Hypothesis Testing Report

## Problem Statement

In this project, we will seek to understand electric car usage by solving for another research question.

We will 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.

We will look specifically at the weekday data and we will select a simple random sample of two postal codes to compare the blue cars taken. Our claim is that in any two given stations, the average number of blue cars taken is never the same.

## Data Description

We will with data from the Autolib electric car-sharing service company to investigate a claim about the blue cars from the provided Autolib dataset. The variable we will be investigating is the blue cars taken in any randomly selected areas identified using the postal codes.

## Hypothesis Testing Procedure

Our claim, as highlighted in our question above, is going to serve as our null hypothesis i.e. in any two given stations, **the average number of blue cars taken is never the same.**

The alternative hypothesis will, therefore, be that; **the average number of blue cars is the same in every station.**

**Our test statistic is a z test:**

**# Test statistic**

**ztest , pval1 = stests.ztest(p1['bluecars\_taken\_sum'], x2= p2['bluecars\_taken\_sum'], value=0, alternative='two-sided')**

**print(float(pval1))**

**if pval1 < 0.05:**

**print("reject null hypothesis")**

**else:**

**print("fail to reject null hypothesis")**

This will compare our result with the set level of significance which is 0.05.

If it is less than the level of significance the null hypothesis is rejected

## Hypothesis Testing Results

The p-value = 4.3784397428011925e-57

This p-value is a very small value and it is less than the level of significance (0.05)

Hence we reject the Null Hypothesis

## 5. Summary and Conclusions

* There is sufficient evidence to conclude that the difference in the means of blue cars taken in postal code 92160 and 93600 is not 0.
* The test was statistically significant.