

# Chicago Crime Reformatter

André Duarte

**LexisNexis Risk Solutions**

September 1, 2017

## Introduction:

# The assignment

The objective of the exercise was to apply ECL as a data cleaning and data analysis tool using the Chicago Crime Database from 2001 to present.

- Clean and format a dataset into an existing database structure
- Perform data analysis to answer questions

## Reformatter exercise

# From raw to clean layout

```
EXPORT lRaw := RECORD
```

```
  STRING ID;
  STRING Case_Number;
  STRING Date;
  STRING Block;
  STRING IUCR;
  STRING Primary_Type;
  STRING Description;
  STRING Location_Description;
  STRING Arrest;
  STRING Domestic;
  STRING Beat;
  STRING District;
  STRING Ward;
  STRING Community_Area;
  STRING FBI_Code;
  STRING X_Coordinate;
  STRING Y_Coordinate;
  STRING Year;
  STRING Updated_On;
  STRING Latitude;
  STRING Longitude;
  STRING Location;
END;
```

```
EXPORT lClean := RECORD
```

```
  UNSIGNED Row_ID; New
  UNSIGNED ID;
  STRING Case_Number;
  UNSIGNED4 Date_Crime;
  UNSIGNED4 Time_Crime;
  UNSIGNED Address_Number; Loss of info
  STRING Address_Street;
  STRING Address_City;
  STRING2 Address_State; New
  STRING Address_Country;
  UNSIGNED Zip_Code; Unknown from data
  UNSIGNED IUCR; Loss of info
  STRING Primary_Type;
  STRING Description;
  STRING Location_Description;
  BOOLEAN Flag_Arrested;
  BOOLEAN Flag_Domestic;
  UNSIGNED Beat;
  UNSIGNED District;
  UNSIGNED Ward;
  UNSIGNED Community_Area;
  UNSIGNED FBI_Code; Loss of info
  UNSIGNED X_Coordinate;
  UNSIGNED Y_Coordinate;
  UNSIGNED Date_Crime_Year;
  UNSIGNED Date_Crime_Month;
  UNSIGNED Date_Crime_Day;
  UNSIGNED Time_Crime_Hour;
  STRING Date_Crime_Season;
  UNSIGNED4 Date_Updated;
  UNSIGNED4 Time_Updated;
  DECIMAL15_12 Latitude;
  DECIMAL15_12 Longitude;
  STRING Location;
END;
```

# Cleaning and transforming the data

- Macros to parse date and time: return a module to have dot-access to many fields (year, month, day, ...)
- Used COUNTER in the TRANSFORM to create unique IDs
- Inconsistencies, missing values, odd value types:
  - Address number anonymized → loss of info
  - IUCR, FBI codes may contain letters → loss of info
  - Time as hhmmss may become hmmss as UNSIGNED
  - Zip code, victim gender, victim name not present in the dataset

# Summary

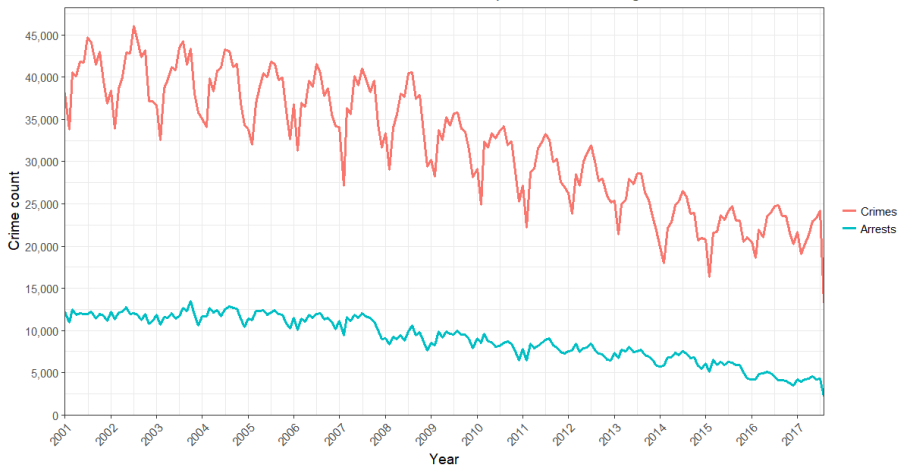
- Fairly straight-forward task
- Need to be careful when defining the cleaned value types
- Use of TRANSFORM, COUNTER, PROJECT, FUNCTIONMACRO, STD.Date, BWR files

# Crime data analysis

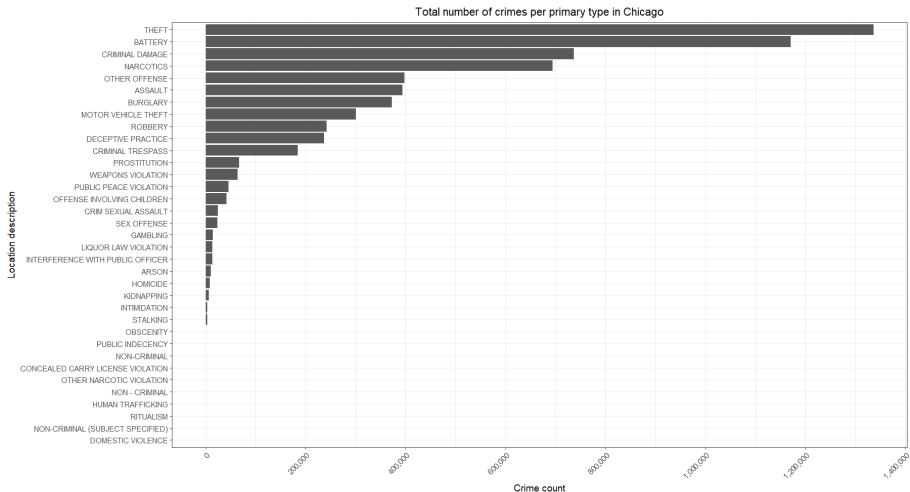


# Evolution of crime in Chicago since 2001

Total number of crimes and arrests per month in Chicago

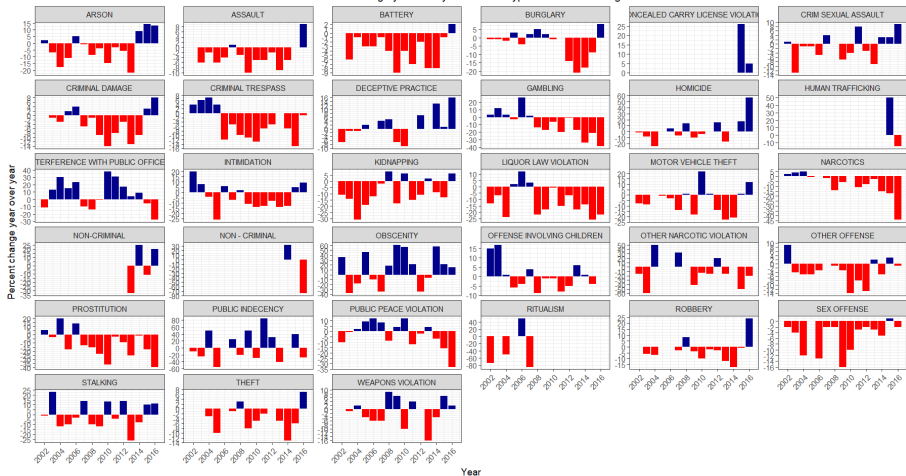


# Most common crimes



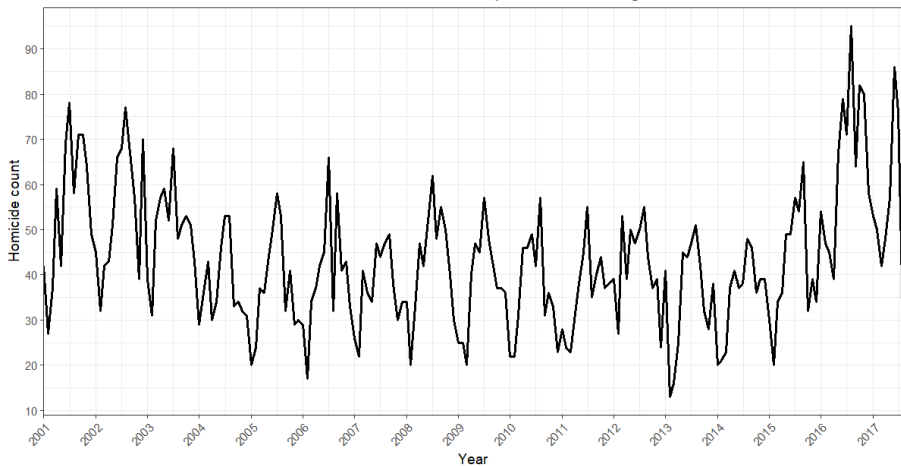
# Most common crimes - year over year

Percent change year over year for each type of crime in Chicago



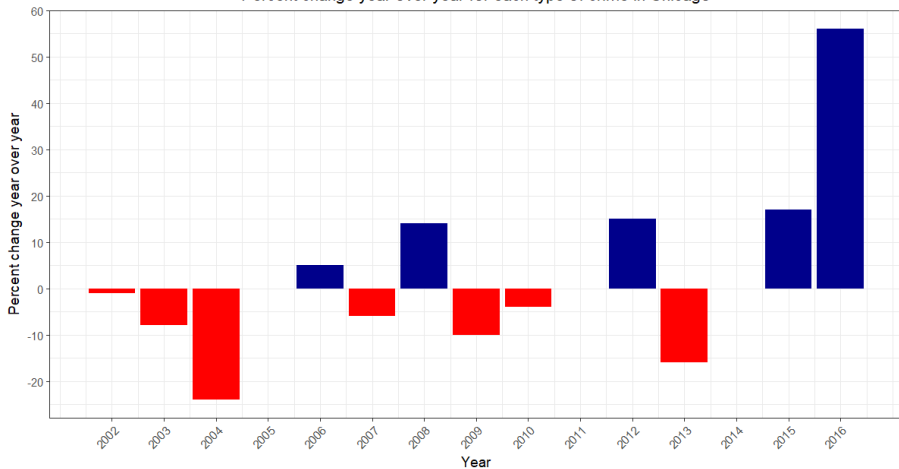
# Homicides

Total number of homicides per month in Chicago



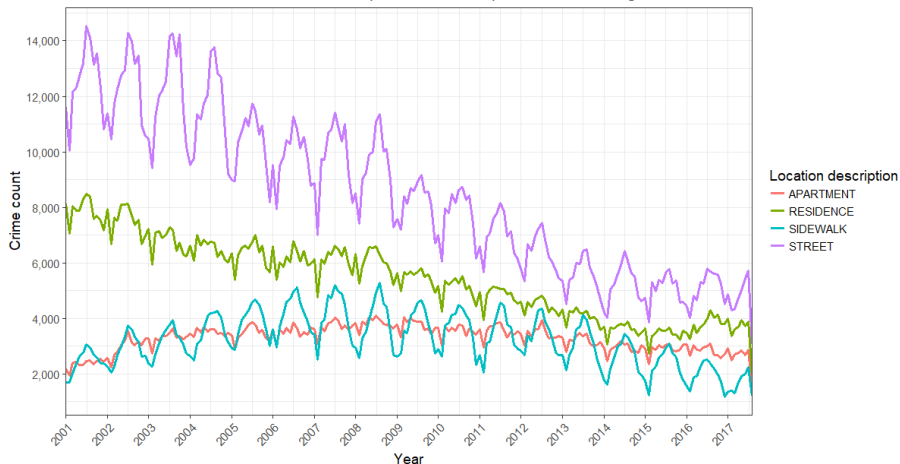
# Homicides

Percent change year over year for each type of crime in Chicago



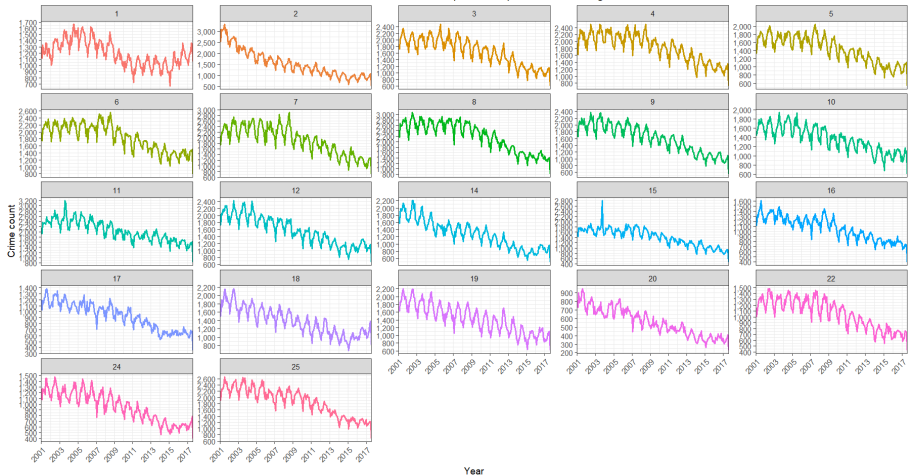
# Where crime happens - locations

Total number of crimes in specific locations per month in Chicago



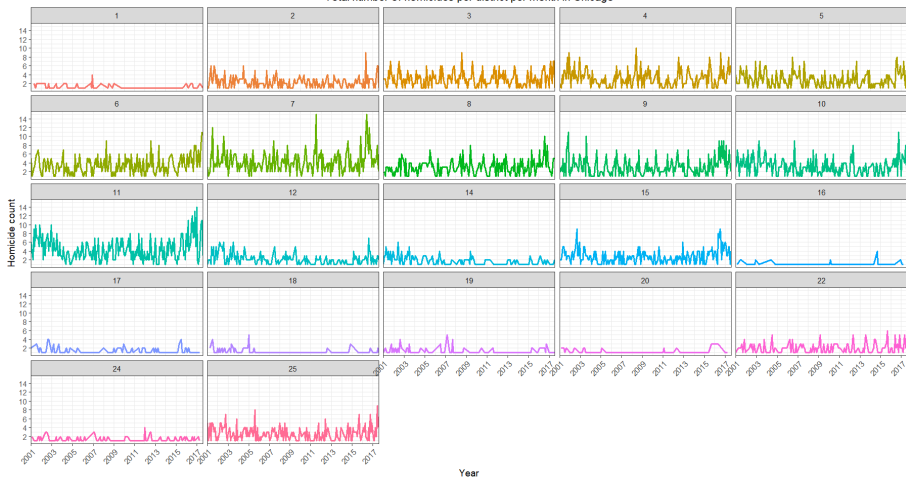
# Where crime happens - districts

Total number of crimes per district per month in Chicago



# Districts on the rise

Total number of homicides per district per month in Chicago





# Summary

- Crime decreasing since 2001, plateau since 2015
- Recent increase in homicides, gun violence
- Rise of gangs and criminal activity to blame
- South side districts are poorer, marginalized, while the northwest and lakeside districts are wealthier, safer

## Conclusion

# Final words

## Reformatter

- Import data from one source into existing layouts
- May have to input some information/take it from elsewhere
- Usually not very complicated in ECL

## Data analysis

- Open-ended task: many possible analyses
- ECL macros and functions are very helpful
- Export tables to another program for plotting

Thank you

# References

- ECL: <https://learn.lexisnexis.com>
- Data: <https://data.cityofchicago.org>
- <http://www.newsweek.com/2016/12/23/chicago-gangs-violence-murder-rate-532034.html>
- <http://www.chicagotribune.com/news/data/ct-shooting-victims-map-charts-htmlstory.html>