# Problem Statement

### A claim says that postal code 1234 have more cars taken within a day compared to postal code 1244.

The Null Hypothesis:

H0 =

The Alternative Hypothesis:

H =

# Data Description

### The sample used for this hypothesis test required two mean from this dataset **link** ( this link lead to Autolib dataset) with its description **link**.

### Below is the description of each column:

|  |  |
| --- | --- |
| **Column name** | **Explanation** |
| Postal code | postal code of the area (in Paris) |
| date | date of the row aggregation |
| n\_daily\_data\_points | number of daily data points that were available for aggregation, that day |
| dayOfWeek | Identifier of weekday (0: Monday -> 6: Sunday) |
| day\_type | weekday or weekend |
| BlueCars\_taken\_sum | Number of blue cars taken that date in that area |
| BlueCars\_returned\_sum | Number of blue cars returned that date in that area |
| Utilib\_taken\_sum | Number of Utilib taken that date in that area |
| Utilib\_returned\_sum | Number of Utilib returned that date in that area |
| Utilib\_14\_taken\_sum | Number of Utilib 1.4 taken that date in that area |
| Utilib\_14\_returned\_sum | Number of Utilib 1.4 returned that date in that area |
| Slots\_freed\_sum | Number of recharging slots released that date in that area |

### This data did have any missing value nor duplicate value. The only part that needs to change was the date column to make it uniform.

# Hypothesis Testing Procedure

### The z-statistics was used for the test statistics. The reason for using it was due to the sample size being greater than 30.

# Hypothesis Testing Result

### The hypothesis was rejected. The p-value was less than the significance level which was 0.05.

# 

### 