Search for life

in the Solar System and beyond

DVIZ Project FS23, Javier Izquierdo and Jannik Bundeli

https://github.com/javihslu/dviz_project.git

Motivation

Our motivation for choosing this project is based on our personal interests and the desire to gain new knowledge about the chosen topic. We are fascinated by the universe and astronomy. The endless expanse of space and the countless mysteries it holds inspired us to dive deeper into this subject. At the same time, we have been able to satisfy our own curiosity and gain new insights about the Universe.

Another driving force is the realization that by visualizing astronomical data and telling a story about the universe, we had a unique opportunity to share complex information in an engaging and easy-to-understand way that could capture the interest of others. This is exactly why we chose data storytelling. Our ultimate goal is to present to the public a new perspective on life.

It has been an exciting journey of exploration and learning, and we are proud to present the result of our efforts.

Targeted Audience

Our project aims at a broad audience, especially people who share our interest in the universe and astronomy. Our story has been deliberately designed so that no specific prior expertise is required to understand the visualizations and information. The design is purposely minimalistic and should not be interpreted as an academic/scientific paper. We are ignoring key knowledge for scholars, in sake of amusement for our audience.

Each visualization is accompanied by a detailed description to convey the context and meaning of the data presented. Additionally, there is a collection of curious and entertaining facts about the subject that we believe are unknown to the audience.

Although the project is accessible, it is still not intended for children. Rather, it is aimed at a wide range of ages, from teenagers to the elderly, who are passionate about the universe and the wonders of astronomy.

Our story is meant as a progressive reading. The first parts shall give context and basic understanding of our position in the universe. We intend to accompany the reader through a selection of topics that finally converge in our key message about life. The ending of the story is more intellectually demanding, due to the increase in complexity of the visualizations. This is done consciously by us, to enhance the strength of our final message by slightly overwhelming the reader. Our goal is for the reader to ponder the challenges of exploration of the universe and to realize the value of life on Earth. “We should praise for what we have.”

Exploratory work

We approached this project with the immature idea of talking about the solar system.
In our exploratory work process, we first looked at how to build our data story logically and coherently about the universe. We researched what data sources were available, particularly on websites such as:

- www.kaggle.com, https://nssdc.gsfc.nasa.gov/planetary/factsheet/
- https://exoplanets.nasa.gov/search-for-life/big-questions/
- https://www.space.com/35695-weirdest-solar-system-facts.html
- https://solarsystem.nasa.gov/planets/venus/overview

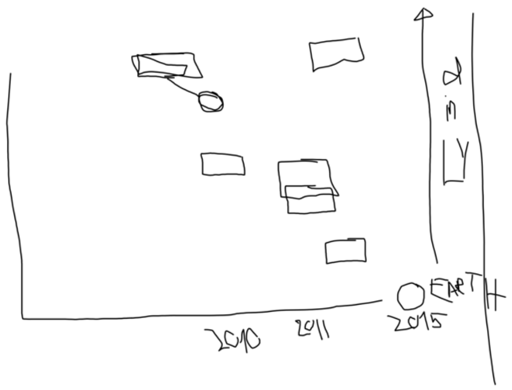
The dataset from Kaggle proved to be extremely rich and offered us many possibilities. In addition, we also had NASA's website where the data was also available, and we could enhance our data sets. We used various websites about curiosities to enrich our story and appeal to the audience’s curiosity.

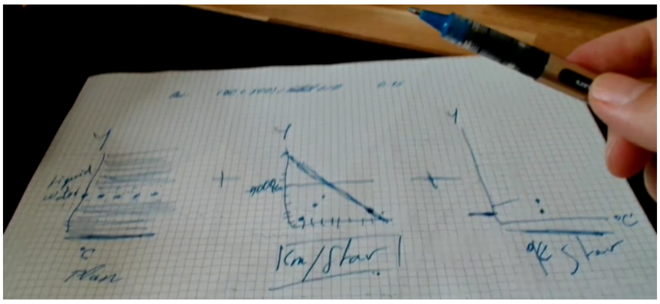
During our research we realized the endless number of stories we could write about, and the unique opportunity to create something not only entertaining but mindful. With this renovated purpose, we looked for inspiration on how to effectively conduct the audience through a narrative.
At this point, we decided to update the title of our story and focus our efforts on “life in the universe.”

This change of scope required further research on the matter and a curation of our sources and ideas. Armed with fresh knowledge and intentions, we carefully started a more refined design. We found great inspiration in the following projects and articles:

- https://exoplanets.nasa.gov/search-for-life/habitable-zone/
- https://ourworldindata.org/grapher/cumulative-exoplanets-by-method
- https://www.watson.ch/wissen/astronomie/578510805-das-sind-die-top-ten-der-exoplaneten-auf-denen-es-leben-geben-koennte
- https://www.esa.int/kids/de/lernen/Unser_Universum/Geschichte_des_Universums/Das_Universum#:~:text=Dazu%20gehören%20Lebewesen%2C%20Planeten%2C%20Sterne,Millionen%20oder%20Milliarden%20Sterne%20enthält.

Once we had decided the sections of the story and our message, we searched for rich visualizations on https://datavizcatalogue.com. This helped us drafting of our data story. Initially, we chose the most suitable charts to fit our data. We analyzed all dimensions in our data, butchered charts and put it all back together again. This stage was extremely demanding and frustrating at times. But we succeeded in displaying as much as needed, but in a clean way.





This process required revisiting the drawing board, searching for new sources, and rediscussing the narrative over multiple times. Only the amount of work we discarded could provide material for 10 projects.

Finally, we opted for a website format rendered with Quarto as the suitable vehicle for our story. We believe our exploratory work has been one of the most critical aspects of the process, and key to developing a comprehensive understanding of how to use the available data and effective communication.

Data Preparation

Our source data was originally very reach and required delicate handling. We have been using large datasets with extreme magnitudes. It required us to carefully study the data and the units. Astronomy researchers have chosen measurements that help their work, but these measurements can be challenging to understand for many people. We had to transform and normalize the data accordingly to make it more accessible. Some data features sometimes combined multiple units, which required feature engineering.

We added interactive visualizations to provide rich information to the audience. Unfortunately, we faced limitations in Quarto and most of this work had to be discarded. As a result, we have created a collection of custom visualizations that are very stable, although static.

Visualizations

Styling

The story style is lean and minimalistic, with as few numbers as possible. We have relied on position encodings for effective comparisons.
Thanks to Quarto, we have been able to unify most of the styling. The body is dark, fonts and shapes are white, and we use bright distinctive colors to highlight the most relevant features. All curiosities and complements to the main story are displayed on the margin within neutral-green boxes.
For this we have used the theme “darkly” from Quarto, some custom .scss, plotly’s color-hue and custom inline styling.

Fig1. Planets Volume

Chart type

Plotly, Subplots, circlify: One level packed circles: To compare the size of the planets in the Solar System. It complements the sidenote “Earth, a dwarf planet” from the margin.

Our aim is to give the audience a fast and accurate sense of the scale of the objects in the Solar System.

Encodings

- **Area:** The elements in the figure respect the real size-relations between the planets in the solar system.
- **Shape:** The circles portray planets
- **Color Hue:** Orange, to highlight the largest of the planets portrayed. And black for the rest of the planets.

Annotation

Only the largest planet is labelled.

Figure 2: Planet signatures

Chart type

Plotly, Pyplot: Multiple subplots with radar charts. Allows a fast comparison of multidimensional features among planets. These features values have been normalized with a standard-scaler. The result is a unique signature for each planet.

Encodings

- **Area:** Represents the unique signature of each planet and the magnitude of the features
- **Position:** With the positioning of the areas, we can quickly find similarities among planets without the need to understand the features. Subplots are distributed by distance from the Sun
- **Color Hue:** Bright distinctive colors in the areas, to enable fast matching.

Figure 3: Mean planet temperatures by distance from the Sun (lightspeed)

Chart type

Plotly Express: Scatter Plot with images and shapes. Allows a fast comparison of multidimensional features among planets, as well as classification thanks to the areas. The scales are customized for a quick comprehension of the dimensions. Distances have been computed into time. The images enable fast identification of the objects.

Encodings

- **Position:** Distance along the x-axis as distance from the Sun. Distance along y-Axis for temperature. Position of the areas for water states.
- **Color Hue:** Bright distinctive colors corresponding to water states, for fast classification within temperature thresholds. Yellow y-axis representing the sun
- **Slope/Distance:** Relation between distance from Sun and water state

Insets

The scatterplot points are images from the planets in the Solar System. Using images instead of colored points allows for fast identification and acts as legend.

Annotations

Each planet displays further details when mousing over them.

Fig4. Cumulative numbers of exoplanets discovered, by method

Chart type

With this graph we would like to show our audience how each discovery method changed over time and how many exoplanets have been discovered. For this purpose, the line chart was suitable as an infographic in a static and explanatory form. We reused the chart from (https://ourworldindata.org/grapher/cumulative-exoplanets-by-method)

Encodings

- **Slope**: How fast the cumulative number of exoplanets changes in respect to years. (No interpolation to see the trend)
- **Spatial position**: The audience can find out in which year how many exoplanets were discovered. With the additional grid lines on the y-axis the lines become more accurate.
- **Color hue**: By coloring the different methods the user gets a comparison and which method accounts for how many discoveries. We used the Set1 color palette from Seaborn, which is well distinguishable.

Fig5. Size and type of exoplanets compared to earth

Chart type

We decided to use a scatterplot, because we want to show the user the distribution of the mass and radius of the exoplanets and allow direct comparison. Furthermore, we created the chart in an interactive way with Plotly for a more exploratory chart.

Encodings

- **Spatial position**: This allows a detailed comparison of data points and recognizes clusters in an overview.
- **Color hue**: The 2nd encoding is the color differentiation for the planet types. This allows the audience to recognize patterns in the distribution. We used the T10 color palette from Plotly Express, which are well distinguishable.
- **Marker**: The habitable exoplanets were marked with a star and the Earth was marked with a diamond to recognize them quickly. The color of the planet type is kept.

Annotation

The labeling of the axes and the title needed to be self-explanatory because the data is normalized and compared to Earth. Each datapoint displays further details when hovering over them.

Fig6. How the top 10 habitable differ to other exoplanets and the Earth

Chart type

With this visualization we want to show the distribution of data from different properties. For this purpose, we created subplots and used Plotly to create an interactive violin plot for each property. With the focus on the shape, the distributions are quickly visible.

Encodings

- **Shape: With the violin plot shape, it becomes quickly clear how the data are distributed.** The wider the violin, the closer the data are to each other.
- **Position:** With the additional column for the values of the exoplanets and the Earth, they can be compared with the distribution.
- **Color hue:** The T10 color palette of Plotly Express can be used to distinguish the different data points of the exoplanets. For consistency, we have used the same palette as in the previous visualization.
- **Marker:** As in the previous visualization, we have marked the Earth with a diamond and the exoplanets with circles.

Annotation

It was important to label all axes in the right place with the correct unit. Each plot displays further details about quartiles, mean etc. when hovering over them.

Bibliography

Alibert, Y. 2013. "Visualization Cumulative Exoplanets, by Method." https://www.aanda.org/articles/aa/abs/2014/01/aa22293-13/aa22293-13.html.

"Astronomical Unit." n.d. https://en.wikipedia.org/wiki/Astronomical_unit#:~:text=The%20astronomical%20unit%20(symbol%3A%20au,)%20or%208.3%20light%2Dminutes.

Banerjee, Sourav. 2022. "Planet Dataset." https://www.kaggle.com/datasets/iamsouravbanerjee/planet-dataset.

"Basic Circle Packing Chart." 2022. https://www.python-graph-gallery.com/circular-packing-1-level-hierarchy/.

"Das Universum." 2022. https://www.esa.int/kids/de/lernen/Unser_Universum/Geschichte_des_Universums/Das_Universum#:~:text=Dazu%20gehören%20Lebewesen%2C%20Planeten%2C%20Sterne,Millionen%20oder%20Milliarden%20Sterne%20enthält.

"Datavizcatalogue." n.d. https://datavizcatalogue.com.

"Discovery Methods." n.d. NASA. https://exoplanets.nasa.gov/alien-worlds/ways-to-find-a-planet/#.

Huber, Daniel. 2015. "Top 10 Potentially Habitable Exoplanets." https://www.watson.ch/wissen/astronomie/578510805-das-sind-die-top-ten-der-exoplaneten-auf-denen-es-leben-geben-koennte.

Laboratory, NASAs Jet Propulsion. 2023. "Planets in Our Solar System." NASA. https://solarsystem.nasa.gov/planets/overview/.

Libraries, Plotly Graphing. 2022. "Radar Charts in Python." https://plotly.com/python/radar-chart/.

Lunar, and Planetary Institute. 2003. "Solar System Sizes." NASA. https://solarsystem.nasa.gov/resources/686/solar-system-sizes.

———. 2023. "10 Need-to-Know Things about Venus." NASA. https://solarsystem.nasa.gov/planets/venus/overview/.

"Planet types." 2022. https://exoplanets.nasa.gov/what-is-an-exoplanet/planet-types/overview/.

polsci. 2021. “Packed Circle Graph with Circlify, Matplotlib and Plotly.” https://gist.github.com/polsci/d9ecd38934f690ded5cd6ca061359814.
RenaudLN. n.d. “How to Subplot Radar Charts with 2 Traces Each.” https://community.plotly.com/t/how-to-subplot-radar-charts-with-2-traces-each/64440/2.
Singh, Harmanjot. 2020. “NASA Exoplanets Exploration Data.” https://www.kaggle.com/datasets/harmanjotsingh12js/nasa-exoplanets-exploration-data.
“The Habitable Zone.” 2021. https://exoplanets.nasa.gov/search-for-life/habitable-zone/.
“The Search for Life.” 2020. https://exoplanets.nasa.gov/search-for-life/why-we-search/.
Vicky Stein, Elizabeth Howell. 2022. “25 Weird and Wild Solar System Facts.” Future US Inc. https://www.space.com/35695-weirdest-solar-system-facts.html.
“Visualization Cumulative Exoplanets, by Method.” 2023. https://ourworldindata.org/grapher/cumulative-exoplanets-by-method.

Work Summary

Date	Hours	Done by	Description
19.04.23	2 each	Javier & Jannik	Setting up Git repository and working environment
22.04.23	4 each	Javier & Jannik	Brainstorming and collect first ideas, exploratory work for datasets
24.04.23	2 each	Javier & Jannik	Coaching, topic decision
29.04.23	3 each	Javier & Jannik	Brainstorming, research, task split up and distribution
30.04.23	3	Javier	Research and exploratory work about my own sections
03.05.23	4	Jannik	Research and exploratory work about my own sections
06.05.23	4 each	Javier & Jannik	Exchange of findings, brainstorming and story drafting with the message we want to communicate
13.05.23	3 each	Javier & Jannik	Research about the tools and the framework we use. Decision for Quarto.
23.05.23	3	Javier	Setting up Quarto and testing.
24.05.23	2	Jannik	Setting up the project report.
27.05.23	4 each	Javier & Jannik	Visualization Drafting, visualization research
30.05.23	2 each	Javier & Jannik	Brainstorming for final story drafting, but changing directions completely (see report)
03.06.23	4 each	Javier & Jannik	Brainstorming for final script
04.06.23	8	Jannik	Fig4 research and coding
	10	Javier	Fig1 research and coding

05.06.23	8	Jannik	Fig5 research and coding
06.06.23	12	Javier	Fig2 research and coding
	2	Jannik	Fig5 adjustments
07.06.23	3 each	Javier & Jannik	Code review and merging work
08.06.23	2 each	Javier & Jannik	Brainstorming and decision to make additional charts, improving the story
09.06.23	12	Jannik	Fig6 research and coding
	10	Javier	Fig3 research and coding
10.06.23	6 each	Javier & Jannik	Merging the code, debug
11.06.23	4 each	Javier & Jannik	Write final script for story and implement it
12.06.23	10 each	Javier & Jannik	Reviewed each other's work and final adjustments
13.06.23	10 each	Javier & Jannik	Prepare submission, create final report and record video
	Javier: 101 **Jannik: 99** **Total: 200**		