

Università Politecnica delle Marche  
Dipartimento di Ingegneria dell'Informazione  
Facoltà di Ingegneria Informatica e dell'Automazione



**Analisi della criminalità dello stato del Maryland (USA)  
mediante i tools Qlik, Tableau e Power BI**

**Docenti:**  
Domenico URSINO  
Gianluca BONIFAZI

**Studenti:**  
Silvia Ciuffreda  
Luca Liberatore  
Beatrice Moliterno

**ANNO ACCADEMICO 2022/2023**

# Indice

<b>1</b>	<b>Introduzione</b>	<b>4</b>
<b>2</b>	<b>ETL</b>	<b>5</b>
2.1	Il dataset . . . . .	5
2.2	Le operazioni eseguite . . . . .	6
<b>3</b>	<b>Qlik</b>	<b>8</b>
3.1	Caricamento del dataset . . . . .	8
3.2	Le analisi effettuate . . . . .	8
3.2.1	La differenza tra reati e crimini . . . . .	8
3.3	L'analisi Temporale dei reati . . . . .	9
3.4	L'analisi Spaziale dei reati . . . . .	10
3.5	Le vittime coinvolte nei reati . . . . .	16
<b>4</b>	<b>Tableau</b>	<b>18</b>
4.1	Dataset e Caricamento Dati . . . . .	18
4.2	Analisi Effettuate . . . . .	18
4.2.1	Panoramica generale . . . . .	18
4.2.2	Tipologie di crimine . . . . .	20
4.2.3	Crimini contro il patrimonio . . . . .	21
4.2.4	Crimini contro la persona . . . . .	25
4.2.5	Crimini contro la società . . . . .	26
4.2.6	Forecasting . . . . .	27
<b>5</b>	<b>Power BI</b>	<b>29</b>
5.1	Caricamento del dataset . . . . .	29
5.2	Le forze dell'ordine dello stato del Maryland . . . . .	29
5.2.1	Le aree gestite dalle forze dell'ordine . . . . .	30
5.3	L'analisi dei crimini legati alle sostanze stupefacenti . . . . .	33
5.3.1	Le sostanze stupefacenti possedute . . . . .	34
5.3.2	Le sostanze stupefacenti vendute . . . . .	37

## Elenco delle figure

1	I valori NULL . . . . .	6
2	Alcuni esempi degli elementi della colonna "Place" . . . . .	7
3	Alcuni esempi degli elementi della colonna "filtered_place" . . . . .	7
4	Logo Qlik . . . . .	8
5	Caricamento del dataset in Qlik . . . . .	8
6	Un esempio di più crimini associati allo stesso reato . . . . .	9
7	Il numero di crimini . . . . .	9
8	Il numero di reati . . . . .	9
9	La funzione per calcolare i crimini . . . . .	9
10	La funzione per calcolare i reati . . . . .	9
11	L'analisi dei reati per anno . . . . .	10
12	L'analisi dei reati per mese . . . . .	10
13	La cartina dello stato del Maryland . . . . .	11
14	L'analisi delle città in cui sono avvenuti i reati . . . . .	11
15	La cartina della contea del Montgomery . . . . .	12
16	L'analisi delle città in cui sono avvenuti i reati nel 2017 . . . . .	12
17	L'analisi delle città in cui sono avvenuti i reati nel 2021 . . . . .	13
18	I principali luoghi dei reati . . . . .	13
19	I principali luoghi dei reati più nel dettaglio . . . . .	14
20	I luoghi maggiormente coinvolti nelle città principali . . . . .	14
21	I luoghi dettagliati che sono stati maggiormente coinvolti nelle città di Silver Spring, Gaithersburg e Rockville . . . . .	15
22	Un campione di elementi del dataset appartenenti a tre tipologie di crimine differente . . . . .	16
23	Un campione di elementi del dataset appartenenti alla tipologia di crimine contro la persona, in cui uno stesso reato è costituito da due crimini . . . . .	16
24	Le vittime dei reati . . . . .	16
25	I reati contro la persona con più vittime . . . . .	17
26	La funzione per calcolare il numero totale di vittime dei reati . . . . .	17
27	Tableau . . . . .	18
28	Tableau - Panoramica generale . . . . .	19
29	Tableau - Crimini contro la persona . . . . .	20
30	Tableau - Tipologie di crimini per anno . . . . .	21
31	Tableau - Formula campo calcolato per fasce orarie . . . . .	22
32	Tableau - Crimini contro il patrimonio . . . . .	23
33	Tableau - Crimini contro il patrimonio, Luoghi . . . . .	24
34	Tableau - Mappa Silver Spring Crimini contro il patrimonio . . . . .	25
35	Tableau - Crimini contro la persona . . . . .	26
36	Tableau - Luoghi dei Crimini contro la persona . . . . .	26
37	Tableau - Crimini contro la società . . . . .	27
38	Tableau - Forecasting . . . . .	28
39	Tableau - Crimini Totali per Mese . . . . .	28
40	Logo Power BI . . . . .	29
41	Le Agenzie di polizia . . . . .	29
42	La mappa geografica delle aree di polizia . . . . .	30
43	I distretti in cui è intervenuta la MCPD . . . . .	31
44	I distretti in cui è intervenuta la RCPD, GPD e TPPD . . . . .	31
45	I quartieri in cui è intervenuta maggiormente MCPD, RCPD, GPD e TPPD . . . . .	32
46	L'area in cui è intervenuta maggiormente MCPD . . . . .	32
47	L'area in cui sono intervenute maggiormente GPD, RCPD e TPPD . . . . .	33
48	I crimini legati alle sostanze stupefacenti . . . . .	34
49	Le sostanze stupefacenti possedute . . . . .	35
50	La marijuana posseduta . . . . .	36
51	La previsione del possesso di sostanze stupefacenti . . . . .	36
52	Le sostanze stupefacenti vendute . . . . .	37

53	Le Agenzie di polizia che si sono occupate del crimine legato alla vendita di Marijuana . . . . .	38
54	I luoghi in cui è stata venduta la Marijuana . . . . .	38
55	La previsione della vendita di Marijuana . . . . .	39

## **1 Introduzione**

Il presente progetto si prospetta di analizzare i dati relativi ai crimini commessi nel Maryland, reperibili al seguente link:

<https://www.kaggle.com/datasets/jgiigii/uscrimesdataset>

In particolare sono state svolte analisi relative a diverse tematiche legate ai suddetti dati. Gli strumenti utilizzati sono tre importanti software della Business Intelligent: Qlik, Tableau e Power Bi.

## 2 ETL

L'ETL ("Extract Trasform Load") è stata la fase preliminare del progetto; questo processo consente di rendere disponibili i dati nel modo più efficiente possibile, affinché risultino perfettamente fruibili dalle applicazioni di analytics utilizzate per generare valore informativo a partire dai dati. La seguente fase è stata svolta in Python, sfruttando l'ambiente di sviluppo Jupyter Notebook.

### 2.1 Il dataset

Il dataset analizzato raccoglie tutti i dati riguardanti i crimini commessi nello stato del Maryland (USA) ed è salvato in un file con formato CSV. Nella sua versione originaria questo è composto da 306'094 righe e 35 campi, i quali sono elencati di seguito:

- *Incident ID*, il codice di identificazione del crimine;
- *Offence Code*, il codice che riconduce al tipo di crimine commesso;
- *CR\_number*, il numero associato al "Criminal Case" assegnati dal tribunale ;
- *Start\_Date\_Time*, la data e l'ora in cui è stato registrato l'inizio del crimine;
- *Start\_Date\_Time*, la data e l'ora in cui è stata registrata la fine del crimine;
- *Dispatch Date / Time*, la data e l'orario in cui sono stati inviate gli agenti di polizia;
- *NIBRS Code*, il codice descrittivo della tipologia di crimine commesso;
- *Crime Name 1*, la categoria di crimine commesso;
- *Crime Name 2*, la descrizione più dettagliata della categoria del crimine descritta in Crime Name 1;
- *Crime Name 3*, la descrizione più dettagliata della categoria del crimine descritta in Crime Name 2;
- *Victims*, il numero di vittime coinvolte nel crimine;
- *Year*, l'anno in cui è stato commesso il crimine;
- *Month*, il mese in cui è stato commesso il crimine;
- *Year-Month*, l'anno e il mese in cui è stato commesso il crimine;
- *Day*, il giorno in cui è stato commesso il crimine;
- *Committed\_At\_Morning*, un valore booleano che definisce se il crimine è stato commesso la mattina o meno;
- *State*, lo stato in cui il crimine è stato commesso;
- *City*, la città in cui il crimine è stato commesso;
- *Place*, il luogo in cui è avvenuto il crimine;
- *Block Address*, l'indirizzo in cui il crimine è stato commesso;
- *Address Number*, il numero dell'indirizzo in cui è avvenuto il crimine;
- *Street Prefix*, il prefisso della strada in cui è avvenuto il crimine;
- *Street Name*, il nome della strada in cui è avvenuto il crimine;
- *Street Suffix*, il suffisso della strada in cui è avvenuto il crimine;
- *Street Type*, la tipologia di strada in cui è avvenuto il crimine;

## 2.2 Le operazioni eseguite

- *Zip Code*, il codice postale associato al luogo in cui è avvenuto il crimine;
- *Latitude*, la latitudine corrispondente al punto esatto in cui è stato commesso il crimine;
- *Longitude*, la longitudine corrispondente al punto esatto in cui è stato commesso il crimine;
- *Location*, le coordinate geografiche che comprendono la latitudine e la longitudine.
- *Agency*, l'agenzia di polizia intervenuta;
- *Police District Name*, il nome del distretto di polizia in cui è stato commesso il crimine;
- *Police District Number*, il numero corrispondente al distretto di polizia;
- *Sector*, un codice che rappresenta il sottoinsieme del distretto di polizia;
- *Beat*, un codice che rappresenta un sottoinsieme del settore;
- *PRA*, "Police Response Area", un codice che rappresenta un sottoinsieme del Beat;

### 2.2 Le operazioni eseguite

Innanzitutto è stato caricato il dataset su Jupyter notebook e sono stati valutati i valori NULL presenti; come mostrato in Figura 1, i campi con il maggior numero di valori NULL sono "Dispatch Date / Time", "Block Address", "Zip Code", "Address Number", "Street Prefix", "Street Suffix" e "End\_Date\_Time". Dunque, queste colonne sono state interamente eliminate, sia per il poco valore informativo che per la loro poca utilità nelle analisi da effettuare. Inoltre, sono state eliminate tutte le righe con almeno un valore nullo.

Incident ID	0
Offence Code	0
CR Number	0
Dispatch Date / Time	49029
NIBRS Code	0
Victims	0
Crime Name1	272
Crime Name2	272
Crime Name3	272
Police District Name	94
Block Address	26206
City	1276
State	0
Zip Code	3179
Agency	0
Place	0
Sector	1530
Beat	1530
PRA	239
Address Number	26109
Street Prefix	292463
Street Name	1
Street Suffix	300662
Street Type	339
Start_Date_Time	0
End_Date_Time	161658
Latitude	0
Longitude	0
Police District Number	0
Location	0
Year	0
Month	0
Year-Month	0
Day	0
Committed_At_Morning	0

Figura 1: I valori NULL

## 2.2 Le operazioni eseguite

I dati raggruppano tutti i crimini che sono stati commessi dalla metà del 2016 alla metà del 2022. Affinchè le analisi fossero comparabili anno per anno, si è ritenuto opportuno eliminare tutti gli anni che risultavano incompleti, ovvero il 2016 e il 2022.

In seguito ad una prima visione generale del dataset, si è notato che nella colonna "Place", per ogni crimine, è riportato sia il luogo generale, che la sua particolarizzazione, separati da un trattino, come è possibile visualizzare nella Figura 2. Per effettuare analisi più approfondite, si è voluto creare una nuova colonna chiamata "filtered\_place" che raggruppasse solo il luogo generale, come riportato in Figura 3.

```
'Parking Lot - Residential',
'Parking Lot - Commercial',
'School/College - DO NOT USE',
'Street - Residential',
'Street - Bus Stop',
'Retail - Mall',
'Retail - Drug Store/Pharmacy',
```

Figura 2: Alcuni esempi degli elementi della colonna "Place"

```
retail
school/college
parkinggarage
Restaurant
Grocery/Supermarket
```

Figura 3: Alcuni esempi degli elementi della colonna "filtered\_place"

Concludendo, il dataset sottoposto alle analisi descritte nei capitoli successivi è costituito da 248'251 righe e 29 colonne.

### 3 Qlik

Qlik Sense è una piattaforma di analisi che, tramite dashboard e grafi personalizzati, permette di effettuare analisi di tipo descrittivo e diagnostico all'interno di quel paradigma conoscitivo che è la Business Intelligence.

Nello svolgimento del progetto si è utilizzata la versione Cloud del software, sfruttando la possibilità che mette a disposizione l'applicativo di creare spazi di lavoro condivisi ed interattivi. In termini tecnici, ci si è avvalsi della capacità dello strumento di svolgere analisi coordinate all'interno della stessa "App", data la possibilità di lavorare su "Fogli" condivisi, con l'obiettivo di operare all'interno di quella fetta di conoscenza che si occupa di Data Science.



Figura 4: Logo Qlik

#### 3.1 Caricamento del dataset

Il loading del dataset all'interno della piattaforma Qlik Cloud è avvenuto all'interno della sezione "Gestione dati" tramite il caricamento del file csv nello spazio del "Catalogo". Dato che il file sorgente è costituito da una sola tabella complessiva il dataset compare come rappresentato nella Figura 5.



Figura 5: Caricamento del dataset in Qlik

#### 3.2 Le analisi effettuate

Le analisi svolte in Qlik sono volte ad ottenere una panoramica generale rispetto i dati a disposizione. In particolare, in tale sezione, è stata effettuata un'analisi riguardo l'andamento temporale, i luoghi e le vittime coinvolte.

##### 3.2.1 La differenza tra reati e crimini

Nei primi approcci al dataset si è subito notato che il codice che identifica il crimine (*Incident ID*) può ripetersi più volte in diverse righe; per tale ragione è stata adottata una nomenclatura che prendesse in considerazione questo aspetto. È stato definito "*Reato*", l'evento accaduto ad una certa data e ora, in un determinato luogo, mentre "*Crimine*", la singola azione che ha violato le leggi in seguito all'evento stesso. Per l'appunto uno stesso reato può interessare crimini multipli differenti. Per comprendere meglio il concetto è stato riportato un esempio visibile in Figura 6. Come si può osservare, l'*Incident ID* di riferimento (201236396) è ripetuto in due righe del dataset, ma i crimini, sebbene registrati nello stesso istante temporale e nello stesso luogo, sono differenti. Infatti, nel caso specifico, andando ad osservare i valori dell'attributo *Crime Name 3*, i crimini in questione sono: "DRUGS - MARIJUANA - POSSESS" e "LIQUOR - POSSESS", i quali corrispondono rispettivamente a due *Offence Code* distinti, che, per

### 3.3 L'analisi Temporale dei reati

l'appunto, identificano la tipologia di crimine. Dunque, l'evento risulta essere unico, ma le azioni illecite sono due. E' importante specificare che questa terminologia è stata adottata, non solo in Qlik, ma anche in Tableau e Power BI.

Incident ID	Offence Code	Crime Name1	Crime Name2	Crime Name3	Start_Date_Time	Latitude	Longitude
201236393	3562	Crime Against Society	Drug/Narcotic Violations	DRUGS - MARIJUANA - POSSESS	04/23/2019 06:07:00 AM	39.086692243	-77.07876594
201236393	4104	Crime Against Society	Liquor Law Violations	LIQUOR - POSSESS	04/23/2019 06:07:00 AM	39.086692243	-77.07876594

Figura 6: Un esempio di più crimini associati allo stesso reato

Una volta fatta tale distinzione sono stati calcolati il numero di crimini e di reati registrati nel dataset, visualizzati due KPI, illustrati in Figura 7 e 8.



Figura 7: Il numero di crimini



Figura 8: Il numero di reati

I crimini sono stati calcolati tramite un semplice conteggio del numero dei valori del campo *Incident ID*, la cui funzione è riportata in Figura 9; i reati sono stati calcolati allo stesso modo, ma eliminando i duplicati del campo stesso e la funzione è riportata in Figura 10.

**Count([Incident ID])**

Figura 9: La funzione per calcolare i crimini

**Count(distinct[Incident ID])**

Figura 10: La funzione per calcolare i reati

Dunque, dai dati riportati si può dedurre che circa l'8% dei dati corrisponde ad un *Incident ID* duplicato. Tutte le analisi svolte in Qlik prendono in considerazione i reati commessi in quanto si è voluto valutare la generalità degli eventi accaduti.

### 3.3 L'analisi Temporale dei reati

Le prime analisi svolte sui reati riguardano il loro andamento nel tempo. Sono stati realizzati due grafici a barre, visibili rispettivamente in Figura 11 e Figura 12, i quali visualizzano, il conteggio dei reati negli anni, dal 2017 al 2021 e il conteggio dei reati nei mesi degli stessi anni.

### 3.4 L'analisi Spaziale dei reati

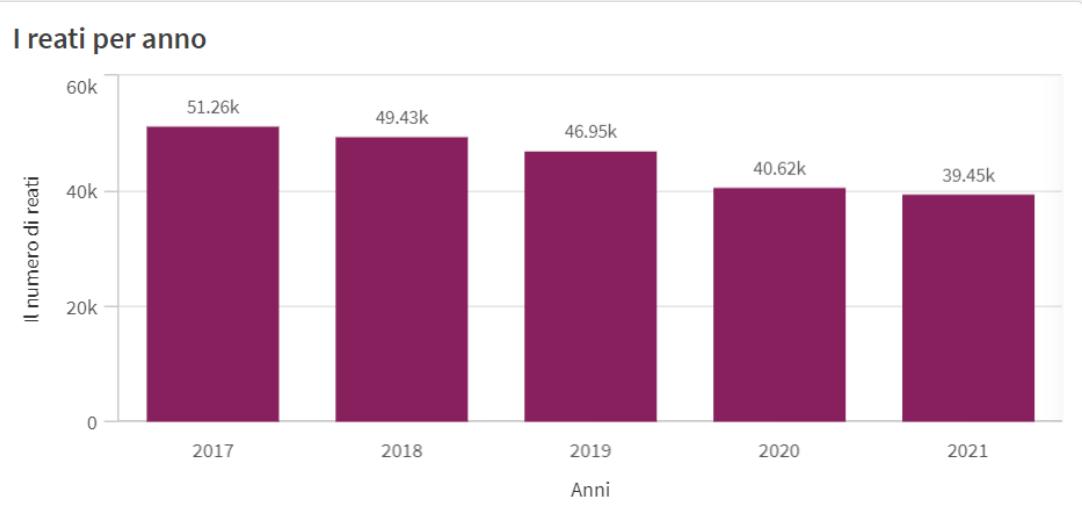


Figura 11: L'analisi dei reati per anno

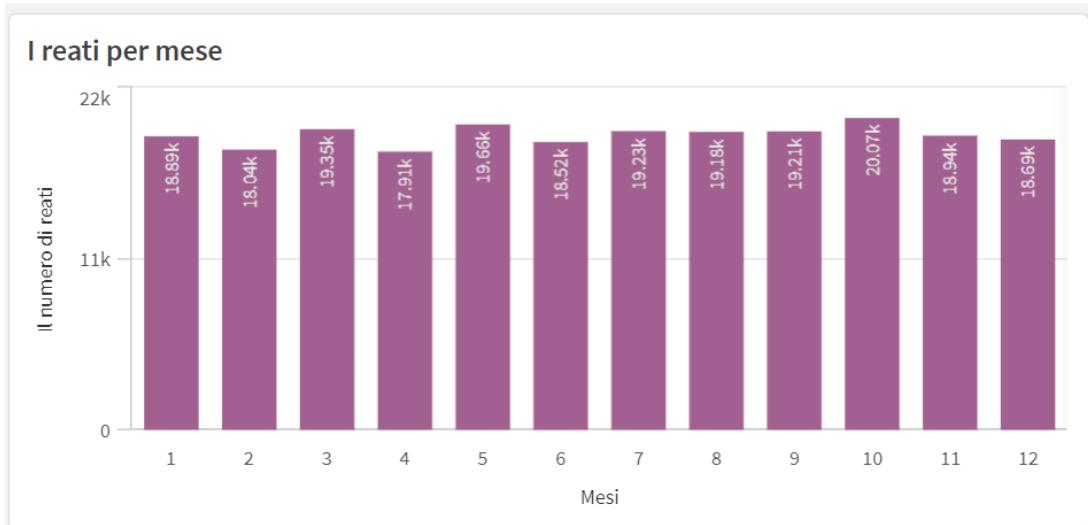


Figura 12: L'analisi dei reati per mese

Nel primo grafico a barre è possibile notare che nel corso del tempo il numero di reati registrati diminuisce, da un valore di 51'000 nel 2017 ad un valore di 39'000 nel 2021. Questa diminuzione è associabile alla pandemia di COVID-19, un'epidemia diffusa a livello globale, iniziata a Dicembre 2019; grazie ad essa molti reati potrebbero esser stati compromessi. Nel secondo grafico a barre, invece, è possibile osservare come i valori dei reati oscillino in un intervallo che varia tra circa 18'000 e 20'000, tra Gennaio e Dicembre degli anni considerati. Il mese di Ottobre registra il maggior numero di casi.

### 3.4 L'analisi Spaziale dei reati

L'analisi descrittiva è proseguita prendendo in considerazione le città e i luoghi in cui sono stati registrati i reati avvenuti tra il 2017 e il 2021.

Il Maryland si divide in 23 contee e una città indipendente chiamata Baltimora, come mostrato nella cartina in Figura 13; queste 23 contee sono composte da 364 città.

Dalla dashboard in Figura 14 si osserva che le città coinvolte sono 51, dunque il 14% dello Stato ha subito reati.

### 3.4 L'analisi Spaziale dei reati

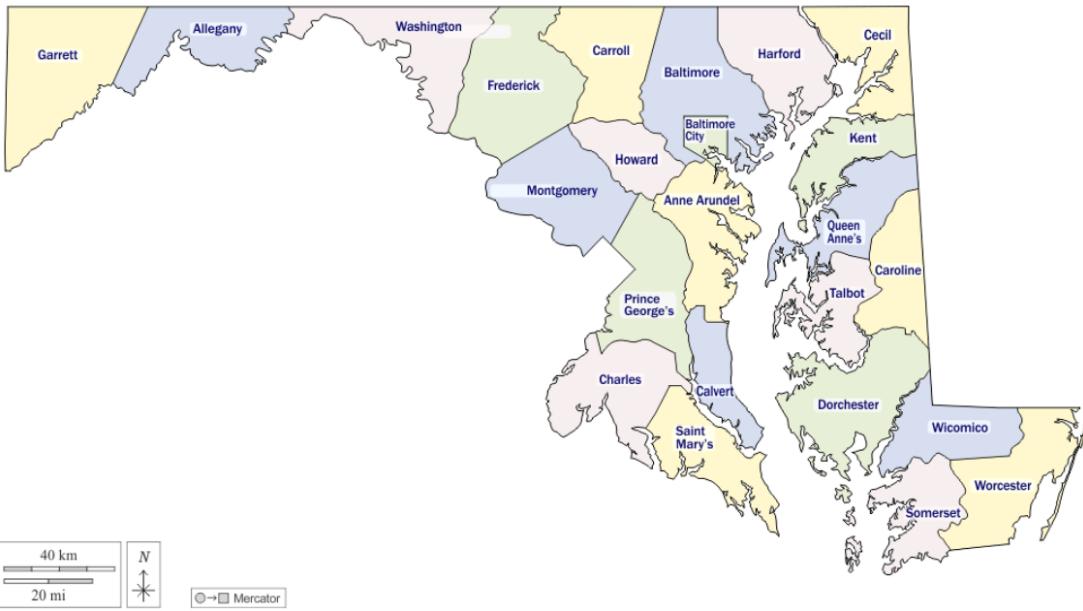


Figura 13: La cartina dello stato del Maryland

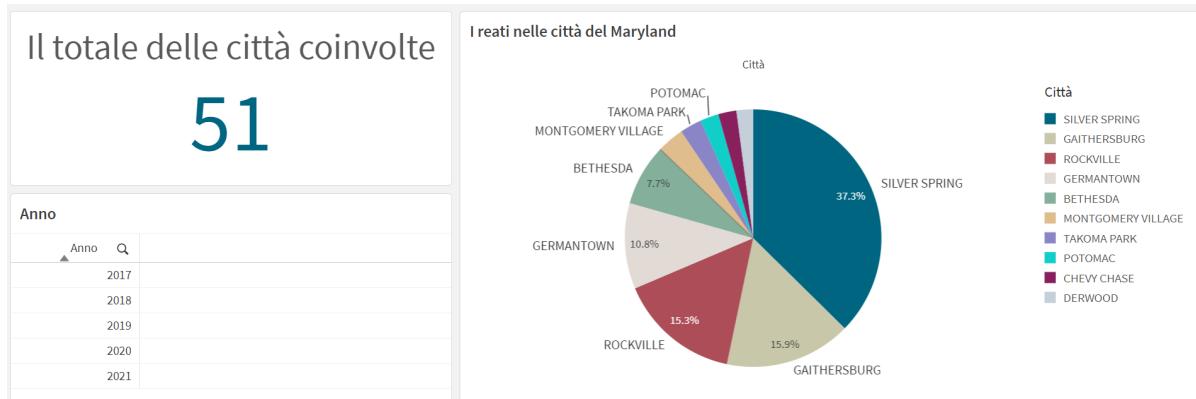


Figura 14: L'analisi delle città in cui sono avvenuti i reati

Nella stessa dashboard è stato inserito un grafico a torta che permette di rappresentare il conteggio dei reati nelle suddette città; si può constatare che quelle più pericolose sono Silver Spring (37,3%), Gaithersburg (15,9%), Rockville (15,3%), Germantown (10,8%) e Bethesda (7,7%). E' interessante notare che queste fanno tutte parte della stessa contea, ovvero la contea del Montgomery e lo si può verificare dalla mappa riportata in Figura 15.

Inoltre, in questo cruscotto è stata inserita una tabella utilizzata come filtro per permettere di selezionare gli anni di interesse. Si è voluto confrontare la concentrazione di reati nel 2017 (Figura 16) e nel 2021 (Figura 17), rispettivamente l'anno in cui sono stati registrati più reati e l'anno in cui ne sono stati registrati meno.

### 3.4 L'analisi Spaziale dei reati



Figura 15: La cartina della contea del Montgomery

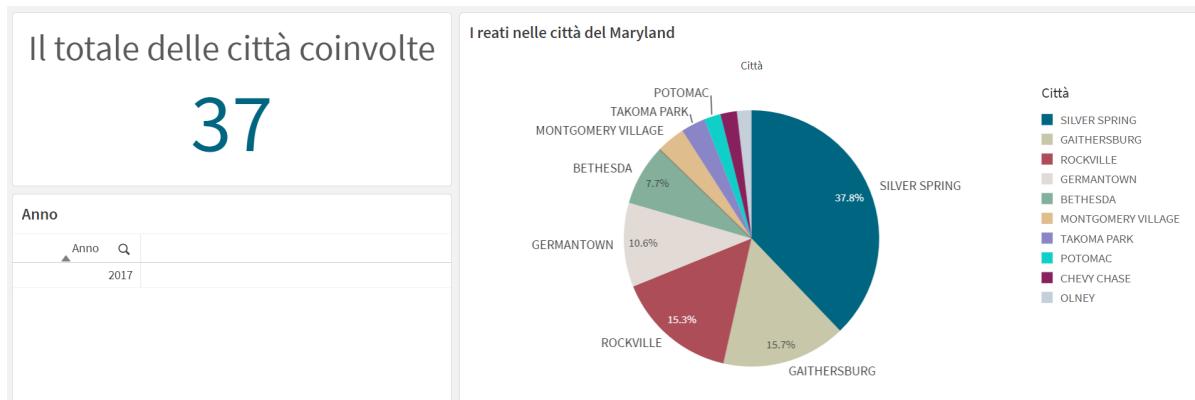


Figura 16: L'analisi delle città in cui sono avvenuti i reati nel 2017

### 3.4 L'analisi Spaziale dei reati

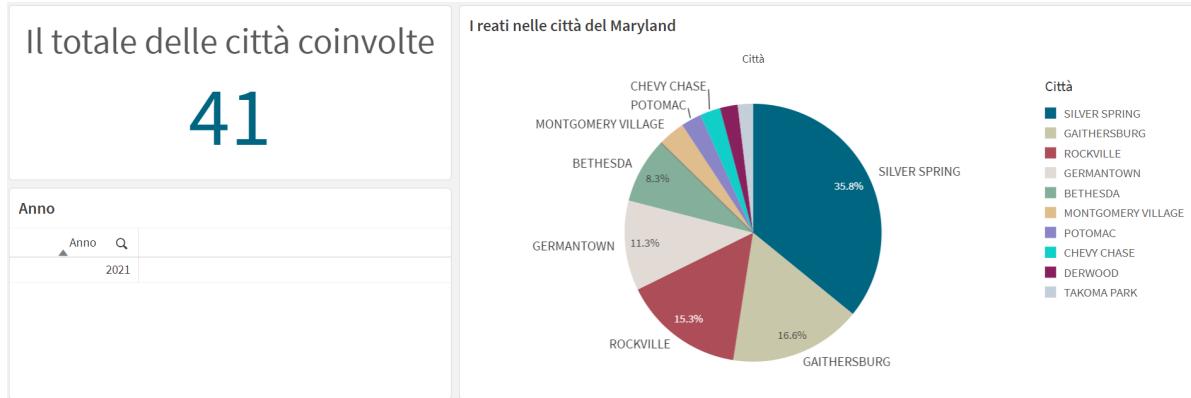


Figura 17: L'analisi delle città in cui sono avvenuti i reati nel 2021

E' possibile notare che nel 2017 il numero di città coinvolte è poco meno rispetto a quelle nel 2021, mentre, le cinque più pericolose sono le stesse. La concentrazione maggiore si presenta sempre a Silver Spring, ma nella suddetta area, nel 2021 si registra una diminuzione dei reati del 2% rispetto al numero registrato nel 2017.

Successivamente, sono stati realizzati alcuni grafici per analizzare i luoghi maggiormente coinvolti nei reati. Il primo è un grafico a barre con visualizzazione orizzontale riportato in Figura 18; esso visualizza il conteggio degli elementi appartenenti all'attributo "filtered\_place", presentato nella fase di ETL del dataset.

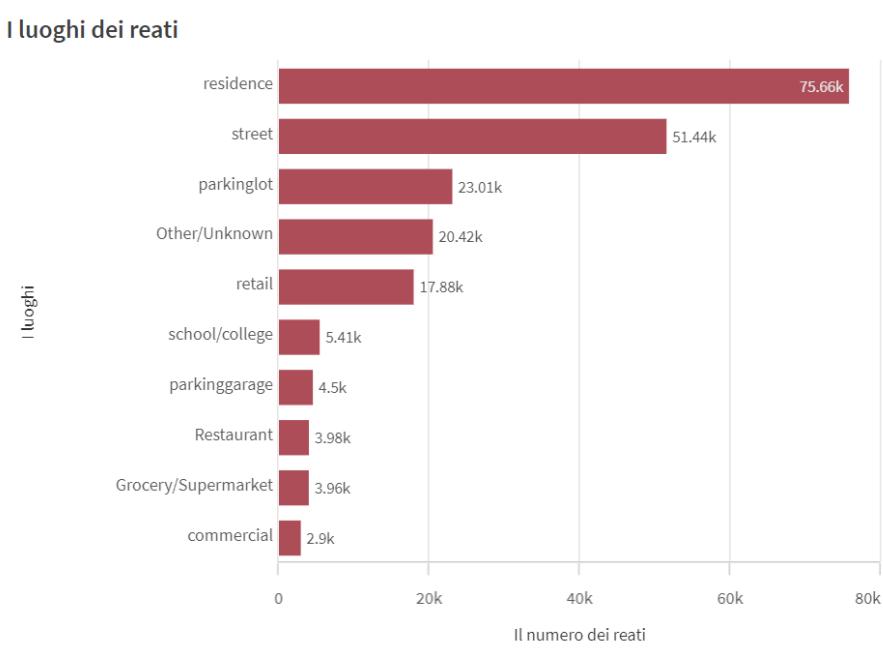


Figura 18: I principali luoghi dei reati

Si può dedurre che i luoghi maggiormente soggetti a reati nelle varie città sono le residenze, le strade ed i parcheggi.

Per approfondire ulteriormente l'analisi spaziale, è stata realizzata una mappa ad albero, considerando i luoghi appena citati come i tre gruppi principali del suddetto grafico. Ciascun gruppo è rappresentato da un rettangolo la cui dimensione è pari al numero di reati avvenuti rispettivamente in residenze, strade e parcheggi; ogni rettangolo, a sua volta, è diviso in tre aree che corrispondono agli elementi del campo Place in cui si sono registrati maggiormente i casi.

### 3.4 L'analisi Spaziale dei reati

#### I luoghi dei reati nel dettaglio



Figura 19: I principali luoghi dei reati più nel dettaglio

Si può constatare che, per ciascun gruppo, i luoghi più frequenti sono le residenze unifamiliari (25'250 reati), i veicoli in strada (24'3600 reati) e i parcheggi delle residenze (11'940 reati).

Proseguendo l'analisi è stato di interesse combinare le informazioni ottenute sulle città e sui luoghi principali; è stata, dunque, realizzata un dashboard in Figura 20, in cui è presente un grafico a barre impilate con visualizzazione orizzontale, accompagnato da una tabella. Il grafico a barre impilate con visualizzazione orizzontale mostra il numero di reati commessi nelle prime tre città che registrano la più alta frequenza di casi, evidenziando per ciascuna i luoghi maggiormente coinvolti. Tali luoghi corrispondono ai segmenti colorati di ciascuna pila. Per avere una visione generale dei dati in questione, è stata riportata una tabella che evidenziasse il numero totale di reati nelle città considerate.

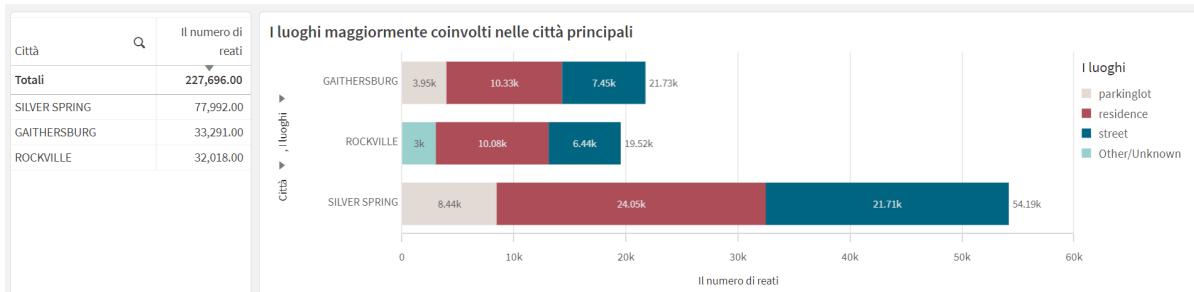


Figura 20: I luoghi maggiormente coinvolti nelle città principali

E' possibile constatare che:

- il 70% dei reati registrati a Silver Spring si concentrano in:
  - residenze (24'050 reati),
  - strade (21'710 reati),
  - parcheggi (8'440 reati);
- il 65% dei reati registrati a Gaithersburg si concentrano in:

### 3.4 L'analisi Spaziale dei reati

- residenze (10'330 reati),
  - strade (7'450 reati),
  - parcheggi (3'950 reati);
- il 68% dei reati registrati a Rockville si concentrano in:
    - residenze (10'080 reati),
    - strade (6'440 reati),
    - luogo sconosciuto (3'000 reati).

Il secondo grafico, che completa l'analisi di questa sezione, è una mappa ad albero, riportata in Figura 21, la quale considera Silver Spring, Gaithersburg e Rockville come i tre gruppi principali; questi sono rappresentati dai tre rettangoli più grandi, ognuno evidenziato da un colore differente, le cui dimensioni sono determinate dal conteggio dei reati in ciascuna zona. Eseguendo l'operazione di drill-down, per ciascuna città sono stati considerati i tre luoghi precedentemente analizzati, che circoscrivono un rettangolo più piccolo. A sua volta, per ciascun luogo, è possibile visualizzare i dettagli dei luoghi stessi, in particolare, i primi tre elementi del campo *Place* in cui si sono registrati maggiormente i casi.



Figura 21: I luoghi dettagliati che sono stati maggiormente coinvolti nelle città di Silver Spring, Gaithersburg e Rockville

Da questo grafico si può dedurre che:

- Silver Spring registra, per ciascun luogo, il maggior numero di reati in corrispondenza di:
  - condomini (8'780 reati) e residenze unifamiliari (8'160 reati),
  - veicoli in strada (9'600 reati) e strade residenziali (6'210 reati),
  - parcheggi residenziali (4'360 reati);
- Gaithersburg registra, per ciascun luogo, il maggior numero di reati in corrispondenza di:
  - condomini (3'570 reati) e residenze unifamiliari (2'810 reati),
  - veicoli in strada (4'470 reati),
  - parcheggi residenziali (2'030 reati);
- Rockville registra, per ciascun luogo, il maggior numero di reati in corrispondenza di:
  - residenze unifamiliari (3'700 reati) e condomini (3'060 reati),
  - veicoli in strada (3'370 reati),
  - luogo sconosciuto (3'000 reati), che non ha particolarizzazioni.

### 3.5 Le vittime coinvolte nei reati

#### 3.5 Le vittime coinvolte nei reati

Per completare la panoramica generale, come ultima analisi sono state analizzate le vittime coinvolte.

In genere per "Vittima" si intende la parte lesa, nonchè chi ha subito un danno; osservando il dataset questo concetto sembra variare in base alla tipologia di crimine, di cui si approfondirà nel prossimo capitolo che tratta le analisi svolte in Tableau. Brevemente, si anticipa che la tipologia di crimine è descritta dal campo *Crime Name 1* e nel presente dataset si evidenziano i crimini contro la persona, contro il patrimonio, contro la società, altri che non rientrano in queste tre categorie, i quali sono contrassegnati con "Other" ed, infine, i "Not a crime", azioni non attribuibili ad un crimine. In Figura 22 sono mostrati tre crimini che, osservando il campo *Crime Name 3*, è possibile affermare riguardino rispettivamente, un'aggressione di secondo grado (crimine contro la persona), un furto d'auto (crimine contro il patrimonio) ed il possesso di attrezzatura per narcotici (crimine contro la società).

Incident ID	Offence Code	CR Number	NIBRS Code	Victims	Crime Name1	Crime Name2	Crime Name3
201122422	1399	17009733	13B		1 Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201138627	1212	170513167		120	1 Crime Against Property	Robbery	ROBBERY - CARJACKING - ARMED
201295084	3550	200027872	35B		1 Crime Against Society	Drug Equipment Violations	DRUGS - NARCOTIC EQUIP - POSSESS

Figura 22: Un campione di elementi del dataset appartenenti a tre tipologie di crimine differente

Ciascuno registra una vittima; mentre la vittima del primo reato si riferisce ad una persona, la vittima del secondo si riferisce all'auto e la vittima del terzo alla società. Dunque, non ha valore informativo analizzare le vittime appartenenti ad una categoria di crimine differente da quella contro la persona e per questo le analisi sono state condotte solo per tale tipologia.

Un ulteriore aspetto da tenere in considerazione è che i crimini considerati potrebbero far parte di un unico reato, come si può osservare dall'esempio in Figura 23; in questo caso, nonostante sia stata registrata una vittima per crimine, la descrizione degli stessi (in corrispondenza del campo *Crime Name 3*) fa presupporre che la vittima sia sempre la stessa. Dunque, questo reato registra una sola vittima ed, in generale, è importante tenere conto di tali duplicati per effettuare un corretto conteggio delle vittime coinvolte nei reati.

Incident ID	Offence Code	CR Number	NIBRS Code	Victims	Crime Name1	Crime Name2	Crime Name3
201328668	1103	210015612	11A		1 Crime Against Person	Forcible Rape	RAPE - STRONG-ARM
201328668	9003	210015612	11C		1 Crime Against Person	Sexual Assault With An Object	SEX ASSAULT - WITH AN OBJECT

Figura 23: Un campione di elementi del dataset appartenenti alla tipologia di crimine contro la persona, in cui uno stesso reato è costituito da due crimini

E' stata realizzata un dashboard (Figura 24) che permettesse di visualizzare le vittime dei reati in questione.

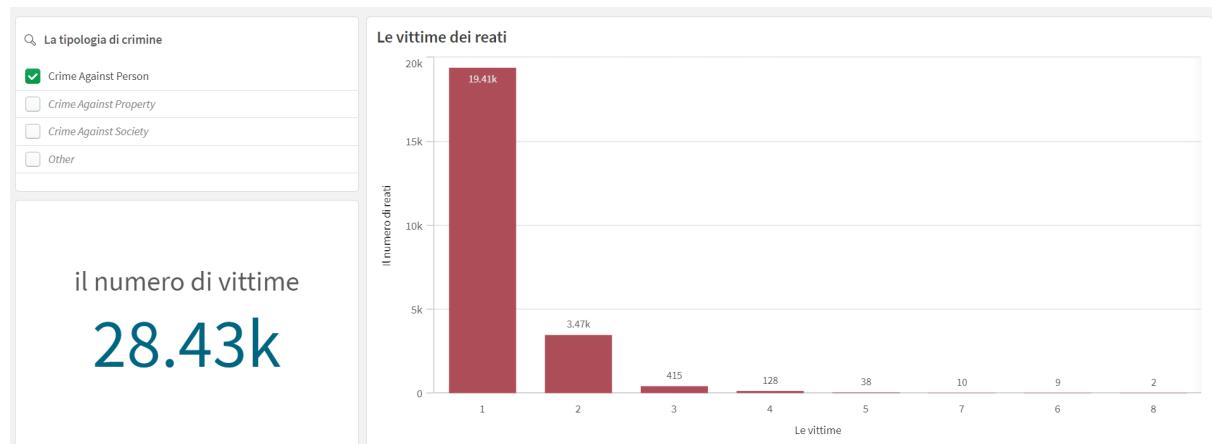


Figura 24: Le vittime dei reati

### 3.5 Le vittime coinvolte nei reati

In alto a sinistra è riportato un filtro per selezionare la categoria dei crimini ed è stata scelta la tipologia di interesse; sulla destra è riportato un grafico a barre che valuta l'occorrenza delle vittime per ciascun reato. E' evidente che la maggior parte dei reati (19'410) ha una sola vittima e che il valore più alto di vittime per reato è 8 che, però, è registrato in soli due eventi. Nella tabella in Figura 25 sono riportati i reati con 7 e 8 vittime e si evince che la maggior parte riguardano un'aggressione di secondo grado, due delle quali avvenute a mano armata.

Incident ID	Victims	Crime Name1	Crime Name2	Crime Name3
201116450	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201146813	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201158596	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201167448	7	Crime Against Person	Human Trafficking, Commercial Sex Acts	HUMAN TRAFFICKING - COMMERCIAL SEX ACTS
201170308	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201241666	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201246411	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201271466	7	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE
201313582	7	Crime Against Person	Aggravated Assault	ASSAULT - AGGRAVATED - GUN
201352791	7	Crime Against Person	All Other Offenses	OBSCENE COMMUNICATION
201202272	8	Crime Against Person	Aggravated Assault	ASSAULT - AGGRAVATED - NON-FAMILY-GUN
201308411	8	Crime Against Person	Simple Assault	ASSAULT - 2ND DEGREE

Figura 25: I reati contro la persona con più vittime

Infine, è stato calcolato il numero totale di vittime dei reati in questione, visibile nel KPI in basso a sinistra della dashboard. Questo valore è stato calcolato utilizzando la funzione in Figura 26 che permette di calcolare la somma dei valori distinti delle vittime per ciascun reato. In particolare, la funzione "*Aggr(Sum(Distinct(Victims)), [Incident ID])*" aggrega i dati per *Incident ID* e somma i valori distinti (non duplicati) di *Victims* per ogni ID; la funzione esterna *Sum()* calcola la somma totale di tutti i valori selezionati dalla funzione *Aggr()*. Dunque, il numero totale di vittime di reati è 28'430.

---

**Sum(Aggr(Sum(Distinct(Victims)), [Incident ID]))**

Figura 26: La funzione per calcolare il numero totale di vittime dei reati

## 4 Tableau

Tableau è uno dei leader nel mercato per la Business Intelligence tra i più affermati e utilizzati per la comprensione e l'analisi dei dati; esso, infatti, si posiziona subito dietro il software Power BI nel Magic Quadrant.

Rispetto al tool QlikSense, ampiamente illustrato nella sezione precedente, Tableau ha la possibilità di una visualizzazione attraverso mappe geografiche, anche se risulta essere meno chiaro e intuitivo in termini estetici ma sicuramente più potente nella gestione ed esplorazione dei dati dai quali estrarre e condividere conoscenza.

Gli utenti, di fatto, possono creare e distribuire un dashboard interattivo e condivisibile, il quale rappresenta le tendenze o gli insight ricavati dai dati. Tableau, inoltre, può collegarsi a file, database relazionali e server di ogni tipo per acquisire ed elaborare i dati, considerato uno dei punti di forza del tool.



Figura 27: Tableau

### 4.1 Dataset e Caricamento Dati

Il dataset impiegato per il progetto di Tableau è lo stesso adottato per QlikSense, una raccolta di dati delle diverse tipologie di crimini denunciati negli Stati Uniti e, in particolare, nello stato del Maryland dal 2017 al 2021.

In maniera analoga a quanto eseguito con Qlik, la prima fase è stata quella relativa al caricamento dei dati mediante l'interfaccia messa a disposizione dal tool stesso.

### 4.2 Analisi Effettuate

Proseguendo con l'esplorazione dei dati iniziata con QlikSense, l'analisi effettuata in Tableau ha l'obiettivo di focalizzarsi sulle diverse tipologie di crimini. Se da un lato con QlikSense si ha una visione d'insieme dei reati commessi nello stato del Maryland, con Tableau si entra ad un livello di dettaglio maggiore concentrandosi sui diversi aspetti che caratterizzano ciascuna tipologia di crimine.

In particolare, in tale sezione inizialmente verranno riportate le informazioni estrapolate da QlikSense in una panoramica generale; successivamente, si passerà all'illustrazione delle tipologie di crimine e le loro caratteristiche ed, infine, si porrà l'attenzione sull'applicazione del forecasting sui risultati ottenuti dall'analisi temporale. Tutto ciò allo scopo di raggiungere l'obiettivo ultimo dell'elaborato, ovvero fornire informazioni utili che possano rendere il dipartimento di polizia più efficiente nella propria missione: prevenire e reprimere il crimine.

#### 4.2.1 Panoramica generale

Prima di iniziare ad analizzare nel dettaglio le tipologie di crimine nello stato del Maryland, è stato creato un dashboard panoramico che riassume le principali statistiche e le informazioni presenti nel modello di dati di Tableau relativi ai crimini commessi (Figura 28).

L'obiettivo del dashboard, infatti, è quello di avere un quadro d'insieme concernente le proprietà che caratterizzano i diversi tipi di crimini dal 2017 al 2021. Da un lato vi è la composizione dei crimini mediante una visualizzazione a bolle, le cui dimensioni variano in base al numero di crimini denunciati. Da ciò si evince chiaramente come i reati contro la proprietà rappresentano il più alto valore contando 112.840 contro i 248.251 dei crimini totali commessi, quasi la metà di tutti i reati; a parità di estensione, seguono

## 4.2 Analisi Effettuate

i crimini contro la società, "altri", i crimini contro la persona ed, infine, "not a crime". Quest'ultimo, non essendo identificato come crimine dal NIBRS, non dispone di tutte quelle caratteristiche intrinseche di un crimine stesso; pertanto, tale tipologia è stata omessa nelle successive analisi. Si ricorda, inoltre, che con il termine "Altri" si intendono tutti quei crimini che non sono classificati nelle tre principali categorie previste dal NIBRS.

Nella seconda metà del dashboard, l'attenzione si riversa su un aspetto interessante utile a prevenire i crimini: due grafici a linee che rappresentano rispettivamente i giorni della settimana e orari più frequenti in cui essi vengono commessi; ciò è stato possibile modificando i livelli del campo *Start\_Date\_Time* secondo la sezione della parte di data di interesse in quanto, riconosciuta come tipo di dato *data e ora*, Tableau permette di selezionare il grado della gerarchia desiderato. Allo scopo di garantire una lettura più semplice ed intuitiva dei grafici a linee, è stata inserita alla linea una colorazione divergente verde-rosso sulla base del conteggio dei crimini.

Alla luce di queste premesse, un'osservazione interessante sull'andamento dei crimini su base settimanale è la presenza di una diminuzione significativa del numero di crimini nel fine settimana rispetto ai giorni feriali, con un picco di eventi criminali nella giornata di venerdì.

Per quanto riguarda i crimini per orario giornaliero, il grafico a linea mostra come l'andamento sia abbastanza irregolare: si evidenzia un picco alle 00:00 per poi registrare un drastico calo nelle ore notturne e nuovamente un rialzo a partire dalle ultime ore del mattino.

Sulla base di tali visualizzazioni, sono state necessarie ulteriori ricerche comportamentali della popolazione, di cui però si rimanda alle sezioni successive.

Infine, in basso a destra viene illustrata, tramite un grafico a barre verticali, la percentuale di crimini commessi per le città dello stato del Maryland. Come mostrato già nella sezione del tool precedente, Silver Spring rappresenta la città con la più alta percentuale. Infatti, essa è tra le aree urbane a più alto rischio di criminalità nella contea di Montgomery.

Il cruscotto, inoltre, risulta fortemente interattivo; infatti, se si seleziona una categoria di crimine nel grafico a bolle, quest'ultima filtra i dati mostrandone, dunque, i dati particolarizzati per essa. Nella Figura 29 è mostrato un esempio, ottenuto dopo aver individuato la categoria di reati contro la persona.

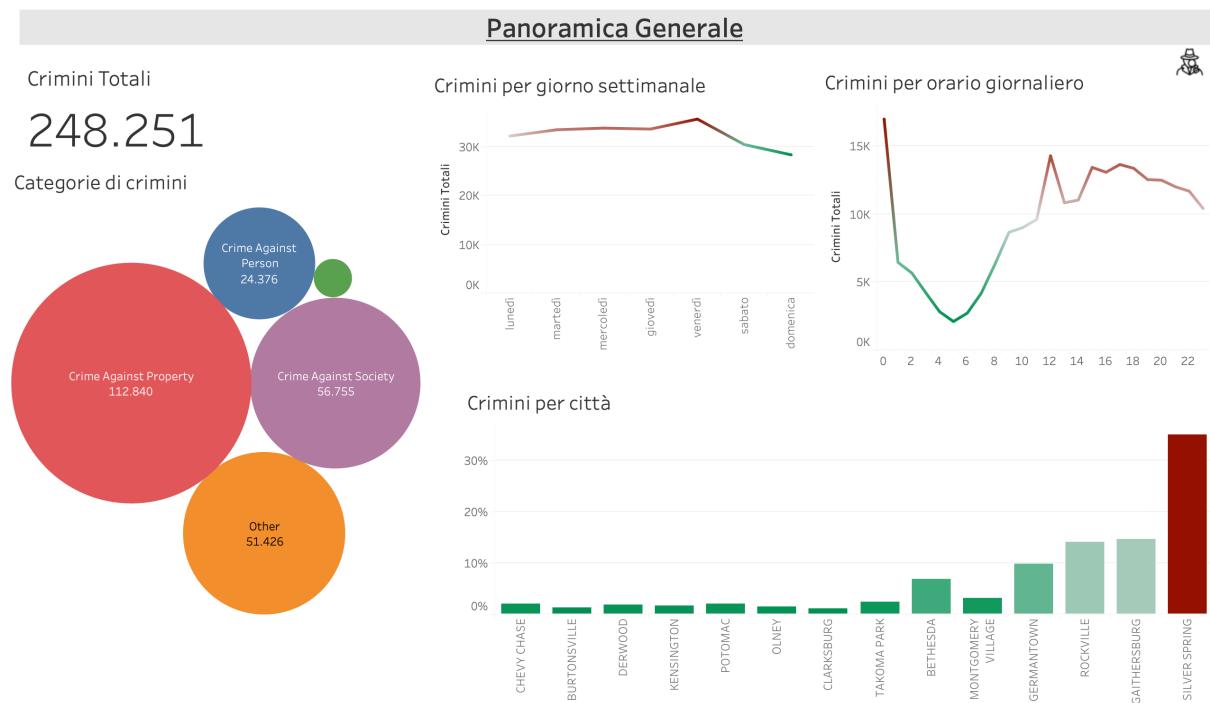


Figura 28: Tableau - Panoramica generale

## 4.2 Analisi Effettuate

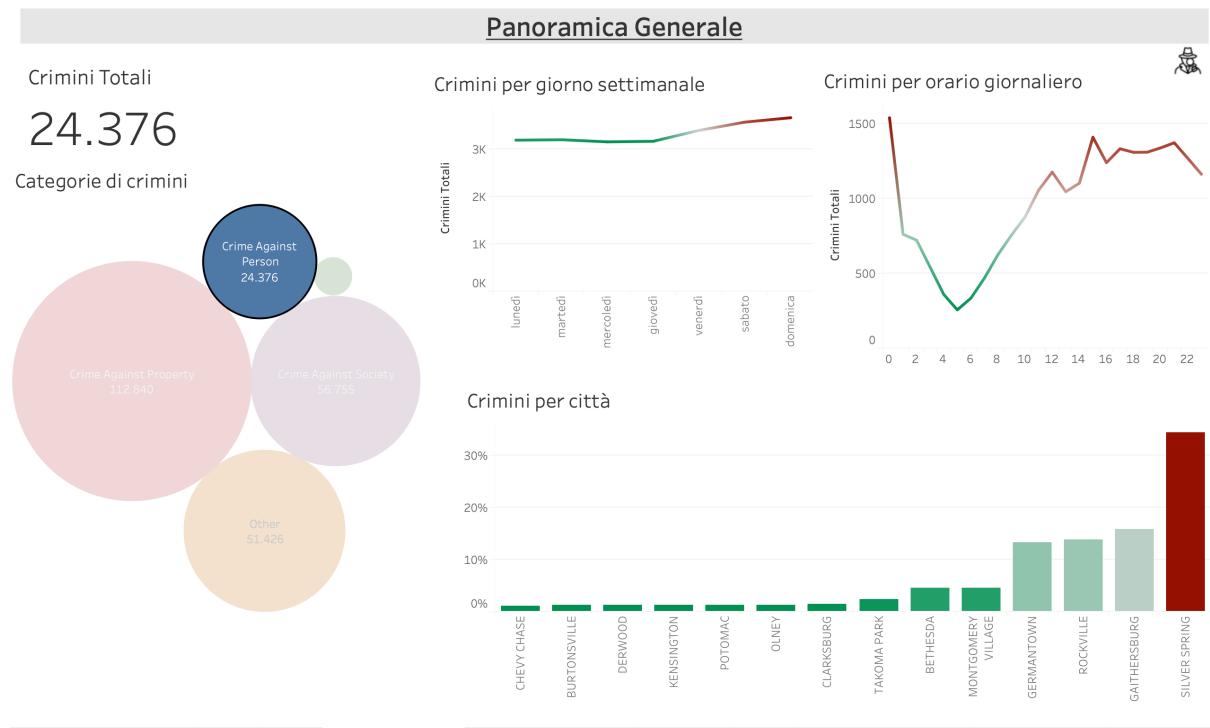


Figura 29: Tableau - Crimini contro la persona

### 4.2.2 Tipologie di crimine

Una volta aver introdotto, in linea generale, il fenomeno della criminalità nella contea del Montgomery, si è proseguito l’analisi concentrandosi sulle diverse tipologie di crimine e da cosa esse sono caratterizzate. Prima di addentrarsi nello studio, si riportano alcune considerazioni sulla distinzione delle categorie effettuata dal sistema di segnalazione delle forze dell’ordine negli Stati Uniti per raccogliere e segnalare dati sui crimini (NIBRS), al fine di rendere più leggibili le visualizzazioni che seguono.

Il NIBRS, infatti, sebbene esistano molti tipi diversi di crimini, esprime gli atti criminali segnalati e denunciati sulla base di tre principali categorie: crimini contro la proprietà, crimini contro la persona e crimini contro la società.

I crimini contro la proprietà, in genere, comportano l’interferenza con la proprietà di un’altra parte. Sebbene possano comportare un danno fisico o mentale per un altro, essi hanno come obiettivo la privazione e l’ottenimento di denaro, beni o altri vantaggi. Molti crimini contro il patrimonio sono, ad esempio, reati di furto tra cui furto con scasso, furto, rapina, furto d’auto e taccheggio.

I crimini contro la società, invece, rappresentano il divieto della società di intraprendere determinati tipi di attività e, tipicamente, sono senza vittime in cui la proprietà non è quindi l’oggetto. Esempi sono il gioco d’azzardo, la prostituzione e le violazioni della droga.

I crimini contro la persona, al contrario, provocano un danno fisico o mentale ad un’altra persona.

Terminate tali premesse, si è proseguito l’analisi con il grafico a barre in pila (Figura 30) in cui viene illustrato la distribuzione delle tipologie di crimini raggruppate per ciascun anno all’interno del range 2017-2021. A parità del totale di crimini, il grafico mostra un leggero declino per gli anni 2018 e 2019 e una sostanziale discesa per gli ultimi anni rispetto al 2017.

In linea con quanto annunciato precedentemente, i crimini contro la proprietà non solo è la categoria più frequente all’interno dell’intero dataset ma è anche quella che mostra il più alto numero di crimini ogni anno. Un altro aspetto che emerge è la notevole diminuzione dei crimini contro la società a partire dal 2020, in concomitanza con l’avvenire della pandemia da COVID-19 e conseguenti lockdown come misura di contenimento in cui la quasi totalità delle attività furono vietate e tra queste anche quelle vietate dalle società.

## 4.2 Analisi Effettuate

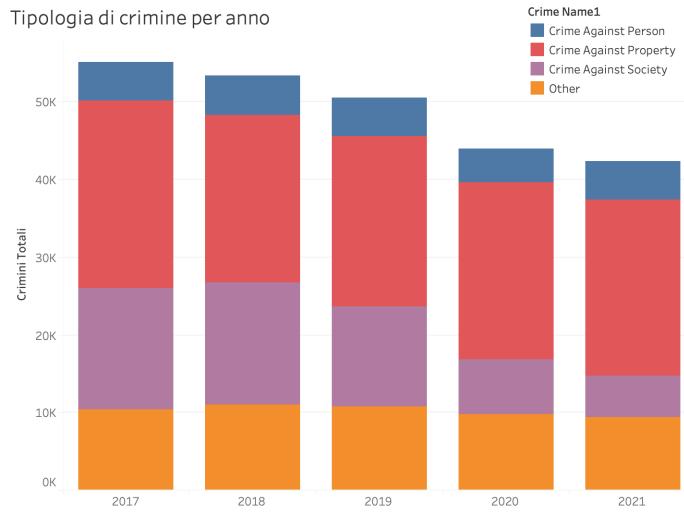


Figura 30: Tableau - Tipologie di crimini per anno

Giusti a questo punto, è stato ritenuto interessante effettuare un'analisi più approfondita delle principali categorie di crimine. Pertanto, nelle successive sezioni verranno esaminate individualmente ciascuna categoria, al fine di ricavare informazioni importanti.

### 4.2.3 Crimini contro il patrimonio

Innanzitutto, si ricorda che i crimini contro la proprietà comprendono tutti quegli atti il cui bersaglio è l'acquisizione di una proprietà materiale posseduta da qualcun altro. Il cruscotto particolarizzato per tale categoria è mostrato nella Figura 32.

Nella parte superiore a sinistra sono presenti, da una parte un indicatore che mostra, in valore numerico, il totale dei crimini per la categoria di riferimento e, dall'altra un filtro su campo data relativo agli anni disponibili per filtrare, appunto, i dati delle visualizzazioni e rendere, dunque, il dashboard interattivo agli occhi dell'utente.

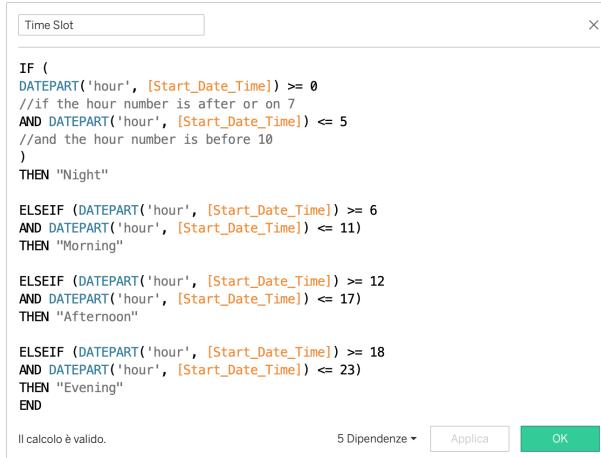
Un grafico interessante è quello a barre orizzontali posizionato nella parte superiore in cui viene mostrata la top 5 sul dettaglio dei crimini contro il patrimonio. Come emerge dal grafico, i furti di veicoli a motore registrano il più alto numero di crimini contro il patrimonio, seguito dalla distruzione, danneggiamento e valdalsimo dei beni, dal taccheggiamento ed, infine, tutti gli altri furti e furti dagli edifici.

Per avere un quadro più chiaro su quando questi crimini principali vengono commessi, si vede l'ultimo grafico mostrato nel dashboard: crimini per fascia oraria.

Il campo "Start\_Date\_Time" presente nel dataset, oltre alla data composta da giorno, mese e anno, espone anche l'ora e il minuto in cui il crimine è stato commesso, in seguito alla ricostruzione avvenuta al momento della registrazione del crimine stesso. Ai fini della leggibilità dell'utente, è stato deciso di raggruppare il tempo in fasce orarie. In tal senso, considerando l'arco giornaliero composto da 24 ore, quest'ultimo è stato diviso in 4 time slot, ciascuna composta da 6 ore: mattina, pomeriggio, sera e notte.

Per replicare l'idea in Tableau, è stato necessaria la creazione di un campo calcolato dal nome *Time Slot* con all'interno l'inserimento della formula mostrata in Figura 31. Tale formula, di fatto, utilizza la funzione *DATEPART()*, la quale consente di restituire una parte specifica di data/ora. Nello specifico caso di studio, essa produce la sezione della data riguardante l'ora e li raggruppa in base alle condizioni desiderate.

## 4.2 Analisi Effettuate



```

Time Slot

IF (
  DATEPART('hour', [Start_Date_Time]) >= 0
  //if the hour number is after or on 7
  AND DATEPART('hour', [Start_Date_Time]) <= 5
  //and the hour number is before 10
)
THEN "Night"

ELSEIF (DATEPART('hour', [Start_Date_Time]) >= 6
AND DATEPART('hour', [Start_Date_Time]) <= 11)
THEN "Morning"

ELSEIF (DATEPART('hour', [Start_Date_Time]) >= 12
AND DATEPART('hour', [Start_Date_Time]) <= 17)
THEN "Afternoon"

ELSEIF (DATEPART('hour', [Start_Date_Time]) >= 18
AND DATEPART('hour', [Start_Date_Time]) <= 23)
THEN "Evening"
END

Il calcolo è valido.      5 Dipendenze ▾    Applica    OK

```

Figura 31: Tableau - Formula campo calcolato per fasce orarie

Ritornando al grafico sui crimini per fasce orarie poc'anzi introdotto, esso è stato ottenuto inserendo proprio il campo calcolato *Time Slot* precedentemente illustrato, al fine di distribuire i tipi di crimini più frequenti nei periodi di tempo corrispondenti. Per ragioni di maggiore leggibilità, si mostrano i tipi di crimini emersi dalla top 5 del grafico della Figura 32.

In linea generale, si evidenzia che i crimini avvengono maggiormente nel pomeriggio e nelle ore serali; al contrario, al mattino e durante le ore notturne tendono a verificarsi meno casi. Entrando nel dettaglio, si può notare che i crimini tendono a distribuirsi in maniera differente nelle diverse fasce orarie, probabilmente una conseguenza del mutevole comportamento della popolazione nelle distinte fasce orarie giornaliere. Nelle ore pomeridiane, ad esempio, il tipo di crimine più frequente è il taccheggio, ovvero il furto di merce che si trova esposta in vendita all'interno di esercizi commerciali, seguita dai furti di veicoli a motore e tutti gli altri furti.

Distruzione, danneggiamento e vandalismo di proprietà, invece, sono i tipici crimini che avvengono nelle ore notturne o nelle prime ore del mattino, causati presumibilmente da gente incosciente indotta dall'uso di alcool e droga soltanto per provare un senso d'onnipotenza.

Al contrario, nelle ore serali il crimine più frequente tende ad essere il furto di veicoli a motore, probabilmente dovuto al rientro della popolazione nella propria abitazione in seguito a una giornata lavorativa.

## 4.2 Analisi Effettuate

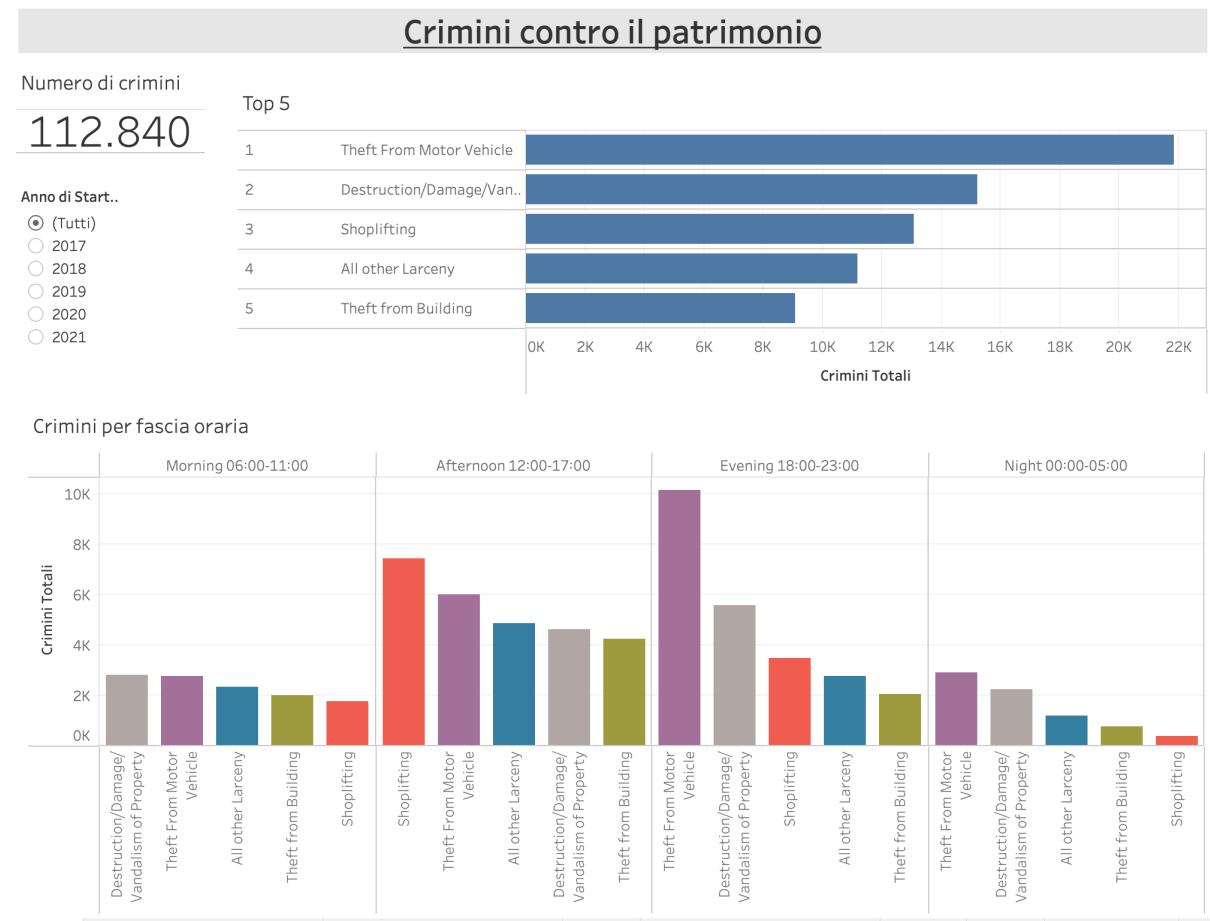


Figura 32: Tableau - Crimini contro il patrimonio

Alla luce di ciò, per approfondire la questione relativa alle località in cui i crimini contro il patrimonio vengono commessi, è stato creato un dashboard di dettaglio sui crimini contro la proprietà sui luoghi in cui vengono registrati gli stessi crimini. (Figura 33). In tale cruscotto, oltre ai grafici in valore numerico del numero di crimini e all'elenco degli anni utile a filtrare le intere visualizzazioni, sono presenti due importanti rappresentazioni raffiguranti i luoghi. In particolare, la tabella sulla sinistra mostra, in valore percentuale, l'elenco dei luoghi e il grafico a torta sulla destra ne illustra il dettaglio; da ciò consegue che le due visualizzazioni sono fortemente interattive.

Come si può notare, il dashboard risulta già filtrato per il luogo con la più alta percentuale di crimini contro la proprietà commessi, ovvero le abitazioni con quasi il 29%. Il dettaglio delle residenze raffigurato nel grafico a torta mostra come il principale bersaglio dei criminali sia il vialetto, luogo in cui vengono parcheggiate le autovetture di proprietà. Questo risulta perfettamente in linea con quanto illustrato nel dashboard precedente dove, appunto, il furto di veicoli a motore risultava essere il crimine più frequente, soprattutto nelle ore serali 18:00-22:00, presumibilmente coincidente con l'uscita serale dei proprietari dalla propria abitazione senza autovettura. In aggiunta, avere un vialetto nella propria abitazione farebbe pensare che la proprietà sia di una famiglia benestante o comunque che vivano in medie condizioni economiche. Allo stesso modo seguono i crimini negli appartamenti e condomini e le abitazioni di una singola famiglia.

Inoltre, oltre alle abitazioni, restano invariante al circa 14% i crimini nei parcheggi, nelle vendite al dettaglio e nelle strade pubbliche.

Come si può notare, i principali luoghi dei crimini risultano in perfetta linea con quanto emerso nel dashboard precedente, in cui si mostrava la top 5 delle categorie dei crimini stessi. Infatti, i vialetti delle abitazioni, i parcheggi, e le strade residenziali fanno chiaramente pensare a furti di auto; al contrario, le vendite al dettaglio, anch'essa con un'alta percentuale, sposta l'attenzione sul taccheggio.

## 4.2 Analisi Effettuate

I restanti luoghi, con una percentuale che va dall'1% al 3% circa, farebbero pensare alla distribuzione degli atti valdalici e alla distruzione degli edifici.

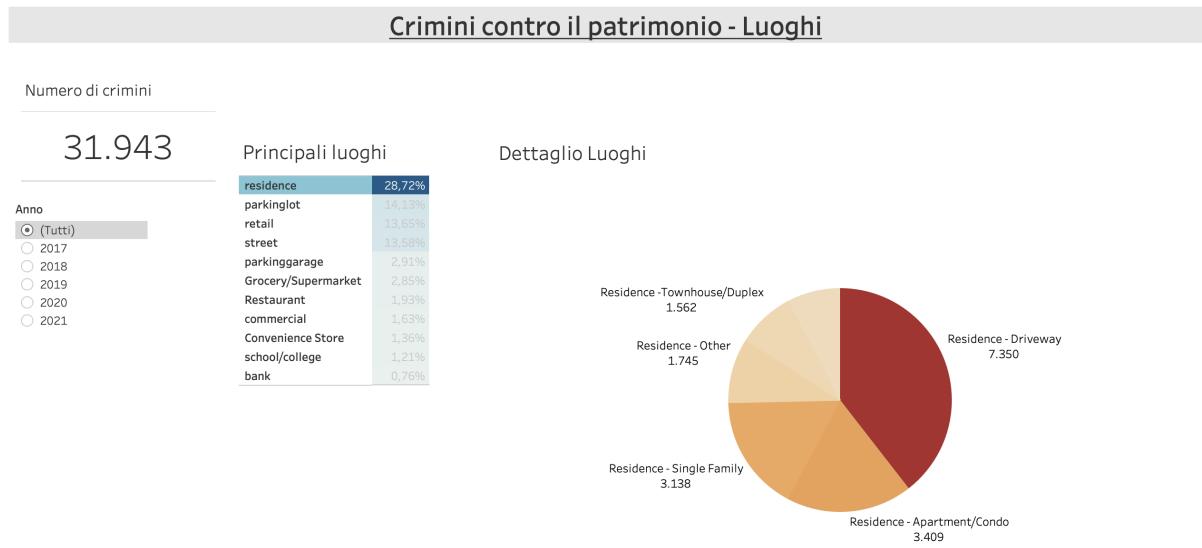


Figura 33: Tableau - Crimini contro il patrimonio, Luoghi

Tuttavia, una volta aver analizzato i principali luoghi in cui i crimini vengono commessi, è necessario essere a conoscenza esattamente in quali territori del Montgomery si concentrano maggiormente, al fine di riuscire ad adottare misure di controllo più restrittive e sorvegliare le aree più inclini ai crimini. Per fare ciò, sono stati utilizzati i campi relativi alla longitudine e latitudine presenti nel dataset per realizzare una mappa che mostri informazioni di dettaglio utili all'adozione di adeguate norme di protezione. Prima di mostrare la mappa, però, è utile tenere a mente di un limite del tool nella realizzazione delle mappe con le coordinate geografiche, in quanto tende ad unire coordinate vicine, anziché mostrare come punti distinti.

Alla luce di ciò e per ragioni di maggiore leggibilità e semplicità, ci si è concentrati sulla città con il più alto tasso di criminalità emersa dai grafici precedenti, ovvero la città di Silver Spring, la cui mappa è mostrata nella Figura 34. Tale mappa mostra il conteggio dei crimini contro il patrimonio delle principali categorie elencate nella legenda, distribuite tra le coordinate geografiche. Come annunciato poc'anzi, Tableau è poco sensibile alle piccole differenze nelle coordinate geografiche, per cui nella mappa i crimini della città di Silver Spring appaiono accorpati in un unico punto.

Quello che però si può constatare è l'alta concentrazione di strutture, residenze, aree adibite a parcheggio nei dintorni che confermano i dati emersi dai dashboard precedenti, in cui il taccheggio e i furti di veicoli a motore si mostrano essere le categorie di crimini contro il patrimonio più frequenti.

Stando a ricerche esterne, diversi sono i fattori che fanno di Silver Spring il leader in alcuni crimini tra le città del Maryland. Tra queste, oltre al fatto che la città ha la popolazione più numerosa, è anche la città con il più alto tasso di disoccupazione; ciò è un ulteriore conferma e spiegherebbe probabilmente il motivo per cui il furto di veicoli a motore e taccheggio sono tra i crimini più frequenti.

## 4.2 Analisi Effettuate

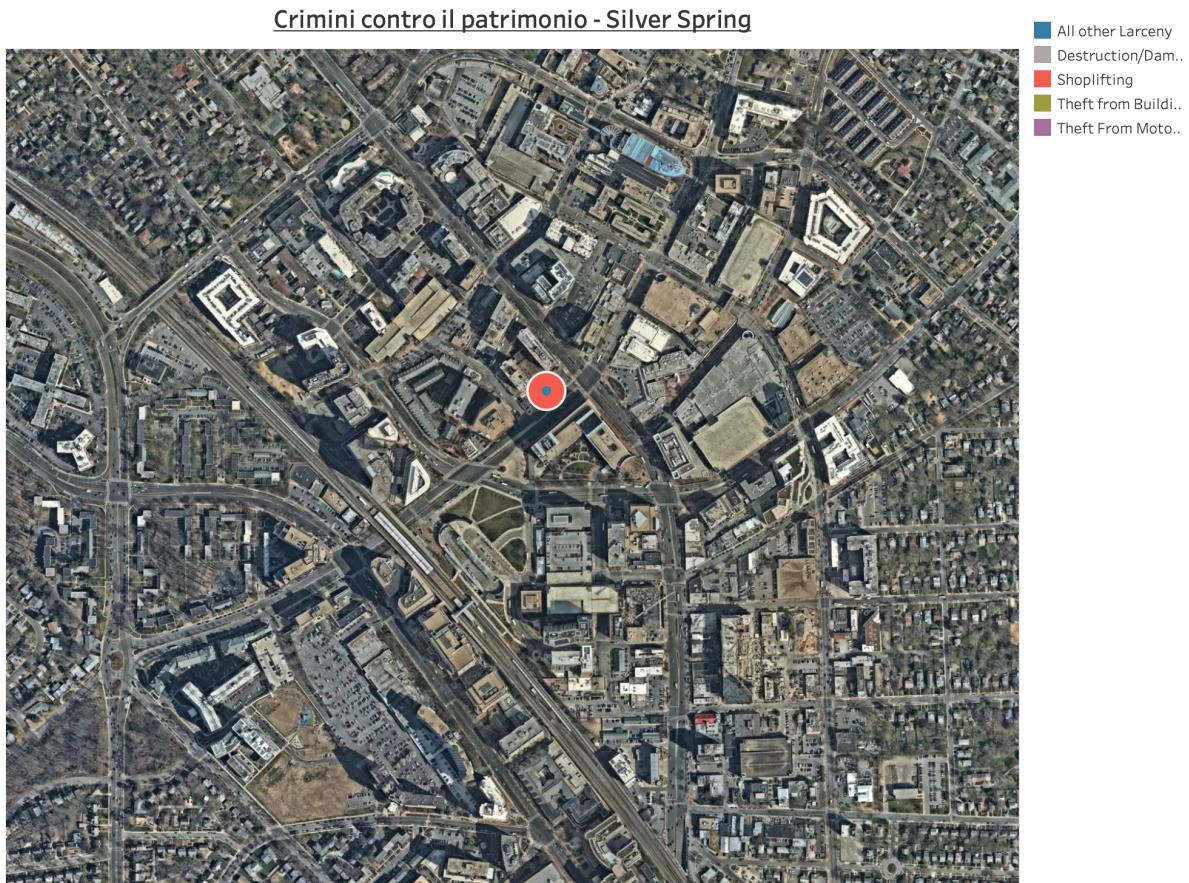


Figura 34: Tableau - Mappa Silver Spring Crimini contro il patrimonio

### 4.2.4 Crimini contro la persona

Una volta aver analizzato i crimini contro la proprietà e tratto le conclusioni ampiamente discusse nella sezione precedente, ci si concentra ora sui crimini contro la persona, il cui dashboard viene mostrato nella Figura 35. E' importante tenere nota del fatto che i grafici del cruscotto sono gli stessi della categoria dei crimini contro il patrimonio; per tale ragione si rimanda alla sezione precedente per la descrizioni delle stesse rappresentazioni.

Si ricorda come con crimini contro la persona si intendono tutti quegli atti in cui viene provocato un danno fisico e, talvolta anche mentale, ad un'altra persona. Tra le categorie mostrate nel grafico a barre in alto a sinistra appare evidente come l'aggressione singola e aggravata siano i tipi di crimini più frequenti, contando quasi la totalità del numero di crimini contro la persona. In numero minore ma che prestano particolare attenzione sono i tipi di crimini legati al palpeggiamento, stupro e sodomia forzati, eventi raccapriccianti che dovrebbero far riflettere in che società si vive e sensibilizzare la popolazione sul tema della sicurezza.

Il grafico dei crimini per fasce orarie mostra come le aggressioni non si concentrano maggiormente in alcune fasce orarie piuttosto che altre bensì sono presenti in tutto l'arco della giornata.

## 4.2 Analisi Effettuate

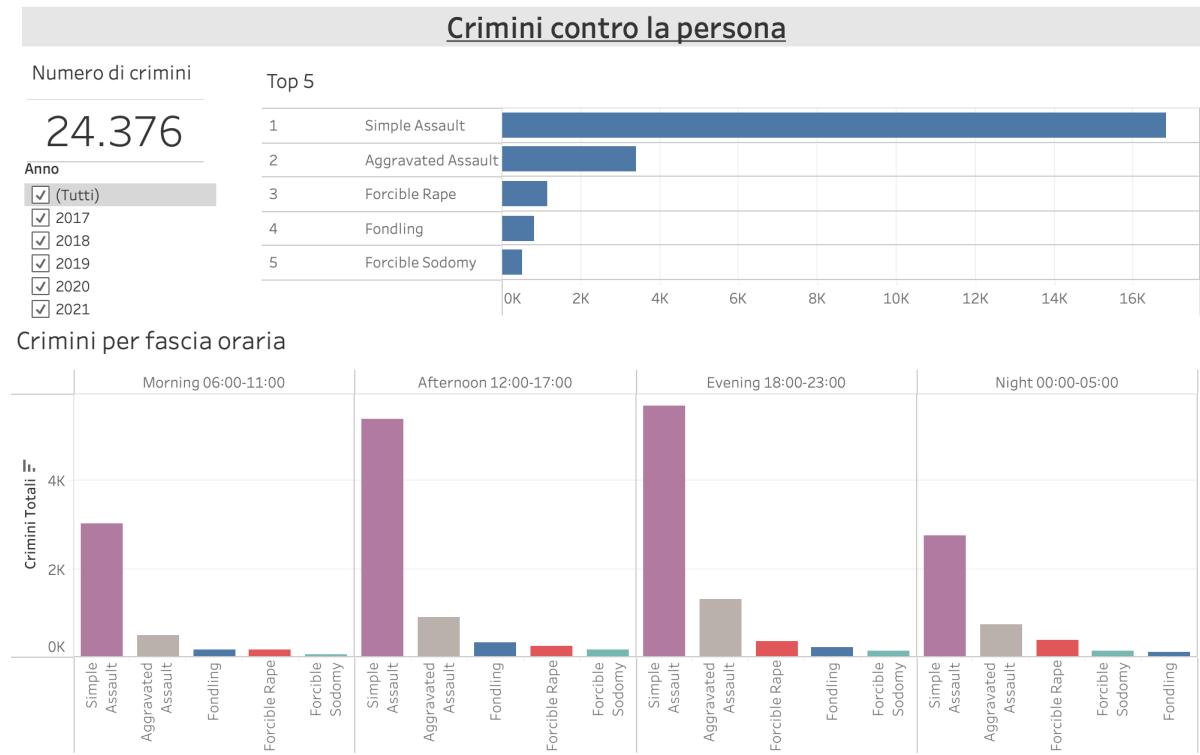


Figura 35: Tableau - Crimini contro la persona

Al fine di trarre conclusioni più esaustive e adottare misure di prevenzione per limitare l'avvenire di crimini contro la persona, si mostrano i grafici in cui si vanno ad analizzare i luoghi in cui tali eventi purtroppo avvengono maggiormente (Figura 36). Com'era in realtà plausibile, i crimini contro la persona avvengono generalmente per strada con una percentuale pari al 58%.

Questo aspetto dovrebbe rivelare la presa di coscienza che le strade sono poco inclini al controllo delle forze dell'ordine nonché prive di telecamere. Traspare chiaramente, dunque, la necessità di attivare un maggior controllo da parte delle forze dell'ordine, oltre all'attivazione di un numero maggiore di telecamere di videosorveglianza in quante più zone possibili.

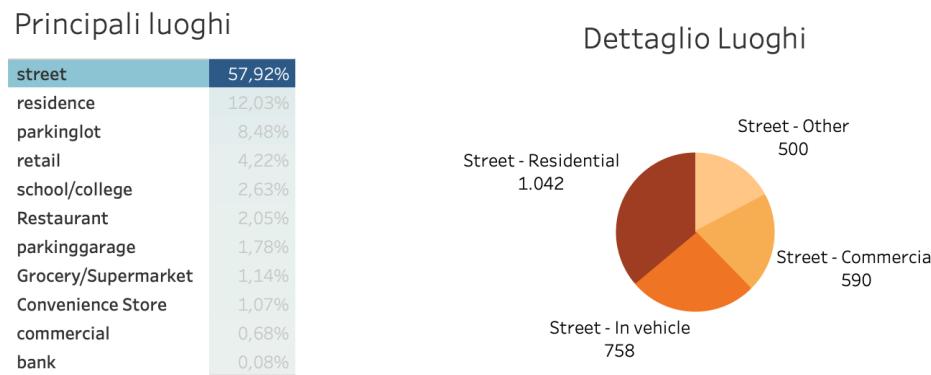


Figura 36: Tableau - Luoghi dei Crimini contro la persona

### 4.2.5 Crimini contro la società

Per terminare lo studio di analisi delle diverse tipologie di crimini, si mostra nella Figura 37 il dashboard sui crimini contro la società.

## 4.2 Analisi Effettuate

Come si evince dal grafico a barre, le violazioni di droga e narcotici seguito dalla guida sotto effetto di sostanze stupefacenti, sono le tipologie che costituiscono quasi la totalità del numero di crimini contro la società. Per quanto riguarda le fasce orarie, nel grafico immediatamente sottostante, si può notare che le violazioni di droga/narcotici ha un andamento pressochè crescente nell'arco della giornata: i crimini sono minori al mattino per poi registrare un picco nelle ore serali. Nelle ore notturne, invece, la tipologia con più grado di frequenza è la guida sotto effetto di sostanze stupefacenti.

Tali aspetti, comunque, verranno approfonditi nello studio di analisi mediante l'ultimo tool di Business Intelligence adottato, Power BI, nella prossima sezione.

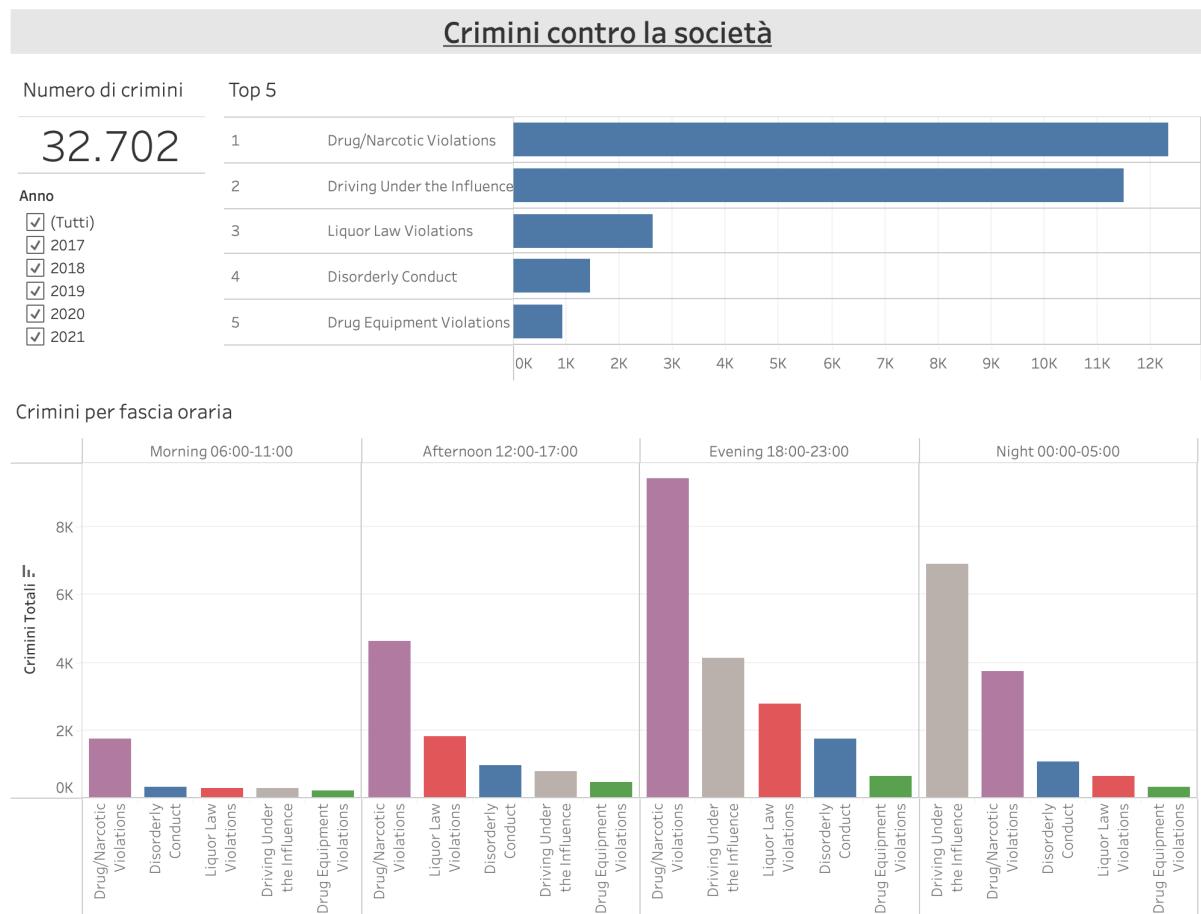


Figura 37: Tableau - Crimini contro la società

### 4.2.6 Forecasting

Tableau, inoltre, offre la possibilità di effettuare previsioni al fine di prendere decisioni affidabili. Per fare questo, il tool sfrutta la funzionalità "Previsione" che permette di stimare lungo una dimensione l'andamento di dati non disponibili, facendo uso dell'algoritmo dell'Exponential Time Smoothing.

Nel caso in questione si è deciso di stimare il numero di crimini per mese dell'anno. Per verificare la validità del sistema sulla serie temporale di interesse è stata fatta una previsione su due anni: dal 2021 al 2023, tenendo conto che tra il 2021 e il 2022 i dati sono contenuti nel dataset.

Nella Figura 38 viene mostrata la previsione, mentre nella Figura 39, costruito a partire da tutti i dati effettivamente disponibili, permette di valutare l'accuratezza. Come si può notare la previsione del primo anno risulta essere pressochè compatibile con i primi mesi dei dati effettivi, e discostandosi dalla realtà nei mesi successivi. Tali risultati, ad ogni modo, risultano pressochè ragionevoli, in quanto un modello predittivo si allena ed effettua previsioni sulla base dei dati che ha a disposizione, trascurando tutti quei fattori esterni che, in qualche modo, potrebbero influenzare in modo cruciale i risultati. In particolare, nel caso specifico di studio, Tableau non può essere a conoscenza della serie di lockdown e delle misure di

## 4.2 Analisi Effettuate

contenimento attuati per contrastare il contagio da COVID-19 e che hanno portato a una diminuzione sostanziale di numerosi crimini.

Inoltre, in entrambi i grafici è presente un filtro per particolarizzare la previsione per tipologia di crimine.

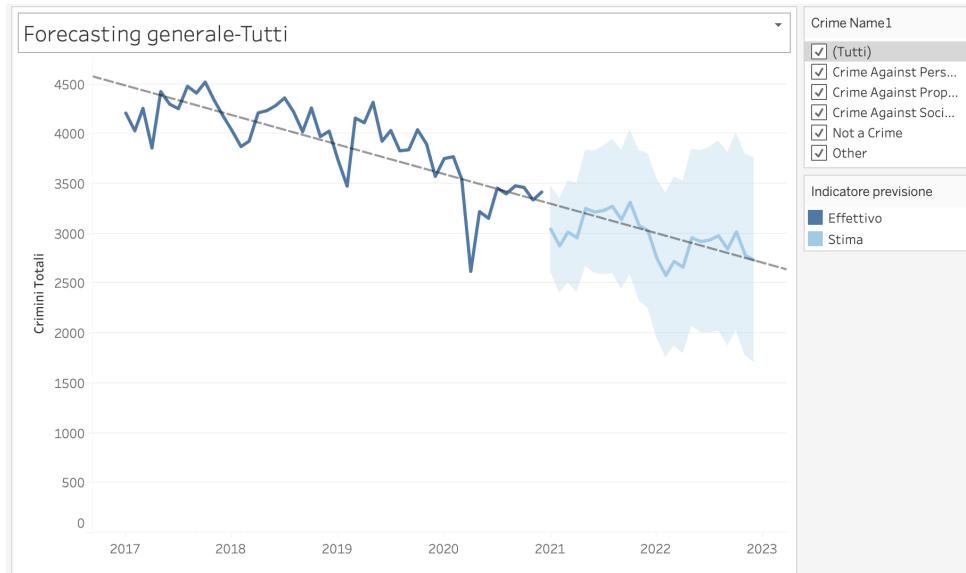


Figura 38: Tableau - Forecasting

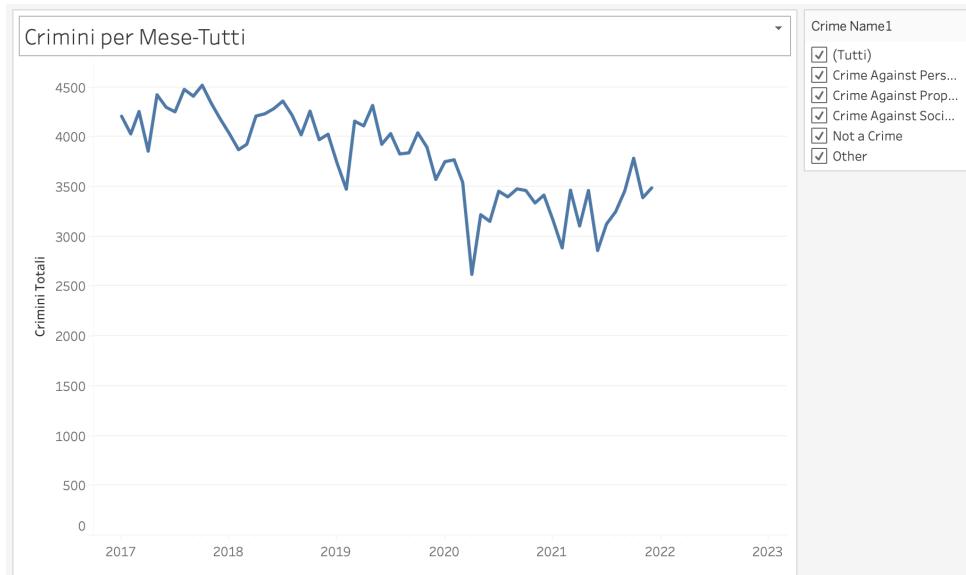


Figura 39: Tableau - Crimini Totali per Mese

## 5 Power BI

Microsoft Power BI è uno dei più importanti strumenti di Self Service Business Intelligence, volto a definire una soluzione all-in-one per l'analisi dei dati, dalla raccolta degli stessi, fino alla realizzazione di report e dashboard per monitorare l'andamento delle performance aziendali, da pc e da dispositivi mobili.



Figura 40: Logo Power BI

In questo capitolo è stato di interesse analizzare quali fossero le Agenzie di polizia maggiormente intervenute in seguito ai reati commessi nello stato del Maryland tra il 2017 e il 2021 e, successivamente, ci si è concentrati sull'analisi dei crimini legati al possesso e alla vendita di sostanze stupefacenti

### 5.1 Caricamento del dataset

Il dataset impiegato per il progetto di Power BI è lo stesso adottato per QlikSense e Tableau. In maniera analoga a quanto eseguito con gli altri due tool, la prima fase è stata quella relativa al caricamento dei dati mediante l'interfaccia messa a disposizione dal tool stesso. Inoltre, non è stata necessaria alcuna operazione di ETL e aggiunte di alcun tipo ai dati sui quali si è lavorato, in quanto l'intera fase di pulizia è stata eseguita in Python, in stato preliminare del progetto.

### 5.2 Le forze dell'ordine dello stato del Maryland

Come già analizzato in Qlik, lo stato del Maryland è costituito da 23 contee e una città indipendente (Baltimore). In seguito ad alcune ricerche si può affermare che le forze dell'ordine nel Maryland si distinguono in Agenzie statali, Agenzie delle contee e Agenzie municipali.

I reati maggiormente registrati tra il 2017 e il 2021 riguardano la contea di Montgomery, dunque, ci si aspetta che le stesse Agenzie di polizia coinvolte in tali crimini, siano situate nel suddetto territorio. Tramite il grafico a torta in Figura 41 è possibile dedurre che le quattro principali Agenzie coinvolte sono:

- *Montgomery County Police Department (MCPD)*, con sede a Gaithersburg, accreditata a livello nazionale, nonché la principale agenzia delle forze dell'ordine della contea di Montgomery;
- *Rockville City Police Department (RCPD)*, responsabile del pattugliamento della città di Rockville, la terza città più grande dello stato del Maryland;
- *Greenbelt Police Department (GPD)*, la principale agenzia delle forze dell'ordine che serve la popolazione entro 17 km<sup>2</sup> dalla città di Greenbelt;
- *Takoma Park Police Department (TPPD)*, la principale agenzia delle forze dell'ordine che serve la popolazione entro 6,5 Km<sup>2</sup> dalla città di Takoma Park.

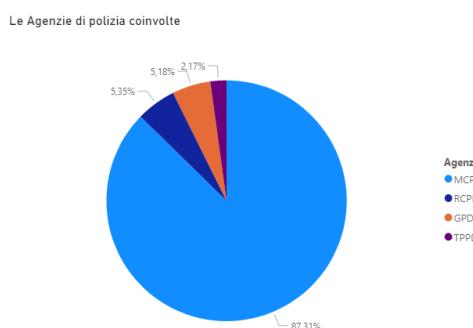


Figura 41: Le Agenzie di polizia

## 5.2 Le forze dell'ordine dello stato del Maryland

La MCPD è l'Agenzia di contea del Montgomery che, quindi, sorveglia l'intero territorio, mentre la RCPD, GPD e TPPD sono agenzie municipali che interessano, anch'esse, la suddetta contea, ma intervengono su zone più ristrette. Questo ci permette di giustificare che la MCPD è l'Agenzia che si occupa della maggior parte dei crimini analizzati, più precisamente dell'87,25% della totalità.

### 5.2.1 Le aree gestite dalle forze dell'ordine

Ciascuna agenzia di polizia ha il compito di sorvegliare ed intervenire su una determinata area che è descritta da un codice denominato "Beat", costituito da 3 valori che descrivono, in ordine:

- *il distretto di polizia*, il quale delinea i confini di azione e si divide in:
  - 1D: distretto di Rockville,
  - 2D: distretto di Bethesda,
  - 3D: distretto di Silver Spring,
  - 4D: distretto di Wheaton,
  - 5D: distretto di Germantown,
  - 6D: distretto di Montgomery Village,
  - 8D: distretto di Takoma Park;
- *il settore*, definito da una lettera dell'alfabeto, dalla A alla T, escluso Q, il quale rappresenta un sottoinsieme di ciascun distretto;
- *il numero distintivo di Beat*, il quale rappresenta un sottoinsieme di ciascun settore.

Ogni Beat, a sua volta, è diviso in "PRA" (*Police Response Area*).

Per comprendere al meglio le suddette aree, è riportata in Figura 42 la mappa geografica della contea del Montgomery divisa per distretti, settori e Beat.

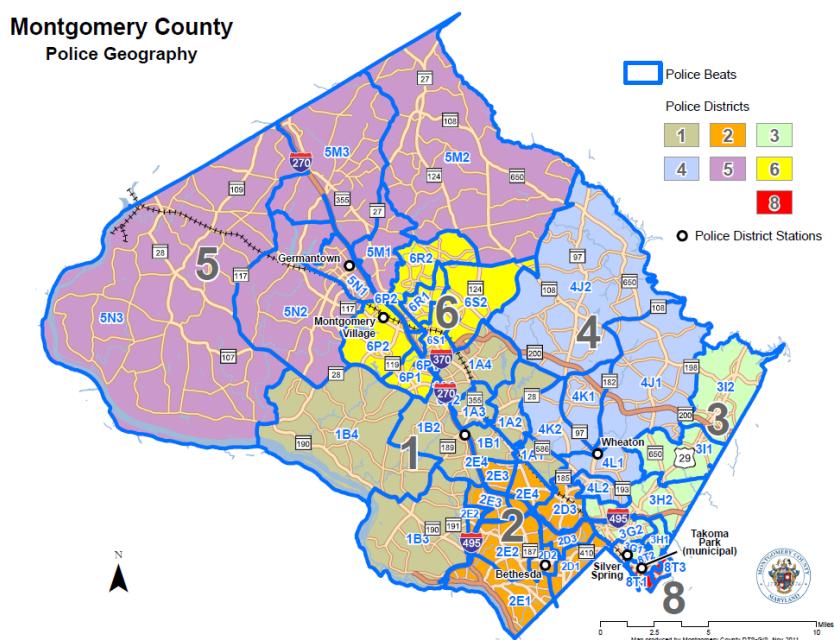


Figura 42: La mappa geografica delle aree di polizia

La MCPD, come anticipato nel paragrafo precedente, è l'Agenzia più importante per la presente analisi esplorativa. Dal grafico a barre in Figura 43 si evince che il distretto su cui tale ente è dovuto intervenire maggiormente è il 3D.

## 5.2 Le forze dell'ordine dello stato del Maryland

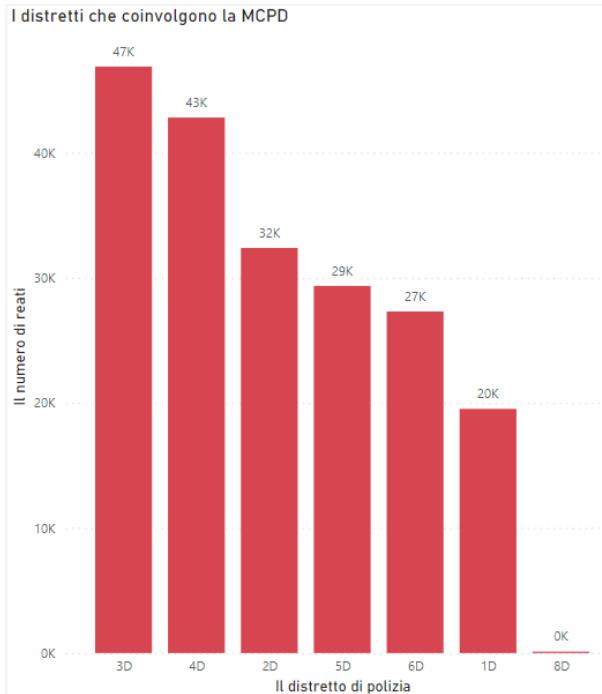


Figura 43: I distretti in cui è intervenuta la MCPD

Mentre, dal grafico in Figura 44, si deduce che la RCPD, GPD e TPPD intervengono maggiormente, rispettivamente, nel distretto 1D, 6D e 8D; tale affermazione, infatti, spiega il motivo per cui l'agenzia di contea agisce meno in corrispondenza di questi distretti.

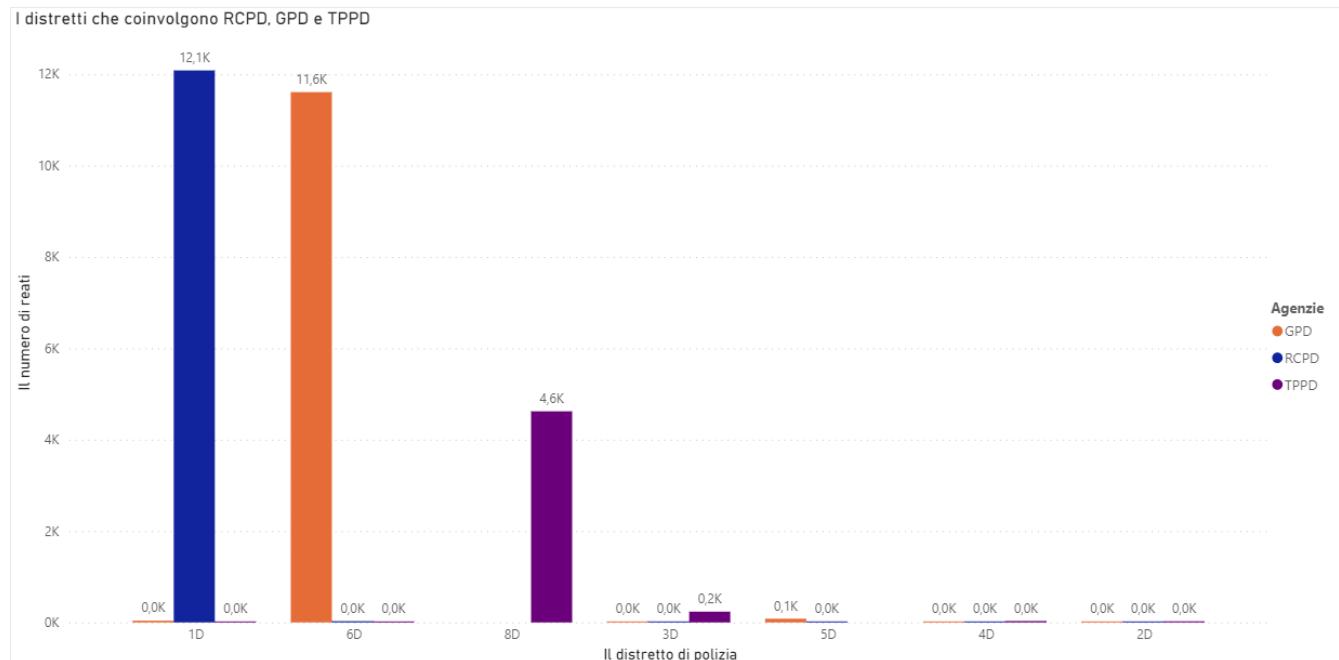


Figura 44: I distretti in cui è intervenuta la RCPD, GPD e TPPD

I PRA in cui sono state maggiormente coinvolte le diverse forze dell'ordine sono riportati in Figura 45; in particolare, nel grafico ad anello è possibile visualizzare le aree con più di 2'000 reati che coinvolgono l'agenzia di contea, invece, nel grafico ad albero è possibile visualizzare le aree con più di 500 reati che

## 5.2 Le forze dell'ordine dello stato del Maryland

coinvolgono le agenzie municipali prese in considerazione. Inoltre, da questi grafici è possibile dedurre che:

- il PRA 7020 è soggetto a frequenti interventi da parte di MCPD,
- il PRA 4190 è soggetto a frequenti interventi da parte di GPD,
- il PRA 2770 è soggetto a frequenti interventi da parte di RCPD,
- il PRA 8040 è soggetto a frequenti interventi da parte di TPPD.

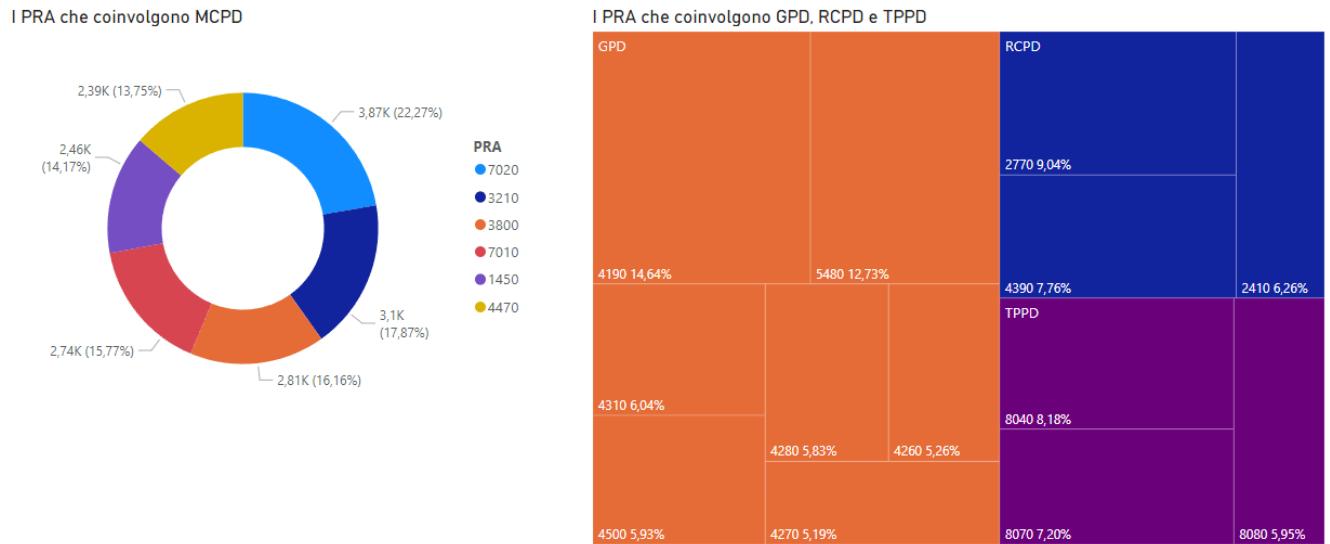


Figura 45: I quartieri in cui è intervenuta maggiormente MCPD, RCPD, GPD e TPPD

Per concludere, sono stati mappati, tramite l'ausilio di cerchi colorati, i PRA appena elencati per comprendere al meglio le zone coinvolte e, le diverse dimensioni di tali cerchi, permettono di capire la differente concentrazione di crimini nella specifica zona.

In Figura 46 è riportata l'area corrispondente al PRA 7020 che si colloca nella città di Germantown (distretto 5D).

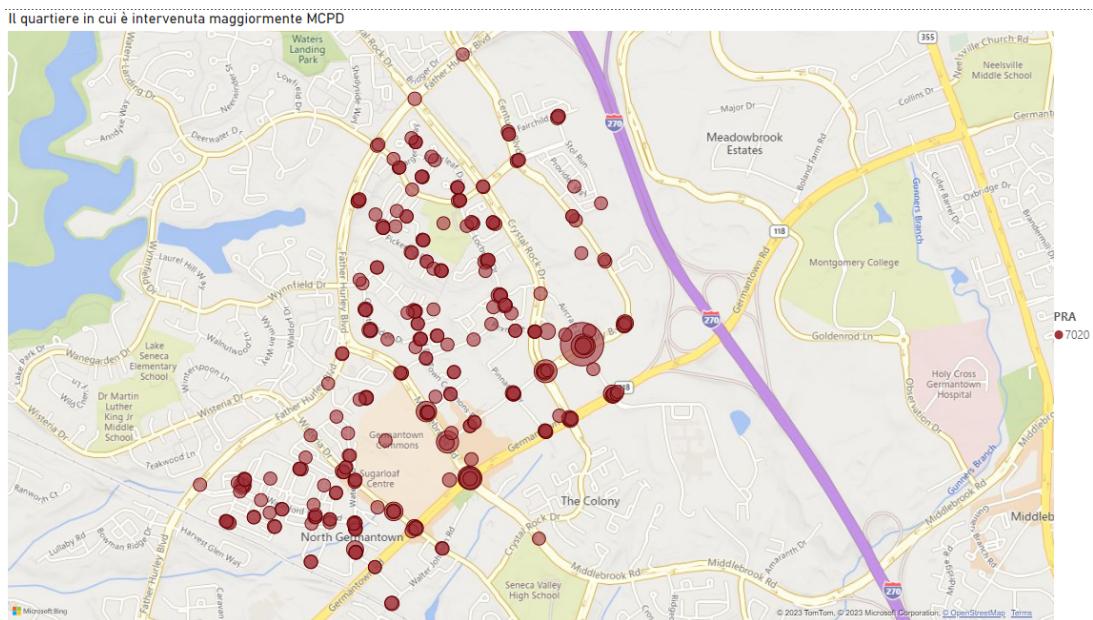


Figura 46: L'area in cui è intervenuta maggiormente MCPD

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

Questo territorio si pone, prevalentemente, ad ovest della *Germantown Road*, l'autostrada statale principale che attraversa la suddetta città e ad ovest dell'*Interstate 270*, l'autostrada interstatale ausiliaria che collega la contea di Frederick con la contea del Montgomery. Si nota una maggiore distribuzione di reati (n°418) in corrispondenza della strada *Century Boulevard*, che si colloca nei pressi della *Germantown Police Station*, motivo per cui, probabilmente, si presta ad una così elevata concentrazione di casi.

In Figura 47 sono riportate le aree corrispondenti, rispettivamente, ai PRA 4190, 2770 e 8040.

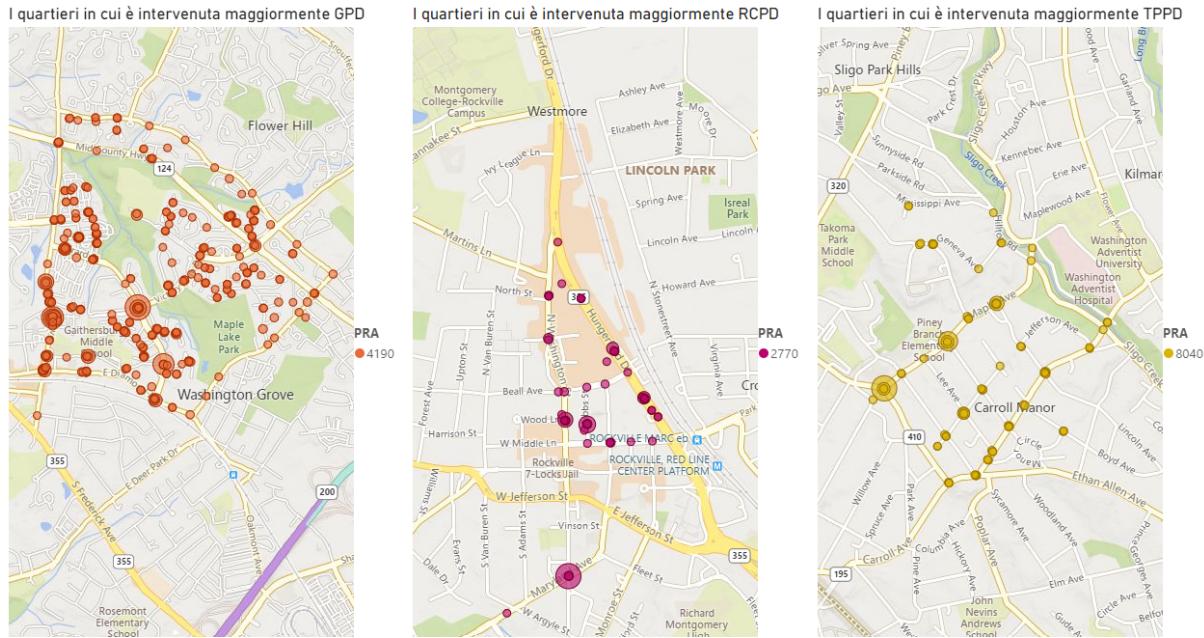


Figura 47: L'area in cui sono intervenute maggiormente GPD, RCPD e TPPD

Nella prima immagine si riporta il PRA in cui è intervenuta maggiormente la GPD e questa si colloca nella città di Gaithersburg (distretto 6D); in particolare, tale area si sviluppa intorno ai parchi di *Maple Lake*, *Kelley*, *Victory Farms Stream Valley* e *Emory Grove Hills*, che insieme costituiscono un'ampia zona verde nella suddetta città. Inoltre, questa zona è caratterizzata dalla presenza della *Gaithersburg Elementary School*, *Gaithersburg Middle School* e *Forest Oak Middle School*. Nella seconda immagine si riporta il PRA in cui è intervenuta maggiormente la RCPD e questa si colloca nella città di Rockville (distretto 1D); quest'area si sviluppa, prevalentemente, sulle strade principali che conducono alla *Maryland Route 355* che funge da arteria principale attraverso le contee di Frederick e Montgomery e che durante il suo percorso cambia nome, infatti, in questo caso, il suo appellativo diventa *Hungerford Drive*. Si nota, però, una particolare concentrazione di casi (n°264), in corrispondenza della *Maryland Ave*, nei pressi della *Christ Episcopal Chuch* e dell'ufficio di amministrazione comunale di Rockville, la *Rockville City Hall*. Infine, nell'ultima immagine, si riporta il PRA in cui è intervenuta maggiormente la TPPD e questa si colloca nella città di Takoma Park (distretto 8D); questo quartiere è circoscritto dalle strade di *Carroll Ave*, *Philadelphia Ave*, *Maple Ave* e la strada panoramica *Sligo Creek Parkway* ed è caratterizzata dalla presenza della scuola elementare *Pinery Branch* ed un centro di ritrovo per giovani e anziani denominato *Takoma Park Community Center*.

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

Una volta analizzate le forze dell'ordine coinvolte nei reati commessi nel Maryland, si è scesi maggiormente nel dettaglio e sono stati analizzati i crimini legati alle sostanze stupefacenti. Per svolgere le seguenti analisi, è stata presa in considerazione la colonna di "Crime Name3", la quale rappresenta una particolarizzazione dei reati descritti in "Crime Name2" e "Crime Name1". Sono stati selezionati tutti gli elementi del suddetto campo che presentavano la parola *Drugs* o *Overdose*.

In Figura 48 è riportato, sulla sinistra, il numero totale di crimini legati alla droga ed il suo valore ammonta a quasi 22'000, mentre, sulla destra, è riportato un grafico ad imbuto in cui si osserva l'elenco

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

dei crimini con la relativa percentuale. In seguito ad un'attenta visione del presente grafico, si evince che in tutto sono riportati 38 crimini e le illegalità maggiormente commesse sono il possesso e la vendita di droga, di cui si tratterà nei prossimi sottoparagrafi.

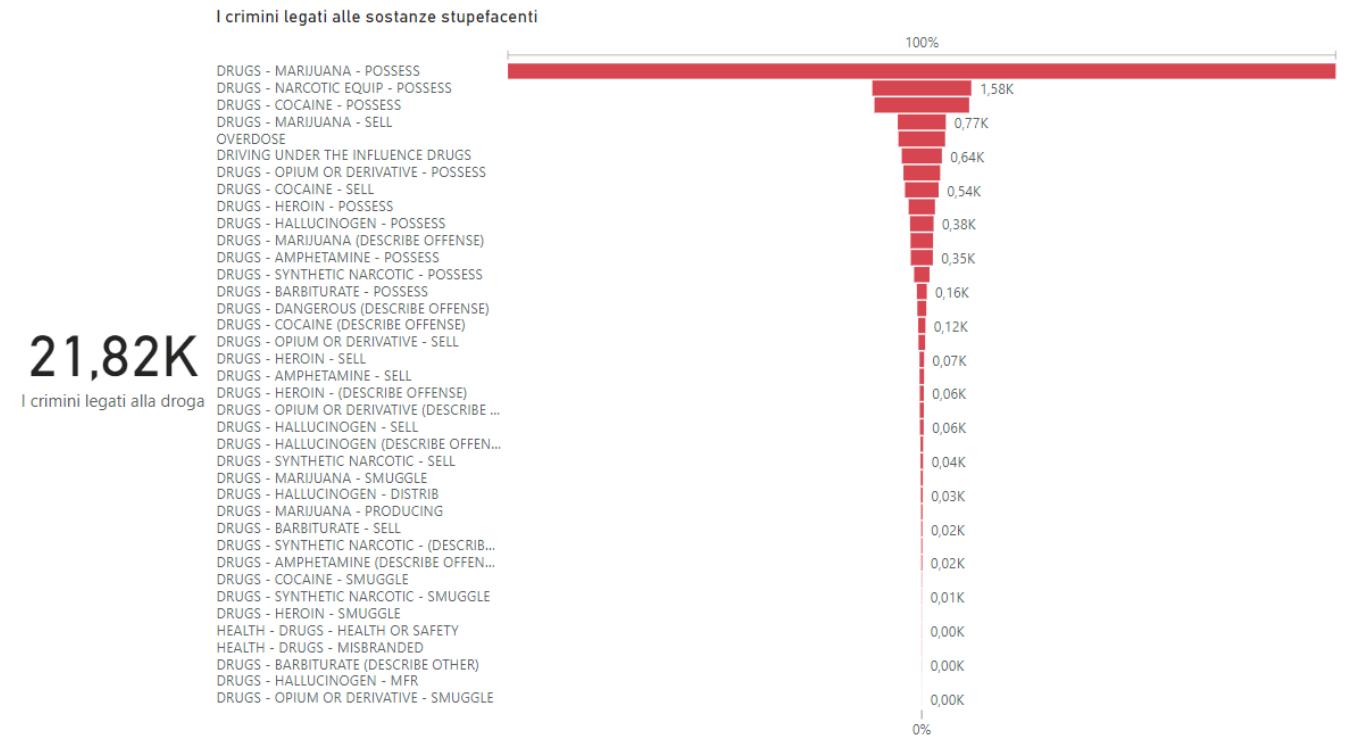


Figura 48: I crimini legati alle sostanze stupefacenti

#### 5.3.1 Le sostanze stupefacenti possedute

In questa sezione è riportata l'analisi svolta sul possesso di sostanze stupefacenti nel Maryland, tra il 2017 e il 2021. In Figura 49 è riportata un dashboard in cui sono sintetizzate le principali analisi.

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

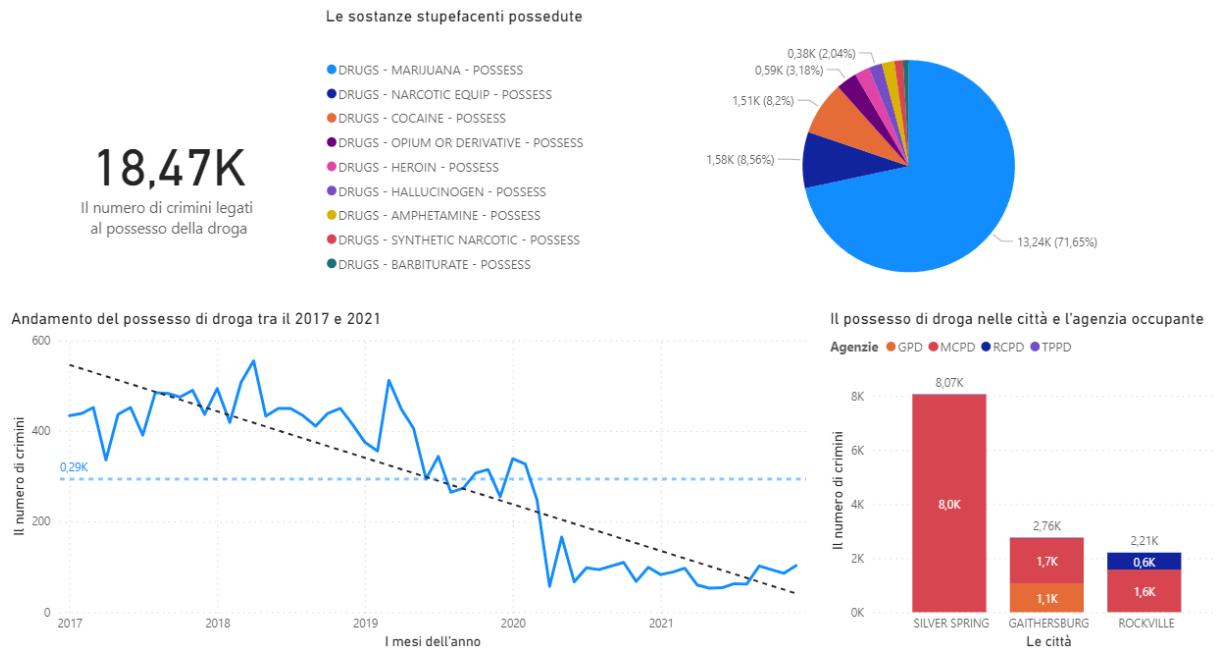


Figura 49: Le sostanze stupefacenti possedute

Il numero di crimini è pari a circa 18'000, quindi, all'85% del totale dei crimini commessi in questo ambito; in particolare, si registra che il 71,65% della droga posseduta riguarda la Marijuana, mentre il restante 28,35% riguarda il possesso di: attrezzatura per produrre, nascondere e consumare droghe illecite (8,56%), cocaina (8,2%), oppio e derivati (3,18%), eroina (2,28%), allucinogeni (2,04%), amfetamina (1,91%), droghe sintetiche (1,33%) e barbiturici (0,86%). Nel grafico in basso a sinistra, si può osservare l'andamento nel tempo del possesso di droga; sull'asse delle x sono riportati i mesi dell'anno in cui sono stati registrati i crimini, mentre sull'asse delle y è riportato il numero di crimini legato alla vendita di droga. Il valor medio dei crimini associati a tale ambito risulta pari a circa 290 e si registrano due picchi in corrispondenza del 1° Aprile 2018 (555 crimini) e del 1° Marzo 2019 (512 crimini); si può notare, però, che il trend sia in calo, in particolare tra 1° Marzo 2020 e il 1° Aprile 2020, in cui si registrano solo 57 crimini e da quella data in poi il numero pare stabilizzarsi, rimanendo sempre al di sotto di 200 crimini. Questo evidente calo molto probabilmente è legato alle restrizioni dovute alla pandemia di COVID-19, perché proprio verso inizio marzo di quell'anno si iniziarono a registrare i primi casi nella contea del Montgomery e, con il veloce aumento di questi, seguirono, verso metà del mese, restrizioni che portarono alla chiusura di spazi aperti al pubblico. Lo stato di emergenza COVID-19 del Maryland terminò il 1° Luglio 2021. Tutto questo potrebbe aver reso più difficoltoso la reperibilità delle sostanze stupefacenti, ma allo stesso tempo si potrebbe esser verificato una inevitabile diminuzione dei controlli da parte delle forze dell'ordine. Infine, nel grafico in basso a destra è riportato un istogramma in cui sono evidenziate le città maggiormente colpite da questo genere di crimine e le agenzie di polizia maggiormente intervenute. E' evidente come Silver Spring sia la città con il maggior numero di casi di possesso di droga e che l'Agenzia di contea sia stata la principale forza dell'ordine intervenuta sullo stesso territorio.

In Figura 50 sono riportati i grafici precedentemente illustrati, ma particolarizzati per il possesso di Marijuana.

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

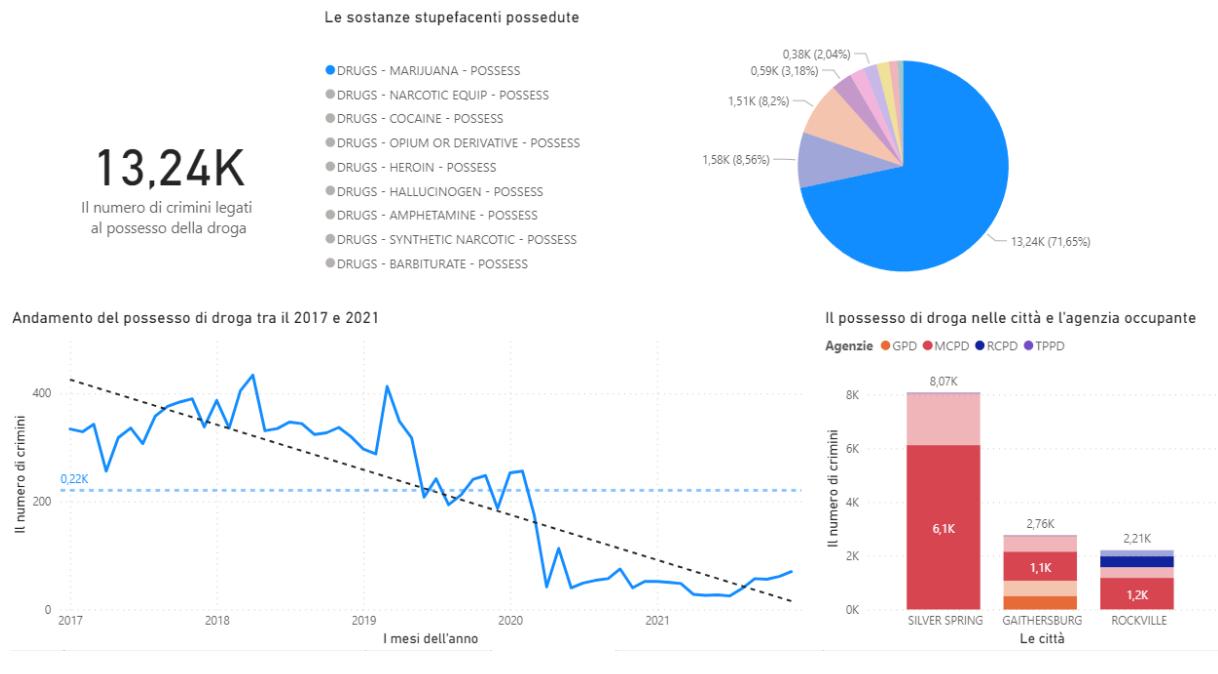


Figura 50: La marijuanna posseduta

Si registrano circa 13'000 casi legati a tale crimine e l'andamento nel tempo della sua fruizione pare molto simile a quello generale; inoltre, il 75% dei crimini commessi a Silver Spring, riguardano il possesso di Marijuana e, come ci si potrebbe aspettare, la MCPD è la principale agenzia ad occuparsene.

Infine, come ultima analisi riguardante tale ambito, è stata effettuata la previsione del possesso di sostanze stupefacenti nel 2022. Come si può osservare in Figura 51, la previsione è stata effettuata per 12 mesi e i risultati sembrano rispettare il trend in calo, in particolare, i reati non sembrano superare i 150.

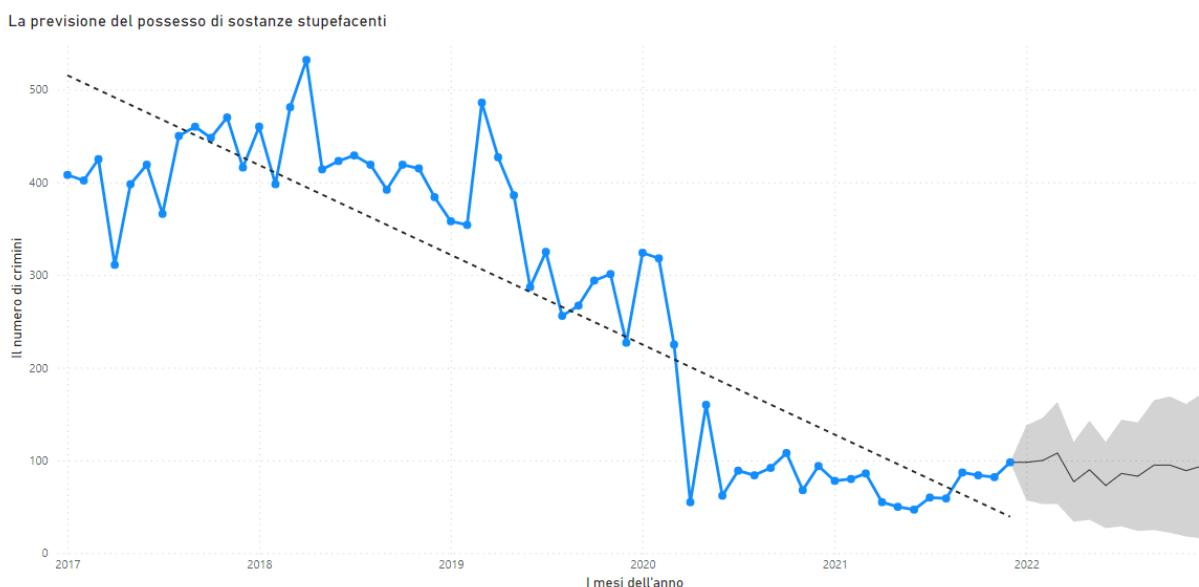


Figura 51: La previsione del possesso di sostanze stupefacenti

Potrebbe essere interessante comprendere se tale andamento, ereditato sicuramente dall'impatto del COVID-19, sia dovuto ad una reale diminuzione del possedimento di droga oppure sia dovuto a minori

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

controlli da parte delle forze dell'ordine. Dunque, se la reale causa fosse attribuita alle agenzie di polizia, questa previsione sarebbe potuta essere un ottimo movente per aumentare le ispezioni.

#### 5.3.2 Le sostanze stupefacenti vendute

Dopo aver analizzato i crimini legati al possesso di sostanze stupefacenti, in questa sezione è riportata l'analisi svolta sulla vendita. In Figura 52 è riportato, sulla sinistra, il numero di crimini legati alla vendita della droga, mentre, sulla destra, l'albero di scomposizione, grazie al quale è possibile effettuare l'operazione di drill-down per risalire, dalla sostanza stupefacente più distribuita nel Maryland, ad una zona specifica in cui le forze dell'ordine sono intervenute maggiormente a causa di tale crimine.

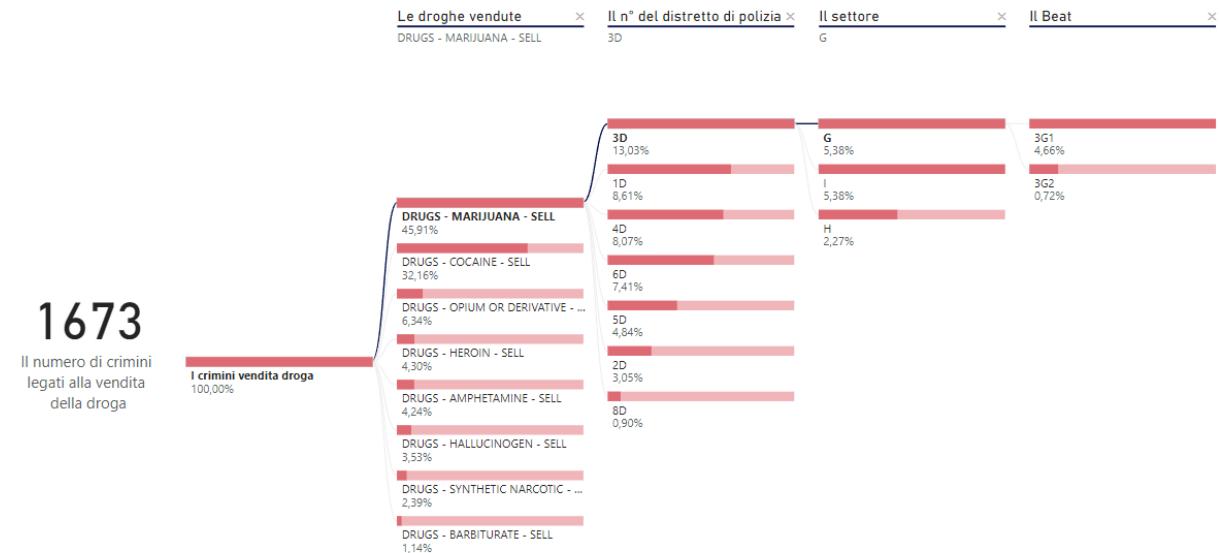


Figura 52: Le sostanze stupefacenti vendute

In totale i crimini commessi sono 1'673 e la Marijuana risulta essere la droga più soggetta al commercio illegale (45,91%); il distretto in cui è stato prevalentemente sventato tale crimine è il 3D e i settori di tale distretto maggiormente coinvolti risultano G e I, entrambi con una concentrazione pari a 5,38%. Prendendo come riferimento la mappa in Figura 42 è possibile dedurre che il settore G corrisponde alla città di Silver Spring, la stessa che registrava il maggior numero di crimini riguardante il possesso di marijuana, mentre il settore I corrisponde ad un'area più periferica, a nord-est rispetto alla stessa città. Dunque, si è preferito prendere in considerazione il settore G per valutare quale fosse il Beat con la maggior concentrazione di casi; il Beat in questione è il 3G1. In Figura 53 è riportata la mappa di tale zona e le agenzie di polizia intervenute sono la MCPD e la TPPD.

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

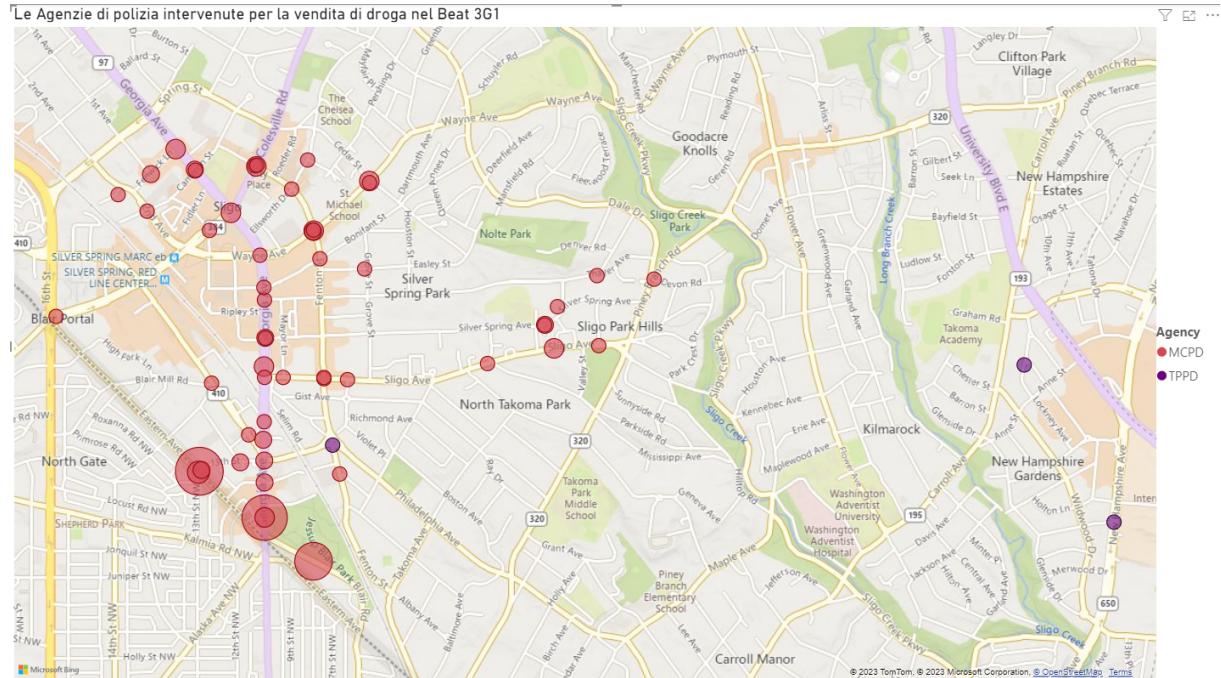


Figura 53: Le Agenzie di polizia che si sono occupate del crimine legato alla vendita di Marijuana

Quest'ultima è intervenuta solo per pochi crimini nella città di Takoma Park, che in generale è localizzata nel settore T, ma probabilmente, quelle singole aree evidenziate in viola non rientravano nei PRA del settore T. Si può osservare una maggiore concentrazione in corrispondenza della *Eastern Avenue* che rappresenta una delle tre strade principali che segna il confine dello stato del Maryland con Washington, DC. Per comprendere meglio i luoghi in cui sono avvenuti questi scambi, si è voluto approfondire mappando le stesse aree precedentemente illustrate, ma utilizzando come legenda il campo "filtered\_place" (Figura 54).

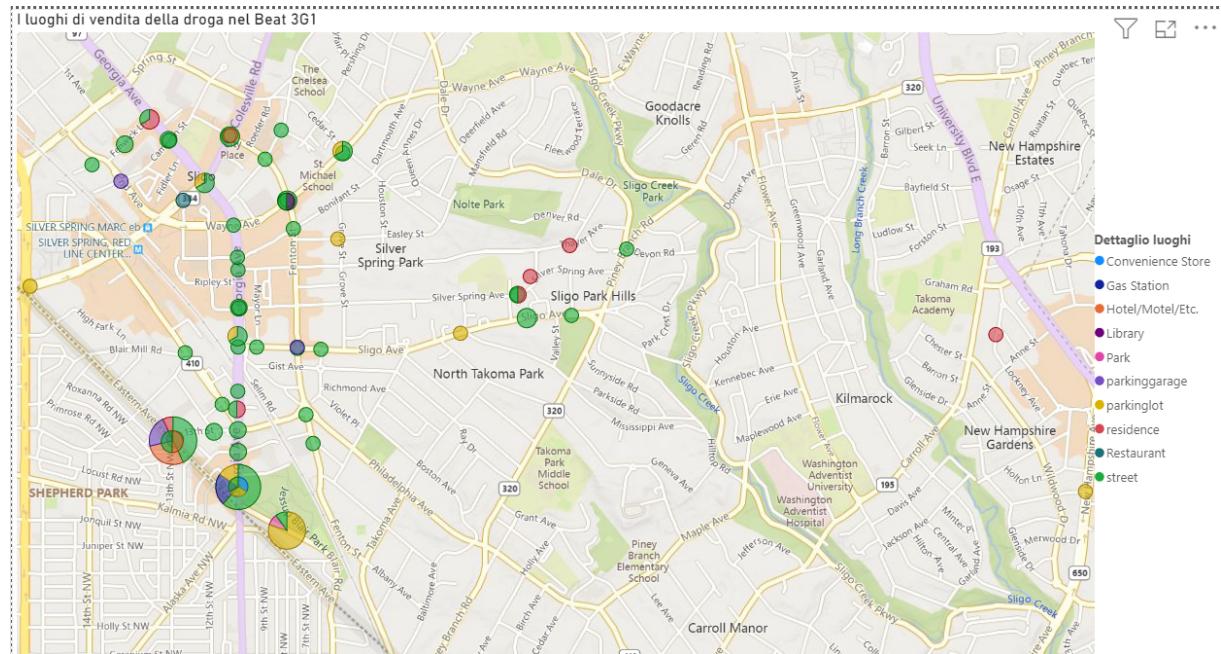


Figura 54: I luoghi in cui è stata venduta la Marijuana

In generale, si evidenzia una notevole concentrazione di casi in strada, in particolare sulla *Georgia Avenue*

### 5.3 L'analisi dei crimini legati alle sostanze stupefacenti

*Avenue*, un'importante arteria nel nord-overst di Washington, DC, che attraversa la contea di Montgomery e Silver Spring. Mentre, sulla *Eastern Avenue* è presente una maggiore varietà di luoghi interessati; è evidente che a questi luoghi corrispondono le stesse coordinate geografiche e questo fa presupporre che la loro localizzazione sia pressochè indicativa e quindi non esatta.

Infine, come ultima analisi, è stata effettuata la previsione della vendita di Marijuana nel 2022, illustrata in Figura 55, dove sull'asse delle x sono riportati i mesi dell'anno in cui sono stati registrati i crimini, mentre sull'asse delle y sono riportati il numero di crimini legati alla vendita di marijuana.

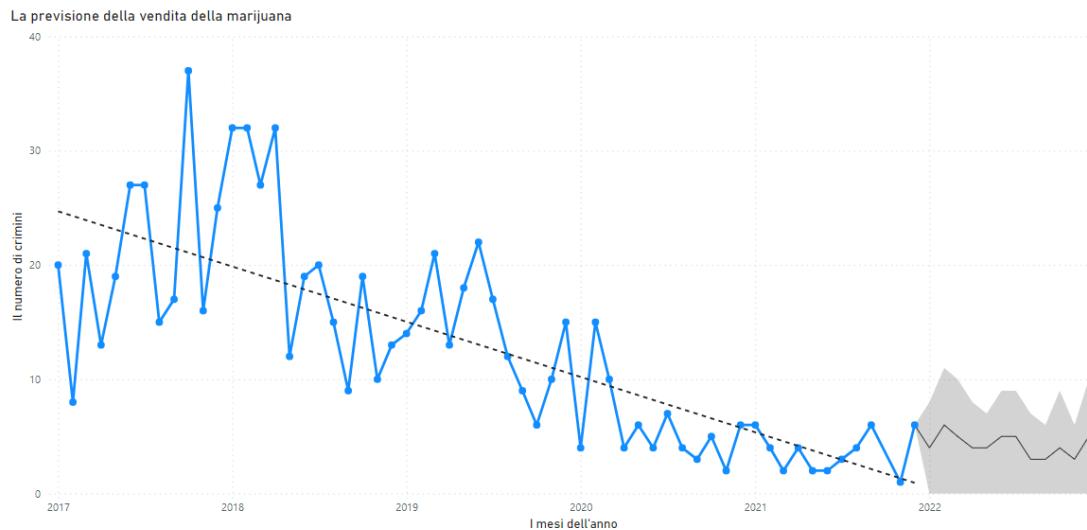


Figura 55: La previsione della vendita di Marijuana

Il trend, tra il 2017 e il 2021, è in calo e i valori risultano diminuire soprattutto in corrispondenza dell'anno 2020 che, come anticipato nel paragrafo precedente, era stato soggetto alla pandemia da COVID-19. Si prevede che nel 2022 i casi mantengano l'andamento ereditato dall'anno precedente. In seguito ad alcune ricerche si è scoperto che proprio il 2022 è stato l'anno in cui hanno approvato che, a partire dal 1° luglio 2023, la cannabis sarà legalizzata. Il Maryland, dunque, diventerà il 20° stato americano a legalizzare la marijuana e questo fa presupporre che la sua vendita illegale diminuirà sempre di più.