TANGIBLE STATISTICS

Annamaria Lisotti and Laura Vicini
IIS Cavazzi, Italy
lisottiannamaria@gmail.com

Data has dramatically scaled up expanding into everyday life. How it is represented and communicated is therefore a key concern for civil society. Statistics has traditionally privileged the visual channel to convey info. This is far from inclusive as it cuts off the visually impaired who rely on other senses to access knowledge and favours visual learners against the established principle that Education should promote all learning styles. Different options of data physicalization are explored to deliver data through multiple sensory channels thus enhancing everyone's comprehension. Sonification; real-time microcontrollers dataflows turned into dynamically interactive Data Art; georeferenced data mapped and 3D printed; AR models for data-journalism: all these activities have a potential for interdisciplinary applications, equipping students with transversal competences and a wealth of digital free tools. They can also all be easily implemented online for distance education, flipped classroom and project-based work enhancing statistical inclusion in a wider outreach context.

INTRODUCTION

The activities presented in this paper took place in the context of the Erasmus+ BeReady Project 2021-'23 funded by the EU under the call "*Supporting the continuation of teaching STEM subjects during the Covid-19 pandemic through project based online practices*". BeReady partners consisted of four Universities and four high schools from different EU countries. Each school had the task to design and test OER on a topic of choice under the common denominator of *Digital Education Readiness in the field of STEM Education*.

Liceo Scientifico at IIS Cavazzi (Pavullo nel Frignano, Italy) - a high school with a science focus and students' age ranging 14-19 - chose to work on Statistics, producing quite innovative results in a specific module about Tangible Statistics. The choice was driven by two considerations. Statistics curricula urgently need upcycling as data literacy is configuring itself as one of the Education pillars for twenty-first century citizens. Data has in fact dramatically scaled up and expanded into everyday life becoming a fundamental interdisciplinary tool for the comprehension of the real world beyond prejudices, clichés and fake news. On the other side however there's very little use for this wealth of data without the ability to communicate and engage. How data is conveyed and represented is therefore a key concern for civil society and has recently led to the widespread popularity of infographics which nevertheless do not transcend the traditionally privileged visual channel. This is far from inclusive as it cuts off all visually impaired-who rely on other senses to access knowledge-and eventually favors visual learners against the established principle that Education should promote all learning styles. Hence the idea of stretching data visualization into a wider concept of data physicalization, offering multisensorial data experiences suitable for any kind of public, from specialists to those who have little or no mathematical knowledge.

IIS Cavazzi students tested a wide variety of different options from interactive Data Art to sonification, with a particular focus on haptics. 3D printing is a new and very powerful tool for bringing intangible concepts to life and with Augmented Reality the physical handling of the solid models can be upgraded into digital manipulation reaching out to a much wider and even distant audience thus successfully closing the loop: from digital (csv files) to real (physical models) and then digital again (AR models). Last but not least another engaging and worth exploring modality is represented by Participatory Statistics which actively involves the public in hands-on representation of the very same data they are delivering while answering to specifically designed interviews.

DATA PHYSICALIZATION

Data Physicalization is a fairly recent term as it was first used in 2015 in an open access article. Since then the new trend has been to explore innovative experiences around data consumption beyond the written and the visual. Little of this ever entered school classrooms, though.

There are many ways of making data "tangible" with different levels of sophistication: from arts & craft and the use of everyday objects to high tech solutions. The common denominator however is to engage the public and convey the message quickly and effectively. To this purpose you can find a wealth of free tools available on the Web, many of them perfectly suited to an educational target.

Sonification

Sonification defines any non-speech audio used to convey information or represent and perceptualize data. Since it is proved that music and sound may induce a strong emotional response in listeners, transforming datasets into auditory pieces seems a very promising method in overcoming data communication barriers, adding new layers of understanding and dramatically enriching data storytelling. Adding the fact that distribution is relatively low cost, simple and fast this modality is potentially appealing to journalists, scientists, designers, educators and anyone who wants to reach audiences in new ways.

Despite thirty years of research though, sonification is still not yet widely spread, at least in schools. However many free tools for sonification with no need to be either coders or musicians are now easily available on the Web. We used TwoTone and Highcharts. Their learning curve is very quick and far from steep, holding very satisfactory results with the possibility to reach sophisticated levels.

First triggered by a quick reflection on digital inclusion as a human right not always respected on the Web (see https://heartheblindspot.org/en/) students were then exposed to a wealth of examples from very different areas to make them revive their acoustic sense and appreciate the potential of this methodology. Being so used to transferring and consuming information via visuals, they seemed quite surprised to be able to interpret this novel language and amazed that it worked rather well offering a new insight in data. The experience suggested that after all rather than an intrinsic difficulty to use different ways of representing data, it is more a matter of being regularly exposed to different sensory stimuli.

As a second step pupils were instructed on the technical aspects of the tools and finally asked to create their own sonification. The option was to either work on “hearing kinematics” with examples of different motions turned into sound, or produce auditory graphs of mathematical functions. Here you can find a sample https://tinyurl.com/83ny7vk6. The student interpreted the parabolic motion of a bullet fired horizontally rendering the contrast between the initial flat slope of the parabola followed by the increasing speed due to the vertical uniformly accelerated motion. Technically he achieved the effect adding multiple audio tracks at different speeds appropriately filtered in sequence.

Interactive Dynamical Data Art

With the advancing of IoT (Internet of Things) and embedded monitoring systems working with microcontrollers and low cost sensors is becoming more and more popular both in schools and among the general public. However when the acquired data gets too much and scrolling too fast on the serial monitor it can be very difficult to grasp and fully perceive changes and the contextualized meaning of all those digits. An interactive visual approach seems much more accessible, immediate and effective, producing what we may call dynamic interactive DataArt that is to say the art of creating responsive, visual data representations that both inform and engage.

Students used Arduino UNO with simple sensors like LDR (Light Dependent Resistor) and thermistors (thermal resistances). They set up the circuit and wrote the code in the Arduino IDE. The sensors’ dataflow stored in the parameters were then translated through Processing free software in colour, dimension, position, etc. of the visualized shapes. Top results are reached when those visuals become synesthetic representations where parameters and their changes are able to convey the feeling of the sensed values. Just to cite a few examples students designed cats’ eyes in the dark getting bigger and yellower as the level of light decreased or a home fireplace whose flame changed in color turning to a more intense red as temperature increased; some others had a ball rolling down the screen main diagonal whose size increased with temperature. The level of sophistication can be easily spiralled along with coding proficiency. Everything is real-time as data acquisition and Processing code are running simultaneously on the same laptop. In this way changes in data are much more easily caught and live interaction is possible. It is also feasible to retrieve data from a csv file and make dynamical objects out of them. Although interaction will no more be an option an effective synesthetic experience can still be reached.

All the students were already used to programming Arduino as they regularly ran Physics homelabs with their own device while they had to be instructed in the use of Processing with 1 hour crash course (they already knew Openprocessing, the web based version of the software from MIT). However the proposed activity is easily reproducible in any class with a preliminary 2 hours for Arduino basics and 2 more for Processing basics.

Data Haptics

Touch is an innate way of gaining knowledge in babies, and the instinct to reach out for things is not really dying as we grow up. 3D printed models aim to convey data in an immediate, intuitive and elementary way: what is lost in precision is gained in clarity, engagement and interaction. We made our first experiments with GeoGebra to recreate through the extrusion of polygonal profiles the 3D models of variable motion (uniform and uniformly accelerated ones). Inspired by blind pupils we wanted to test whether "feeling" graphs in 3D would enhance understanding of all students, particularly those who seem to experience difficulties in reading 2D kinematic graphs. Work was run in 2 steps: *phase 1*- students drew the graphs, reflected on their characteristics, 3Dprinted the models. p*hase 2*-_students blindfolded examined through tactile exploration their schoolfellows' models then they interpreted and drew the corresponding graphs. Results were satisfactory; two thirds of the students thanks to the tactile immersive experiential approach together with the augmented concentration during the exploration seemed to gain a better understanding of motion graphs; a touch of gamification in the team challenge helped as well.

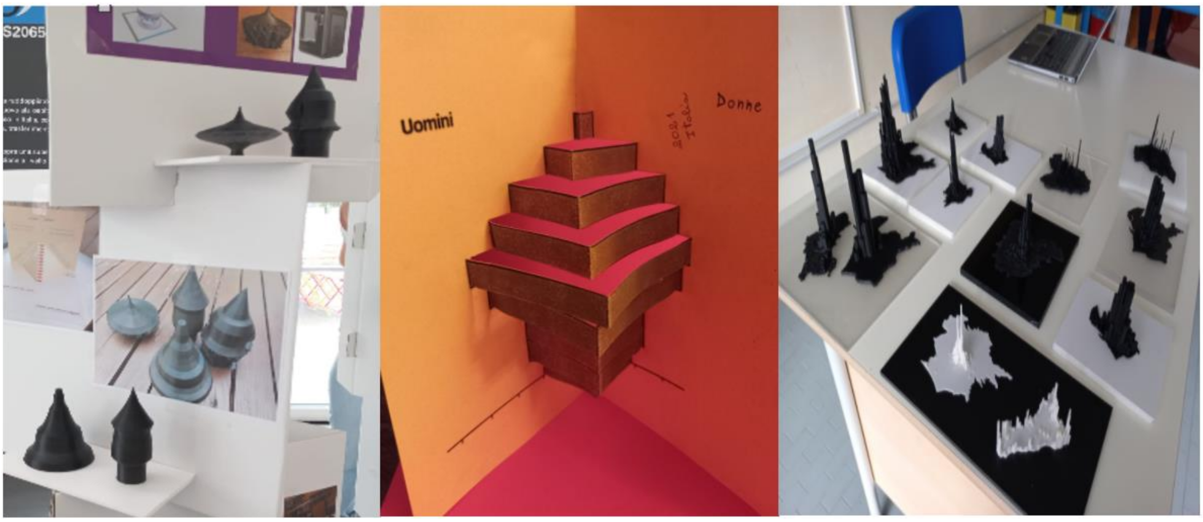

Figure 1. *left:* 3d printed age pyramids; *centre:* kirigami age pyramid; *right:* 3d printed models of mapped geolocalized data

There was a time in the past when Art was normally used as a complementary support in explaining Science and materialize concepts, just think of anathomy tables or painted herbals. Explanation greatly beneficiated by the aesthetical fascination of the artifacts. The same enchantment can be triggered also in Statistics. Students noticed that the 3D printed models with no exception first captivated the public attention because they looked beautiful and only then fully revealed their informative nature.

As they investigated the research question about migration flows and demographic trends, students stumbled on the clean and sinuous lines of the State of the World Age Pyramids by the French artist and designer Mathieu Lehanneur. It was love at first sight and they immediately looked for a way to replicate them. They successfully experimented with the free dynamic mathematic software GeoGebra: first they imported the 2D images of the pyramids taken from the web, then they traced the profile with the spline tool and within the 3D view menu produced the corresponding rotational solid which was finally exported as stl to Tinkercad for refining and cleaning before slicing (Cura) and 3D printing. While admiring and manipulating the pyramids the public is able to grasp the whole picture in seconds, easily make comparisons, formulate questions naturally inspired by the shapes and find most of the answers in the artifact as well. Ideally this manipulation should eventually bring to further reflection and even action. For instance the alien disk shape of Qatar 2022 triggered a critical reflection on the kind of immigration politics is implemented in the country and how respectful of workers' rights this can be. Comparison came quite naturally with what's now happening in EU which on one side is experiencing a fast aging population and consequently a shortage of workforce but on the other seems to resist immigration or at the very least have a very conflictual approach to it. The discussion involved many disciplines beyond Math and ICT while the heart of the discussion took place in L1 and L2, Geography, History, Social Studies, Civics with a touch of Art.

The same aesthetic fascination was experienced with Kirigami Statistics which offers a typical process of transforming data values into physical properties. It was all inspired by a recent article on kiriPhys or Kirigami Physics. First students were quickly introduced to kirigami technique, then they were given examples of kirigami models and asked to identify all the possible varying parameters. In such a process aesthetics - although still very important to captivate attention – is always functional. As a rule in tangible data all geometrical and material properties of the artifacts encode data, each one matching a different parameter. As the paper cut and fold constructions of kirigami present quite a number of properties embedding of the largest and most varied info is possible and complex dataset can be easily shown. Once again Kirigami Statistics may offer an intriguing and immersive data experience involving a plurality of senses from visual to haptic in a playful and fun process appealing to young students and more generally to a non-expert public.

Participatory Statistics

Quite a different multi-sensorial data experience which builds not only on individual emotions and aesthetics of the fruition but in collective effort is Participatory Statistics. It is a sort of living hands-on questionnaire where the public answers not just ticking boxes on a paper or filling an interview but actively engaging in completing puzzles, stretching threads around knobs, stacking pieces of different colors and shapes, lifting petals of paper flowers to express their liking, just to cite a few examples. In such a setting results can be immediately seen and read as they grow. The goal besides harvesting data in a cost-efficient and engaging way is also to produce learning, induce reflection and eventually empower people with control of their own data and the data of their community on topics of high interest. In this case data physicalization is not an object but rather a shared experience, where visitors take on the double role of protagonists and recipients of the visualization. What is designed top-down therefore is not the object but rather the stage on which participants will move contextually building data and its analytics.

Introducing participatory statistics practice in a school environment has a double value. On one side it can be an efficient tool in between gamification and challenge-based learning with a focus on prodding curiosity and urging students to delve into a topic for better answering. In this case results can configure as a kind of formative assessment. As a second option students can be asked to choose a topic of interest and design their own set of questions together with all the materials to run the experience. We successfully tested both ways. First we set up three boards with knobs containing questions on EU and opportunities for Youth and the public (60 Erasmus students from four different countries) had to thread its way with a woolen yarn choosing the different options (knobs).Answers were immediately visualized in the emerging cobweb. As a second step we asked our pupils how they would improve the experience if they had a free hand and we challenged them to design their own participatory statistical survey with the only limit of using the same boards and knobs.

Geolocalized Data: from Qgis to 3d printed and AR models

While addressing socioeconomic issues it is quite common to work with geolocalized data. Inspired by a post of the European Data Journalists Association on QGis models we decided to look for ways to reproduce any data values on maps with a 3D printer. In spite of a thorough online search no specific educational material was found so we asked for help to an ex-student who's now an environmental engineer (see co-author) which helped us in developing a simplified but fully detailed protocol for schools briefly summarized here. All used tools are free ones: QGis, FreeCad, Blender, Cura. 1. Choose the source files: csv with data (including latitude and longitude columns) and geojson for the base pedestal shape. 2. Import and dissolve the geojson file (thus eliminating all unnecessary internal borders) and save to SVG. 3. Upload the CSV file on QGis and upon choosing a matching Reference System all data will be visualized with dots on the map. 4. Filter the dots according to the analysis needs. 5. Apply a hexagonal grid and intersect with the base; count how many dots are in each hexagon. The corresponding value will be proportionally extruded as height of the pillars with hexagonal base in Blender, 6. Add a negative minimal buffer to avoid hexagons overlapping. Save as shp. 7.Import the svg file in FreeCad, clean it and extrude the base saved as stl. This passage is necessary as Blender is not able to import svg files. 8. A QGis extension needs to be installed in Blender first. Then import both files and snatch together the hexagonal pillars and the base. 9. Once Blender stl file is ready and saved, Cura will be used for the slicing and the model will be finally delivered to the 3D printer. In step 2 if data are aggregated in regions or districts no dissolution nor grid will be necessary, just use *Vector → Analysis tools→ Count data in the polygons*.

However powerful these physical models have an intrinsic limitation: you need to be there where they are to touch and manipulatively explore them. To extend their potentialities and reach over to a wider audience the 3D printed objects were eventually turned into AR models. Such result was achieved with the use of the free apps 3DScanner (iOS) https://apps.apple.com/it/app/3d-scanner-app/id1419913995 and polycam (Android and iOS) https://poly.cam/. With an average of ninety shots from the smartphone camera in appropriate lighting the result is already rather satisfying and can be easily improved with more shots. A key factor is the right illumination set: in our case it was achieved with three lamps at the vertex of a triangular area where the model was standing. The shots were taken at different angles moving slightly in a continuous and increasing with 4 degrees steps approximately to regularly complete the circle around the model. See one of the prototypes by scanning the QR code in fig.2 with your smartphone and observe the model under each possible angle by rotation and zooming.

The augmented reality object is thought as the perfect complement to Data Journalism essays . A QR code embedded into the newspaper page provides to readers the access to the AR visualization. The solid models can actually be manipulated, rotated, zoomed to appreciate specific details exactly as it would be done with the physical object.

Figure 2. QR code to access the AR 3D printed data model

RESULTS

Within the BeReady Erasmus+ project we worked with 14 -19 years old students of a general science focused high school to design and test innovative practices in data communication thus creating multisensory data fruition experiences. Activities were quite diverse ranging from sonification to 3D printed models not to forget kyrigami statistics, interactive Data Art, data journalism papers with embedded AR models and participatory statistics. Our goal was to contribute to a new vision for statistical education and data literacy for responsible citizens of the XXI century.

Statistics rather than a section of the Math curriculum should be considered as a trans disciplinary tool for all subjects, a very powerful aid to correctly interpret the World around us: a tool for objective knowledge, informed decision making and subsequent action.

Our approach is an inclusive one since each technique is enriching data with a new layer of knowledge and interpretation, thus enhancing data consumption of individuals with a wide variety of learning styles from visually impaired to auditory and kinesthetics. But this modality is not inclusive for end users only: as it strongly appeals to students' creativity it can motivate those who have designers' and makers' approaches to learning.

Moreover there's an added value in this kind of undoubtedly time consuming projects. To design, develop and deliver effective outreach experiences you need to reflect and distill the datasets into their major characteristics. It's a metacognitive process of analysis and selection: what is not important and can be overlooked, what is the eventual main message, how it can be conveyed through the physical characteristics of the different models, how people will actually interact with the models. This reflection will induce in pupils a long-lasting and in-depth knowledge and the transversal competences developed as well as the meaningful use of the variety of digital tools mastered in the process will enrich the students' curriculum and hopefully become useful in many future contexts both at school and beyond..

Further on, with a perspective more focused on the entire secondary cycle, we may say that what is time consuming for regular class work is on the contrary very well suited to project-based autonomous work even in an online environment and is definitely calling for a teaching methodology adjustment, a reorganization of learning times and spaces in a more flexible and flipped classroom modality, empowering students with the responsibility of their own knowledge building.

All the materials produced are going to be published as OER in the BeReady project website https://www.beready.pw.edu.pl/ and uploaded in Scientix repository https://www.scientix.eu/projects/project-detail?articleId=1581500. They include detailed teachers' guides, tutorials of the tools, presentations, students' sheets and lists of resources. Most files are editable to ease customization.

REFERENCES

Arduino Education https://www.arduino.cc/education/visualization-with-arduino-and-processing

Corona A. (2019) https://datajournalism.com/read/longreads/lets-get-physical-how-to-represent-data-through-touch

Daneshzand, Forooozan & Perin, Charles & Carpendale, Sheelagh. (2022). KiriPhys: Exploring New Data Physicalization Opportunities. *IEEE Transactions on Visualization and Computer Graphics.* PP. 1-11. 10.1109/TVCG.2022.3209365. https://tinyurl.com/mvb7x6zm

Jansen Y., Dragicevic P., Isenberg P, Alexander, Karnik A., et al. (2015). Opportunities and Challenges for Data Physicalization. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI), ACM, Apr 2015, New York, NY, United States.* ff10.1145/2702123.2702180ff. ffhal-01120152f

Lehanneur M. https://www.mathieulehanneur.fr/project/state-of-the-world-sculptures-297