



Visualisation of “The UCS Satellite Database”

Shanmugapriya Murugavel

Student No: R00195696

For the module – Data Analytics & Visualization (DATA9005)

as part of the

**Master of Science in Data Science and Analytics, Department of
Mathematics, 7th March 2022**

Declaration of Authorship

I, Shanmugapriya Murugavel, declare that this assignment and the work presented in it are my own. I confirm that,

- This work was done wholly or mainly while in candidature for the Masters' degree at Munster Technological University
- Where any part of this thesis has previously been submitted for a degree or any other qualification at Munster Technological University or any other institution, this has been clearly stated
- Where I have consulted the published work of others, this is always clearly attributed
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this project report is entirely my own work
- I have acknowledged all main sources of help
- I understand that my project documentation may be stored in the library at MTU, and may be referenced by others in the future

Signed: Shanmugapriya Murugavel

Date: 07 March 2022

1. Introduction:

This report is prepared after performing analysis on “**The UCS Satellite Database**” [1]. The database contains a list of active man-made satellites that revolve around the Earth’s orbit. Natural satellites such as the moon and artificial satellites such as the International Space Station, orbits around the Earth. These satellites play a significant role in our day-to-day lives.

The dataset contains information like Name of the satellite, Country of the operator, Owner, Users, Purpose, Launch Mass, Expected Lifetime, Launch Site and Launch Vehicle, etc.,

“One of the best ways to explore and try to understand a large dataset is with visualisation. Place the numbers into a visual space and let your brain or your readers’ brain find the patterns.” – Nathan Yau[2]

Taking the above lines into consideration, the work is been carried out for exploring the satellites that were launched into the earth’s orbit. John Tukey, one of the most favourite statistician and the father of exploratory data analysis strongly believed in discovering the unexpected through pictures.

2. Importance of Data Visualization and Data Story Telling:

“Your data may hold tremendous amounts of potential value, but not an ounce of value can be created unless insights are uncovered and translated into actions or business outcomes” [3]. Data Storytelling is considered as the essential data science skill everyone needs. Data visualization expert Stephen Few said, *“Numbers have an important story to tell. They rely on you to give them a clear and convincing voice.”* Any insight worth sharing is probably best shared as a data story. Modern-day storytelling is often associated with the popular TED conference [4] series and its slogan of “Ideas Worth Spreading.”

From the dataset 1, we find that 350 satellites are currently revolving around the Earth’s Orbit. Let’s find some interesting facts about Satellites by finding out like: Why these Satellite were launched?, Who launched them?, When was the first Satellite launched, Which year launched the most number of Satellites, What purpose are these used for, Who are the end users of these Satellites, How were these launched?, Where they were launched?

The main idea behind choosing this Satellite dataset was to explore the way in which technology has evolved over the years. The growth in technology is predominantly seen due to the large number of satellites being launched for various purposes (*answer for question a*). The dataset 1 is not widely used for visualisation by anyone. Whereas, dataset 2 [7] there are a few other fellow members who have done visualisations using Python (*answer for question b*).

3. Data Visualisation:

R Shiny app highlighting the 5W questions to answer about Satellites

Shiny is a web-application based framework which is designed using R code. It mainly consists of the front-end designs where the user interacts in the UI part and the server section which renders our back-end activities and displays the output back to the user. The method of 5W or the questioning method is used to analyse the dataset. The shiny app is designed in such a way that each tab answers one of them.

Implementing plots using plotly, ggplot2, choropleth was carried out easily. But integrating those plots into the R Shiny app was a time-consuming process. There were plots where only country codes can be used for marking on the world map, there a separate dataset 3 – CountryCodes.csv was loaded and joins (*answer to d*) were used to plot in the choropleth map (figure 2).

3.1: When were these many Satellites launched?

To have a clear understanding of number of satellites launched we must understand when the first Satellite was launched, how many are launched every year. Here comparison is made year-wise and based upon different countries. This highlights the major theoretical consideration stated by Edward Tufte in his talk (*answer for question e*).

3.1.1: Lollipop Plot showing year-wise count of satellites launched

When these Satellites were launched?

The year of launch of satellites are represented using two plots. Select the type of plot to display.

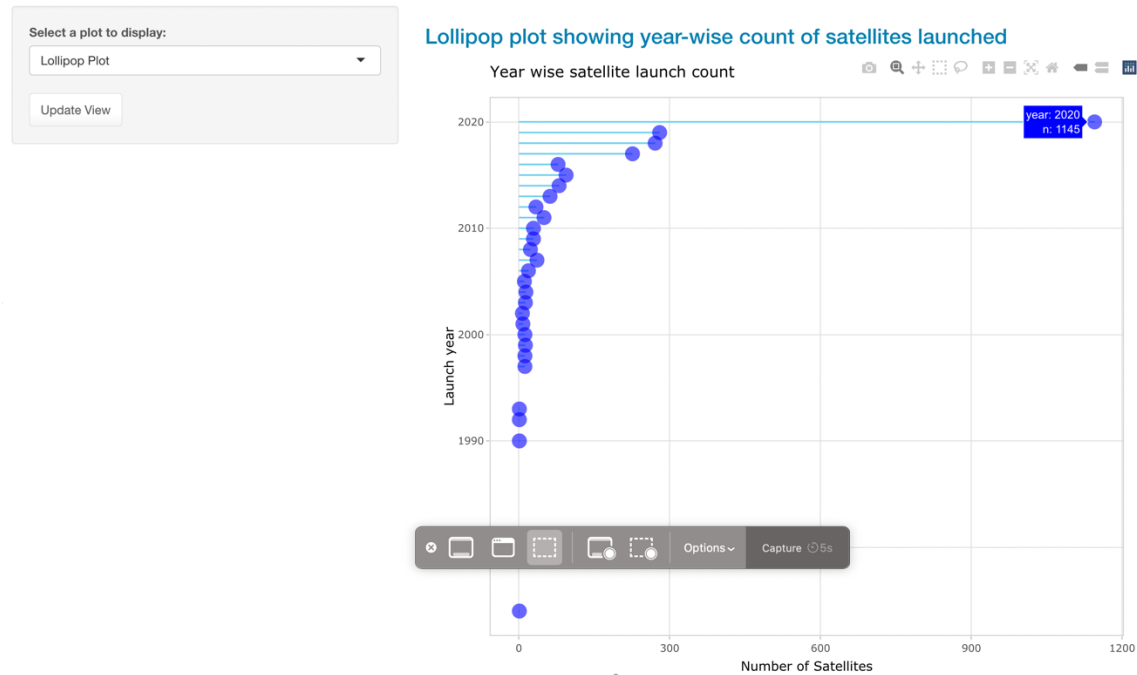
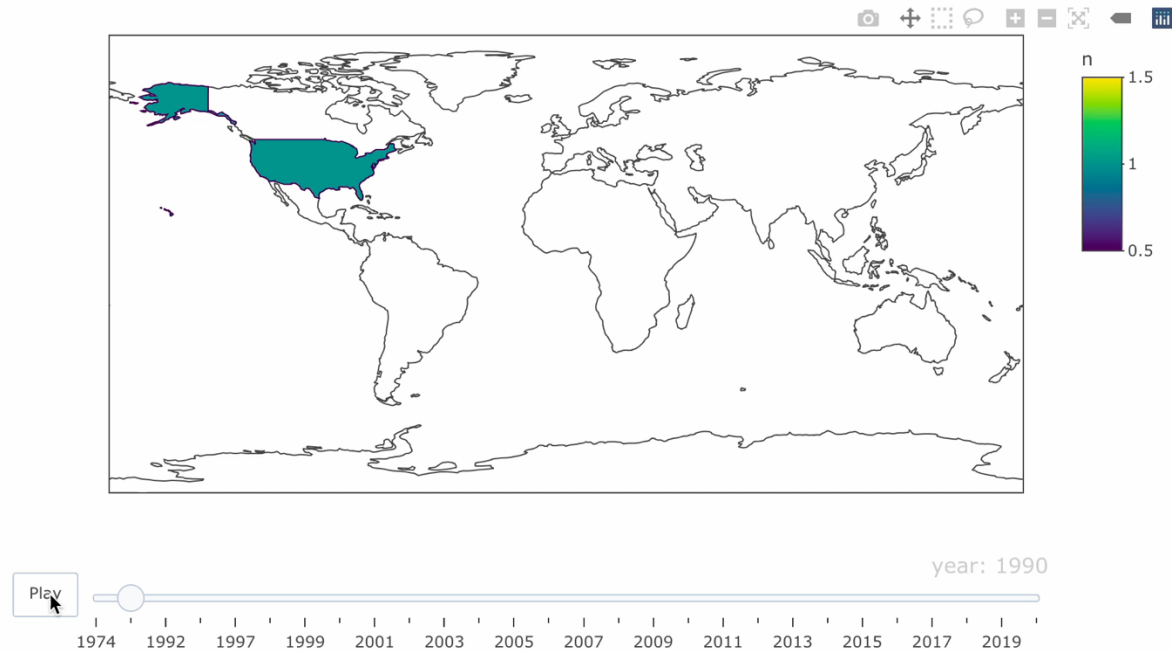


Figure 1: Lollipop plot representing number of Satellites launched yearly

From the above visual we can infer that in the year 1974 the first Satellite was launched. As years passed by gradually the number of Satellite launches are also increasing. In 2020, the highest number of Satellites were launched.

3.1.2: Choropleth Map showing year-wise count of satellites launched for different country

Choropleth map showing year-wise count of satellites launched in animation



Video 1 (*plays on double-clicking the above image*) : Choropleth map indicating number of satellites launched according to year

The above Choropleth plot helps us to show comparisons (*major principle of Edward Tufte*) of the number of satellites that were launched by various countries in different time periods. The animation the graph provides helps us to differentiate and compare them country-wise (*answer to question e*).

3.2: What time of the year, most of the launches are made?

To further understand the spread over each month a percent stacked bar plot is used for visualization.

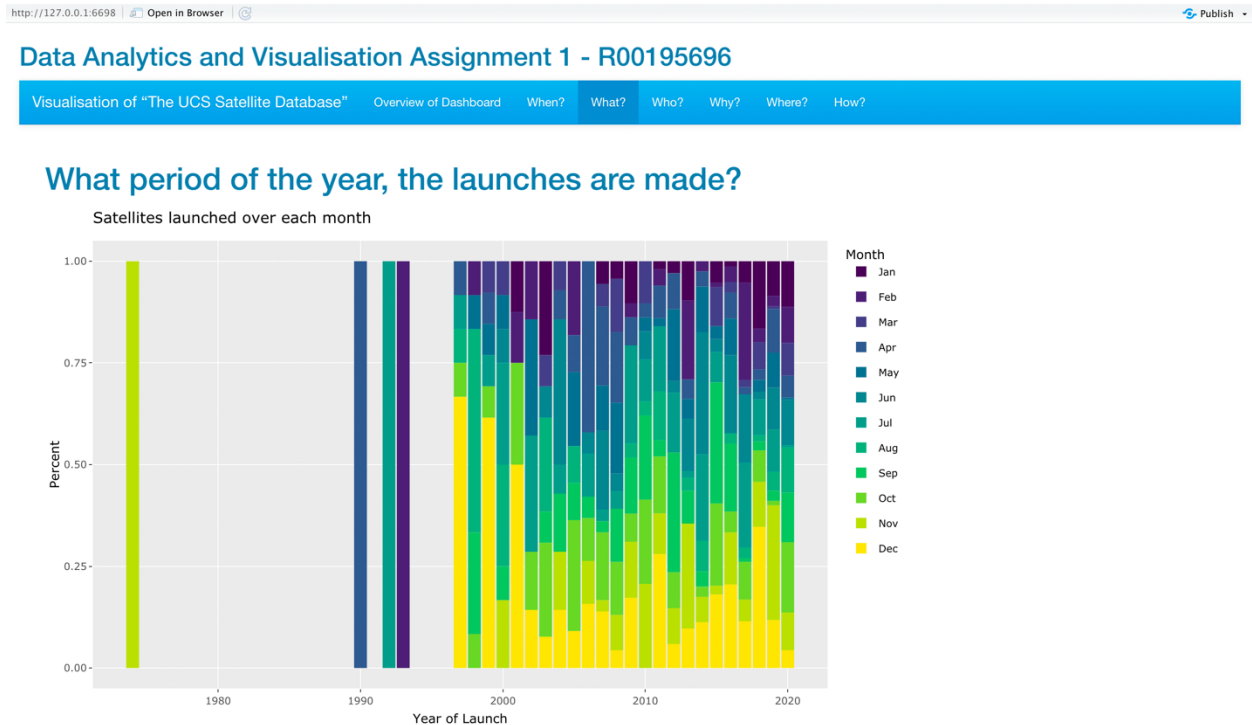


Figure 2: Percentage of Satellites launched over the month of each year

Here the colour yellow seems to be dominant which indicates a greater number of launches are done over the year end for the past 2 decades. In the year 2020, contribution of each month seems to be equal.

3.3: Why are these Satellites launched?

While there are thousands of new satellites being launched every year, let's find more about the purpose of launching these.



Figure 3: Wordcloud representing various purposes of Satellite launch

From the above plot, most of the satellites are used for Communication purpose followed by Earth Observatory and Technology Development. Apart from these major purposes there are many other fields for which they are launched over the years.

Advancing further let's see the major classifications under the Earth Observation purpose.

Different purposes of Earth observing satellites

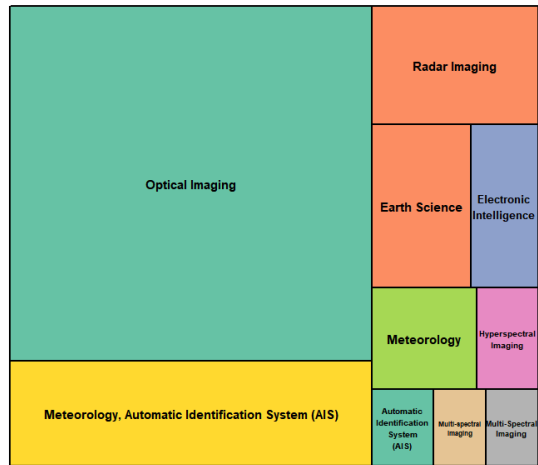


Figure 4: Treemap representing categories under the Earth Observing Satellites
From the above plots, we can find that Optical Imaging tops the list.

3.4: Who launches these Satellites?

After knowing when and why they are launched let's find who are the launchers.

Data Analytics and Visualisation Assignment 1 - R00195696

Visualisation of "The UCS Satellite Database" Overview of Dashboard When? What? Who? Why? Where? How?

Who launched these satellites?

Number of satellites launched across the world

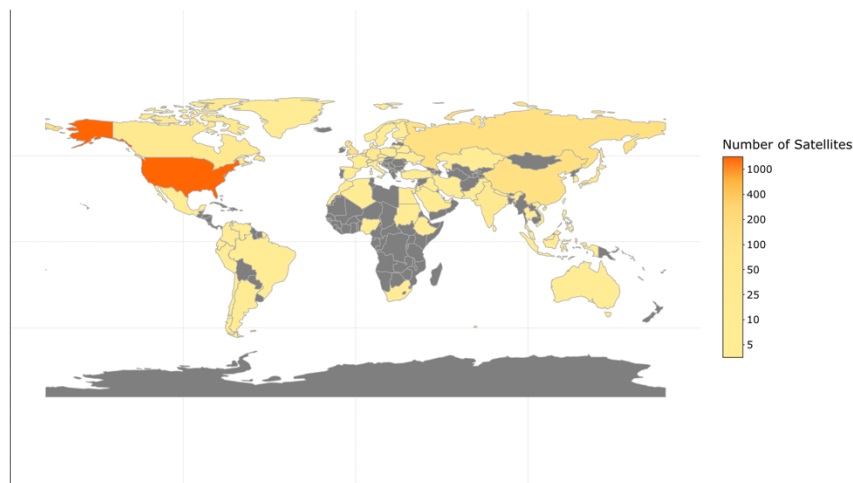


Figure 5: Map showing Satellites launched by different countries across the world.

From the above interpretation of launches every year, it was significantly seen that the United States had an increasing count. To verify the assumption, the total number of launches are marked in the world plot above inspired from AnyChart [8]. Darker-coloured countries are areas that had relatively higher number of launches. Here, USA has launched the greatest number of Satellites, while there are few countries that are yet to launch any. The different shades of yellow for different countries shows the distribution of launches done by them.

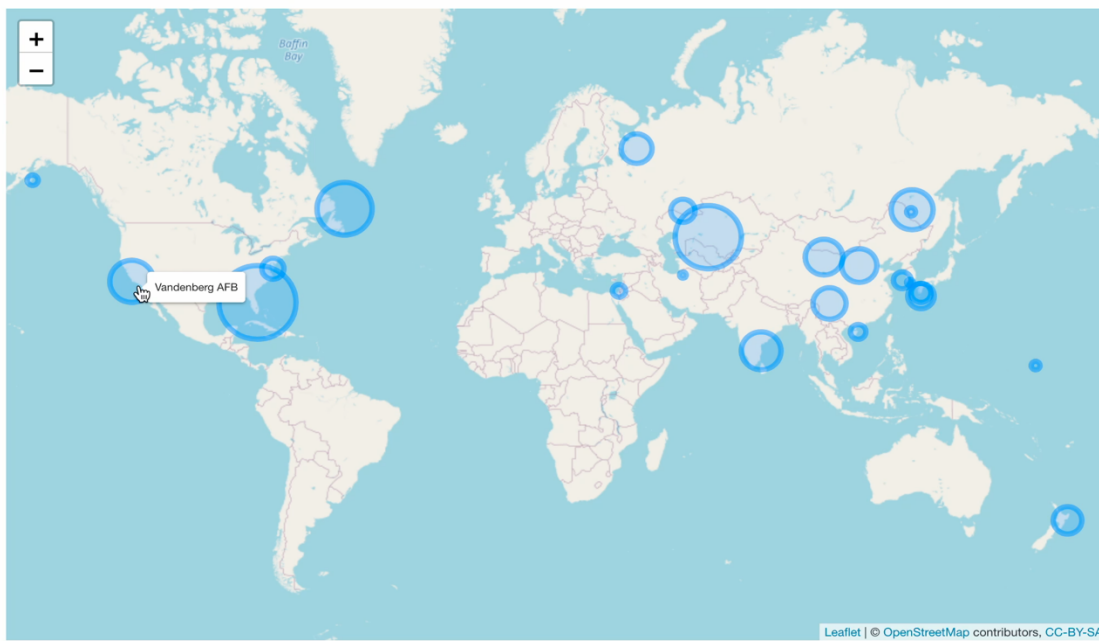
3.5: Where these Satellites are launched?

The site for launching the spacecraft is called the spaceport or the cosmodrome. The leaflet plot is used to mark all the spaceports in various countries. Here the radius of the points gives us the number of launches. Bigger the radius, higher the launches.

Data Analytics and Visualisation Assignment 1 - R00195696

Visualisation of "The UCS Satellite Database" Overview of Dashboard When? What? Who? Why? **Where?** How?

Satellite Spaceports Around the World



This is a plot of all the spaceports where the satellites are being launched every year.

Video 2 (plays on double – click): Leaflet plot showing spaceports and its related information

Conclusion and Future Works:

The satellite data gave amazing insight about the different satellites and their purposes. There were answers found for some of the Who's and Why's. But there is column with city name of the launch site then weather data could be combined with the satellite data and more insightful visualisations can be created. Like the most appropriate or common weather that satellites are launched (*answer to g*).

Reference:

- [1] **UCS Satellite Database:** Source of the database. <https://www.ucsusa.org/resources/satellite-database#.Vq8gsubGF31>.
- [2] **Visualize This: The FlowingData Guide to Design, Visualization and Statistics.** A book written by **Nathan Yau**. Pg. xvi
- [3] **Data Storytelling: The Essential Data Science Skill Everyone Needs:** Article by Forbes <https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/?sh=5d7e9b0952ad>
- [4] **TED Ideas Worth Spreading:** <https://www.ted.com/talks>
- [5] **R Graph Gallery** - <https://www.r-graph-gallery.com/301-custom-lollipop-chart.html#horiz>
- [6] **Data Visualization – Tableau:** <https://www.tableau.com/learn/articles/data-visualization>
- [7] **All Space Missions from 1957 – Dataset 2** taken from Kaggle <https://www.kaggle.com/agirlcoding/all-space-missions-from-1957>
- [8] **Any Chart website -** https://www.anychart.com/products/anymap/gallery/Maps_General_Features/World_Choropleth_Map.php
- [9] **Making Maps with R – Nico Hanh** https://bookdown.org/nicohahn/making_maps_with_r5/docs/introduction.html
- [10] **High Quality Graphics in R** - <https://web.stanford.edu/class/bios221/book/Chap-Graphics.html>