

USING JUPYTER NOTEBOOKS FOR VIEWING AND ANALYSING GEOSPATIAL DATA: TWO EXAMPLES FOR EMOTIONAL MAPS AND EDUCATION DATA

Gabriele Silveira Camara¹, Silvana Philippi Camboim¹, João Vitor Meza Bravo²

¹ Graduate Program in Geodetic Science, Federal University of Parana, Brazil - camaragabriele@gmail.com, silvanacamboim@gmail.com

² Institute of Geography, Graduate Program in Geography and Graduate Program in Agriculture and Geospatial Information, Federal University of Uberlandia - jvbravo@gmail.com

KEYWORDS: Open Science, Jupyter Notebooks, Google Colab, Education Data, Emotional Maps.

ABSTRACT:

This article presents two applications developed using Jupyter Notebook in the Google Colab, combining several Python libraries that enable an interactive environment to query, manipulate, analyse, and visualise spatial data. The first application is from an educational context within the MAPFOR project, aiming to elaborate an interactive map of the spatial distributions of teachers with higher education degrees or pedagogical complementation per vacancies in higher education courses. The Jupyter solutions were applied in MAPFOR to better communicate within the research team, mainly in the development area. The second application is a framework to analyse and visualise collaborative emotional mapping data in urban mobility, where the emotions were collected and represented through emojis. The computational notebook was applied in this emotional mapping to enable the interaction of users, without a SQL background, with spatial data stored in a database through widgets to analyse and visualise emotional spatial data. We developed these different contexts in a Jupyter Notebook to practice the FAIR principles, promote the Open Science movement, and Open Geospatial Resources. Finally, we aim to demonstrate the potential of using a mix of open geospatial technologies for generating solutions that disseminate geographic information.

1. INTRODUCTION

Scientific research is how modern society develops knowledge about the world and its phenomena, answering questions by testing hypotheses with valid methods. However, making impactful science is difficult once there is an increasing need for (super)budgets to collect data, buy equipment, and publish research results in relevant scientific vehicles (Cantrell and Collister, 2019; Tennant et al., 2016).

For instance, we are experiencing a challenging scenario promoted by the COVID-19 pandemic. The virus has reached populations worldwide in a few weeks, also due to the global interconnection of our society (Castells, 2009; Mas-Coma et al., 2020). The technology we developed during the last decades is responsible for that globalisation: and it is an outcome of scientific research of our society. However, despite all this intellectual development, it is noticeable that the virus spread has found a freeway into the people's misinformation (Apuke and Omar, 2021; van der Linden et al., 2020).

We understand that the lack of knowledge about Covid-19 could be filled by the large dataset of scientific research we – scientists – rapidly built in such a chaotic scenario. Meanwhile, this gap is being occupied by "fake news" (van der Linden et al., 2020), and it has happened because scientific research is made inaccessible for people outside academia and, sometimes, also restricted to researchers beyond the "paywalls" (Cantrell and Collister, 2019).

Moreover, here we accept that fake news is being spread once people cannot find reliable sources for refuting such absurd claims. Thus, the fake news phenomenon is considered one of those variables with negative impacts on the COVID-19 pandemic, promoting mistakes that lead people to take ineffective treatments and avoid vaccines or safety measures (van der Linden et al., 2020). So then, how could we – scientists

- modify this scenario? Somehow, delivering scientific knowledge as fast as we could and accessible for general people. However, as we told before, the scientific knowledge is still unobtainable, even for scientists behind the paywall (Papin-Ramcharan and Dawe, 2006; Poulin, 2004; Scheliga and Friesike, 2014).

Nevertheless, there is an increasing attempt to empower "the availability of scholarly works to read and reuse" (Cantrell and Collister, 2019), here called the open access or open science movement (Kathawalla et al., 2021). Open Science does not have a formal definition (Arabito & Pitrelli, 2015; Vicente-Saez and Martinez-Fuentes, 2018), but it is a term that has its core based on knowledge, with required characteristic, like transparency, accessibility, shareability, collaboratively-developed (Vicente-Saez and Martinez-Fuentes, 2018), credibility and reproducibility (Kathawalla et al., 2021). This knowledge comprises scientific research and outputs, code, data, results, publications, information, and ideas (Vicente-Saez and Martinez-Fuentes, 2018).

The open concept has assumed different meanings over time and depending on the context. There is a consensus that the "open" emerged from "free", but it crossed the boundaries of "free", meaning rights, access, use, transparency, participation, and openness (Pomerantz and Peek, 2016). Furthermore, to support open science, it is necessary to rethink the research landscape and practice the FAIR principles, which means ensuring research is findable, accessible, interoperable, and reusable (Bruce and Cordewener, 2018). The open concept into geospatial information science has increased in importance since the beginning of the 21st century (Sui, 2014). Also, several free and open geospatial solutions have emerged into the context of web 2.0 technologies, once more individuals were interested in disseminating geospatial information into the internet (Elwood et al., 2012; Griffin and Fabrikant, 2012).

These sorts of solutions have created opportunities for the open science development into geospatial science (Sui, 2014) and produced such a fertile environment for disseminating the use of geographic information in several contexts. For example, the open geospatial information played a key role in the context of the COVID-19 pandemic by enabling governments to create policies and adopt strategies for holding the virus dissemination into the territories (Franch-Pardo et al., 2020).

Further, the rise of open geospatial data - open and interoperable - has been accompanied by the emergence of open GIS software, standards, and methods (Sui, 2014). In this context, the Open Geospatial Consortium (OSGeo) promotes software and standards for disseminating open geospatial information. The combination of open and interoperable geospatial technologies/resources - encouraged by the agents involved with OSGeo - plays an essential role in geospatial data science. This rise has not only taken place in the fields of health sciences but also is consolidating itself in the research and practice of the most diverse themes, for the sustainable management of cities, the environment, education, and several other aspects where the territorial understating is fundamental for knowledge building.

Nowadays, widely used open-source software that allows disseminating, accessing, and creating geospatial information - such as QGIS - may have extensions and customisations made with the Python language. Remarkably, Python is the most popular coding language used since 2018, holding this position for five years (Carbannelle, 2020). Tools and libraries for data analysis, including spatial applications, have been developed within an environment of a growing trend in adopting free software solutions through the decade. The Jupyter is an example of a non-profit and open-source project to data science and scientific computing (Project Jupyter, 2021), enabling the embodiment of the FAIR principles, being a tool for open science (Randles et al., 2017). Similarly, the Google Colaboratory (or Colab) is a Jupyter cloud environment maintained by Google that allows for a shared development environment with no client-side software installation required (Google, 2021a).

This article aims to present two different applications developed with Python in Google Colab to manipulate and visualise data in an Open Science environment. The projects are 1 - MAPFOR: an interactive geospatial query tool to visualise supply and demand of teacher training in Paraná State; and 2 - the Collaborative Emotional Mapping, a tool to explore maps of emotions in urban mobility and represent them using emojis. Therefore, here we demonstrate the potential of using a mix of open geospatial technologies for generating solutions that disseminate geographic information. For both study cases, the data is acquired from a spatial database through widgets interactions and allows the cartographic visualisation of phenomena we mapped. In the following sections, we present an overview of the used libraries, a context, the methods applied, and results achieved from both applications. Finally, in the conclusion section, we discuss the results, the notebook's approach, limitations, and future works.

2. GEOSPATIAL SOLUTIONS USING JUPYTER

Jupyter is a computational notebook browser-based tool that supports workflows, code, data, and visualisations (Randles et al., 2017; Perkel, 2018), combining user's input and output in

the same client-side. Likewise, Google Corporation launched the Colab, a hosted Jupyter notebook provided by a cloud-based service with no client-side software installation required (Google, 2021b; Carneiro et al., 2016).

The Jupyter notebooks in Colab were chosen to develop these applications due to the interaction, open-source, and shareability, promoting an even more collaborative environment. Additionally, the notebook could be hosted in GitHub or shared as an online link, where it is possible to set the permissions of access, ensuring security issues if the projects handle database credentials, as the applications developed in this research.

Nowadays, there are several examples of geospatial applications using computational notebooks. For example, Kiran et al. (2018) use the Jupyter Notebook as a tool for remote access and processing the DataCube API to geoprocessing satellite images, such as NDVI generation and unsupervised classification. Another example is the open-source implementation of a protocol for identifying problems in continuous movement data (Graser, 2021). Moreover, these geospatial Jupyter notebooks could interact with other open source initiatives like the Open Street Map, as in Boeing's urban street network analysis (2019).

Consequently, new libraries are emerging to optimise geospatial applications, such as Leafmap (leafmap.org), a Python package for performing advanced geospatial analysis in Jupyter Notebooks. Leafmap is developed upon several data science frameworks and geospatial packages widely used, such as Folium, Ipyleaflet, Ipywidgets, and Geopandas.

The Python libraries have considerable importance in the development of applications in notebooks. Hence, these tools enable the connection with external elements, such as databases, visualisation of graphic and cartographic solutions, manipulation of datasets, among other functions. Several libraries are native to Jupyter notebooks in Colab since the environment is executed in a cloud-based runtime. However, some particular libraries need to be installed in the runtime environment. Table 1 presents the Python libraries used in the development of both applications of this research.

Libraries	Application	Source
psycopg ipywidgets	Both	https://www.psycopg.org/ https://pypi.org/project/ipywidgets/
folium		https://python-visualization.github.io/folium/
configparser		https://pypi.org/project/configparser/
json	Emotional maps	https://docs.python.org/3/library/json.html
openrouteservice		https://github.com/GIScience/openrouteservice
matplotlib jenkspy	Mapfor	https://matplotlib.org/ https://pypi.org/project/jenkspy/

Table 1. Libraries used in the development of the research and their sources.

The main libraries used were:

- Psycopg: A PostgreSQL adapter for the Python programming language;

- Ipywidgets: Interactive HTML widgets for Jupyter notebooks and the IPython kernel;
- Folium: Manipulates data in Python and visualises it on a Leaflet map;
- ConfigParser: Read data from files, strings, and dictionaries;
- Json: Package to manipulate JSON data, as encoding and decoding;
- Openrouteservice: API that provides services using free geographic data from OpenStreetMap;
- Matplotlib: Library to create graphs in Python.
- Jenkspy: Compute the limits of classes using the natural breaks method.

Among the cited libraries, openrouteservice for the Collaborative Emotional Mapping and jenkspy of the MAPFOR are not native to Google Colab; thus, they were installed.

3. APPLICATION ONE: MAPFOR - INTERACTIVE MAP OF BASIC EDUCATION TEACHER TRAINING

The educational system in Brazil faced many challenges, such as illiteracy, dropout rates, grade repetition, and inefficiency in teaching training (Bomeny, 2003; Schwartzman, 2005). In addition, there is a significant gap in the availability of educational data, despite being essential for the development of public policies to tackle educational inequalities in Brazil (Gazeli, 2012).

Therefore, in 2018 the Brazilian Ministry of Education (MEC) developed the PARFOR project (in Portuguese Programa Nacional de Formação de Professores da Educação Básica). This project aims to provide higher or complementary guidance for teachers who lecture without superior formal education or who already have higher education, but teach in different areas that diverge from their formation or even bachelor teachers without basic teacher training (Brasil, 2018).

Due to the national program, it was established by the Federal University of Paraná, the MAPFOR project (in Portuguese, Mapeamento da Formação dos Professores do Estado do Paraná), which aims to ensure the educational quality offered in all stages of primary education in the State of Paraná (Camara and Camboim, 2020). In this project, state universities, especially the Federal University of Paraná, seek to map the supply and demand for undergraduate courses for teacher training to meet the specific demands of each region.

Besides the geolocation of 9511 schools, divided between different administrative dependencies, such as primary school, high school, federal institutions, and private schools (Camara and Camboim, 2020), we developed a spatial visualisation to research spatial distributions of professors with higher education degrees or pedagogical complementation concerning other variables vacancies in higher education courses. Furthermore, considering the multidisciplinary team of MAPFOR project, as professors, designers, cartographers, and developers, we choose to develop a Jupyter notebook for thematic maps instead of a traditional Geographical Information System to establish better communication within the research team.

3.1 Methods

The spatial data of Paraná State municipal boundaries were from an open data source produced by IBGE (in Portuguese

Instituto Brasileiro de Geografia e Estatística). Furthermore, the educational data used is also from another open data source provided by INEP (In Portuguese Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), an institution associated with the Brazilian Ministry of Education, which produces the Scholar Census since 1995. In this project, the educational data used were the percentage of graduate professors per municipal boundaries and the information about the institutes of higher education, such as vacancies in higher education courses, both were from the Census of 2018. Finally, we selected as base map the OpenStreetMap tile layer. As already mentioned, the platform chosen for the project's development was the Jupyter notebooks through the Google Colab platform.

The data was stored in a PostgreSQL database with PostGIS extension. In order to connect the database with the Jupyter notebook, we used the Python library Psycopg, enabling the query and manipulation of data through the notebook. Furthermore, using the Ipywidgets library, tools have been added to the interface to allow users to interact with data and maps and create interactive queries. Consequently, through a combo box, the user can choose which higher education course he wants to see the number of vacancies compared to the percentage of graduated teachers.

Considering that data queries are customised, developing a method to symbolise the map dynamically is necessary. For example, the number of vacancies in undergraduate courses in teacher training was represented in proportional symbol maps. Furthermore, as the classification of numeric data changes with each query, we use the Jenkspy library to redo the classification by Jenks' method based on the user's choice. Thus, it was possible to automatically classify and determine the ranges of the data classes for each course.

Finally, the maps were developed using the Folium library. It was determined that the better way to visualise the spatial data and their symbology was through the OGC standard WMS (Web Map Service). Thus, the data has been allocated from the database on the map server Geoserver, allowing storing the layer's symbology in SLD (Styled Layer Descriptor).

This graphic (Figure 1) represents how the elements that compose this application interact with each other, enabling the system's reproducibility.

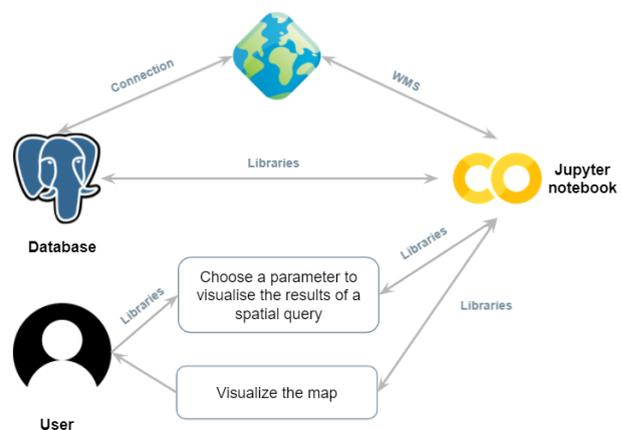


Figure 1. The system architecture of MAPFOR project.

3.2 Results

The Google Colab developed for MAPFOR project is available in [Github](https://github.com/GabrieleCamara/Mapfor/blob/main/maps_mapfor.ipynb) (https://github.com/GabrieleCamara/Mapfor/blob/main/maps_mapfor.ipynb).

After connecting with the database, the user can choose a higher education course (Figure 2a). The parameter is passed to the database and returns the number of vacancies of that course chosen. This data is classified using the natural breaks (Jenks) method. Figure 2b presents a graphic that illustrates the distribution of the data from a higher education course and the limit of classes determined by Jenks library.

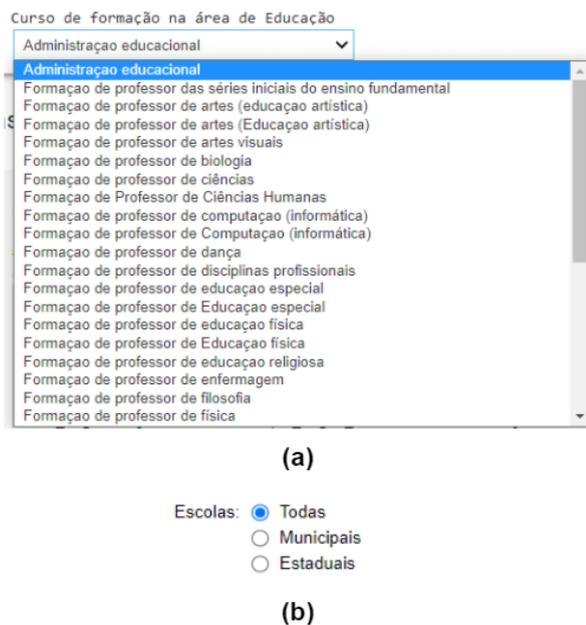


Figure 2. (a) Widget of higher education course. (b) Widget of options of aggregation of the percentage of graduate professors.

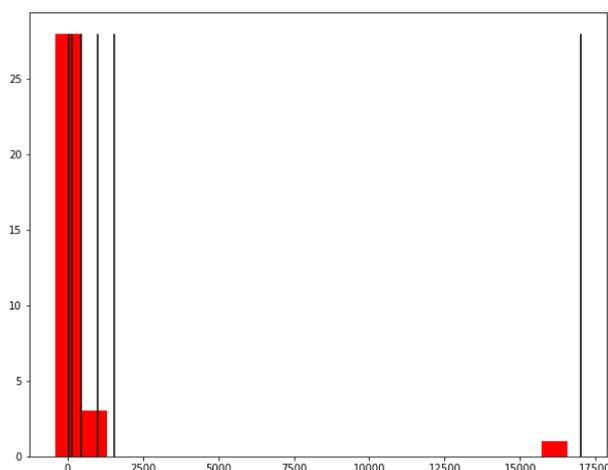


Figure 3. Graphic generated by Jenks library with the data distribution and the limits of classes.

Then, the user chooses to normalise the percentage of graduate professors aggregated by the total number of primary schools in

the municipal boundaries, the total number of primary schools, or the number of high schools (Figure 2b). Therefore, after executing cells that pass the parameter chosen to the database and Geoserver, the user visualises the map (Figure 4).

As it is an interactive map, the data can be presented in an aggregated form in a choropleth map with the 399 municipalities of Paraná, when the maximum level of distance is activated, until progressively closer to the individual school location. The choropleth map of municipal boundaries represents the percentage of graduate teachers, and the proportional punctual symbols represent the number of vacancies in the high educational course.

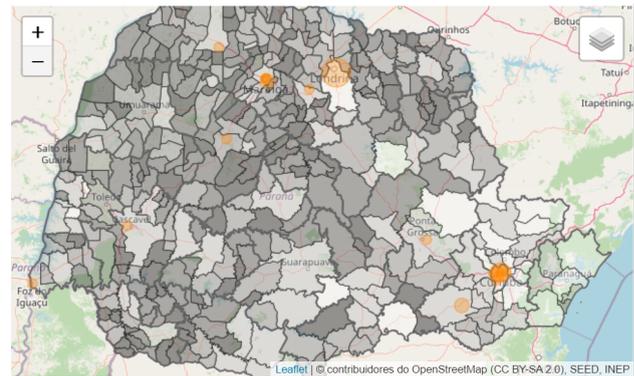


Figure 4. Percentage of graduate teachers aggregated with municipalities of Paraná, and the proportional punctual symbols represent the number of vacancies in the high educational course.

4. APPLICATION TWO: COLLABORATIVE MAP OF EMOTIONS IN THE CONTEXT OF URBAN MOBILITY

The Emotional Cartography collects and represents the emotions associated with an environment to understand their connection between individuals and places (White, 2007; Nold, 2009; Gartner, 2012). Thus, collaborative emotional mapping allows the representation of the emotions or sentiments experienced in a specific location of the space according to the individual's emotional bond, generating information with the environment's experience (Tuan, 1974; Camara, Camboim, and Bravo, 2021). This emotional information is being applied in the development of urban policies, including those on urban mobility, where the citizens contribute through their engagement in the Citizen-Centered Perspective, contributing as sensors (Fathullah and Willis, 2018; Goodchild, 2007). According to Camara, Camboim, and Bravo (2021), collaborative emotional mapping could be a tool to identify issues related to urban mobility.

Thus, a case study was developed with participants of an intermodal challenge in the city of Curitiba, the capital of Paraná state Brazil, for emotional mapping in the context of urban mobility, through the collection and representation of data using emojis. The participants indicated the emojis that represented their emotions when traveling along the path taken in the different modes (Camara, Camboim, and Bravo, 2021).

4.1 Methods

It was registered 426 points associated with positive (52%) and negative/neutral (48%) emotions. The data was collected through a paper map and later were organised and vectorised in QGIS software using as reference the Open Street Map, where each emoji was attributed to a line corresponding with a street. We modelled and implemented a spatial database in PostgreSQL with the extension PostGIS.

As in the MAPFOR project, we used the `psycopg` library to connect the application and the database, and the `ipywidgets` enabled the user to interact with the system. The options of widgets are from Python functions that get information from the database.

Figure 5 illustrated the elements that compose the system and intent the reproducibility of the application.

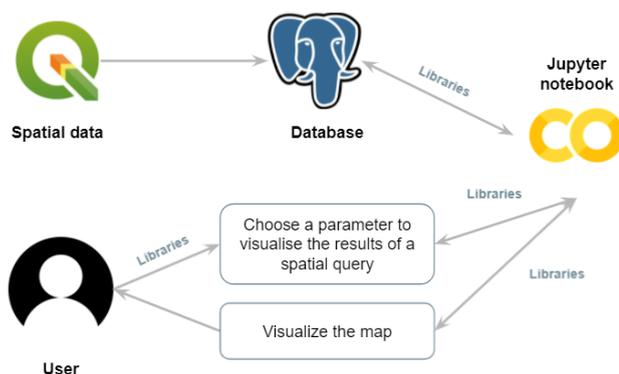


Figure 5. The system architecture of Collaborative Emotional Mapping Project.

To visualise the spatial data in the map, we used the tools of the Folium library instead of the WMS service, such as in application one. Consequently, the spatial information to be compatible with Folium tools must be in geojson format. Thus, it is necessary to develop a transformation between the PostGIS geometry WKT (Well-Known Text), returned in response to a query, and the Geojson format. This conversion was done using the `row_to_json` tool in a spatial query and the `json` library to decoding the data (Figure 6).

```
cursor.execute("""SELECT row_to_json(fc)
FROM ( SELECT 'FeatureCollection' As type, array_to_json(array_agg(f)) As features
FROM (SELECT 'Feature' As type, ST_AsGeoJSON(lg.geom_wgs)::json As geometry,
row_to_json((SELECT l FROM (SELECT gid) As l)) As properties
FROM emoc_selec As lg ) As f ) As fc;""")
```

Figure 6. SQL query that transforms a selected spatial data in WKT geometry to Geojson format.

The symbology of emotions is made through emojis from Emojitwo open-source set, available under a CC-BY 4.0 license ([emojitwo.github.io](https://github.com/emojitwo)). Each point on the street axis was customised with URL icons, a property of Folium's markers. Therefore, the list of emojis was stored in a Github repository, enabling a URL from the image, together with the final result of Jupyter notebooks developed.

We developed three types of maps; each one is within a Jupyter section. As mentioned before, the user has to run the cells sequentially to run the application, send the parameter chosen to the database, and execute the database responses.

To establish routes based on emotions desired, customising the sensory experience with the urban space, we developed the third map of application that enables the user to determine a route and then returned the streets within the emotionally mapped route. The directions are generated by Openroute Services (Table 1), as defined in Section 2.

4.2 Results

The Jupyter notebook developed for the collaborative emotional map is available in Github (https://github.com/GabrieleCamara/emotional_maps/blob/master/visualizer_emotional_maps.ipynb) in order to keep the project open-source. However, to interact with the application, it has to be opened in Google Colab.

The first map developed, the user chooses the emotion in the widget and sees the streets that emotion was assigned represented with the emoji (Figure 7). The user can activate and deactivate five layers, the emotions attributed into lines and points, all the paths taken in the collection of the data, and the streets classified with the gender of the participant who assigned that emotion.

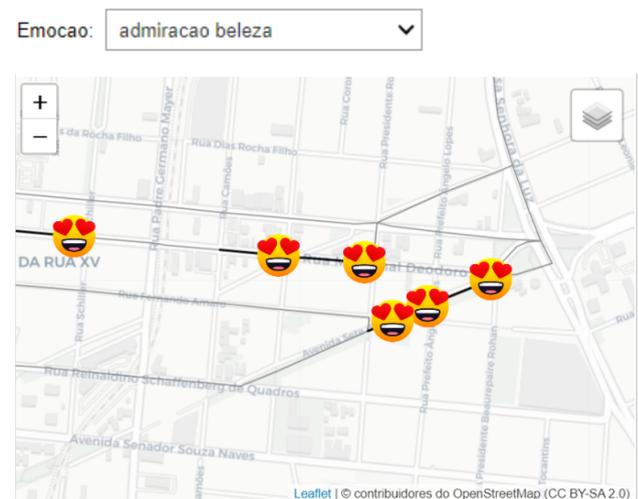


Figure 7. Streets assigned with the emotion Admiration (Beauty).

On the second map, it is possible to choose a street name and visualise which emojis were assigned to it to represent an emotion (Figure 8). The emojis are divided into three layers: emojis classified as positives extracts, emojis classified as negatives/neutral, and the street was chosen. Each emoji is also assigned in layers, thus the user can activate and deactivate the emojis individually.

Finally, in the third map, the user can type two addresses to define a route. Thus the application queries on the database if any street of the route was emotionally mapped, then return which emojis were associated with the street (Figure 9).

It is possible to activate and deactivate a group layer of negative/neutral emojis (Figure 9a) and positive emojis (Figure 9b). Besides, the blue marker represents the starting point and the green marker the ending point of the chosen route.

5. CONCLUSION

This article presents two applications in different contexts using Jupyter Notebooks as the environment of interaction, analysis, and visualisation of spatial data and integrating the notebook with Python libraries made it possible to create an interface for spatial queries on a database through widgets, allowing users access, exploring, and analysing data without specific knowledge, like SQL. The geospatial libraries used were also essential to visualise the spatial data, enabling the map update according to the parameter chosen by users. Additionally, the possibility of install additional libraries, such as Jenkspy and Openrouteservices, enriches the power of data manipulation and analysis.

The use of Jupyter Notebooks in several contexts, including the ones presented in this research, shows how it is possible to apply this approach to perform spatial visualisation and analysis in a research environment in an open way, with the advantage of allowing the monitoring and improvement of the code simultaneously by the team involved. Furthermore, the notebooks reinforce the open science and the principles for digital objects: Findable, Accessible, Interoperable, Reusable (FAIR), always aiming the reproducibility and interoperability.

Other advantages of projects that follow this approach are the openness of development processes and better communication and collaboration among a multidisciplinary research team, the possibility of reproducibility of methods, and the popularisation of codes in data science.

On the other hand, Jupyter also has limitations. The notebooks are great for exposing and modifying code, working and teaching content through coding, developing prototypes, and publishing peer-reviewed articles, for example. However, this environment is orientated to data scientists, students, and researchers with minimal programming languages and notebooks background. Despite the possibilities of interaction of user/application, projects developed in Jupyter Notebooks are not aimed at the end-user, such as a teacher searching about the training course in the case of study of MAPFOR, or a person doing a daily route based on emotions, in the case of study of collaborative emotional mapping. Thus, for these cases where the interface must be intuitive to not specialists user, it has to be developed type of solution.

Although Google Colab is an excellent solution for building collaborative projects, as it does not need local computing resources and promotes sharing through Google's already disseminated tools, the cloud platform has the disadvantage of being inflexible in the options it offers to manipulate the server. For a project with a greater capacity to use different libraries and customisations, it would be advised to use its own Jupyter server.

Finally, Jupyter notebooks will continue to have great potential for the future, considering the increase of open resources in the last two decades. Thus, with the growth of the Open Science movement and FAIR principles incorporation into the landscape of research, more open research and libraries are created and derived from others, establishing a virtuous cycle and expanding the contexts of application. The possibility of interaction in a single environment between the code and its outputs makes the learning curve fast, and the creation of codes is incorporated into the routine of teachers, students and researchers from distinct fields.

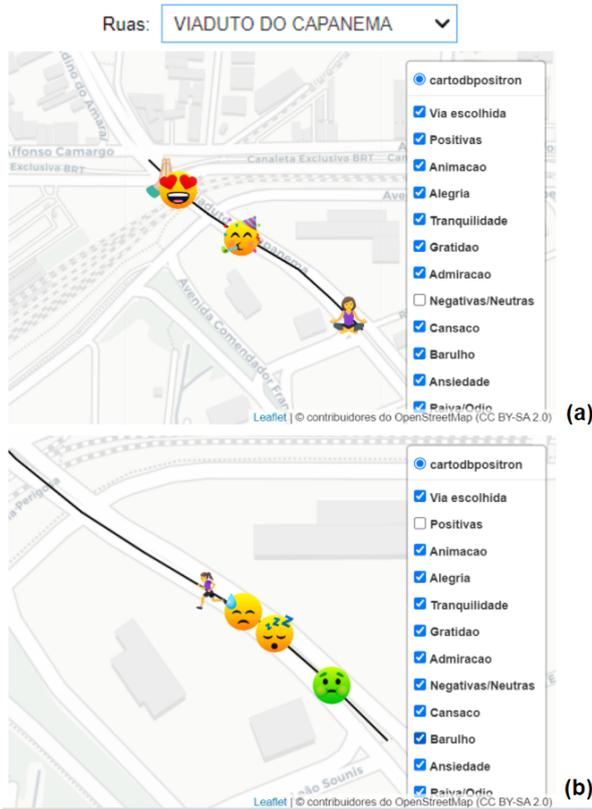


Figure 8. (a) Viaduto do Capanema assigned with positive emotions. (b) Viaduto do Capanema assigned with negative/neutral emotions.

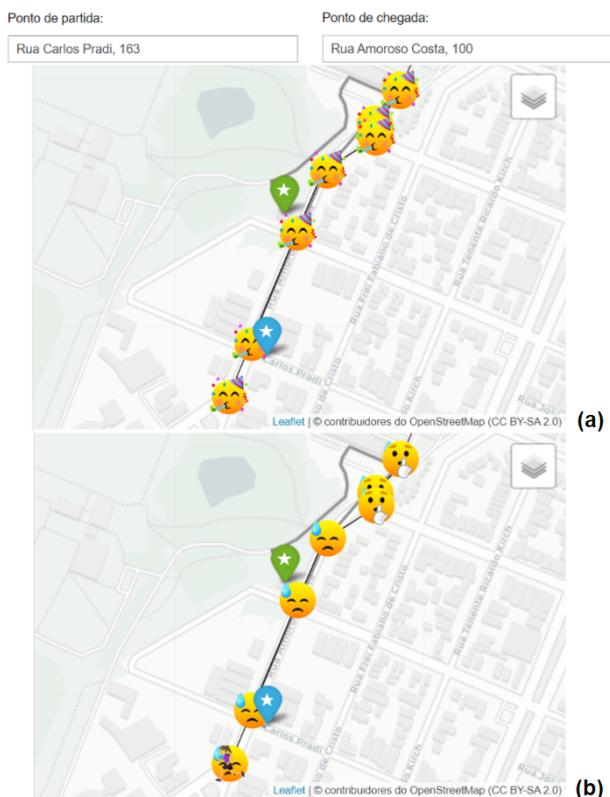


Figure 9. (a) Negative/neutral emotions of a route between two addresses. (b) Positive emotions of a route between two addresses.

REFERENCES

- Apuke, O.D., Omar, B., 2021: Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telemat. Informatics*, 56(101475). doi.org/10.1016/j.tele.2020.101475.
- Arabito, S., Pitrelli, N., 2015: Open Science training and education: Challenges and difficulties on the researchers' side and in public engagement. *Journal of Science Communication*, 14(4).
- Boeing, G., 2019: Urban Street Network Analysis in a Computational Notebook. *The Journal of ERSA*, 6(3), 39-51. doi.org/10.18335/region.v6i3.278
- Bomeny, H., 2003: Quando os números confirmam impressões: desafios da educação brasileira. *CPDOC*, 29f.
- Bruce, R., Cordewener, B., 2018: Open Science is all very well but how do you make it FAIR in practice? *LSE. blogs.lse.ac.uk/impactofsocialsciences/2018/07/26/open-science-is-all-very-well-but-how-do-you-make-it-fair-in-practice/* (12 June 2021)
- Camara, G.S., Camboim, S.P., Bravo, J.V.M., 2021: Collaborative Emotional Mapping as a Tool for Urban Mobility Planning. *Boletim de Ciências Geodésicas*, 27(e2021011). doi.10.1590/s1982-21702021000s00011.
- Camara, G.S., Camboim, S.P., 2020: A importância da disponibilização de dados educacionais na INDE: A experiência do MAPFOR - Mapeamento de Formação dos Professores no Estado do Paraná. *Anais 2º Simposio Brasileiro de Infraestrutura de Dados Espaciais*, 111-112.
- Cantrell, M.H., Collister, L.B., 2019: The status quo bias and the uptake of open access. *First Monday*, 24(7). doi.org/10.5210/fm.v24i7.10089
- Carbonnelle, P., 2020: PYPL PopularitY of Programming Language, CC BY 3.0. pypl.github.io/PYPL.html (12 June 2021)
- Carneiro, T., Nóbrega, R.V.M., Nepomuceno, T., Bian, G.B., Albuquerque, V. H.C., Filho, P.P.R., 2016: Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. *IEEE*, 4. doi.org/10.1109/ACCESS.2018.2874767
- Castells, M., 2009: *The Rise of the Network Society*. The Rise of the Network Society, 1. doi.org/10.1002/9781444319514
- Elwood, S., Goodchild, M.F., Sui, D.Z., 2012: Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Ann. Assoc. Am. Geogr.*, 102, 571–590.
- Franch-Pardo, I., Napoletano, B.M., Rosete-Verges, F., Billa, L., 2020: Spatial analysis and GIS in the study of COVID-19. *A review. Sci. Total Environ.* 739(140033). doi.org/10.1016/j.scitotenv.2020.140033
- Gazeli, P., 2012: Plano Nacional de Educação: Implicações para a Educação Infantil. *Revista Exitus*, 2(2).
- Google, 2021a. O que é o Colaboratory? colab.research.google.com/notebooks/intro.ipynb (12 June 2021).
- Google, 2021b. Colaboratory: Frequently asked Questions. research.google.com/colaboratory/faq.html (14 June 2021).
- Graser, A., 2021. An exploratory data analysis protocol for identifying problems in continuous movement data. *Journal of Location Based Services*, 15(2), 89-117. doi.org/10.1080/17489725.2021.1900612.
- Griffin, A.L., Fabrikant, S.I., 2012: More Maps, More Users, More Devices Means More Cartographic Challenges. *The Cartographic Journal*, 49, 298–301. doi.org/10.1179/0008704112z.00000000049
- Vicente-Saez, R., Martinez-Fuentes, C., 2018: Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research*, 88, 428-436. doi.org/10.1016/j.jbusres.2017.12.043
- Kathawalla, U.K., Silverstein, P., Syed, M., 2021: Easing Into Open Science: A Guide for Graduate Students and Their Advisors. *Collabra Psychol*, 7(1). doi.org/10.1525/collabra.18684
- Kiran, A., Gupta, P.K., Jha, A.K., Saran, S., 2018: Online Geoprocessing using Multi-Dimensional Gridded Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Science - ISPRS TC V Mid-term Symposium "Geospatial Technology - Pixel to People"*. IV(5). doi.org/10.5194/isprs-annals-IV-5-29-2018
- Mas-Coma, S., Jones, M.K., Marty, A.M., 2020: COVID-19 and globalisation. *One Heal*, 9(100132). doi.org/10.1016/j.onehlt.2020.100132
- Papin-Ramcharan, J.I., Dawe, R.A., 2006: Open access publishing: A developing country view. *First Monday*, 11(6). doi.org/10.5210/fm.v11i6.1332
- Perkel, J.M., 2018: Why Jupyter is data scientists' computational notebook of choice. *Nature*, 563, 145-146. doi.org/10.1109/ACCESS.2018.2874767
- Pomerantz, J., Peek, R., 2016: Fifty shades of open. *First Monday*, 21(5). doi.org/10.5210/fm.v21i5.6360
- Poulin, D., 2004: Open access to law in developing countries. *First Monday*, 9(12). doi.org/10.5210/fm.v9i12.1193
- Project Jupyter, 2021. jupyter.org (12 June 2021)
- Randles, B.M., Pasquetto, I.V., Golshan, M.S., Borgman, C.L., 2017: Using the Jupyter Notebook as a Tool for Open Science: An Empirical Study. *IEEE/ACM Joint Conference on Digital Libraries (JCDL)*, 1-2. doi: 10.1109/JCDL.2017.7991618.
- Robinson, A.C., Demšar, U., Moore, A.B., Buckley, A., Jiang, B., Field, K., Kraak, M.-J., Camboim, S.P., Sluter, C.R., 2017: Geospatial big data and cartography: research challenges and opportunities for making maps that matter. *International Journal of Cartography*, 3, 32–60. doi.org/10.1080/23729333.2016.1278151

Scheliga, K., Friesike, S., 2014: Putting open science into practice: A social dilemma? *First Monday*, 19(9). doi.org/10.5210/fm.v19i9.5381

Schwartzman, S., 2005: Os desafios da educação no Brasil. *Editora Nova Fronteira*.

Sui, D., 2014: Opportunities and Impediments for Open GIS. *Transactions in GIS*, 18, 1–24. doi.org/10.1111/tgis.12075

Tennant, J.P., Waldner, F., Jacques, D.C., Masuzzo, P., Collister, L.B., Hartgerink, C.H.J., 2016: The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, 5(632). doi.org/10.12688/f1000research.8460.3

Touya, G., Brando-escobar, C., 2013: Detecting Level-of-Detail Inconsistencies in Volunteered Geographic Information Data Sets. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 48(2), 134–143. doi.org/10.3138/carto.48.2.1836

van der Linden, S., Roozenbeek, J., Compton, J., 2020: Inoculating Against Fake News About COVID-19. *Frontiers Psychol*, 11, 2928. doi.org/10.3389/fpsyg.2020.566790