**Problem Statement**

Autolib’, a French company in Paris, was an electric car sharing service company that closed down on 31st July, 2018. We have been tasked to work as a Data Scientist for the Autolib electric car-sharing service company to investigate a claim about the blue cars from the provided Autolib dataset.

In an effort to do this, we need to identify some areas and periods of interest via sampling stating the reason for the choice of method, then perform hypothesis testing with regards to the claim that we will have made.

Our random variables to be used for the hypothesis testing are the day types, Bluecars returned and the recharging slots freed.

**Hypothesis Testing**

Null Hypothesis: The number of Bluecars returned is equal to the number of recharging slots freed on the weekend.

Alternative Hypothesis: The number of Bluecars returned is not equal to the number of recharging slots freed on the weekend.

Our hypothesis seeks to find out if there are an equal number of freed recharging slots for the Bluecars to ensure that there is efficiency and business continuity.

**Data Description**

We have one main dataset available for this project. A detailed description of the dataset is provided as follows:

* **Autolib’ Dataset -** this dataset contains data on the address, stations’ information, types of cars and availability, charging information and status.

**Hypothesis Testing Procedure**

Hypothesis testing was done on a colab notebook. The process was as follows:

1. Found and dealt with outliers, anomalies, and missing data within the dataset.
2. Plotted appropriate univariate and bivariate summaries recording our observations.
3. Implemented the solution by performing hypothesis testing.

Our hypothesis testing wants to find out if there are enough recharging slots freed at the weekend to ensure the Bluecars are able to be charged efficiently and there is business continuity.

The test statistic used was the t-test because:

The variables, bluecars returned and recharging slots freed, are independent observations.

Random sample was used to select the sample and sample mean was calculated. Therefore the sampling distribution of sample means is normally distributed.

The significance level used is 0.05 or 5%.

**Hypothesis Testing Results**

The number of Bluecars returned is not equal to the freed recharging slots.

The test statistic value is 2.3 and the p-value is 0.03. Therefore, we reject the null hypothesis and accept the alternative hypothesis.

The sample means were:

Bluecars Returned = 124.65

Slots freed = 24.05

**Summary and Conclusions**

The evidence from the research shows that the recharging slots freed are not equal to the number of Bluecars returned. Autolib company should ensure there are enough recharging slots to charge the returned Bluecars in the appropriate time/duration to ensure the business runs smoothly.