Exploratory Data Analysis (EDA) of the Performance of Philippine Higher Education Institutions in the Electronics Engineering Licensure Examinations

Elaine R. Rivera, PECE, MS ME, Professor, Associate Dean
ECE Department, School of Engineering and Architecture
Saint Louis University, Baguio City, Philippines.
elaine_r_rivera@yahoo.com

***Abstract* - A significant indicator of quality in engineering education is the results of licensure examinations. Examination results as well as the performance results of Higher Education Institutions (HEIs) are made available through the Professional Regulations Commission (PRC) website and other publications within a week that examinations are administered. Through the years PRC has in its databank an increasing accumulation of large data of examination results that need to be statistically and graphically summarized to unravel hidden patterns which could be significant basis for important decisions, strategies and plans. The objective of this paper is to give summary representations of the performance of HEIs in the Electronics Engineering Licensure examinations from the year 2012 to 2016. Results may be used as basis for strategic decisions such as but not limited to: evaluation or reformation of the monitoring of performance of HEIs in the licensure examinations, and the restructuring of accreditation requirements pertaining to the performance of HEIs in Electronics Engineering licensure examinations. This study made use of the descriptive research methodology using the data mining tasks which are: Data Preparation, Descriptive Data Modeling, Model Evaluation and Output Presentation. Results show that in all ten examinations from the year 2012-2016, the average passing rates of majority of HEIs are less than the national passing rates. Both the national passing rates and the average passing rates of HEIs have been fluctuating but mostly hovering between 30 to 40% over the ten examinations. Moreover, the most frequent occurring passing rate of HEIs with less than 25 takers is 0%. There is then a need to monitor and investigate HEIs with less than 25 takers to probe on the causes of low passing rates.**

Keywords: **Exploratory Data Analysis, Electronics Engineering Licensure Examination Results, Licensure Performance**

I. INTRODUCTION

Engineering Education primary mandate and responsibility is to supply aptly skilled graduates to the workplace. A significant indicator of quality in engineering education is the results of licensure examinations. This indicator of outcomes is particularly substantial because it is a quantitative measure and mostly is directly related to the knowledge and skills of an institution's graduates. The licensure exams serve as a uniform measurement device that attempts to maintain curriculum consistency among academic programs between institutions, and consequently licensing boards have a high degree of control over program length and curriculum content. A highly credible performance of an institution's graduates in licensure examinations over the last five years is one major criteria for accreditation purposes.

In the Philippines there is a strong coupling between licensing through the Professional Regulations Commission (PRC), and the engineering curricula as approved by the Commission on Higher Education (CHED). Examination Results as well as the performance results of Engineering Schools are made available within a week that examinations are administered. This is in accordance with R.A. 8981 otherwise known as PRC Modernization Act of 2000 Section 7(m) "To monitor the performance of schools in licensure examinations and publish the results thereof in a newspaper of national circulation"

The PRC website also releases performance of schools for that particular licensure examination. The performances of the schools are in terms of the passing percentage rates of first time takers, repeaters, and the overall passing percentage rates [1]. These data are being used by Higher Education Institutions (HEIs) in monitoring the licensure examination performance of their respective graduates. However, although performances of institutions in terms of passing percentages are given, summary statistics and visual generalizations are not included which could have revealed HEIs' performance patterns.

The objective of this paper is to summarize the main characteristics, mostly through visual methods, of the performance of HEIs in the Electronics Engineering (ECE) Licensure Examinations over the five year period 2012-2016. Graphical summaries are almost universally sought to augment algebraic summaries because graphics can portray numerous data values simultaneously, while algebraic summaries often sum over important attributes of the data or fail to suggest important patterns. Graphical displays and graphical analysis is central to an Exploratory Data Analysis (EDA) [2]. The role of graphics in EDA is summed up by the statement: "The greatest value of a picture is when it forces us to notice what we never expected to see" [3].

II. METHODS

This paper made use of the descriptive research methodology. In particular, it made use the data mining tasks which are: Data Preparation, Descriptive Data Modeling and Output Presentation. These tasks served as the major steps in the process of the study. It also followed the Input- Process-Output correlation for a systematic flow.

Inputs of the study

The Inputs of the study are the data on the PRC Engineering Licensure Ratings of all Higher Education Institutions in the Philippines offering Electronics Engineering. The data that were released by PRC on the performance of schools in the licensure exams reflect the following: a) name of school, b) number of first time takers, repeaters, and their combined total, and c) ratings of passers for first timers, repeaters and their combined ratings. The data that was taken were for the two examinations per year for the period 2012 to 2016.

Process

There were three major steps in the process of this part of the study: Data Preparation, Descriptive Data Modeling using Exploratory Data Analysis, and Model Evaluation. The first step: *Data Preparation* necessitated Data Selection such that only the data on the performance of HEIs in the licensure examinations for ECE were considered covering the two examinations per year for five years from 2012 to 2016. The data in the PRC website was in PDF format which required data transformation into Excel format to be able to use its Data Analysis feature.

The second step was the *Descriptive Data Modeling using Exploratory Data Analysis* where the data is presented through Descriptive Statistics e.g.: mean, median, mode, standard deviation; and Data visualization using scatterplots and histograms. These Exploratory Data Techniques were used because they are most appropriate when summarizing big data and in searching for hidden patterns [4]. To explore and mine the data for possible hidden representations, the researcher categorized the HEIs in terms of whether their passing rate is above or below the national passing rates for each of the ten examinations. These additional data were used as the groupings in the scatterplots.

III. RESULTS

The results of the EDA of the performance of HEIs in the ECE Licensure examinations are presented in two major parts: the Descriptive Statistics and the Data Visualizations using histograms, boxplots, and scatterplots.

Descriptive Statistics Table

The performance of ECE HEIs in the licensure exams as revealed in Table 1 shows that the highest National Passing Rate for ECE was in April 2012 while the lowest was in September 2014. Out of the ten examinations, eight have rates in the 30 to 40% range. It is alarmingly to note that the mean and median of the passing rates of ECE HEIs are lower than the national rates. With 186-208 participating HEIs it is equally alarming that the most frequently occurring passing rate which is the mode, is 0%. The standard deviation values also reveal that passing rates are spread out from the mean.

Table 1. Passing Rates of Electronics Engineering (ECE) HEIs

ECE Passing Percentage	Apr 2012	Oct 2012	Apr 2013	Dec 2013	Mar 2014	Sept 2014	Apr 2015	Oct 2015	Apr 2016	Oct 2016
National Passing	53.53	51.57	37.21	34.5	35.24	31.59	34.95	39.94	36.95	40.36
Mean	47.8	36.8	37.2	23.8	25.9	23.6	27.5	28.7	29.9	29.1
Median	50.0	36.4	33.3	20.6	23.8	19.0	25	26.7	26.0	25.4
Mode	0	0	0	0	0	0	0	0	0	0
Standard Deviation	30.7	28.3	27.4	22.5	25.0	23.9	26.1	24.1	23.8	25.0
Skewness	0.0	0.2	0.6	0.8	0.8	1.2	1.0	0.5	0.6	0.8
Minimum	0	0	0	0	0	0	0	0	0	0
Maximum	100	100	100	100	100	100	100	100	100	100
Count	190	203	192	200	194	197	186	197	199	208

Data Visualizations

To augment the limitations of the descriptive Statistics table in exploring data, data visualization is needed. Figure 1 shows the scatterplots of the Passing Rates against the Total Takers of ECE HEIs for the ten examinations from 2012-2016. Reflected in the scatterplots are the grouping of the ECE HEIs in terms of whether their rates for each exam is above or below the national rates.

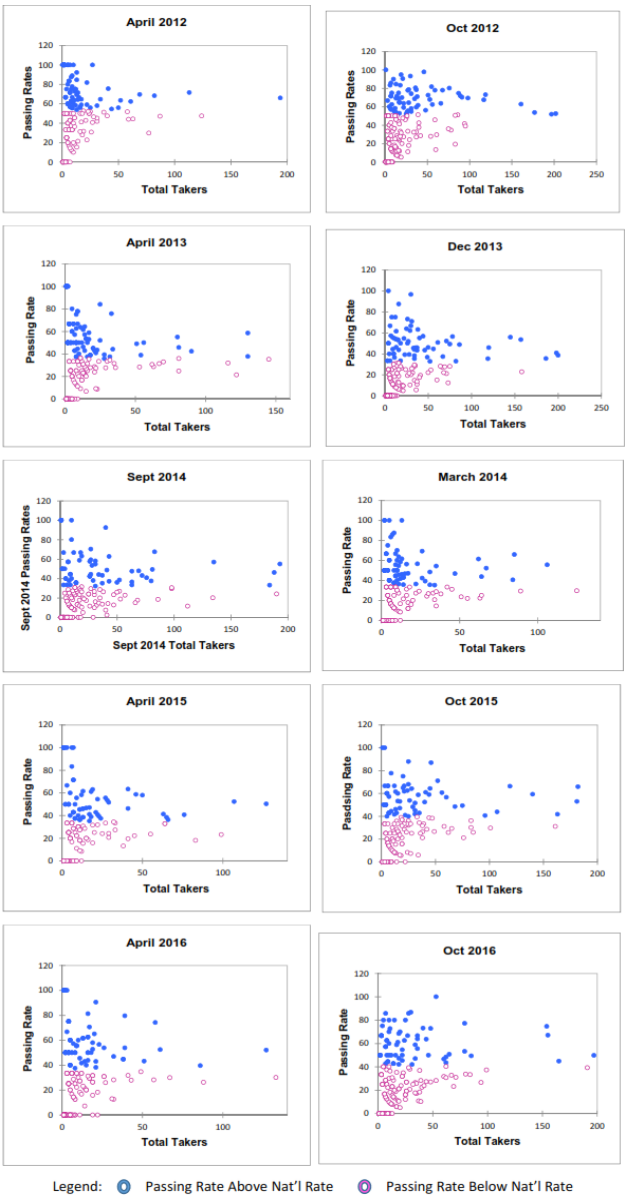

Figure 1. Scatterplots of the Passing Rates against the Total Takers of ECE HEIs for the ten examinations from 2012-2016

The ten scatterplots show similar patterns. There are clusterings in the region with less than 50 takers and in the region below the 30% passing rates. The scatterplots also show that ECE HEIs with 0% passing rates are those with less than 50 total takers.

Figure 2 shows the histograms of the passing rates of ECE HEIs over the ten examinations from 2012 to 2016. In nine of the ten examinations, the passing rate with the most frequency ranges from 0 to 10%. The shape of the distribution in almost all the examinations is skewed to the right which means that only few HEIs have passing rates above the mean and the median rates.

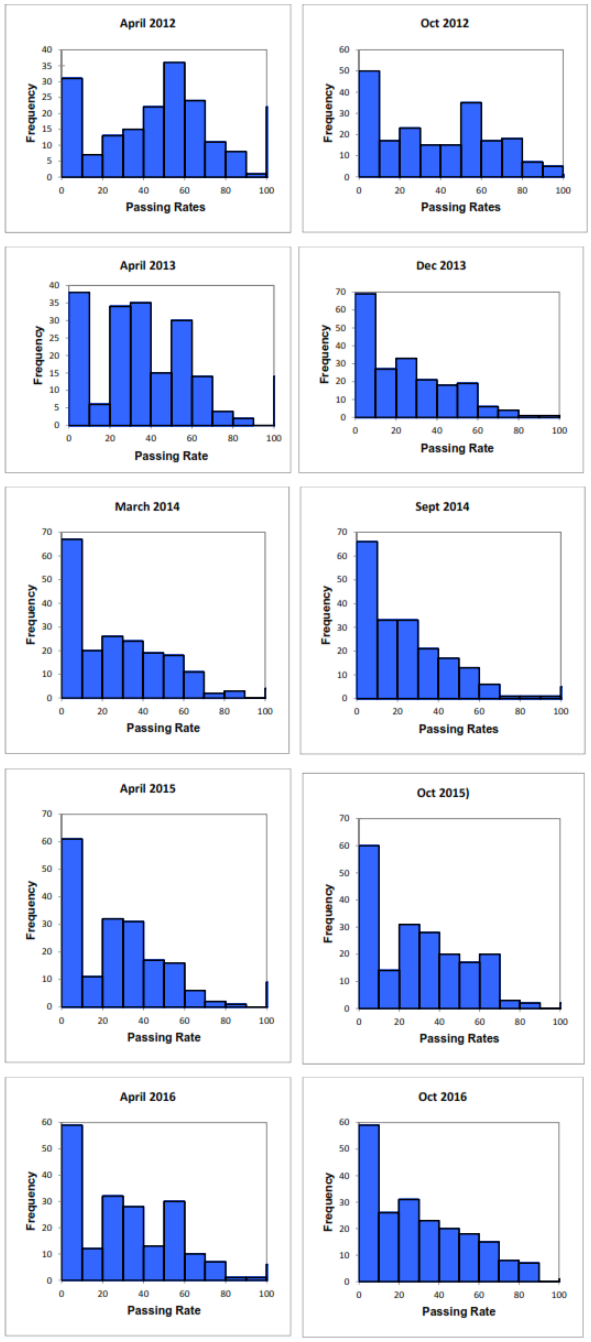

Figure 2. Histograms of the Passing Rates of ECE HEIs for the ten examinations from 2012-2016

IV. DISCUSSIONS

Exploratory data analysis is either non-graphical or graphical. Non-graphical methods generally involve calculation of summary statistics, while graphical methods summarize the data in a diagrammatic or pictorial way. Although there are guidelines about which EDA techniques are useful in what circumstances, there is an important degree of looseness and art to EDA. What is important is the hidden patterns that are revealed through EDA.

The use of Descriptive statistics to summarize and describe the overall engineering licensure examination ratings of HEIs for the period of 2012-2016 proved to be insufficient in

exposing hidden patterns in the data. The disadvantage of Descriptive Statistics tables is that people are not very good at looking at a column of numbers and then determining important characteristics of the data. Looking at numbers is perceived to be tedious, boring, and/or overwhelming [5]. The Descriptive statistics table for ECE over the period of 2012-2016 showed that the mean, and median are below the national passing rates. Also, the mode was 0% in almost all the examinations implying that many HEIs have 0% as their passing rate. These findings using descriptive statistics did not show hidden patterns such as whether the passing rates vary when the number of takers of an HEI is below or more than a specified number. It also failed to show the distribution of the number or frequency of HEIs with respect to the passing rates. Graphical Exploratory data analysis techniques were further done as an aid in this situation.

V. CONCLUSION

Based on the findings of the study, the following conclusions were obtained:

a. The Descriptive Statistics on the overall performance of ECE HEIs in the licensure examinations for the period 2012 to 2016 did not show hidden patterns in the data but gave summaries as follows:
   a1. Majority of performance of HEIs are lower than the national passing.
   a2. The most frequent passing rate of HEIs is 0%.

b. Data Visualizations on the performance of ECE HEIs in the Licensure Examinations for the period 2012 to 2016 revealed hidden patterns which are as follows:
   b1. There are clusterings in the region with less than 50 takers and in the region below the 30% passing rates.
   b2. The frequency of distribution of the passing rates are mostly skewed to the right.

VI. RECOMMENDATIONS

The researcher hopes that the output of this study would be used by School Administrators, Accrediting Bodies, CHED and other concerned agencies, in the monitoring and evaluation of Engineering HEIs and the formulation of policies and guidelines to improve the performance of HEIS in licensure examinations. Investigations on HEIs with 1- 25 takers and with passing rates below the national passing rates are highly recommended. This simple study revealed the need to reform the monitoring and the restructuring of accreditation requirements pertaining to the performance of HEIs in engineering licensure examinations so that an increase in the weight of an HEI's performance in the licensure examinations be considered in the evaluation criteria.

VII. FUTURE WORK

The researcher is currently updating the data on this study to include the results of licensure examinations for 2017. This is with the intention of conducting a thorough Clustering of HEIs using different algorithms such as k-means clustering and agglomerative hierarchical clustering. Moreover, the researcher is also working on a regression analysis for predictive models for the performance of HEIs in the Engineering licensure examinations for ECE as well as for CHE, CE, EE and ME programs.

REFERENCES

[1] Philippine Regulations Commission Website

[2] Bell, J. (2001). Visualising Multilevel Models: The Initial Analysis of Data. Retrieved from: http://www.cambridgeassessment.org.uk/images/109679-visualising-multilevel-models-the-initial-analysis-of-data.pdf

[3] Tukey, J. (1993). Exploratory Data Analysis: Past, Present, and Future. Retrieved from: http://www.dtic.mil/dtic/tr/fulltext/u2/a266775.pdf

[4] Han, J., Kamber, M. (2006), Data Mining: Concepts and Techniques, Morgan- Kaufmann Academic Press,San Francisco. Retrieved from: http://web.engr.illinois.edu/~hanj /bk2/toc.pdf

[5] Chambers, J. M., (1983). Graphical Methods for Data Analysis. Belmont, CA: Wadsworth. doi: 10.1080/02664768400000024

[6] Chatfield, C. (1985). The initial examination of data. *Journal of the Royal Statistical Society*, Series A., 148, 214-53. Retrieved from: http://www.jstor.org/stable /2981969?seq=1#page_scan_tab_contents