Classifying the Public Universities of Colombia than Offer Programs of Industrial Engineering Based on State Tests Using k-Means Clustering

Mendoza Mendoza Adel Alfonso
3i+d Research Group
Universidad del Atlántico, Colombia

Herrera Acosta Roberto José
Gestión de la Calidad Research Group
Universidad del Atlántico, Colombia

Cabarcas Reyes Juan Carlos
3i+d Research Group
Universidad del Atlántico, Colombia

Abstract

The SABER PRO tests are standardized tests performed by students in their senior year of academic programs in Colombia and are used as indicators of the strengths and weaknesses of university education processes. This paper proposes a method for classifying public universities in Colombia that offer university programs in industrial engineering using standardized SABERPRO tests through cluster analysis with K-means. The research made it possible to classify the universities into three groups.

Keywords: Pruebas SABERPRO, K-means, clustering
1. Introduction

The concept of student performance is quite complex, dealing with different aspects, both individual and collective, which determines learning and academic work. In this sense, previous studies have explored several factors such as, for example, general academic performance, aptitude, previous knowledge of mathematics, gender, age, motivation, effort, among other causes that are associated with the performance of students in university level courses, [1], [2], [3], [4], which implies ensuring that there are many characteristic factors in a learning process that as long as they are identified with a reasonable index of precision can be managed and that they can be an indicator of the quality of university academic programs. For this reason, the State Examination on the Quality of Higher Education (SABERPRO) was created in Colombia as a measure of the academic performance of students in higher education; test that has been regulated and and it is an additional requirement for obtaining a degree. [5]. According to Ley 115 of 1994 [6], the education system in Colombia is made up of preschool education levels, basic education, which includes the primary and secondary cycles, secondary education, and higher education, noting that for the State it is essential to measure quality through a series of tests that have been implemented to monitor education in these stages of training [7] and that serve as the basis for decision-making regarding policies to improve the system. For this reason, in recent years, most universities have been investing in physical infrastructure, as well as in technological resources, and improvements have been made in all internal processes, which has been reflected in the increase in accredited programs, but this is no guarantee of greater progress in terms of educational quality. Therefore, it is not possible to establish whether the resources allocated to education are being used efficiently, or if the current amount of these resources is not sufficient to achieve optimal levels of quality, hence the need to evaluate the current performance of the education system. [8].

Due to the importance of the academic programs of industrial engineering in the social and business context of any country, since it is a basic pillar in the improvement and implementation of productive processes, this research provides answers to the following questions: How can the industrial engineering programs of public universities in Colombia be classified? How can the classification of industrial engineering programs of public universities in Colombia be visually analyzed based on the results of state tests? In accordance with the above, the general objective of this research is to present a methodology for classifying industrial engineering programs, identifying the existing relationships between the results of state tests by specific and generic competencies.
Clustering goal to extract natural groupings hidden in data, so that the observations belonging to a group are very close to each other and away from the observations located in another cluster [9]. These algorithms can be classified in partitions, based on density, based on networks and hierarchies. Of them, the most used is the partitioning algorithm K-means for its simplicity and effectiveness. [10]. Partitioning algorithms such as k-means iteratively refine a set of k clusters, density-based algorithms are capable of generating arbitrary sized clusters and dealing with extreme values, network-based algorithms reduce clustering space to cells within a network, allowing efficient grouping of large data sets and hierarchical algorithms can be agglomeration or divisive, in the repeatedly agglomerated form two clusters are merged, while in the repeatedly divisive form one cluster is divided into two [11]. By now, a variety of clustering algorithms k-means has been applied to educational field [12].

2. Methodology

The aim of the research is to classify the public universities that offer the industrial engineering program, based on the results of the SABERPRO state tests, in addition to generating information visualization processes that allow rapid analysis and interpretation of the results. This classification was carried out using the k-means grouping technique using the SPSS software. The available database of the Colombian Institute for the Evaluation of Education (ICFES) was used as the primary source of information. ICFES is the entity in charge of offering the service of evaluation of education at all levels as support to the Ministry of Education in carrying out the state examinations. [13]. The analysis was carried out with the results of the SABERPRO test for the years 2016 and 2017, due to the fact that from 2016 the Scale of Results of the State Examination on the Quality of Higher Education was changed to give effect to with Resolution 892 of 2015. [14]. Table 1 shows the institutions considered in the study.

Table 1. Universities considered in the study

<table>
<thead>
<tr>
<th>Universidad del Atlántico</th>
<th>Universidad de Pamplona</th>
<th>Univ. Nacional Abierta y a Distancia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universidad de Antioquia</td>
<td>Universidad del Magdalena</td>
<td>Univ. Nacional sede Medellín</td>
</tr>
<tr>
<td>Universidad Distrital de Bogotá</td>
<td>Universidad del Valle</td>
<td>Univ. Nacional sede Bogotá</td>
</tr>
<tr>
<td>Universidad de Córdoba</td>
<td>Univ. Francisco de Paula Santander</td>
<td>Univ. Tecnológica de Pereira</td>
</tr>
<tr>
<td>Universidad de Cundinamarca</td>
<td>Universidad Industrial de Santander</td>
<td></td>
</tr>
<tr>
<td>Universidad de La Guajira</td>
<td>Universidad Militar</td>
<td></td>
</tr>
</tbody>
</table>

The ICFES in the SABERPRO tests examines generic competencies (which are considered common to all careers) and specific competencies in each discipline, in the case of industrial engineering programs these competencies are as follows.

Specific competencies: Formulation of engineering projects (SC1); Mathematics and statistics (SC2); Design of production and logistics systems (SC3).
Generic competencies: Written communication (GC1); Quantitative reasoning (GC2); Critical reading (GC3); Citizen competencies (GC4); English or Second language (GC5). Table 2 shows the values of the variables for each university.

### Table 2. Values of the variables for each university

<table>
<thead>
<tr>
<th>University</th>
<th>SC1</th>
<th>SC2</th>
<th>SC3</th>
<th>SC4</th>
<th>GC1</th>
<th>GC2</th>
<th>GC3</th>
<th>GC4</th>
<th>GC5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlántico</td>
<td>161</td>
<td>156</td>
<td>164</td>
<td>158</td>
<td>179</td>
<td>167</td>
<td>159</td>
<td>177</td>
<td></td>
</tr>
<tr>
<td>Antioquia</td>
<td>173</td>
<td>163</td>
<td>175</td>
<td>164</td>
<td>189</td>
<td>177</td>
<td>170</td>
<td>173</td>
<td></td>
</tr>
<tr>
<td>Distrital de Bogotá</td>
<td>179</td>
<td>182</td>
<td>186</td>
<td>164</td>
<td>200</td>
<td>186</td>
<td>178</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>Córdoba</td>
<td>157</td>
<td>155</td>
<td>161</td>
<td>158</td>
<td>179</td>
<td>163</td>
<td>156</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>Cundinamarca</td>
<td>154</td>
<td>143</td>
<td>146</td>
<td>150</td>
<td>165</td>
<td>153</td>
<td>146</td>
<td>149</td>
<td></td>
</tr>
<tr>
<td>Guajira</td>
<td>141</td>
<td>133</td>
<td>135</td>
<td>148</td>
<td>150</td>
<td>139</td>
<td>140</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td>Pamplona</td>
<td>149</td>
<td>140</td>
<td>144</td>
<td>149</td>
<td>162</td>
<td>145</td>
<td>138</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>Magdalena</td>
<td>158</td>
<td>150</td>
<td>157</td>
<td>153</td>
<td>173</td>
<td>161</td>
<td>153</td>
<td>154</td>
<td></td>
</tr>
<tr>
<td>Valle</td>
<td>175</td>
<td>165</td>
<td>181</td>
<td>161</td>
<td>192</td>
<td>178</td>
<td>171</td>
<td>182</td>
<td></td>
</tr>
<tr>
<td>Francisco de Paula Santander</td>
<td>158</td>
<td>152</td>
<td>154</td>
<td>156</td>
<td>172</td>
<td>157</td>
<td>152</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td>Industrial de Santander</td>
<td>166</td>
<td>160</td>
<td>170</td>
<td>157</td>
<td>185</td>
<td>168</td>
<td>159</td>
<td>169</td>
<td></td>
</tr>
<tr>
<td>Militar</td>
<td>158</td>
<td>153</td>
<td>155</td>
<td>162</td>
<td>173</td>
<td>160</td>
<td>153</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>Nacional abierta y a distancia</td>
<td>144</td>
<td>135</td>
<td>134</td>
<td>147</td>
<td>151</td>
<td>144</td>
<td>139</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td>Nacional sede Manizales</td>
<td>170</td>
<td>162</td>
<td>168</td>
<td>162</td>
<td>186</td>
<td>173</td>
<td>166</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>Nacional sede Medellín</td>
<td>173</td>
<td>176</td>
<td>180</td>
<td>166</td>
<td>197</td>
<td>182</td>
<td>173</td>
<td>183</td>
<td></td>
</tr>
<tr>
<td>Nacional sede Bogotá</td>
<td>182</td>
<td>185</td>
<td>190</td>
<td>174</td>
<td>204</td>
<td>191</td>
<td>183</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>Tecnologica de Pereira</td>
<td>154</td>
<td>145</td>
<td>152</td>
<td>157</td>
<td>170</td>
<td>159</td>
<td>156</td>
<td>166</td>
<td></td>
</tr>
</tbody>
</table>

### 3. Results

#### 3.1 Determining the number of groups

The partition of the universities in groups will be made by means of the application of K-means using the software SPSS. For this purpose, the number of suitable groups will be determined initially, for which we use the method of the square sum of errors (SSE) of the distances between each element of a group and its centroid; in such a way that an appropriate number of groups would be the one in which the decrease of the SSE is more drastic. [15]. Indeed, the SSE difference between N=2 and N=3 is 4850, while the SSE difference between N=3 and N=4 is 1578, so the number of groups for our analysis is three, as can be seen in figure 1.

Once we have determined that three is the number of groups initially most suitable for our data set, we subject the values of the eight competencies of the SABRPRO tests to the k-means algorithm.
3.2 Identification and visualization of clusters

In the process of finding clusters, it was detected that three groups are clearly identified; see figure 2. The first cluster is made up of the following universities: Antioquia, Distrital de Bogotá, Nacional de Bogotá, Nacional de Medellín and Valle. The second cluster is made up of the following universities: Atlántico, Córdoba, Francisco de Paula Santander, Industrial de Santander (UIS), Magdalena, Military, Nacional de Manizales and Tecnológica de Pereira. The third cluster is made up of the following universities: Cundinamarca, Guajira, Nacional Abierta y a Distancia (UNAD) and Pamplona.

Figure 2. Visualization of the K-means analysis.
3.3 Group analysis

The analysis of the three groups in relation to the eight parameters considered in the grouping algorithm is structured in the following essential points, and a summary can be consulted in Table 3.

**Group 1**

The universities classified in Group 1 represent the best results of the SABERPRO tests; the representative element is the Universidad Nacional de Bogotá. This group is characterized by being universities that have institutional accreditation and accreditation in industrial engineering programs, as well as the universities of this group are in the top positions in different rankings. Within their curricula, the components of basic sciences, production and operations are outstanding. It should also be noted that these are the universities with the highest budgets and are located in the main cities of the country. This group will be labeled as high-performance universities, corresponding to 29% of the total.

**Group 2**

The universities classified in Group 2 represent the intermediate results of the SABERPRO tests; the representative element is the Universidad Nacional de Manizales. This group is heterogeneous in terms of accreditation, some universities have institutional accreditation while others have accreditation in industrial engineering programs. In this group stands out universities with technology and engineering approaches. This group will be labeled as medium-performance universities, corresponding to 47% of the total.

**Group 3**

Universities classified in Group 3 represent the lowest results of the SABERPRO tests. This group is characterized by being universities that do not have institutional accreditation and do not have accreditation in industrial engineering programs. Within their curricula, this group emphasizes Management and Administrative Engineering. The universities of this group are located in intermediate cities of the country. This group will be labeled as low-performing universities, corresponding to 24% of the total.

Table 3. Characteristics of the groups

<table>
<thead>
<tr>
<th>Cluster</th>
<th>SC1</th>
<th>SC2</th>
<th>SC3</th>
<th>GC1</th>
<th>GC2</th>
<th>GC3</th>
<th>GC4</th>
<th>GC5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>176</td>
<td>174</td>
<td>182</td>
<td>166</td>
<td>196</td>
<td>183</td>
<td>175</td>
<td>184</td>
</tr>
<tr>
<td>2</td>
<td>160</td>
<td>154</td>
<td>160</td>
<td>158</td>
<td>177</td>
<td>163</td>
<td>157</td>
<td>166</td>
</tr>
<tr>
<td>3</td>
<td>147</td>
<td>138</td>
<td>140</td>
<td>148</td>
<td>157</td>
<td>145</td>
<td>141</td>
<td>141</td>
</tr>
</tbody>
</table>
Conclusions

The classification proposal of the public universities that offer the academic programs of Industrial Engineering in Colombia identified three big groups, the visualization of the results obtained by means of the algorithm k-means show the intra-group cohesion and the intergroup separation which is an indicator of validation. This information provides the management of the universities, both at the Faculty and Rector level, with a general overview of the level of competence of the programs, allowing the establishment of positioning strategies. The results obtained can be contrasted with the employability statistics of the graduates of each university.

References


Received: September 12, 2018; Published: October 2, 2018