STATISTICAL ANALYSIS OF DEMOGRAPHIC DATA AND STUDENT PERFORMANCE IN THE COURSES OF THE BACHELOR'S DEGREE PROGRAM AT THE DEPARTMENT OF COMPUTER, INFORMATICS AND TELECOMMUNICATIONS ENGINEERING - INTERNATIONAL HELLENIC UNIVERSITY WITH ORACLE APEX STATISTICS

ATHANASIOS ANGEIOPLASTIS, ELEFTHERIOS PANAGIS, SOTIRIOS TSAKIRIDIS, ALKIVIADIS TSIMPIRIS, DIMITRIOS VARSAMIS

Department of Computer, Informatics and Telecommunications Engineering, International Hellenic University, Serres, Greece
lefter.panag@gmail.com, aagiop@gmail.com, atsimpiris@ihu.gr, dvarsam@ihu.gr

Undoubtedly, statistical analysis is a powerful tool for understanding, examining and interpreting complex phenomena and data in academic institutions and drawing also valuable conclusions. In the context of this work, the main characteristics and data related to the undergraduate study program of International Hellenic University (IHU) Department of Computer, Informatics and Telecommunications Engineering (CITE) were analyzed, using the tools provided by Oracle APEX. The aim is to explore trends, learner profiles, academic outcomes and other aspects of the curriculum. This work provides important results regarding the demographic data and performance of the students of the IHU department. ORACLE APEX tools helped us to easily transform and clean the data, group them, and display the results online. This analysis provides the impetus to improve the weaknesses identified in the curriculum courses and increase its effectiveness.

1   Introduction

In contemporary society, statistical analysis constitutes a powerful tool for comprehending and interpreting data, enabling the extraction of reliable conclusions. It is a distinguished scientific field that forms the basis for taking correct decisions in various domains, from economics and business to health and education. Regarding the educational process, an increasing number of educational institutions and entities acknowledge the contribution of statistical analysis [Ferrandino, 2016]. This is because the insights it can provide are invaluable and contribute to understanding the behavior and performance of students.

A progressively increasing number of statistical research studies are being conducted by educational institutions at all levels, focusing on the characteristics and daily lives of students, aiming to explain or even predict their performance. Various factors are examined, such as socioeconomic background, gender, or even the use of mobile phones [Lepp, Barkley, Karpinski, 2014], and how these correlates with school or academic performance [Farooq, Chaudhry, Shafiq and Berhanu, 2011].

Such studies [Angeioplastis, Tsimpiris, Varsamis, Baggia, & Leskovar, 2023], served as the catalyst for a statistical analysis of the characteristics of undergraduate students in the Department of Computer, Informatics, and Telecommunications Engineering at IHU. This analysis was tailored to meet the requirements of Greek educational institutions, with a primary focus on the profiles through which students were admitted to this specific undergraduate program.

2   Dataset

The final Oracle APEX database used consists of 9097 records, forming the complete set of 60 courses offered in the undergraduate program of the Department of Computer, Informatics, and Telecommunications Engineering at IHU. It includes all grades and efforts from 383 students, both active and inactive, throughout all years of the program's operation, along with other information crucial for statistical analysis, such as place of origin, gender, method of admission to the department, year of enrollment, and more. The data were processed anonymously with respect to the personal information of all students.

A. Angeioplastis et al.: Statistical Analysis of Demographic Data and Student Performance in the Courses of the Bachelor's Degree Program at the Department of Computer, Informatics and Telecommunications Engineering - International Hellenic University with Oracle APEX

15

3   Oracle Application Express (APEX)

Oracle Application Express (APEX) offers several benefits for creating statistics and managing web-based information systems. Here are some key insights:

1. **Rapid Application Development**: APEX is known for its rapid application development capabilities, allowing users to create rich web applications quickly, based on data within the Oracle database [Austwick, 2013].
2. **Efficiency and Productivity**: APEX enhances developer productivity by providing tools like the Application Synopsis tool, which includes various features for analyzing applications, thereby minimizing efficiency lags [Srinivas, Biswas, & Srinivasan, 2010].
3. **Web-Based Information System Advantages**: The platform offers significant advantages for web-based information systems, such as ease of use, integration of components, fundamental concepts, and a structured framework for development [Liu Hong-xing, 2008].
4. **Low-Code Development**: APEX assists in developing applications with little or no code, making it ideal for creating modern data-driven applications quickly and efficiently [Veerasamy, 2022].
5. **Global Deployment and Customization**: APEX is suited for developing applications that can be deployed globally, respecting linguistic and cultural differences, and allows customization to meet specific application needs [Scott, Buytaert, Cannell, et al., 2011].

4   Methodology

At first, we focused on the process of data extraction from the Department's database from 2019 until current's season data. The next step involved cleaning, processing and encoding the data, having as a main goal the categorization of the data based on their characteristics. Creating a database in the Oracle APEX environment helped to consolidate all the individual tables and avoid redundancy in the data. Oracle APEX tools then helped in creating graphs and descriptive statistical analysis of the results. This was followed by conducting statistical tests, which were performed using the open-source software Jamovi and its libraries [Jr Miller and G. Rupert, 1997], [The jamovi project, 2023]. The statistical tests conducted included the Independent Samples t-test for 2 independent samples [J. Fox and S. Weisberg,

2018] and a continuous variable, followed by Analysis of Variance (ANOVA) [R Core Team, 2013].

5   Results

5.1   Descriptive Statistics

In the Department of Computer, Informatics and Telecommunications Engineering – IHU a total of 383 students have been enrolled from 2019 until the time of the research. Their distribution based on their biological gender is 60 females (15.67%) and 323 males (84.33%), indicating a overwhelming preference of males in choosing this specific department.
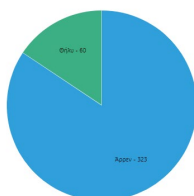

Figure 1: Distribution of biological gender of all students
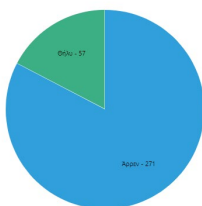**Θήλυ (Green color) → females**
Source: Own

Figure 2: Distribution of biological gender of all <u>active</u> students
Άρρεν (Blue color) → males
Source: Own

A. Angeioplastis et al.: Statistical Analysis of Demographic Data and Student Performance in the Courses of the Bachelor's Degree Program at the Department of Computer, Informatics and Telecommunications Engineering - International Hellenic University with Oracle APEX

17

The active students in the department are 324, meaning that out of the total 383, 59 students are either inactive or have left the department (15.40%). As evident in Figure 2, the number of women among active students is 55 (16.97%), while men are 269 (83.03%). Almost identical percentages compared to the total number of admissions indicate that the rate at which students remain inactive or leave is the same for both men and women.

The demographic data regarding the students' regions of origin is of particular interest. Knowing the data for a total of 318 students, the results shown in Figure 3 have been extracted. Specifically, students coming from the Thessaloniki area account for 27.04%, and those from the Serres region constitute 23.90%. Collectively, students from these two areas make up more than half of the department's students, with a percentage of 50.94%.

Figure 3: Distribution of students' place of origin
Source: Own

Furthermore, it's worth noting that a total of 231 students, accounting for 72.64%, originate from the Central and Eastern Macedonia regions. This indicates that the department attracts students from the same or surrounding Regional Units.

The method of student admission to the department was also examined. From the analysis of data for all 383 entrants, whether active students or not, the following results emerged: 78% of students were admitted to the department through the General High School nationwide exams, 12% through nationwide exams for Vocational High Schools, 5% through a 10% special admission process, while those

admitted through placement exams or other categories (5%, athletes, foreigners, minorities, transfers) constitute 3% each.
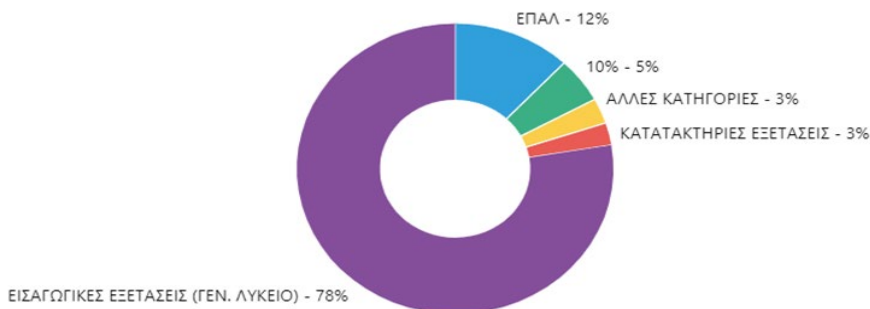

Figure4: Percenages of students according to the methon of admission
Source: Own

- **ΕΙΣΑΓΩΓΙΚΕΣ ΕΞΕΑΤΣΕΙΣ (ΓΕΝ. ΛΥΚΕΙΟ) → General High School**
- **ΕΠΑΛ → Vocational High School**
- **ΚΑΤΑΤΑΚΤΗΡΙΕΣ ΕΞΕΤΑΣΕΙΣ → Placement exams**
- **ΑΛΛΕΣ ΚΑΤΗΓΟΡΙΕΣ → Other categories**

5.2   Students' Performance

5.2.1   Success percentages for academic courses

The analysis of student performance constitutes the most significant part of this study. In Figure 4, the success rates of students for each course offered in the department are detailed. With blue color we can see the total number of students that have passed the course, while with green we can see the number of students that have passed the course with their first attempt.
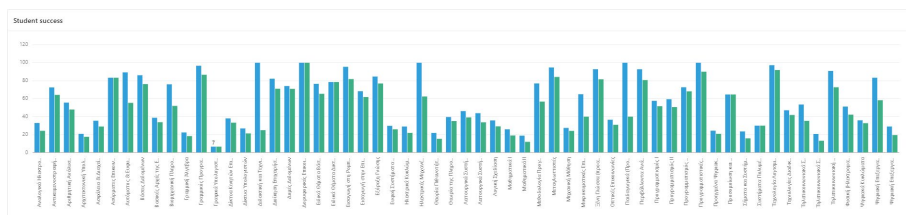
A. Angeioplastis et al.: Statistical Analysis of Demographic Data and Student Performance in the Courses
of the Bachelor's Degree Program at the Department of Computer, Informatics and Telecommunications
Engineering - International Hellenic University with Oracle APEX   19

Figure 5: Students' success percentages for each academic course
Source: Own

The overall success rates of students in each course were examined, as well as the percentage of students who manage to pass each course "on the first attempt", meaning during the semester in which the course is scheduled in the curriculum. The results highlight the courses with the lowest and highest success rates, where deeper reasons for students' failure to succeed in certain courses should be investigated in a future research.

5.2.2   Correlation between students' grades and biological gender

It was deemed appropriate to examine whether there is a correlation between the biological gender of students and the grades they achieve. To draw conclusions, the independent samples t-test was selected. This is a statistical test used to examine whether there are statistically significant differences between two independent groups or two samples measured for a continuous variable.

Table 1: Statistics of biological gender

	Biological gender	Sample size	Mean	Standard deviation
Grade	Male	7306	3.93	2.79
	Females	1686	4.04	2.72

In our case, the independent groups consist of biological genders (Male, Female), and the continuous random variable is the grade, which takes values in the continuous interval [0,10]. The results are shown in Tables 1 and 2. The field F refers to the value of the F-test conducted using the Levene method [Schultz, 1985]. The Levene F-test is a statistical test used to assess whether the variances (differences)

among multiple groups are statistically significant. The mean grade for females is 4.04, slightly higher than that of males, which is 3.94, while the standard deviation differs by approximately 0.07 between the two groups. The null hypothesis is defined as the assumption that there are no statistically significant differences in the means of grades between males and females with a significance level of 5%.

Table 2: Independent samples t-test

		F	Sig.	t	df	Sig.(2-tailed)
Grade	Equal variances assumed	3.06	0.08	-1.36	8990	0.174
	Equal variances not assumed		-1.38	1.38	2570.15	0.167

Because the significance level of the test (referred to as p-value and denoted as Sig. in the tables) is 0.17 > 0.05, this means that there is insufficient evidence to reject the null hypothesis. Therefore, there is no statistically significant difference in the mean grades between males and females.

5.2.3   Correlation between students' grades and method of admission

It was considered important to investigate whether the method students were admitted to the Department of Computer, Informatics and Telecommunications - IHU affects their performance. The statistical test used for drawing conclusions was the Analysis of Variance (ANOVA) with the Bonferroni method [Nakagawa, 2004]. The ANOVA (Analysis of Variance) statistical test allows the assessment among multiple groups simultaneously and is particularly useful when determining if there are significant differences between many conditions or experiments. The Bonferroni method is an adjustment method for significance levels in multiple comparisons between groups. If the p-value of each comparison is less than the corresponding adjusted significance level, then the null hypothesis is rejected, and it is concluded that there is statistical significance.

A. Angeioplastis et al.: Statistical Analysis of Demographic Data and Student Performance in the Courses
of the Bachelor's Degree Program at the Department of Computer, Informatics and Telecommunications
Engineering - International Hellenic University with Oracle APEX

21

Table 3: Descriptives statistics for the comparison between grade - method of admission

	N	Mean	Standard deviation	Typical error	95% Confidence Interval for the Mean Value		Minimum	Maximum
					Lower Bound	Upper Bound		
GHS	7489	4.00	2.76	0.03	3.94	4.06	0.00	10.00
VHS	854	3.36	2.75	0.09	3.18	3.55	0.0000	10.00
GHS 10%	366	4.04	2.90	0.15	3.74	4.34	0.0000	10.00
Other	283	4.37	3.06	0.18	4.01	4.73	0.0000	10.00
Total	8992	3.95	2.78	0.02	3.90	4.01	0.0000	10.00

Students were divided in 4 categories. These categories are:

- GHS → General High school nationwide exams
- VHS → Vocational High school nationwide exams
- GHS 10% → General High school with the method of 10%
- Other → Other categories included (5%, athletes, foreigners, minorities, transfers)

Table 3 presents basic statistics for these four categories. Of significant importance is the "Mean" field, displaying the average grades of students in each admission category. It is observed that the mean grade for students coming from vocational high schools is approximately 3.37, noticeably lower than the other average scores.

This is confirmed by the Bonferroni test in Table 4. The null hypothesis of the test is that there are no statistically significant differences in students' grades based on their method of admission, with a significance level of 5%. It is observed that the p-value < 0.001 for all comparisons in the "Vocational High School" category compared to the other categories. This means that the null hypothesis "there are no statistically significant differences in students' grades based on their method of admission" is rejected, and the conclusion is drawn that the grades of students coming from vocational high schools are lower compared to the grades of other students. In the remaining categories, the p-value is greater than 0.05, indicating that there are no statistically significant differences in the grades of students between these specific categories.

Table 4: Bonferroni method for the correlation between grade - method of admission

(I) Method of admission	(J) Method of admission	Διαφορά Μέσων τιμών (I-J)	Τυπικό σφάλμα	Sig.	95% Διάστημα Εμπιστοσύνης Κάτω φράγμα	Άνω φράγμα
GHS	VHS	0.6352	0.1003	<0.001	0.3704	0.9000
	GHS 10%	-0.0391	0.1487	1.000	-0.4316	0.3532
	Other	-0.3692	0.1682	0.169	-0.8131	0.0747
VHS	GHS	-0.6352	0.1003	<0.001	-0.9000	-0.3704
	GHS 10%	-0.6744	0.1735	<0.001	-1.1324	-0.2164
	Other	-1.0044	0.1905	<0.001	-1.5073	-0.5016
GHS 10%	GHS	0.0391	0.1487	1.000	-0.3532	0.4316
	VHS	0.6744	0.1735	<0.001	0.2164	1.1324
	Other	-0.3300	0.2199	0.801	-0.9103	0.2502
Other	GHS	0.3692	0.1682	0.169	-0.0747	0.8131
	VHS	1.0044	0.1905	<0.001	0.5016	1.5073
	GHS 10%	0.3300	0.2199	0.801	-0.2502	0.9103

6   Conclusion

In the Department of Computer, Informatics and Telecommunications Engineering at IHU., there are a total of 324 active students, with the distribution of their biological gender being 55 females (15.67%) and 269 males (83.03%). The majority of students come from the regions of Serres and Thessaloniki (50.94%), while the department mainly attracts students from the central and eastern Macedonia regions with a total percentage of 72.64%. Most of students have been admitted to the department through nationwide exams (76.76%). There is no statistically significant difference in the performance of students based on their biological gender. On the contrary, it is statistically confirmed that students coming from vocational high schools achieve significantly lower performance compared to other students. Further research on the department's statistics, as well as the development of models predicting student behavior, will be presented in future work.

A. Angeioplastis et al.: Statistical Analysis of Demographic Data and Student Performance in the Courses of the Bachelor's Degree Program at the Department of Computer, Informatics and Telecommunications Engineering - International Hellenic University with Oracle APEX

23

References

Angeioplastis, A., Tsimpiris, A., Varsamis, D., Baggia, A., & Leskovar, R. (2023, March). Data analysis of student's performance on "Database" course with Oracle Apex statistics. In Conference on Organizational Science Development Interdisciplinarity Counts (p. 5).

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2014). The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. Computers in Human Behavior.

Austwick, T. (2013). Using Oracle Apex securely. Network Security, 2013, 19-20.

Schultz, B. B. (1985). Levene's test for relative variation. Systematic Zoology, 34(4), 449-456.

Hong-xing, L. (2008). Developing web-based information system using Oracle APEX. Computer Knowledge and Technology.

Ferrandino, J. A. (2016). Student achievement in undergraduate statistics: The value of encouraging failure. Journal of the Scholarship of Teaching and Learning, 16(6), 1-18.

Fox, J., & Weisberg, S. (2018). An R companion to applied regression. Sage Publications.

Miller Jr, M., & Rupert, G. (1997). Beyond ANOVA: Basics of applied statistics. John Wiley & Sons Inc, CRC Press.

Farooq, M. S., Chaudhry, A. H., Shafiq, M., & Berhanu, G. (2011). Factors affecting students' quality of academic performance: A case of secondary school level. Journal of Quality and Technology Management, 7(2), 1-14.

Scott, J., Buytaert, N., Cannell, K., D'Souza, M., Gault, D., Gielis, D., Hartman, R., Kubicek, D., Mattamal, R., McGhan, D., Mignault, F., Petrus, T., Rimblas, J., & Ruepprich, C. (2011). Expert Oracle Application Express.

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. Behavioral Ecology, 15(6), 1044-1045.

Srinivas, S., Biswas, A., & Srinivasan, J. (2010). An application synopsis tool for database applications developed using oracle application express. Proceedings of the..., 113-118.

The jamovi project. (2023). jamovi (Version 2.3) [Computer software]. https://www.jamovi.org/.

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Veerasamy, D. (2022). A pragmatic step to deploy low-code web apps on apex cloud services for emerging business assistance. i-manager's Journal on Software Engineering.