Big Data Analytics System for Fact/Data-driven Decision Making

Dr David Chen
Division of Computer Science and Informatics
School of Engineering
London South Bank University
Email: chend@lsbu.ac.uk

Project Background and Objective

- The Cooperative Local Investment Plans (CLIPs) of Lambeth council
  - Provide an evidence-based area plan bringing together an overview of the council's strategies and services that relate to the built environment (transport, parks, education, etc).
  - Bring together information about existing assets (including physical and organisational spatial assets), existing demographic and business intelligence (such as demographic data and categories and levels of deprivation), investment made and investment planned by the council and other public and private partners.
  - To enable the community to provide their views and identify priorities for spend of the Neighbourhood Funding.

Data & Data Sources

- Multiple and Diverse, including
  - NOIMS (National Online Manpower Statistics)
  - Census
  - Gov.uk
  - London Data Store
  - Annual business survey
  - School survey
  - IMD (Index of Multiple Deprivation)
  - Met Police
  - Health Stats
  - …

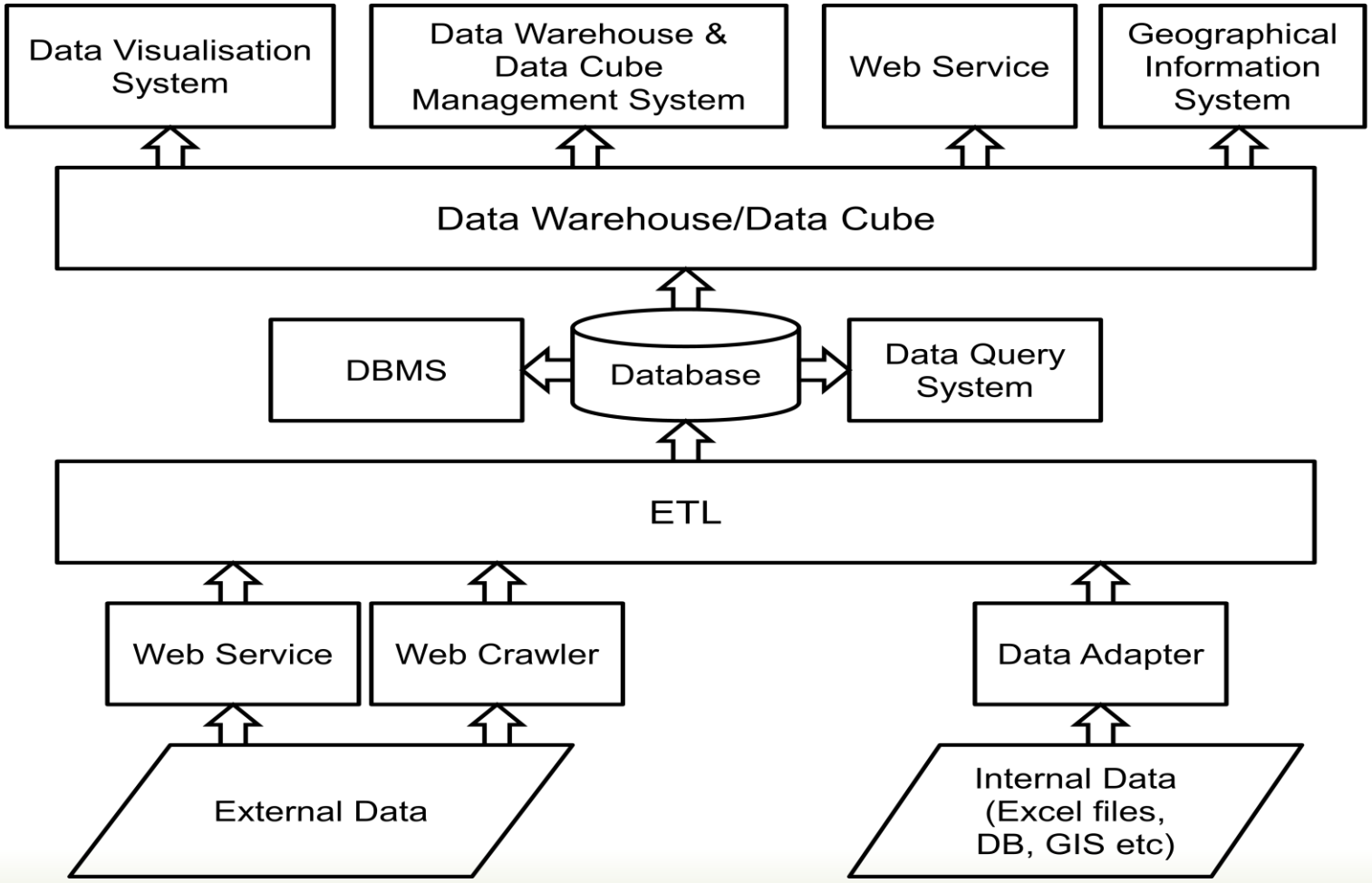
Data & Data Sources

- Graphical information is essential
  - Latitude and longitude
  - Postcode
  - LSOA code (Lower Layer Super Output Areas)
  - Ward code
  - LA code (Local authority code)
  - The council is particular interested in LSOA level's data
- Time
- Use of APIs

Data & Data Sources

- The target data sets
  - Demographics (Census: Population, Ethnicity, Disability, Gender, Tenure, Household composition)
  - Employment (NOIMS, Census, and Annual business survey: Working age, Unemployment by age and gender, Employment by industry, Year-on-year change in unemployment, NEETs, Jobseekers,)
  - Deprivation (NOIMS, IMD, Census, and School survey: Child and working tax credits, Free school meals, IMD scores, DLA claimants, Benefit claimants)
  - Health care (Health stats: Childhood obesity, Percentage of people with respiratory problems, number of hospital admissions, live expectancy, deaths)
  - Transport (Gov.uk: Road accident instances, Year-on-year change in road accident instances)
  - Crime (Met police: Average monthly total crimes, Year-on-year change in total crimes), etc.

Systems Architecture



Project Deliverables

- Tableau-based solution
  - Enable quick and easy data visualization in an interactive and hierarchical way
  - Engage the community in the planning process
  - View multiple measures based on polygon maps
- SQL Server-based solution
  - Parallel with the Tableau-based solution
  - More powerful functionality
  - Can be linked to Tableau
- A set of APIs

Project Deliverables

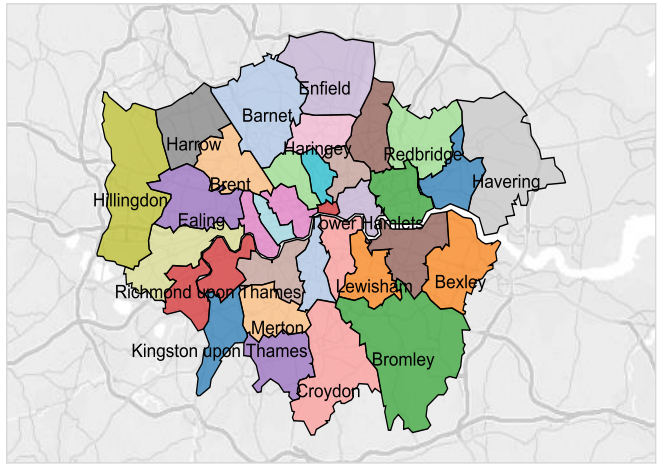
Please select a borugh

- Barking and Dagenham
- Barnet
- Bexley
- Brent

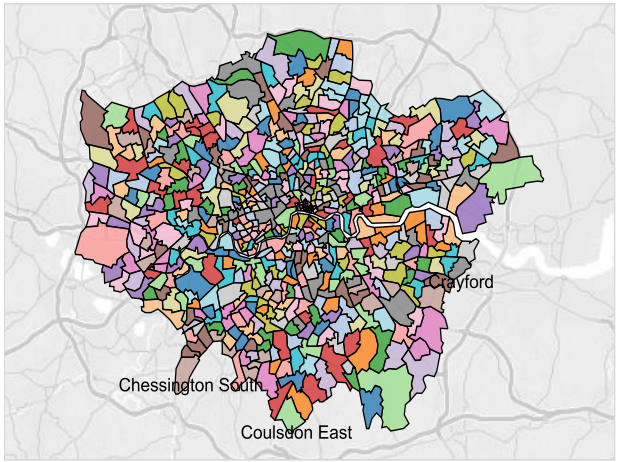
Please select a ward

- Abbey
- Abbey Road
- Abbey Wood
- Abingdon
- Acton Central
- Addiscombe
- Addison
- Aldborough

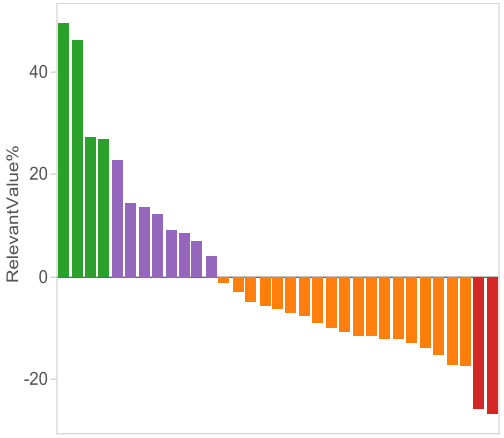
LAD_Map



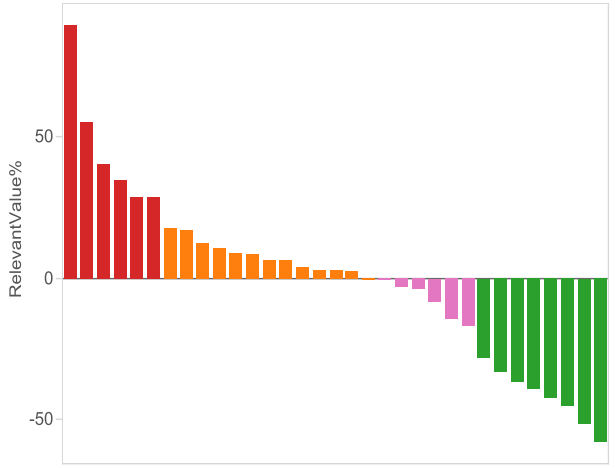
Ward_Map



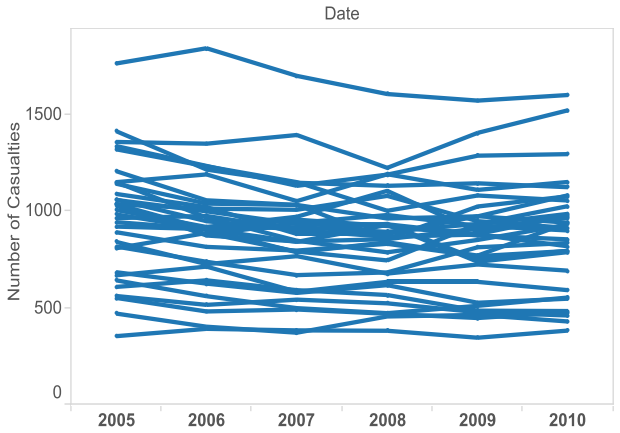
Household_Income in London by Borough (%)



No of Road Casualties in London by Borough (%)



Yearly No of Casualities in London by Borough



Project Deliverables

Please select a borugh

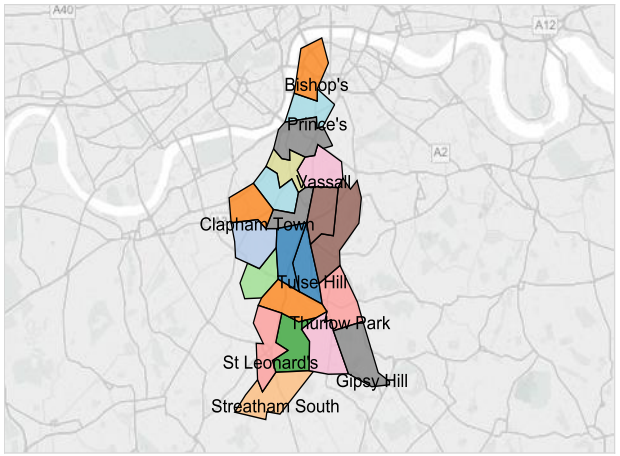
Lambeth

Please select a ward

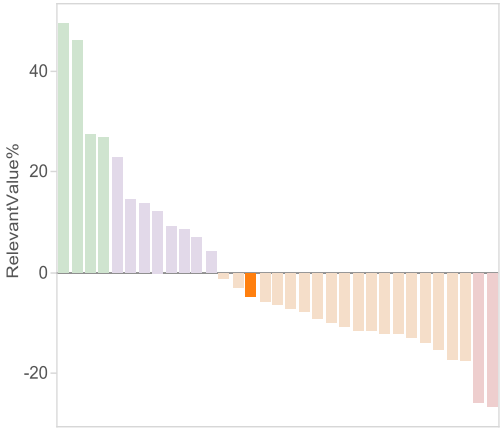
- Bishop's
- Brixton Hill
- Clapham Common
- Clapham Town
- Coldharbour
- Ferndale
- Gipsy Hill
- Herne Hill

LAD_Map

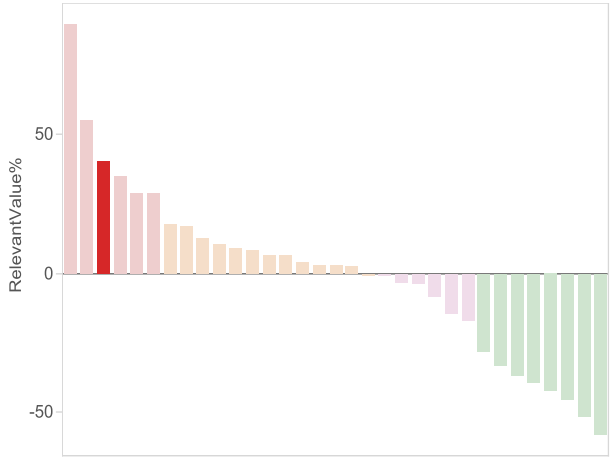
Ward_Map



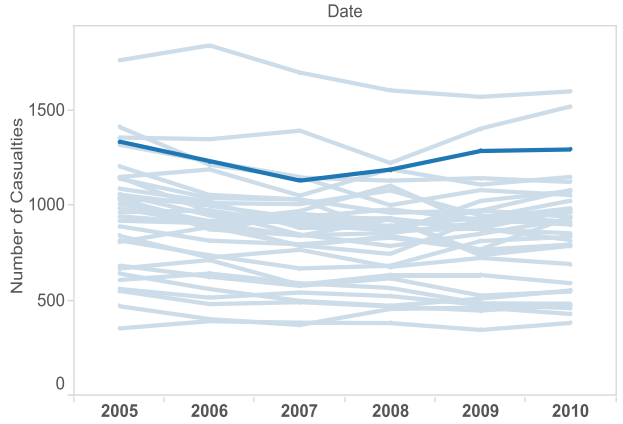
Household_Income in London by Borough (%)



No of Road Casualties in London by Borough (%)

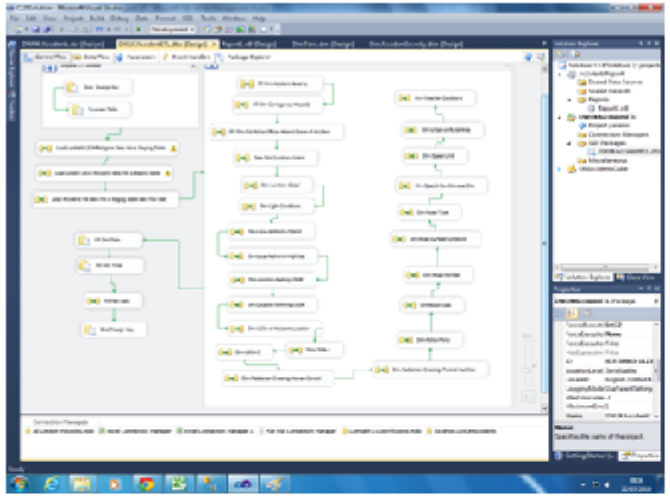


Yearly No of Casualities in London by Borough

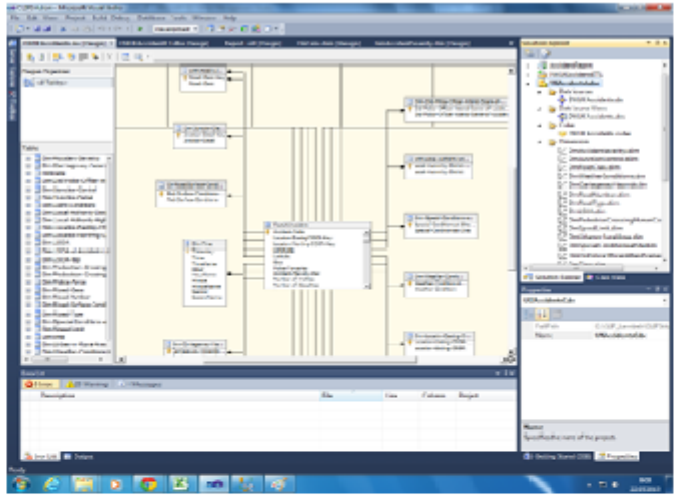


Project Deliverables

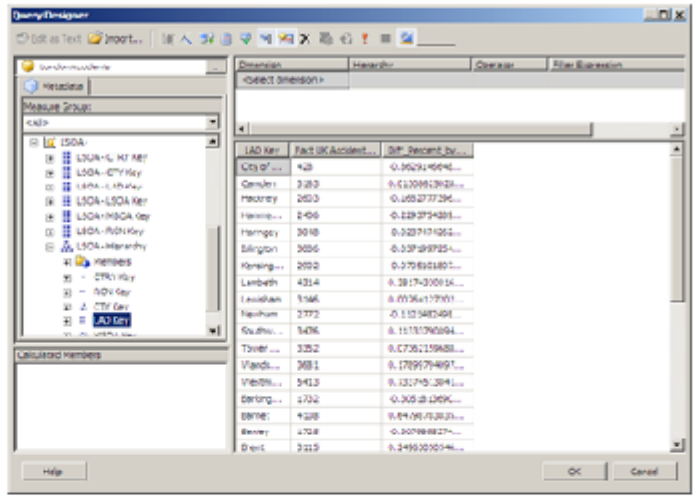
SSIS: ETL Process



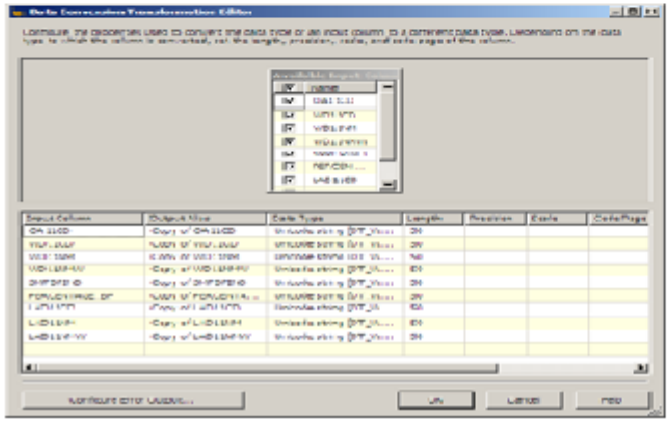
SSAS: Create Cube and Dim



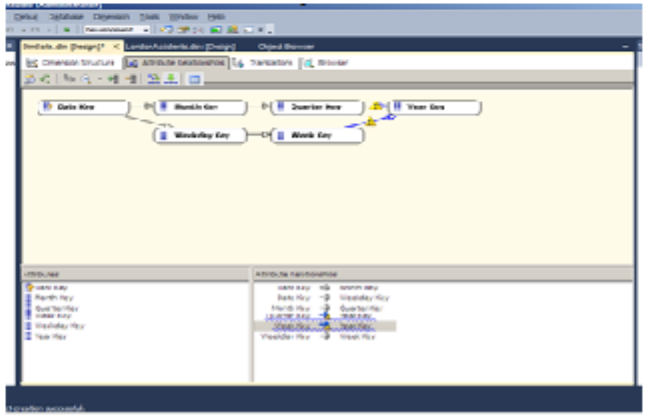
SSRS: Create Report



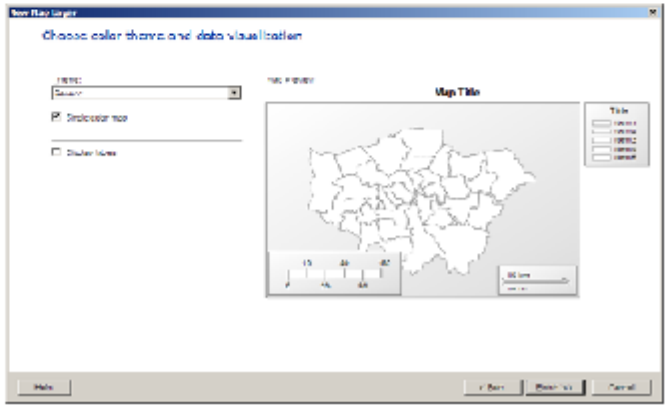
ETL: Data Conversion Transformation



Dimension Hierarchy

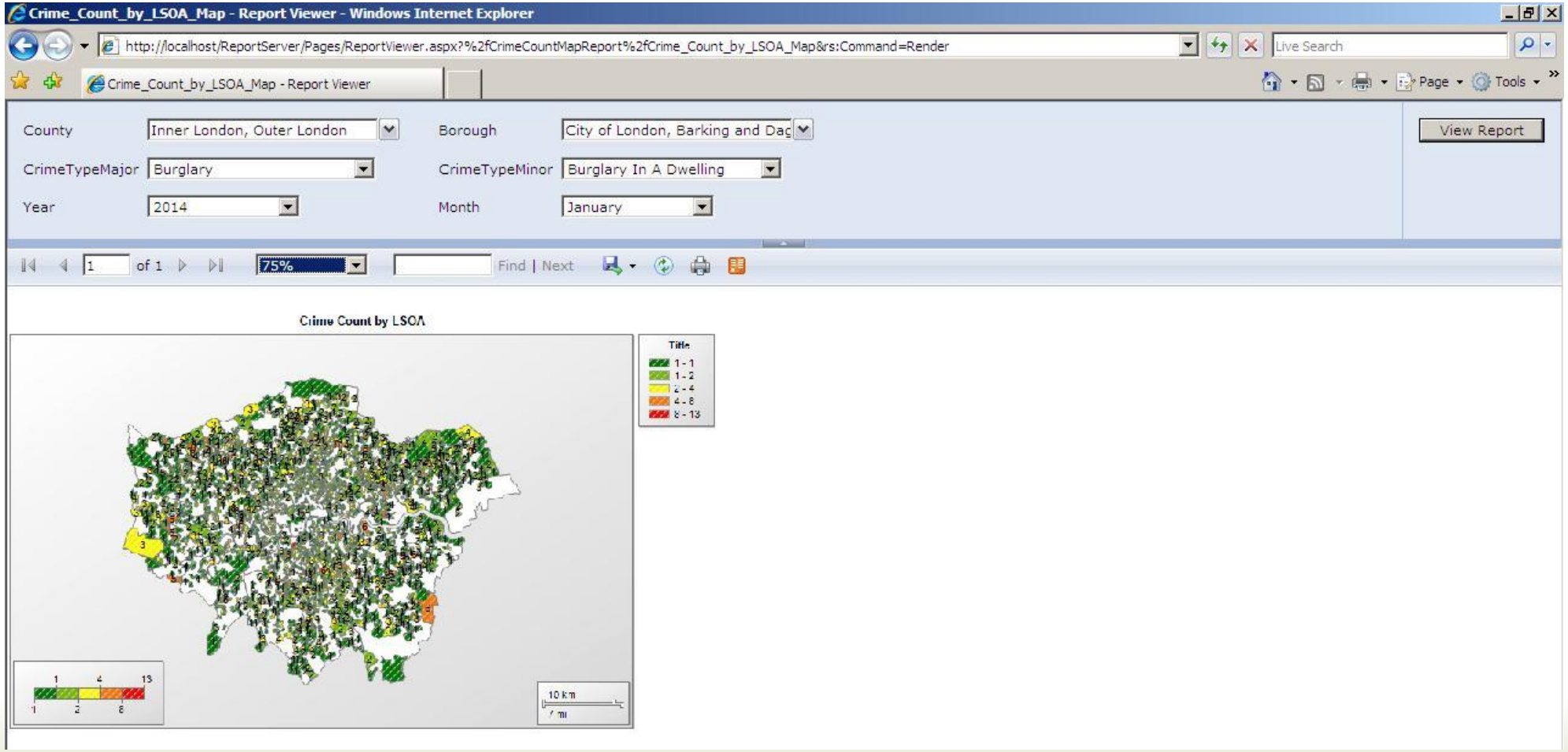


Load Spatial Data



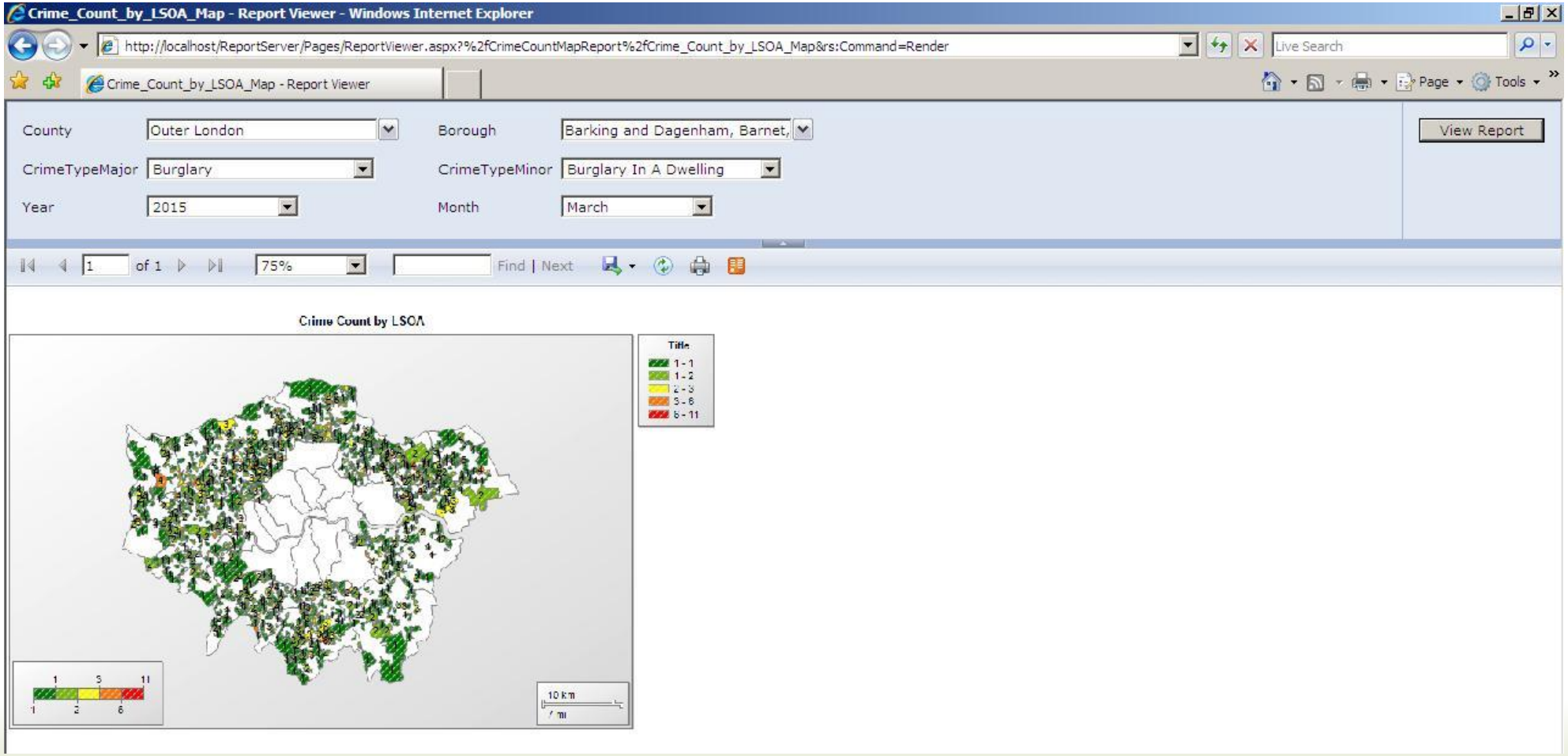
Project Deliverables

Crime count of London



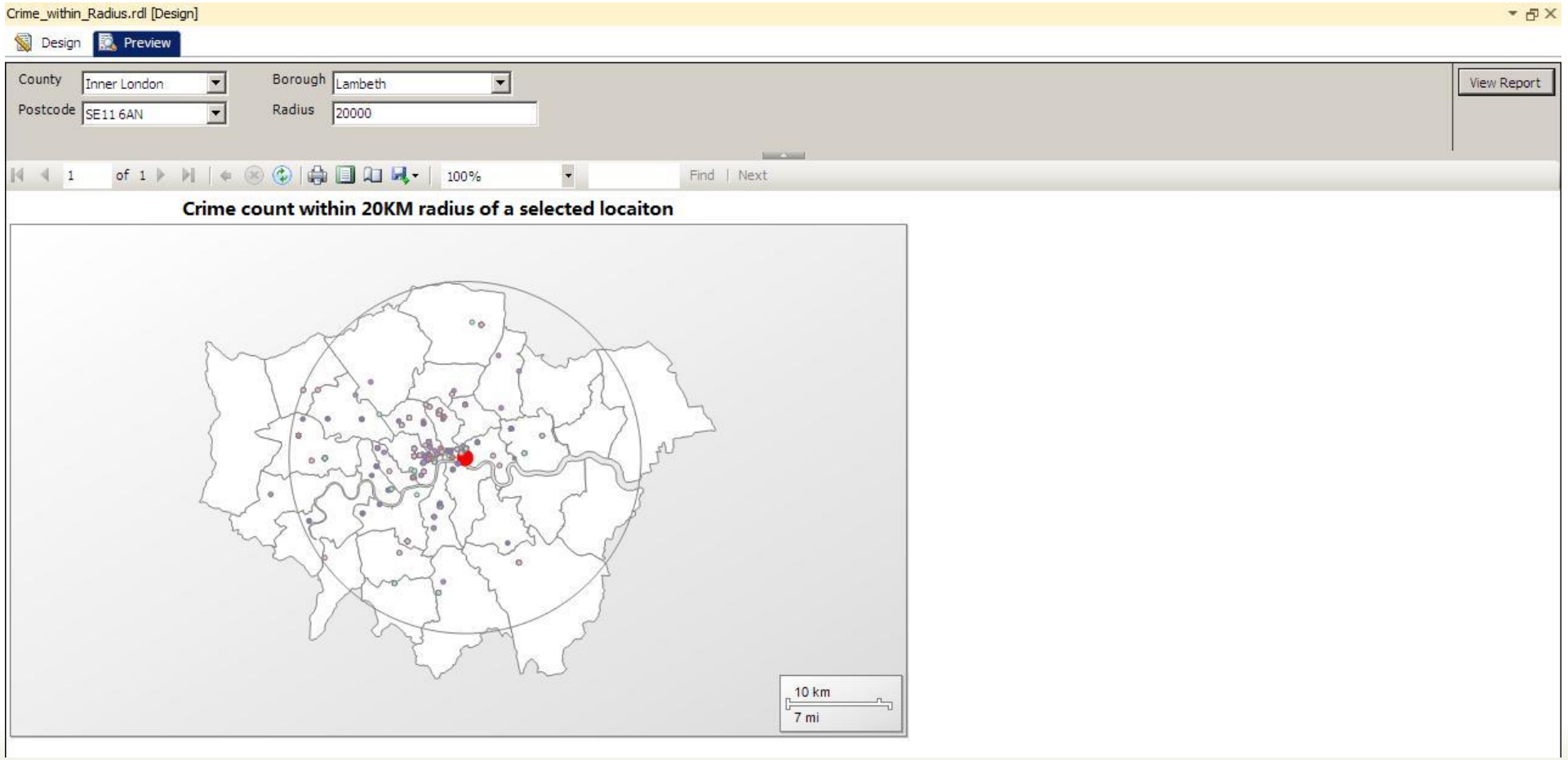
Project Deliverables

Crime count of outer London



Project Deliverables

Crime count within 20KM radius of a selected location



Key Lessons

- Appropriate Infrastructure and architecture for data pre-processing is essential
- Using APIs to allow automatic update of new data
- Data authority
- Data consistency and accessibility
- A centralised data hub needed
- Involvement of the council

Acknowledgement

- Thanks to the staff from Lambeth council: Trisha Boland, Daro Clark, and Tom Brown.
- Thanks to my team: Dr Bo Li, Isakh Weheliye, Brahim El Boudani, Teng Ma, Elias Borchani, Romain Vignaux, and Sean Jarkas.
- Thanks to Howard Thomas, the University Research, Enterprise and Innovation at London South Bank University.

Questions & Comments