MS) and Exceeded Standard (ES). The count of each of these assessments was plotted against the completion of AFOD (Yes, No, N/A).
    - A Mann-Whitney Test was applied to determine if there was a statistical difference in the assessments of those that completed AFOD 5 and those that did not.
- **Did the completion of AFOD 5 result in higher marks in Staff Duties?**
    - Staff Duties was assessed as Below Standard (BS), Met Standard With Difficulty (MSWD), Met Standard (MS) and Exceeded Standard (ES). The count of each of these assessments was plotted against the completion of AFOD (Yes, No, N/A).
    - A Mann-Whitney Test was applied to determine if there was a statistical difference in the assessments of those that completed AFOD 5 and those that did not.
- **Did more years of experience result in higher marks on APOC?**
    - An XY plot was created of Assessed Average versus student's years of service.

- Years of service was grouped/binned into those with 0-11 years of Service and 12+ years of service. (N.B. Years of service was captured for only one of the serials of APOC. A third bin of "No Data" was created.) A bar graph depicting the numbers in each of these bins was created.
- A box plot can be used to visibly represent the assessed average versus the years of service bins (0-11, 12+, No Data).
- A Mann-Whitney Test was applied to determine if there was a statistical difference in the course average versus the years of service bins (0-11, 12+)

- **Do Some Military Occupations do better than others?**
  - A bar graph showing the number of personnel in each Military Occupation Structure Identification (MOSID) was created.
  - A box plot can be used to visibly represent the assessed average versus the MOSID.
  - A Kruskal-Wallis test was used to determine if the mean assessed mark was different between the MOSIDs.

- **Self-Assessment of staffing ability versus mark obtained in staff planning.**
  - Students were asked in a survey at the beginning of the course to assess their own staff planning abilities (Poor, Fair, Good, and Very Good).
  - Routine Staffing was assessed during the course as Below Standard (BS), Met Standard With Difficulty (MSWD), Met Standard (MS) and Exceeded Standard (ES). A bar graph showing the count of the Routine staffing assessments versus the student's self-assessment was created.

Discussion of results

The results discussed below are only for APOC 1601 and 1701. As more data is added the results may change.

Did the completion of Air Force Officer Development (AFOD) 5 result in higher marks on APOC?

The statistical test indicated that there is no statistical significance between the means of marks of those students who completed AFOD 5 and those that did not. Completion of AFOD *has not* been an indicator of higher marks on APOC.

Did the completion of AFOD 5 result in higher marks on the Service Paper?

The statistical test indicated that there is no statistical significance between the service paper results of those students who completed AFOD 5 and those that did not. Completion of AFOD *has not* been an indicator of higher marks on the service paper.

Did the completion of AFOD 5 result in higher marks in Staff Duties?

The statistical test indicated that there is no statistical significance between the staff duties results of those students who completed AFOD 5 and those that did not. Completion of AFOD *has not* been an indicator of higher marks on the service paper.

Did more years of experience result in higher marks on APOC?

The statistical test indicated that there is no statistical significance between the means of marks of those students who had less than 12 years of service and those that had 12 or more. Years of service *has not* been an indicator of higher marks on APOC.

Do Some Military Occupations do better than others?

The statistical test indicated that there is no statistical significance between the means of marks of students from different MOSIDs. MOSID *has not* been an indicator of higher marks on APOC.

Self-Assessment of staffing ability versus mark obtained in staff planning

The student's self-assessments on their staff planning ability had no relationship with their assessed results. The results indicate that some students underestimated while others overestimated their abilities; no particular trend was detected.

Conclusion

The results reported on in this letter are only from APOC 1601 and APOC 1701. It is recommended that more data be collected in order to develop a more complete understanding of the statistical significance of the results.

Prepared by: Bruce, J, Chapman. OR&A Branch, CFAWC (DRDC – Centre for Operational Research and Analysis).

Annex A   Annex A: Results of Analysis

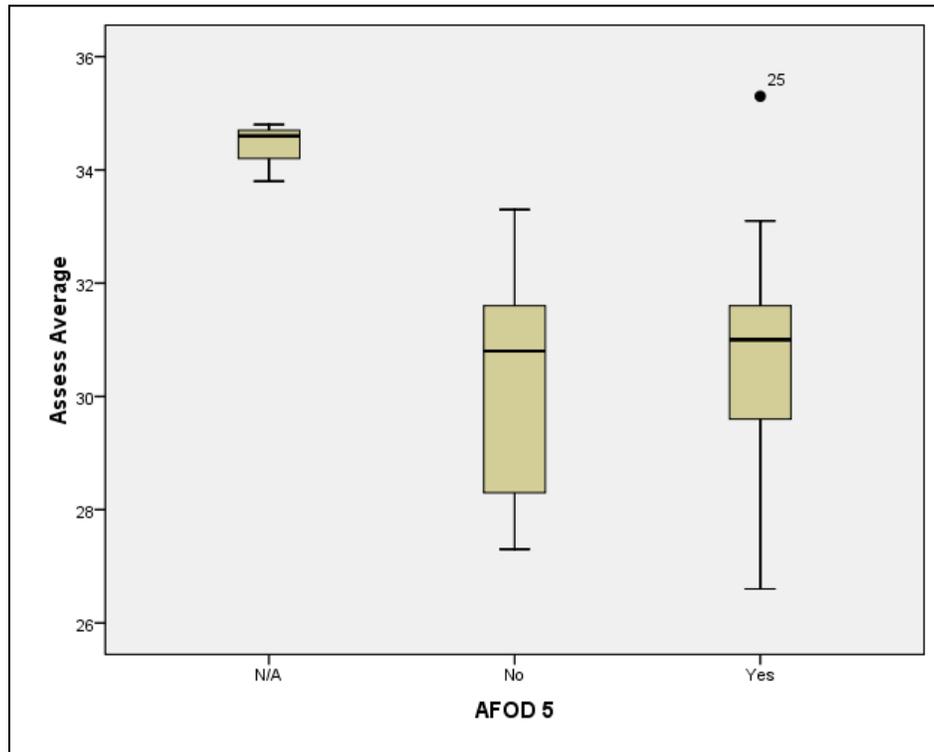**Completion of AFOD 5 and course results in planning**

Figure 1: Completion of AFOD 5 and assessed average.

At the .05 alpha significance level, is there a difference in the Planning mark for personnel who completed AFOD 5 (Yes vs No).

$H_0: \mu_d = 0$ $\qquad\qquad$ $H_1: \mu_d \neq 0$

H_0 is rejected if $t < -2.021$ or $t > 2.021$ (for 40 degrees of freedom).

Table 1: Group Statistics.

	AFOD 5	N	Mean	Std. Deviation	Std. Error Mean
Assess Average	Yes	29	30.71	1.710	.318
	No	18	30.22	2.028	.478

Table 2: *Independent Samples Test.*

		Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Assess Average	Equal variances assumed	2.634	.112	.879	45	.384	.485	.551	-.625	1.595
	Equal variances not assumed			.845	31.592	.405	.485	.574	-.685	1.654

t = 0.879 and therefore the Null hypothesis is rejected. There is no statistically significant difference between the Planning marks of those who completed AFOD 5 and those who have not.

Completion of AFOD 5 and course results in service paper
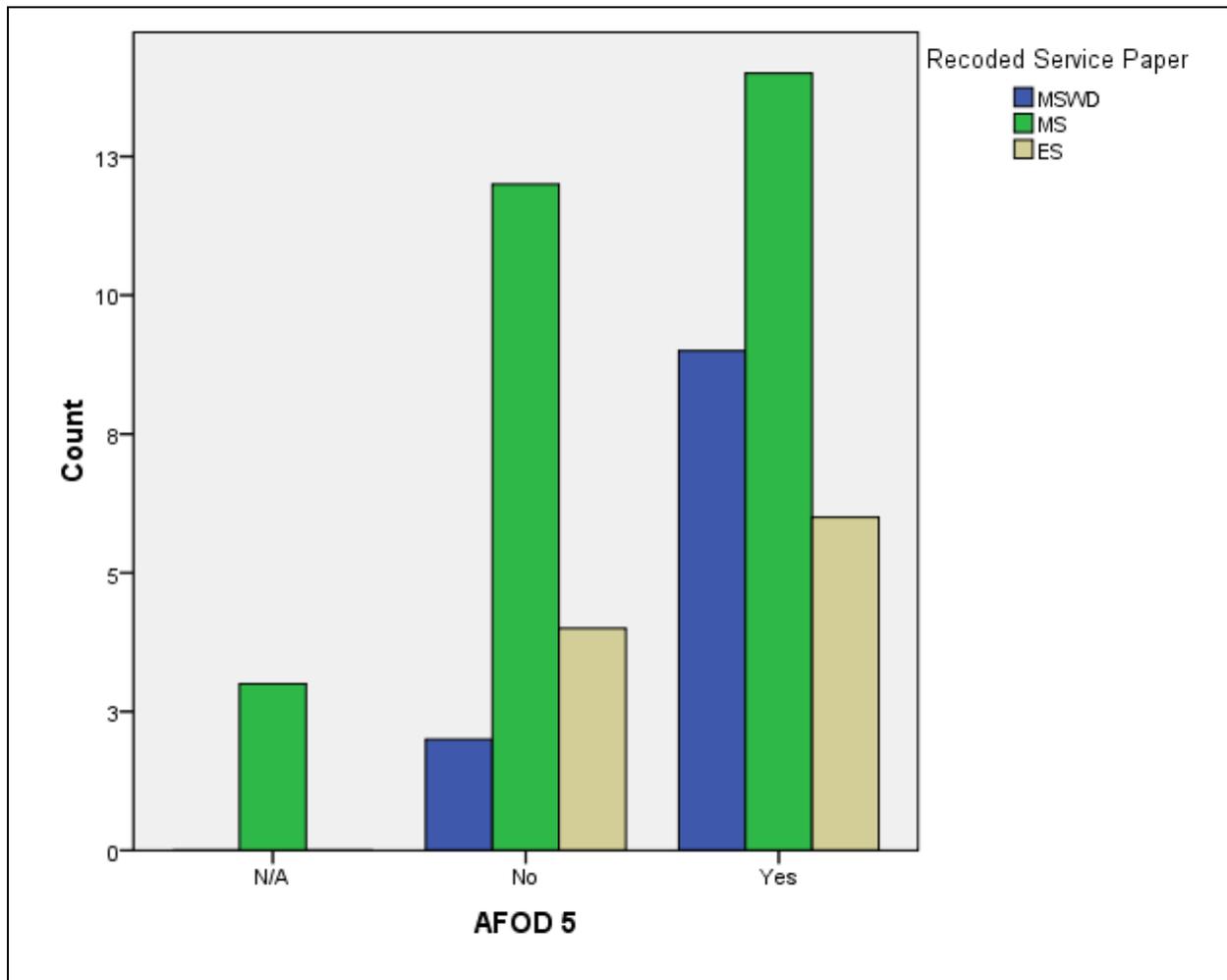

Figure 2: Completion of AFOD 5 and service paper results.

AFOD was recoded into N/A = 0, No = 1 and Yes = 2.

Service paper results were recoded as BS = 1, MSWD = 2, MS = 3 and ES = 4.

The Mann-Whitney *U* test for independent samples was used. At the .05 alpha significance level, is there a difference in the service paper marks for personnel who completed AFOD 5 (Yes vs No, N/A not included).

$H_0: \mu_d = 0$ $\qquad\qquad$ $H_1: \mu_d \neq 0$

H_0 is rejected if Asymp. Sig. (*p*) < 0.05.

Mann-Whitney Test

Table 3: Ranks.

	Recoded AFOD5	N	Mean Rank	Sum of Ranks
Recoded Service Paper	1	18	26.44	476.00
	2	29	22.48	652.00
	Total	47		

Table 4: Test Statistics[a].

	Recoded Service Paper
Mann-Whitney U	217.000
Wilcoxon W	652.000
Z	-1.071
Asymp. Sig. (2-tailed)	.284

a. Grouping Variable: Recoded AFOD5.

The Asymp. Sig (2-tailed) = 0.284 which is greater than 0.05 the null hypothesis cannot be rejected and therefore the completion of AFOD has no bearing on the results of the service paper.

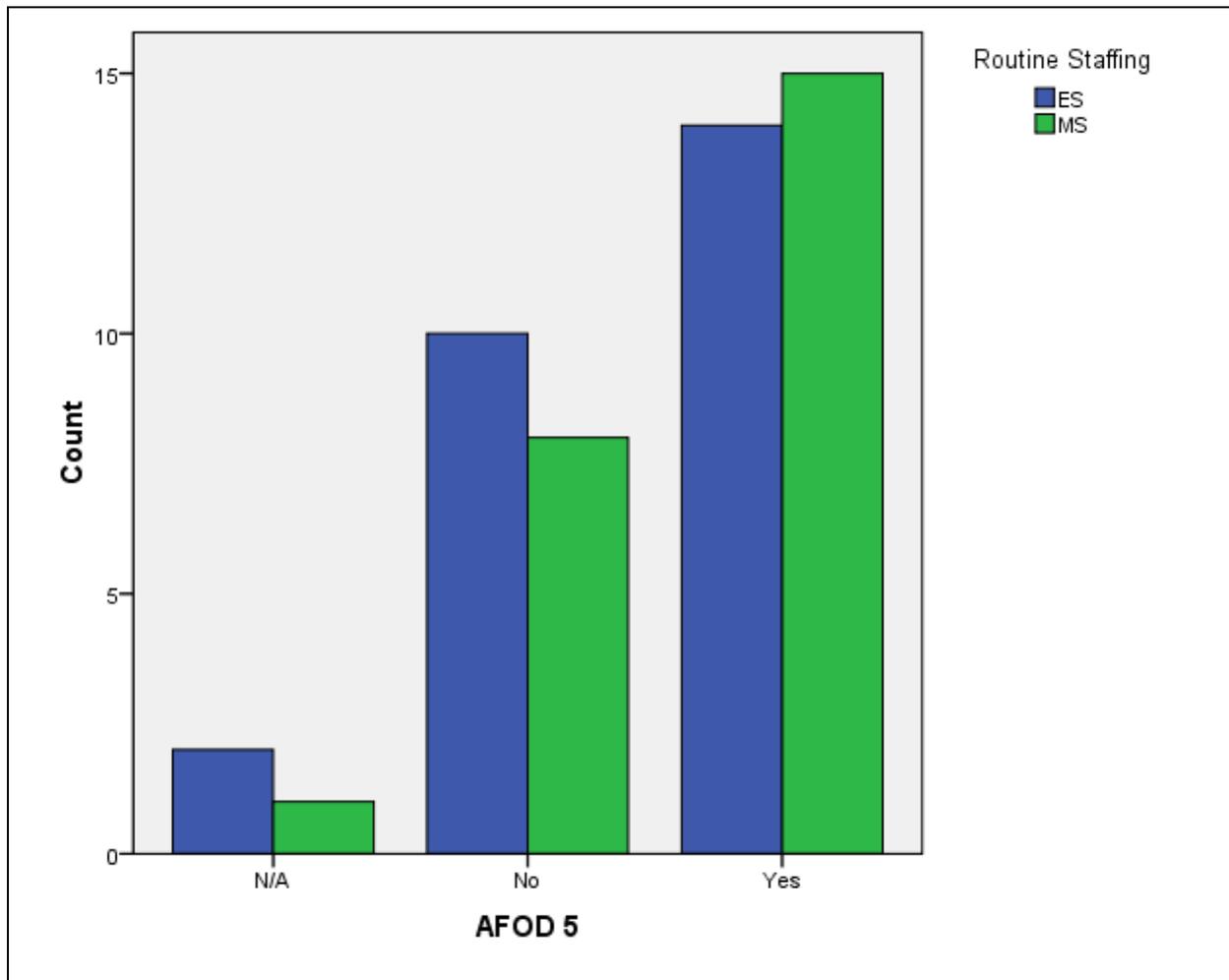**Completion of AFOD 5 and course results in staff duties**

Figure 3: AFOD 5 completion and staff duties assessed results.

AFOD was recoded into N/A = 0, No = 1 and Yes = 2.

Service paper results were recoded as BS = 1, MSWD = 2, MS = 3 and ES = 4.

The Mann-Whitney *U* test for independent samples was used. At the .05 alpha significance level, is there a difference in the staff duties marks for personnel who completed AFOD 5 (Yes vs No, N/A not included).

$H_0 : \mu_d = 0$ $\qquad\qquad$ $H_1 : \mu_d \neq 0$

H_0 is rejected if Asymp. Sig. (*p*) < 0.05.

Mann-Whitney Test

Table 5: Ranks

	Recoded AFOD5	N	Mean Rank	Sum of Ranks
Recoded Routine Staffing	1	18	24.67	444.00
	2	29	23.59	684.00
	Total	47		

Table 6: Test Statistics[a].

	Recoded Routine Staffing
Mann-Whitney U	249.000
Wilcoxon W	684.000
Z	-.294
Asymp. Sig. (2-tailed)	.769

a. Grouping Variable: Recoded AFOD5.

The Asymp. Sig (2-tailed) = 0.769 which is greater than 0.05 the null hypothesis cannot be rejected and therefore the completion of AFOD has no bearing on the results of the service paper.
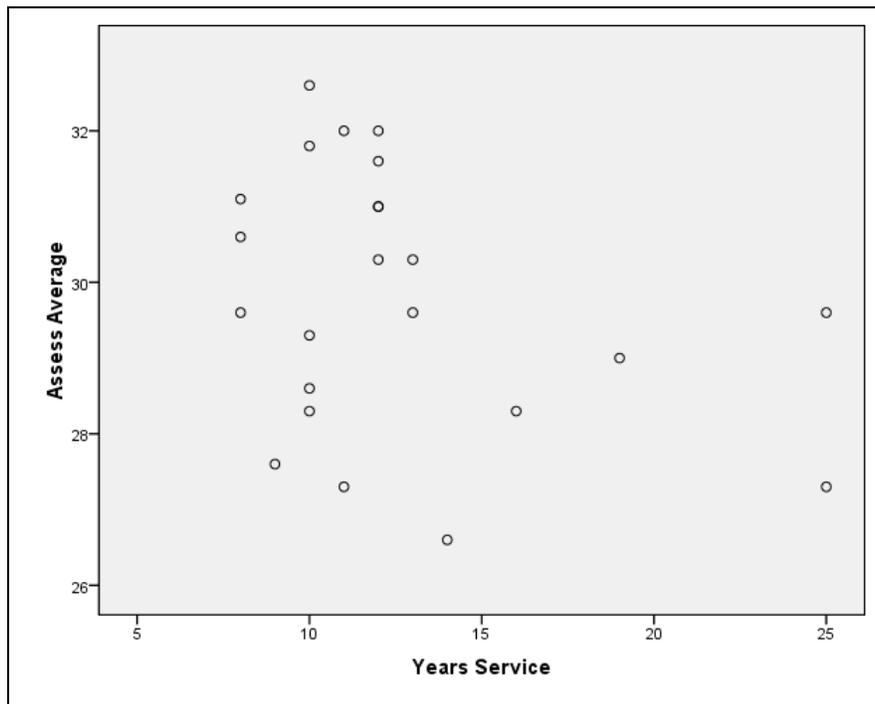
Years of experience and course results

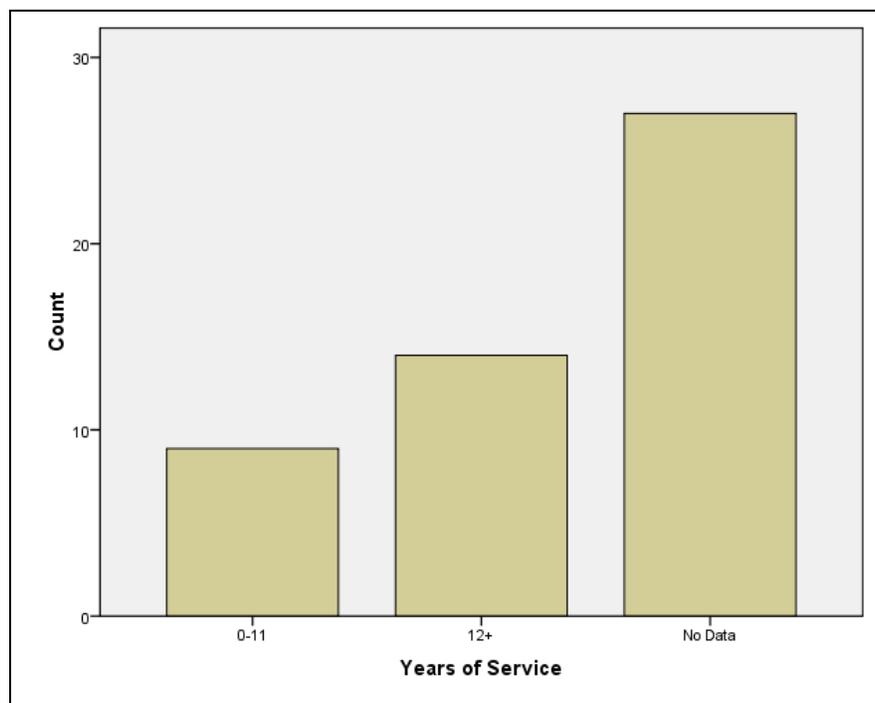*Figure 4: Assessed average vs years of service.*
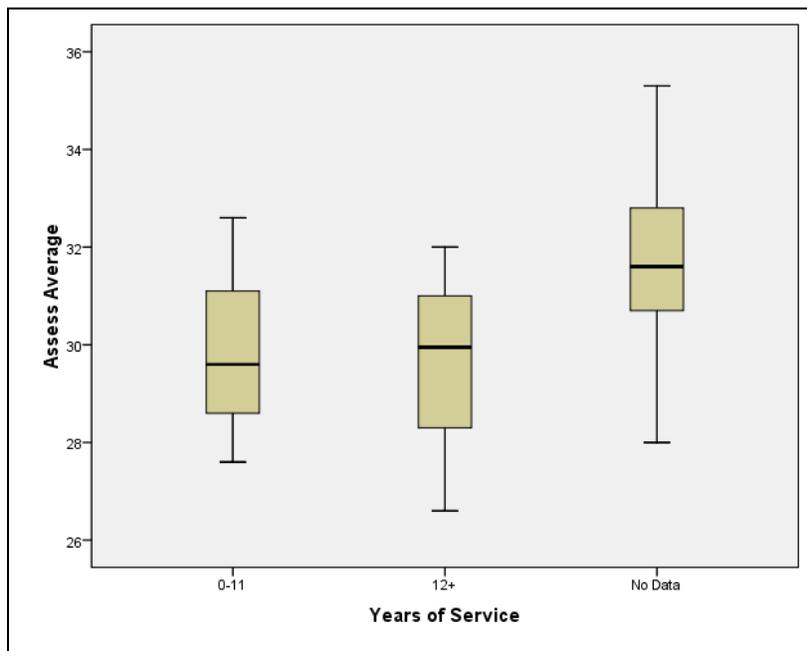

Figure 5: Number of students by years of service.

Figure 6: *Assessed average vs years of service.*

Mann-Whitney Test

Table 7: Ranks.

	Coded Yrs Service	N	Mean Rank	Sum of Ranks
Assess Average 1		9	12.50	112.50
	2	14	11.68	163.50
	Total	23		

Table 8: Test Statistics[b].

	Assess Average
Mann-Whitney U	58.500
Wilcoxon W	163.500
Z	-.284
Asymp. Sig. (2-tailed)	.776
Exact Sig. [2*(1-tailed Sig.)]	.781[a]

a. Not corrected for ties.

b. Grouping Variable: Coded_Yrs_Service.

The Exact Sig. = 0.781 which is greater than 0.05 the null hypothesis cannot be rejected and therefore there is no difference averages of personal with less than 12 years of service or 12 or more years of service.

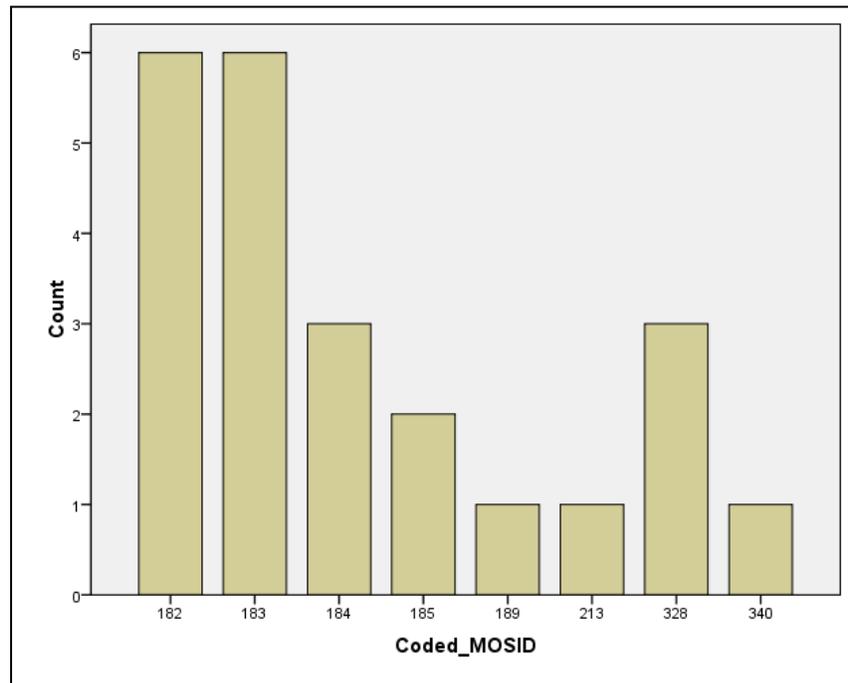**MOC and crse results (do some MOC fair better than others?).**
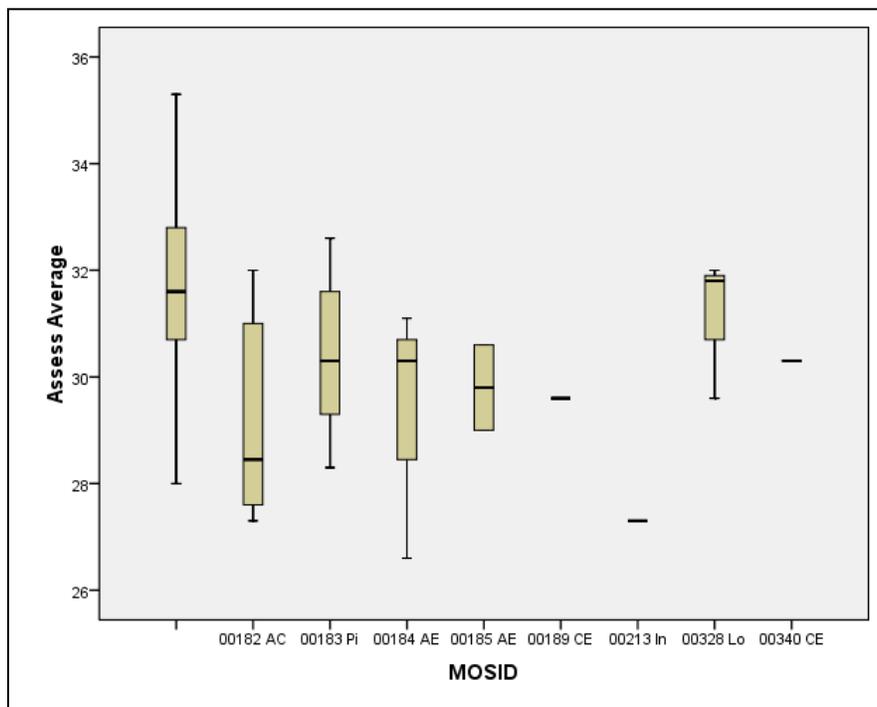

Figure 7: Numbers of students by MOSID.

Figure 8: Assessed average by MOSID.

Kruskal-Wallis Test

Table 9: Ranks.

	Coded_MOSID	N	Mean Rank
Assess Average	182	6	9.50
	183	6	14.00
	184	3	10.83
	185	2	11.50
	189	1	11.00
	213	1	2.50
	328	3	17.50
	340	1	13.50
	Total	23	

Table 10: Test Statistics[a,b].

	Assess Average
Chi-Square	5.466
df	7
Asymp. Sig.	.603

a. Kruskal Wallis Test.

b. Grouping Variable: Coded_MOSID.

The Asymp. Sig = 0.603 which is greater than 0.05 the null hypothesis cannot be rejected and therefore the is no statistical significant difference between the mean course marks by MOSID.

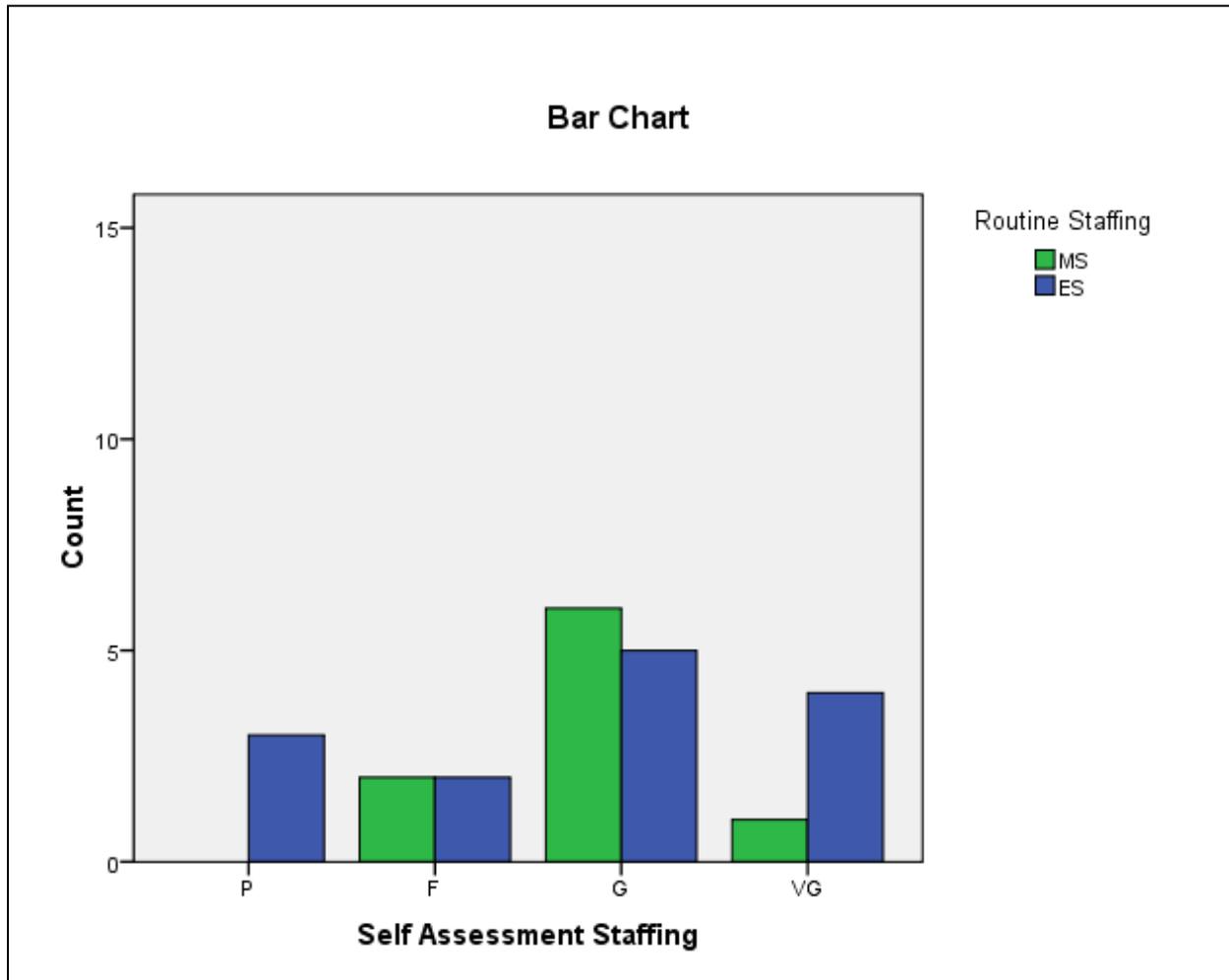**Self-assessment of staffing ability vs course results in Routine Staffing**

Figure 9: Self-assessment of staffing ability and assessed staff work.