1. **Business Understanding**

Autolib is an electric car sharing company managed by the Bolloré group enterprise that was started in Paris in the year 2011 by Marie Bolloré and closed in the year 2018 because the company got into bad debt. The service was introduced to complement an existing bike-sharing service.

Autolib wishes to expand its operations within Europe and beyond. The operations team has been tasked by the senior leadership team to develop a strategy that will guide in this process

We are going to investigate the mean difference of the blue cars taken between two different postal codes. The assumption is the mean difference between Blue Cars picked on weekdays is equal to zero.

H0 : The difference in the mean of blue cars taken in postal code 75006 and 75012 is 0

H1 : The difference in the mean of blue cars taken in postal code 75006 and 75012 is not 0

OR

H0 : (U1 = U2)

Ha : (U1 ≠ U2)

Confidence interval = 95%

Level of significance = 5%

If the p\_value obtained from the sample size is less than our level of significance which is 0.05, fail to accept the null hypothesis as there will be enough evidence that there is a significant difference between the means and accept the alternative.

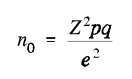
1. **Data Understanding**

The data provided consists of different electric cars the company has and since we are only interested in the Blue cars, we are going to drop the irrelevant data and focus only on the blue cars. The total records provided are 16085 with 9 columns. You can find the [description](http://bit.ly/DSCoreAutolibDataset) of the columns and the [data](http://bit.ly/DSCoreAutolibDatasetGlossary) used respectively in the embedded links.

**Sample size**

For analysis we are going to select a sample population from each postal code . Every postal code has approximately 156 records each. Since the focus is on two postal codes, the total records that can be fetched is 312 records for investigation.

We are going to determine the sample size using the Cochran’s formula:



e is the desired level of precision (i.e. the margin of error) - we take 95% confidence interval which equates to *0.05* level of significance

The z -score for 95% confidence interval equates to = 1.96

From the formula we need approximately 30 records from each postal code to ensure that the data is normally distributed.

**Z Test.**

When to use Z test:

* When your sample size is greater than 30. Otherwise, use a t test.
* Data points should be independent from each other. i.e, one data point isn’t related or doesn’t affect another data point.
* Your data should be normally distributed. However, for large sample sizes (over 30) this doesn’t always matter.
* Your data should be randomly selected from a population, where each item has an equal chance of being selected.
* Sample sizes should be equal if at all possible.

1. **Hypothesis testing procedure**

**STEP 1 : Stating the Hypothesis**

H0 : The difference in the mean of blue cars taken in postal code 75006 and 75012 is 0

H1 : The difference in the mean of blue cars taken in postal code 75006 and 75012 is not 0

OR

H0 : U1 = U2

Ha : U1 ≠ U2

**STEP 2 : Level of significance**

α = 0.05

**STEP 3 : Calculating the test statistic & Applying the conditions for an appropriate decision.**

The calculated p\_value = 3.571771542382906e-172

**STEP 4 : Evaluating our results**

The p-value = 3.571771542382906e-172

This p-value is a very small value and it is less that the level of significance (0.05), hence we Reject the Null Hypothesis

**STEP 5 : Interpreting the results**

There is sufficient evidence to conclude that the difference in the means of blue cars taken in postal code 75012 and 75006 is not 0.

The test was statistically significant.

**Conclusion**

Based on the hypothesis test, we conclude that the difference between the average number of cars taken from stations with postal code 75015 and the average number of cars taken from stations with postal code 75016 is statistically significant; therefore we reject the null hypothesis for the alternative hypothesis.