A New Statistical Adventure with Exploristica 2.0

Martins, J. Pinto[1], and Campos, Pedro[1,2]
Statistics Portugal
FEP, Universidade do Porto and INESC TEC

Abstract

In this work we introduce Exploristica 2.0, a new web-based (and physical) travelling exploratory exhibition consisting of various interactive modules with the aim of bringing the fundamentals of Statistics and Probability, and improve the general standard of statistical literacy among young people. The exhibition offers several interactive modules, inspired in different episodes told by Sebastian Odd, a detective interested in solving the mysteries of statistics, and related to the phases of the statistical process: selection and collection of data, description of data, data estimation, and interpretation. The new version 2.0 is based on storytelling and is available since 2020, where new physical and digital modules can be played locally or in distance mode. In the "Virtual Submarine", participants embark on a journey through an underwater ecosystem where they need to collect a sample of a new reptile species. Students sit in a chair, wear VR goggles and hold hand controls. That way, players feel as if they are sitting in the cockpit of a submarine and able to have a 360º view. In this work we provide an introduction to Exploristica 2.0, describe the several games, and focus on the box plots experiments of the virtual submarine.

1. Introduction

Edutainment materials are designed to educate through entertainment, and seem to be really important in the teaching of science. Although there is not a consensus on whether there are specific benefits on undergraduate research experiences (Linn, et al., 2015), there are many situations where educators use practical experiments with manipulatives in the teaching of Biology, Chemistry, Physics, and Mathematics. There are some companies that produce kits with experiments that can be used in schools or sold in specialized stores. However, there are not many physical materials available for teaching Statistics. One can use sets of dice (see, for example, statistics manipulatives - Learning Resources (2021)), cards, sample surveys, or tools. In 2003, many resources were already available to teach statistics and probability: there are, indeed, many electronic resources, made of Java applets or other software tools that let students explore and interact. Statistics Online Computational Resource (SOCR), developed at UCLA (the University of California, Los Angeles.), is a good example of resources that can be used to teach statistics (NETWATCH, 2013). Using these resources, students can see how modifying parameters affects the mean, median, variance, and other measures. A list of different resources used in several countries in the world is available in the ISLP, the International Statistical Literacy Project, website (ISLP, 2021). Another important list of resources is MyStatsLab, intended to engage high education students with immersive content, tools, and experiences (Pearson, 2021). But there are other important initiatives related to electronic edutainment that it is worth mentioning for the teaching of Statistics. Census at Schools (see the Irish page a CensusatSchool, 2021), is an international project that started in the UK in 2000. Several countries became involved in this initiative containing electronic resources, surveys to be used in the classroom. In Ireland, for example, CensusatSchool was used during 2016 to teach about the census and about how to effectively to use census and other statistical data as a key part of education. Statistical offices also produce materials to improve aiming at improving students' understanding of statistics and, consequently, their civic participation as consumers and producers of data. Statistics Canada pioneered the creation of a network between a statistical office and schools. Statistics Canada's Education Outreach Program, recognized internationally, aimed to develop statistical literacy and data management skills among youth: In the mid 1980s Statistics Canada began creating an electronic teaching resource—called E-STAT—that would bring together census and time series

data for schools. About the same time, workshops on Statistics Canada resources were offered to professors and teachers-in-training at university faculties of Education (Townsend, 2003).

Statistics Portugal was one of the first statistical agencies in Europe to create links with educational communities. ALEA (www.alea.pt) has started in 1999 as the result of a cooperative project involving Statistics Portugal, the Portuguese Ministry of Education, and a Secondary School (ALEA, 2021). Two courses (one in Probability and another one in Statistics) are available, as well as many other resources, including cartoons, a guided tour to the history of statistics, data, games, and competitions (most part of the contents is presented in Portuguese). Another good example of the use of everyday statistics to teach statistics is "Against all odds", a video instructional series on statistics for college and high school classrooms and adult learners, with the emphasis on "doing" statistics, (Annenberg Learner, 2017).

2. Exploristica

Exploristica (version 2.0) (https://exploristica.scl.pt/ is a new itinerant exploratory exhibition, virtual and physical, consisting of various interactive modules (including tangible manipulatives and computer manipulatives) with the aim of bringing the fundamentals of Statistics and Probability to basic and secondary schools. The previous version (version 1.0) of Exporistica has been created in 2013 by the Portuguese Statistical Society and been supported by Ciencia Viva, the Portuguese agency for the dissemination of scientific culture.
The main concepts to be taught are: survey samples and census, location and dispersion measures (mean, median, standard deviation), graphs (box and whiskers, bar plots, histograms), random and non-random sampling, relative frequency, and probability. Exploristica describes five important phases of the statistical process - Select, Collect, Describe, Estimate and Interpret - and it is organized in six modules that offers their contents in the form of games and other interactive experiences. The target group of Explotistica 2.0 are teachers and students of upper basic and secondary school (12 to 18 years old).
All those modules (games) were developed and made available in three environments:

a. physical, with touch screens an other physical interfaces, like joysticks, buttons, pumps, and so on;
b. internet, browser based. Everyone, since connected to the internet, can play the games using a computer and a browser (https://exploristica.scl.pt/)
c. Mobile app, either IOS or Android systems. Each module/game can be free downloaded as an app, which is available on the app store.

Exploristica 2.0 was supported by DIGICOM a project of the European Statistical System (ESS). The goal of DIGICOM was to modernise the communication and dissemination of European statistics. This is achieved by exploring and developing innovative dissemination products and services based on experiences within the ESS and concrete needs of European statistics users.

3. *Modules that make up Exploristica 2.0:*

Watch - It is necessary to operate a watch manufacturing machine. Some watches from this factory are defective and it is necessary to control the machine pressure and evaluate the quality control of the whole production. It is intended for players to be aware of statistical measures that are mostly used in industry control charts (mean, median and standard deviation).

Dr. Odd – This is a Cluedo style game, where probability studies should help to solve a crime. Dr. Odd was murdered in his mansion. Five weapons, which must be discovered, may have been used in this crime. On the other hand, there are six suspects of being the murder. Knowing the probabilities of each weapon being the murder weapon and each suspect using each weapon,

players must calculate which of the suspects is most likely to be the murderer. This module aims to explain the use of conditioned probabilities and the usefulness of the Venn diagram in this context.

Conga – With elections approaching, the president of the parish council, D. Stimacione (the villain), wants to know if he can be re-elected. With the help of Conga, the gorilla, a sample is collected, firstly for convenience and then at random, showing how different the results are in these two types of sample data collection, when trying to estimate the expected number of votes for D. Stimacione.

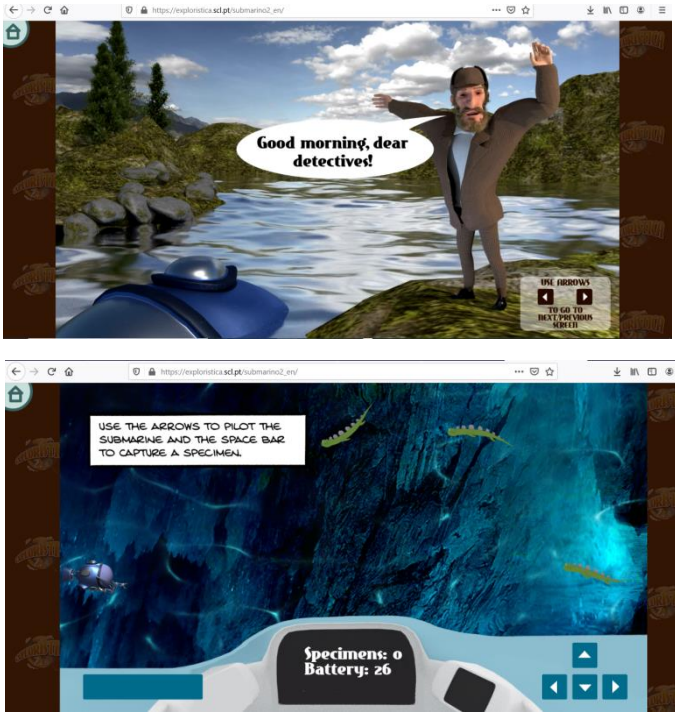
The Key – This module consists of a transparent dome. Inside the dome, there are 100 little cubes on a surface with a circle drawn inside a square. The player presses and holds a button that throws and shuffles around the little cubes as they fly inside the dome and fall randomly. The player counts how many cubes fell outside the circle. The player verifies that the relationship between the number of cubes in both areas is Pi/4. The player must repeat the procedure to verify that the relation between the cubes inside and outside the circle is roughly constant.

The challenge is to relate the result to keys that have numbers inscribed, to find the one that opens a secret door.

Patchwork - After a bank robbery, pieces of cloth are found in a nearby garden and several suspects are arrested whose pants were torn by a dog as they ran through that garden. The player must collect pieces of fabric in the place so that they can be used as a representative sample. On the basis of the pattern showed by this sample he tries to identify the robber looking at his torn clothing.

Submarine – A submarine trip in a lake where the user must pilot the submarine and catch specimens (one at a time) with a claw at the front of the submarine.Participants weigh, measure, and identify the sex and age of each specimen they capture (and then return to water). After this data collection, an explanation is presented on how to draw a diagram of extremes and quartiles.

The Interactive Submarine is based on a submarine trip in a lake where the user must pilot the submarine and catch specimens of the *Exploristicum* reptile (one at a time) with a claw at the front of the submarine.Participants weigh, measure, and identify the sex and age of each specimen they capture (and then return to water). After this data collection, an explanation is presented on how to draw a diagram of extremes and quartiles.This game can be played on line (Fig 1) or with virtual



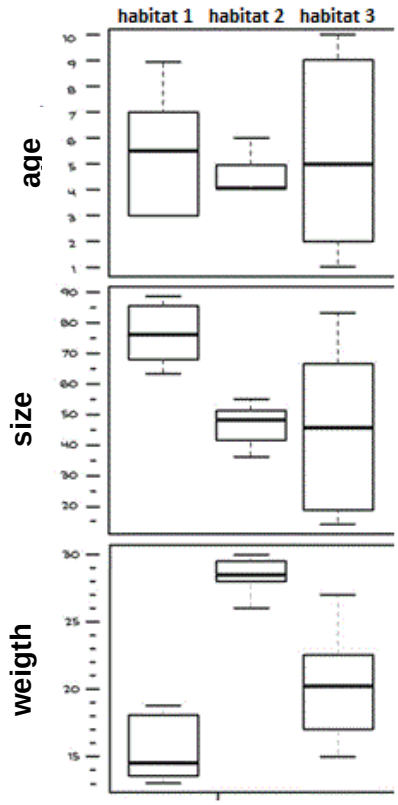


Fig. 1. Display of Virtual Submarine as an online game (left) and the boxplots of population data

reality goggles (Fig. 2). An app may be downloaded to the mobile phone devices, so that the Virtual Submarine can be played in a classroom within the context of a Maths class. Samples of Google cardboard's-type viewers will be also available. The Google Cardboard is a simple viewer that anyone can build or buy for use with a head mount for a smartphone to encourage interest and development in Virtual Reality applications. So, users can either build their own viewer from simple, low-cost components using specifications published by Google, or purchase a pre-manufactured one. To use the Submarine App, users have to download it and run it on their phone, place the phone into the back of the viewer, and view content through the lenses.

There are three different subspecies of the reptile in the same lake, which are distinguished by their main characteristics (measure weight, age, size and sex). Therefore, participants have to produce boxplots of these variables (see Fig.1 , right) and compare them with the boxplots of the population data in order to identify the subspecies that have been collected. Players finding out the real subpopulation can win the game. Although the use of boxplots make them particularly difficult for young students to use in authentic contexts, they are indeed, a very important graphical representation to be disregarded as a statistical tool. Besides, boxplots can somehow be viewed as likelihood measures, since they allow comparing sample data with population. Despite not being a "parametrical" likelihood, this notion of similarity between sample and population is the same we can find in the work of Royall (1977), that is connected with the law of likelihood and is called interpretation of statistical evidence provided by data.

References

ALEA, 2021, Acção Local Estatistica Aplicada, accessed in March, 22nd, 2021, available at: www.alea.pt

Annenberg Learner, 2021, Against All Odds: Inside Statistics, video series (some are availble at: https://www.learner.org/series/against-all-odds-inside-statistics/

CensusatSchool, 2021, accessed in March, 22nd, 2021, available at: http://www.censusatschool.ie/en/home

ISLP 2021, web resources accessed in March, 22nd, 2021, available at: https://iase-web.org/islp/Resources.php?p=New_Resources

Learning Resources, 2021, webpage accessed in March, 22nd, 2021, available at: https://www.learningresources.com/search.do?query=statistics

Linn, M., Palmer, E., Baranger, A., Gerard, E., Stone, E., 2015, Undergraduate research experiences: Impacts and opportunities, Science 06 Feb 2015: Vol. 347, Issue 6222, 1261757, DOI: 10.1126/science.1261757

NETWATCH, 2003, EDUCATION: Statistics Starter Kit, Science, Issue of 5 December 2003, Volume 302, Number 5651

Royall, R., M., (1977), Statistical Evidence: A likelihood paradigm. Chapman and Hall, Teaching about Approximate Confidence Regions Based on Maximum Likelihood Estimation

Townsend, M., 2003, Developing statistical literacy in youth: Statistics Canada's education outreach program, Proceedings of Statistics Canada Symposium 2003, Challenges in Survey Taking for the Next Decade