KEYSTATS

Managing Crime

Strategy and Optimization through Analytics

Prepared by KEYSTATS, Inc. for the New York City Police Department

Table of Contents

Section I: Introduction

Introduction

This document represents a proposal for Managing Crime in the 5 Boroughs and 77 Precincts of New York City through:

1. Leveraging and linking past system wide data (*i.e. CompStat, NYPD, NYC Criminal Justice Agency*) to better understand characteristics associated with crime.

2. Measurable and "trackable" methods to understand the impact of efforts undertaken to address crime, and refine current initiatives as required.

3. Incorporating results of current initiatives into the development of refined strategies and tactics (*Feedback Learning cycle*)

Introduction (*cont*)

The proposed framework for Managing Crime consists of:

- Understanding characteristics associated with Crime (*Murder, Rape, Robbery, Fel. Assualt, Burglery, Gr. Larceny, G.L.A., Petit Larceny, Miscd. Assualt, Misd. Assault, Misd. Sex Crimes, Shooting Vic., Shooting Inc. – from Police Dept, City of New York, CompStat, Weekly Report, Vol. 25, No. 5*) inclusive of data available agency wide at a borough, precinct , or operational level.

- Determining future Crime Activity – level and frequency at an operational level.

- Assessing required resources for minimizing Crime Activity by looking at past policies on Crime Reduction and their subsequent impacts – inclusive of Type of Effort, level of spending, and other relevant factors impacting Crime Reduction.

- Determining best allocation of resources available (*manpower, funding, etc.*) to reduce crime to an effective level without wasting dollars.

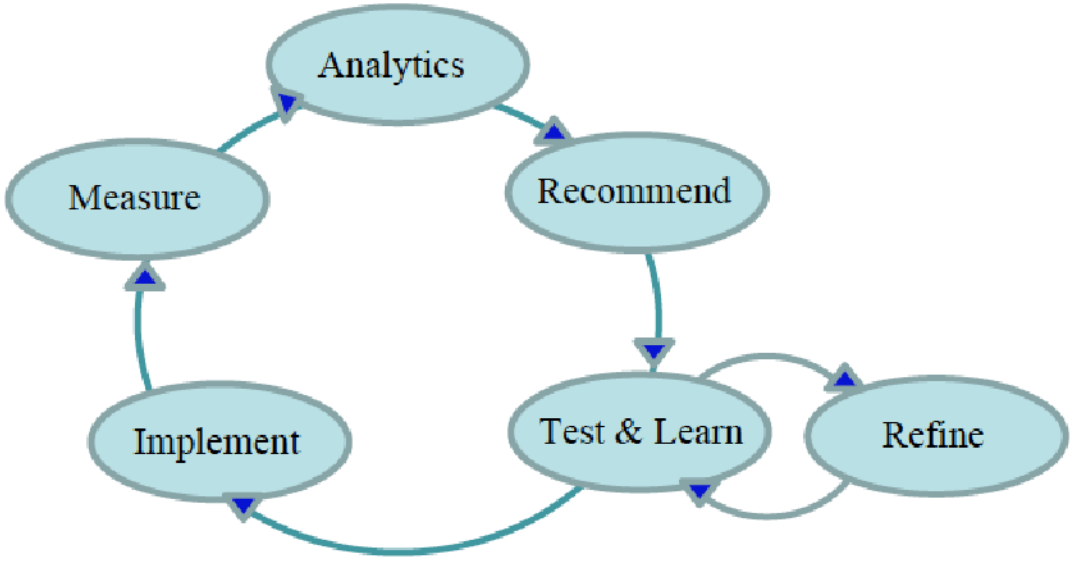
Data Sources

The basis for understanding crime and allocating resources accordingly depends on available data from a variety of sources.

- Crime data provides insight into:
  - Where crime is likely to occur
  - Who is likely to commit the crime
  - When the crime is likely to occur (time *of year, seasonality*)
  - What type of crime is likely to occur
- Crime data consists of linking together:
  - Past criminal activity (by region)
  - Past criminal activity from various independent sources:
    - Court databases (sentencing, level & severity of crime)
    - Past criminal records
    - Other available sources (DMV, and transactional records, e.g., Credit Card, *Bank Accounts, etc.*)

- **Optimal Allocation of resources depends on linking Crime Activity with Tactics deployed by NYPD primarily in terms of:**
  - Funding By Various Initiatives
  - Manpower

Operational Model for Analytics



- The strategy for analysis consists of a "test-and-learn" approach which is constantly being reviewed at the analytics stage as recommendations are executed.
- The foundation is built on a constant feed of new learnings to fuel the analytics
- The basis is measurement of tactics at an operational level.

Illustrative

Background and Objective

- NYPD is currently executing proactive strategies to combat violent and non violent crimes in the borough of Brooklyn
- Crime rates on average are the highest in Brooklyn as compared to all other boroughs over the last 18 months
- NYPD wants to reduce crime rates in most categories for borough of Brooklyn by 6% to fall in line with the other boroughs. This will be accomplished by:

1. Detailed profile analysis of crime activity in Brooklyn vs. other boroughs
2. Statistical modeling to predict crime likelihood and timing (*by crime category*) at precinct , zip code, or individual level.
3. Forecasting crime rate for subsequent 18 month period along with forecasting cost and mix of resources required.
4. Formulate and prioritize crime management strategies based on analysis.

Objective:

Reduce Category Crime rates in Borough of Brooklyn

Actively deploy Policing strategy based on results from analysis

6% Rate Reduction

Illustrative

Overall Profile of Crime in the 5 Boroughs

NYPD Crime Category	Borough					
	Bronx	Brooklyn	Manhattan	Queens	Staten Island	Total
All Crimes Across Boroughs	21.4%	29.7%	24.7%	19.9%	4.3%	100.0%
Crimes Within Boroughs						
Murder	0.1%	0.1%	0.0%	0.1%	0.0%	--
Rape	0.2%	0.3%	0.2%	0.3%	0.2%	--
Robbery	3.5%	3.9%	2.3%	3.4%	2.1%	--
Fel. Assault	4.3%	3.9%	2.3%	3.4%	2.6%	--
Burglary	2.3%	3.4%	2.0%	4.1%	2.7%	--
Gr. Larceny	4.8%	7.3%	11.6%	7.3%	5.0%	--
G.L.A.	1.1%	1.5%	0.5%	2.1%	0.9%	--
Petit Larceny	11.7%	13.4%	18.9%	13.8%	15.2%	--
Misd. Assault	10.6%	9.7%	6.8%	9.6%	9.0%	--
Misd. Sex Crimes	0.6%	0.7%	0.7%	0.8%	0.4%	--
Other-Non Seven Major Felony Offenses	9.5%	10.4%	8.8%	10.4%	9.6%	--
Other-Misdemeanor	41.2%	34.8%	37.4%	33.7%	35.8%	--
Other-Violation Offenses	10.0%	10.7%	8.5%	11.3%	16.5%	--
Total	100.0%	100.0%	100.0%	100.0%	100.0%	--

Section II: Discriptive Analytics

Descriptive Methodology

- Our descriptive methodology is data intensive and involves following the stepped approach described below:

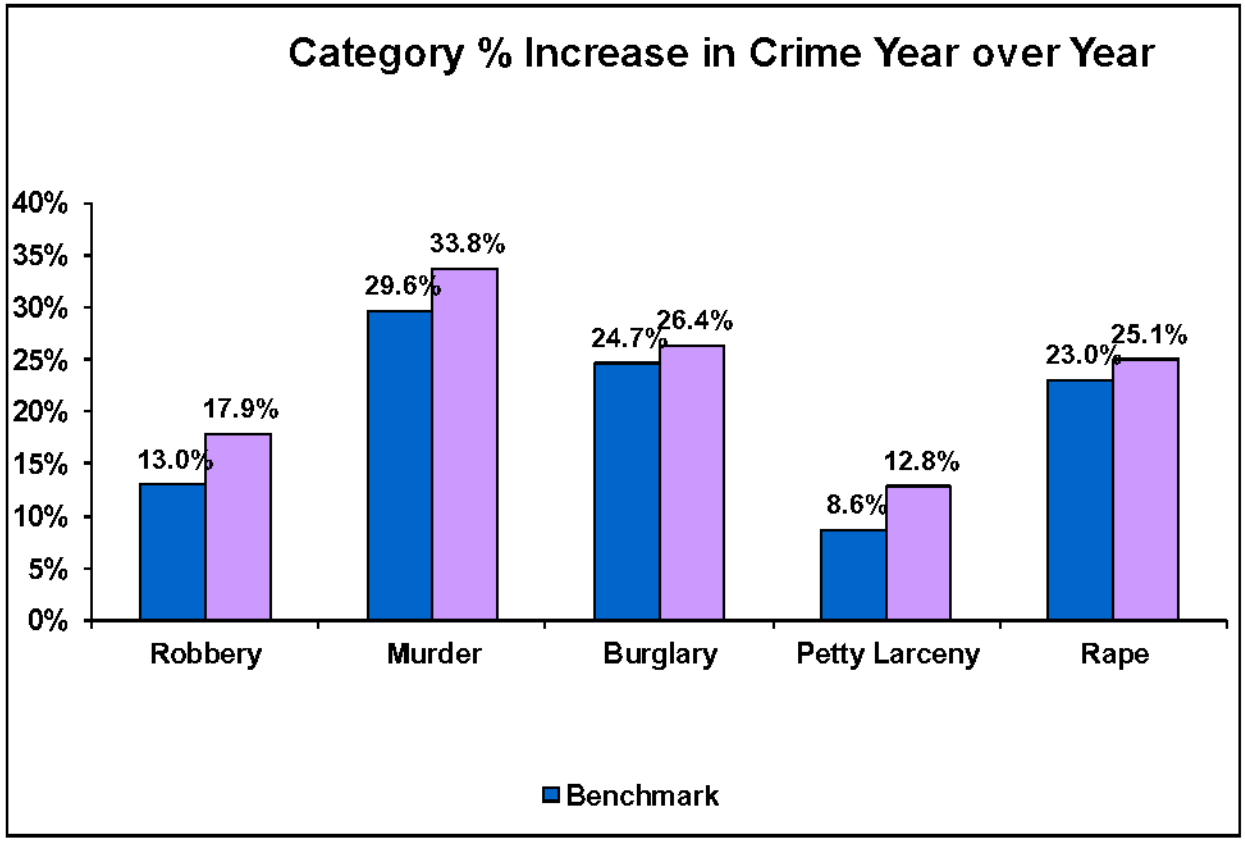
Step 1	Step 2	Step 3
Gain Access to all Data Sources	**Develop valid benchmark for baseline KCI* comparison**	**Execute Profile Analysis**
• Evaluate all available data source documentation • Finalize & Prioritize relevant data sources (*Court Records, Past Crime Data indicative of Level & Severity of Crime*) • Gain access to data	• Create weighted stratified samples for each borough using several demographic data points (*i.e. gender, age, education*) • Combine each borough sample to form the baseline benchmark sample	• Identify KCI's of interest • Calculate magnitude of KCI gap versus the benchmark • Prioritize KCI's based on gap size, estimated resource requirement, ease of implementation

* KCI – Key Crime Indicators

Illustrative

Crime Category Rate of Change

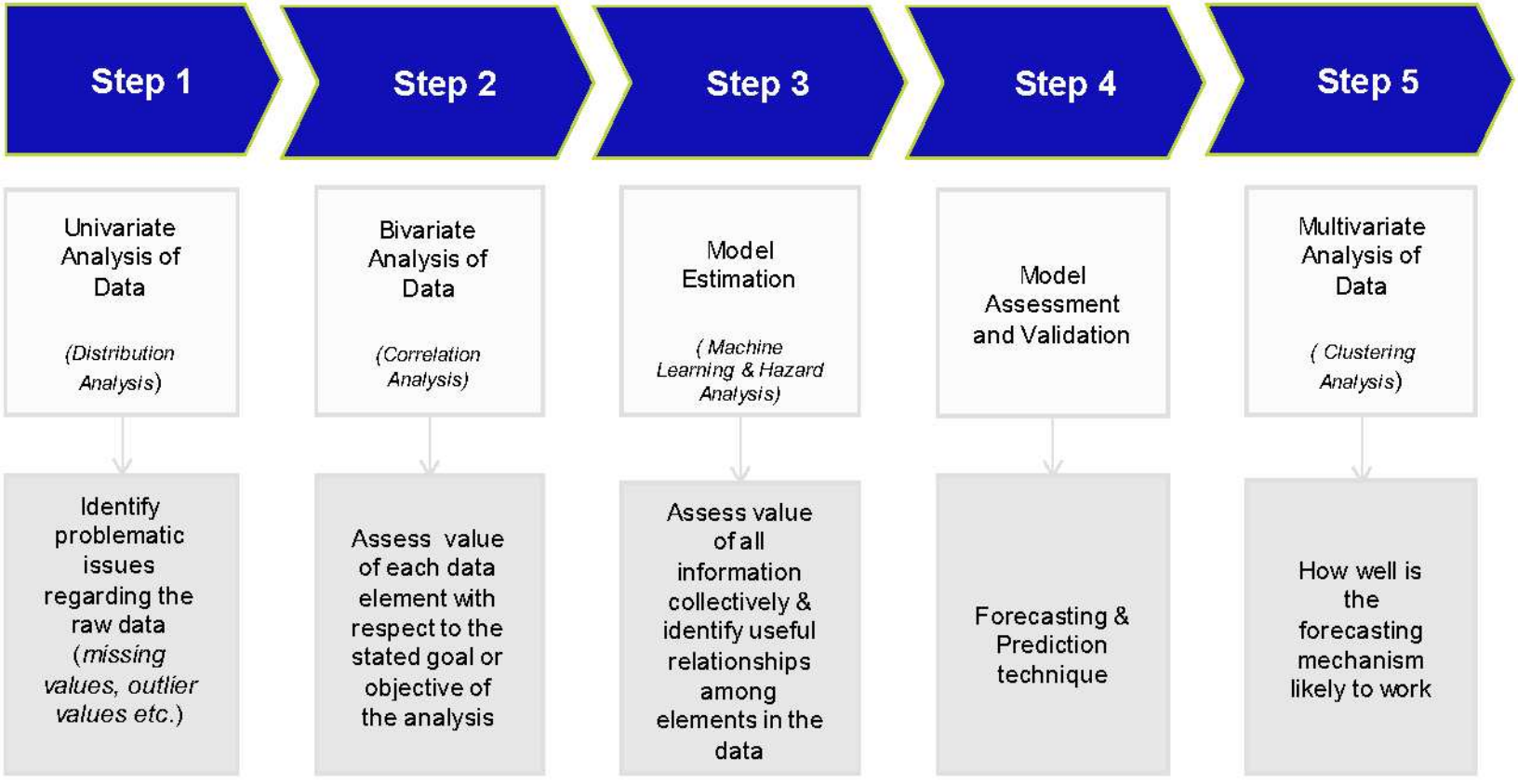
- Brooklyn has higher % increase in crime across five targeted categories , particularly *Petty Larceny* where there is an almost 50% increase compared to benchmark over a 12 month time period.



Section III: Predictive Analytics (*Statistical Modeling*)

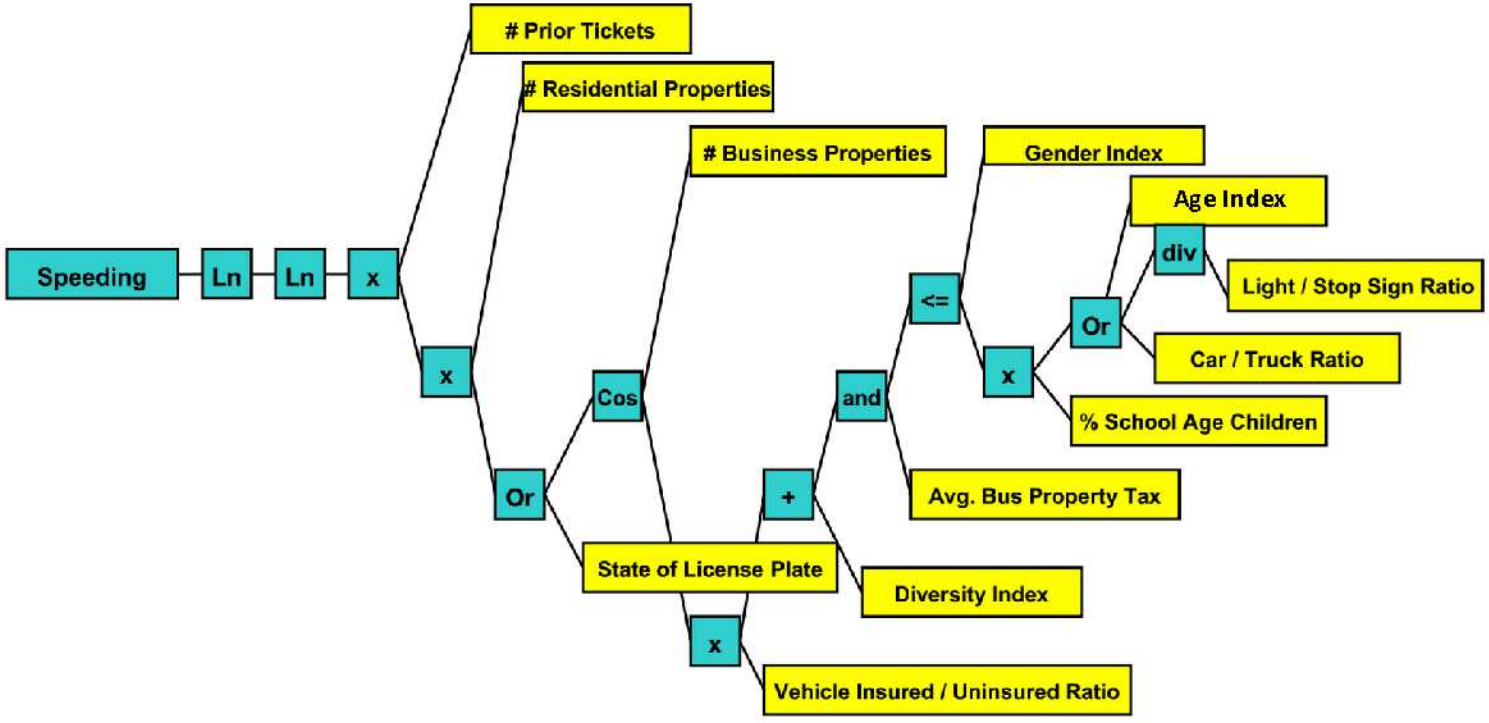
Statistical Methodology

- Our approach to developing a successful prediction model for NYPD in borough of Brooklyn is based on the five step approach described below:



Illustrative

Speeding Prediction Model via Genetic Algorithms



Genetic Algorithms fall under the umbrella of Artificial intelligence. The method uncovers complicated relationships in the data that are highly related to predicting speeding. The model produces a prediction probability at measurement level.

Illustrative

Speeding Prediction Model – Key Factors

- A plus sign indicates a positive relationship with likelihood to speed whereas a minus sign indicates a negative relationship with likelihood to speed

Top Predictive Variables	Sign	Importance
# Previous Tickets	+	20.69%
State of License Plate	+	19.36%
Vehicle Insured to Uninsured Ratio	-	15.03%
Car to Truck Ratio	+	9.84%
# Residential Properties in Area	-	7.84%
Age Index	-	5.91%
Gender Index (Male)	+	5.60%
Time of Day Ticket Issued	-	3.72%

Illustrative

Speeding Prediction Model – Key Factors (cont)

- A plus sign indicates a positive relationship with likelihood to speed whereas a minus sign indicates a negative relationship with likelihood to speed

Top Predictive Variables	Sign	Importance
# Business Properties in Area	-	3.69%
Stop Light to Stop Sign Ratio	-	3.36%
% School age Children	-	3.03%
Avg. Business Property Tax	-	3.00%

Predictive variables at above 95% significance level

Illustrative

Speeding Prediction Model – Benefit

- Model shows effectiveness in its ability to discriminate between zip codes that have a high level of speeding drivers and those that do not have a high level of speeding drivers

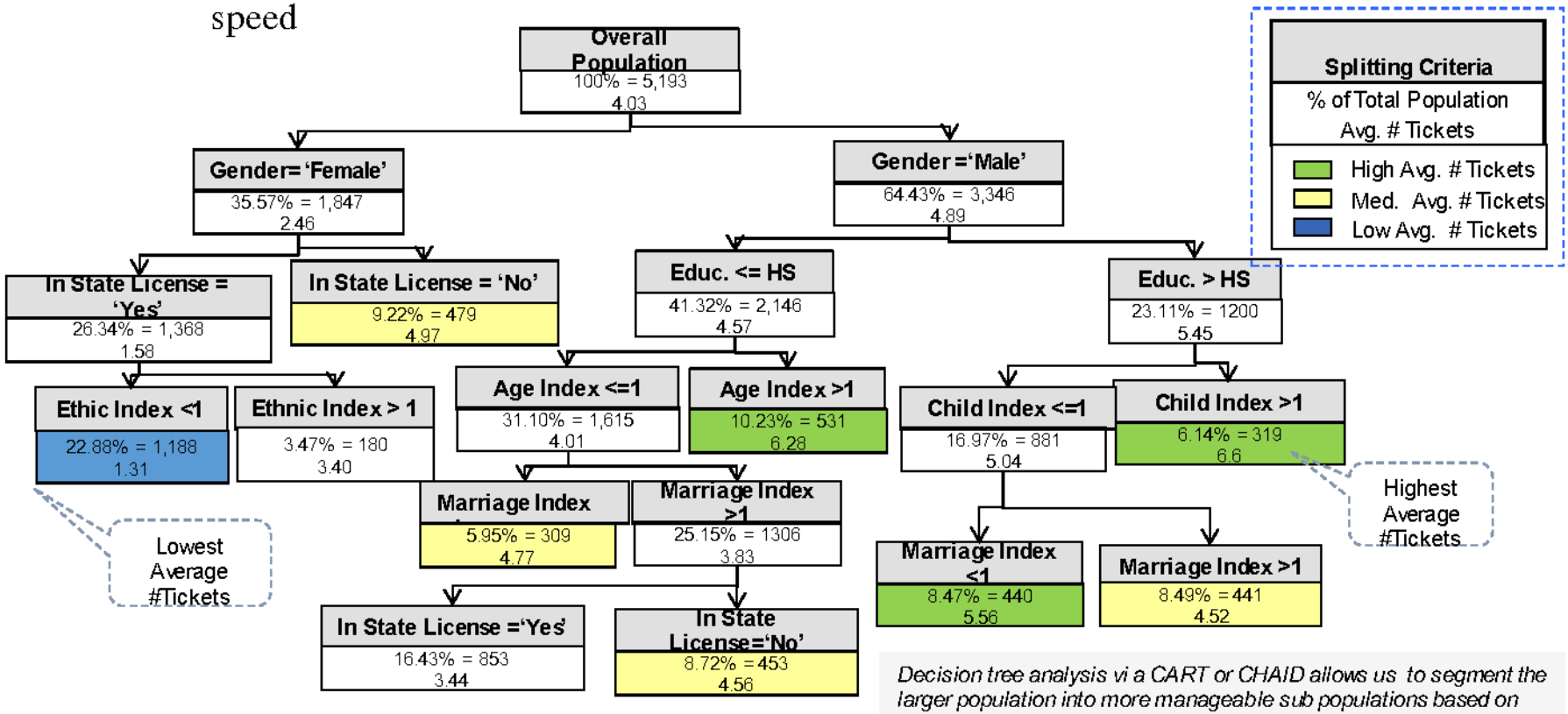
- The cumulative lift shows the degree of benefit one would get by allocating additional traffic resources to a specified group of Brooklyn zip codes versus current allocation levels
  - Brooklyn zip codes in quartile 1 are likely to have 3.09 times as many speeding drivers relative to the entire borough. Thus increased allocation of traffic resources can:
    - Reduce the # of speeding drivers given a higher police presence
    - Increase speeding ticket revenue generated *given the increased police presence.*

Quartile	Incr.	Incr.	Incr.	Model Index
	# Speed Tickets	Speed Ticket- Rate	# Queens zip codes	Benefit
1	3,327	66.54%	5	309
2	1,643	32.86%	5	231
3	1,297	21.94%	5	127
4	1,051	14.02%	5	90

Illustrative

Decision Tree Analysis - Demographics

- **Critical factors: Gender, Education, Marriage Index, Children Index, License State.**
  - Highly educated males with children tend to speed more than any other segment.
  - Female in state drivers from ethnically diverse areas are the least likely cohort to speed



Decision tree analysis vi a CART or CHAID allows us to segment the larger population into more manageable sub populations based on characteristics that are related to speeding tickets issued

Section IV: Optimization
(*Forecasting*)

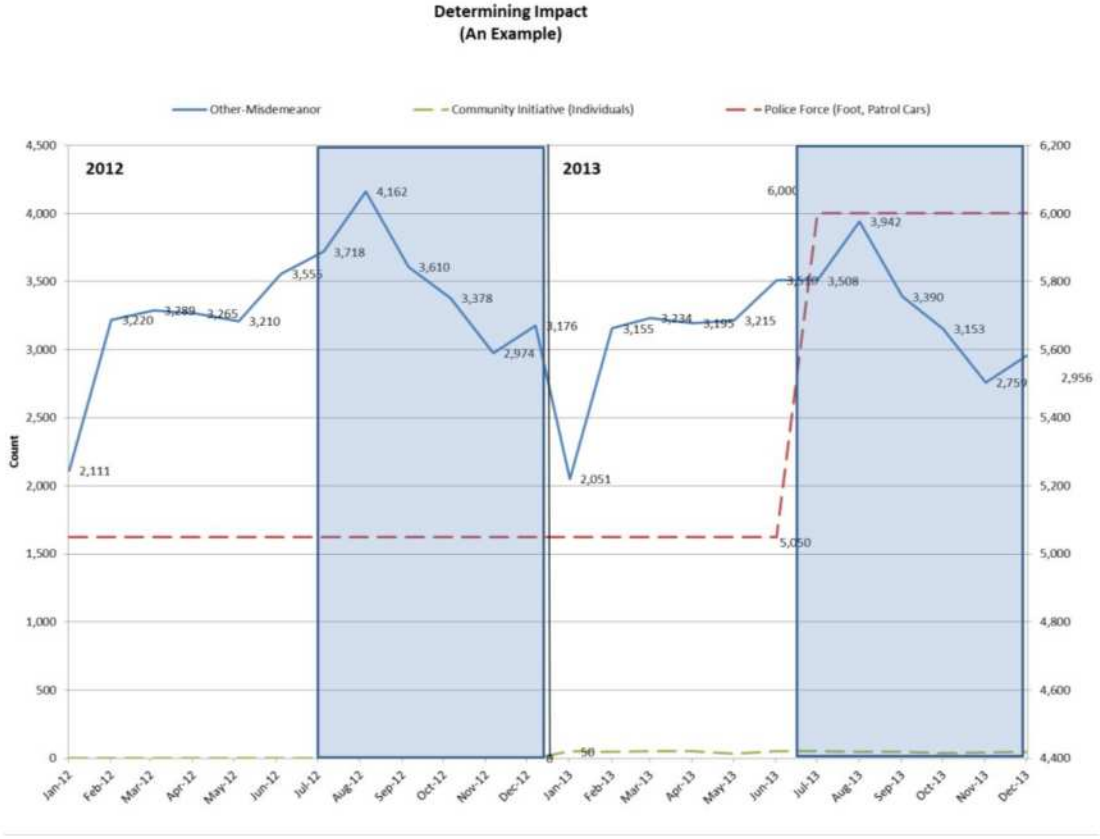
Forecasting – Policy Analysis

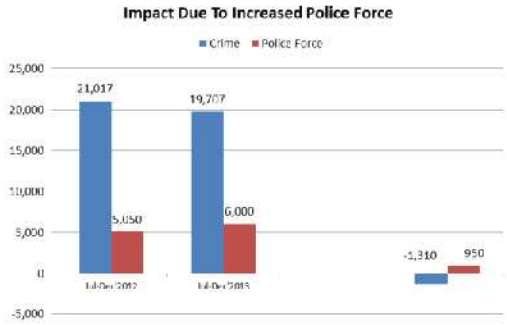
- **Optimizing the allocation of resources towards crime management depends on the impact new initiatives have on key measures of crime activity**
  - The relationship between the implementation of initiatives and the rise or fall of crime provides insight into the impact of these initiatives
  - Initiatives could be policy changes (*new laws, increased or decreased manpower or funding*)

- **To look at these in a simultaneous manner allows us to determine the contribution the initiatives made towards impacting crime rates.**

Illustrative

Forecasting Simulation – Policy Analysis



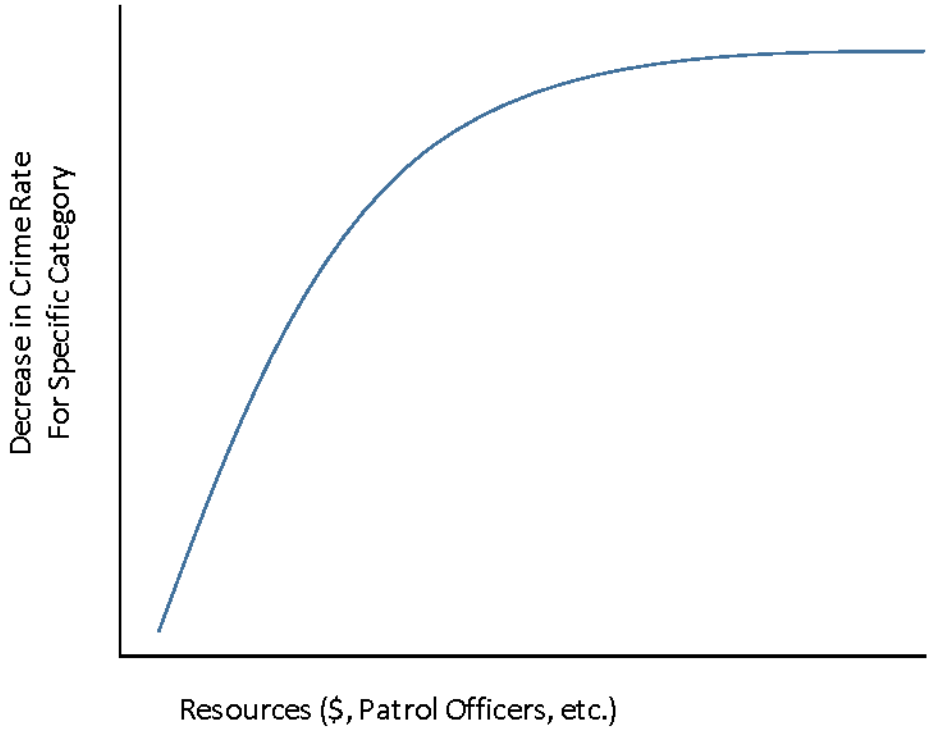


An 18.8% increase in Police Officers results in a 6.2% decrease in crime.

- **As an example, we consider the impact of increased manpower along with a community initiative have on Misdemeanor crimes.**
  - An association between crime rate and introduction of an initiative is sought.

Illustrative

Forecasting – Policy Analysis



There is a point of diminishing returns as investment increases.

- *The optimal point is where $'s invested in managing crime does not yield the same amount*
- *"Amount," refers to the inherent quantifiable value of the crime – or as surrogate, the crime rate for a particular crime type.*

- **Optimal allocation of resources is where we see no decrease in crime activity as a result of increasing resources (e.g., funding, manpower, etc.)**
  - The point of diminishing returns will be different for each tactic.

Illustrative

Forecasting – Policy Analysis

Impact of Tactic On Crime			Total FY 2014	Jan-14	Feb-14	Mar-14	Apr-14	May-14	Jun-14	Jul-14	Aug-14	Sep-14	Oct-14	Nov-14	Dec-14
Tactic	Assumptions	Comments	Full Fiscal Year	Formulas --->											
Police Force	Tactic	Count of Police Officers	66,300	5,050	5,050	5,050	5,050	5,050	5,050	6,000	6,000	6,000	6,000	6,000	6,000
Crime - Projected	Crime By Type	Resulting impact on Crime	37,948	2,081	3,105	3,264	3,145	3,245	3,460	3,538	3,892	3,420	3,103	2,789	2,906
Crime - Actual	--	Change in Crime As A Result of Increased Police Count	38,068	2,051	3,155	3,234	3,195	3,215	3,510	3,508	3,942	3,390	3,153	2,759	2,956
Crime - Variance (Projected Vs. Actual)	--	--	99.7%	101.5%	98.4%	100.9%	98.4%	100.9%	98.6%	100.9%	98.7%	100.9%	98.4%	101.1%	98.3%

- **A tracking mechanism can be put in place to quickly determine impact in the short term to address immediate tactical concerns.**
  - By looking at the change on a seasonal basis from one year to the next for example, we can quickly determine the expected impact of policy changes (e.g., increased man power).
  - We can also determine the level of the tactic (increased man power) to address an expected increase in crime as well.

- **This enables addressing issues in the short term, as we work on refining it through a more comprehensive econometrics approach.**

Section V: Strategy Implementation

Illustrative

Key Findings for Policing Strategy

- Our analysis show that ……..
- Further, this supports the departments desire to ……..