

# Analisi dei crimini a Chicago dal 2001 ad oggi



di: *Del Nero Claudio*  
*Goggia Alessio*  
gruppo: *Sparkiamoci in due*



# Caso di studio

Tra gli open data dell' US city open data census abbiamo identificato la città di Chicago, conosciuta come la città di Al Capone ai tempi del proibizionismo.

Nel 2011 più omicidi di Baghdad (4 volte più numerosi rispetto a New York e 2 volte più di Los Angeles!)



# Il dataset

ID	Case Number	Date ↓	Block	IUCR	Primary Type	Description
11002920	JA330961	07/01/2017 11:59:00 PM	075XX S ESSEX AVE	502P	OTHER OFFENSE	FALSE/STOLEN/ALTERED TRP
11002900	JA330936	07/01/2017 11:50:00 PM	063XX N FAIRFIELD AVE	1310	CRIMINAL DAMAGE	TO PROPERTY
11002933	JA330960	07/01/2017 11:50:00 PM	091XX S LOWE AVE	1811	NARCOTICS	POSS: CANNABIS 30GMS OR LE
11002991	JA330958	07/01/2017 11:50:00 PM	114XX S PRINCETON AVE	143A	WEAPONS VIOLATION	UNLAWFUL POSS OF HANDGUN
11003000	JA330939	07/01/2017 11:45:00 PM	063XX S LOWE AVE	031A	ROBBERY	ARMED: HANDGUN

# Il dataset

Location Description	Arrest	Domestic	Beat	District	Ward	Community Area	FBI Code
VEHICLE NON-COMMERCIAL	✓		0421	004	7	43	26
RESIDENCE			2413	024	50	2	14
STREET	✓		2223	022	21	73	18
STREET	✓		0522	005	34	49	15
SIDEWALK			0723	007	16	68	03

# Il dataset

X Coordinate	Y Coordinate	Year	Updated On	Latitude	Longitude
1194173	1855449	2017	07/08/2017 03:49:39 PM	41.758285829	-87.56393345
1156849	1941918	2017	07/08/2017 03:49:39 PM	41.996399909	-87.698378208
1173573	1844368	2017	07/08/2017 03:49:39 PM	41.728358902	-87.639757297
1176329	1829030	2017	07/08/2017 03:49:39 PM	41.686207879	-87.630120372
1173121	1863031	2017	07/08/2017 03:49:39 PM	41.779582501	-87.640862796



# Problema

Come sfruttare tutti questi dati?

What's in this Dataset?

Rows

6.37M

Columns

22

Each row is a

Reported Crime

Alcune domande a cui vorremmo rispondere:

- quali sono i crimini più commessi?
- dove si sono verificati?
- ci sono relazioni tra i tipi di crimini commessi e i luoghi fisici in cui sono accaduti?
- quali sono i distretti più colpiti e i crimini che più li caratterizzano?
- possiamo prevedere se in un distretto avverrà un crimine?
- è possibile fornire una visualizzazione di essi sulla mappa della città?



# Soluzione e tecnologie utilizzate

1) memorizzazione dataset



2) analisi base



3) analisi avanzate



4) visualizzazione dataset







```
> db.crimes.find().pretty()
{
  "_id" : ObjectId("59414b8813100a0fadea18ee"),
  "ID" : 2474022,
  "Case Number" : "HH805953",
  "Date" : "11/27/2002 05:07:00 PM",
  "Block" : "026XX S KEDVALE AVE",
  "IUCR" : 320,
  "Primary Type" : "ROBBERY",
  "Description" : "STRONGARM - NO WEAPON",
  "Location Description" : "SIDEWALK",
  "Arrest" : "false",
  "Domestic" : "false",
  "Beat" : 1031,
  "District" : 10,
  "Ward" : 22,
  "Community Area" : 30,
  "FBI Code" : 3,
  "X Coordinate" : 1149141,
  "Y Coordinate" : 1886156,
  "Year" : 2002,
  "Updated On" : "04/15/2016 08:55:02 AM",
  "Latitude" : 41.843537135,
  "Longitude" : -87.728180227,
  "Location" : "(41.843537135, -87.728180227)"
}

{
  "_id" : ObjectId("59414b8813100a0fadea18ef"),
  "ID" : 2474024,
  "Case Number" : "HH805575",
  "Date" : "11/27/2002 02:29:42 PM",
  "Block" : "076XX S STEWART AVE",
  "IUCR" : "051A",
  "Primary Type" : "ASSAULT",
  "Description" : "AGGRAVATED: HANDGUN",
  "Location Description" : "STREET",
  "Arrest" : "true",
  "Domestic" : "false",
  "Beat" : 623,
  "District" : 6,
  "Ward" : 17,
  "Community Area" : 69,
  "FBI Code" : "04A",
  "X Coordinate" : 1174952,
  "Y Coordinate" : 1854202,
  "Year" : 2002,
  "Updated On" : "04/15/2016 08:55:02 AM",
  "Latitude" : 41.755314051,
  "Longitude" : -87.634413175,
  "Location" : "(41.755314051, -87.634413175)"
}

{
  "_id" : ObjectId("59414b8813100a0fadea18f0"),
  "ID" : 2474026,
  "Case Number" : "HH802466",
  "Date" : "11/25/2002 06:00:00 AM",
  "Block" : "020XX W CULLERTON ST",
  "IUCR" : 325,
  "Primary Type" : "ROBBERY",
  "Description" : "VEHICULAR HIJACKING",
  "Location Description" : "STREET",
  "Arrest" : "false",
  "Domestic" : "false",
  "Beat" : 623,
  "District" : 6,
  "Ward" : 17,
  "Community Area" : 69,
  "FBI Code" : "04A",
  "X Coordinate" : 1174952,
  "Y Coordinate" : 1854202,
  "Year" : 2002,
  "Updated On" : "04/15/2016 08:55:02 AM",
  "Latitude" : 41.755314051,
  "Longitude" : -87.634413175,
  "Location" : "(41.755314051, -87.634413175)"
}
```



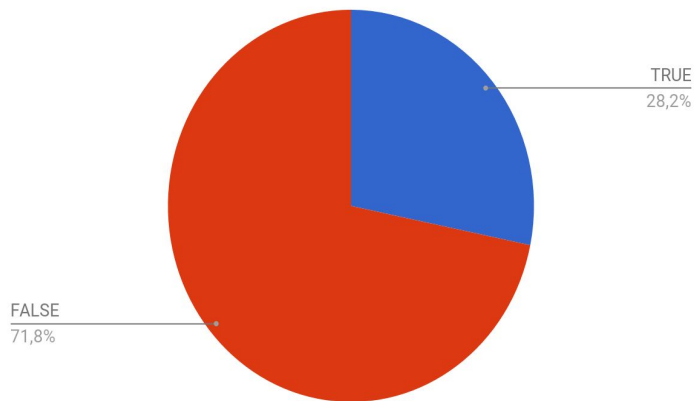
## Fase 2 - analisi base



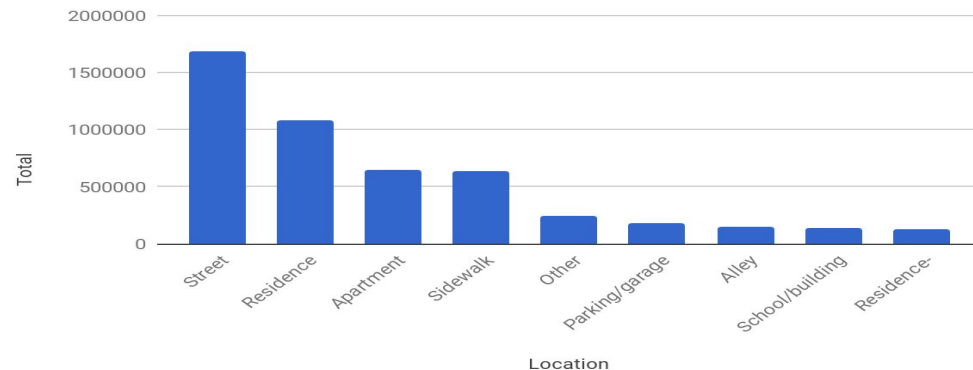


## Fase 2 - analisi base - risultati (1)

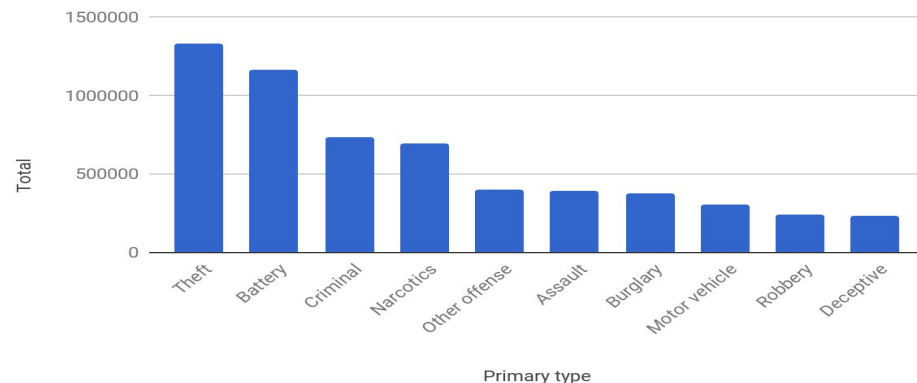
Percentuale arresti/non arresti



Top10 location più colpite

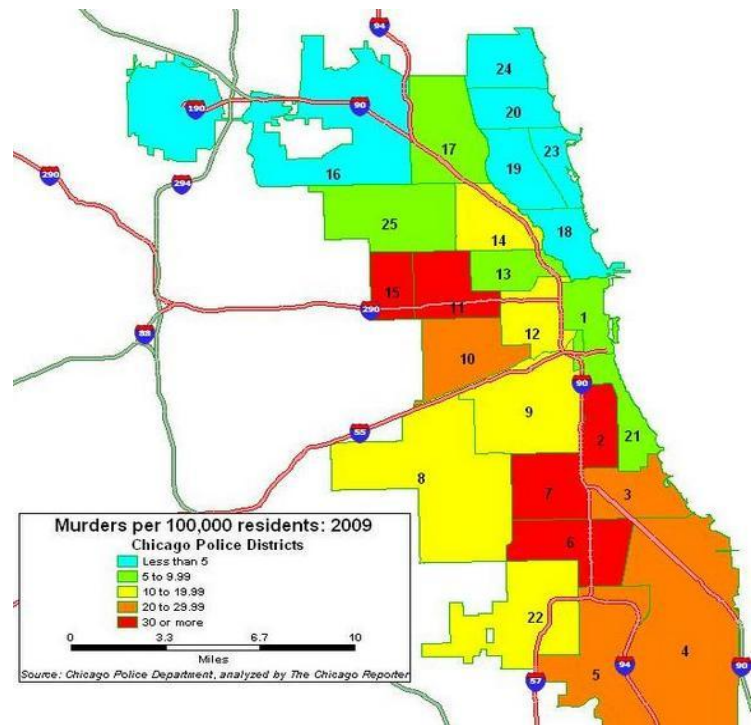
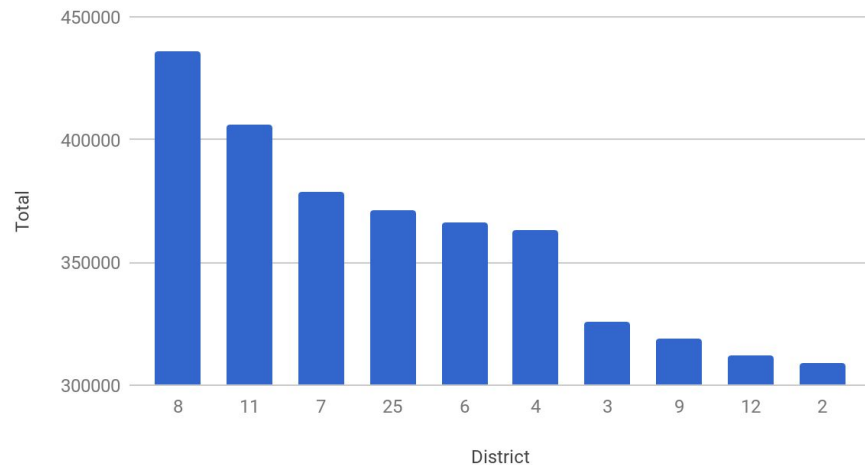


Top10 crimini più commessi



## Fase 2 - analisi base - risultati (2)

Top10 distretti con più crimini





## Fase 2 - analisi base

- crimini più commessi per ogni distretto
- crimini nel corso degli anni
- crimini per ogni mese dell'anno
- luoghi fisici in cui sono avvenuti per ogni distretto

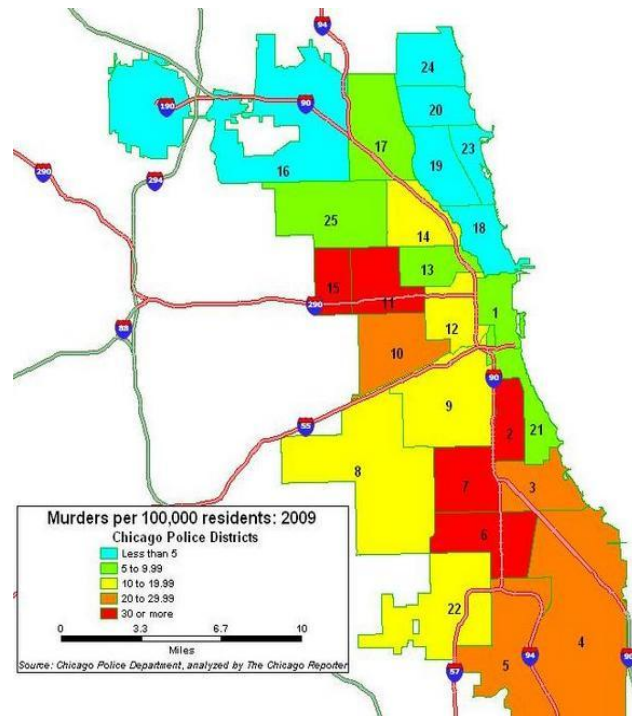
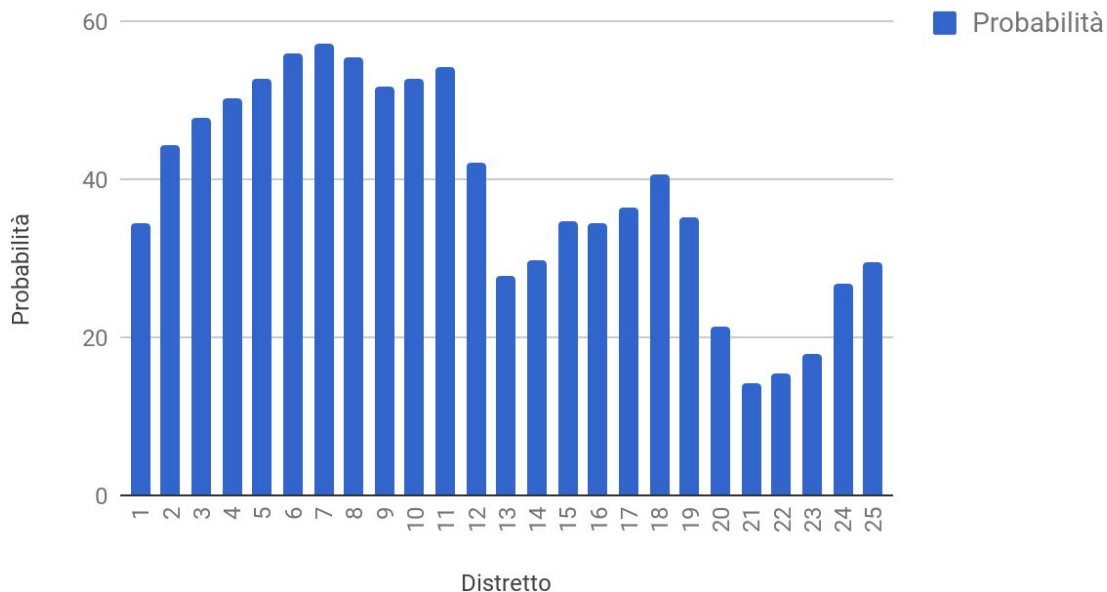


## Fase 3 - analisi avanzate

- kernel density estimation
- testing Kolmogorov-Smirnov

# Fase 3 - kernel density estimation - risultati

Probabilità stimata di crimini nei distretti sulla base dell'anno 2017







## Fase 3 - testing Kolmogorov-Smirnov - risultati

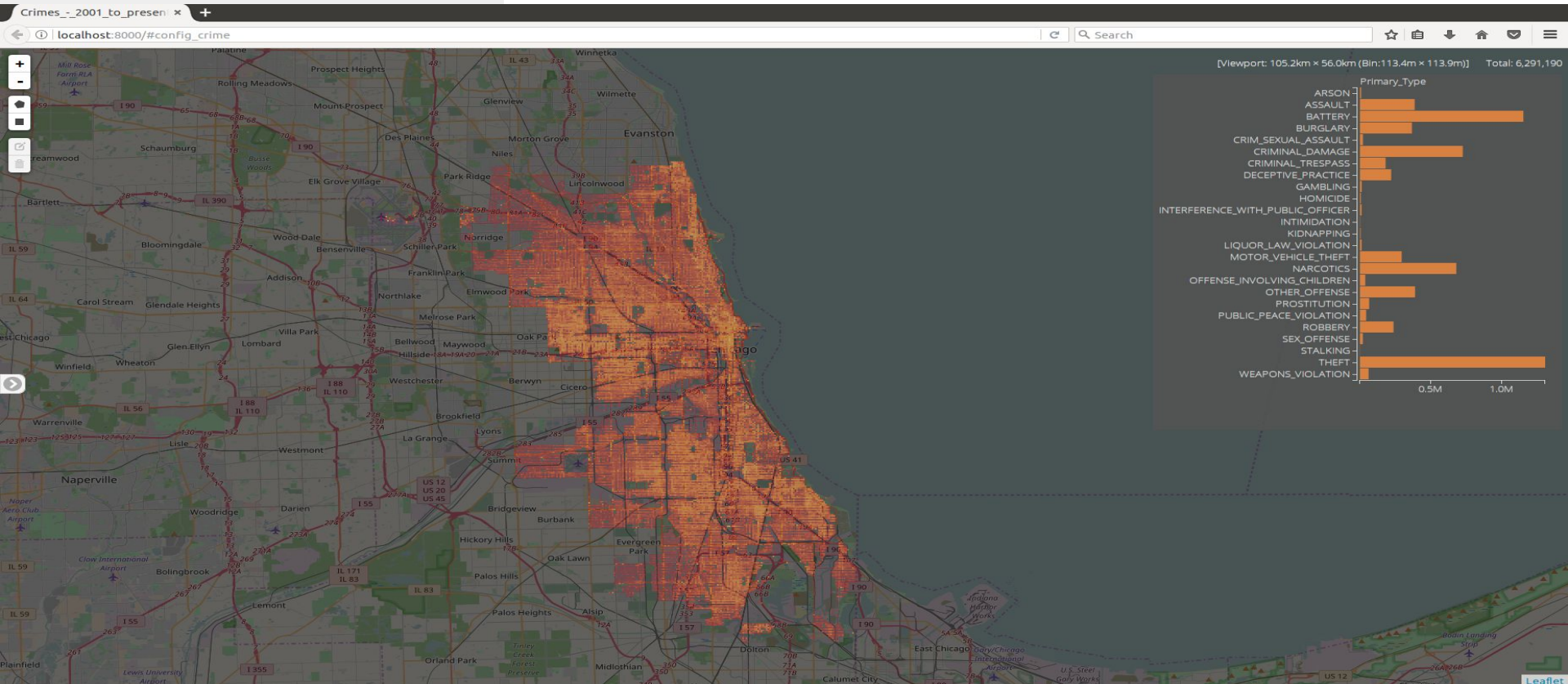
Kolmogorov-Smirnov test summary:

Very strong presumption against null hypothesis

# Nanocubes

Fast visualization of large spatiotemporal datasets

## Fase 4 - visualizzazione

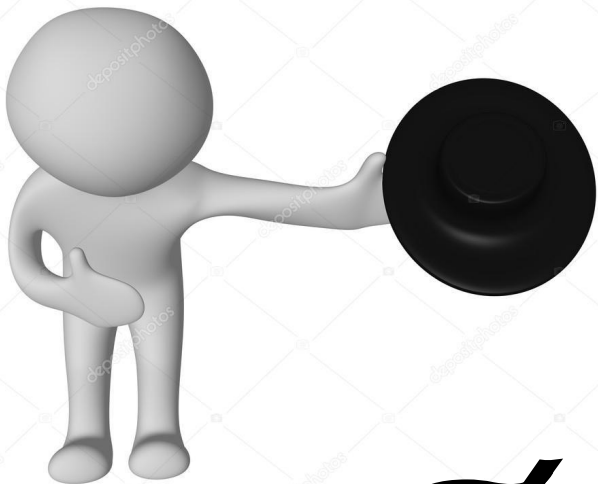


Roma, 17 Luglio 2017

Analisi dei crimini a Chicago dal 2001 ad oggi

Sparkiamoci in due

16



*The end*

