

A fair allocation algorithm for predictive police patrolling

Isabella Rodas Arango¹, Mateo Dulce Rubio¹, and Álvaro J. Riascos Villegas^{1,2}

¹Quantil

²Universidad de los Andes

Abstract

We address the tradeoff of developing good predictive models for police allocation vs. optimally deploying police officers over a city in a way that does not imply an unfair allocation of resources. We modify the fair allocation algorithm of [1] to tackle a real world problem: crime in the city of Bogotá, Colombia. Our approach allows for more sophisticated prediction models and we show that the whole methodology outperforms the current police allocating mechanism in the city. Results show that even with a simple model such as a Kernel Density Estimation of crime, one can have much better prediction than the current police model and, at the same time, mitigate fairness concerns. Although we can not provide general performance guarantees, our results apply to a real life problem and should be seriously considered by policy makers.

Keywords

Predictive Policing, Algorithmic Fairness, Optimal Allocation of Police, Censored Data.

1 Introduction

Police agencies around the world have widely started to include data mining techniques in their decision-making processes. In particular, machine learning algorithms have been designed and implemented in recent decades to predict crime events in cities, which then serve as input for the allocation of scarce law enforcement resources [2]. These predictive patrol models have become a standard in the way police operate in major cities around the world [3].

To take advantage of these developments, Colombia's government has recently introduced changes in the way patrolling is carried out in Bogotá, its capital city. First, they announced the development of machine learning models to predict high-impact crime events tailored to the city's social and criminal dynamics [4]. On the other hand, they are modifying the patrolling strategy from predetermined static geographic units to new dynamic units [5]. The present article is framed in the second effort and investigates how to reassign police patrols to the newly defined patrol areas using machine learning models as input in a fair resource allocation algorithm with partial observation of crime (i.e., censored data).

The use of predictive models to guide the patrolling of a city has caused strong criticism from various organizations, as they have been shown to generate biased predictions and disparate impacts

on disadvantaged populations. [6] empirically demonstrated that using predictive models trained with biased police data to assign patrols, and then retraining the models including the observed data, generates a feedback loop bias that does not allow the models to recover the underlying crime distribution. This is often the case: police data reflects historical policing patterns and captures complex dynamics between criminality, law enforcement agents and the society.

Some approaches have been proposed to mitigate the undesired effects of these models. For the problem of feedback loop due to censored data [7] stands out. In this paper the authors show that filtering the observed crimes used to retrain the models avoids feedback loop (under the assumption of observing crimes at the true rate of occurrence) and allows the model to recover the true criminal distribution. The proposed filter consists of including a crime in the re-training dataset with probability inverse to that of going to the region in which the crime was observed.

Other proposals add regularization components to the objective function of statistical models to penalize for deviations from some particular definition of fairness. For instance, [8] modifies a widely used model based on self-exciting point processes and penalizes the likelihood function so that the number of police patrols exposed to each racial group is proportional to the share of the race in the total population. In [9] the authors modify a GANs model and include in the learning function a term that penalizes for model miscalibration, arguing that a poorly calibrated model generates over- and under-policed areas. Despite the proposed approaches, none is currently widely accepted by the academic community.

In this paper we build on another approach to this problem proposed by [1]. In their article the authors propose an algorithm for the fair allocation of limited resources, in particular police patrols. The authors consider an allocation as fair if the probability of a crime being discovered in every pair of regions does not differ by more than a threshold α . This notion is similar to the well studied notion of equal opportunity in the machine learning literature (see [10]). With this in mind, the proposed algorithm iteratively constructs allocations to maximize the expected number of crimes discovered by the police by ruling out unfair allocations, i.e., where the difference in the probability of observing crimes between some pair of neighborhoods exceeds α .

Following [1] we allow for censored data in the sense that not all crimes are discovered, but only in proportion to the number of police and crimes per unit area. If there is more police in a particular region than crimes, then all of these are discovered. Otherwise, crimes are randomly discovered in a region depending on the number of police patrols and crimes. This setting allows for a more realistic framework than traditional approaches that ignore altogether this obvious partial observability phenomena in crime data.

We extend the algorithm proposed by [1] in two main ways. First, the authors assume that crime occurrence follows a Poisson distribution, which does not hold for our data for Bogotá. Consequently, we use a Kernel Density Estimation (KDE) model to flexibly learn the distribution of crime occurrence in each region and use it in the allocation algorithm of [1]. Second, we modify the greedy assignment rule such that the algorithm assigns an additional patrol to the region with the highest probability of exhibiting more crimes than its current allocation, rather than the region with the highest marginal probability of having an additional crime.

We run the modified algorithm for different fairness thresholds α and we compare our results with an approximation to the current allocation strategy used in Bogotá, a naive allocation in which patrols are assigned according to the proportion of crime in each region, and with a greedy allocation without taking into account any fairness requirement. In particular, we found that the current allocation method used by Colombia's government taking into the account administrative small regions (a total of 1051 in Bogotá) called *cuadrantes* is particularly unfair and inefficient in comparison to all the methodologies we propose in this paper.

The rest of the article is organized as follows. In the following section we detail the problem setup and the data sources. Section 3 presents a summary of the proposed algorithm in [1], illustrates the use of KDE methodology to flexibly estimate the crime distribution, and discusses the proposed modifications to the algorithm. Section 4 describes the implementation of the algorithm, and presents the results of police allocation in Bogotá against baseline allocation scenarios. The last section concludes and discusses some limitations of the present work and possible future research directions.

2 Problem setup

Bogotá is a relatively wealthy and developed city with 8 million inhabitants. In 2010, the National Plan for Community Policing by Quadrants was created (PNVCC, by its Spanish abbreviation) in which the city was divided into 1051 quadrants (*cuadrantes*) as basic units for police patrolling. Each quadrant is assigned a police patrol with two officers for each police shift [11]. Bogotá thus has 1051 patrols that it allocates uniformly among its 1051 quadrants.

Recently, after an increase in the number of criminal events in the city, the Mayor's office of Bogotá announced modifications to this patrolling system including the introduction of dynamic quadrants to tackle insecurity in the city. This change implies a reduction in the number of quadrants which raises the central question of this paper: *How should the 1051 patrols available in Bogotá be reallocated?*

To answer this question we use the fact that Bogotá is administratively divided into 19 localities (*localidades*),¹ each of them with a Police Station that is in turn in charge of a certain number of quadrants. The spatial distribution of localities and quadrants in Bogotá can be seen in Figure 1. It should be noted that each PNVCC police quadrant belongs to one, and only one, locality.

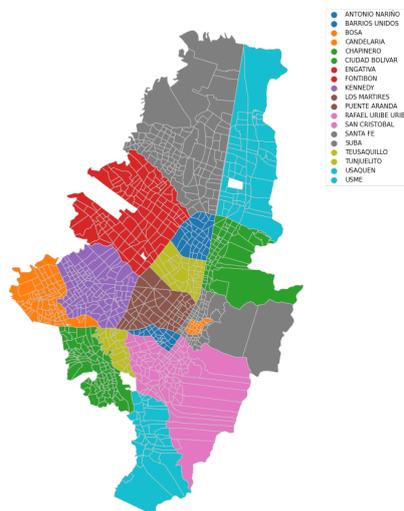


Figure 1: Spatial distribution of localities and quadrants in Bogotá.

With this in mind, we refine our central question to: *How should the 1051 patrols available in Bogotá be reallocated among Bogotá's 19 localities?*

¹Bogotá is divided into 20 localities but we ignored Sumapaz locality as it is entirely rural and has no quadrants assigned according to the PNVCC.

2.1 Data

We used records of robberies in the city officially registered by the Metropolitan Police of Bogotá. The complete dataset consists of robbery occurrences registered anywhere in Bogotá during 2018 and 2019 which consists of 1,366,712 robbery records.

We aggregate this data weekly for each locality to emulate the weekly allocation of the different patrols in Bogotá. Every locality has a different weekly crime distribution that can be detailed in Figure 2 that shows histograms for the weekly number of crimes recorded for the 19 localities. The empirical crime distributions do not follow a Poisson distribution, therefore the weekly robbery count for each locality is used to train a kernel density estimation model to learn the crime distribution for each urban area.

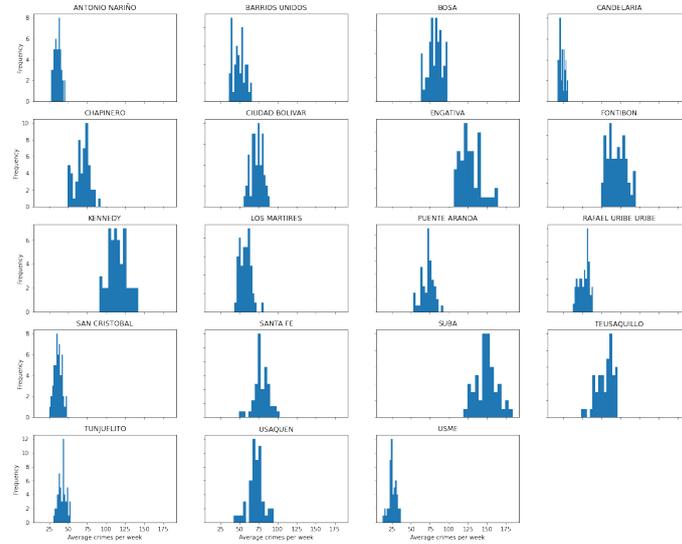


Figure 2: Number of crimes recorded per week for the 19 localities in Bogotá.

3 Methodology

This section summarizes the fair algorithm for learning in allocation problems proposed in [1]. The algorithm maximizes the expected number of crimes discovered in the city by taking into account exclusively fair police allocations. In detail, the authors define an allocation as fair if the probability of discovering a crime for any two regions does not differ by more than a parameter α (equation (1)). In other words, an allocation is fair if a crime is observed in each region with (almost) equal probability. This notion is similar to the well studied notion of equal opportunity in the machine learning literature [10]. To formalize this fairness notion, let

$$f_i(v_i) = \mathbb{E}_{c_i \sim C_i} \left[\frac{\min(v_i, c_i)}{c_i} \right]$$

be the probability of discovering a crime in region i when v_i police officers are allocated to this region, where C_i is the true distribution of crimes in the region. Implicitly, we are assuming that $\min(v_i, c_i)$ crimes are actually observed by the police. Hence, the discovery model assumes that each police officer is able to observe one, and only one, crime, and allows for censored data in the sense that not all crimes that occur are actually discovered by the police. With this notation, the

fairness requirement is expressed as:

$$|f_i(v_i) - f_j(v_j)| \leq \alpha, \quad \forall i, j \in \mathcal{G}, \quad (1)$$

To compute this expected value, the authors assume that crime occurrences follow a Poisson distribution C_i for each geographical region i .

The proposed algorithm uses the fact that once a region i is assigned v_i units, this assignment uniquely defines bounds on the admissible allocations of the other regions: $f_j \in [f_i - \alpha/2, f_i + \alpha/2], \forall j$. It then assigns to each region the minimum number of police officers such that $f_j \geq f_i - \alpha/2$ and check that this allocation does not exceed the total number of units available. If so, the remaining units are assigned according to a greedy rule that assigns an additional patrol to the region in \mathcal{G} with the highest marginal probability of having an additional crime:

$$\operatorname{argmax}_{i \in [\mathcal{G}]} (\mathcal{T}_i(v_i^t + 1) - \mathcal{T}_i(v_i^t)),$$

where $\mathcal{T}_i(v_i^t) = \mathbb{P}_{c_i \sim C_i}[c_i > v_i^t]$ is the right tail of the Poisson distribution for region i . These tails are also used to compute the expected number of crimes discovered for an allocation $v = (v_1, \dots, v_g)$ among the regions $i \in \mathcal{G}$:

$$\chi(v) = \sum_{i \in \mathcal{G}} \sum_{c=1}^{v_i} \mathcal{T}_i(c). \quad (2)$$

The detailed algorithm, taken from [1], is depicted in Figure 3.

```

Input:  $\alpha, C$  and  $\mathcal{V}$ .
Output: An optimal  $\alpha$ -fair allocation  $w^\alpha$ .
 $w^\alpha \leftarrow \vec{0}$ .
 $\chi_{\max} \leftarrow 0$ .
for  $i \in [\mathcal{G}]$  do
   $v \leftarrow \vec{0}$ .
  for  $v_i \in \{0, \dots, \mathcal{V}\}$  do
    Set  $v_i$  in  $v$  and compute  $f_i(v_i)$ .
     $ub_i \leftarrow v_i$ .
     $lb_i \leftarrow v_i$ .
    for  $j \neq i, j \in [\mathcal{G}]$  do
      Update  $lb_j$  and  $ub_j$  using  $f_i(v_i), \alpha$  and  $C_j$ .
       $v_j \leftarrow lb_j$ .
    if  $\sum_{i \in [\mathcal{G}]} v_i > \mathcal{V}$  then
      continue.
     $\mathcal{V}_r = \mathcal{V} - \sum_{i \in [\mathcal{G}]} v_i$ 
    for  $t = 1, \dots, \mathcal{V}_r$  do
       $j^* \in \operatorname{argmax}_{j \in [\mathcal{G}]} (\mathcal{T}_j(v_j + 1) - \mathcal{T}_j(v_j))$  s.t.  $v_j < ub_j$ .
       $v_{j^*} \leftarrow v_{j^*} + 1$ .
     $\chi(v) = \sum_{i \in [\mathcal{G}]} \sum_{c=1}^{v_i} \mathcal{T}_i(c)$ .
    if  $\chi(v) > \chi_{\max}$  then
       $\chi_{\max} \leftarrow \chi(v)$ .
       $w^\alpha \leftarrow v$ .
return  $w^\alpha$ .

```

Figure 3: Algorithm to compute an optimal fair allocation. From [1].

3.1 Extension of the algorithm

We first relax the assumption that the distribution of crime follows a Poisson distribution [1], as the empirical distributions of the number of crimes in Bogotá per locality do not follow such a

distribution (Figure 2). Consequently, we use a Kernel Density Estimation to estimate in a flexible way the weekly distribution of crime in each locality of Bogotá. This distribution is then used to calculate the expected number of crimes discovered in each locality $\chi(v)$ given a police assignment v using equation 2. Note that the right tail of the crime distribution $\mathcal{T}_i(v_i) = \mathbb{P}_{c_i \sim \hat{C}_i}[c_i > v_i]$ is now computed with respect to the learnt distributions \hat{C}_i .

Furthermore, we modified the greedy rule to allocate the available patrols. Instead of using the rule proposed by [1]

$$\operatorname{argmax}_{i \in \mathcal{G}} (\mathcal{T}_i(v_i^f + 1) - \mathcal{T}_i(v_i^f))$$

in which an additional police unit is assigned to the region in \mathcal{G} with the highest probability mass in its crime distribution between v^f and $v^f + 1$, we assign patrols with the rule:

$$\operatorname{argmax}_{i \in \mathcal{G}} (\mathcal{T}_i(v_i^f + 1))$$

such that the police allocation is increased by one unit in the region with the highest probability of crime occurrence greater than v^f . This change allows police to be assigned to regions with a low probability of observing exactly $v^f + 1$ crimes but with a high probability of observing a number of crimes greater than $v^f + 1$.

4 Results: Police allocation for Bogotá

We fit a Kernel Density Estimation (KDE) model for each locality to estimate the weekly crime distribution. The bandwidth for each of the kernels in the different localities was chosen using 10-fold cross validation. We then run the proposed modified algorithm to allocate 1051 police patrols among the 19 localities in Bogotá for different fairness thresholds α .

We compare our results with (i) an approximation to the current allocation used in Bogotá based on the number of quadrants in each locality, (ii) a naive allocation in which patrols are assigned according to the proportion of crime in each region, and (iii) a greedy allocation without taking into account any fairness requirement. We further study each of the changes introduced to the algorithm and compare the solution obtained to (iv) the original [1] algorithm assuming Poisson distributions (but computing the expected number of crimes using the distribution estimated with KDE), and (v) with the greedy allocation rule used by [1].

The allocations obtained by the different allocation strategies outlined above are shown in Figure 4. The current allocation based on the number of quadrants in each locality differs the most from all the allocations found. The naive allocation and the proposed modified algorithm have similar solutions as both take into account the crime distribution throughout the city. The proposed algorithm generates an allocation with a higher expected number of observed crimes an $\alpha \approx 0.02$, that is with a maximum difference in the probabilities of discovering crime between two localities of approximately 2%. Although the solution of the different methodologies does not differ much and most localities end up having a similar number of patrols assigned, slight changes in the solutions of each of the models represent either a fairer model or a higher expected number of crimes being discovered. The Appendix provides the solutions obtained in Table 1.

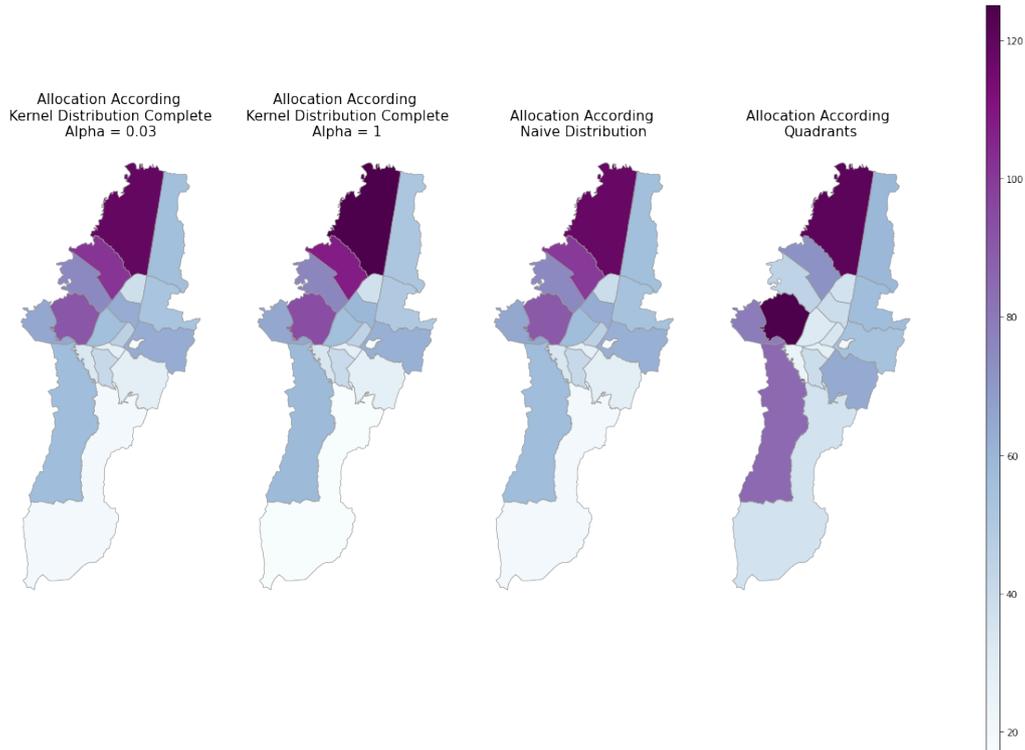


Figure 4: Allocation of police patrols using different methodologies.

To evaluate the different allocation configurations we use the expected number of crime discovered associated to the resulting allocation:

$$\chi(v) = \sum_{i \in \mathcal{G}} \sum_{c_i}^{v_i} \mathcal{T}_i(c),$$

where \mathcal{T}_i is the right tail of the distribution estimated through KDE.

The results for different values of the fairness parameter α are shown in Figure 5. Both models using the kernel density functions (with different greedy allocation rules) exhibit a similar behaviour and seem to stabilize for $\alpha < 0.03$. The Poisson model has a different behavior as the expected number of crimes discovered decreases as the α value increases. We believe that this behavior is a consequence of the greedy allocation rule used in [1]. It should be noted that the Naive model in which patrols are assigned according to the proportion of crime in each region have a low α value (0.022) representing a fair allocation, but has a lower expected number of crimes discovered than the proposed model.

On the other hand, the kernel density model that takes into account the right tail of the estimated distribution to allocate the remaining police units seems to have a better performance than any of the other models for all the α values. This model allocates the patrols fairly while maximizing the average number of crimes discovered overall. The current allocation based on the number of quadrants in each locality has an $\alpha = 0.54$ and an average crime discovered equal to 983.59, representing an unfair allocation with a low expected number of crimes discovered (not shown in figure 5). This is the most deficient allocation out of all the models tested.

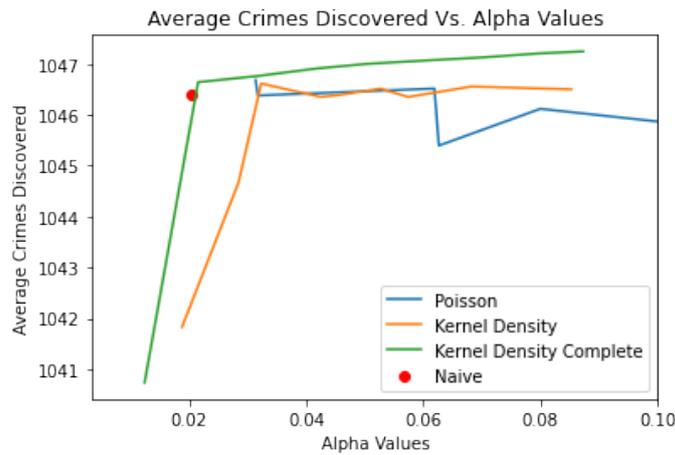


Figure 5: Allocation for the different methodologies over different fairness thresholds α .

5 Conclusions

Policy makers armed with new machine learning tools are making progress in crime prediction. Nevertheless, algorithms may be biased and reproduce existing police patrolling biases that may have disparate impacts for different populations. This paper shows in a real setting how to take a simple predictive policing algorithm (e.g., a KDE model) and use it for fairly allocating scarce police resources in such a way that we still have a good predictive model and mitigate the unfair allocation of police resources as described in the paper. This paper addresses a central question in predictive police patrolling for any city independent of their police deployment strategy.

We extend the algorithm proposed by [1] in two main ways. We first relax the assumption that the distribution of crime follows a Poisson distribution, as the empirical distributions of the number of crimes in Bogotá per locality do not follow such a distribution (Figure 2). Consequently, we use a Kernel Density Estimation (KDE) model to flexibly learn the distribution of crime occurrence in each region and use it in the allocation algorithm of [1]. Second, we modify the greedy assignment rule such that the algorithm assigns an additional patrol to the region with the highest probability of exhibiting more crimes than its current allocation, rather than the region with the highest marginal probability of having an additional crime.

We run the modified algorithm for different fairness thresholds α and compare our results with an approximation to the current allocation strategy used in Bogotá, a naive allocation in which patrols are assigned according to the proportion of crime in each region, and with a greedy allocation without taking into account any fairness requirement. The results clearly show that the current allocating police strategy in Bogotá is dominated by our allocating mechanism even when we use a simple machine learning model to make crime predictions (KDE): the model discovers more crimes and mitigates unfair allocations.

However, at least two issues are left for future research. First, although the modified algorithm [1] works well with real data for Bogotá, which clearly suggests policy makers should take it seriously, we do not provide general performance guarantees as [1] does in the particular setting in which crime follows a Poisson distribution. This is clearly unrealistic in our case as shown in 2 and given all the evidence from more sophisticated machine learning algorithms. Second, in this model censored data plays a passive role since, by construction, we do not allow discovered crimes to bias our algorithm (there is no training as police is deployed). This means we do not allow for

feedback loop and take our model as ground truth. Clearly this is not the case in a real setting where one has to discover the true model and at the same time deploy police that discovers crimes. We leave this learning problem of sophisticated machine learning models for future research.

References

- [1] H. Elzayn, S. Jabbari, C. Jung, M. Kearns, S. Neel, A. Roth, and Z. Schutzman, “Fair algorithms for learning in allocation problems,” in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 170–179.
- [2] G. O. Mohler, M. B. Short, S. Malinowski, M. Johnson, G. E. Tita, A. L. Bertozzi, and P. J. Brantingham, “Randomized controlled field trials of predictive policing,” *Journal of the American statistical association*, vol. 110, no. 512, pp. 1399–1411, 2015.
- [3] S. Greengard, “Policing the future,” *Communications of the ACM*, no. 3, pp. 19–21, 2012.
- [4] B. mayor’s office. (2019) In bogotá, crimes will be predicted with artificial intelligence. [Online]. Available: <https://bogota.gov.co/mi-ciudad/seguridad/modelo-de-prediccion-de-crimenes-en-bogota>
- [5] Semana. (2020) Bogota: dynamic quadrants, among the new measures to reduce insecurity. [Online]. Available: <https://www.semana.com/nacion/articulo/bogota-cuadrantes-moviles-entre-las-nuevas-medidas-para-frenar-la-inseguridad/682206/>
- [6] K. Lum and W. Isaac, “To predict and serve?” *Significance*, vol. 13, no. 5, pp. 14–19, 2016.
- [7] D. Ensign, S. A. Friedler, S. Neville, C. Scheidegger, and S. Venkatasubramanian, “Runaway feedback loops in predictive policing,” *arXiv preprint arXiv:1706.09847*, 2017.
- [8] G. Mohler, R. Raje, J. Carter, M. Valasik, and J. Brantingham, “A penalized likelihood method for balancing accuracy and fairness in predictive policing,” in *2018 IEEE international conference on systems, man, and cybernetics (SMC)*. IEEE, 2018, pp. 2454–2459.
- [9] C. Urcuqui, Christian, S. Moreno, Juan, A. Montenegro, Carlos, J. Riascos, Alvaro, and M. Dulce, “Accuracy and fairness in a conditional generative adversarial model of crime prediction,” *7th International Conference on Behavioural and Social Computing (BESC 2020)*, Bournemouth, UK., 2020.
- [10] M. Hardt, E. Price, and N. Srebro, “Equality of opportunity in supervised learning,” *In Advances in neural information processing systems*, pp. 3315–3323, 2016.
- [11] C. N. Police. (2010) Institutional strategy for citizen security: National plan for community policing by quadrants. [Online]. Available: <https://www.oas.org/es/sap/dgpe/innovacion/banco/ANEXO%20I.%20PNVCC.pdf>

Appendix

| Locality | Poisson | Kernel Density | Kernel Density Complete | Naive | Quadrants |
|--------------------|---------|----------------|-------------------------|-------|-----------|
| Antonio Nariño | 29 | 28 | 28 | 28 | 29 |
| Barrios Unidos | 39 | 38 | 38 | 39 | 36 |
| Bosa | 65 | 64 | 65 | 65 | 79 |
| Candelaria | 18 | 18 | 18 | 18 | 21 |
| Chapinero | 54 | 53 | 53 | 54 | 57 |
| Ciudad Bolívar | 57 | 56 | 57 | 57 | 85 |
| Engativá | 100 | 101 | 101 | 100 | 72 |
| Fontibón | 75 | 75 | 75 | 75 | 44 |
| Kennedy | 91 | 91 | 91 | 90 | 125 |
| Los Mártires | 45 | 44 | 45 | 45 | 34 |
| Puente Aranda | 57 | 56 | 56 | 57 | 32 |
| Rafael Uribe Uribe | 41 | 40 | 41 | 41 | 39 |
| San Cristobal | 29 | 29 | 29 | 29 | 64 |
| Santa Fé | 62 | 63 | 63 | 62 | 54 |
| Suba | 117 | 119 | 119 | 118 | 121 |
| Teusaquillo | 63 | 64 | 63 | 63 | 39 |
| Tunjuelito | 33 | 33 | 33 | 34 | 25 |
| Usaquén | 56 | 57 | 56 | 56 | 59 |
| Usme | 20 | 20 | 20 | 20 | 36 |

Table 1: Number of patrols allocated to each locality given the different methodologies using $\alpha = 0.03$.