Chapter 2

Thinking with Data: The Role of ALEA in Promoting Statistical Literacy in Portugal Pedro Campos

"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write". (H.G. Wells)

2.1 Introduction: the statistical literacy view

Statistical information is nowadays an essential form of knowledge necessary for full citizenship. In carrying out their activity, Statistical Offices are responsible not only for producing, disseminating and analysing statistical information but also for ensuring that this information is understood as objectively as possible. One of Statistics Portugal's main goals consists of disseminating data and promoting the general use of statistical information with objectivity and comprehensibility. The improvement of statistical literacy is a direct consequence of this attitude that Statistics Portugal defines as an important goal in their activity plans.

In modern societies, statistical literacy is an important asset for full citizenship and ensures the best possible provision of a service of public utility. Indeed, to understand the meaning of a text, we need to be able to do more than just spell out the words it contains in the same way that in the reading of statistical information we need the understanding of the concepts and methodologies used in its preparation. We need to improve statistical thinking, or in other words: we nee to know how to "read" the data.

Some efforts have been made in last years with the aim of improving the way Statistics Portugal writes to the public (mass media and society, in general): in 1998 the format and the style of Press Releases changed. It consists on daily information that is a summary of information through which Statistics Portugal disseminate the earliest results obtained through the various statistical operations, as well as other relevant data within the scope of statistical activity. The improvement of the writing style, aiming at increasing the clarity and the objectivity of the statistical message has been achieved with the help of an inside training program in Statistical Portugal.

Later, in 1999, as an answer to the requests of the educational communities, a new project was set up for the specific purpose of providing tools related to understanding, using and teaching of statistics. The ALEA (Local Action of Applied Statistics - <u>www.alea.pt</u>) is aimed essentially at primary and secondary schools, but it is also an important resource for supporting interdisciplinary projects, being, simultaneously, of benefit to many other groups of people.

In section 2 we introduce the view of statistical literacy in Portugal and proceed in section 3 with a description of the components of ALEA, focusing on statistical competitions that were created to improve significant statistical learning and Activaleas - statistics at a glance. Later, in section 4, we present some indicators in order to systematize the assessment of the use of ALEA. Finally in section 5, we give an overview of ALEA in the future.

2.2 Statistics Portugal and "thinking with data"

Statistics Portugal is committed to increase statistical literacy. Since the last decade, the promotion of statistical literacy is one of the strategic goals of the institute. Consequently, Statistics Portugal is aware of the importance that the Teaching of Statistics plays in Primary and Secondary Schools. It is, in fact, one of the most important instruments in achieving this goal.

After analysing the curricula of several introductory courses in Statistics, we observe that Statistics is often seen as a branch of Mathematics rather than a problem-based matter. There are indeed courses in which it is suggested that the students read the texts before the lessons, during which only questions regarding the data sets to analyse are discussed. From this perspective, lesson time is used to discover statistical principles and apply statistical techniques. Thus, the students are asked to analyse and use authentic data, some gleaned from available sources and other through class surveys or experiments. Therefore, statistical thinking or reasoning will come before statistical methods. Following Smith (1998), good statistics should, of course, be identified with purity and mathematical precision, but they should also be associated with "careful thinking".

Creating experimental and learning environments based on ICT, namely through the analysis of real data is an important way to improve the statistical thinking and reasoning, as well as technological literacy. The ALEA project, by placing a vast quantity of available official statistical information at the disposal of the students and by allowing the students to create their own documented databases, complies with this objective.

ALEA is to be viewed as a project which aims at providing both teachers and students of primary and secondary education with teaching materials for the study of Statistics. These materials include a web page containing several educational resources, "statistical" entertainment, and downloadable material. This project is a joint idea of Statistics Portugal and Tomaz Pelayo Secondary School, bringing together The Portuguese Ministry of Education and the Statistical office.

The ALEA project includes several online resources, CD-ROM, paper publications, b-learning training courses, workshops, etc., covering different topics, from statistical theory, real data and class-room problems to entertainment, games and jokes. In the next session we give an overview of some of ALEA's features, including these and other contents.

2.3 ALEA: that's Edutainment!

The web site is the core of ALEA. That is where all the contents are available, and where the users can interact with the statistical resources. Two statistical courses are available online: "Statistical concepts" which presents elementary concepts of descriptive statistics and "Probability concepts" which provides concepts and exercises on probability. A b-learning environment was also created, based on the two statistical courses available in ALEA. Other online resources are included, such as statistical information about Portugal (until county level), the European Union (with interactive maps and statistics for all countries) and an interactive statistical summary of the countries having Portuguese as the official language. A set of handbooks (Dossiers & Resources) dealing with specific topics, such as Environment Statistics, Census, History of Statistics, Statistical Software, Surveys, etc. is also available. There is also a special area for entertainment (Fun with Statistics!) with several games and cartoons; the Challenges, a competition where students may participate; Focus on Statistics, a page containing news with a statistical content found in the press, and a section named "Facts & Names in Statistics", with the most significant names and dates of the history of statistics. There is also a special area (Your Surroundings) where the users can choose a territorial area (from the whole country to a single county) and see statistics for that particular geographical unit.



Fig. 1 – ALEA Home page

The triangle is used as metaphor for ALEA's structure. In this type of graph, the vertices represent Statistics, Data and Entertainment. The different "parts" of ALEA are placed in between these three vertices according to their specificity. Actually, some efforts were made so that ALEAs' main structure is a balanced content where both teachers and students can take part in solving everyday problems giving rise to reflections of a statistical nature. Our aim is that it also gives an opportunity to use interdisciplinary activities by using and serving other subjects besides mathematics.



Fig. 2 – ALEA's Structure

Geography, Economics, History, Sociology are some of the subjects for which ALEA has specific contents. In what concerns Mathematics, topics such as Descriptive statistics, Graphs (Histograms, Bar plots, etc.) and Probability and Inference Statistics are covered by ALEA with some detail. In Table 1 we present the sections containing elements of Statistics considered in the official curriculum of Mathematics at the secondary level, and the corresponding contents included in ALEA.

* Available only in the Portuguese Version

Official Curriculum	ALEA		
• What is Statistics : goals and definitions;	Course on Statistical Concepts		
 Notes on the history and evolution of statistics 	Dossier VI - Notes on the history of statistics		
• Statistical Phenomena; examples from real life.,	Focus on Statistics		
	Daily News*		
	Official Data: Portugal in Figures, EuropAlea*, etc		
	Data in School: Virtual Gallery*		
• Basic concepts: Census, Surveys, Population, Sample.	Course on Statistical Concepts		
• Organization and interpretation of statistical data (qualitative and quantitative):			
- Data, tables and graphs	Dossier IV – Statistics with Excel*		
- Location measures	Course on Statistical Concepts		
- Spread measures	CalcALEA		
- Bivariate distributions	Course on Statistical Concepts		
Probability and Inference	Probability Concepts*		



2.3.1 Statistical Competitions to improve significant statistical learning

As mentioned above, the reading of statistical information needs that the reader understands the concepts and the methodology used in its preparation. That is indeed something that Statistics Portugal is committed to do, in order to promote the increase of statistical literacy. As stated by Fink (2006) "students will always learn something, but good teachers want their students to learn something important or

significant, rather than something relatively Therefore, one particular resource that has been recently is ALEA's Challenges. Being often used by and students, ALEA's Challenges is a Portuguese competition containing everyday life problems based It is oriented towards the Primary and Secondary aiming at increasing the reading ability of tables and problem showed in Figure 3 is an example of such



insignificant". introduced both teachers language on daily news. students graphs. The a competition:

we asked the students to analyse the graphs published in a Portuguese newspaper concerning the difference of wages between women and men in Europe. Women earn lower wages compared to men and besides, only 32.1% of the women in the European Union ascend to top positions in the firms they work. A comparison between that percentage in some European countries is depicted in a bar plot. In a different bar plot, the difference between men and women wages in percentage in shown for the same set of countries. The fist horizontal bars in the graphs represent the average values for European Union (UE 25).

Problema de Nível 1				
Faz uma leitura e interpretação dos gráficos da figura ao lado e indica: 1. Quais são os países da União Europeia (UE) em que a percentagem	Mai pagas Só 32,1% das mulheres na União Europeia ascendem a cargos de topo nas empresas mas, em média, as trabalhadoras do sexo feminino aufrem salários inferiores em 15% aos dos homens. Em Portugal, essa diferença não ultrapassa os 5 por cento (Dados para 10 países da UE)			
de mulheres em cargos de gestão é superior à média da UE 25?	% de mulheres em carges Diferença salarial entre homens de gestão (1º semestre 2005) e mulheres em % (2004)			
2. Qual é o país em que a diferenca	UE 25 ALEMANHA	32		15
salarial entre homens e mulheres em	DINAMARCA	23		17
% e a mais baixa?	FINLÂNDIA	30		20
	FRANÇA GRÉCIA	37		10
	IRLANDA POLÓNIA	30		
		34		5
	Fonte: Eurostat		and in the local division of the local divis	

Fig. 3 – Example of ALEA Challenge

Based on this information, we create two different types of questions, with different degrees of difficulty. The questions are grouped according to the education level - basic and secondary. For the basic level, the questions are:

- What are the countries of the UE 25 for which the percentage of women in top positions is greater than the average of the European Union?
- In what country is the difference of wages between men and women lower (in percentage)?

For the secondary level, a different question was posed. Students must consult another graph and compare some indicators such as life expectancy, education level, employment rate, and rate of employment in the fields of Exact Sciences, Maths and Computer Science, considering women in Portugal and in the European Union. We ask the students to comment a piece of writing that sates that Portuguese women are less qualified than the average in Europe, although they get better jobs and obtain higher results in the fields of Exact Sciences. In some cases they can even overcome the difference of earnings when compared to men.

With this and other experiences available in the "ALEA Challenges" (many other are available at <u>http://www.alea.pt/html/desafios/html/desafios.html</u>)¹, we try to promote the understanding of the context as it is crucial to the reading of numbers. Nobody will be capable to realize the intensity of a number if he or she does not "feel" it. It must have some kind of meaning for the reader. Later on in this chapter, we give a summary of the tax of response and make an overview of the participation in ALEA Challenge competition.

2.3.2 Activate Statistics in your mind: statistics at a glance

Another resource we have made available recently is ActivAlea. ActivALEAs are learning-by-doing assignments containing tasks, comments and self-test questions in order to systematize the basic statistical concepts in the classroom. The idea behind ActivALEAs is that sometimes you do not how to accomplish a trivial task in statistics. For instance, the graphical presentation of information is very important in Initial Data Analysis and one must be aware of the main characteristics of the data in order to proceed with more sophisticated statistical techniques. However, you do not have to read an entire book on introductory statistics if you want to produce a bar chart or an histogram. For that purpose, we developed small sheets containing short explanations for some frequently asked questions in order to systematize the basic skills of statistics. The last ActivALEA published so far (number 8) describes the sensibility that the user must have to deal with the differences between the mean and the median. In particular, we start by computing the mean and the median of the weights in a sample of 10 students collected in the classroom. Then we add the number 2 to each observation and multiply each observation by 2. Furthermore, we compute the mean and the median again. Later we ask the

students to change one of the sample values and analyse the result. What happens? The mean changes, the median remains unaffected. This activity is used to study the basic properties of the mean and the median well as the robustness of these two statistics. With three four exercises, these properties are exemplified in a short classroom experience using real data.



The other learning-by-doing assignments are indicated in Figure 4.

¹ Only in the Portuguese Version



- 1. Population, Sample and Descriptive Statistics
- 2. Association between qualitative variables
- 3. A graph is worth more than a thousand words?
- 4. Association between quantitative variables: the correlation coefficient
- 5. Random experiments
- 6. Frequency tables
- 7. Bar charts
- 8. Mean or Median?

Fig. 4 – ActivALEAs titles available so far

2.4 Assessment and accesses

2.4.1 How is ALEA used by teachers?

As teachers play an important role in statistical education, we need to know whether ALEA has been a useful and efficient resource for teachers up to now. Is ALEA used in the classroom context? Is it used for personal training? Do teachers with background in mathematics use ALEA in the same way compared to teachers that have no background in mathematics?

A survey has been made in order to answer to some of these questions (Martins, et al, 2007). This survey was targeted to the teachers that use the resources available in ALEA's web page². One of the conclusions of the study is that the majority of the teachers with Mathematical qualifications use ALEA more often than others. On the other hand, the teachers with different qualifications (Natural Sciences, Economy, etc) use ALEA more frequently for the specific purpose of "Personal training and learning concepts".

² The questionnaire used in the survey included 12 questions that can be clustered in three different groups: (1) scientific, academic and demographic background of the individuals in the sample, (2) reasons regarding the use of ALEA and finally, (3), we asked the teachers to rank a set of topics available in ALEA and to weight its contribute to the leaning and teaching of statistics. The target population of the survey is the group of teachers that use ALEA as a resource. A data base of registered users of ALEA being simultaneously teachers at any level of teaching was used as the source of information for addressing the questionnaire. About one thousand questionnaires have been sent and, for technical reasons 200 have not been delivered. The response rate was small (approximately 13%) although it is framed within this kind of electronic survey.



Fig. 5– Use of ALEA according to the teachers' background

Simultaneously, we needed to know what the circumstances and contexts involving the use of ALEA are. Therefore, to simplify the process of collecting the information, we have previously classified the contexts in four categories: (1) In the context of the subject matter; (2) School projects; (3) School clubs and (4) To provide educational/pedagogical support. For every option, a rank between 1 (few) and 3 (much) was used to codify the possible answers. As it is a multiple response question, the teachers could choose more than one option. Figure 5 contains a summary of the answers to this question.

Concerning teachers with Mathematical qualifications, ALEA is frequently used "In the context of the subject matter" and rarely used in "Schools clubs". On the other hand, teachers with different qualifications, use ALEA more often "In the context of the subject matter" and "To provide educational/pedagogical support".



background



Fig. 6– Circumstances and contexts involving the use of ALEA according to the teachers' background according to the teachers' background

2.4.2 Evolution of the number of visits

Considering now the overall statistics of ALEA, daily information is collected automatically by the web server and used to evaluate the evolution of the visits and page views in last years. Those statistics are important to assess the use of the resources in ALEA. Figure 5 shows the evolution of the unique visitors and visits from 2003 to 2008.

This partial analysis contains the last five years of ALEA's activity, but it is enough to conclude about the growth in the recent past. In 2007 ALEA has been visited by 370 thousand different visitors and we estimate that in 2008 this number will rise to about half million. Although we were not able to confirm the origin of the accesses, we believe that a great part of this growth is a result of the increase in the accesses of the students and teachers to the Internet during classroom activities due to the importance that statistics is attaining in school curricula.





2.4.3 Interaction with the users

There is more than one way through which it is possible to interact with ALEA. There is a mailbox in ALEA where messages are automatically generated by a web form that allows for comments and requests. Most of the messages are "technical" requests, such as how to compute a confidence interval, or how to explain the meaning of the Variance. However the number of messages coming from users not related to schools is increasing considerably. Many of them are related to the sample size in surveys or to statistical information. Actually, Statistics is, to a greater extent, a matter of citizenship, so this fact is not surprising. ALEA's mailbox receives more than 500 messages per year, some of them being from other Portuguese spoken countries (mainly from Brasil). An increasing number of messages are coming now from several different countries with requests for translating materials to their home languages.

But interaction with users is more than electronic asynchronous communication: we are also organizing blended-learning courses oriented to the teachers of Mathematics where the resources of ALEA are used. At the same time, we have other face-to-face activities such as Workshops, meetings, etc. so that we can receive direct feedback from the users. Since 1999, ALEA has already been present in many different places. The diffusion strategy is based on three types of events:

- 2.4.3.1 Courses and Workshops included in statistical events such as Statistics and Data Analysis congresses – (Meeting of the Portuguese Statistical Society, SPE 1997 and 2002; Meeting of the Society of Classification and Data Analysis: JOCLAD 1998, 2002)
- **2.4.3.2 ALEA meetings** organized by the Portuguese Ministry of Education, with classes of 25 to 30 teachers of different areas, at maximum, where ALEA is presented and some activities are practised (Secondary School in Matosinhos 2002, Secondary School in Guimarães 2003)
- **2.4.3.3 ALEA Forum** a national one-day meeting with invited speakers with invited presentations related to the use of ALEA (Porto, 2006)

Two distance learning courses have been launched in the last years: A course on Descriptive Statistics (e-learning course opened worldwide in December 2003) and a Distance Learning Workshop for the Teaching of Statistics (four editions - 2004, 2005, 2006 and 2007). The main goal in performing these courses is to reinforce the contacts of ALEA with the schools and with the public, using the materials already available in the web page.

2.5 ALEA in the future: learning is more significant!

Data analysis is getting more and more important in the curriculum of the basic level (5th to 9th grades). Therefore, we are starting a new area devoted to the basic learning and teaching of Statistics in those levels of education in order to provide resources that are useful to teachers. ActivALEA's 6 (Frequency tables) and 7 (Bar charts) have already been created with this purpose.

The Maths curriculum of these grades considers that teachers have to relate statistical issues with the issues studied in other subjects, and with themes related to the national or international reality. Simultaneously, the student must get involved, and have a critical opinion about wrong graphs, statistical "lies" or badly selected samples. So the learning of statistics is becoming more significant because the data used to test the procedures is related to the themes that students are learning in other subjects.

This is basically the idea of the ALEA Challenges, where students have to read an article in the news and have to answer a couple of questions or then have to write a comment.

We are also creating a prize for the best statistical project. It is a competition that aims at promoting the collection and analysis of data, as in the spirit of *thinking with data*. Following Snee (1993), "*Data collection and analysis is the heart of statistical thinking*. *Data collection promotes learning by experience and links the learning process to reality*". The works submitted to competition take the form of projects providing the students with experience in formulating questions, defining problems, formulating hypotheses and operational definitions, planning experiments and *surveys*, collecting data and, regarding the best way to deal with measurement errors, draw up data summaries, analyse them, how to communicate discoveries and plan experiments and how to correlate the ideas suggested by the discoveries.

2.6 References

Fink, D., (2003), "*Creating Significant Learning Experiences*", Jossey-Bass, eds. Lovett, M. and Shah, P. (editors), (2007), "*Thinking with Data*", Lawrence Erlbaum Associates

Martins, R., Graça-Martins, E., Oliveira, E., (2007), How can teachers promote Statistical Literacy? Teaching and learning statistics with ALEA – A Survey, in *Proceedings of the 56th Session of ISI*, Lisbon

Smith, G., (1998), "Learning Statistics by Doing Statistics", in *Journal of Statistics Education*, v.6, n.3

Snee, R. D., (1993), "What's Missing in Statistical Education?" in *The American Statistician*, 47, 149-154.