Knowledge Discoveries on Performance of IGNOU Science Graduates through Data Mining

PANKAJ KHARE, ANURAG SAXENA and SURESH GARG
Indira Gandhi National Open University, New Delhi

Abstract: *In this paper we have applied the data mining technique for knowledge discoveries on IGNOU science graduates by analyzing the relevant records with a view to facilitate science educators to apply the findings to support pedagogy in laboratory courses. A model construct "cuboid deep drilling", has been done to discover the most popular lab course combination.*

The paper also brings out the selectivity applied by students in opting lab courses, their preference towards different science disciplines, and their performance in practical examinations. In particular, about 37% students are active and only one-fourth of them attempt lab courses. Chemistry is the most popular and highest scoring subject among all disciplines offered by IGNOU.

Introduction

Empirical research into student performance is important for introducing innovations and reforms in the fast changing global scenario in higher education. This is particularly true of a new, dynamic and vibrant system like open and flexible learning, which offers need-based relevant programmes and courses. However, availability of reliable students' data and its authentic analysis has attracted little attention of researchers. In such a situation, impressionable conclusions are likely to be arrived at and there is a lurking fear that these may create many myths that could be detrimental to the growth of the system. As such, it can become an impediment not only for informed policy decisions but also for their implementation, and impact assessment of the system. It is therefore extremely important for any institution to analyse data of various programmes on offer and derive conclusions based on authentic research about how well its 'products' I) are received and valued by the employer groups/society, ii) perceive their own strengths, skills and competencies vis-à-vis their peers, and iii) judge the value-addition by the system in their career. With a view to know answers to some such questions, we have analysed the date about the science students of Indira Gandhi National Open University (IGNOU), which offers education through distance mode.

This study uses data mining and knowledge discovery techniques. Data mining technique employs an array of data analysis tools on critically important data. The

Indian Journal of Open Learning, 2003, 12(1), 29-45
ISSN 0971-2690, Printed in India,

process involves collection, exploration and selection of data (Prabhu, 2002) to discover knowledge in the form of patterns and associations (Borgelt and Kruse, 2002). Then a predictive model based on the patterns is postulated and is used to make valid predictions. Finally, the model is verified empirically (Berry and Linoff, 2000, Han and Kamber, 2001).

The IGNOU enrolled over three hundred thousand students in the year 2003 and the active enrollment has now crossed one million. This has provided great impetus to teaching learning of all disciplines at the tertiary level. The offer of bachelor of science programme, launched since 1991, has thrown several key questions like parity of curriculum and acceptability of the products of open system and the conventional system, workload of students and their performance, retention and success rates. While the experience gained so far has helped to remove most of the doubts about the system as well as its products, preliminary researches on academic issues have settled initial skepticism of conventional academia (Garg et.al., 1992; Vijayshri et.al., 1998). In this work, a data set of approximately 32,500 records of undergraduate science students, computerized over a period of 11 years (from 1992 to 2003) has been taken as base data from the data repository of IGNOU. Of these, 6,948 records are of those students who attempted at least one laboratory course. In this paper, we report our analysis on knowledge discovery on students' performance based on practical courses.

Distance Science Education

It has been realized since long that for socio-economic development, creation of scientific temper and culture of logical thinking are most important ingredients. It means that greater demand has to be placed on science education to improve the quality of life (Jegede, 2001). For science programmes offered through distance mode, Chandra (1995) reviewed the utility of practical component – the integral component of science curriculum – by looking at the aims of lab teaching and experimental projects. Laboratory training allows the learners to become familiar with materials, and equipment, and gain insight of natural phenomena. This also enables acquisition of laboratory skills and understanding of the processes through which scientific knowledge is accumulated and scientific principles are understood. An important outcome of experimental work is the enhancement of motivation to learn (Rao, 2002).

Since acquiring hand-on skills and competencies is the most important pedagogical issue in science education, there should be no compromise with quality in assessment of science lab courses, particularly at the tertiary level. Kannan et al (1998) have argued that to infuse a sense of professionalism and competence in science graduates, the assessment pattern should be designed to measure the degree of attainment of experimental as well as cognitive abilities. It necessitates process assessment as well as product assessment.

Lab Courses at IGNOU

The lab courses are integral component of any science programme offered by an institution either on-campus or through the distance mode. In B.Sc. programme of

IGNOU, the lab courses, though mandatory, have been designed as stand-alone courses (Vijayshri et.al, 1998; Khan et.al., 2000). Just as computing courses provide opportunities to experience the abstraction and design paradigms, through hands-on experience in sciences it has been recognized that students must have access to appropriate lab resources. In the face-to-face (F2F) mode, the institution traditionally provides this access. However, an open university like IGNOU creates an access to countrywide network of laboratories by utilizing the infrastructure of other educational institutions for conduct of lab courses under the supervision of experienced counsellors. In a way, the practical courses of an Open University are conducted in F2F mode.

As per the curriculum design of the B.Sc. programme of IGNOU, there are 42 theory (11 in Physics, 7 in Chemistry, 10 in Life Science and 14 in Mathematics) and 12 practical courses (3 in Physics, 5 in Chemistry and 4 in Life Sciences) as science electives. These courses are given weightage in terms of credits (one credit is equivalent to 30 hours of learner academic activity (IGNOU Profile, 2003)) and provide autonomy to learners to map their study according to their interest. The theory courses worth 36 credits in Physics, 26 credits in Chemistry, 48 credits in Life Sciences and 52 credits in Mathematics (Total 162 credits) are on offer.
The lab courses are worth 42 credits (12 credits in Physics, 14 credits in Chemistry and 16 credits in Life science), as shown in Table 1.

Table 1: List of lab courses

Physics :	1.	PHE-03(L) – Physics Laboratory – I (On offer from 1992)	4 Credits
	2.	PHE-08(L) – Physics Laboratory – II (On offer from 1992)	4 Credits
	3.	PHE-12(L) – Physics Laboratory – III (On offer from 1996)	4 Credits
Chemistry :	4.	CHE-03(L) – Chemistry Lab – I (On offer from 1992) (To be taken with theory course CHE-01 – Atoms and Molecules)	2 Credits
	5.	CHE-07(L) – Chemistry Lab – II (On offer from 1993)	2 Credits (To be taken together)
	6.	CHE-08(L) – Chemistry Lab – III (On offer from 1993)	2 Credits
	7.	CHE-11(L) – Chemistry Lab – IV (On offer from 1994)	4 Credits
	8.	CHE-12(L) – Chemistry Lab – V (On offer from 1995)	4 Credits
Life Sciences :	9.	LSE-04(L) – Laboratory Course – I (On offer from 1993)	4 Credits
	10.	LSE-08(L) – Laboratory Course – II (On offer from 1995)	4 Credits
	11.	LSE-11(L) – Animal Diversity Lab (On offer from 2001) (To be taken with theory course LSE-09 & LSE-10 – Animal Diversity I & II)	4 Credits
	12.	LSE-14(L) – Plant Diversity Lab (On offer from 2001) (To be taken with theory course LSE-12 & LSE-13 – Plant Diversity I & II)	4 Credits

In order to major in a discipline, a student has to complete 48 credits in that discipline. A student must earn at least 25% of the total elective credits opted for in disciplines with experimental component (Physics, Chemistry and Life Science) in laboratory courses to successfully complete the requirement of B.Sc. programme. For example, if a student opts for 40 credits in physics and chemistry or chemistry and biology or physics and biology or physics, chemistry and biology, at least 10 credits out of 40 should be earned only in laboratory courses. (In any discipline, not less than eight credits should be earned). The lab courses are conducted during summer and autumn vacations, evenings or weekends so that in-service persons can attend laboratory courses without difficulty. A lab course requires full-time presence of the learner at the study centre for a period equivalent to one or two weeks. During this time, a student has to work for around 60 to 120 hours, depending on the credit weightage. Around 80% of this time is spent on experimental work and the remaining time is used for doing the calculations, preparation of reports, viewing or listening to the video/audio programmes (Programme Guide, 2002). An innovative feature of the practical work is that the work done on any day is evaluated on the next day and it carries 70 percent weightage of the total award.

Tools and Techniques Employed

Data Mining

Data mining is a multidisciplinary field, drawing upon work done in areas like database technology, artificial intelligence, machine learning, neural networks, statistics, knowledge-based systems, knowledge acquisitions, information retrieval, and data visualization (Fayyad et.al., 1996, Han and Kamber, 2001). Data mining is a repository of data analysis tools that can identify patterns and relationship in data that are otherwise not known. It then uses this information to make predictions (Prabhu, 2002). Data Mining is clearly not replacement of statistics, rather it is an extension of it. It makes use of advances in computing to apply statistical techniques to a bigger dataset.

Any data mining exercise begins with the description of data through statistical measures such as averages and percentages, tabulations and cross tabulations, graphs and charts. It then makes predictions by using an array of classification and regression methods such as Neural Networks, Decision Trees, Regression, Discriminant Analysis and Boosting (Berry and Linoff, 2000, Han and Kamber, 2001, Borgelt and Kruse, 2002), etc.

To be able to use data mining tools to analyze the data, a researcher has to build a data mart, which puts the data in a form suitable for analysis. The data mart can be classified into two groups – Multidimensional Online Analytical Processing (MOLAP) and Relational OLAP. In multidimensional data mart, the numeric data can be sliced and diced in a manner independent of data modeling constraints of a DBMS. The ROLAP or relational OLAP data mart may contain both text and numeric data and be supported by RDBMS (Prabhu, 2002). The data drawn from various sources have to be cleaned, arranged and stored to form a data warehouse for useful knowledge discoveries. The refining process is an important aspect in consolidating the data, and the same is true in present educational database. The data under analysis is a ROLAP data mart, i.e., the data fields in a database have inter-dependence and need interaction for extracting the information. For example, the key field of registration number controls multiple data sets for interaction, examination results in multiple fields (assignments, term-end or lab marks), etc.

Data Refining

The data collected from the central repository was found to have a difficult accumulation problem. At IGNOU, the student records are kept in various database record files. The software used is FoxBASE. Unique enrolment number controls each record, which is normally a numeric field with length of 9 digits. First two digits represent the year of registration. The data is maintained in four basic databases, namely,

1. address file, containing basic information about the students;
2. elective file, containing choice of electives taken by the students;
3. theory result file, containing students' assignments and term-end marks and their result; and
4. practical file, containing grades of laboratory exercises.

There are other databases having information of material dispatch, student support, convocation, etc.

The data collected from these databases were first merged into a single file. Various fields were found to be intermingling and records provided inconsistent information. Using small programs, written in FoxBASE, data refining was carried out to fill the gaps in data, removing inconsistencies and noises. At the end, manual refining was also done and a ROLAP data mart was created.

Data Description

The total data taken into account comprised 32,551 unrefined records of undergraduate science students who took admission in B.Sc. Programme of IGNOU from 1992 to 2003. The year-wise registration of the students is given in Figure-1, which reveals that registration showed linear increase up to 1996. This was followed by a 'ripple effect' for the next four years, which is reflective of its stabilisation. From 2000 onwards, the enrolment have registered significant rise. This could be due to re-registration facility provided by the university to those learners who were unable to successfully complete 96 credits in the stipulated period of 8 years. Also, the rising credibility of the programme, increasing reach of the University to geographically inaccessible regions and growing demand for higher education by eligible student population over the years countrywide seem to have contributed to this increase in registration.

Out of total 32,551 students, 5,044 students (15.5%) registered in 2003. These records are not included in the analysis because they have not yet got a chance to take a practical course. Among 27,507 students, 10,049 (36.53%) have attempted at least one science theory elective or laboratory practical exam. The data further reveals that 6,948 students (25.26%) have attempted various lab courses since the launch of the programme. Thus, for the present study, 6,948 student records formed the basis of our analysis. Henceforth, it will be referred to as "data".

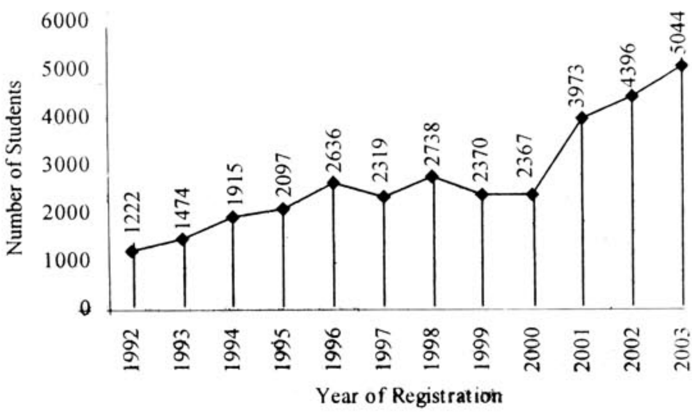


Figure-1 : Line chart showing year-wise registration of BSc students

Table 2 : Year-wise number of students attending lab courses

Year	PHE3L	PHE8L	PH12L	CHE3L	CHE7L	CHE8L	CH11L	CH12L	LSE4L	LSE8L	LS11L	LS14L
1992	95	26	21	136	55	44	17	14	37	16	0	0
1993	149	44	16	213	86	73	47	36	72	34	0	0
1994	276	85	58	459	128	104	91	79	114	53	0	0
1995	364	133	64	460	160	150	146	151	140	82	0	0
1996	380	134	92	692	234	212	191	184	238	183	2	0
1997	389	167	123	544	271	272	213	199	249	197	0	0
1998	449	221	149	699	345	314	260	259	371	284	2	1
1999	378	162	86	577	277	222	166	163	297	192	54	4
2000	255	87	19	459	166	113	77	37	210	85	19	6
2001	203	18	0	340	136	86	10	0	258	24	0	1
2002	20	0	0	29	7	4	0	0	43	0	0	0
Total	**2958**	**1077**	**628**	**4608**	**1865**	**1594**	**1218**	**1122**	**2029**	**1150**	77	**12**

Analysis and Discussions

Course Options

The data was summarized to classify the options of laboratory courses. Table-2 provides year-wise details of lab courses attempted by students.

Figure-2 shows attendance of students in different lab courses. As may be noted, the attendance was maximum in the first level of lab courses (PHE3L, CHE3L and LSE4L) for all three disciplines. The number of students opting higher-level lab courses declined in later years. For example, maximum number of students attended practical for CHE3L, whereas there is a clear drop in attendance for CHE7L, CHE8L, and CHE11L to CHE12L. This is most likely due to the fact that the first level courses are standalone and most generic, whereas higher-level lab courses are specialized.

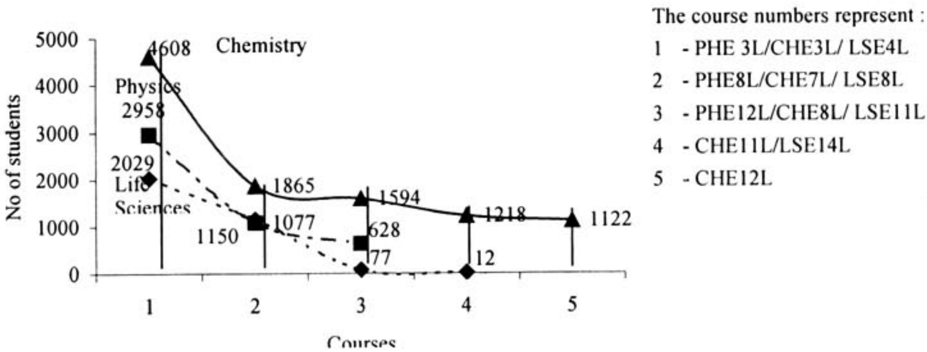


Figure-2 : Appearance of students in lab courses

Table-2 also reveals that the weighted ratio of distribution, i.e., credit values of the discipline lab courses and number of students attempting these is, Physics : Chemistry : Life Science :: 1.90 : 3.64 : 1.00. That is, if one student opts Life Science Lab, two opt Physics Lab and four opt Chemistry lab (percentage-wise, Physics - 29.08%; Chemistry - 55.63% and Life Sciences - 15.29%). An important finding has been that though the programme structure binds students to opt CHE7L and CHE8L together, a combination analysis revealed that students still do not attempt both lab courses. A microanalysis of data carried out to find the details of this fact showed that only 872 students attempted both the lab courses together, whereas 1,865 students attempted CHE7L and 1,594 students attempted CHE8L. Such an inappropriate combination can lead the students into unwarranted difficulties and it would be desirable to clearly lay down necessary instructions and inform students at the time of re-registration or declaration of results but before completion of the programme.

To find out the various combinations opted by the students among the disciplines, a matrix was developed (Table-3).

Data distilled from the warehouse reveals that :

Students are very selective; 29.52% opted only Chemistry, 17.24% only Physics and 5.80% only Life Sciences lab courses. 37.06% students attempted only one lab course from a single discipline (20.08% Chemistry; 12.41% Physics and 4.58% Life Sciences). 12.22% students attempted two lab courses. However, very few (1.53%) students attempted three or more lab courses (1.50% students attempted four lab courses and 0.20% students attempted five lab courses) in a single discipline. 47.50% students attempted lab courses from more than one discipline.

This analysis indicates that more than half students who attempted lab courses preferred to limit themselves to only one discipline. Chemistry was the first preference and Physics was second. The choice of Life Sciences was third. This is not surprising because these courses have been put on offer recently.

Of 47.50% students who attempted lab courses in more than one discipline, 19.37% students took Chemistry with Physics as combination, whereas 18.11% students took Chemistry with Life Sciences and 5.80% students took Physics with Life Sciences combination. 8.36% students took lab courses for all the disciplines. While the former two combinations are traditional, the third option shows awareness of emerging fields. However, the last combination comprises the group of **degree seekers with no preferences.**

Table 3 : Lab course option matrix

Life Sc	Physics	Chemistry						Grand Total
		0	1	2	3	4	5	
0	0		1395 (20.08%)	462 (6.65%)	76 (1.09%)	104 (1.50%)	14 (0.20%)	2051 (29.52%)
	1	862 (12.41%)	644 (9.27%)	131 (1.89%)	65 (0.94%)	81 (1.17%)	12 (0.17%)	1795 (25.83%)
	2	307 (4.42%)	119 (1.71%)	54 (0.78%)	62 (0.89%)	52 (0.75%)	18 (0.26%)	612 (8.81%)
	3	29 (0.42%)	17 (0.24%)	33 (0.47%)	21 (0.30%)	61 (0.88%)	1 (0.01%)	162 (2.33%)
0 Total		1198 (17.24%)	2175 (31.30%)	680 (9.79%)	224 (3.32%)	298 (4.29%)	45 (0.65%)	4620 (66.49%)
1	0	318 (4.58%)	392 (5.64%)	118 (1.70%)	84 (1.21%)	118 (1.70%)	46 (0.66%)	1076 (15.49%)
	1	67 (0.96%)	42 (0.60%)	35 (0.50%)	32 (0.46%)	32 (0.46%)	14 (0.20%)	222 (3.20%)
	2	7 (0.10%)	13 (0.19%)	7 (0.10%)	13 (0.19%)	18 (0.26%)	13 (0.19%)	71 (1.02%)
	3	5 (0.07%)	1 (0.01%)	7 (0.10%)	7 (0.10%)	19 (0.27%)	17 (0.24%)	56 (0.81%)
1 Total		397 (5.71%)	448 (6.45%)	167 (2.46%)	136 (1.96%)	187 (2.69%)	90 (1.30%)	1425 (20.51%)
2	0	80 (1.15%)	61 (0.88%)	155 (2.23%)	110 (1.58%)	146 (2.10%)		552 (7.94%)
	1	6 (0.09%)	19 (0.27%)	26 (0.37%)	17 (0.24%)	32 (0.46%)	7 (0.10%)	107 (1.54%)
	2	1 (0.01%)	5 (0.07%)	18 (0.26%)	15 (0.22%)	47 (0.68%)	22 (0.32%)	108 (1.55%)
	3			7 (0.10%)	10 (0.14%)	82 (1.18%)		99 (1.42%)
2 Total		87 (1.25)	85 (1.22%)	206 (2.96%)	152 (2.19%)	307 (4.42%)	29 (0.42%)	866 (12.46%)
3	0	5 (0.07)	3 (0.04%)	19 (0.27%)	6 (0.09%)			33 (0.47%)
	1		1 (0.01%)		1 (0.01%)			2 (0.03%)
	2			1 (0.01%)		1 (0.01%)		2 (0.03%)
	3							
3 Total		5 (0.07%)	4 (0.06%)	20 (0.29%)	7 (0.10%)	1 (0.01%)		37 (0.53%)
4	0							
	1							
	2							
	3							
4 Total								
Grand Total		1687 (24.28%)	2712 (39.03%)	1073 (15.44%)	519 (7.47%)	793 (11.41%)	164 (2.36%)	6948 (100.00%)

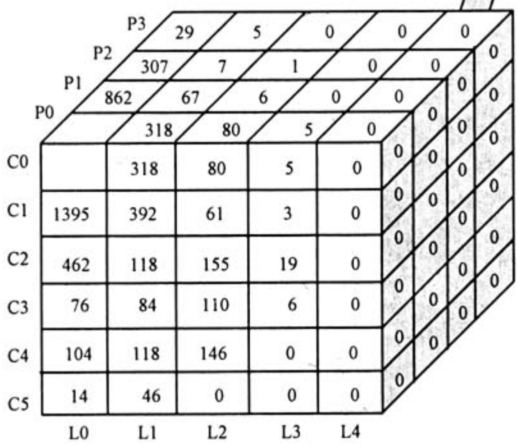
* *The combination not opted by any student is kept blank. Shaded area indicates most popular combinations opted by students among the disciplines.*

Most popular combination from amongst Physics & Chemistry disciplines is one Physics and one Chemistry (1P and 1C) lab courses (9.27%), followed by two Chemistry and one Physics (2C and 1P) course (1.89%) and two Physics and one Chemistry lab courses (1.71%).

When all three disciplines are taken into account, a combination of three Physics, four Chemistry and two Life Science (3P, 4C & 2LS) lab courses is **most popular** (1.18%). No student has attempted all four-lab courses of Life Sciences so far.

From the bouquet of 12 lab courses available, 1.74% student opted a maximum of nine lab courses; 1.40% attempted eight, 2.5% attempted seven, 5.33% attempted six, 7.33% attempted five, 7.99% attempted four, and 8.55% attempted three lab courses. This means that more than 65% of students opted for one or two lab courses and that too at the first level. It suggests that learners are serious and heed to the advice of the teachers given through programme guide and induction meetings.

The same matrix can be redrawn as *multidimensional data cube* (Figure-3). To extract the relevant knowledge, one may have to drill down or roll up the data warehouse. Figure-4 shows *drilled data cuboids*.



Legend :

P0, P1, P2, P3 : Number of lab courses attempted by a student in Physics. 0 indicates no attempt any lab in the discipline.

C0, C1, C2, C3, C4, C5 : Number of lab courses attempted by a student in Chemistry

L0, L1, L2, L3, L4 : Number of lab courses attempted by a student in life Sciences.

Figure-3: Multidimensional data cube

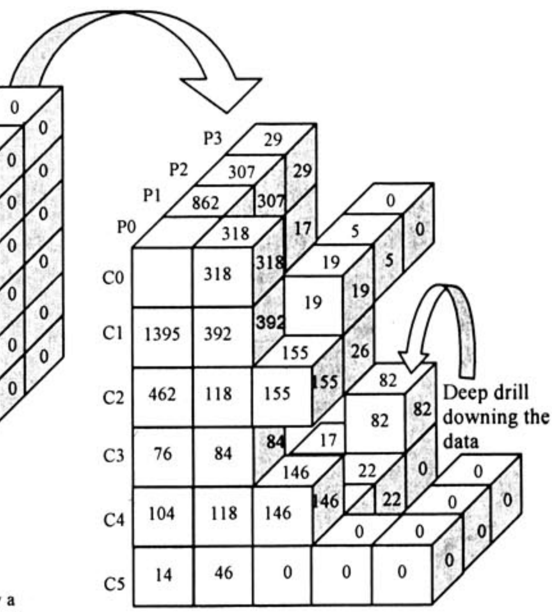


Figure-4: Deep drill in data cube to extract information

The multidimensional data cube – a metaphor for multidimensional data storage – is a classical tool for three-dimensional (3-D) data presentation without confining data to 3-D. In the present case, disciplines of Physics, Chemistry and Life Sciences constitute the three dimensions. In Figure-3, Life Sciences has been taken along *x*–

axis, Chemistry along *y*–axis and Physics along *z*–axis. Each beam of the cube thus represents Chemistry lab courses in combination with Life Science lab courses while Physics option remains constant. Similarly, each pillar represents combination of Physics with Life Sciences while Chemistry option remains constant. To read the cube, for example, the blank cubes in the illustration represent number of students who have not opted any lab course. It is kept blank, as this study is not taking such students into account. Putting matrix in the cubic form facilitates visibility of all three dimensions. This also enables drilling to abridge and secure the important hidden knowledge. To illustrate this, an open cast drilling has been done in the cube at the surface and in the deep. We find that the **most popular** combination opted by the students in multiple disciplines resides deep in the cube (Figure-4). The cuboids are to be removed from all directions to reach the level of visibility. In the present case, the most popular three-discipline combination (3P, 4C and 2LS) has 82 students.

Summary

Data of 6,948 students in various lab courses over a period of eleven years has been mined to discover the performance of students in lab courses of IGNOU's B.Sc. programme. The data was drilled to understand the grading structure of lab courses in all the three disciplines. Table-4 summarizes year-wise and lab course-wise marks of an average student.

Table-4 : Aggregate percentage of marks obtained by students' in lab courses

Year	PHE3L	PHE8L	PH12L	CHE3L	CHE7L	CHE8L	CH11L	CH12L	LSE4L	LSE8L	LS11L	LS14L
1992	70.71	70.38	75.05	72.42	73.25	73.27	76.47	76.36	73.06	76.56		
1993	72.93	70.05	76.56	71.38	69.42	71.34	73.36	75.78	71.11	71.00		
1994	68.61	70.76	72.62	71.77	71.18	72.68	75.37	76.35	74.13	74.04		
1995	69.12	71.06	70.53	72.58	73.74	75.05	75.13	75.09	75.89	73.78		
1996	70.44	71.04	72.18	72.51	71.74	71.70	72.40	73.02	73.14	72.62	60.00	
1997	71.87	71.60	71.97	72.22	72.31	74.41	75.81	74.47	74.78	71.19		
1998	70.37	72.32	72.79	74.01	73.30	75.51	74.28	74.75	74.65	73.09	70.50	60.00
1999	70.10	70.17	72.72	73.23	72.38	72.90	76.70	74.34	73.00	73.01	73.20	64.75
2000	73.04	73.60	68.68	72.86	73.84	71.14	75.92	72.03	74.46	72.87	78.05	73.67
2001	68.73	68.59		72.42	72.34	74.08	75.00		71.97	72.71		58.00
2002	70.75			76.66	70.57	74.00			73.16			
Average	**70.61**	**70.96**	**72.57**	**72.91**	**72.19**	**73.28**	**75.04**	**74.69**	**73.58**	**73.09**	**70.44**	**64.10**

The following knowledge surfaces on analyzing the data in Table-5 :

1. On an average, students obtained 71.95% marks in lab courses. This could be due to two basic reasons that directly influenced the performance, namely, (a) the lab courses are performed under direct guidance of the instructor(s) in F2F situation, and (b) preparation prior to attending the lab sessions and instant feedback.

2. Students obtained more marks in Chemistry lab courses.

3. In a discipline, students obtained higher percentage in PHE12L in physics, CHE12L in chemistry, and LSE4L in life sciences. Among all lab courses, students perform very poorly in third level life sciences courses.

The aggregate percentages of marks obtained in lab courses by students, who have completed B.Sc. programme are given in Table-5.

A comparative analysis of Table-4 and Table-5 reveals that on an average, successful students obtained 74.34% mark, 2.38% more than the average of all students, in lab courses. Moreover students got higher marks in Physics and Life Science (as in Physics, these students got 72.59% against 71.38% and in Life Sciences, 75.61% against 70.30% aggregate marks), whereas in Chemistry, there is not much difference in the two categories of students (73.38% against 73.62%).

Table-5 : Aggregate percentage of marks obtained by students completed B.Sc. programme

Year	PHE3L	PHE8L	PH12L	CHE3L	CHE7L	CHE8L	CH11L	CH12L	LSE4L	LSE8L	LS11L	LS14L
1992	79.25	74.33	76.33	80.50	74.00	60.00	54.00	57.00	75.00	69.00		
1993	81.00	72.50	68.00	79.25	72.00	60.67	76.00	72.00	72.00	82.00		
1994	68.10	67.20	65.00	74.92	66.45	70.40	74.50	77.29	79.29	78.60		
1995	73.78	74.90	76.58	74.86	76.77	77.38	77.93	78.60	79.45	70.69		
1996	69.47	71.42	72.26	74.02	72.14	72.61	71.71	73.43	75.91	74.22		
1997	70.52	71.58	72.23	73.23	72.74	72.63	75.55	73.57	76.20	73.83		
1998	70.30	71.64	74.09	73.49	73.33	74.49	75.16	74.89	74.27	73.16	69.00	
1999	74.43	77.44	79.46	76.17	75.69	76.13	80.30	76.42	74.68	78.15	77.29	76.00
Average	**71.31**	**72.57**	**73.88**	**74.26**	**73.38**	**73.92**	**75.35**	**75.00**	**75.51**	**74.20**	**76.73**	**76.00**

In Life Sciences, successful students obtained better marks in LSE11L and LSE14L courses.

To elaborate performance further, a microanalysis was conducted to obtain range of marks of the students who attempted lab courses. To this end, the following mark ranges were created,

a) R1 – less than 45 marks,

b) R2 – 46 to 60 marks,

c) R3 – 61 to 75 marks, and

d) R4 – above 75 marks.

In Tables-6, 7 and 8, we have given details of the range micro-analysis of Physics, Chemistry and Life Sciences, respectively.

Table-6 : PHYSICS : Year-wise and marks obtained range-wise students performance (in percentage)

Year	PHE3L				PHE8L				PHE12L				Physics Averages			
	<45	46 to 60	61 to 75	>75	<45	46 to 60	61 to 75	>75	<45	46 to 60	61 to 75	>75	<45	46 to 60	61 to 75	>75
1992	4.21	15.79	40.00	40.00	7.69	3.85	34.62	53.85		9.52	23.81	66.67	5.95	9.72	32.81	53.50
1993	1.34	13.42	36.24	48.99	2.27	20.45	38.64	38.64			50.00	50.00	1.81	16.94	41.63	45.88
1994	3.99	19.93	46.01	30.07	2.35	16.47	43.53	37.65	3.45	8.62	50.00	37.93	3.26	15.01	46.51	35.22
1995	6.04	19.51	45.05	29.40	0.75	20.30	38.35	40.60	1.56	25.00	29.69	43.75	2.79	21.60	37.70	37.92
1996	2.37	14.21	51.05	32.37	1.49	17.16	44.78	36.57	0.00	15.22	48.91	35.87	1.93	15.53	48.25	34.94
1997	3.08	10.03	51.41	35.48	3.59	14.37	37.13	44.91	0.00	15.45	43.90	40.65	3.34	13.28	44.15	40.35
1998	3.12	15.81	48.78	32.29	0.45	12.22	47.06	40.27	0.67	5.37	55.03	38.93	1.41	11.13	50.29	37.16
1999	0.79	18.78	44.97	35.45	1.85	18.52	45.06	34.57	1.16	16.28	40.70	41.86	1.27	17.86	43.58	37.29
2000	2.35	13.73	42.35	41.57	1.15	8.05	45.98	44.83		21.05	42.11	36.84	1.75	14.27	43.48	41.08
2001	10.84	17.24	34.48	37.44	5.56	5.56	77.78	11.11					8.20	11.40	56.13	24.27
2002		15.00	40.00	45.00										15.00	40.00	45.00
	3.81	**15.77**	**43.67**	**37.10**	**2.72**	**13.69**	**45.29**	**38.30**	**1.71**	**14.56**	**42.68**	**43.61**	**3.17**	**14.70**	**44.05**	**39.33**

Table-7 : CHEMISTRY : Year-wise and Marks obtained range-wise students performance (in percentage)

Year	CHE3L				CHE7L				CHE8L				CHE11L				CHE12L				Average			
	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76
1992	1.47	8.82	51.47	38.24	3.64		61.82	34.55	2.27	11.36	36.36	50.00		11.76	23.53	64.71		7.14	42.86	50.00	1.48	7.82	43.21	47.50
1993	4.23	15.49	39.91	40.38	2.33	22.09	45.35	30.23		19.18	43.84	36.99	2.13	8.51	42.55	46.81		8.33	33.33	58.33	1.74	14.72	41.00	42.55
1994	3.05	17.65	34.20	45.10	3.13	17.97	40.63	38.28	1.92	17.31	40.38	40.38	1.10	5.49	43.96	49.45		8.86	32.91	58.23	1.84	13.46	38.42	46.29
1995	1.52	13.26	41.30	43.91	1.88	7.50	43.75	46.88	0.67	8.67	37.33	53.33	0.68	6.85	35.62	56.85	0.66	3.97	47.68	47.68	1.08	8.05	41.14	49.73
1996	3.47	10.69	43.21	42.63	3.85	8.97	45.73	41.45	2.83	12.26	41.98	42.92	2.62	13.61	35.08	48.69	0.54	7.61	50.00	41.85	2.66	10.63	43.20	43.51
1997	2.57	10.66	45.40	41.36	2.58	10.33	40.22	46.86	1.10	8.46	40.07	50.37	0.47	7.51	35.21	56.81	0.50	8.54	39.20	51.76	1.45	9.10	40.02	49.43
1998	1.57	7.15	43.49	47.78	2.32	8.41	45.80	43.48	1.27	4.78	43.31	50.64	0.77	6.92	50.38	41.92	0.77	3.86	44.02	51.35	1.34	6.22	45.40	47.03
1999	1.56	6.93	46.45	45.06	2.89	9.75	43.68	43.68	3.15	6.76	46.40	43.69		4.22	34.34	61.45	0.61	6.75	39.26	53.37	1.64	6.88	42.03	49.45
2000	3.27	6.75	47.06	42.92	5.42	4.22	27.11	63.25	7.96	11.50	37.17	43.36		3.90	42.86	53.25	2.70	13.51	43.24	40.54	3.87	7.98	39.49	48.66
2001	5.29	7.65	41.76	45.29	4.41	8.09	41.18	46.32	3.49	8.14	39.53	48.84			50.00	50.00					3.30	5.97	43.12	47.61
2002			41.38	58.62			85.71	14.29			50.00	50.00											59.03	40.97
	2.80	**10.51**	**43.24**	**44.66**	**3.24**	**10.81**	**47.36**	**40.84**	**2.74**	**10.84**	**41.49**	**46.41**	**1.29**	**7.64**	**39.35**	**52.99**	**0.97**	**7.62**	**41.39**	**50.35**	**2.04**	**9.08**	**43.28**	**46.61**

Table 8 : LIFE SCIENCES : Year-wise and marks obtained range-wise students performance (in percentage)

Year	LSE4L				LSE8L				LSE11L				LSE14L				Averages			
	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>76	<45	46 to 60	61 to 75	>75
1992	8.11	8.11	43.24	40.54			50.00	50.00									8.11	8.11	46.62	45.27
1993	2.78	15.28	47.22	34.72		11.76	55.88	32.35									2.78	13.52	51.55	33.54
1994	1.75	10.53	46.49	41.23		13.21	41.51	45.28									1.75	11.87	44.00	43.26
1995	2.86	4.29	38.57	54.29		12.20	37.80	50.00									2.86	8.24	38.19	52.14
1996	0.42	11.34	42.86	45.38	1.64	10.38	45.90	42.08		50.00	50.00						1.03	23.91	46.25	43.73
1997	0.40	8.43	43.37	47.79	1.52	12.69	49.24	36.55									0.96	10.56	46.31	42.17
1998		5.93	46.09	47.98	0.70	5.63	57.04	36.62						100.00			0.70	37.19	51.57	42.30
1999	1.68	10.44	44.78	43.10	1.56	8.85	45.83	43.75		7.41	53.70	38.89		25.00	50.00	25.00	1.62	12.92	48.58	37.68
2000	0.95	7.14	44.29	47.62		8.24	56.47	35.29			26.32	73.68			83.33	16.67	0.95	7.69	52.60	43.32
2001	3.49	8.91	44.57	43.02		4.17	70.83	25.00						100.00			3.49	37.69	57.70	34.01
2002		18.60	44.19	37.21														18.60	44.19	37.21
	2.49	**9.91**	**44.15**	**43.90**	**1.36**	**9.68**	**51.05**	**39.69**		**28.70**	**43.34**	**56.29**		**75.00**	**66.67**	**20.83**	**2.43**	**17.30**	**47.96**	**41.33**

Microanalysis of these tables reveals that:

1. In Physics, 44.05% students fall in R3, 39.33% students in R4, and 3.17% fall in R1. In 1992, 1993 and 2002, more students were in R4 rather than in R3. The gap widened between R3 and R4 in 2001.
2. In Chemistry, R4 is more prominent, where 46.61% students are categorized. R3 is found to be second common (48.28%) and the total of R3 and R4 cover approximately 90% students. This trend is visible in all the academic cycles, but R2 crossed 10% in 1993, 1994 and 1996. R1 remained only with 2.04% students. A very large cluster, i.e., 95%, falls in R3 and R4.

3. In Life Sciences also, R3 is more prominent with 47.96% students, while 41.33% students fall in R4. Here, like Physics and Chemistry, R1 remained very low with only 2.43% students. In 1995, R4 dominated R3 with a difference of more than 14%. Similarly, around 90% students are in R3 and R4.

It is important to note that most students are scoring very high marks in the lab courses. Our analysis shows that, on an average, the score is in-between 70–75%, but the overall average percentage of marks of successful candidate is about 64%. Such a "good performance" can be an indicator of good hands-on training. This also suggests that the students need greater support and attention in theory courses in terms of counseling, evaluation and assessment. It would be worthwhile for the faculty to revisit its strategies and asses their effectiveness in totally. That is, how far the anxieties of distance learners are affecting their self-study, particularly in theory courses, which have some sort of hierarchy.

Conclusion

Drilling the data can unveil specific knowledge base and patterns in educational institutions to support learning activities or to understand the needs of a learner with respect to timely support and targeted action to improve learning. The datasets are converted into a data warehouse by refining and systematizing. In this paper, we have used **Cuboid deep drilling techniques** on multidimensional fields to surface knowledge on the performance of IGNOU science graduates in laboratory courses. We find that :

- About 37% students are active and only one-fourth of them attempted lab courses. Of the active students, the weighted distribution in lab courses, among the disciplines is 29.08% Physics, 55.63% Chemistry and 15.29% Life Sciences.

- 29.52% students opted only Chemistry, 17.24% students only Physics and 5.80% students opted only Life Sciences lab courses. 47.50% students attempted lab courses from more than one discipline.

- The most popular combination among disciplines is three labs of physics, four labs of chemistry and two labs of life sciences. The performance in lab courses is quite satisfying and Chemistry lab courses are more scoring.

These knowledge discoveries are expected to sensitise in-house faculty to introspect, re-examine and reflect on its innovative structure of practical courses with a view to make sure that practicals are used to enhance scientific skills.

References

Berry, M.J. and Linoff, G.S. (2000) Mastering Data Mining – The Art and Science of Customer Relationship Management, Singapore : John Wiley.

Borgelt, Christian and Kruse, Rudolf (2002) Graphical Models : Methods for Data Analysis and Mining, Singapore : John Wiley.

Chandra, S. (1995) Practical Content in Technical Education by Distance Mode, Paper presented at the AAOU VIII Annual Conference, New Delhi, 20-22 Feb, 1, pp 371-85.

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996) Advances in Knowledge Discovery & Data Mining, Cambridge : MIT Press.

Garg, S., Vijayshri and Panda, S. (1992) A Preliminary Study of Student Workload for IGNOU Physics Elective Courses, *Indian Journal of Open Learning,* 1(2), 19-25.

Han, J. and Kamber, M. (2001) Data Mining: Concepts and Techniques, San Francisco: Morgan Kaufmann.

IGNOU Profile (2003), Indira Gandhi National Open University, 10-12, 47.

Jegede, O.J. (2001) Making Science Globally Owned, Accessible and Relevant in a Postmodern World, Staff and Educational Development International, 5(1), 71-89.

Kannan, S., Goel, V. and Raghvan, V.V. (1998) Need for Quality in Assessment - Part I : Concept of Assessment of Science Laboratory Courses in Open Learning System, *Indian Journal of Open Learning*, 7(2), 191-97.

Khan, A.R., Panda, S. and Garg, S. (2000) Learner-Centered Delivery in Distance Education: Challenges and Possibilities, Paper presented at the Open and Distance Learning in New Millennium – ICDE Asian Regional Conference, New Delhi, Nov 2000.

Prabhu, C.S.R. (2002) Data Warehousing: Concepts, Techniques, Products and Applications, New Delhi: Prentice-Hall.

Programme Guide for Bachelor of Science (B.Sc.) Programme, Foundation, Elective and Application-Oriented Courses, Indira Gandhi National Open University, September 2002, 3-5, 10.

Rao, S. (2002) Women In Science Education: Enhancing Access Through Distance Mode, Paper presented at the XV Annual Conference of the Asian Association of Open Universities, New Delhi, Feb 2002.

Vijayshri, Garg, S. and Panda, S. (1998) 'Teaching Science Education at the Indira Gandhi National Open University', in Garg, Suresh et. al. (Eds) *Science and Technology through Open and Distance Education,* SEDS Series, New Delhi : Aravali Books International.

Pankaj Khare is Deputy Director, Indira Gandhi National Open University, New Delhi.

Anurag Saxena is Reader, School of Management Studies, Indira Gandhi National Open University, New Delhi.

Suresh Garg is Professor of Physics and currently Pro-Vice Chancellor, Indira Gandhi National Open University, New Delhi.