Study of Factors Influencing the Performance of Students with the Analysis of Variance (ANOVA): Case of Morocco

Raouf Radouane
Assistant Professor
Faculty of Economics and Law
Mohammed V-Souissi University
Department of Economics
Rabat, Morocco

Echaoui Abdellah
Assistant Professor
Faculty of Economics and Law
Mohammed V-Souissi University
Department of Economics
Rabat, Morocco

Abstract

In this paper, we investigate a statistical study to highlight some pertinent indicators contributing to the performance of students in Moroccan universities with open access. Such problem has attracted a great attention in higher education and many studies are devoted to this problematic to get some insights about the factors influencing directly or indirectly the efficiency of students' productivity. Our approach is based on real data collected from many national different universities. From a sample of 901 students (S3 & S5) belonging to these institutions, we have performed ANOVA test by using the SPSS software to examine the impact of each variable which can be taken into account for explaining the dependent variable (student performance). It is worth noting that our methodology can be improved by other advanced statistic methods which will be the subject of our future work.

Key Words: Higher education, student performance, bi-variate analysis, ANOVA...

This work is done in the framework of a research project on higher education in Morocco and was conducted by the team ERADIASS of FSJES-Souissi, Rabat. The authors wish to thank all the members.

1. Introduction

Many students in open access academic institutions, who have difficulties in learning and problems to continue their education, are sometimes headed to failure and loss. In this sense, two theoretical models exist [Finnie and *al* (2012)]. The first one which focuses on the perseverance is the "integration of students" model made by Tinto (1975, 1993) where he integrated several sociological and economic factors that lead the student to abandon or persevere. The second model concerns the "downsizing" made by Bean and Metzer (1985) where the authors consider other external factors to the institutions.

In Morocco, there is evidence that several indicators of academic performance (success rate, graduation rate, insertion rate, attrition rate) are at half-mast. In this context, ERADIASS team (Team for Research in Computing and Data Analysis Applied to Social Sciences) of the FSJES of Mohammed V-Souissi University, has been involved since 2011 in the framework of a project funded by the university to study the determinants of student achievement in basic license.

Using a survey made beforehand with a sample of 1,500 students from several academic institutions, this project has allowed us to create a usable data base on SPSS. After correction and data cleansing, our sample includes responses from 901 students. About 80 questions were addressed to students. These questions form, from a hypothetical manner, the set of variables that can explain the performance of students (the explanatory variables).

The variable used to characterize the performance of students is the simple arithmetic average of the averages for the various semesters, which is named "Average-Total".

Our approach is in line with the work of Tinto (1975, 1993). Using the database of the survey, we analyze the impact of each variable on the average student. We specifically look for the variables that explain the better performance of students and those who have no influence on their performance.

The approach followed is to perform a bi-variate analysis using statistical comparison of means. The bi-variate analysis reveals the set of explanatory variables that are statistically significant and impact directly on the dependent variable. To consolidate our results, the analysis of variance (ANOVA) was used as test.

The analysis of variance is a family of methods to express and interpret the differences observed between groups for the same variable (ANOVA Analysis of Variance). This type of analysis enters within the general framework of the linear model, where a quantitative variable is explained by a qualitative variable.

ANOVA is used to check if there are differences between subgroups by studying their variance [M. CARECANO et al (2010)]. The null hypothesis is tested by the F-test in SPSS. The significance level used was 5%. We test the hypothesis of the absence of a relationship between two variables by examining the p-value corresponding to F calculated. If the p-value is less than 0.05, we reject the null hypothesis [J. Straford and P. Bodson (2006)] .

The homogeneity of the group variance and normality of the data are checked successively using the Levene test (null hypothesis in which the variances are equal in groups) and Kolmogorov -Smirnov test (null hypothesis: the normal distribution).

Tests and the analysis are performed by groups of variables in accordance with their order and coding given in the survey. We limit our work to the variables that seem most relevant. We are interested in the study of related variables:

2 - Personal information
3 - Previous studies
4 - Conditions of study
5 - Teaching tools
6 - Social conditions
7 - Perspectives

2. Personal Information

In this block of variables; where there are age, sex, marital status, availability to conduct the studies, the origin institution and the field, only how to conduct studies are statistically significant (sig = 0.002 <0.05), which may explain the difference in means. The test result shows that full-time students are more successful than others.

Average-Total / Availability to lead the studies

Avail./Study	Average	N	Standard-deviation
1	11,33	670	1,0075243
2	11,01	117	1,1425899
Total	11,28	787	1,0341338

Table ANOVA

		sum of squares	df	Average of squares	F	Signification
Average - total * Avail./Study	Inter-groupes Combined	10,029	1	10,029	9,479	,002
	Intra-class	830,545	785	1,058		
	Total	840,574	786			

3. Previous studies

3.1 Type of bachelor

Against all odds, the type of bachelor (Bac) has no effect on the average student (sig =0.52). The students seem to adapt to the requirements of the university and the type of Bac (scientific, technical, letter ...) has no effect on their performance.

Average-total

TypeBac	Average	N	Standard deviation
1	11,27	347	1,1413
2	11,33	347	,8945
3	11,18	90	1,0422
4	11,74	3	1,6890
Total	11,29	787	1,0290

Table ANOVA

			Sum of squares	df	Average of squares	F	Signification
Total-Average * TypeBac	Inter-groupes	Combined	2,370	3	790	745	525
	Intra-class		829,975	783	1,060		
	Total		832,345	786			

3.2 Mention of the bachelor

The mention of the bachelor directly affects student's performance. It is statistically significant (F = 5.507 with sig = 0.001), and helps explain the overall average. Students having good mention (grade) get better results than others.

Average-Total/ Mention of Bachelor

Mention_bachelor	Average	N	Standard Deviance
1	11,19	500	,9823
2	11,43	235	1,0688
3	11,63	48	1,2362
Total	11,28	785	1,0339

Table ANOVA

			Sum of squares	df	Average of squares	F	Signification
Total-Average * Mention_Bachelor	Inter-groupes	Combind	17,361	3	5,787	5,507	,001
	Intra-class		820,762	781	1,051		
	Total		838,123	784			

3.3 Establishment of origin

The overall average is not affected by the establishment of origin of the student. A slight difference in means between the private and the public institutions. Students with a foreign Bac arrive in third place. Anyway, the test is not significant (sig = 0.226) and therefore we cannot say that the establishment of origin explains or influences the overall average.

Total-Average

Estab_Origin	Average	N	Standard-deviation
1	11,28	652	1,0579
2	11,34	128	,9094
4	10,46	5	,3384
5	10,33	1	.
Total	11,29	786	1,0336

Table ANOVA

			Sum of squares	df	Average of squares	F	Signification
Total-Average * Estab_Origin	Intergroupes	Combined	4,655	3	1,552	1,455	,226
	Intra-class		834,121	782	1,067		
	Total		838,776	785			

4. Conditions of studies

This variable block is a set of factors related to study conditions that may influence the results of the student.

4.1 Understanding level of courses

The test shows that the level of understanding courses affects student performance. This variable is statistically significant (sig = 0.000). Students haven't difficulty in assimilating courses get better results.

Average-Total/ Understanding of courses

Cmprhcourses	Average	N	Standard-deviation
1	11,06	105	1,0296
2	11,18	469	,9233
3	11,60	216	1,1864

Table ANOVA

			Sum of Squares	df	Average of Squares	F	Signification
Total Average * Cmprhcourses	Inter-groupes	Combined	32,638	2	16,319	15,818	,000
	Intra-class		811,910	787	1,032		
	Total		844,548	789			

4.2 Obstacles to understanding

This variable provides information about the type of barriers to understanding. The result of this test shows that students who have language problems have lower averages than others (sig = 0.000). This result reinforces the test for the understanding of the courses. Students with language problems, a priori they have difficulty understanding.

Total-Average/Obstacles

Obstacles	Average	N	Standard-deviation
1	11,1977	250	,91853
2	11,2130	237	,98450
3	11,5801	178	1,08305
Total	11,2846	723	1,0389

Table ANOVA

			Sum of squares	df	Average of squares	F	Signification
Total-Average * Obstacles	Inter-groupes	Combined	28,727	4	7,182	6,852	,000
	Intra-class		814,348	777	1,048		
	Total		843,074	781			

4.3 Schedule

The test results indicate that the use of time (too busy or not) has no impact on student performance (sig = 0.593).

Total-Average

Schedule	Average	N	Standard-deviation
1	11,2670	443	,9570
2	11,3068	345	1,1290
Total	11,2844	788	1,0353

Table ANOVA

		Sum of squares	df	Average of squares	F	Signification
Total-Average * Schedule	Inter-groupes Combined	,307	1	,307	,286	,593
	Intra-class	843,294	786	1,073		
	Total	843,601	787			

4.4 Attendance rate

The frequency of presence of students has a real effect on the overall average. Being regularly in courses (mode 3) significantly improves the average. The test is highly significant (sig = 0%). The variable "Attendance rate" clearly explains the performance of students. It's, more or less, an obvious fact confirmed by the statistical test.

Total-Average

Presence	Average	N	Standard-deviation
1	11,0758	70	1,27330
2	11,1598	387	,990459
3	11,4770	329	1,00415
Total	11,2834	787	1,03636

Table ANOVA

		Sum of squares	df	Average of squares	F	Signification
Total-Average * Presence	Inter-groupes Combined	22,935	3	7,645	7,289	,000
	Intra-class	821,270	783	1,049		
	Total	844,205	786			

4.5 Presence in tutorials

This variable provides information about presence or not in tutorials (TD). The statistical study shows that 40% of students regularly attend the TD, 45% occasionally and 15% never attend. The study of the comparison of means shows that this variable has an impact on student performance. Students who regularly attend sessions TD have a higher average compared to others. The test is significant (sig = 0.000).

Total-Average

PresenceTD	Average	N	Standard - deviation
1	11,0728	105	1,0461
2	11,2023	367	1,0128
3	11,4678	312	1,0242
Total	11,2906	784	1,0315

Table ANOVA

		Sum of Squares	df	Average of squares	F	Signification
Total-Average * PresenceTD	Inter-roupes Combined	17,634	2	8,817	8,444	,000
	Intra-class	815,551	781	1,044		
	Total	833,186	783			

5. Teaching tools

In this group of variables, we think that access to a number of educational tools will allow students to improve their performance. These tools are multiple, we use two variables: the computer ownership and remedial teaching (private lessons).

5.1 Computer ownership

The difference between owning a computer or not is statistically significant (F (1,780) = 8.633, p = 0.035 <5%) and helps to explain the performance of student. We note that students who own a computer have in general a higher average than those who do not have one.

Total-Average

Computer	Average	N	Standard-deviation
1	11,3255	683	1,0160
2	11,0008	99	1,1057
Total	11,2844	782	1,0327

Table ANOVA

			Sum of squares	df	Average of squares	F	Signification
Total-Average * Ordinateur	Inter-groupes	Combined	9,119	1	9,119	8,633	,003
	Intra-class		823,890	780	1,056		
	Total		833,008	781			

5.2 Remedial teaching

This variable provides information on the fact that the student follows private lessons, tutoring, or not. Descriptive statistics show a slight advance of those who do not go to tutoring, but this observation is not confirmed by the test (sig = 0.37).

Total-Average

Remed-teach	Average	N	Standard-deviation
1	11,2087	104	1,0843
2	11,3031	673	1,0263
Total	11,2905	777	1,0340

Table ANOVA

			Sum of squares	df	Average of Squares	F	Signification
Total-average * CrSt	Inter-groupes	Combined	,803	1	,803	,751	,387
	Intra-class		828,985	775	1,070		
	Total		829,788	776			

6. Social Conditions

We find in this group all exogenous factors that can influence directly or indirectly the performance of students.

6.1 The type of student housing

The overall average is not influenced by the type of accommodation of the student, be it family, campus, tenant or collocation, the average does not vary significantly (sig = 0.3).

Total-Average

TypLog	Average	N	Standard-deviation
1	11,3108	565	1,0410
2	11,2466	188	,9962
3	10,7854	13	,8222
4	11,3021	21	1,1711
Total	11,2866	787	1,0315

Table ANOVA

		Sum of squares	df	Average of Squares	F	Signification
Total-Average * TypLog	Inter-groupes Combined	3,904	3	1,301	1,224	,300
	Intra-class		783	1,063		
	Total		786			

6.2 Educational level of parents

The statistical test doesn't show any significant difference of mean regarding the level of parental education (sig = .19). Our hypothesis was to admit that the level of education of parents can have a positive effect on student performance is not confirmed by the ANOVA test.

Total-Average

LevPare	Average	N	Standard-deviation
1	11,3226	204	1,1192
2	11,3227	407	,9947
3	11,1550	160	1,0407
Total	11,2879	771	1,0393

Table ANOVA

		Sum of squares	df	Average of squares	F	Signification
Total-Average * Levpare	Inter-groupes Combined	3,561	2	1,780	1,651	,193
	Intra-class	828,269	768	1,078		
	Total	831,830	770			

7. Prospects

This group of variable includes, for student, perspectives in terms of employment after his university studies, pursuit of higher education or concerns about the labor market. It seems, according to the statistical tests performed, that all these variables have no impact on student performance.

Results about concerns inherent to labor market

Total-Average

Labormarket	Average	N	Standard-deviation
1	11,2795	601	1,0648
2	11,3069	181	,9198
3	11,4065	5	,9200
Total	11,2866	787	1,0314

Table ANOVA

		Sum of squares	Df	Average of squares	F	Signification
Total-Average * labormarket	Inter-groupes Combined	,176	2	,088	,083	,921
	Intra-class	836,065	784	1,066		
	Total	836,241	786			

Conclusion

In conclusion, we find, after statistical analysis, almost 50 % of the tested variables have no significant influence on student performance. In addition, some results are cons-intuitive as the case of the variable "type of bachelor" that has no effect on the overall average. In this sense, we remained vigilant about the interpretation of results and without explaining the causes of these facts. The problems of accuracy of the data and the conditions of their collection may be responsible for such contingencies.

However, the factors that explain the overall average, such as the availability to conduct the studies, the presence in courses and presence in tutorials, the mention of bachelor, understanding the course, the problem of language and other, came to confirm what generally recommends the lecturers in these universities.

The approach by the bi-variate statistical analysis is not without limitations, it must be complemented by other statistical tools. It may be that variables which are not explained one by one, if interaction is introduced can change the result. In this context, and in line with this work, a MANOVA (Multivariate Analysis Of Variance) will be more appropriate.

Références

BEAN, J.P., & METZNER, B.S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Review, 55*, 485–540.

M. CARECANO et al (2010). Analyse de données avec SPSS. Pearson.

FINNIE, R., S. CHILDS et T. QIU (2012). Persévérance aux études postsecondaires : Nouvelles données pour l'Ontario, Toronto, Conseil ontarien de la qualité de l'enseignement supérieur.

J. STRAFORD et P. BODSON (2006).L'analyse multivariée avec SPSS, Presses de l'université de Québec.

TINTO, V. (1993), Leaving College: Rethinking the causes and cures of student attrition. USA: University of Chicago Press.

TINTO, V. Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 1975, 45, 89-125.