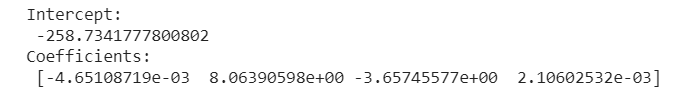
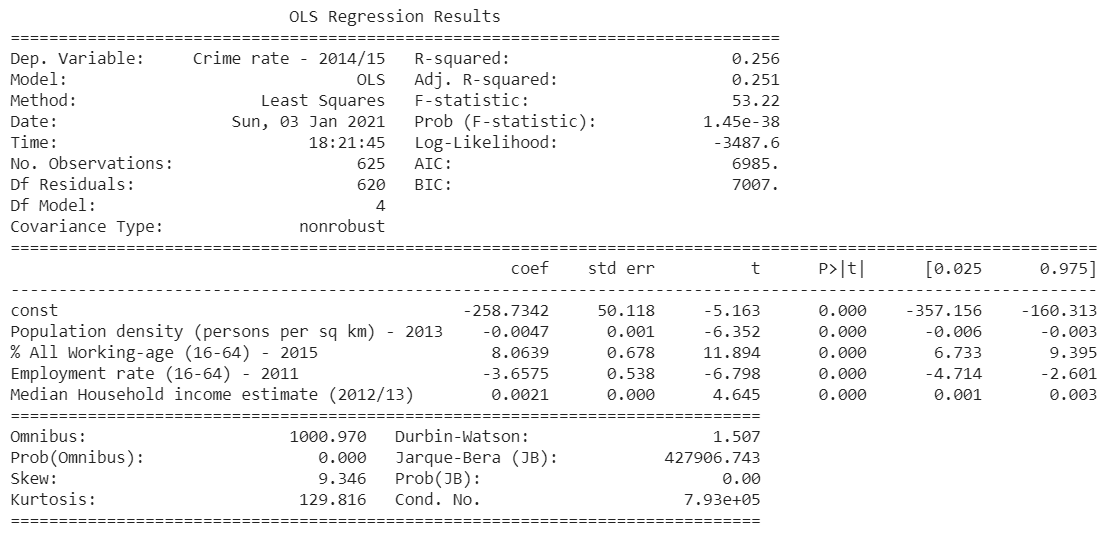
**INTERPRETING THE DATA OUTPUT OF MLR**

All Wards:

The core output is this:



The intercept is the y-intercept of line of best fit, coefficients are discussed below.



Firstly the R-squared value is 0.256, meaning that it is not a very well fitting model (maybe the reality is not linear, but quadratic, or cubic.)

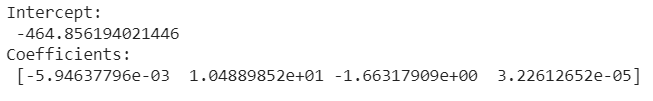
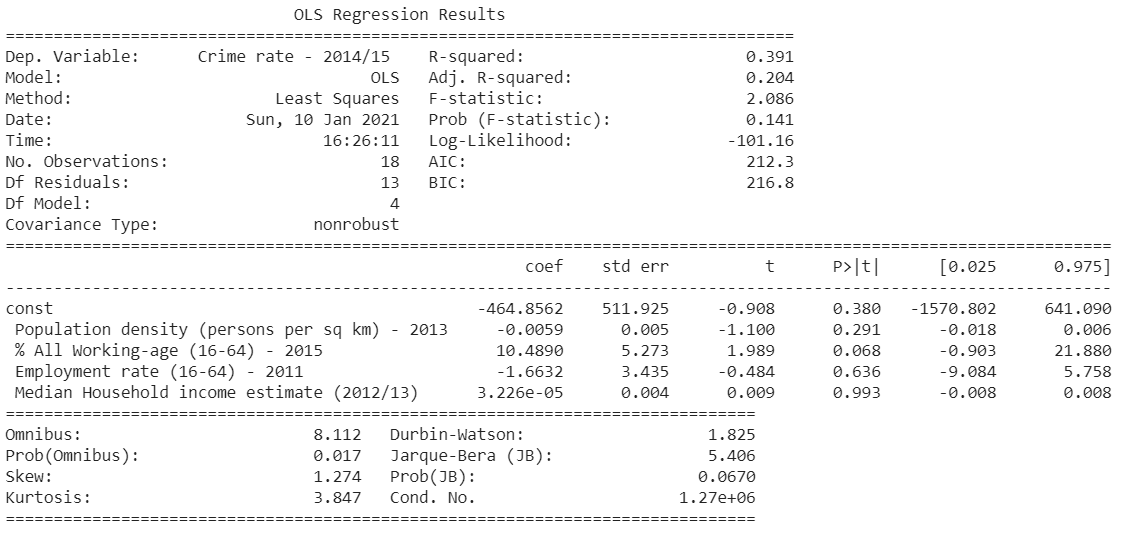
Secondly the coefficients (crime rate is measured as number of crimes reported per 1000 people) :

* Population density: -0.0047 (-0.00465108719) meaning that every additional 1 person in a km­2 area, crime rate decreases by 0.0047; or every additional 1000 people per km2 decrease crime rate by 4.651.
* % working age: every 1% increase in the working age population increase crime rate by 8.064.
* Employment rate: every 1% increase in employment rate decrease crime rate by 3.657.
* Median house income: every £1 extra increase in median household income increases crime rate by 0.0021, or every additional £1000 means an increase of 2.106.

Please be wary that only correlation is indicated, not causation.

Thirdly we are confident about the statistical significance of the coefficients as their t-values are all very high, well outside the 5% two-tailed critical values, meaning that we are very sure that these factors do indeed have correlation with crime rate.

Camden:

Firstly the R-squared value is 0.391, meaning that it is an okay fitting model, better than the model for all wards.

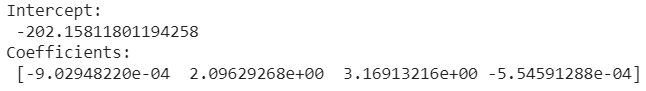
Secondly the coefficients (crime rate is measured as number of crimes reported per 1000 people) :

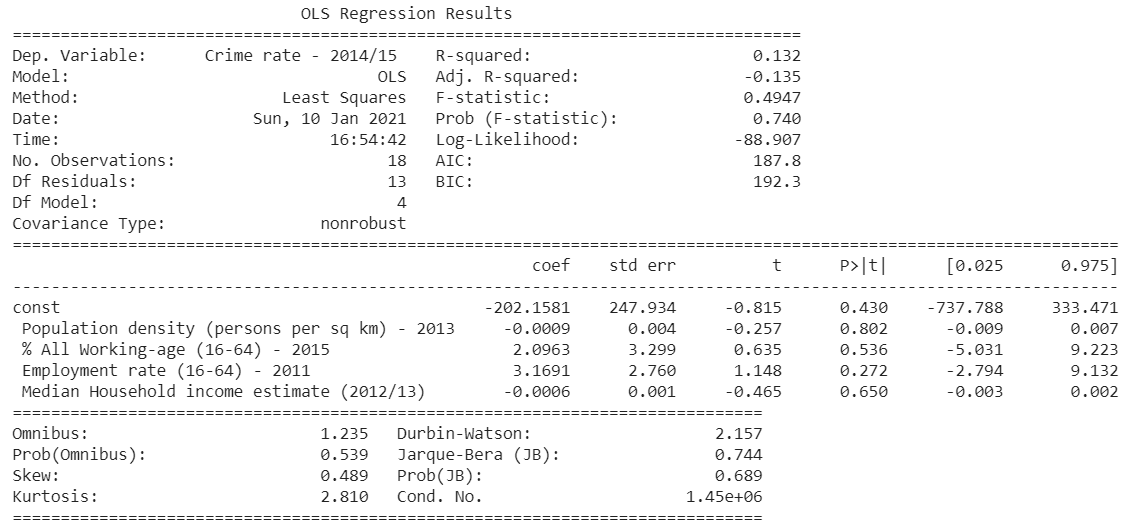
* Population density: -0.0059 (-0.00594637796) meaning that every additional 1 person in a km­2 area, crime rate decreases by 0.0059; or every additional 1000 people per km2 decrease crime rate by 5.946.
* % working age: every 1% increase in the working age population increase crime rate by 10.489.
* Employment rate: every 1% increase in employment rate decrease crime rate by 1.6632.
* Median house income: every £1 extra increase in median household income increases crime rate by 0.00003226, or every additional £1000 means an increase of 0.0323.

Please be wary that only correlation is indicated, not causation.

Thirdly on statistical significance: none of the coefficient, as well as the constant is not statistically significant at 5%. The coefficient for working age population has a confidence level of 85% (2 tailed)

Kensington and Chelsea:





Firstly the R-squared value is 0.132, meaning that it is a badly fitting model, worse than the Camden or the all wards model.

Secondly the coefficients (crime rate is measured as number of crimes reported per 1000 people):

* Population density: -0.0009 (-0.00090294822) meaning that every additional 1 person in a km­2 area, crime rate decreases by 0.0009; or every additional 1000 people per km2 decrease crime rate by 0.903.
* % working age: every 1% increase in the working age population increase crime rate by 2.0963.
* Employment rate: every 1% increase in employment rate increase crime rate by 3.1691.
* Median house income: every £1 extra increase in median household income increases crime rate by 0.0006 (0.000554591288), or every additional £1000 means a decrease of 0.6.

Please be wary that only correlation is indicated, not causation.

Thirdly on statistical significance: none of the coefficient, as well as the constant is not statistically significant.

Summary:

The following table compares the effect of each variable on crime rate (measured by number of crimes reported per 1000 people.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | R2 | Constant | Population density (1000 per km2) - 2013 | % All Working-age (16-64) - 2015 | Employment rate (16-64) - 2011 | Median Household income estimate (per £1000) (2012/13) |
| All wards | 0.256 | -258.7 | -4.6512 | 8.0639 | -3.6575 | 2.1060 |
| Camden | 0.391 | -464.9 | -5.9464 | 10.4889 | -1.6632 | 0.0323 |
| Kensington and Chelsea | 0.132 | -202.2 | -0.9029 | 2.0963 | 3.1691 | -0.5546 |

**How well the data fits the trendline**

As we can see in the table, the R2 of all sample ranges are not particularly high, meaning a line of best fit does not represent the sample’s distribution well.

**What the coefficient tells us**

For population density, Camden shows the largest correlation between it and crime, while Kensington and Chelsea do not show much correlation.

For percentage of working age population, Camden shows the largest correlation between it and crime, while Kensington and Chelsea do not show much correlation.

For employment rate, all wards and Camden show a negative correlation between it and crime, while Kensington and Chelsea shows a positive correlation.

For median household income, all wards show a positive correlation between it and crime, while Camden shows very little correlation, and Kensington and Chelsea shows a negative correlation.

**Confidence level of the data**

As the sample size for Camden and Kensington and Chelsea is relatively small, the variance (as therefore the standard deviation) is particularly large. The coefficients are not statistically significant to 95%.

As for all wards, there is a much larger sample. The variance is very small and they are statistically significant to 99.999%