**Ticket to Mars: Estimating the Cost and Timeline of Human Travel to Mars**

**Name:** Cory Golladay

**Course:** DSC 680: Applied Data Science

**Date:** July 2025

**Business Problem**

As interest in Mars colonization grows, driven by companies like SpaceX and agencies like NASA, projections about cost and launch timelines are often based more on vision than data. This project aims to estimate the realistic cost and potential launch windows for future manned Mars missions using a regression model and Monte Carlo simulation. These insights can help policymakers, researchers, and investors better understand the magnitude of resources required and the uncertainty involved.

**Background and History**

Missions to Mars have progressed from unmanned flybys and rovers to plans for human exploration in the coming decades. NASA’s Artemis program and SpaceX’s Starship platform have laid groundwork for future crewed missions. However, delays and cost overruns in projects like the ISS and SLS suggest that grounded, data-driven forecasting is essential for mission planning.

**Data Explanation**

A synthetic dataset of 250 hypothetical Mars missions was created using domain-informed ranges for:  
- Payload Mass (10–100 tons)  
- Development Time (4–15 years)  
- Launch Year (2026–2040)  
- Vehicle Type: Falcon Heavy, Starship, SLS  
- Estimated Mission Cost (dependent variable)  
  
Each variable was grounded in published estimates from NASA, SpaceX, and the Planetary Society.

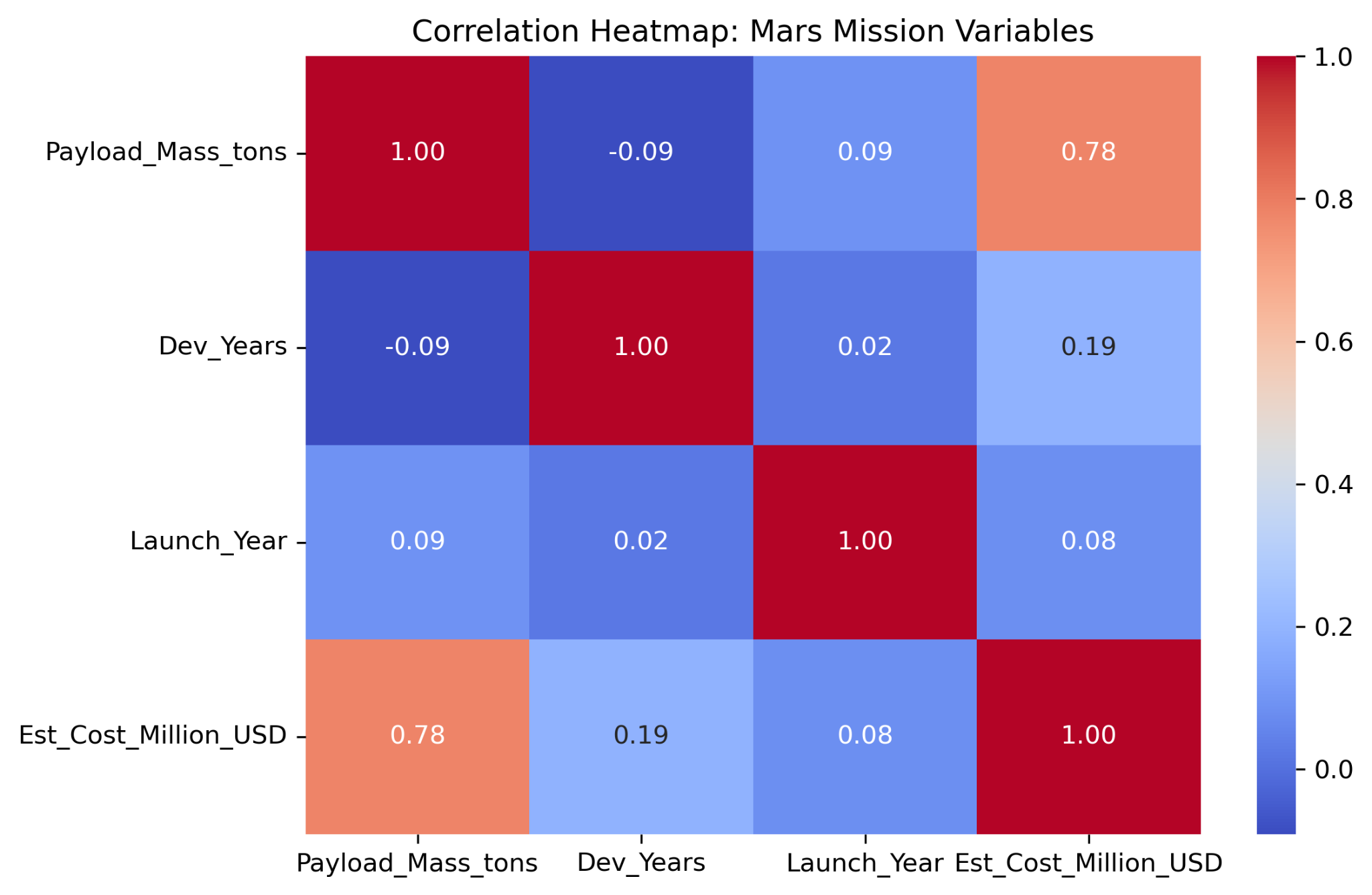


Figure 1. Correlation Heatmap: Cost was most strongly correlated with payload mass.

Methods

- EDA: Summary stats, heatmaps, boxplots, scatter plots  
- Multivariate Linear Regression: Predicted mission cost using payload mass, dev years, and vehicle type  
- Monte Carlo Simulation: Modeled 1,000 future missions to estimate a distribution of cost outcomes

Analysis and Results

The regression model showed strong performance with R² = 0.91 and RMSE ≈ $519M, indicating reliable predictive accuracy. Vehicle type had a substantial impact: SLS missions were most expensive, while Starship missions offered cost savings.

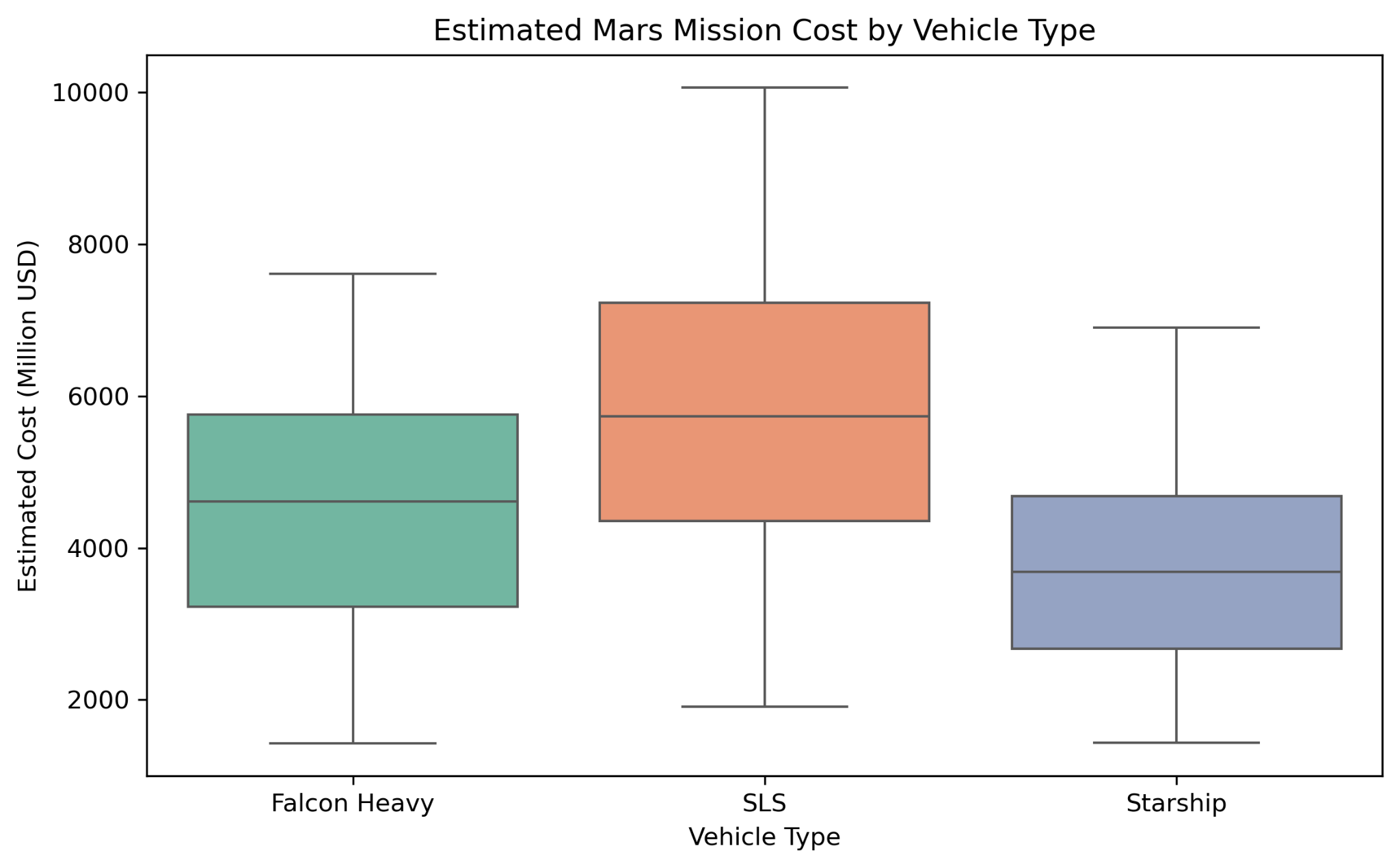


Figure 2. Estimated Mars Mission Cost by Vehicle Type.

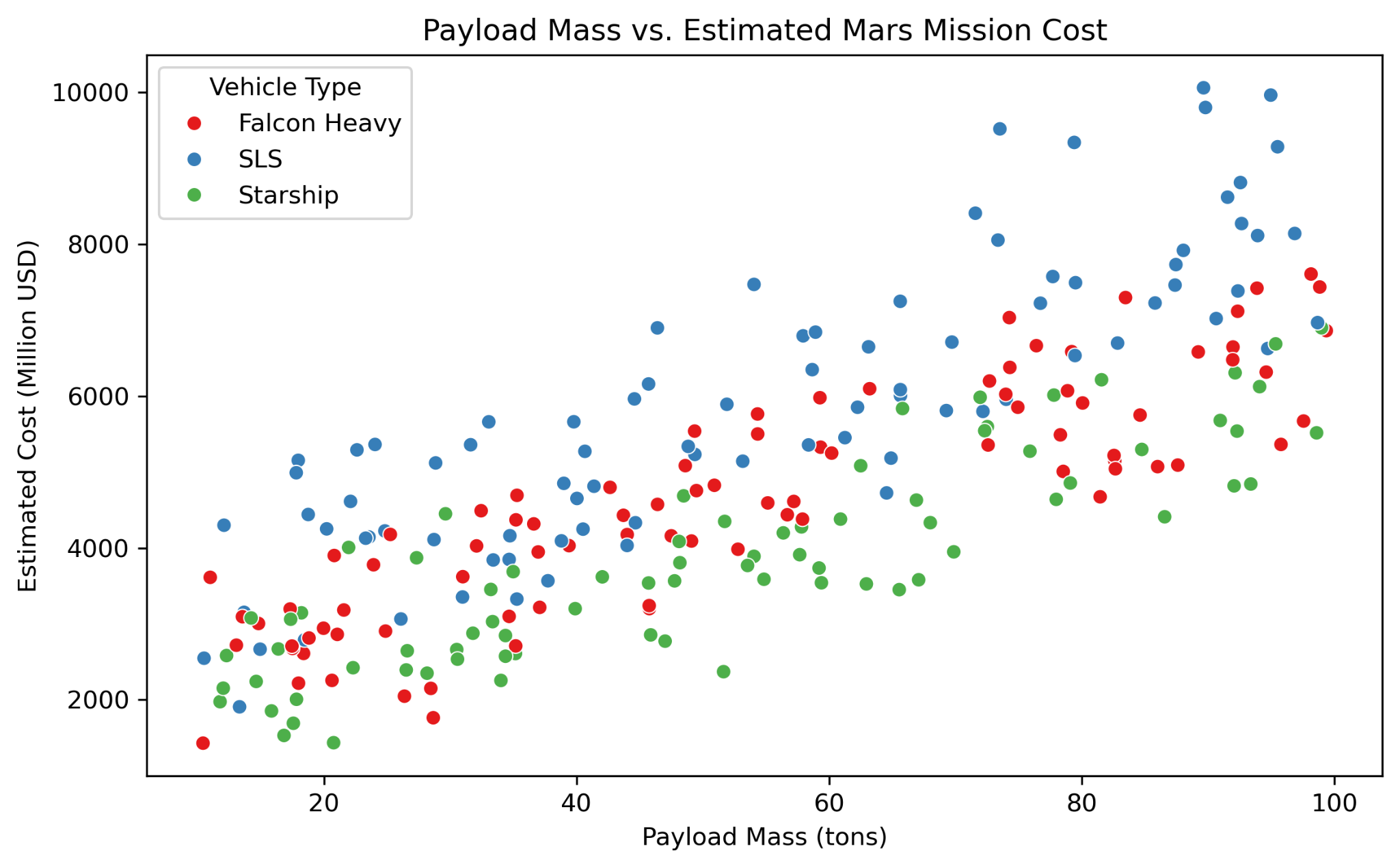


Figure 3. Payload Mass vs. Estimated Mars Mission Cost.

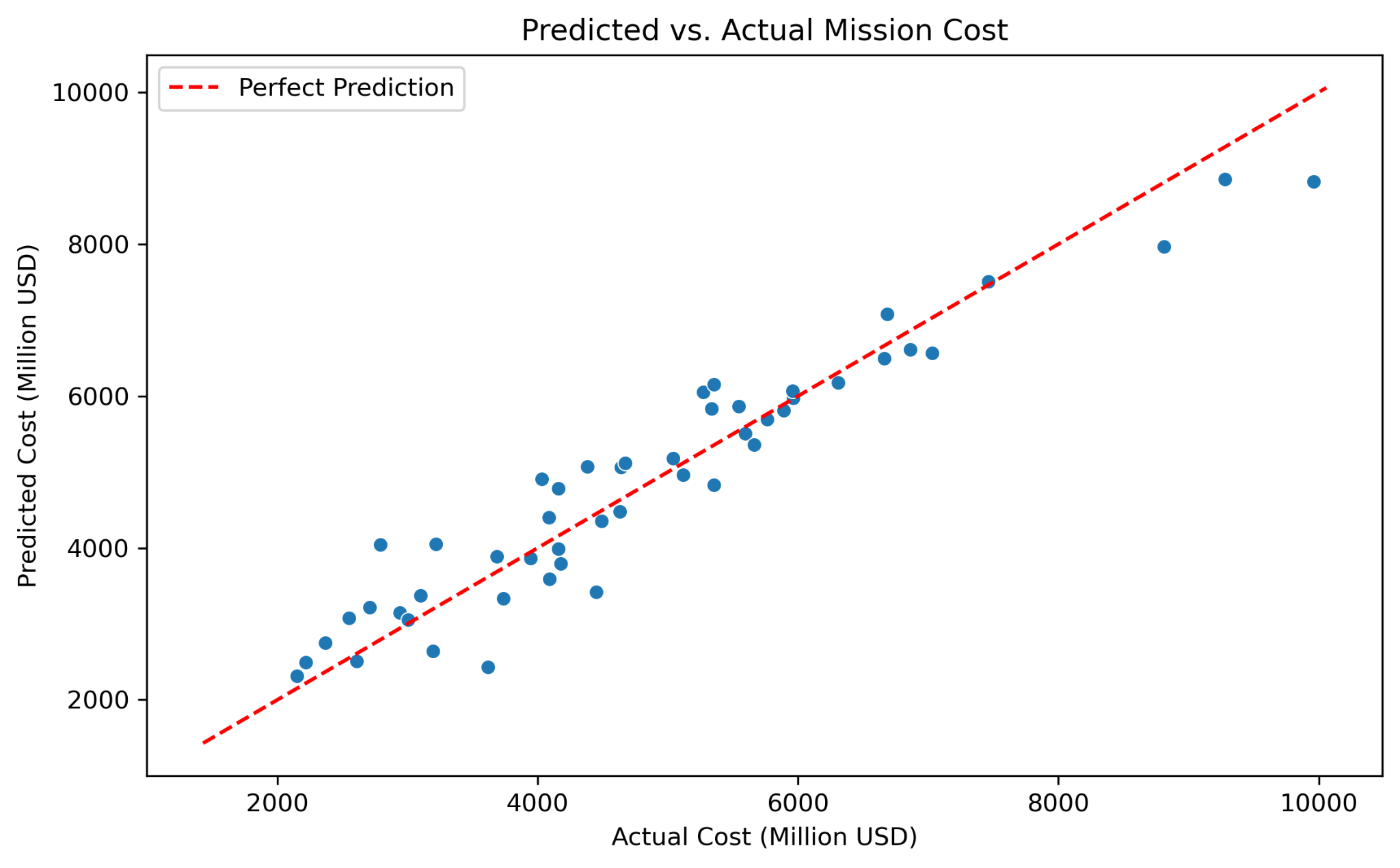


Figure 4. Predicted vs. Actual Mission Cost.

The Monte Carlo simulation revealed that most Mars missions are expected to cost between $3.6B and $6.2B, with outliers ranging from $850M to $9.4B.

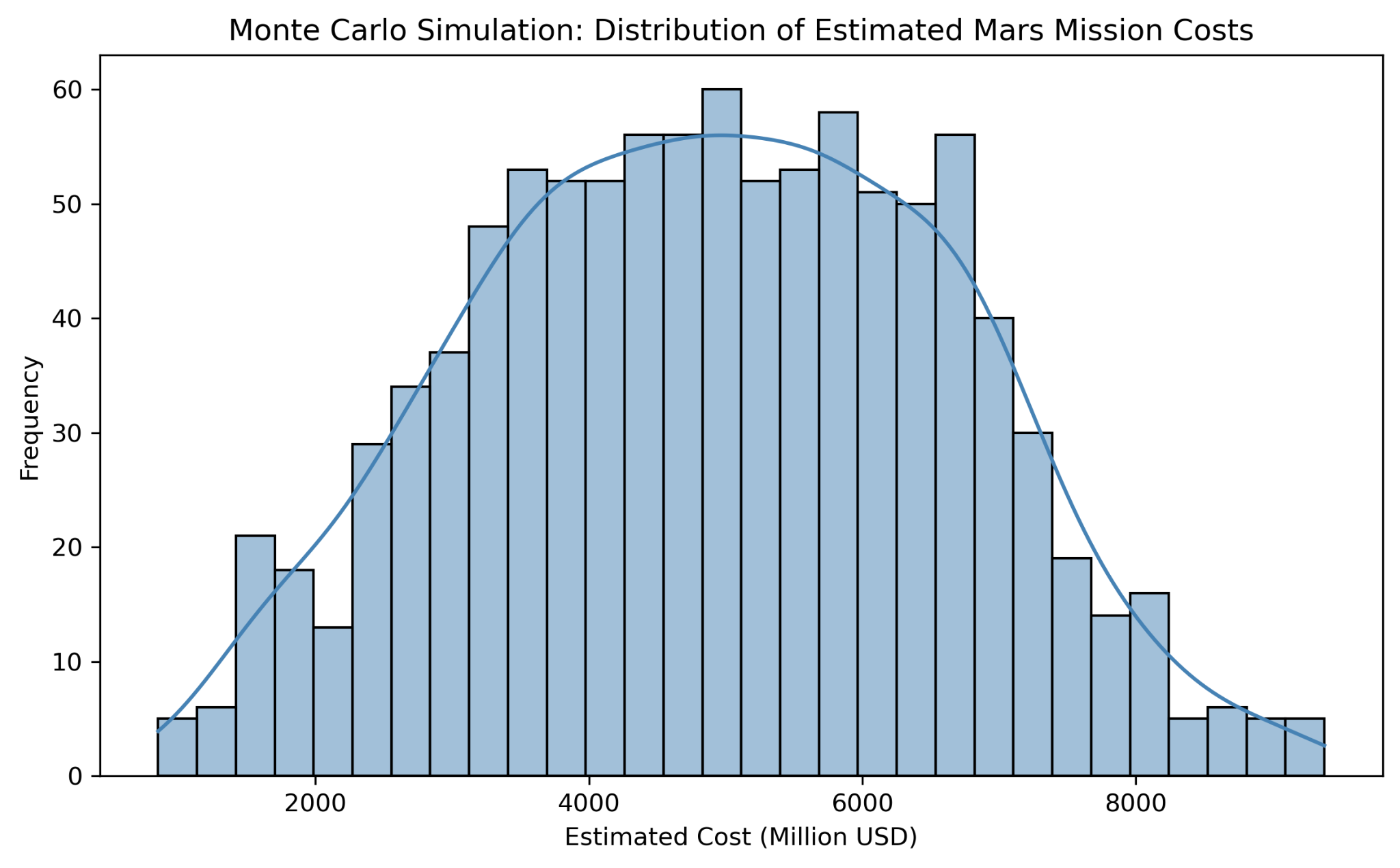


Figure 5. Monte Carlo Simulation: Distribution of Estimated Mars Mission Costs.

**Conclusion**

This analysis provides a realistic, data-driven forecast of Mars mission costs. While current ambitions are bold, this modeling highlights the importance of accounting for uncertainty. Starship appears the most cost-effective option among the vehicles considered.

**Assumptions**

Cost formulas reflect historical launch and development dynamics. Simulated variables are based on published engineering and budget estimates. Monte Carlo simulation assumes normally distributed variation in mission factors.

**Limitations**

The dataset is synthetic and informed by secondary sources, not real NASA logs. Private sector cost data (e.g., SpaceX) is often speculative or undisclosed. Technological breakthroughs could shift all variables dramatically.

**Future Uses / Applications**

Adaptable to lunar missions, satellite constellations, or asteroid mining. Can be expanded to include time-to-launch forecasting or failure probability. Useful for investors or governments modeling cost risk scenarios.

**Recommendations**

Continue improving transparency in space project costs. Use probabilistic modeling like this for policy planning. Fund comparative studies across propulsion technologies.

**Implementation Plan**

Publish this framework as a public policy memo or simulation tool. Share results with space-focused organizations or think tanks. Integrate new real-world mission data as it becomes available.

**Ethical Assessment**

Weighing Mars funding against Earth needs is essential. Potential planetary contamination is a serious environmental concern. Equity must be considered—who gets to go, and who benefits?

**References**

NASA Mars Exploration Program. (2025). https://mars.nasa.gov  
SpaceX Starship Updates. (2025). https://www.spacex.com  
Our World in Data. (2025). https://ourworldindata.org  
The Planetary Society. (2024). Mars mission cost data  
U.S. GAO Reports on NASA Budgeting

**Appendix**

**1. How were the variables in your simulated dataset chosen and validated?**  
The variables were selected based on publicly available data and trends from NASA, SpaceX, and the Planetary Society. I used realistic ranges for payload mass, vehicle type, development time, and mission cost to mirror the scale and complexity of real-world Mars planning.

**2. Why did you choose regression and simulation methods over others like time series forecasting?**  
Regression was ideal for identifying how specific factors—like payload and vehicle type—impact cost. Monte Carlo simulation allowed me to model uncertainty and explore a range of possible outcomes. Time series forecasting wasn’t suitable because no recurring historical data exists for manned Mars missions.

**3. How will your results change once you incorporate real-world data?**  
Real data would likely refine model accuracy and highlight cost outliers tied to specific mission designs. The underlying patterns may hold, but cost variance and vehicle performance could shift the predicted range.

**4. What are the limitations of using publicly available data from sources like NASA and SpaceX?**  
The main limitation is that detailed cost breakdowns—especially from private companies—aren’t always transparent. NASA provides audit reports, but SpaceX data is often aspirational or incomplete, which limits full validation.

**5. How did you estimate development years in the absence of actual timelines?**  
I used historical data from missions like the ISS, Artemis, and Perseverance to set a realistic range of 4–15 years. These values reflect typical planning, engineering, and testing cycles for large-scale space missions.

**6. What assumptions were made in your Monte Carlo simulation model?**  
The simulation assumed that mission costs are normally distributed around a realistic mean, with variability introduced through changes in payload, dev time, and vehicle type. It also assumed consistent cost behavior across vehicle classes.

**7. How do you ensure your findings are accessible to non-technical stakeholders?**  
I used simple language, clear visuals, and explained statistical results without technical jargon. Visuals like the Monte Carlo histogram and regression scatter plots were chosen specifically to communicate uncertainty and trends clearly.

**8. What are the most significant cost drivers in your current analysis?**  
Payload mass, development duration, and vehicle type are the key cost drivers. SLS adds significant cost compared to Starship or Falcon Heavy, and longer development timelines also raise total mission budgets.

**9. How will this model be updated if new technologies or breakthroughs occur?**  
The model can be retrained with updated data, especially if new propulsion systems or mission architectures are introduced. Future iterations could include additional variables like reusability, in-orbit refueling, or international collaboration.

**10. What role do ethics play in deciding whether Mars colonization is a worthy investment?**  
Ethics play a major role—there are trade-offs between investing in space and addressing needs on Earth. There are also environmental concerns, such as planetary protection, and equity questions about who benefits from space exploration.