|  |
| --- |
| How could UK Police Resources be better placed? |
| Ali Kokaz |
|  |
| Through ingestion of public records and databases, crimes over a 5 year period (2011-2015) were plotted and analysed, with areas of concern outlined. |
| Contents  [**Introduction** 1](#_Toc461172681)  [**Data sources & ingestion** 1](#_Toc461172682)  [**Database cleansing & normalization** 1](#_Toc461172683)  [**Results & Analysis** 3](#_Toc461172684)  **Crime location case study: London………………………………………………………….………8**  [**Conclusion** 10](#_Toc461172685)  [**Appendix** 11](#_Toc461172686) |
|  |

How could UK Police Resources be better placed?

Ali Kokaz

# Main Facts and Figures

• • •

Total crime decrease of 15% from 2011 to 2015.

40% increase in violent and sexual crimes between 2011-2015.

Burglary shows biggest improvement, decreasing by 85% in the 5-year period, with London being the main driver behind the improvement.

Anti-Social behavior crimes remain the biggest offenders, accounting for 48% of total crimes.

## **Introduction**

Software knowledge and expertise gained over the past 6 weeks were focused into analyzing the crime habits and trends over a 5- year period.

I will outline the process and techniques used to process this data throughout the project, I will also analyse and discuss my findings later in this report.

## **Data sources & ingestion**

Using <https://data.police.uk/>, crime data was downloaded, providing crime types, location and outcomes. Stop & search data was also collected.

Population data was ingested from both the Office of National Statistics (ONS) in addition to the Home Office, which also provided police force numbers.

Data was loaded onto a SQL database using custom made “for each” loops on Visual Studio, allowing the software to iteratively search each file and upload the data from the associated .CSV documents.

## **Database cleansing & normalization**

Some of the data provided by the Police was found to be “dirty”, with missing columns or values shifted under other columns. In order to minimize their effect on results and maximize accuracy of findings, two approaches were considered for this data.

1. Where possible, data was recovered through altering the columns and re-importing the data into the database, after removing the duplicates in the database. (an example of this type of data is available in the appendix).
2. Data missing the primary vital information (location, type or date) were deemed incomplete and were removed. (a total of 154,306 data rows were removed through this method).
3. Data where offender was found “not guilty” were also removed (police expenditure on cases found “not guilty” are significantly lower).
4. Data outside the range of 2011-2015 was also removed, partly to reduce the sheer number of data present, and also to focus on full year period.

Unnecessary columns, (such as last outcome category, location, LSOA name, reported by, context, Crime ID) were deleted from the main table using the ALTER TABLE, DROP COLUMN commands.

A new, main table was created, and would be linked to numerous other look-up tables within SQL. The new central table will contain ID numbers relating to the look-up tables, maximizing performance by making the table very thin data wise. The data-base relationship diagram is given in figure 1 to show how the look-up tables link to the central table.



**Fig 1: Data-base relationship diagram showing how the different look-up tables link to the central table.**

Look-up tables were created using CREATE TABLE commands, with the IDENTITY (1,1) first column to produce an ID column, and INSERT INTO was used to feed values into the newly created tables, where SELECT DISTNICT was used to return the distinct value types, (i.e. the discrete crime types).

Examples of these look up tables (and associated syntax), along with an extract of the main table can be found in the appendix.

Using these methods, along with creating indexes (through the project explorer) for the relevant ID’s, produced a query time enhancement from 4:58 mins to 20 seconds.

In addition, views focusing on the primary analysis topics of this project were created, allowing data to be called up more efficiently and be imported into Excel for analysis.

## **Results & Analysis**

3 main topics were focused upon during the analysis of this project, being:

* Time: monthly and yearly crime evolution rates were looked at, through crime number, type and region evolution.
* Location: Crimes were grouped under the constabulary responsible, since this is where the budgets will be split into. Maps of the crime locations were created to give a visual representation of crime hotspots to suggest where to adequately place spending.
* Type: crimes were grouped by type, giving a good insight into the breakdown of severity of crimes. (anti-social behaviour vs Rape for example).

#### Crime by time

total crime numbers show a gradual yearly decrease, with a dramatic decrease from 2011 to 2012. A secondary reason for this decrease is the change in documentation procedures introduced. This trend is displayed in figure 2.

**Figure 2: Yearly crime evolution. A noticeable decrease is present between 2011 to 2012, however a slight increase is seen towards more recent times.**

Looking in more detail at this trend, by plotting monthly crime numbers, again displays this downward trend in crime numbers, however a seasonal spike can be seen in the data, with crimes generally increasing in the summer months. This trend is displayed in Figure 3.

**Figure 3: A plot of country-wide monthly crime numbers. The solid line represents actual crime numbers, while the dashed line represents a linear moving average, showing general decrease in crimes.**

One worrying aspect of crime evolution through the 5-year period can be noticed when looking at crime type number evolution. The 16 crime types given by the police data were merged into 4 main categories, in the following manner:

|  |  |
| --- | --- |
| Theft | Bicycle theft |
| Burglary |
| Robbery |
| Shoplifting |
| Theft from person |
| Other theft |
| Violence | Violent crimes |
| Violence and sexual offences |
| Criminal damage and arson |
| Possession of weapons |
| Drugs & disorder | Anti-social behaviour |
| Drugs |
| Public disorder |
| Disorder |
| Vehicle crime |
| Other | Other Crimes |

Plotting the evolution of these crime types shows a remarkable decrease in burglary and other crimes, however violent crimes show a worrying year-on-year increase since 2013, rising by 15.7% since then. Drugs and disorder crimes also show an increase; which experts believe is down to the increasingly relaxed view society is taking on drugs. These crime evolutions are shown graphically in Figure 4.

**Figure 4: a yearly plot of crime numbers by crime type. There is a sharp decrease in both burglary and other crimes, with a small increase in both Violent and Drug crimes.**

#### Location

Arguably the most interesting variable in crime rate was the location or regions crime occurred at. Firstly, crimes were split into their respective constabularies and number of crimes consequently counted. As expected the London area trumps every other in that respect, however this does not paint the full picture.

In order to give a clearer depiction of police force efficiency, population numbers along with crime rates per person were calculated. This threw up some interesting results, as then London metropolitan areas dropped to 3rd in the list, with Cleveland Police and South Yorkshire topping the tables.

Investigating further, police force numbers were taking into account, to give a police officer: population ratio. From there, crimes per officer, per 100,000 population figures were produced, reflecting the amount of crimes each officer had to deal with, relative to the population.

In this data, which is covered in Figure 5, Metropolitan Police remarkably drop to lowest on the list, suggesting under-efficiency by the police officers in dealing or reporting crimes, with South Yorkshire Police coming 1st, suggesting more funds need to channelled there to address the under policing in effect in those regions.

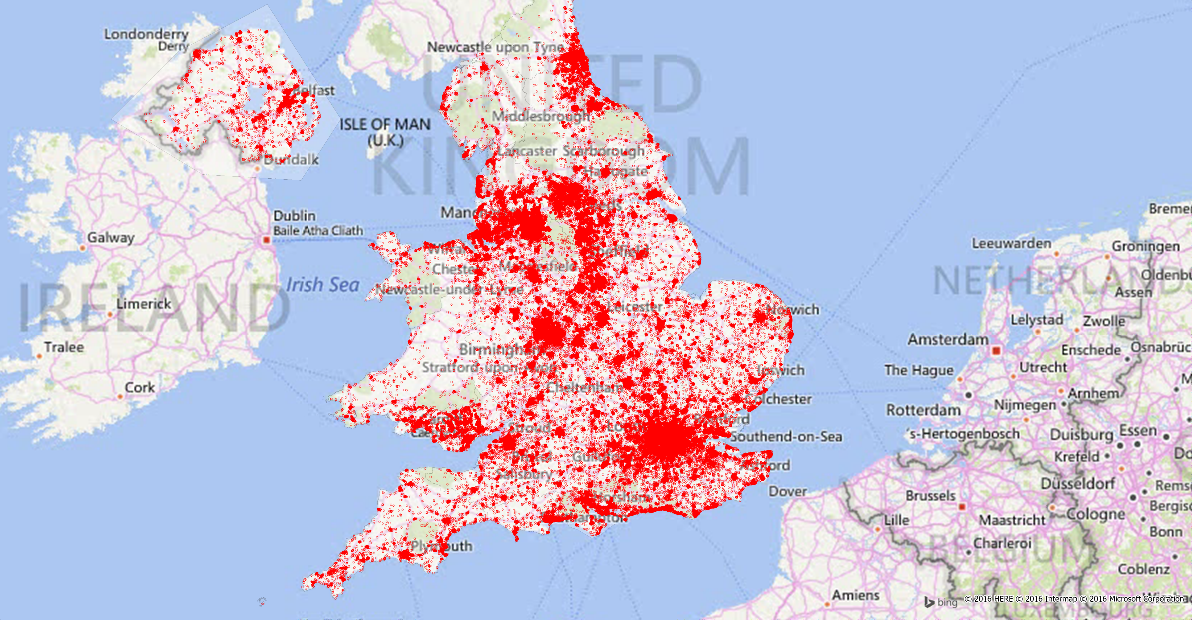
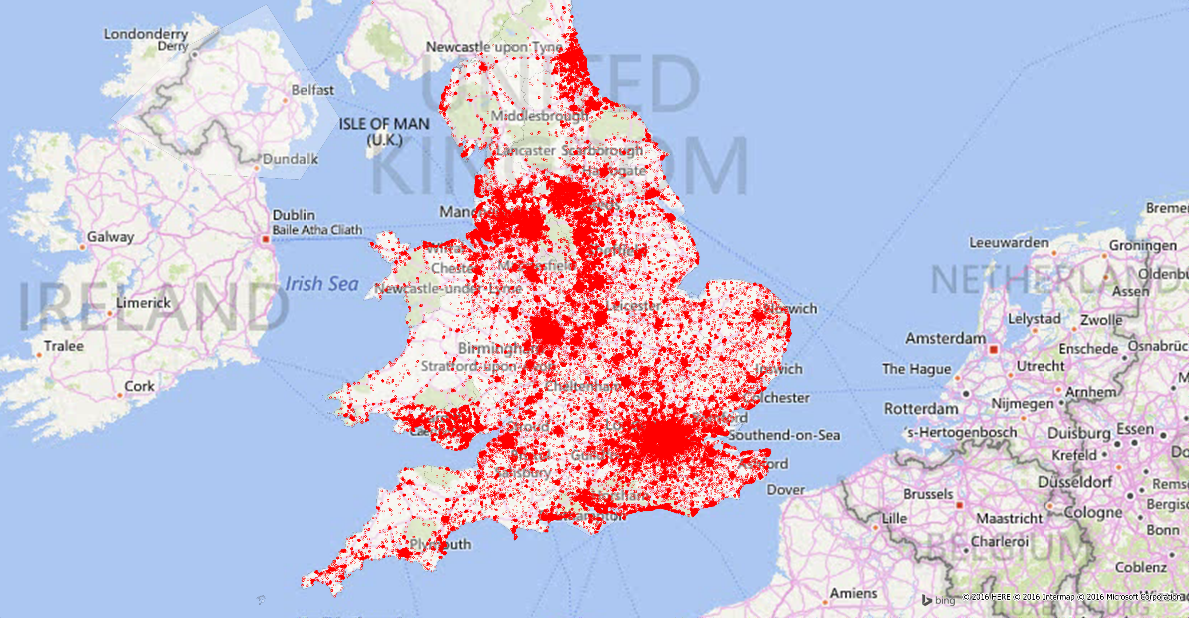
**Figure 5: Graph displaying number of crimes per officer, per 100,000 people, grouped by Constabulary, shown by the solid bars. The dashed line represents the national average. Clearly visible are the 2 Yorkshire constabularies and Cleveland police being over-run with crimes per officer.**

[](Crime%20Heat%20Maps.xlsx)Crime heat maps were produced to give a visual representation; these can be found by clicking the link below.

[Crime Heat Maps](Crime%20Heat%20Maps.xlsx)

Go to the insert tab 🡪 choose 3D maps 🡪 and then choose the crime type heat map you would like to view.

It’s interesting to note that against common beliefs and theories, the location of crimes does not change radically with the time of year. Theories have previously suggested that as summer approaches, crime move outwards and towards the coastlines to reflect the movement of people vacationing, however as seen in Figure 4, no discernible difference can be seen from the timing of year on location of crimes.



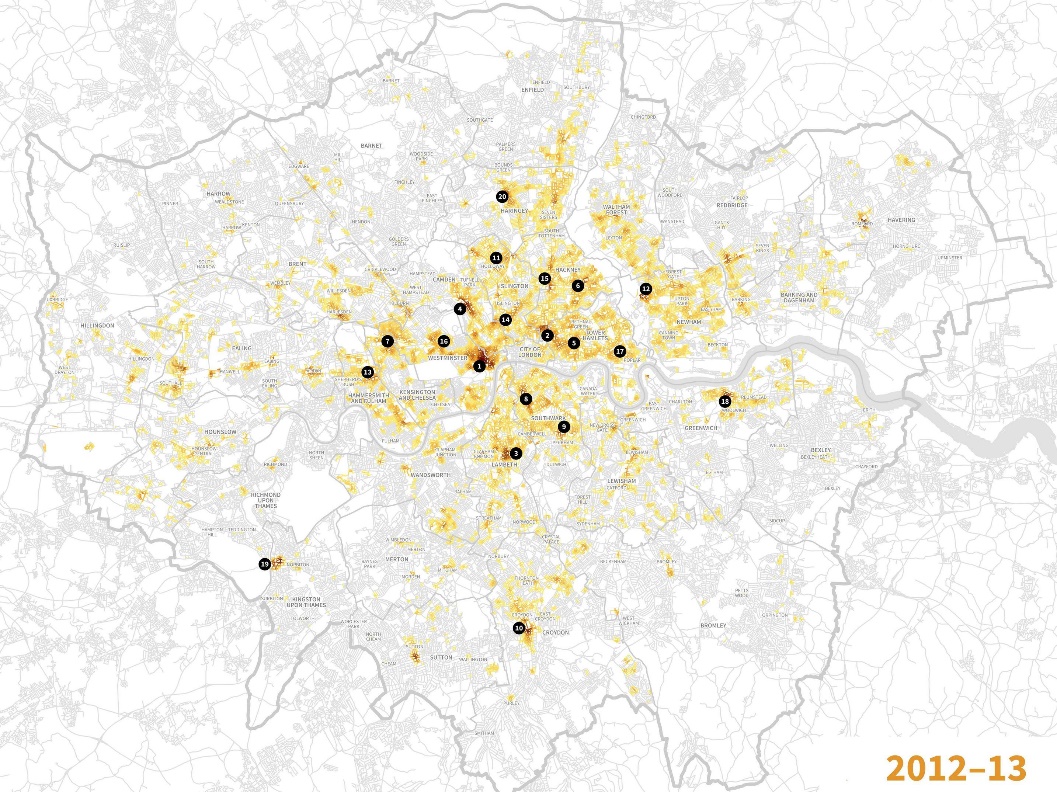
**Figure 4: A visual comparison of the crime locations occurring nationally between summer and winter. Notice the remarkable similarity between the two maps, showing little difference to crime locations vs time of year**

**2015-07**

**2015-01**

#### Crime Location Case Study: London

Using open crime data from [police.uk](http://data.police.uk/) and map data from [OS OpenData](http://www.ordnancesurvey.co.uk/oswebsite/products/os-opendata.html), I’ve created a map of violent and sexual crime hotspots in London, based on 135,000 crimes that occurred between April 2012 and March 2013, shown in Figure 6.



**Figure 6: Crime hot-spots in London over the period of 2012-2013. The full resolution version available for browsing is enclosed in the host document.**

Central London is in a class of its own: seven of the top ten violent-crime hotspots fall within an area between Camden Town in the north, Poplar in the east, Brixton in the south and Hyde Park in the west. The area around Soho and Leicester Square has the highest concentration of violent crime anywhere in London, at-least partly because it has vast numbers of people passing through it to visit shops and, especially, bars. The same applies to the next three hotspots on the list: Shoreditch, Brixton and Camden Town.

**Figure 7: Violent crimes in Croydon and Kingston, 2012-2013.**

Several of London’s suburban centres seem to suffer a lot of violent or sexual crime while being surrounded by low-crime areas. In the cases of Croydon (particularly around the High Street), Kingston (around the train and bus stations) and Romford (South Street), this is probably due to these areas attracting lots of late-night revellers from surrounding suburbs, as anyone who’s spent a Friday night in any of them can attest. In the cases of Romford and Croydon—but not Kingston—these small areas of crime have given an entire borough a bad reputation, which must be galling if you live in ultra-leafy Hornchurch or Shirley.

  
Before the major road-building programmes of the 1960s, people got in and out of London along long single-carriageway roads. Many of these live on as local high streets, several of which stand out as hot routes. They usually have lots of small shops, pubs and restaurants, so there are plenty of potential offenders and victims around.

**Figure 8: Old arterial routes in London showing heightened crime rates.**

I have also built two interactive web-apps showing crime evolution in London boroughs through 2014 vs 2015 by crime type, while also ranking the boroughs from best to worst.

To visit these web-apps, go to:

[Evolution of Crimes in London Boroughs (excl. City)](http://bl.ocks.org/kaltera/raw/401791a2a303c0e8cabc/)

[Best and Worst London boroughs by Crime Type (excl. City)](http://bl.ocks.org/kaltera/raw/cd8bf3228373e06100a7/)

[Another Nice Website displaying national crime rates](http://www.newhousepad.com/)

#### Crime by type

Type of crime committed has been a consistent topic in this project, and previous data have all or nearly all been categorized by crime. However, some further insight is still needed into the predominant crime types.

Drugs & disorder crimes account for the majority of crimes committed, with most of those being anti-social behaviour or public disorder crimes. These are described by police as being “minor” crimes, taking up more police time than effort and money, and generally slow the police force down.

In order to fully maximise the efficiency of police forces, and to minimise the number of “serious” offences, police forces need to focus separate efforts and lighter weight cheaper forces into dealing with these offences.

A break-down of crimes by types is shown in figure 9.

**Figure 9: Pie chart representing the percentage contribution of crime types to total crime, 2011-2015. Notice the majority share of Drugs & Disorder crimes.**

#### Stop & Search

The results of stop & search outcomes returned intriguing results. From the ingested data, crime was filtered by outcome for the total 5-year period.

First thing to note is that many reports are incomplete and scattered, meaning the data is largely dirty and difficult to work with. However, the results showed a very clear trend, the majority of stop & searched resulted in nothing. No findings, or arrests.

**Figure 10: The outcomes of stop & search reports carried out nationally between 2011-2015. There is a remarkable domination by nothing found outcomes, questioning the effectiveness of stop & searches as a crime reduction method.**

The failures of stop & search methods to find any meaningful conclusion to question the need for so many of them (15 million stop & searches carried out between 2011-2015) and do suggest that there are budget savings to be made from the cutting down or scrapping of most stop & search methods.

## **Conclusion**

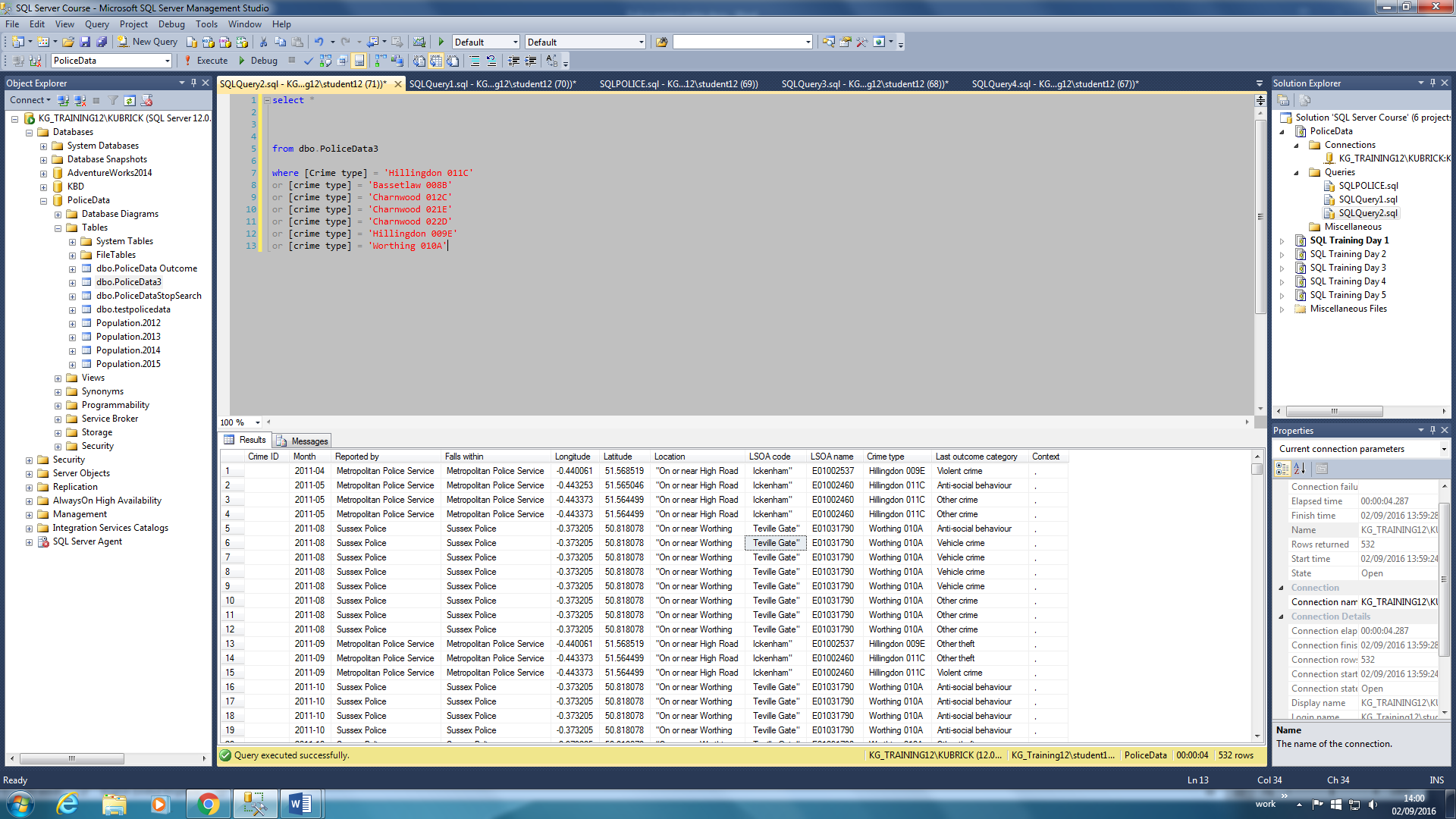
The general decrease pattern for total crime shows that budgets, generally are being put forwards in the right direction and are being effective. However, some areas, especially Yorkshire and Cleveland, show some early signs of needing attention, and must be focused on.

The project overall was successful, with the main tasks and objectives set out being met. In future, improvements could be including adding interactive features to the results (building animations to show crime evolution) to make this report more user friendly.

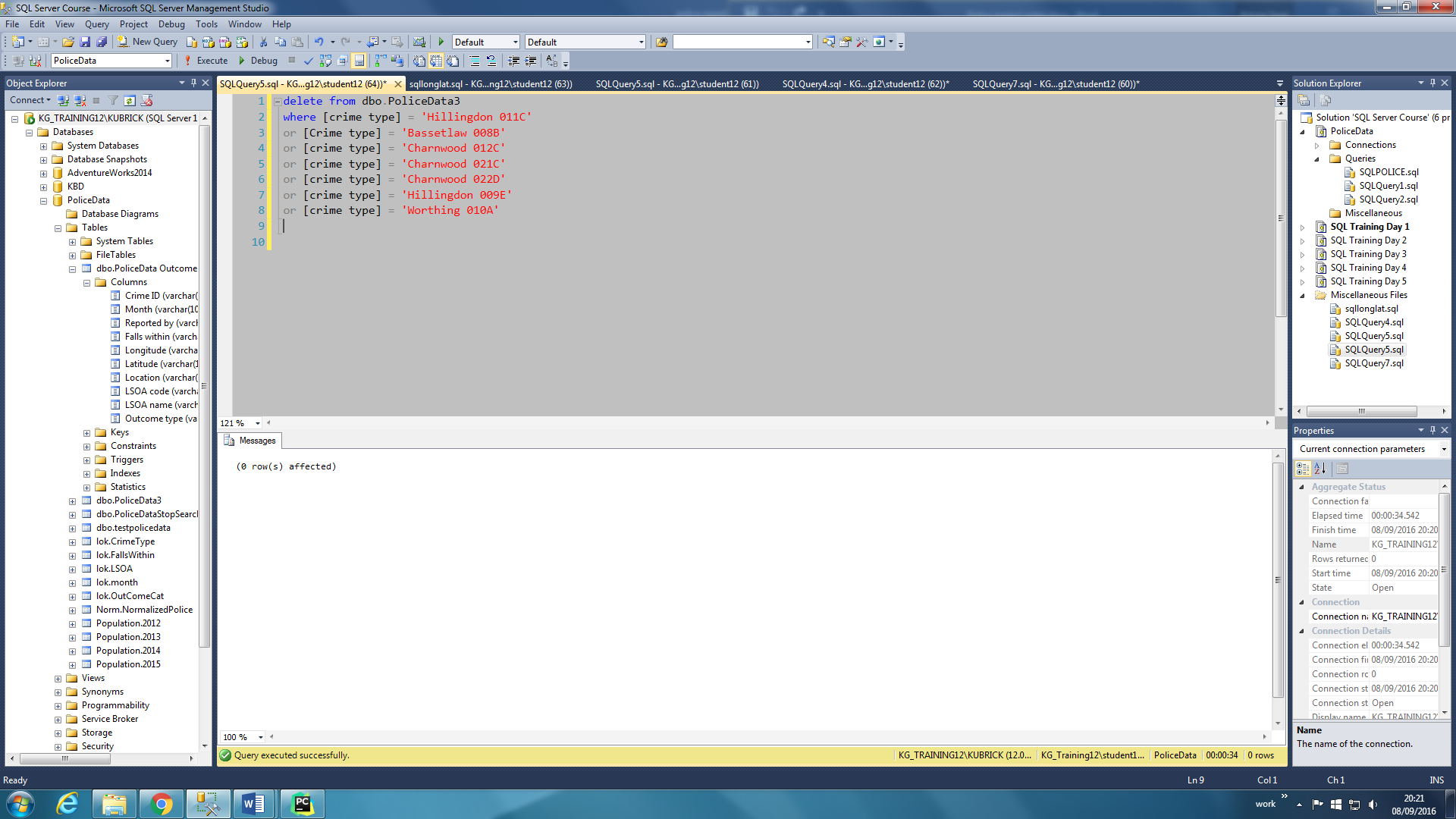
In addition, primary causes for crimes in certain areas could be investigated (i.e. the link between quality of schooling and crime rate in region), I am sure as the Big data industry improves and matures, these aspects can be more efficiently and effectively looked at.

## **Appendix**

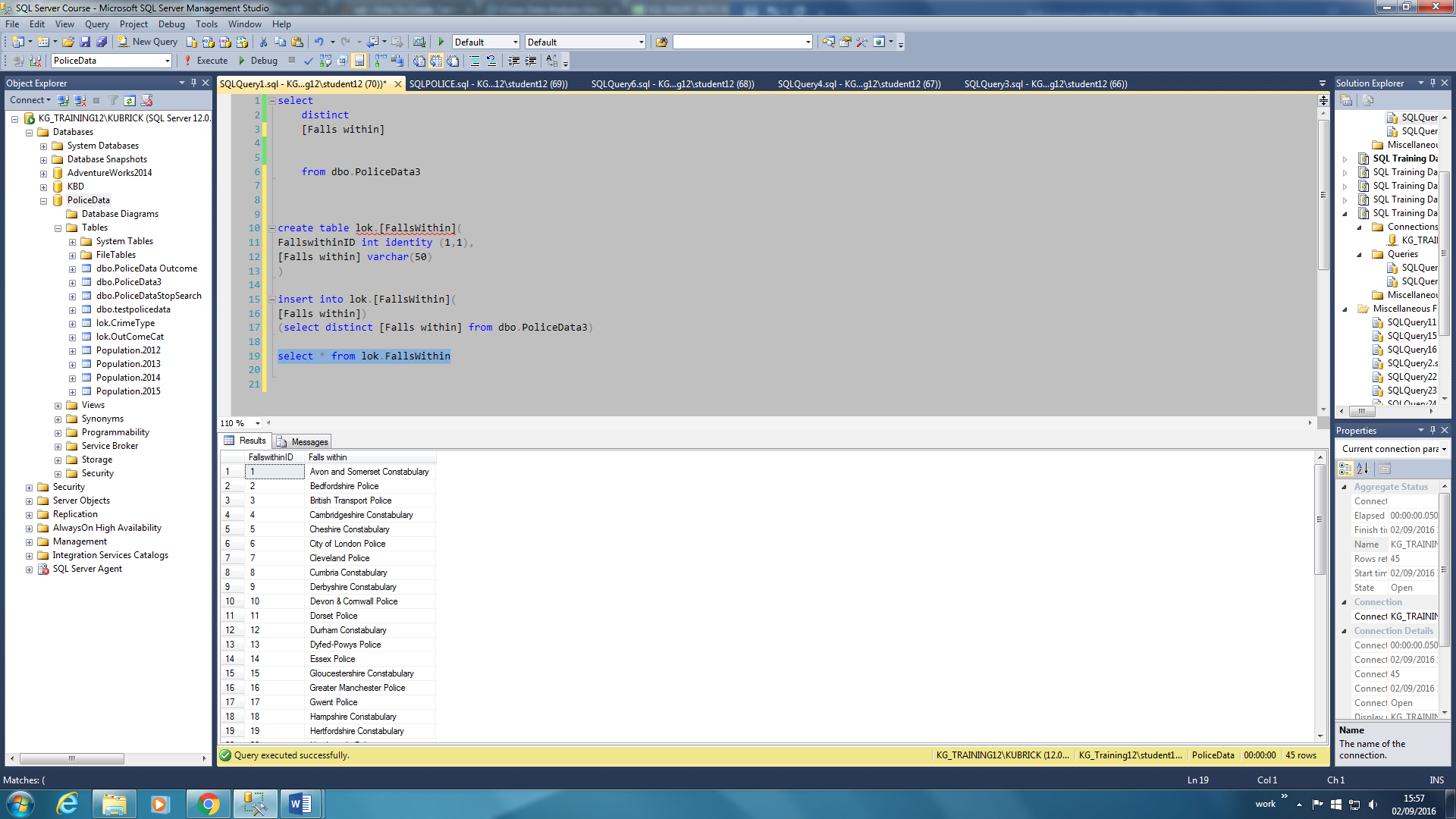
To View the data used in this project, please see the attached Excel documents in this project folder.



Above is an example of some of the dirty data deleted from the main database, notice how from the LSOA CODE onwards, the values are shifted to the right, as a result of a comma in the string values in LOCATION, and since the file is a CSV, it interpreted the next half of the string as a new value in the next column.

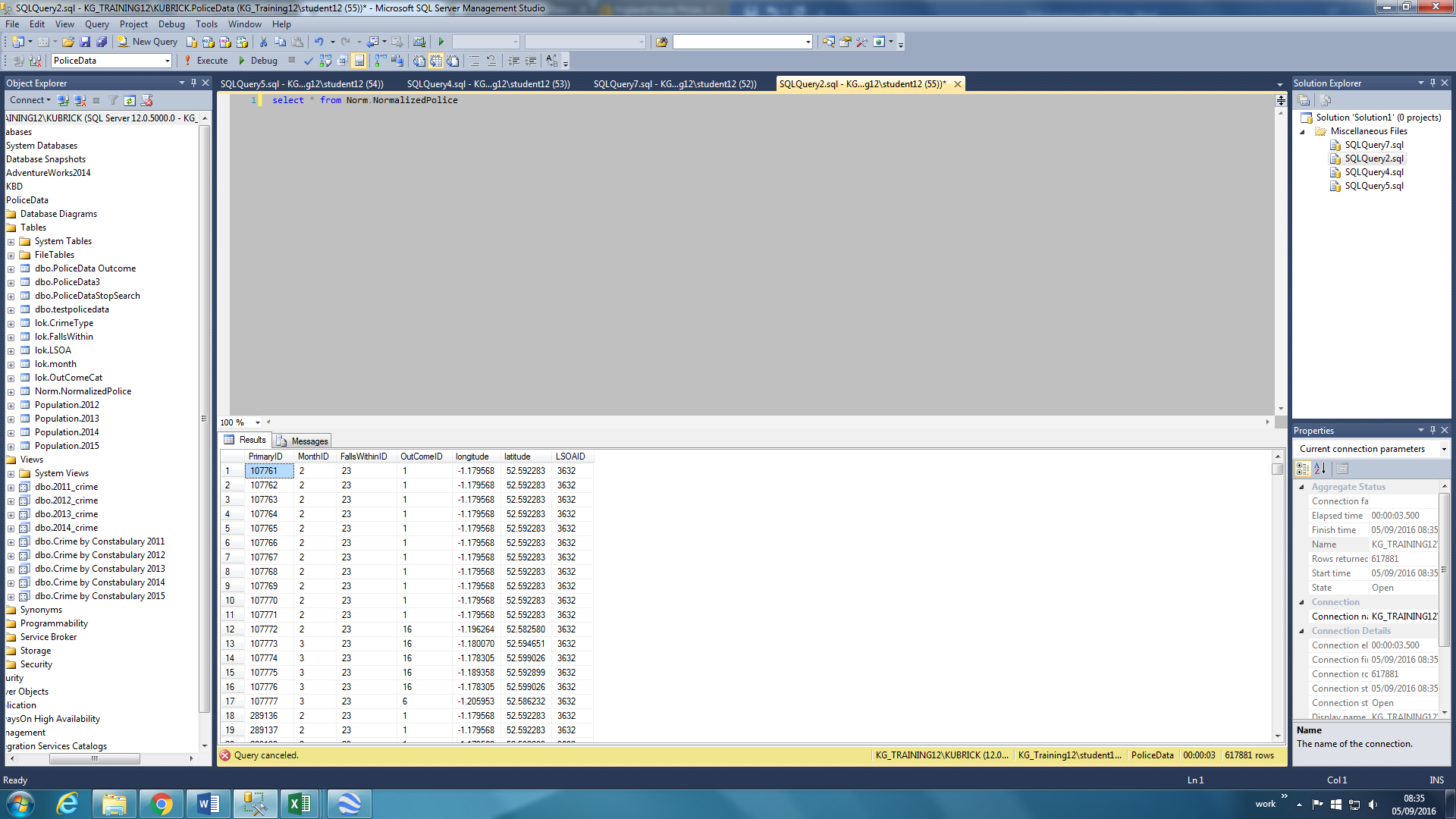


Example of the code used to delete the dirty rows.

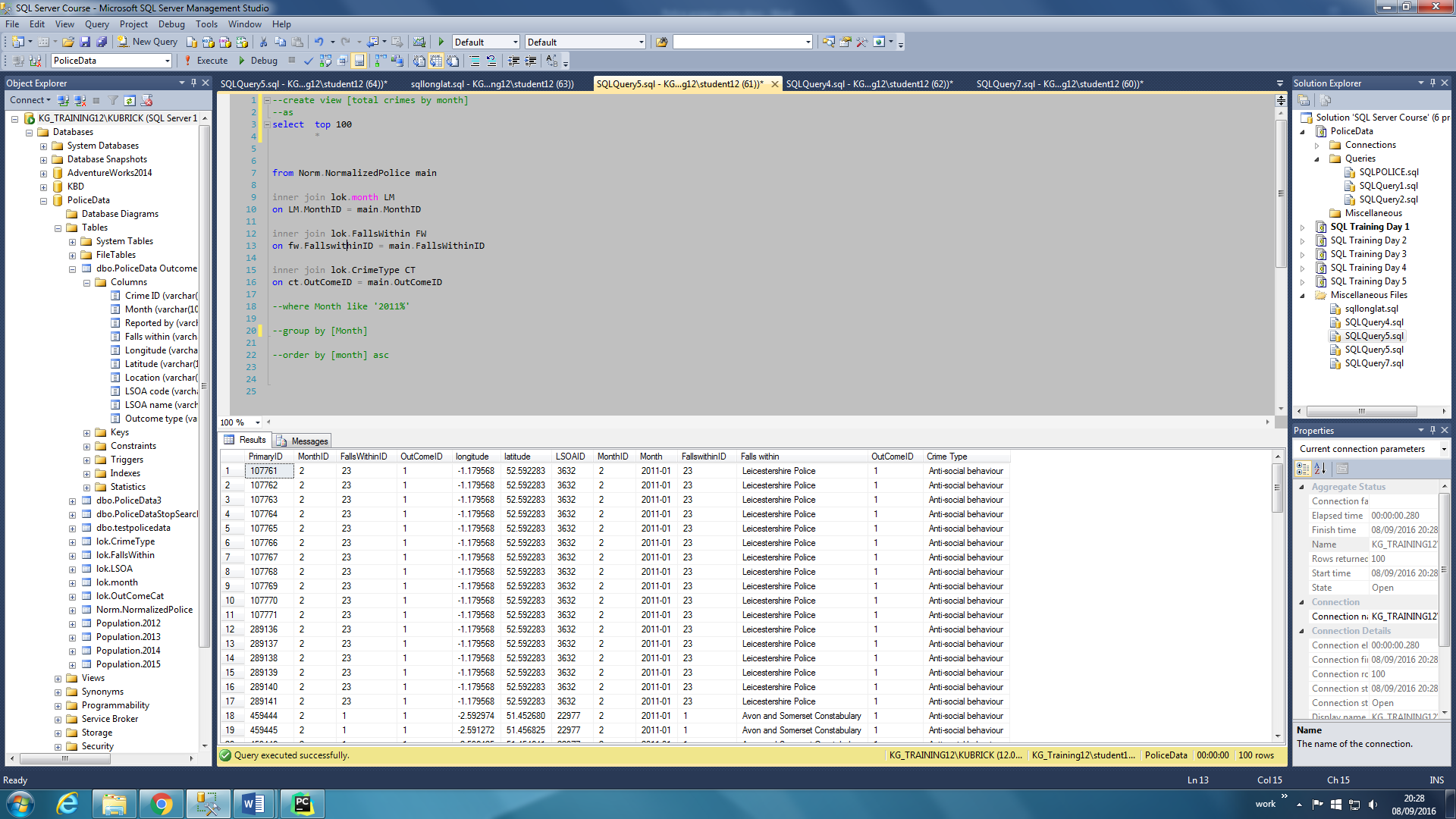


Look-up tables, used in the normalised data-base, were created using the following syntax.

Below is an extract of the final normalised central table. While the table is very long (30,000,000 rows) it is very thin, with values being very small, only needing to reference the look-up tables linked to them.



When connecting to the central table, and for it to make any meaningful sense, the following syntax was used. Notice the heavy use of inner joins linking the look-up tables to the central table.



Views were creating using the following syntax. To the right is a list of some of the various views created.

