Star Travel - Future of the Human Race

1 Introduction

As climate change becomes increasingly more urgent, the topic of space travel and transporting humans to other planets has become a "hot" topic for discussion. Specifically, we see ideas related to space travel in order to inhabit other planets in popular works of science fiction.

In this report, we aim to determine the plausibility of star travel by using machine learning techniques to analyze various factors that would go into successful space travel. Such factors include the rate of space rocket technology with respect to speed and the rate of resource use. Using this, we want to determine if it will be feasible for humans to travel to other planets in a reasonable amount of time.

To perform this analysis, our research data can be narrowed down to resource production or collection; the consumption of the resources; data from nearby planets; and the recorded speeds of rockets across the years. This data should theoretically provide us with an estimated time of travel to nearby planets and possibly the maximum speed we could achieve in the lifetime of planet Earth itself before we have depleted all of the resources. The resources we are focusing on are the commonly used energy sources such as gas, oil, coal, and uranium.

2 Data

2.1 Planet Data

Planet data was downloaded as a CSV from the NASA Exoplanet Archive [14]. Using Pandas, the CSV was partially read to extract the headings and the planetary database in separate reads. NASA calculates distance as parsecs, so this column was converted to lightyears with the help of Pandas operations on columns as that is conventionally more intuitive to understand in terms of scale.

2.2 Resource Consumption

We have collected datasets for the production, consumption, and remaining reserves of coal, gas, and oil [37][39][40]. The datasets included many countries and varying units of measurement, so to keep the data consistent, the units were converted to TWh. Furthermore, we will be looking at the Earth's dataset as a whole, rather than individual countries. In addition, the consumption of those resources also emits CO2, which can cause harm to our planet, hence, we have collected data on CO2 emission and level as well[37][41][42].

Uranium was collected as it is a zero-emission CO2 power source [32]. It is also a possible future propulsion source for future space travel [31]. Data was collected for the current yearly amount mined as well as the yearly amount required for power plants[27][28].

We also included water and monitored the refresh and consumption rate to determine if there is a point where we will run out. Furthermore, in order to travel large distances in space, we need to make sure we have extra supply to put on a spaceship for the group undergoing interstellar travel. Data on water renewable sources and the withdrawal rate was collected [24][25].

2.3 Speed Data

In order to make predictions about the future speed at which we can travel, we can consider the rate at which humans have been able to improve vehicle speed over time. More specifically, we can look at the velocity of past rockets and model how they have improved over time to make our predictions.

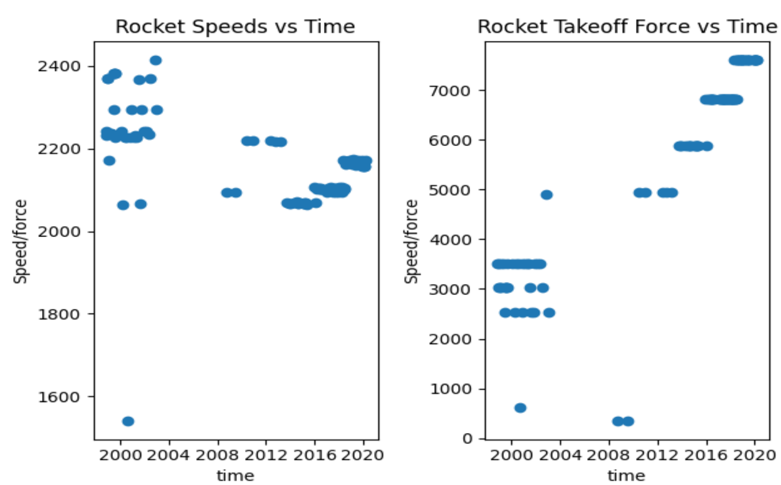
Unfortunately, space agencies do not report on maximum velocities of their space crafts to the public, hence this type of data is not feasible. However, data on the lift off thrust of various rockets was available. The force at which a rocket is launched will determine its final velocity when it reaches space. This is due to the absence of opposing forces when a rocket leaves the Earth's atmosphere. That is, the velocity at which a rocket exits the Earth's atmosphere is the velocity it will continue to travel. This velocity is directly proportional to the initial speed of the rocket, its mass, and the distance it traveled. Using the force, we can compute speed as follows:
First, by Newton's second law, we have the following,

$$f = ma - mg \ \rightarrow a = \frac{f + mg}{m}$$

Where m and a is the mass and acceleration of the rocket respectively. Then we can use the following kinematic equation with the initial velocity at 0 (as the rocket begins at rest), to approximate the top speed of the rocket.

$$v_f^2 = v_i^2 + 2ad \ \rightarrow \ v_f^2 = 0 + 2\left(\frac{f + mg}{m}\right)d \ \rightarrow \ v_f = \sqrt{2\left(\frac{f + mg}{m}\right)d}$$

Where d is the distance from the launch site to the cusp of the Earth's atmosphere. The data we are considering comes from Kaggle [20], which details various flight information for rockets from Arianespace, Boeing, Brazilian Space Agency, European Space Agency, Martin Marietta, Space X, and US Air Force between 1998 and 2020. However, this data did not contain each rocket's mass but we were able to manually go to the Wikipedia page (and one NASA data sheet [4]) of each rocket to grab its mass and insert it as a new column.[1]-[3],[5]-[8],[10]-[13],[16]-[19][21] Furthermore, we used the distance of 100 km from sea level to the edge of the Earth's atmosphere minus the height of each launch site above sea level as our distance input, d. This height above sea level was acquired by submitting the coordinates of each launch site into Google Earth. Given that much of this data had to be manually inputted, these calculations were done directly in Excel before being moved to Pandas.
When considering which data to use, we removed failed launches. There were 3 outliers of force data that were much greater than the rest. Upon further inspection, we note all three data points came from Space X's Falcon Heavy [13], but there was no plausible reason to ignore them, unless they drastically affect the model.



We observe from the plot above that the rocket speed data actually seems to decrease over time. We note that this might have occurred due to errors in calculations as we did not consider wind resistance, which can vary based on a rocket's design, or the change in gravity as the rocket moves away from the Earth's center. Furthermore, while we could add the payload mass given to the rocket mass (available in the Kaggle dataset), some payload masses were unavailable or 'confidential'. Hence, we will consider force only for our analysis as it is directly related to velocity in some capacity.

3 Methods

3. 1 Planetary Data

Using Stellar classifications, a table with less than 15 light years was formed consisting of the closest super earth, earth-like and dissimilar to earth planets - the stellar types are K, G, M, respectively [2].

3. 2 Fitting Lift Off Force

Observing the plot of the lift off force, we can estimate that it will have a polynomial fit. Hence, we can use a polynomial regression from Python's sci-kit learn library to create a model. To apply this method, we need to transform the dates from the data into a format that is suitable for applying regression. That is, we can convert the dates into integers by determining the number of seconds that have passed since some epoch. As the earliest data point comes from November 1998, we can use January 1st 1998 as our epoch. From the plot, a quadratic fit looks most likely. We will use a polynomial of degree 2 as higher order polynomials may predict negative values for the future. Furthermore, we can apply this fit to a training set and validation set to consider any over fitting.

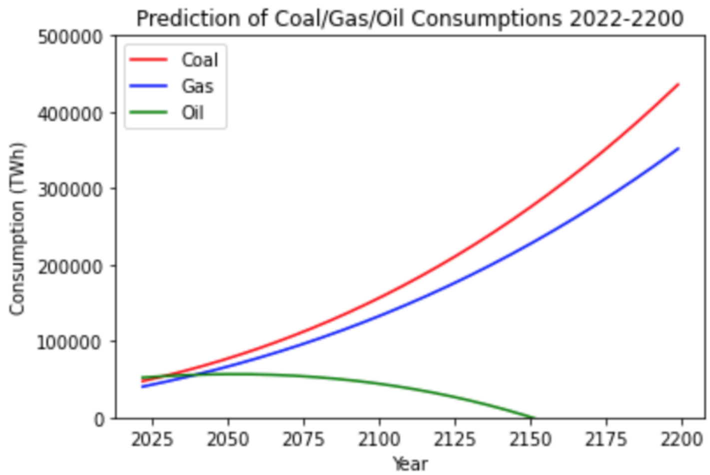
3. 3 Resources

After the initial cleanup, there is not much transforming needed since the date values are in the form of years, which are integers. To analyze the data, similar to the lift off force data, the plots of resources data also hints at a polynomial fit. We applied polynomial regression with a maximum degree of 5 on the data sets. We used sklearn.model_selection.train_test_split to split the data into training and validation sets. The training set was used to fit the model, and the model is scored on the validation set. The scores are all fairly good (mostly above 0.95) and there is not much difference between the training and validation sets, thus we can assume our model is fairly good and overfitting isn't an issue.

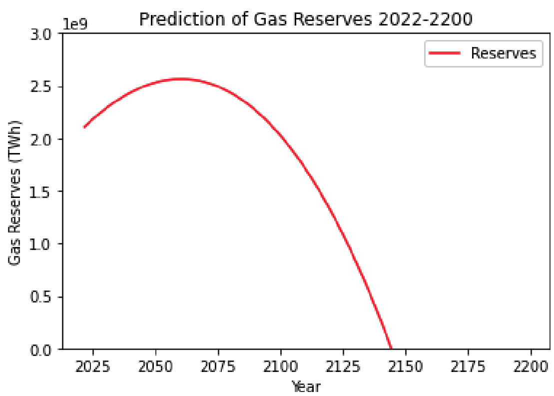
The water data was grouped by renewable water and withdrawal categories per year globally. There was a spike in renewable water data around the year 1992. Research showed that in the 1990s global efforts were made to develop methods to obtain fresh water [33][34]. Performing a linear regression test on the pre-1992 data obtained a p-value significantly greater than 0.05. The data for both the renewable water and water withdrawal data post-1992 showed a p-value significantly less than 0.05 which meant we could reject the null hypothesis of zero slope. The linear regression model was plotted to see the trend for future years.

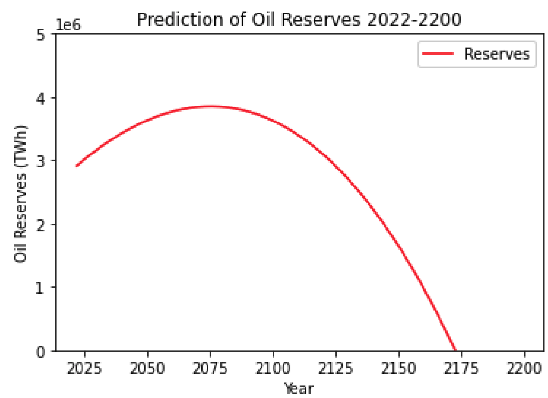
4 Results and Analysis

4.1 Resources

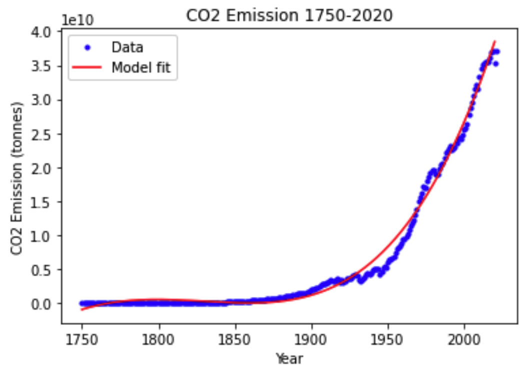


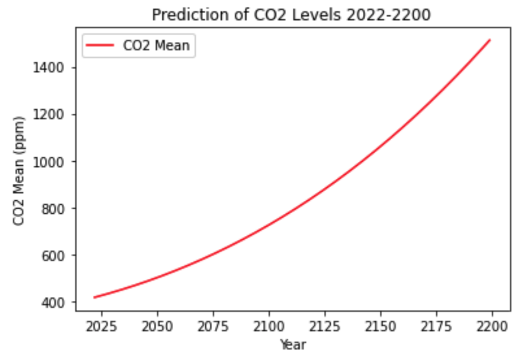
Our model predicts that coal and gas consumption will still be increasing, while our consumption of oil will be 0 in 2151, this prediction is most likely due to the recent trend in switching to electric cars, and realistically our oil consumption will still continue on.



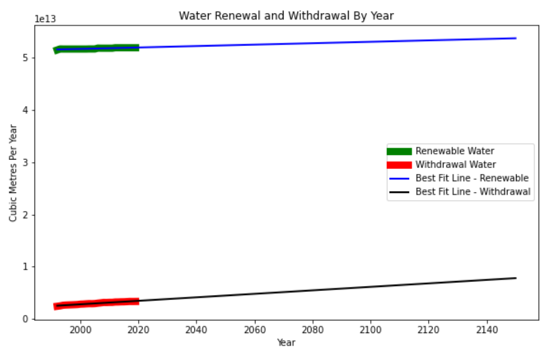


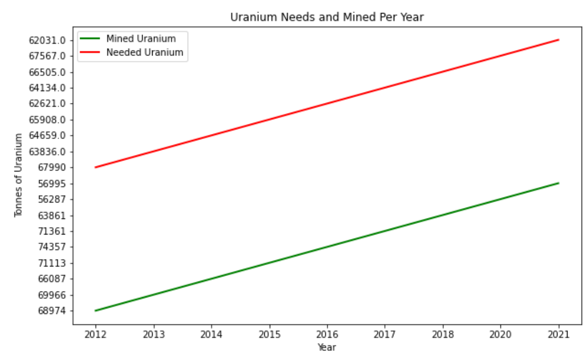
Our model predicts that gas reserves will empty at the year 2145, while our oil reserves will empty in 2173. The only data on coal reserves that we could find is in 2008, at that time we had 7.718e+06 TWh worth of coal. Going by the current consumption rate, we would have consumed 7.760e+06 TWh of coal from 2008 to 2099, thus running out of coal before 2100.





CO2 emissions have grown massively. Our model predicts that we will hit 500 ppm by 2050, and 1000 ppm in 2143. The linear correlation between coal, gas, and oil consumption, and CO2 emission is 0.86, 0.87, and 0.75 respectively, thus it can be assumed that our growing CO2 emission is caused by our vast increase in the consumption of these resources.





The water plot (left) indicates that we will not have a global water shortage for the foreseeable future.

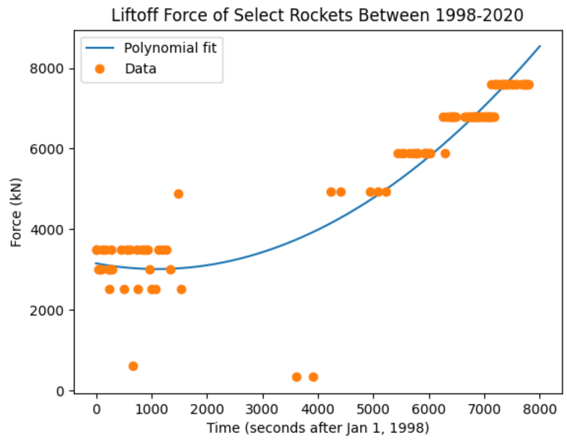
However, the Uranium plot (right) shows that we will be starved for Uranium throughout Earth history for both power plant needs and eventual nuclear propulsion methods for space travel.

4.2 Force Fit

Applying the polynomial regression to the force data, we get very low training and validation scores (about 0.4 and 0.3 respectively), hence we will ignore the outliers for a better fit. The model without the 3 outliers (force less than 10000 kN), is as follows.

$$F = 0.0000829x^2 - 0.0334x$$

When applying this model to our training data, we find the model returns a score of 0.91. This non-near-perfect result is expected as our data is very sparse between 2004 and 2008 and may be missing some features that can aid in a better fit.
The validation score was 0.83, which indicates there was likely some over fitting happening.



Given this model, we can say that as time goes on, our capabilities of producing more take off force for rockets has improved.

We can use our model to make predictions about the force and speed we will be able to reach by the dates when each resource runs out.

Year	2050 (CO_2 - 500 ppm)	2143 (CO_2 - 100 ppm)	2145 (gas)	2173 (oil)	2100 (coal)
Force (kN)	31494.41	231234.69	237636.09	336485.28	115105.26
Speed (m/s)	1405.16	1437.47	1438.50	1454.213	1418.78

*Data changes on run time of program due to randomness of machine learning models

4.3 Travel Time

The table displays the number of months given the departure year to travel to nearby planets based on their classification of Stars. We see despite the advancements in rocket speeds, the number of months to travel to habitable planets is not feasible.

Planet Name	Spectral Type	Distance [light years]	Year 2050 [Months]	Year 2100 [Months]	Year 2143 [Months]	Year 2145 [Months]	Year 2173 [Months]
Proxima Cen b	M5.5 V	4.243909	1.087307e+07	1.050632e+07	1.062867e+07	1.062106e+07	1.076869e+07
HD 20794 b	G8 V	19.578427	5.016071e+07	4.846881e+07	4.903325e+07	4.899814e+07	4.967918e+07
HD 20794 c	G8 V	19.578427	5.016071e+07	4.846881e+07	4.903325e+07	4.899814e+07	4.967918e+07
HD 20794 d	G8 V	19.578427	5.016071e+07	4.846881e+07	4.903325e+07	4.899814e+07	4.967918e+07
HD 20794 e	G8	19.578427	5.016071e+07	4.846881e+07	4.903325e+07	4.899814e+07	4.967918e+07
HD 26965 b	K0 V	16.423422	4.207746e+07	4.065821e+07	4.113169e+07	4.110224e+07	4.167353e+07
eps Ind A b	K2 V	11.867414	3.040479e+07	2.937925e+07	2.972138e+07	2.970010e+07	3.011291e+07

5 Conclusion

Using the data collected we determined that the critical dates for planet Earth in terms of hospitality and resource availability indicates that we need to consider extraterrestrial alternatives. The dates we used for prediction are those related to the depletion of each resource and air pollution due to excess CO_2.

These dates combined with the model used to predict technological growth in speed of rockets allowed us to determine the estimated travel times to nearby planets. However, we note that the needs of the planet for Uranium is not sufficient to waste on rockets unless we globally consume the uranium within nuclear warheads [30]. Even if we were to factor in nuclear propulsion, 50% of the travel time would not make a difference. The unfortunate outcome is that we will need to invest in hibernation methods as the travel time far exceeds the average human lifetime in months.

6 Limitations

There are a lot of factors that could dictate technological growth. We saw huge amounts of growth in the space sector during the Cold War era. Technology as a whole such as modern medicine was adapted from the World Wars[3]. These additional factors evident in world history can control the development of technology. The pandemic had also created huge amounts of delays for resources required for chip manufacturers such as Intel, AMD, and NVidia to even provide newer technology for research and development. For information such as water, we cannot account for changing weather patterns such as long duration droughts.

Data availability was another limitation. Data availability varies by countries where the USA and Canada are readily available compared to other countries in Europe or Asia. Furthermore, there were many limitations to the amount of data related to rockets available on the internet for the public. Even media packages for various rockets used vague language with respect to speed. In the future, given more time, we could collect more detailed data related to each launch, i.e. payload, and various other factors that could impact a rocket's velocity such as computing air resistance.. This way, we could make more accurate predictions about the rocket's velocity.

As for resources, we cannot predict whether there would be new reserves discovered, a new way found to artificially produce these resources, or perhaps new resources discovered/created that can replace the current ones.

7 Project Experience Summary

Angelika Nowlan

- Cleaned, transformed, and prepared data using Microsoft Excel and Pyhton's Pandas library to produce data that is ready to be used for analyzing.
- Applied machine learning techniques through Python's SciKit learn library to produce models for data and make relevant predictions
- Collaborated closely with team members using GitHub and Google Docs to integrate results and produce a technical report that communicates our problem solving process.

Yvonne (Yi Nuo) Wang

- Cleaned and prepared data using Python's Numpy and Pandas
- Analyzed data using machine learning through Python's SciKit learn library
- Created models to make relevant predictions and plots/graphs
- Collaborated in a team of 3 using Github and Google Docs to integrate our results/learning and produce a technical report

Shane Eastwood

- Organized related data from various sources using Pandas dataframes and related functions to group
- Created descriptive visuals with Matplotlib library
- Determined statistical significance by analyzing with statistics using Scipy stats
- Predicted future data points based on machine learning models
- Processed data into presentable results using Python

8 Data Sources and References

[1] "Ariane 1" *Wikipedia*, Wikimedia Foundation, 18 June 2022, https://en.wikipedia.org/wiki/Ariane_1

[2] "Ariane 4" *Wikipedia*, Wikimedia Foundation, 22 April 2022, https://en.wikipedia.org/wiki/Ariane_4

[3] "Ariane 5" *Wikipedia*, Wikimedia Foundation, 2 Dec. 2022, https://en.wikipedia.org/wiki/Ariane_5

[4] "Delta II Data Sheet", *Office of Safety and Mission Assurance,* NASA, 20 Oct. 2018, https://sma.nasa.gov/LaunchVehicle/assets/delta-ii-data-sheet.pdf

[5] "Delta III" *Wikipedia*, Wikimedia Foundation, 14 Sept. 2022, https://en.wikipedia.org/wiki/Delta_III

[6] "Delta IV" *Wikipedia*, Wikimedia Foundation, 24 Nov. 2022, https://en.wikipedia.org/wiki/Delta_IV

[7] "Delta IV Heavy" *Wikipedia*, Wikimedia Foundation, 9 Dec. 2022, https://en.wikipedia.org/wiki/Delta_IV_Heavy

[8] "Falcon 1" *Wikipedia*, Wikimedia Foundation, 25 Nov. 2022, https://en.wikipedia.org/wiki/Falcon_1

[9] " 'Goldilocks' stars may be 'just right' for finding habitable worlds." *NASA*, https://www.nasa.gov/feature/goddard/2019/k-star-advantage. [Accessed: 20-Oct-2022].

[10] "Falcon 9" *Wikipedia*, Wikimedia Foundation, 15 July. 2022, https://en.wikipedia.org/wiki/Falcon_9_v1.0

[11] "Falcon 9 v1.0" *Wikipedia*, Wikimedia Foundation, 9 Dec. 2022, https://en.wikipedia.org/wiki/Falcon_9

[12] "Falcon 9 Full Thrust" *Wikipedia*, Wikimedia Foundation, 17 Nov. 2022, https://en.wikipedia.org/wiki/Falcon_9_Full_Thrust

[13] "Falcon Heavy" *Wikipedia*, Wikimedia Foundation, 9 Dec. 2022, https://en.wikipedia.org/wiki/Falcon_Heavy

[14] "NASA exoplanet archive," *NASA Exoplanet Archive.* [Online]. Available: https://exoplanetarchive.ipac.caltech.edu/. [Accessed: 15-Oct-2022].

[15] "The scientific and technological advances of world war II." *The National WWII Museum*. Available: https://www.nationalww2museum.org/war/articles/scientific-and-technological-advances-world-war-ii. [Accessed: 25-Nov-2022].

[16] "Titan 23G" *Wikipedia*, Wikimedia Foundation, 7 Sept. 2022, https://en.wikipedia.org/wiki/Titan_23G

[17] “Titan IIIA” *Wikipedia*, Wikimedia Foundation, 22 April. 2022, https://en.wikipedia.org/wiki/Titan_IIIA

[18] “Titan IIIB” *Wikipedia*, Wikimedia Foundation, 7 Sept. 2022, https://en.wikipedia.org/wiki/Titan_IIIB

[19] “Titan IIIC” *Wikipedia*, Wikimedia Foundation, 23 Nov. 2022, https://en.wikipedia.org/wiki/Titan_IIIC

[20] *Rosetabares*. “Space Flight Mission Status.” Kaggle. https://www.kaggle.com/datasets/rosetabares/spacemissionsflightstatus

[21] “VLS-1.” *Wikipedia*, Wikimedia Foundation, 12 Nov 2022, https://en.wikipedia.org/wiki/VLS-1

[22] “History of war,” Cold War Technology - History of War in the 20th Century. [Online]. Available: https://publish.uwo.ca/~acopp2/historyofwar/coldwar/technology.html.

[23]D. Reynolds, “Science, technology, and the Cold War,” Cambridge University Press, 2010. https://www.cambridge.org/core/books/abs/cambridge-history-of-the-cold-war/science-technology-and-the-cold-war/A22272BBB260D2E8E9BC6B98A7708C5A

[24] H. Ritchie and M. Roser, “Water Use and Stress,” Our World in Data, Nov. 2017, Accessed: Dec. 09, 2022. [Online]. Available: https://ourworldindata.org/water-use-stress#renewable-freshwater-resources

[25] Fao.org, 2017. https://www.fao.org/aquastat/en/databases/maindatabase/

[26] “Open Data Platform,” data.footprintnetwork.org. https://data.footprintnetwork.org/#/countryTrends?cn=5001&type=BCtot

[27] “World Uranium Mining - World Nuclear Association,” world-nuclear.org. https://world-nuclear.org/information-library/nuclear-fuel-cycle/mining-of-uranium/world-uranium-mining-production.aspx

[28] “Electricity production from nuclear sources (% of total) | Data,” data.worldbank.org. https://data.worldbank.org/indicator/EG.ELC.NUCL.ZS

[29] “World Nuclear Power Reactors | Uranium Requirements | Future Nuclear Power - World Nuclear Association,” world-nuclear.org. https://world-nuclear.org/information-library/facts-and-figures/world-nuclear-power-reactors-and-uranium-requireme.aspx

[30] “Military Warheads as a Source of Nuclear Fuel | Megatons to MegaWatts - World Nuclear Association,” *world-nuclear.org*. https://world-nuclear.org/information-library/nuclear-fuel-cycle/uranium-resources/military-warheads-as-a-source-of-nuclear-fuel.aspx

[31] “Earth To Mars In 100 Days? The Power Of Nuclear Rockets,” Universal-Sci. https://www.universal-sci.com/headlines/2019/7/3/earth-to-mars-in-100-days-the-power-of-nuclear-rockets

[32] U.S. Energy Information Administration, “Nuclear power and the environment - U.S. Energy Information Administration (EIA),” www.eia.gov, Jan. 15, 2020.
https://www.eia.gov/energyexplained/nuclear/nuclear-power-and-the-environment.php#:~:text=Nuclear%20power%20reactors%20do%20not

[33] “Freshwater Resources | National Geographic Society,” education.nationalgeographic.org.
https://education.nationalgeographic.org/resource/freshwater-resources

[34] Rowan Jacobsen,Ensia, “Israel Proves the Desalination Era Is Here,” Scientific American, Jul. 29, 2016.
https://www.scientificamerican.com/article/israel-proves-the-desalination-era-is-here/

[35] “Nuclear & Uranium Data - U.S. Energy Information Administration (EIA),” www.eia.gov.
https://www.eia.gov/nuclear/data.php

[36] “USDA ERS - Natural Resources Datasets,” www.ers.usda.gov.
https://www.ers.usda.gov/about-ers/partnerships/strengthening-statistics-through-the-icars/natural-resources-datasets/

[37] Our World in Data, “Our World in Data,” Our World in Data, 2019. https://ourworldindata.org

[38] “SRIS - Spent Fuel and Radioactive Waste Information System,” sris.iaea.org.
https://sris.iaea.org/home

[39] “Home | DataHub - Frictionless Data,” datahub.io. https://datahub.io

[40] “data.world | The Cloud-Native Data Catalog,” *data.world*. https://data.world

[41] D. Smith, “Is Carbon Dioxide Harmful to People?,” *learn.kaiterra.com*, Jul. 17, 2019.
https://learn.kaiterra.com/en/air-academy/is-carbon-dioxide-harmful-to-people

[42] N. Jones, “How the World Passed a Carbon Threshold and Why It Matters,” *Yale E360*, Jan. 26, 2017.
https://e360.yale.edu/features/how-the-world-passed-a-carbon-threshold-400ppm-and-why-it-matters