Final Project: SpaceX Payload Mass

INTRODUCTION

SpaceX is a private American aerospace company that designs and manufactures rockets and spacecraft. They are well-known for designing, building, and recovering the first fully-reusable first stage of an orbital rocket, known as the *Falcon 9*.

Using launch and payload mass data, I will investigate whether SpaceX increased its average payload mass of the Falcon 9 (Block 5) in 2020 relative to 2019. More formally, my hypotheses are as follows:

$$H_0: \bar{x}_{2020} - \bar{x}_{2019} = 0$$

$$H_a: \bar{x}_{2020} - \bar{x}_{2019} > 0$$

where $\bar{x}_{2020}$ and $\bar{x}_{2019}$ are the average payload mass for 2020 and 2019, respectively.

It is important to note that SpaceX offers ride-sharing services. Therefore, I will only be analyzing the total payload mass for a single launch, rather than individual payloads and their respective masses.

METHODS

The analysis was conducted in Python 3.13 using the following packages: Pandas, NumPy, Matplotlib, SciPy, and Math. Launch data for 2019–2020 were treated as random samples that represent larger population of SpaceX missions. The code and relevant documents that were used to perform this analysis can be found here.

The difference in means and standard error of the difference in means are calculated as $\bar{x} = \bar{x}_{2020} - \bar{x}_{2019}$ and $SE = \sqrt{\frac{s_{2020}^2}{n_{2020}} + \frac{s_{2019}^2}{n_{2019}}}$, respectively. A 95% two-sided confidence interval was constructed for the difference in means, and a one-sided hypothesis test with $\alpha = 0.05$was performed to assess whether the 2020 mean exceeds the 2019 mean.

RESULTS

Summary statistics for each year are shown in Table 1.

Table 1: Summary Statistics for SpaceX Payload Mass in 2019 and 2020 (in kg)

	2019	2020
Mean	8,030.08	11,402.39
Median	6,728	15,410
Standard Deviation	5098.45	5608.13

To gain more insight into the distribution of payload mass in 2019 and 2020, it is helpful to view a box and whisker plot, as seen in Figure 1.

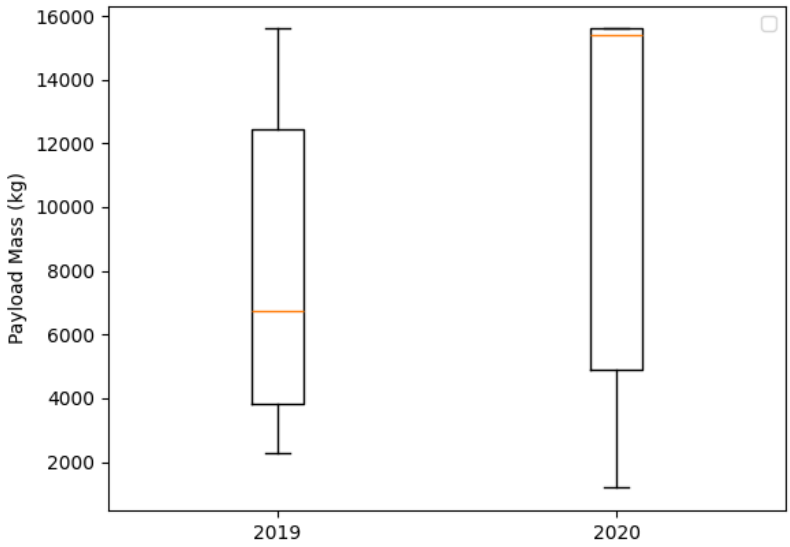

Figure 1: Boxplots for SpaceX Payload Mass in 2019 and 2020

The difference in means and the standard error of the difference in means are $\bar{x} = 3{,}372.31$ and $SE = 1{,}879.79$, respectively. The resulting 95% confidence interval for the difference in means is (-452.161, 7196.777) and the one-sided hypothesis test yields $p = 0.04099$.

DISCUSSION

In Table 1, it's notable that the mean payload mass and standard deviation increased in 2020. It is also notable in Figure 1 the boxplot for 2020 shows how closely quartiles 3 and 4 are packed, whereas quartiles 1 and 2 are spaced far apart. Each quartile in the boxplot for 2019 appears to be similar in spread.

Table 1 and Figure 1 are useful to gain a quick and rough idea of trends, however they do not provide enough evidence to be statistically significant. We now turn to the confidence interval and hypothesis test. The confidence interval for $\bar{x}_{2020} - \bar{x}_{2019}$ spans zero, meaning the observed difference could plausibly be due to sampling variability alone. However, the one-sided hypothesis test yields $p = 0.04099 < 0.05$, allowing us to reject the null hypothesis and conclude that the mean payload mass in 2020 was likely higher.

This conclusion depends on the previous assumption that the sample launch is representative of all Falcon 9 Block 5 missions. Non-random sampling or missing launches could bias results. These risks can be reduced by ensuring a random data selection and increasing the number of observations.