Insights from the 50 Startup Assignment

1. **Correlation Heatmap:** This chart shows the relationships between numerical features.

There is a strong positive correlation between "R&D Spend" and "Profit," it suggests that higher R&D spending is associated with higher profits.

1. **Pair plot:** This chart shows scatter plots and distributions of numerical features.

There is a upward trend in the scatter plot between "Marketing Spend" and "Profit," it suggests that higher marketing spending is associated with higher profits.

1. **Box Plots:** These charts show the distribution of a numerical feature for different categories of a categorical feature.

Median "Profit" is higher for companies in "New York" compared to other states, it suggests that companies in New York tend to be more profitable.

1. **Distribution Plots:** These charts show the distribution of a numerical feature.

The distribution of "Profit" is skewed to the right, it means there are a few companies with very high profits.

1. **Bar Plots:** These charts show the average value of a numerical feature for different categories of a categorical feature.

The Average "Profit" is higher for companies in "California" compared to other states, it suggests that companies in California tend to be more profitable.

1. Based on correlation and regression analysis**, R&D Spend is a significant predictor of Profit.** And **Marketing Spend may also have an impact on Profit.**
2. Based on ANOVA analysis, **there may be a statistically significant difference in Profit across States.**  But it is not that significant since the p-value is around 0.05.
3. Regression

Simple regression

Profit = 49032.90 + (0.85 \* R&D Spend)

R2 value = 0.947

R&D spending is a significant driver of profit for startups.

Multiple regression

Profit = 45542.39 + (0.78 \* R&D Spend) + (0.04 \* Marketing Spend)

R square value: 0.952

Both R&D and marketing efforts are important for driving profitability.  A larger coefficient indicates a greater impact means R&D spending is high compared to marketing spending.