Analysis on All Space Missions from 1957

Mustafa Batuhan Ceylan
Computer Engineering
Galatasaray University
Istanbul, Turkey
ceylan.mustafabatuhan@gmail.com

I. Introduction

This document analyses the all space missions from 1957 to make predictions about the future of spaceflight and make inferences about its sustainability and interest using spending, reliability and other indicators.

II. Data and Analysis

In this section, we will try to find answers to the questions we have come up with in our previous report in which we did an exploratory data analysis on the subject dataset.

A. Did space travel got cheaper since 1957?

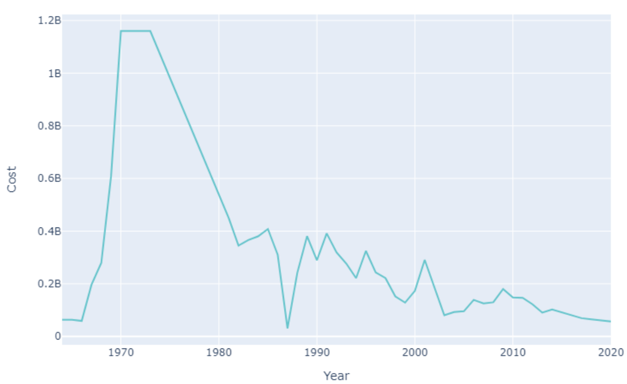

Fig. 1. Yearly spending

At the first glance, this graph seems like a good indicator for price of space travel. However, this is not a great indicator due to the fact that government and public interest in space travel was declining after the cold war.

The mission cost over time calculated per launch is a far better indicator for this and gives an easily deducible answer to our question about the affordability of space travel. Despite the fact that, mission costs are seldom specified in the dataset due to various political and bureaucratic reasons, e.g., cost deduction from yearly agency budget or top secret programs.

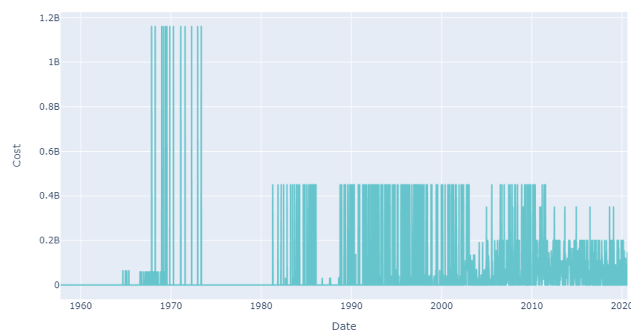

Fig. 2. Launch Cost per Launch

While the cost data has a lot of unfilled values, the trend is visible that launch costs are going down in general. Billions of dollars per launch costs are now replaced with comparatively cheaper 500M to 10M dollars per launch.

B. Did space travel become more reliable since 1957?

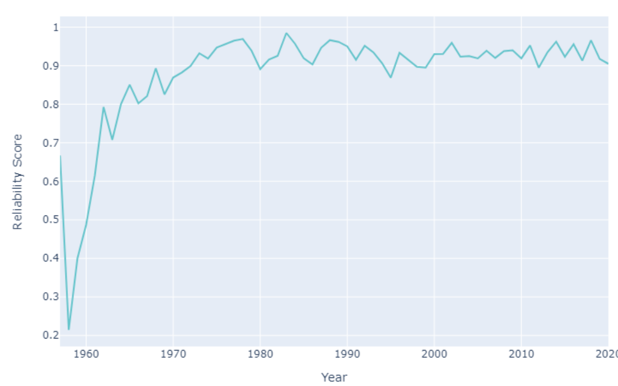

Fig. 3. Reliability Over Time

The graph shows a steep increase in rocket reliability between 1960s and 1970s. Then the increase stagnates over time with negligible local maximums and minimums throughout the years.

In the graph, the value 'Reliability Score', is calculated per year using the following formula:

$$ReliabilityScore(\tau) = \frac{SuccessfulMissions(\tau)}{AllMissions(\tau)} \quad (1)$$

In the equation, the variable τ be defined as a time interval or any other categorical feature, e.g., Company Name or Vessel Model. The reliability score will be calculated with the same formula throughout this paper. For a good measure, we grouped 'Partial Failure' along with 'Prelaunch Failure' and 'Failure' as complete failures since the dataset does not clarify what it considers a 'Partial Failure'.

C. What is the most reliable space company?

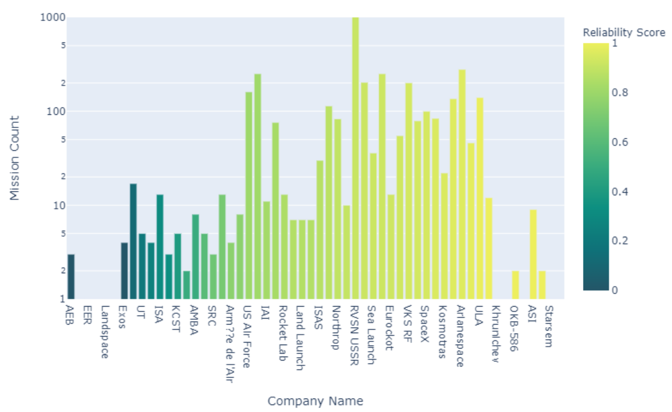

Fig. 4. Reliability Per Company Along with Mission Counts

The graph above is sorted according to the reliability scores of each company (increasing from left to right). The bars represent mission counts for each company. This reveals a major problem with the scores. If a company has done 1 mission and it was not a fail, the group will have an unbelievable reliability score of 1.

If we remove any company that has less than 10 launches from the dataset for the comparisons sake, we can get a much clear picture.

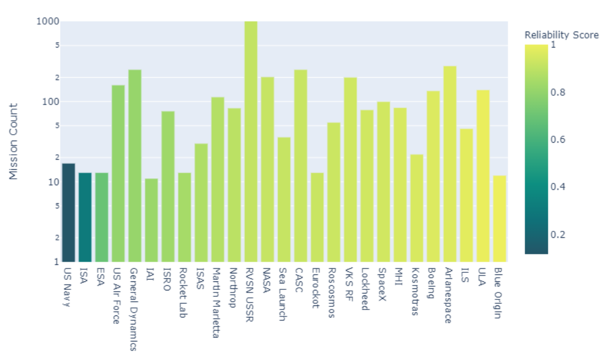

Fig. 5. Reliability Per Company Along with Mission Counts (Mission Count > 10)

As we can see Blue Origin, an up and coming space company has a perfect track record with their low launch count of 12. A more "field tested" company ULA is trailing Blue Origin with their near perfect reliability score of 0.99 with 140 missions on to their name.

D. What is the most reliable vessel?

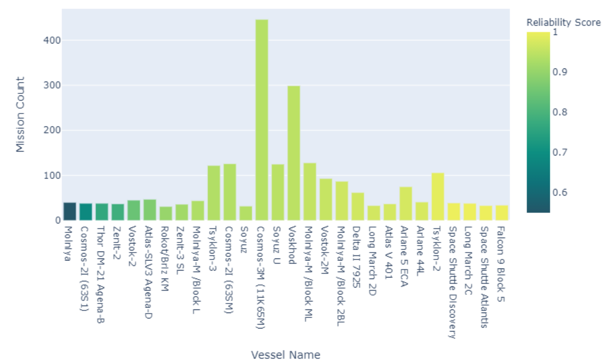

Fig. 6. Reliability Per Vessel Along with Mission Counts (Mission Count > 30)

The most reliable vessel in the graph above is the SpaceX's Falcon 9 Block 5 with a perfect reliability score on its 34 missions. Considering the fact that, the Falcon 9 has flown refurbished considerable amount of times, it is already proving itself as a highly reliable vessel. Not just in the short term and numerical sense but also in a more realistic scenario where vessel re-usability is necessary to cut costs.

E. Are the average launches per month/year increasing?

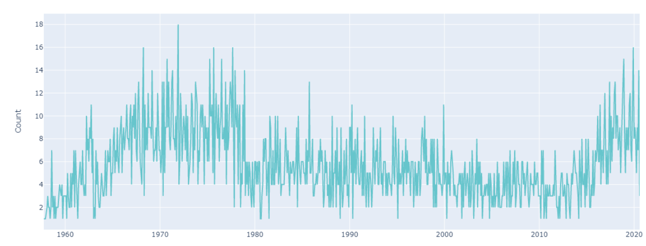

Fig. 7. Launches per Month (Raw)

From the figure, we can see the trend moving towards the first space race era numbers, despite being quite noisy.

For a cleaner figure on the launches we can do an additive decomposition on this time series data then combine its trend and seasonal parts, meaning we remove the residual data points.

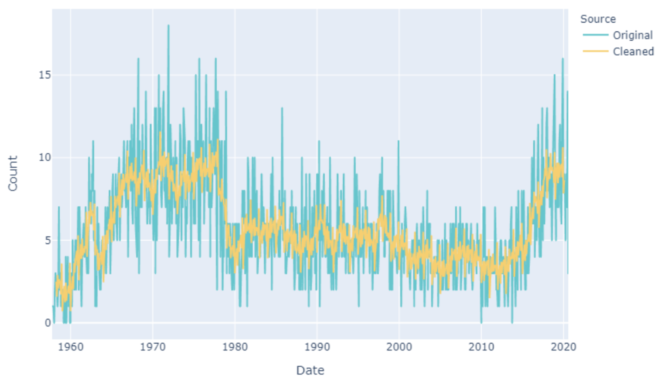

Fig. 8. Launches per Month (Cleaned)

This in turn gives us a much more solid understanding of this time series data. While the sudden increase of launches in the past decade seems to have stagnated, the real world promises for future Moon and Mars missions from multiple space agencies and private space companies point towards even more launches in the foreseeable future.

F. What state is the private space agencies at?

This question is fairly vague on purpose since during the EDA phase, we were not sure how this feature ('Private or State') would turn out. However, we will now analyse the launch counts for privately funded missions over time get a perspective over this feature.

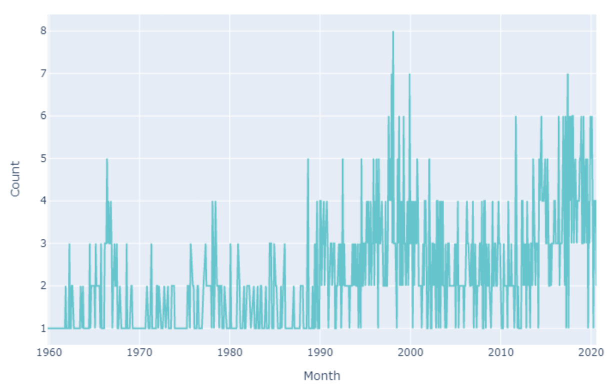

Fig. 9. Privately Funded Launches per Month (Raw)

Above graph shows the launches over time that are funded by private ventures. Which depicts a clear increase in the past 2 decades with peaks around year 2000 and a new trend developing in the last decade. But the small number of launches per month gives us a fairly noisy graph which we can clean up using time series decomposition like before. We can combine the trend and seasonal parts of this data as we did for all launches.

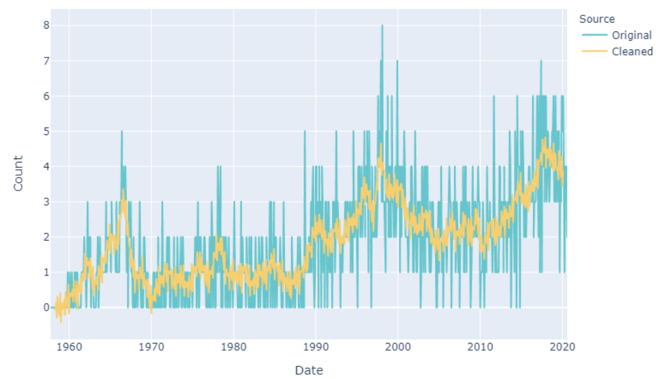

Fig. 10. Privately Funded Launches per Month (Cleaned)

After time series decomposition we can clear the noise from the series to get a better view of the private interest in space. There are 3 major local peaks around 1967, 2000 and 2018. The first is the obvious cold war era space race period, which also saw its fair share of private interest despite the huge government interest. The recent two are the peaks from space shuttle missions and more recently SpaceX with their re-usability campaign.

G. Are government funded missions decreasing in amount?

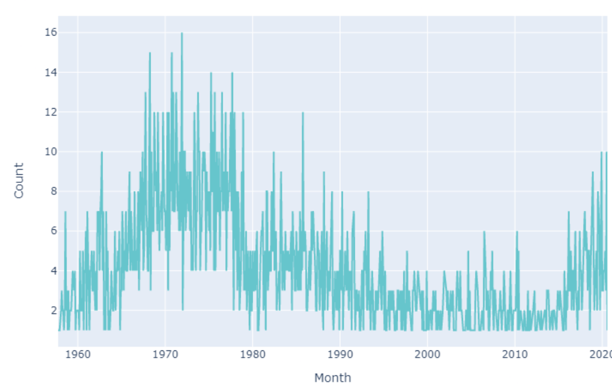

Fig. 11. State Funded Launches per Month (Raw)

While the recent interest in space by government agencies is increasing, the is a very important factor in play that needs to be considered before making a claim.

First, the recent American administrational changes on how space missions are handled. Now, NASA does not produce their own rockets, nor directly fund the missions, however they merely are customers to private ventures like ULA, SpaceX, Boeing etc. which means data is shown as state funded even though state is just a customer for private ventures.

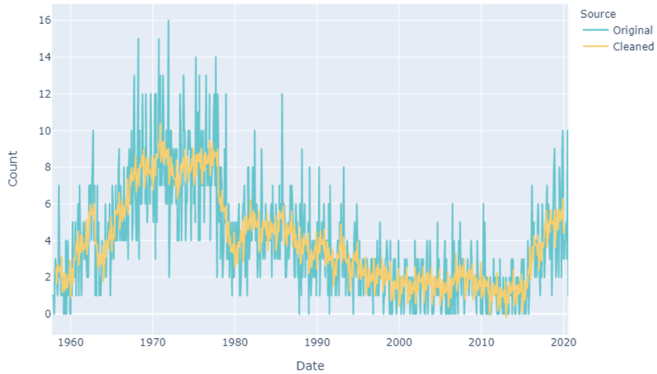

Fig. 12. State Funded Launches per Month (Cleaned)

H. Are missions with private funding more reliable than missions with state funding?

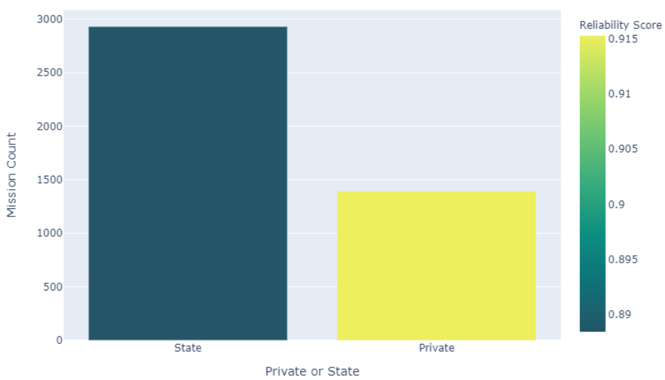

Fig. 13. Reliability Per Funding Along with Mission Counts

The graph above depicts a clear picture with the comparatively higher reliability score of privately funded missions over state funded missions. With privately funded missions having a 0.91 reliability score and state funded missions having 0.88 reliability score in total.

III. Further Analysis and Prediction

A. Predicting Future Launches

To predict a monthly launch number we could easily fit an ARIMA (Autoregressive Integrated Moving Average) model which automatically does regression over the given time series data to predict future trends.

1) Predicting Future State Funded Missions:

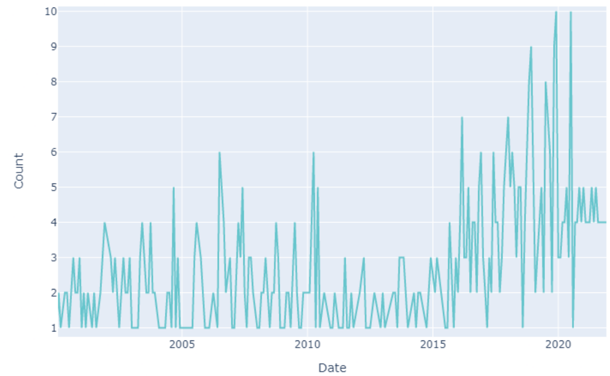

Fig. 14. State Funded Launches per Month (Prediction)

The above graph has the past 10 years, plus the year 2021 as the prediction outcome from our ARIMA model. The model can predict a few first months of 2021 then has a constant rate of 4 launches per month.

Upon further analysis the trend of the first few months of 2021 show a definite increase, and with real world promises from NASA for future Mars and Moon missions we can see an increase or a possible stagnation with the new American administration which seems to have less interest in further space missions.

2) Predicting Future Privately Funded Missions:

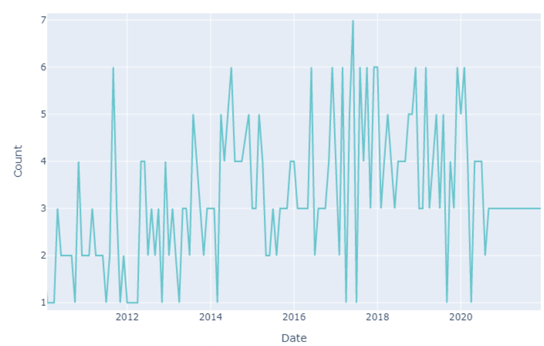

Fig. 15. Privately Funded Launches per Month (Prediction)

The prediction of private ventures have more volatility compared to state funded missions due to the fact that, there are comparatively less privately funded missions. But the first few months of late 2020 show us stagnation in the number of launches. then it halts to a constant value of 3.

B. Cost Comparison Between Private and State Funded Missions

To find if privately funded missions are cheaper than state funded missions we first sampled 400 random missions from a cleaned dataset where all unspecified costs were removed from both privately and state funded missions. Then we looked at the descriptions of both samples to get a basic understanding on both of them:

Funding	Mean	STD	Min	Max
Private	104 338 500.00	71 445 646.95	7 000 000.00	350 000 000.00
State	187 551 075.00	246 062 826.05	5 300 000.00	1 160 000 000.00

From the table above we can find that privately funded missions seem to be cheaper. However to claim this we first need to make an ANOVA model between these two samples:

	sum_q	df	F	PR(F)
C(Funding)	1.384867e+18	1.0	42.188488	1.457169e-10
Residual	2.619491e+19	798.0	NaN	NaN

From the analysis we can deduce that the ANOVA model has significance ($p < 0.05$), meaning privately funded space missions are cheaper than state funded missions.

IV. Conclusion

The conclusion we can draw from this analysis is that, we cannot be sure about our future predictions in a long term with the available context provided from this dataset. We might need more real world data, like orbital positions of each astronomical bodies, current developments of companies and so on.

When we come back to our three original hypothesis from the first EDA report we can make this three statemens:

A. Space mission costs are trending downwards throughout time:

For this hypothesis we can easily reject the null hypothesis because of Fig.2. And conclude that our hypothesis is true.

B. Missions are getting more and more reliable. Meaning amount of failed missions are down all across the vessels:

For this hypothesis we can reject the hypothesis, however the improvement is not astronomical at this current state of rockets. More future data is required to output a more confident conclusion on this hypothesis.

C. Current increase in the amount of missions will continue in the future:

As we have tried to predict with our ARIMA model, the number of missions are predicted be increasing for following few months after the dataset, after that our model predicts a constant rate which gives us the idea that the predictability of monthly launch counts are pretty short-term, no matter the origin of funding. Since this was the same behaviour for 3 ARIMA models. We also tested SARIMA model but since it didn't portray a different outcome we decided to go with simpler ARIMA model in this report.

Luckily, we can cross reference the predictions for about past 4 months. Since this dataset lacks the past 4 months of 2020 at the time of creating this report. When we do that, we can see that our models are under-performing compared to real world data, but not by much.

In 2020 there were in total 104 orbital flights however our model predicted 97. But when we only consider last 4 months our model has predicted %25 less missions than real world data. While our model was incorrect, our assumption was correct that launches would increase. Since about %40 of all flights in 2020 were done in the past 4 months. This in turn means that number of missions are increasing, even more so than our predicted values.