National Incident-Based Reporting System (NIBRS) Data: A Practitioner's Guide

Jacob Kaplan

2021-05-30

Contents

8 Property and Window Property Segment   66

List of Figures

Chapter 1

Preface

The average American had, in 2019, about a 1 in 20,500 chance of being murder, 1 in 1,223 chance of being robbed, and a 1 in 64 chance of having something they own stolen.[1] Getting these numbers is extremely simple. We take the number of crimes reported to the police and divide it by the number of people living in the United States that year. For example, there were about 16,000 murders in 2019 and 328 million people in the county - 16,000 / 328 million = ~1/20,500. You'll more commonly see this - in news articles, in political speeches, in research articles, on TV, etc. - reported as the rate per 100,000 people but that's just a matter of conversion, the numbers are the same. This is, however, totally wrong. It assumes - and let's now just pretend that there's no underreporting of crimes to the police - that every single person has the exact same risk of victimization. We know this is wrong intuitively. There are the "bad parts of town" or people who "run with the wrong crowd." Research in criminology backs this up by finding that crime is generally concentrated among a small group of people and within a small geographic area (usually a very small number of streets or neighborhoods in a city). From surveys that ask if people have been victims of a crime we also know that victimization rates differ by age, race, gender, income, and city type. Indeed, think of a personal characteristic (e.g. risk tolerance, athleticism, frequency outdoors) and there will probably be large differences in the likelihood of being a victim within these groups.

So why do people so frequently talk about crime as rates per total population? Why assume that everyone has equal risk of being a victim? The major reason, I think, is that the main FBI dataset on crime, the Uniform Crime Reporting (UCR) System Data doesn't provide any information about crime victims other than for homicide and hate crime victims.[2] Victimization surveys are nearly all at state or national levels so trying to use their results

---

[1]This is according to data from the FBI's 2019 Crime in the United States report.

[2]There is a small amount of information available for property theft victims, but not enough to do these kinds of victimization rate calculations

for specific city-level rates will have very high levels of imprecision, likely enough to make the results unhelpful. So by necessity people tend to assume equal risk since there's not good data to do so otherwise. Doing so, however, is a large mistake in my opinion. It both overstates and understates risk, often very drastically. For example, many murders are concentrated among young minority men in impoverished areas of urban cities. Their risk of being murdered is many times higher than you'd expect from simply dividing murders by city population overall. And the risk for other people in the city is far smaller than you'd expect from the naive risk value.

While UCR data is the predominant dataset in criminology and the one that truly guides America's perceptions of crime, there is actually one that is far more detailed and allows us do a much better (though far from perfect) measure of crime victimization by a number of victim traits. This dataset, called the National Incident-Based Reporting System - and often called by its abbreviation NIBRS - provides demographics information about each victim of crime. Like the UCR this is a dataset from the FBI which standardizes information from agencies in the United States so it is comparable across agencies and over time (with some caveats). This data includes the age, race, gender, and the relationship of the victim to the offender for every known victim. We also have info about the offense such as which crimes occurred (and unlike the UCR, incidents can have multiple offenses), where and when it occurred, and what weapons (if any) we used. And we have demographic information about offenders, and whether they were arrested (including the type of arrest such as warrant arrest). From this we can figure out victimization rates based on certain victim characteristics, though not all of them that we may think are important.

Of course, the research opportunities from this highly detailed data extend far beyond mere victimization rates. Personally I've used the data to examine topics ranging from marijuana decriminalization and domestic violence injuries to daylight saving time's effect on crime clearance. As of this writing in early summer 2021, Google Scholar has over 1,500 results for NIBRS research since 2017. As the FBI has said that starting in 2021 NIBRS data will completely replace UCR data, this is crime dataset of the future (though I believe that UCR will still be heavily used for many years to come). For graduate students and early career researchers, this is likely the dataset that you'll use for your entire career.

1.1 Goal of the book

By the end of the book you should have a firm grasp on NIBRS data and how to use it (or as is occasionally the case, choose not to use it) properly. However, this book can't possibly cover every potential use case for the data so make sure to carefully examine the data yourself for your own particular use.

I get a lot of emails from people asking questions about this data so my own goal is to create a single place that answers as many questions as I can about the data.[3] As the FBI has moved to only use NIBRS data starting in 2021, I expect the uses of this data - and thus the number of emails I get - to grow very quickly. This is an increasingly popular dataset used by criminologists (and by other fields studying crime) and yet there are still occasions where papers are using the data incorrectly.[4] So hopefully this book will decrease the number of misconceptions about this data, increasing overall research quality.

Since manuals are boring, I'll try to include graphs and images to try to alleviate the boredom. That said, I don't think it's possible to make it too fun so sorry in advanced. This book is a mix of facts about the data, such as how many years are available, and my opinions about it, such as whether it is reliable. In cases of facts I'll just say a statement - e.g. "NIBRS data began in 1991". In cases of opinion I'll temper the statement by saying something like "in my opinion…" or "I think".

1.2   Structure of the book

This book will be divided into nine chapters: this chapter, an intro chapter briefly summarizing each segment files and going over overall issues with NIBRS data, and seven chapters each covering one of the seven UCR datasets. Each chapter will follow the same format: we'll start with a brief summary of the data such as its possible uses and pitfalls. And then, we'll cover the important variables included in the data and how to use them properly (including when to not use them at all) - this will be the bulk of each chapter.

1.3   Citing this book

If this data was useful in your research, please cite it. To cite this book, please use the below citation:

Kaplan J (2021). *National Incident-Based Reporting System (NIBRS) Data: A Practitioner's Guide.* https://nibrsbook.com/.

BibTeX format:

```
@Manual{ucrbook,
  title = {National Incident-Based Reporting System (NIBRS) Data: A Practitioner's Guide
```

---

[3]Writing also keeps away the boredom.

[4]Though given that the data is fairly complicated and requires good programming knowledge, the bar is higher to use it. So there are far fewer bad uses of this data than there is for UCR data.

```
  author = {{Jacob Kaplan}},
  year = {2021},
  url = {https://nibrsbook.com/},
}
```

1.4   Pronunciation

This data is usually just called NIBRS, and generally there's no distinction between segment files since they work in unison as they are pieces of the overall criminal incident. "NIBRS" is generally pronounced as "NIE-BERS". It rhymes with "HIGH-BERS". I've also heard it pronounced - usually by non-academics - using a soft i like in "timber" so it sounds like "nih-bers". I prefer the "NIE-BERS" saying but it really doesn't make a difference.

1.5   Sources of NIBRS data

There are a few different sources of UCR data available today. First, and probably most commonly used, is the data put together by the National Archive of Criminal Justice Data (NACJD)). This a team out of the University of Michigan who manages a huge number of criminal justice datasets and makes them available to the public. If you have any questions about crime data - NIBRS or other crime data - I highly recommend you reach out to them for answers. They have a collection of data and excellent documentation available for UCR data available on their site here. They've also put together what they call "Extract Files" which are where they merged some of the NIBRS segments together, saving you the effort of doing so yourself. These extract files essentially take every potential unit of analysis - incident, victim, offender, and arrestee (some crimes have no victims, only arrestees) - and merge it with the segment which has info about the incident such as the time of day or the outcome, and information about the reporting agency. This source only has data through 2016 which means that the most recent years (NIBRS data is available through 2019) of data are (as of this writing) unavailable.

Next, and most usable for the general public - but limited for researchers - is the FBI's official website Crime Data Explorer. On this site you can chose an agency and see annual crime data (NIBRS data is at the day-level so this is very aggregated data) for certain crimes (and not even all the crimes actually available in the data). This site has only a small subset of the available data and is already aggregated so you're dealing with a subset of data in a unit of analysis that you may not want. For example, this site lets you see the annual age of offenders for certain crimes in age brackets such as aged 20-29. As the data provides

the exact age (in years) of each offender, this is much less useful than the full data. The crimes on this site are also limited to only the eight "Index Crimes" (Murder, rape, robbery, aggravated assault, arson, burglary, theft, and motor vehicle theft) so are only a tiny share of the crimes actually reported in NIBRS data. For more on what Index Crimes are, please see here. This data source is potentially okay for the general public but only provides a fraction of the data available in the actual data so is really not good for researchers.

Finally, I have my own collection of UCR data available publicly on openICPSR, a site which allows people to submit their data for public access. For each of these datasets I've taken the raw data from the FBI and read it into R. Since the data is only available from the FBI as fixed-width ASCII files, I created a setup file (we'll explain exactly how reading in this kind of data works in the next chapter) and read the data into R and saved the files in R and Stata files for easy use. The main advantage is that all my data has standard variable names and column names and can be read into modern programming languages R and Stata (this is also true of recent NACJD years, but early years come as fixed-width ASCII files). The downside is that I don't provide documentation other than what's on the openICPSR page and only provide data in R and Stata format so people using languages such as SAS or SPSS cannot use this data.[5]

1.5.1   Where to find the data used in this book

The data I am using in this book is the cleaned and concatenated data that I put together from the raw data that the FBI releases. The raw data that the FBI releases is available here. The data that I have released is available on the data hosting site openICPSR here. I am hosting this book through GitHub which has a maximum file size allowed that is far smaller than these data, so you'll need to go to openICPSR to download the data; it's not available through this book's GitHub repo.

1.6   Recommended reading

While this book is designed to help researchers use this data, the FBI has an excellent manual on this data designed to help police agencies submit their data. That manual, called the "2019 National Incident-Based Reporting System User Manual" provides excellent definitions and examples of many variables included in the data. In this book when I quote the FBI, such as defining a crime, I quote from this manual. The manual is available to download as a PDF on the FBI's site and I've also posted it on my GitHub page in case the FBI ever

---

[5]I am not sure if SAS or SPSS can read in R or Stata data files.

takes down the PDF.[6] The link on my GitHub page is here. I highly recommend that you read this manual before using the data. That manual, alongside this book which tries to explain when and how the agencies don't follow the manual, will provide a solid foundation for your understanding of NIBRS data.

1.7 How to contribute to this book

If you have any questions, suggestions, or find any issues, please email me at jkkaplan6 [@] gmail.com. For more minor issues like typos or grammar mistakes, you can edit the book directly through its GitHub page. That'll make an update for me to accept, which will change the book to include your edit. To do that, click the edit button at the top of the site - the button is highlighted in the below figure. You will need to make a GitHub account to make edits. When you click on that button you'll be taken to a page that looks like a Word Doc where you can make edits. Make any edits you want and then scroll to the bottom of the page. There you can write a short (please, no more than a sentence or two) description of what you've done and then submit the changes for me to review.

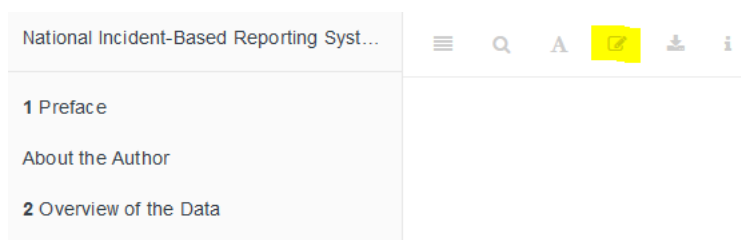Figure 1.1: The edit button for how to make edits of this book.

---

[6]This is far more likely to happen as a result of standard government changing a site and forgetting to update the link rather than intentionally making the manual unavailable.

About the Author

Jacob Kaplan holds a PhD and a master's degree in criminology from the University of Pennsylvania and a bachelor's degree in criminal justice from California State University, Sacramento. He is the Chief Data Scientist at the Research on Policing Reform and Accountability RoPRA at Princeton University. His current research portfolio includes evaluating police policy and reforms, place-based crime prevention, measuring spatial crime concentration, and simulating how firing 'bad apples' affects police complaints and uses of force. In the past he's written on the effect of marijuana decriminalization on domestic violence, how increasing the number of police officers affects prison trends, how outdoor lighting affects crime and perception of safety, and public perceptions of forensic science techniques. He is the author of several R packages that make it easier to work with data, including fastDummies and asciiSetupReader. His website allows for easy visualization of crime-related data and he has released over a dozen crime data sets (primarily FBI UCR data) on openICPSR that he has compiled, cleaned, and made available to the public.

For a list of papers he has written (including working papers), please see here.

For a list of data sets he has cleaned, concatenated, and made public, please see here.

For a list of R packages he has created, please see here.

Chapter 2

Overview of the Data

Nearly a century ago the FBI started collecting data on crime that occurred in the United States as a way to better understand and respond to crime. This data, the Uniform Crime Reporting (UCR) Program Data, is a monthly count of the number of crime incidents (in cases where more than one crime happens per incident, only the most serious crime is included) in each police agency that reports data.[1] Other than for homicides, only the number of crimes that occurred is included. So we know, for example, the number of robberies in a city but nothing about who the victims or offenders were, when that month (day or time of day) the robberies occurred, or the type of location where they happened. To address these limitations the FBI started a new dataset in 1991, the National Incident-Based Reporting System data, which is known by its abbreviation NIBRS, and is the topic of this book.

NIBRS data provides detailed information on every crime reported to the police, including victim and offender demographics, whether the offender was arrested (and type of arrest or type of "exceptional clearance"), the crime date and hour, victim-offender relationship, and the crime location (as a location type, not the exact address). It also covers a far wider range of crimes than UCR data did. With the exception of UCR data on assaults against police officers, all NIBRS data can be converted back to UCR data, making it fully backwards compatible - and, therefore, comparable to UCR data. In many ways NIBRS data is a massive improvement to UCR data. We now have a far deeper understanding of crime and this has led to an explosion of research that allows a far more detailed analysis of crime and crime-policies than the blunt UCR data.

However, this is a major limitation to this data: most agencies don't use it. According to the FBI only about 8,500 police agencies, covering about 45% of the US population, reported NIBRS data in 2019 (the latest year currently available), fewer than half of the about 18,000

---

[1]This data has been expanded since it began in 1929 to include information on arrests, hate crimes, and stolen property. For more on this, please see my book Uniform Crime Reporting (UCR) Program Data: A Practitioner's Guide at ucrbook.com.

police agenices in the United States. This is an even larger problem that it seems are the agencies that do report - especially in earlier years of the data - are disproportionately small and rural. So we're missing out of data from major cities. A number of states don't have any agencies reporting, making this data relatively biased at least in terms of geography and city size. **Even so, the FBI has said that they are moving entirely to NIBRS data starting in 2021, and will no longer even collect UCR data.** While NIBRS can be converted to UCR data, meaning we can have consistent statistics over time, for agencies that don't report to NIBRS, we have no information on their crimes. In effect, unless the majority of agencies suddenly switch to NIBRS - which, given the high level of detail relative to UCR data is a costly and timely switch - we will be flying blind for most crime in the country.

So there are really three major problems with NIBRS data, both related to the lack of reporting. First, we are potentially looking at a massive loss of data when UCR data ends in 2020 - it takes over a year for data to be released so even though I'm writing this is Spring 2021, 2019 UCR and NIBRS data are the latest years available. Especially given the huge crime changes during 2020 - and whose violent crime increases and continuing into 2021 - losing standardized crime data for most cities is a very bad thing. The second problem is that even if suddenly all agencies do start reporting in 2021, we'd only have a single year of data available. Even for agencies that already report, we generally don't have too many years of data for them. This really limits the kind of research since we can do since it's hard to know if a finding is based on a trend or is just a weird outlier without having many years of data available. This means that for the next several years at least we'll be mostly using NIBRS data as UCR-like datasets, aggregated to the month- or year-level so we can compare it with UCR data from the past. Luckily, this problem will be alleviated the longer we wait as more years of data will become available.

The final issue is that this data is massive. A single year of 2019 data, with <50% of agencies, and few large agencies, reporting has about 6.5 million crime incidents recorded. Since each crime incident can have multiple victims, offenders, and crimes, there are more rows for these datasets. Once all agencies report - though it's doubtful that'll ever occur - we're looking at tens of millions of rows per year. And even now if we wanted to look at a decade of data we're going to be dealing with over 50 million rows of data. So this data requires both good hardware - a strong laptop or server is necessary - and good programming skills, which most academics sorely lack.

While people generally refer to NIBRS just as "NIBRS data" it is actually a collection multiple different datasets all corresponding to a single crime incident. For example, if you care about victim info you'll look in the victim file called the "Victim Segment" (each of the datasets are called "Segments" since they are part of the whole picture of the crime incident) and likely will merge it with other data, such as when are where the crime occurred which

is in the "Offense Segment". In most cases you'll merge together multiple datasets from the NIBRS collection to be able to answer the question that you have. This means that you'll need to understand how to deal with multiple datasets, and subset and merge them as needed.

Relative to the FBI's UCR data there are far fewer "weird things" in NIBRS data. Still, we'll cover instances of the "weirdness" in the data, such as the why crime always goes up on the 1st of the month, or why there are more crimes at noon than at nearly all other hours of the day. We'll also be discussing how much of the detailed information that should be available in the data is missing, and when that affects which questions we can answer.

A word of caution. To date fewer than half of agencies report NIBRS data. As they rush to comply with the FBI's order that they only will accept NIBRS data, there will likely be more mistakes made and erroneous data included in NIBRS data later than is covered in this book, which ends with 2019 data as 2019 is the most recent year available. So while I always urge caution when using any data - caution that should be accompanied by a thorough examination of your data before using it - NIBRS data from 2020 and beyond merits extra attention.

2.1 NIBRS gives you choices (which means you need to think)

A major benefit of UCR data is that you have very limited choices. If you wanted to measure crime you only choice was to use their monthly aggregated city-level data. This makes working with the data relatively easy, even though what work you could do was limited. NIBRS data takes an opposite approach. It provides detailed data and largely leaves it up to the users for what to do with it. This flexibility is also a curse.

2.1.1 Choosing a unit of analysis

Consider, for example, that you want to measure rape in a city. Even if we somehow solve the issue of victims not reporting their rapes, we still have a few different ways of even measuring rape. Let's use an incident where four men rape a single woman as an example. If you're interested in measuring robbery you could do so in several different ways, each of which addresses a different part of crime measurement and will lead to different answers to your questions: the number of crime incidents, the number of victims, the number of offenders, and the number of crimes. First, we can follow the old UCR measure of incident-level and say that this is one rape since only one crime incident occurred (even though there were

multiple offenders). Second, we could look at the victim-level, which again is one rape as there was only one victim. Or at the offender-level, which now has four rapes since each offender would be responsible the robbery. Finally we could look at the offense-level. This data actually makes it unclear what act each offender actually committed so we'd likely measure it as one offense per offender in each incident.[2]

2.2 Crimes included in NIBRS

2.2.1 Index crimes

2.3 NIBRS does not have unfounded crimes

2.4 A summary of each segment file

2.5 The data as you get it from the FBI

---

[2]For four people involved in a rape, do we care if all four actually committed the crime? Does it matter if only three committed the rape and one merely stood by?

Chapter 3

Administrative and Window Exceptional Clearance Segment

The Administrative Segment provides information about the incident itself, such as how many victims or offenders there were. In practice this means that it tells us how many other segments - offense, victim, offender, and arrestee segments - there are for this particular incident. It also has several important variables at the incident-level such as what hour of the day the incident occurred and whether the incident date variable is actually just the date the incident was reported. Finally, it tells us whether the case was cleared exceptionally and, if so, what type of exceptional clearance it was. This can tell us, for example, how many crimes was cleared because the offender died or the victim refused to cooperate. As the UCR data doesn't differentiate between normal clearances (i.e. arrest the offender) and exceptional clearances, this provides a far deeper understanding of case outcomes.

3.1 Important variables

3.1.1 The incident date or report date

3.1.2 Hour of incident

An extremely important aspect of crime data is when exactly the crime occurs. If, for example, crime always spikes when the local high school ends their day that would likely indicate that high school students are involved with crime (both as victims-offenders). In my own research on daylight saving time-crime I only care about the sunset hours, which is when daylight saving time would affect outdoor lighting. When crime happens also would
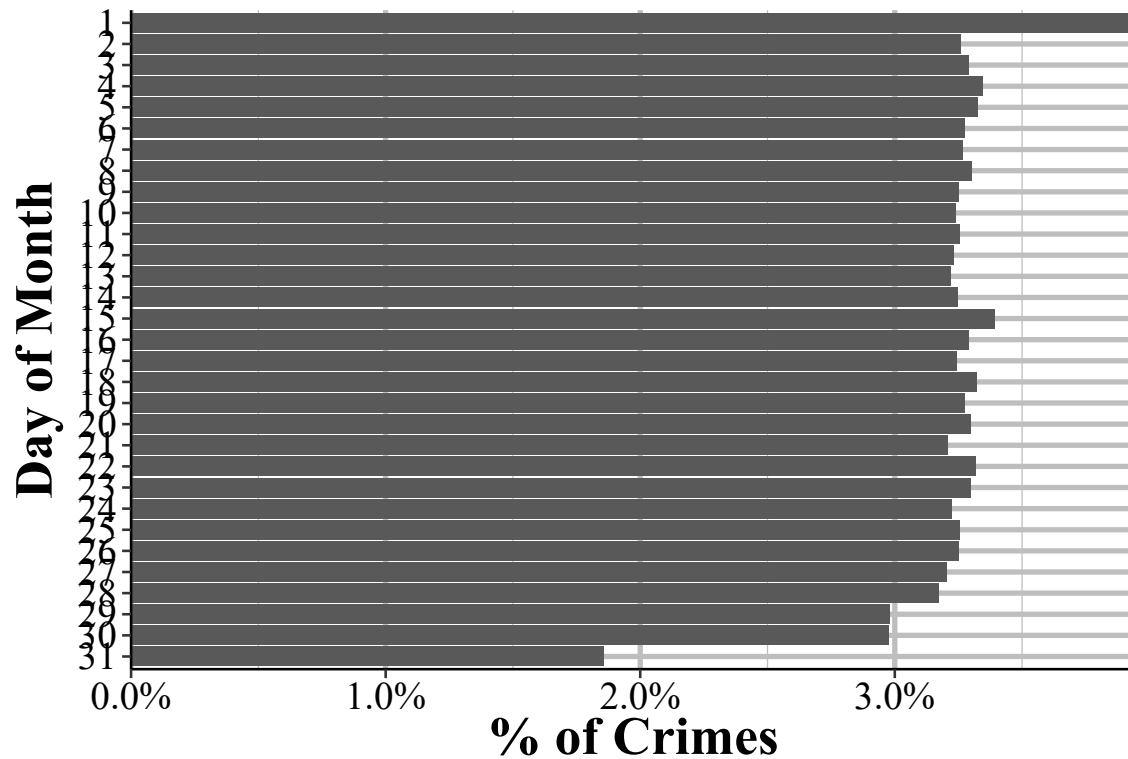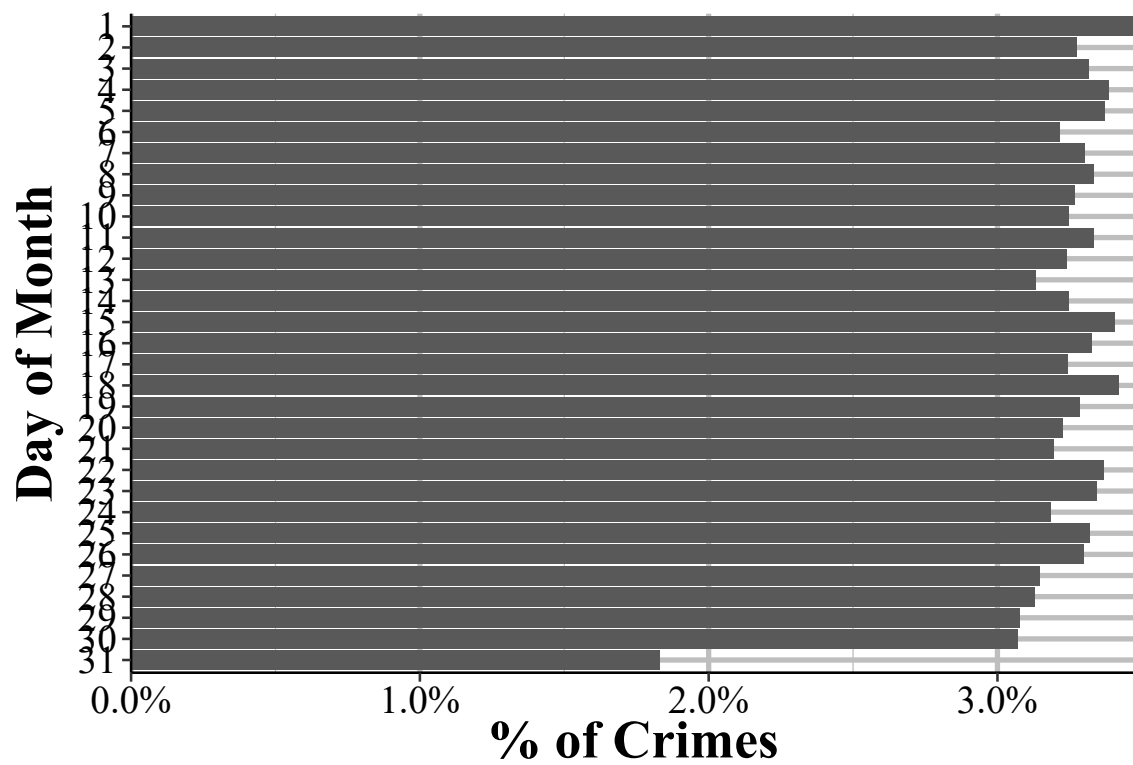
Figure 3.1: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.

affect police behavior as they'd likely increase patrol during times of elevated crime. Luckily NIBRS data does have the time of each incident, though it's only at the hour level.

3.1.3   Exceptional clearance

3.1.4   Number of other segments

3.1.4.1   Offense segments

3.1.4.2   Victim segments

3.1.4.3   Offender segments

3.1.4.4   Arrestee segments

Figure 3.2: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.
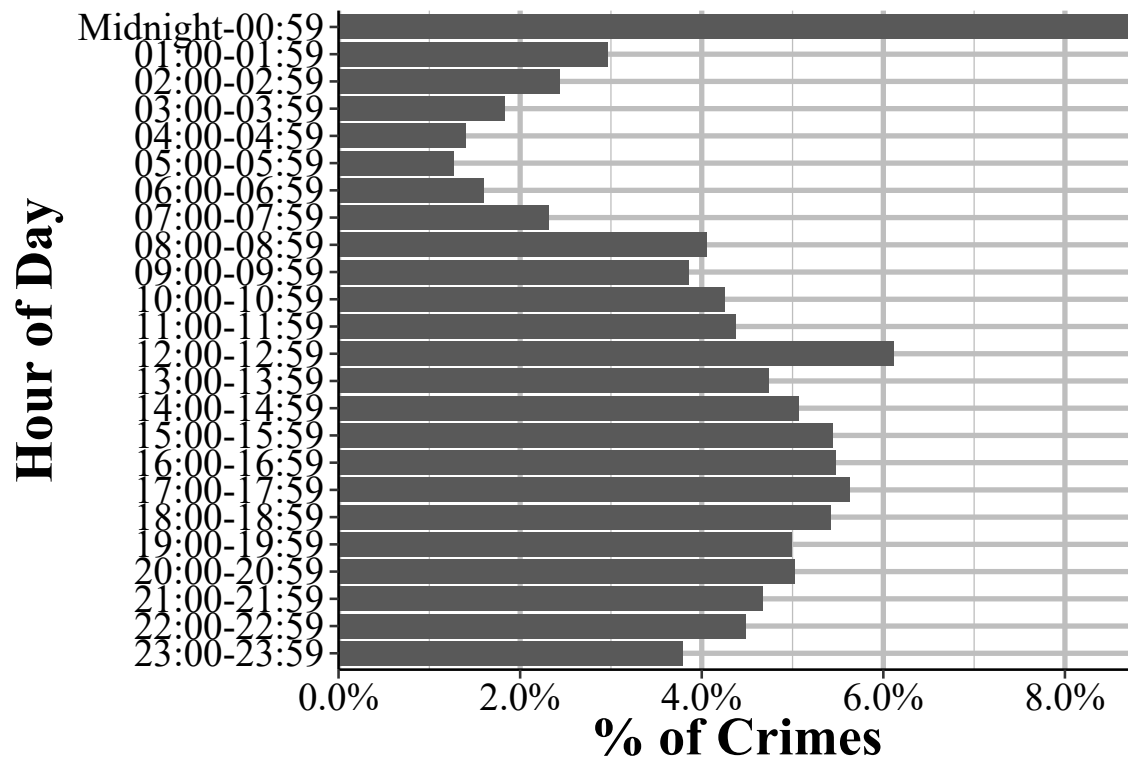
Figure 3.3: The percent of crimes that are reported each hour for all agencies reporting to NIBRS in 2019.
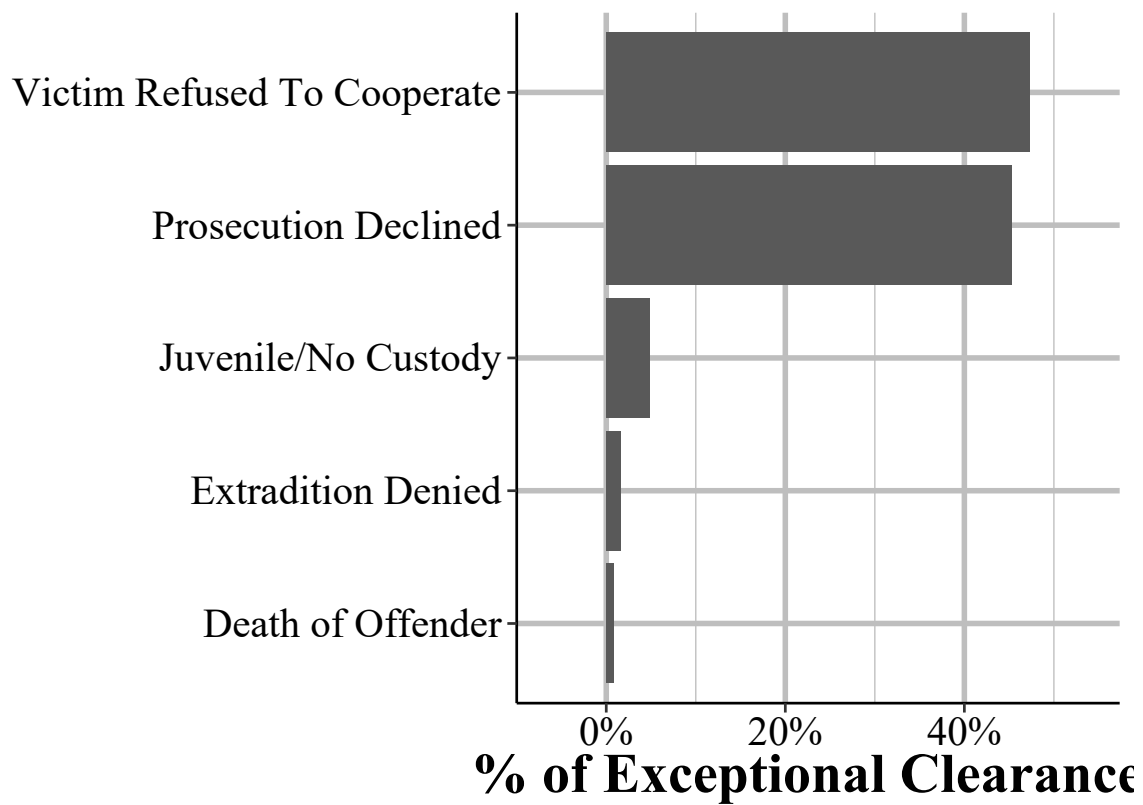
Figure 3.4: The percent of crimes that are reported each hour for all agencies reporting to NIBRS in 2019.
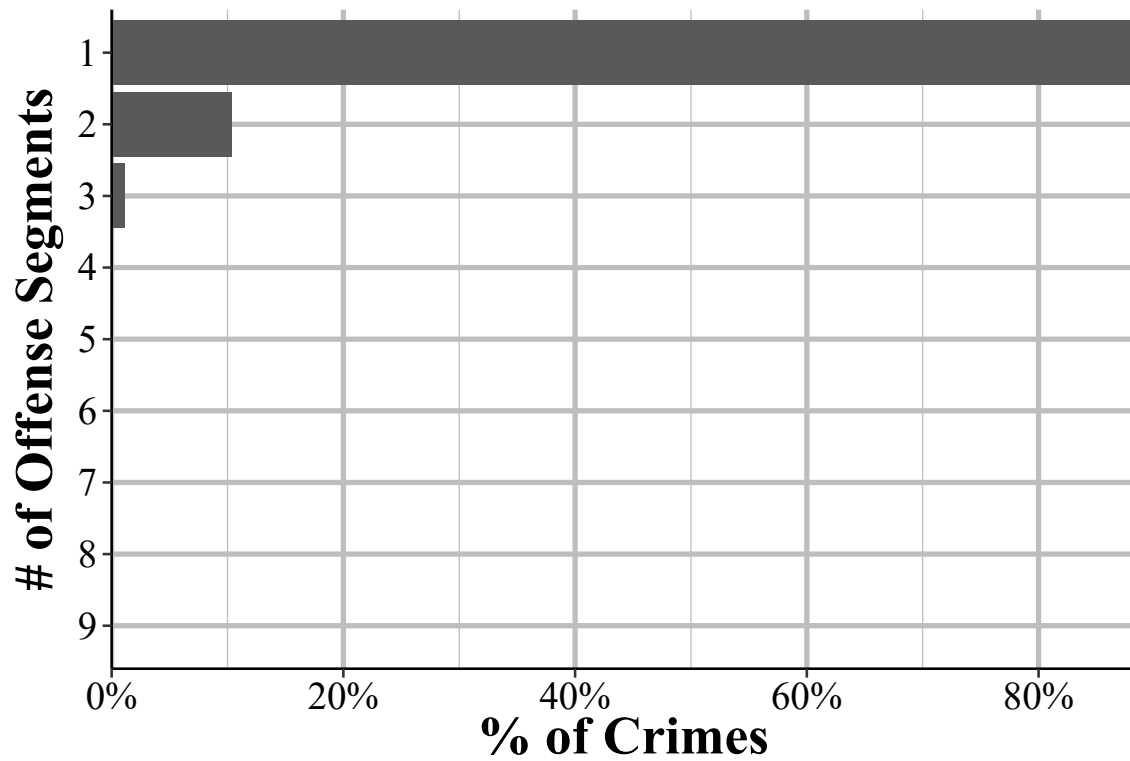
Figure 3.5: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.
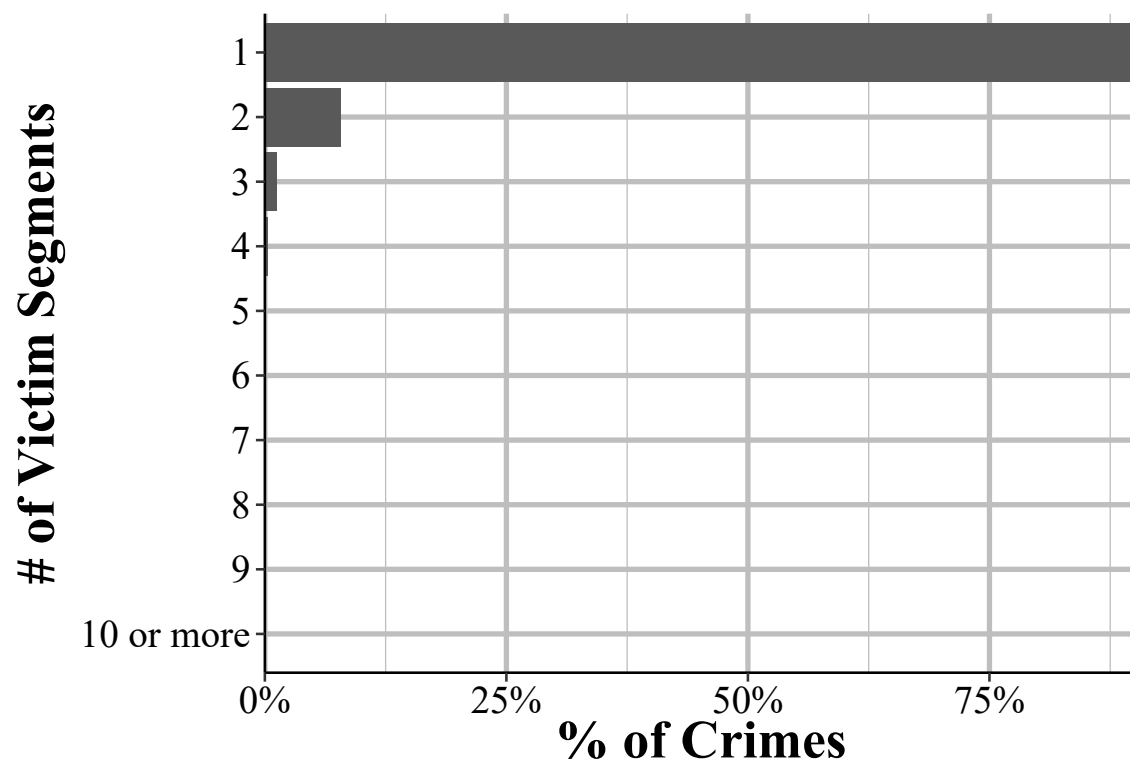
Figure 3.6: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.
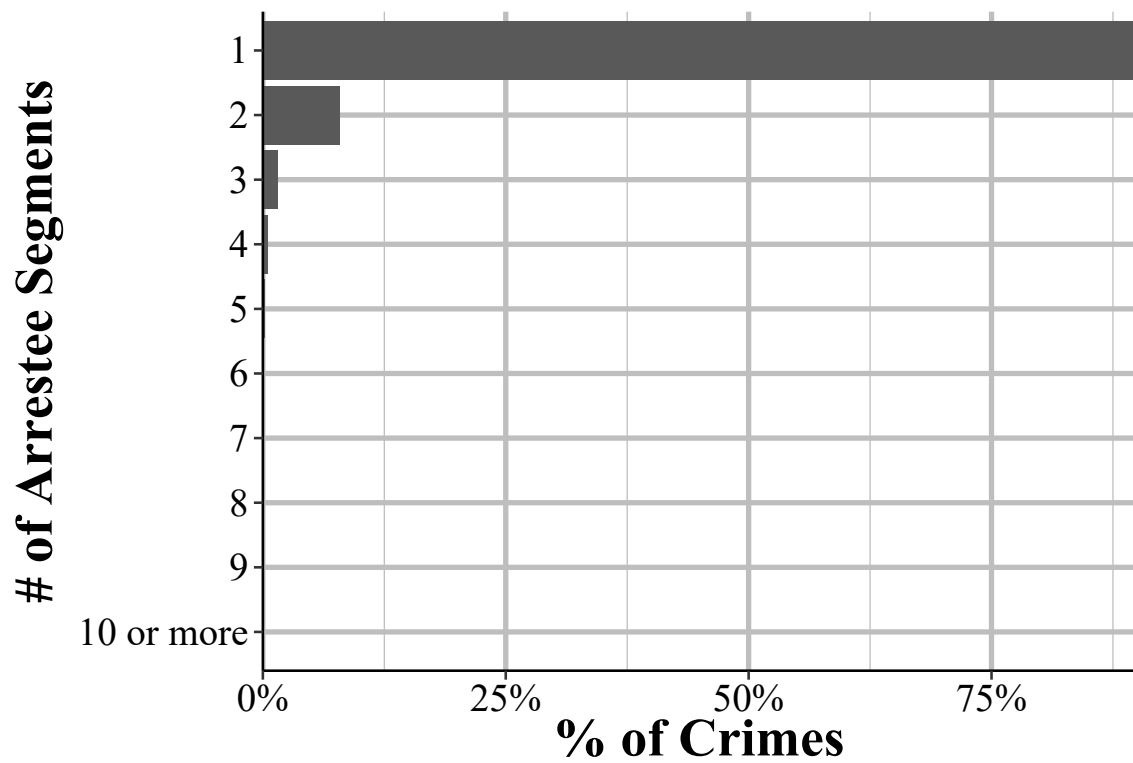
Figure 3.7: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.
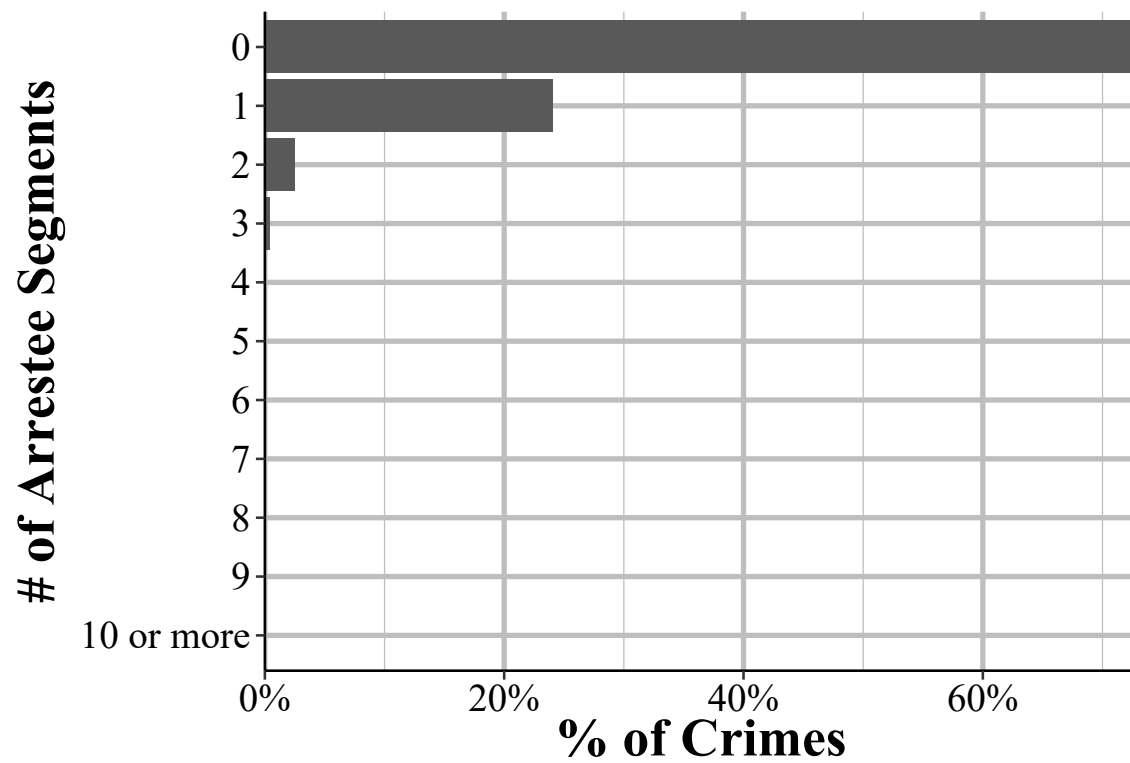
Figure 3.8: The percent of crimes that are reported each day of the month for all agencies reporting to NIBRS in 2019.

Chapter 4

Offense Segment

4.1 Important variables

4.1.1 Crime category

\begin{figure}

Show 100 ▼ entries   Search: [   ]

	Crime Category	# of Offenses	% of Offenses
1	Simple Assault	944,601	12.70%
2	Destruction/Damage/Vandalism of Property	821,523	11.05%
3	All Other Larceny	810,138	10.89%
4	Drug/Narcotic Violations	710,822	9.56%
5	Theft From Motor Vehicle	550,202	7.40%
6	Shoplifting	487,848	6.56%
7	Burglary/Breaking And Entering	476,554	6.41%
8	Drug Equipment Violations	327,715	4.41%
9	Motor Vehicle Theft	304,964	4.10%
10	Intimidation	279,040	3.75%
11	Aggravated Assault	271,444	3.65%
12	Theft From Building	212,507	2.86%
13	False Pretenses/Swindle/Confidence Game	195,003	2.62%
14	Credit Card/Atm Fraud	134,527	1.81%
15	Weapon Law Violations	134,182	1.80%
16	Counterfeiting/Forgery	112,766	1.52%
17	Theft of Motor Vehicle Parts/Accessories	111,384	1.50%
18	Robbery	92,935	1.25%
19	Identity Theft	68,515	0.92%
20	Stolen Property Offenses (Receiving, Selling, Etc.)	58,449	0.79%
21	Impersonation	57,920	0.78%
22	Fondling (Incident Liberties/Child Molest)	46,141	0.62%
23	Rape	46,049	0.62%
24	Embezzlement	24,166	0.32%
25	Kidnapping/Abduction	22,757	0.31%
26	Pornography/Obscene Material	19,814	0.27%
27	Arson	15,500	0.21%
28	Wire Fraud	15,373	0.21%
29	Sodomy	11,046	0.15%
30	Animal Cruelty	9,956	0.13%
31	Pocket-Picking	9,527	0.13%
32	Purse-Snatching	7,487	0.10%
33	Prostitution	7,309	0.10%
34	Murder/Nonnegligent Manslaughter	6,095	0.08%
35	Extortion/Blackmail	5,583	0.08%
36	Statutory Rape	5,430	0.07%
37	Theft From Coin-Operated Machine Or Device	4,617	0.06%
38	Sexual Assault With An Object	4,599	0.06%
39	Hacking/Computer Invasion	2,765	0.04%
40	Assisting Or Promoting Prostitution	2,283	0.03%
41	Purchasing Prostitution	1,159	0.02%
42	Welfare Fraud	940	0.01%
43	Human Trafficking - Commercial Sex Acts	915	0.01%
44	Incest	762	0.01%
45	Operating/Promoting/Assisting Gambling	613	0.01%
46	Negligent Manslaughter	561	0.01%
47	Bribery	544	0.01%
48	Gambling Equipment Violations	351	0.00%
49	Justifiable Homicide	300	0.00%
50	Betting/Wagering	260	0.00%
51	Human Trafficking - Involuntary Servitude	142	0.00%
52	Sports Tampering	7	0.00%

Showing 1 to 52 of 52 entries   Previous 1 Next

{

4.1.2   Offense subtype

\begin{figure}

	Crime Subcategory	# of Offenses	% of Offenses
1	Possessing/Concealing	1,074,646	44.35%
2	None/Unknown Gang Involvement (Mutually Exclusive)	1,040,062	42.92%
3	Distributing/Selling	100,708	4.16%
4	Using/Consuming	90,049	3.72%
5	Buying/Receiving	46,692	1.93%
6	Cultivating/Manufacturing/Publishing	26,289	1.08%
7	Operating/Promoting/Assisting	12,446	0.51%
8	Transporting/Transmitting/Importing	7,821	0.32%
9	Simple/Gross Neglect (Unintentionally, Intentionally, Or Knowingly Failing To Provide Food, Water, Shelter, Veterinary Care, Hoarding, Etc.)	6,996	0.29%
10	Other Gang	6,482	0.27%
11	Exploiting Children	5,448	0.22%
12	Intentional Abuse And Torture (Tormenting, Mutilating, Poisoning, Or Abandonment)	2,770	0.11%
13	Juvenile Gang Involvement	2,582	0.11%
14	Animal Sexual Abuse (Bestiality)	104	0.00%
15	Organized Abuse (Dog Fighting And Cock Fighting)	86	0.00%

Show 100 entries   Search:

Showing 1 to 15 of 15 entries   Previous 1 Next

{

- Aggravated Assault
- Animal Cruelty
- Counterfeiting/Forgery
- Drug Equipment Violations
- Drug/Narcotic Violations
- Fondling (Incident Liberties/Child Molest)
- Gambling Equipment Violations
- Intimidation
- Kidnapping/Abduction
- Murder/Nonnegligent Manslaughter
- Negligent Manslaughter
- Pornography/Obscene Material
- Rape
- Robbery

- Sexual Assault With An Object
- Simple Assault
- Sodomy
- Stolen Property Offenses (Receiving, Selling, Etc.)
- Weapon Law Violations

\begin{figure}

Show 100 ▾ entries   Search:

	Crime Subcategory	# of Offenses	% of Offenses
1	Simple/Gross Neglect (Unintentionally, Intentionally, Or Knowingly Failing To Provide Food, Water, Shelter, Veterinary Care, Hoarding, Etc.)	6,996	70.27%
2	Intentional Abuse And Torture (Tormenting, Mutilating, Poisoning, Or Abandonment)	2,770	27.82%
3	Animal Sexual Abuse (Bestiality)	104	1.04%
4	Organized Abuse (Dog Fighting And Cock Fighting)	86	0.86%

Showing 1 to 4 of 4 entries   Previous 1 Next

{

4.1.3   Drug, alcohol, or computer use
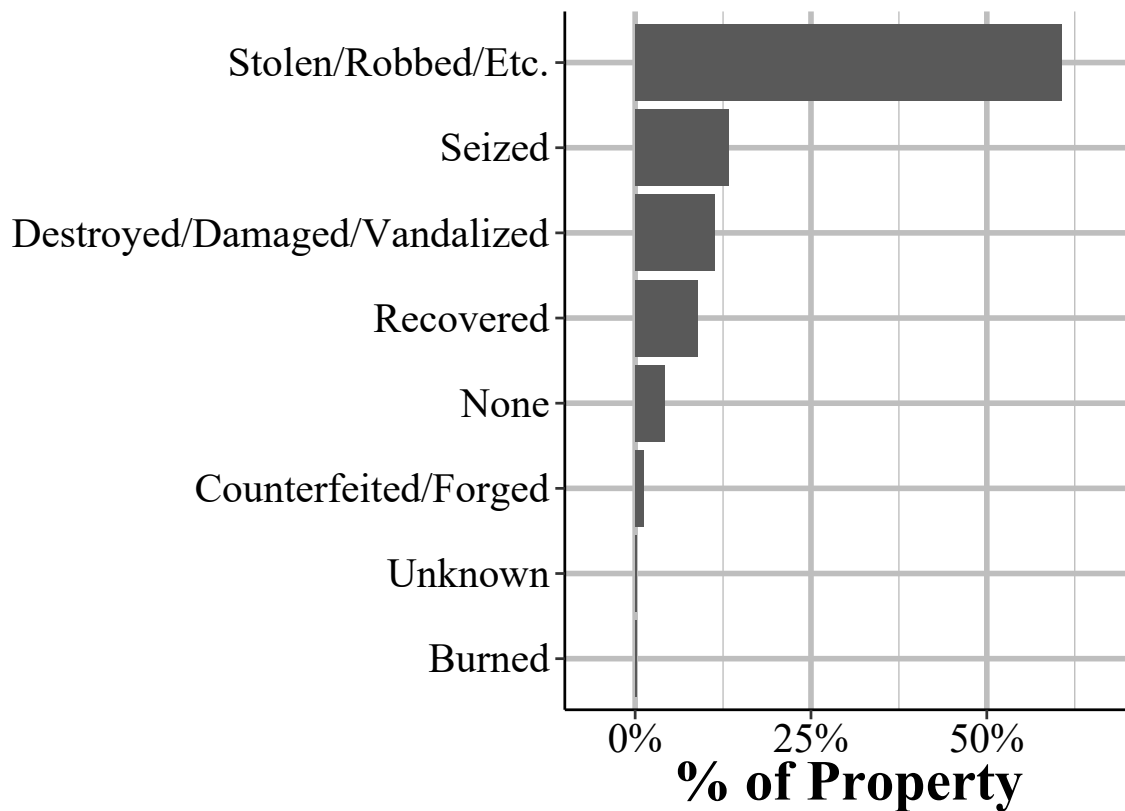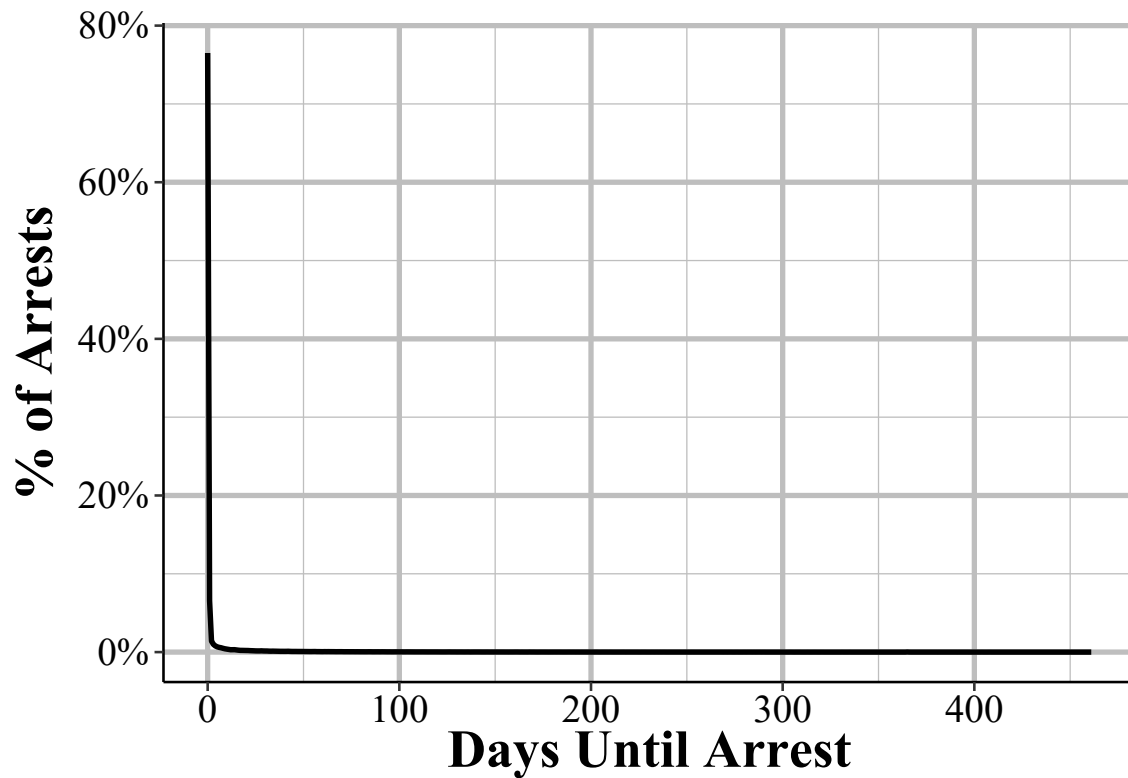
4.1.4   Crime location

\begin{figure}

Show 100 ▼ entries   Search: [   ]

	Crime Location	# of Offenses	% of Offenses
1	Residence/Home	3,001,159	40.36%
2	Highway/Road/Alley	1,188,722	15.99%
3	Parking Lot/Garage	637,986	8.58%
4	Department/Discount Store	373,338	5.02%
5	Other/Unknown	349,477	4.70%
6	Convenience Store	174,952	2.35%
7	Grocery/Supermarket	171,702	2.31%
8	Commercial/Office Building	159,915	2.15%
9	Specialty Store (Tv, Fur, Etc.)	144,856	1.95%
10	Restaurant	132,347	1.78%
11	School - Elementary/Secondary	119,901	1.61%
12	Hotel/Motel/Etc.	118,313	1.59%
13	Service/Gas Station	103,901	1.40%
14	Drug Store/Doctors Office/Hospital	74,861	1.01%
15	Bank/Savings And Loan	67,959	0.91%
16	Government/Public Building	59,118	0.80%
17	Bar/Nightclub	55,220	0.74%
18	Park/Playground	54,226	0.73%
19	Field/Woods	40,011	0.54%
20	School - College/University	36,947	0.50%
21	Jail/Prison	35,363	0.48%
22	Construction Site	34,744	0.47%
23	Rental Storage Facility	34,630	0.47%
24	Shopping Mall	32,532	0.44%
25	School/College	31,330	0.42%
26	Cyberspace	29,830	0.40%
27	Air/Bus/Train Terminal	28,409	0.38%
28	Church/Synagogue/Temple	26,470	0.36%
29	Auto Dealership New/Used	20,643	0.28%
30	Liquor Store	15,237	0.20%
31	Industrial Site	11,044	0.15%
32	Gambling Facility/Casino/Race Track	9,688	0.13%
33	Shelter - Mission/Homeless	7,252	0.10%
34	Community Center	7,191	0.10%
35	Lake/Waterway	7,020	0.09%
36	Farm Facility	5,992	0.08%
37	Arena/Stadium/Fairgrounds/Coliseum	5,256	0.07%
38	Camp/Campground	5,251	0.07%
39	Abandoned/Condemned Structure	4,705	0.06%
40	Atm Separate From Bank	4,597	0.06%
41	Daycare Facility	4,260	0.06%
42	Amusement Park	3,429	0.05%
43	Dock/Wharf/Freight/Model Terminal	3,383	0.05%
44	Rest Area	2,236	0.03%
45	Tribal Lands	414	0.01%
46	Military Installation	273	0.00%

Showing 1 to 46 of 46 entries   Previous   1   Next

{

4.1.5   Weapons

\begin{figure}

	Weapon Used	# of Offenses	% of Offenses
1	Personal Weapons (Hands, Feet, Teeth, Etc.)	918,134	57.84%
2	None	161,623	10.18%
3	Handgun	149,146	9.40%
4	Other	90,385	5.69%
5	Knife/Cutting Instrument (Ice Pick, Screwdriver, Ax, Etc.)	71,311	4.49%
6	Firearm (Type Not Stated)	58,920	3.71%
7	Unknown	57,430	3.62%
8	Blunt Object (Club, Hammer, Etc.)	36,536	2.30%
9	Motor Vehicle	17,278	1.09%
10	Rifle	7,696	0.48%
11	Shotgun	5,896	0.37%
12	Asphyxiation (By Drowning, Strangulation, Suffocation, Gas, Etc.)	4,984	0.31%
13	Other Firearm	4,633	0.29%
14	Drugs/Narcotics/Sleeping Pills	1,239	0.08%
15	Explosives	949	0.06%
16	Fire/Incendiary Device	715	0.05%
17	Poison (Include Gas)	475	0.03%

Show 100 entries   Search:

Showing 1 to 17 of 17 entries   Previous 1 Next

{

- Aggravated Assault
- Extortion/Blackmail
- Fondling (Incident Liberties/Child Molest)
- Human Trafficking - Commercial Sex Acts
- Human Trafficking - Involuntary Servitude
- Justifiable Homicide
- Kidnapping/Abduction
- Murder/Nonnegligent Manslaughter
- Negligent Manslaughter
- Rape
- Robbery
- Sexual Assault With An Object
- Simple Assault

- Sodomy
- Weapon Law Violations

4.1.6   Automatic weapons

This variable only tells you if the weapon is automatic
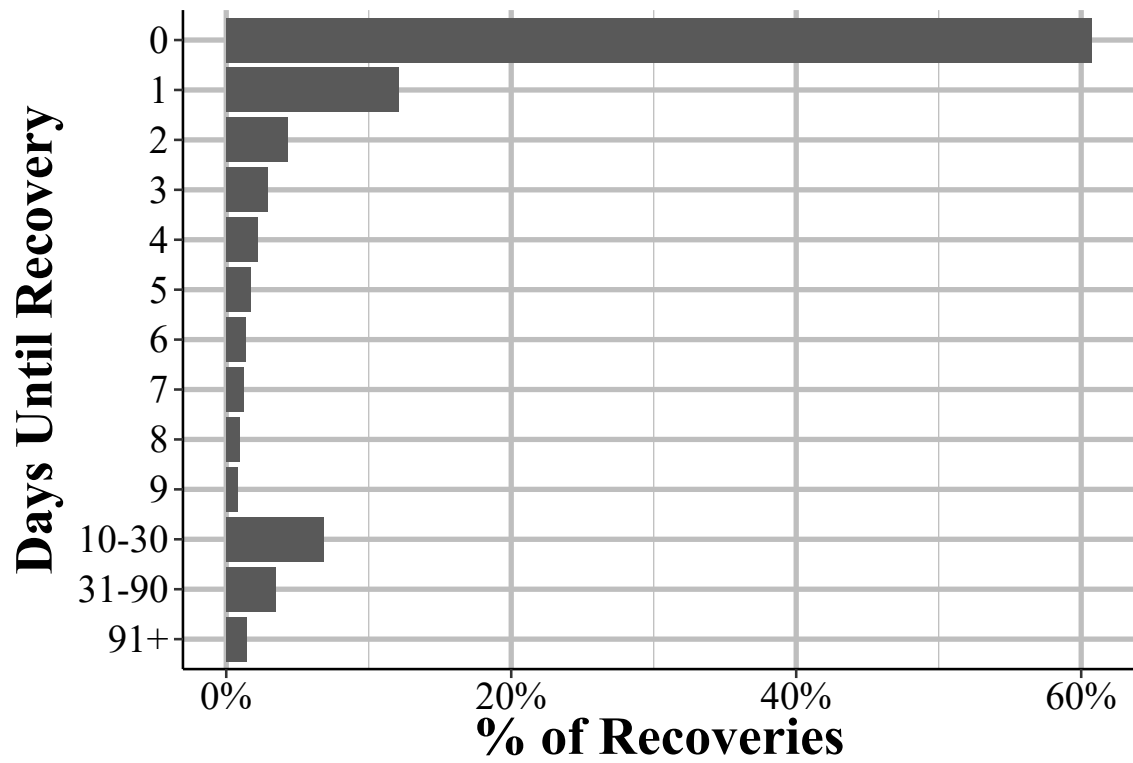
4.1.7   Hate crime indicator (bias motivation)

\begin{figure}

Show 100 ▼ entries   Search: [   ]

	Bias Motivation ⬍	# of Offenses ⬍	% of Offenses ⬍
1	No Bias Motivation	7,372,214	99.14%
2	Unknown Bias Motivation	59,360	0.80%
3	Anti-Black	1,309	0.02%
4	Anti-White	573	0.01%
5	Anti-Male Homosexual (Gay)	357	0.00%
6	Anti-Hispanic	310	0.00%
7	Anti-Jewish	249	0.00%
8	Anti-Homosexual (Both Gay And Lesbian)	213	0.00%
9	Anti-Other Ethnicity/National Origin	210	0.00%
10	Anti-American Indian Or Alaskan Native	125	0.00%
11	Anti-Mental Disability	107	0.00%
12	Anti-Asian/Pacific Islander	101	0.00%
13	Anti-Multi-Racial Group	97	0.00%
14	Anti-Islamic (Muslim)	86	0.00%
15	Anti-Female Homosexual (Lesbian)	70	0.00%
16	Anti-Transgender	68	0.00%
17	Anti-Sikh	66	0.00%
18	Anti-Arab	64	0.00%
19	Anti-Other Religion	63	0.00%
20	Anti-Gender Non-Conforming	59	0.00%
21	Anti-Female	51	0.00%
22	Anti-Eastern Orthodox (Greek, Russian, Etc.)	50	0.00%
23	Anti-Physical Disability	44	0.00%
24	Anti-Catholic	42	0.00%
25	Anti-Other Christian	40	0.00%
26	Anti-Native Hawaiian Or Other Pacific Islander	30	0.00%
27	Anti-Protestant	26	0.00%
28	Anti-Multi-Religious Group	25	0.00%
29	Anti-Bisexual	22	0.00%
30	Anti-Male	17	0.00%
31	Anti-Heterosexual	12	0.00%
32	Anti-Mormon	7	0.00%
33	Anti-Hindu	7	0.00%
34	Anti-Buddhist	6	0.00%
35	Anti-Atheism/Agnosticism	5	0.00%
36	Anti-Jehovahs Witness	5	0.00%

Showing 1 to 36 of 36 entries   Previous   1   Next

{

\begin{figure}

Show 100 ▼ entries   Search: [   ]

	Bias Motivation	# of Offenses	% of Offenses
1	Anti-Black	1,309	28.99%
2	Anti-White	573	12.69%
3	Anti-Male Homosexual (Gay)	357	7.91%
4	Anti-Hispanic	310	6.86%
5	Anti-Jewish	249	5.51%
6	Anti-Homosexual (Both Gay And Lesbian)	213	4.72%
7	Anti-Other Ethnicity/National Origin	210	4.65%
8	Anti-American Indian Or Alaskan Native	125	2.77%
9	Anti-Mental Disability	107	2.37%
10	Anti-Asian/Pacific Islander	101	2.24%
11	Anti-Multi-Racial Group	97	2.15%
12	Anti-Islamic (Muslim)	86	1.90%
13	Anti-Female Homosexual (Lesbian)	70	1.55%
14	Anti-Transgender	68	1.51%
15	Anti-Sikh	66	1.46%
16	Anti-Arab	64	1.42%
17	Anti-Other Religion	63	1.40%
18	Anti-Gender Non-Conforming	59	1.31%
19	Anti-Female	51	1.13%
20	Anti-Eastern Orthodox (Greek, Russian, Etc.)	50	1.11%
21	Anti-Physical Disability	44	0.97%
22	Anti-Catholic	42	0.93%
23	Anti-Other Christian	40	0.89%
24	Anti-Native Hawaiian Or Other Pacific Islander	30	0.66%
25	Anti-Protestant	26	0.58%
26	Anti-Multi-Religious Group	25	0.55%
27	Anti-Bisexual	22	0.49%
28	Anti-Male	17	0.38%
29	Anti-Heterosexual	12	0.27%
30	Anti-Mormon	7	0.16%
31	Anti-Hindu	7	0.16%
32	Anti-Buddhist	6	0.13%
33	Anti-Atheism/Agnosticism	5	0.11%
34	Anti-Jehovahs Witness	5	0.11%

Showing 1 to 34 of 34 entries   Previous | 1 | Next

{

Chapter 5

Offender Segment

As might be expected, the Offender Segment provides information about who the offender is for each incident, though this is limited to only demographic variables. So we know the age, sex, and race of each offender but nothing else. This means that important variables such as criminal history, ethnicity, socioeconomic status, and motive are missing. In the Victim Segment we learn about the relationship between the victim and offender, and in the Offense Segment we learn which weapon (if any) the offender used. So there is some other data on the offender in other segments but its quite limited. This data has one row per offender so incidents with multiple offenders have multiple rows. In cases where there is no information about the offender there will be a single row where all of the offender variables will be "unknown". In these cases having a single row for the offender is merely a placeholder and doesn't necessarily mean that there was only one offender for that incident.

5.1 Important variables

5.1.1 Demographics

5.1.1.1   Age
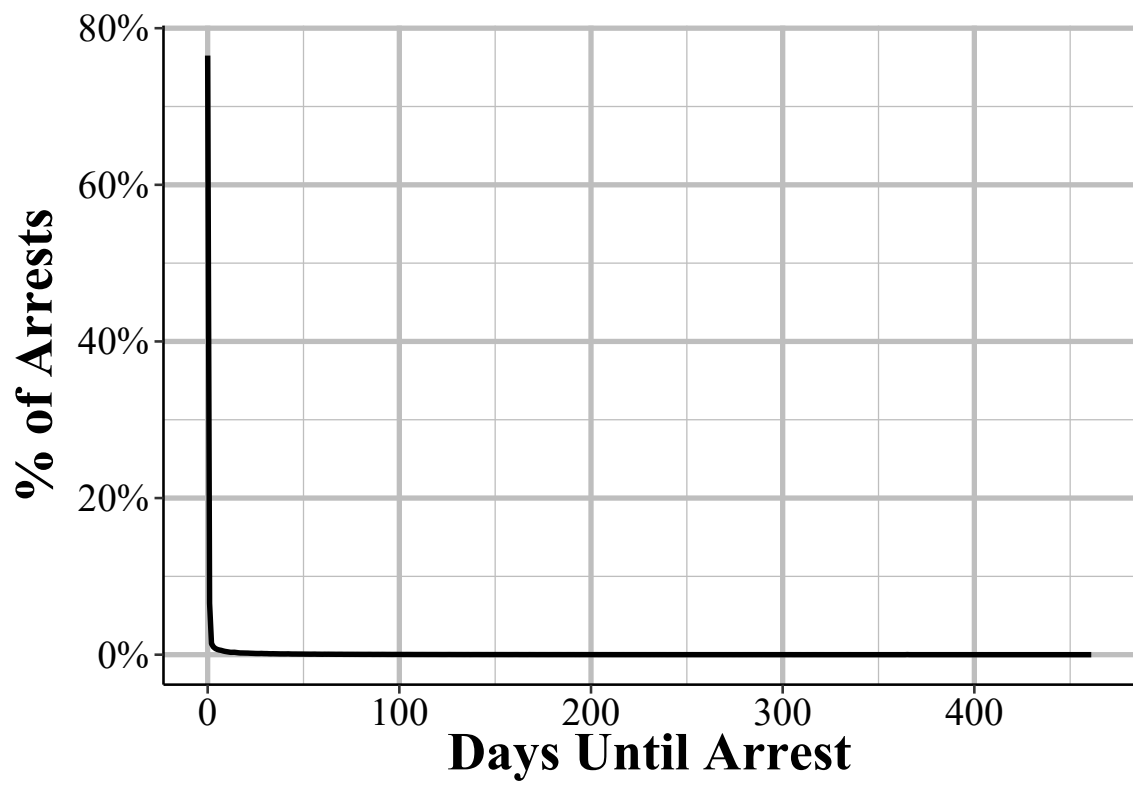
5.1.1.2   Sex
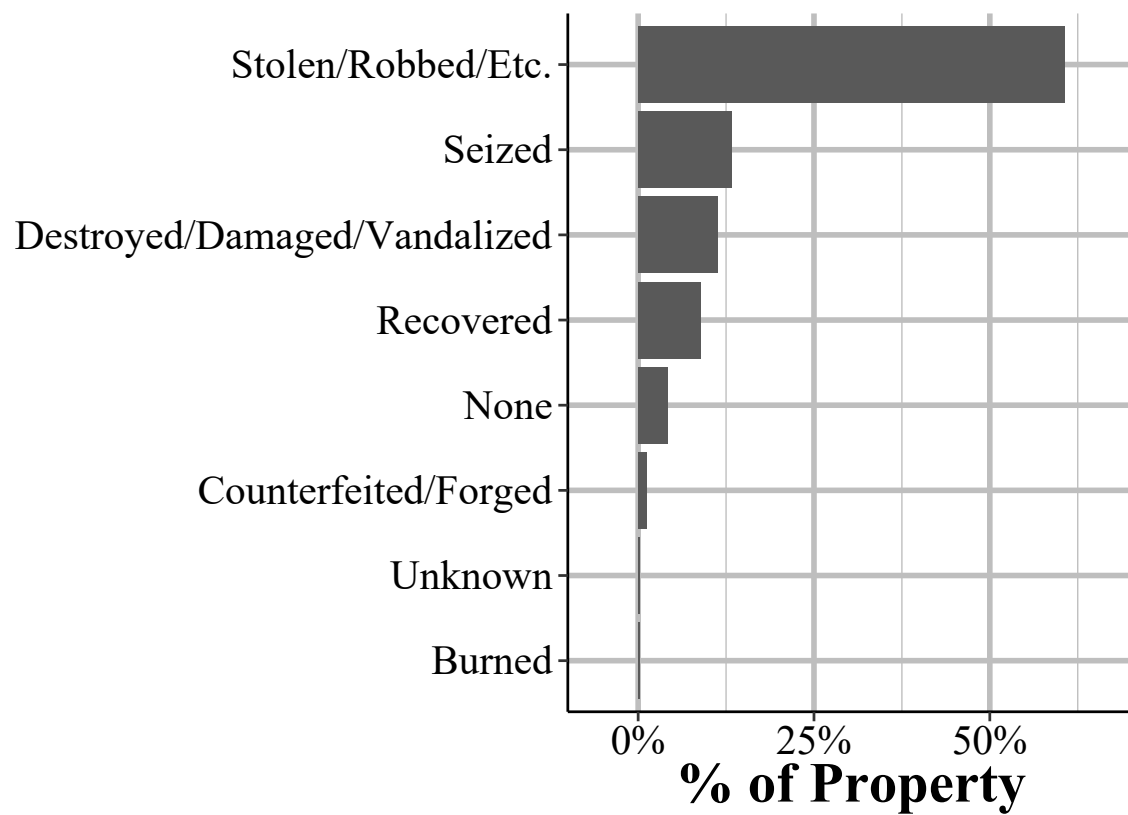
5.1.1.3   Race
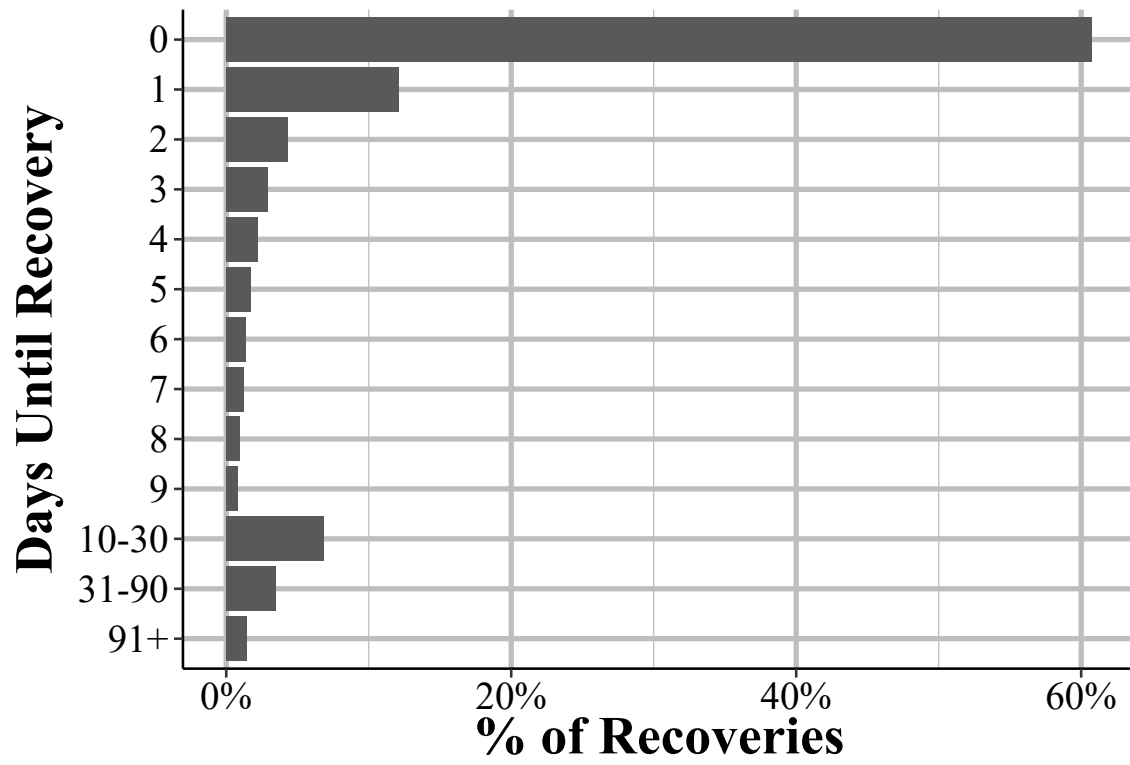
Chapter 6

Victim Segment

6.1 Important variables

6.1.1 Crime category
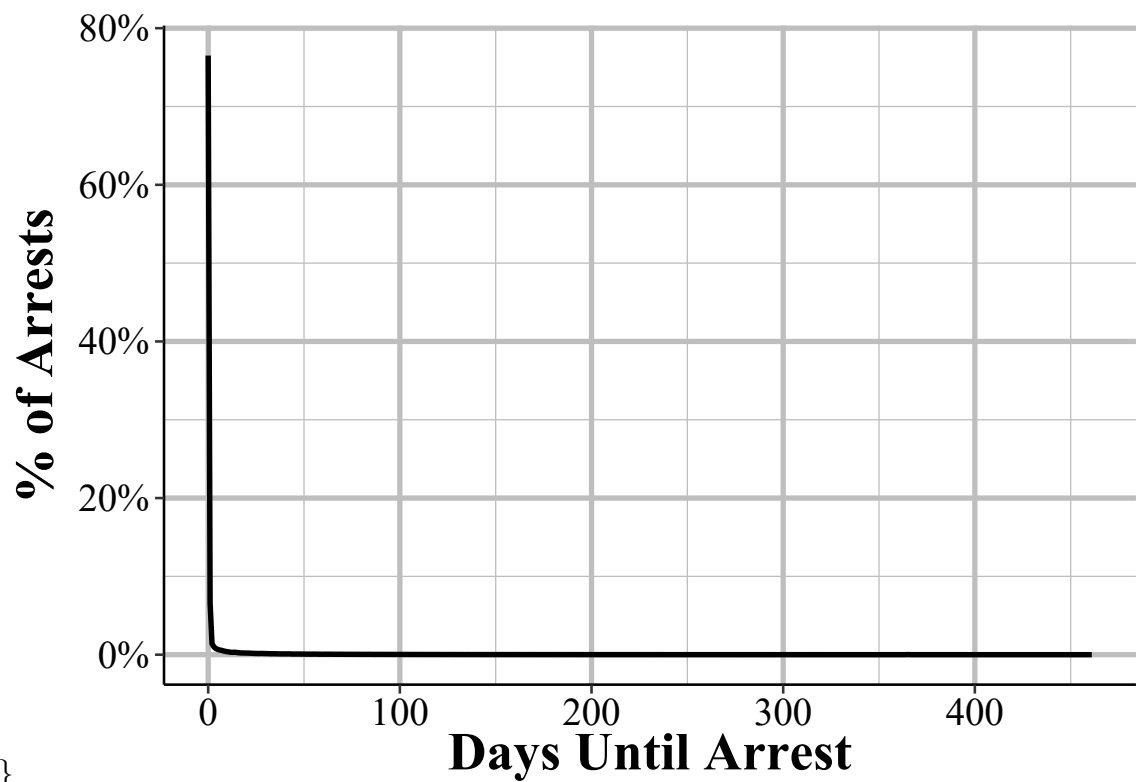
\begin{center}

6.1.2   Victim type
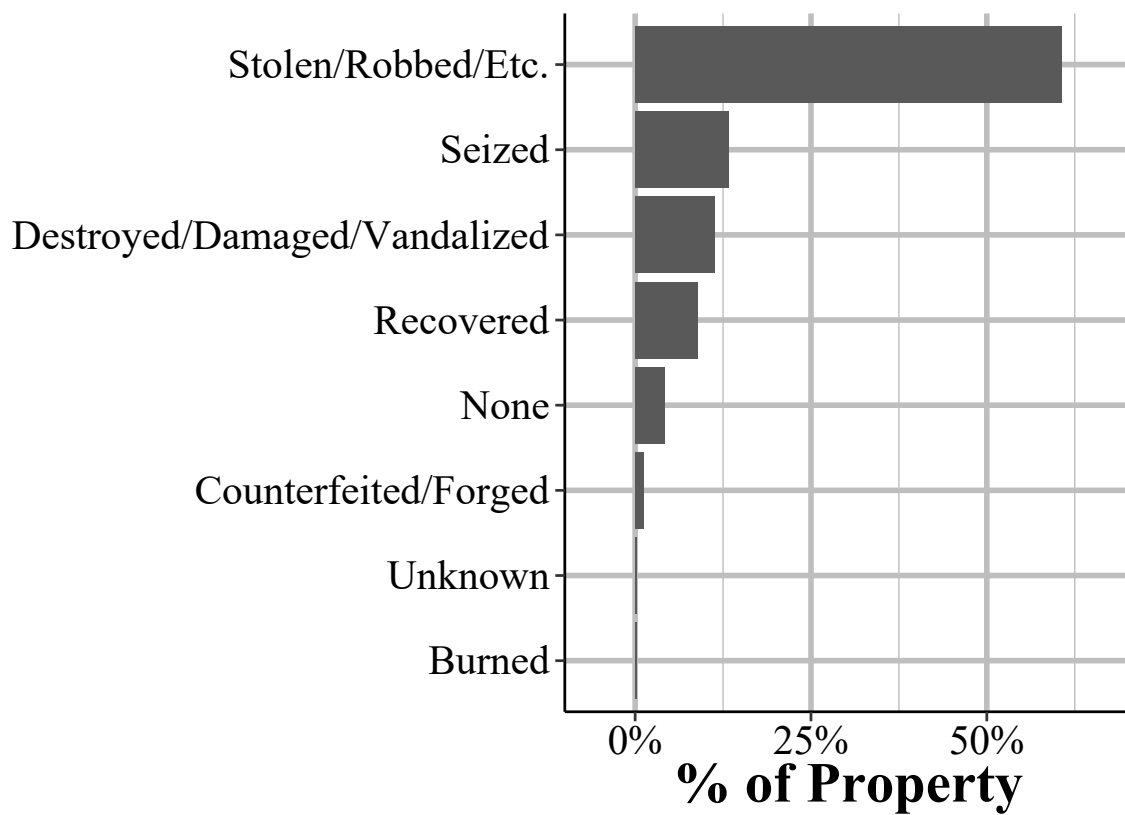
6.1.3   Injury
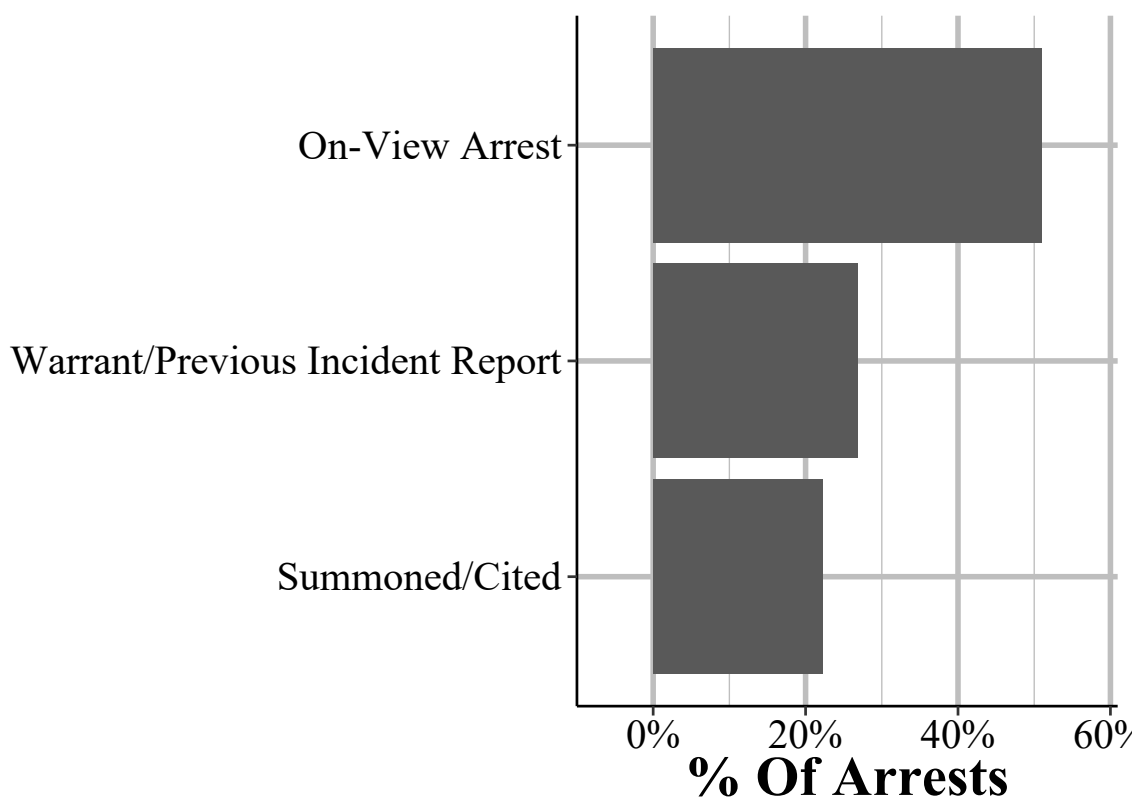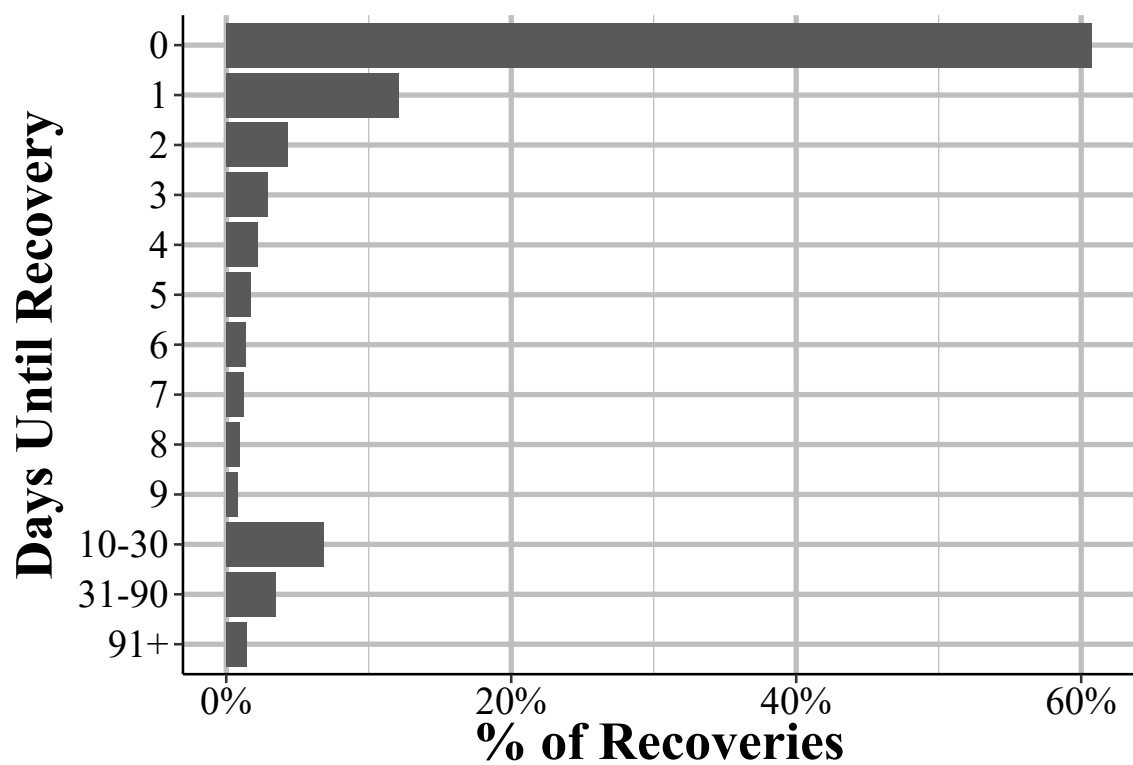
6.1.4   Relationship to offender
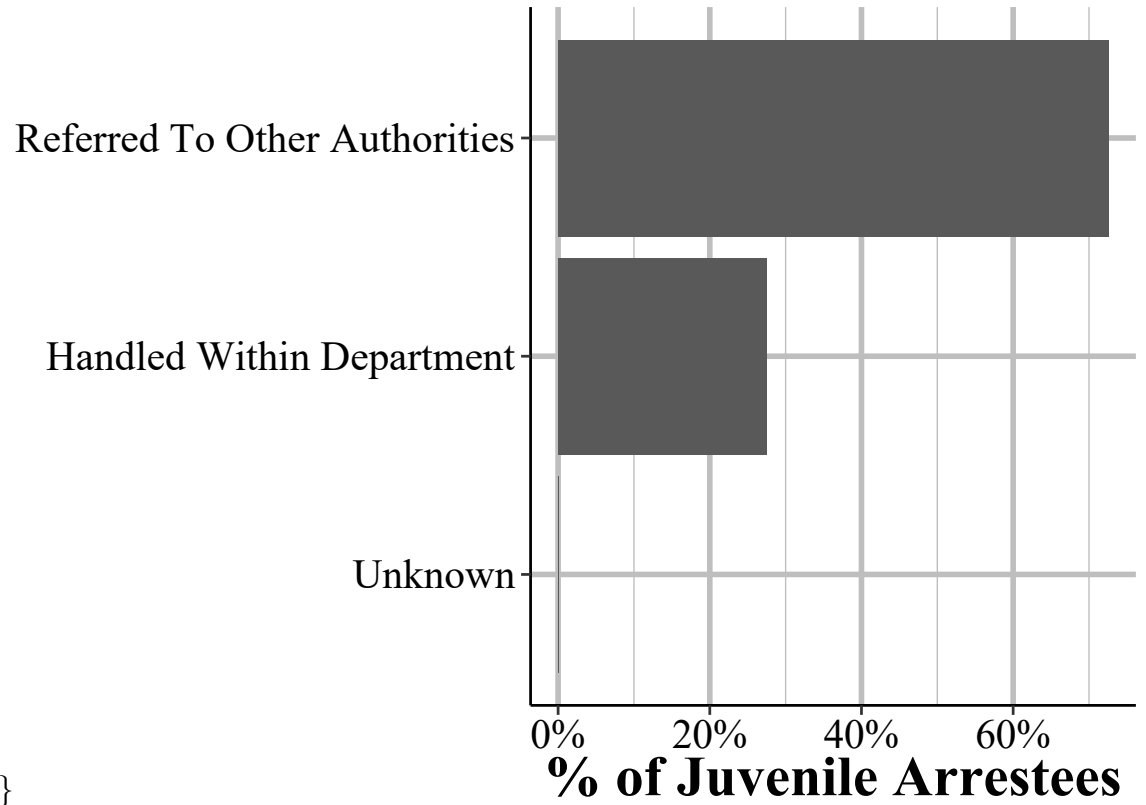
\begin{center}

6.1.5   Residence status

Only for when victim is individual or law enforcement officer

victim$resident_status_of_victim
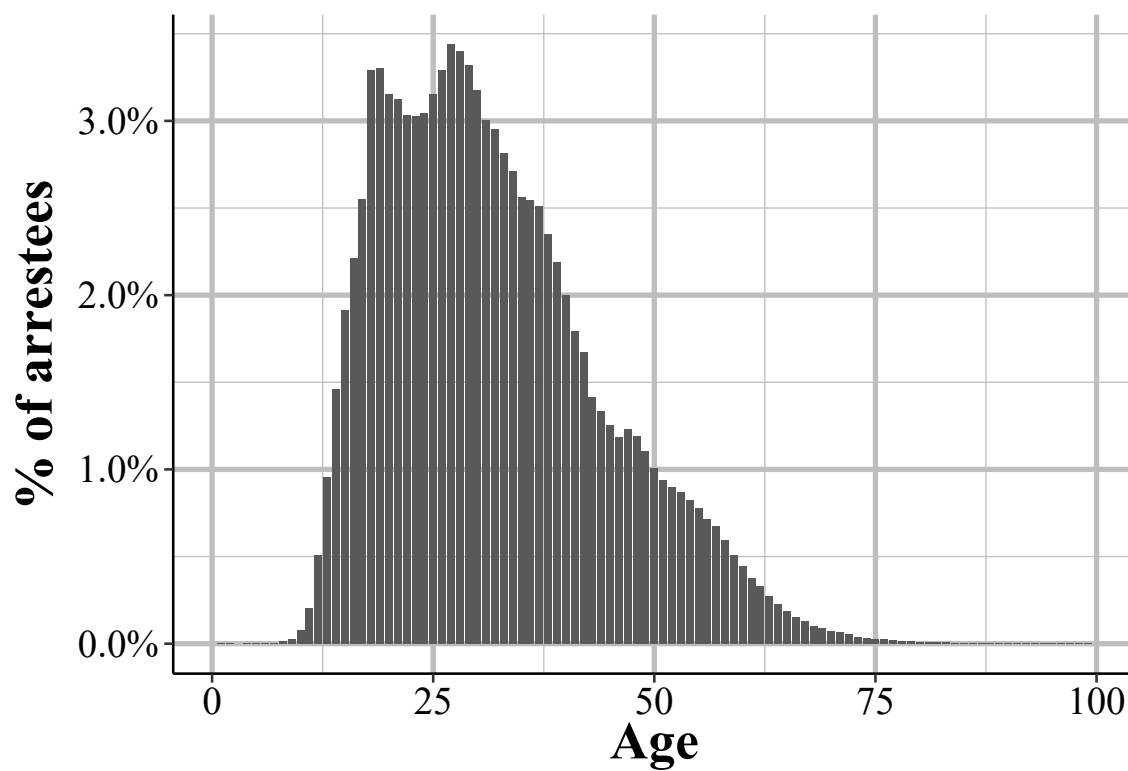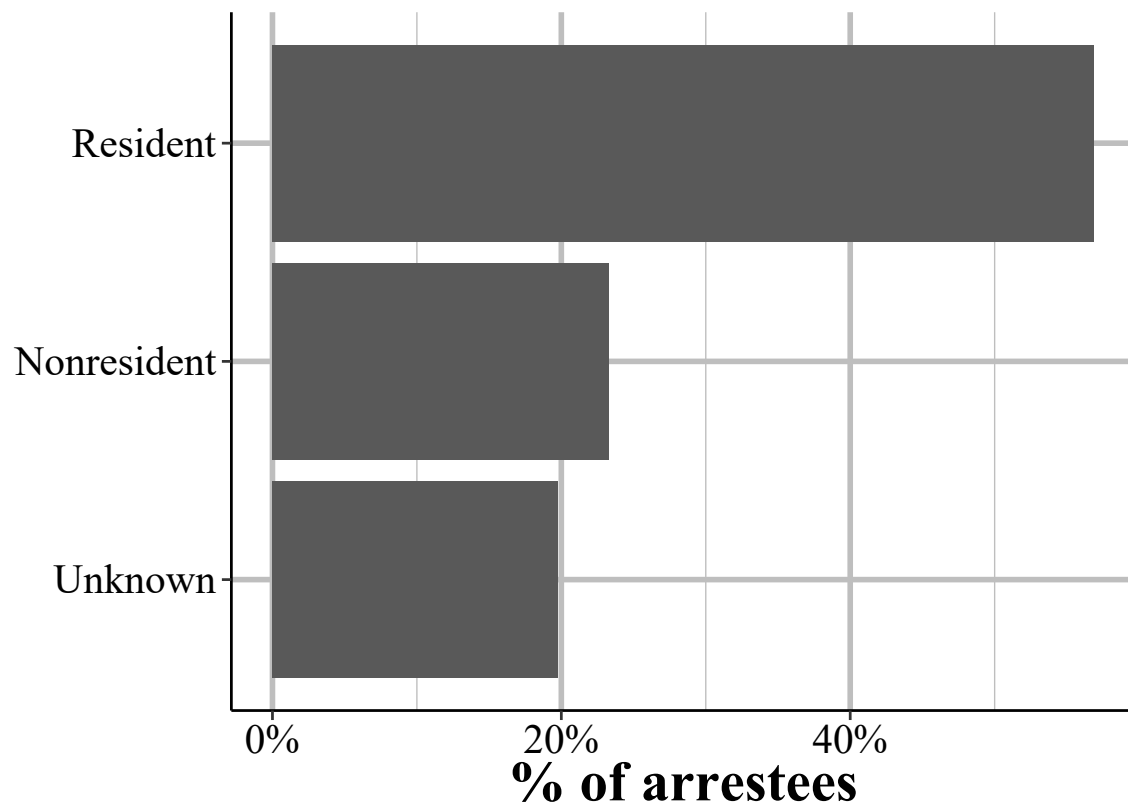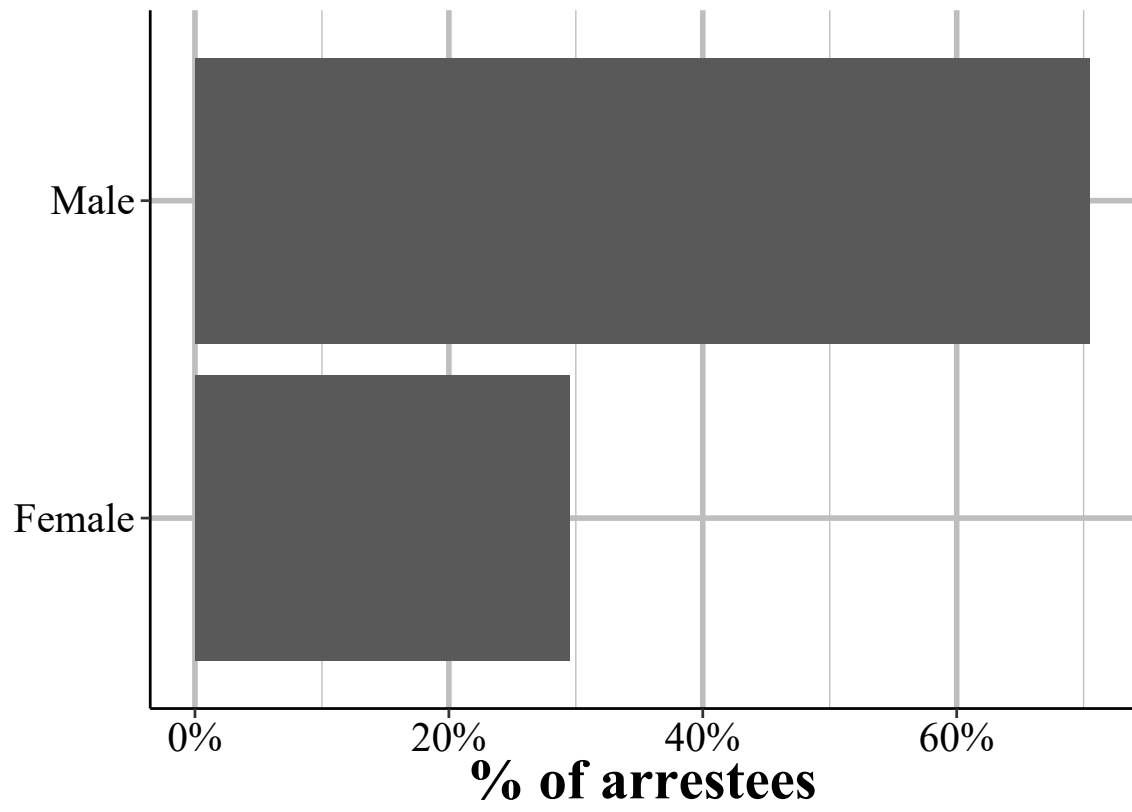
% of arrestees

6.1.6 Aggravated assault and homicide circumstances
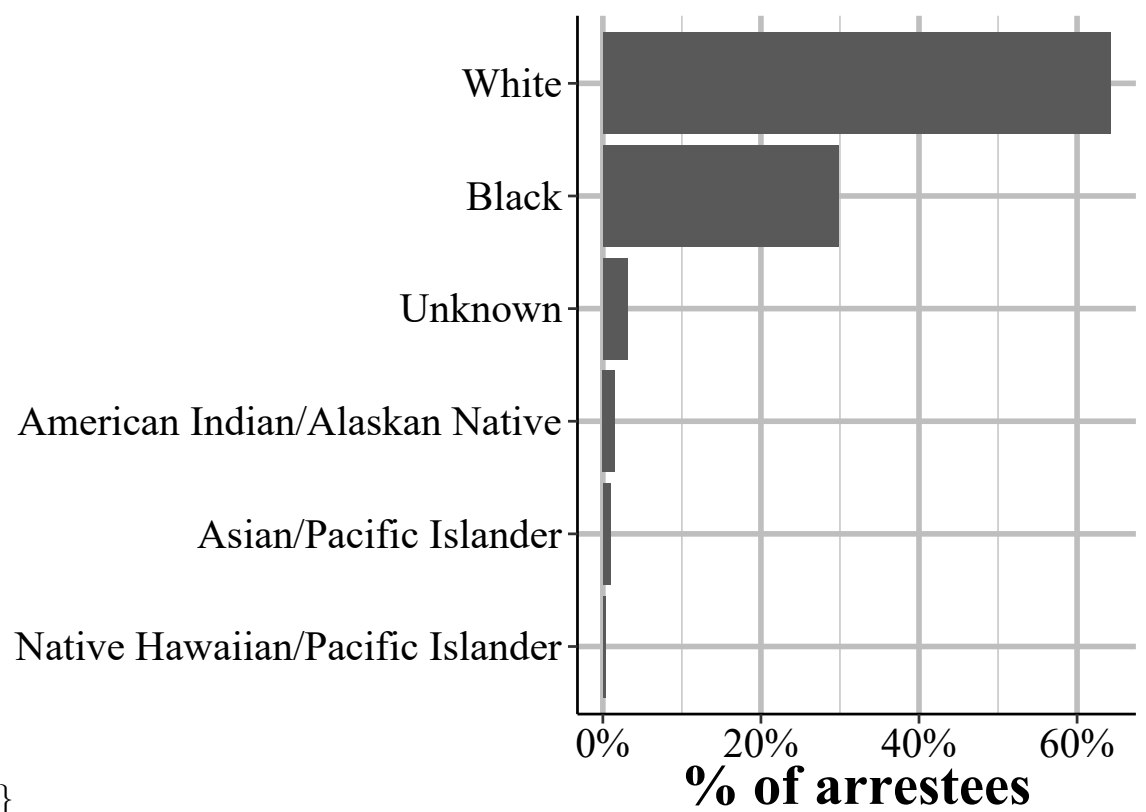
\begin{center}

6.1.7   Justifiable homicide circumstance
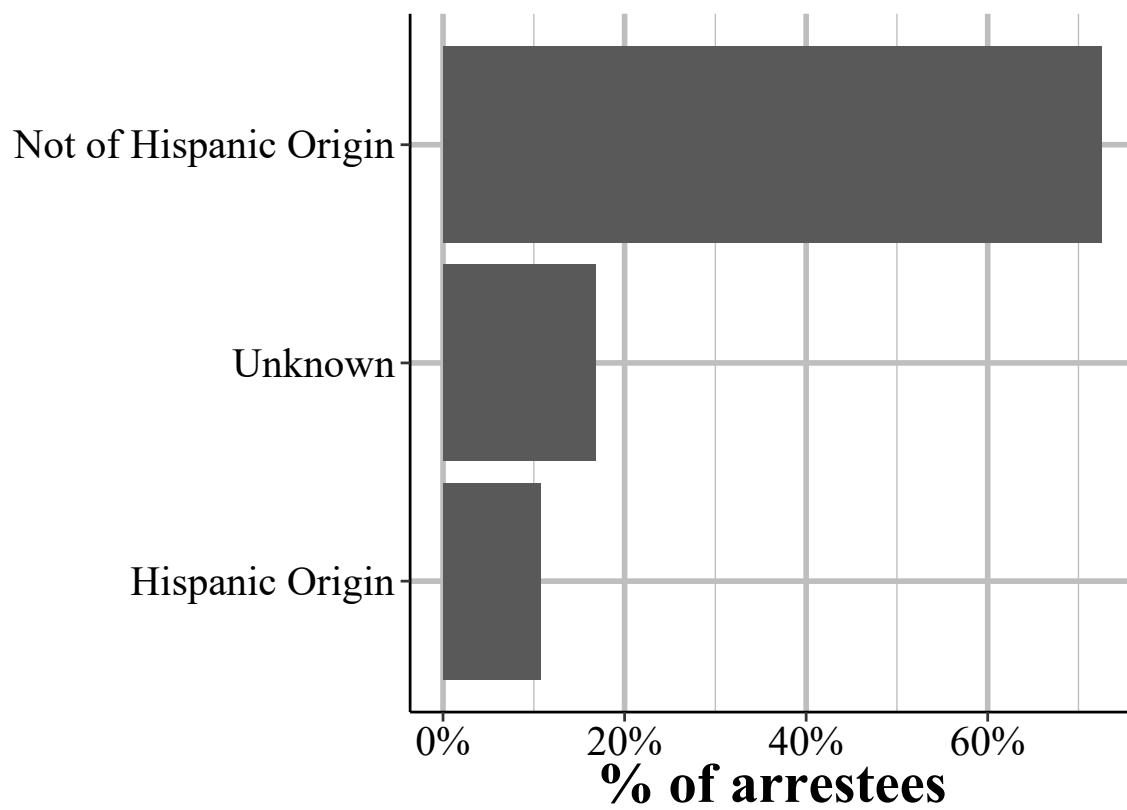
6.1.8 Age
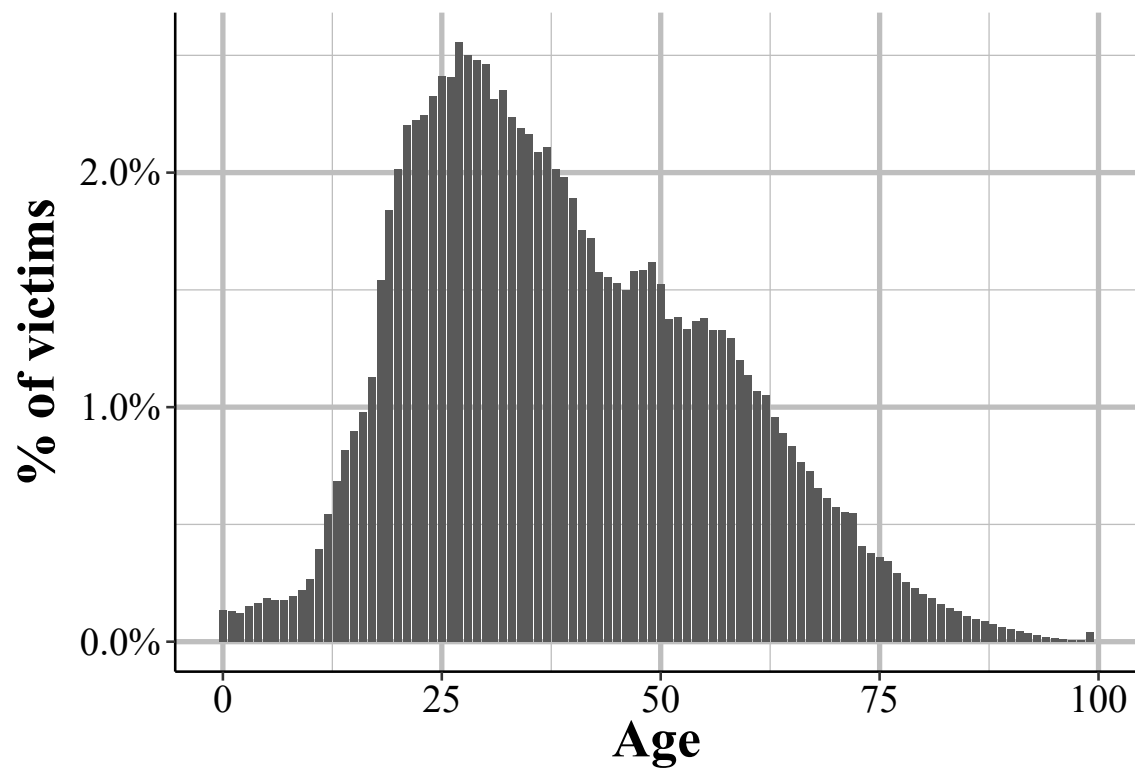
6.1.9　Sex
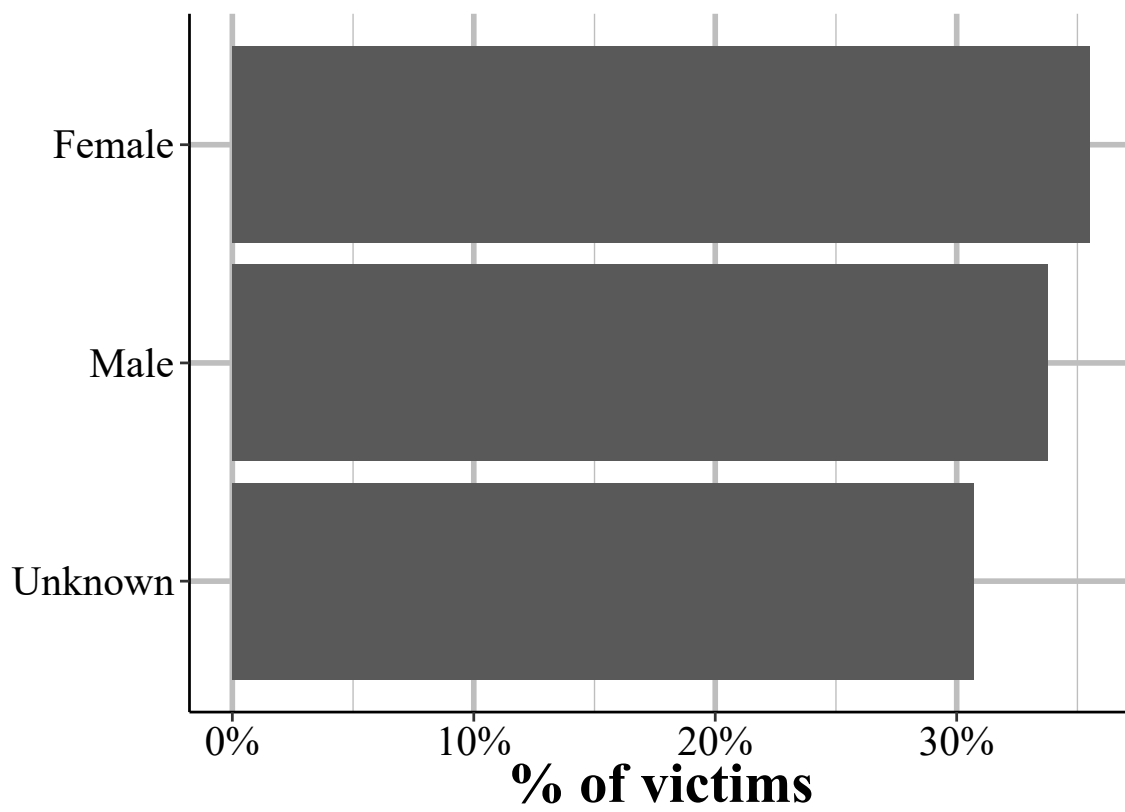
6.1.10   Race
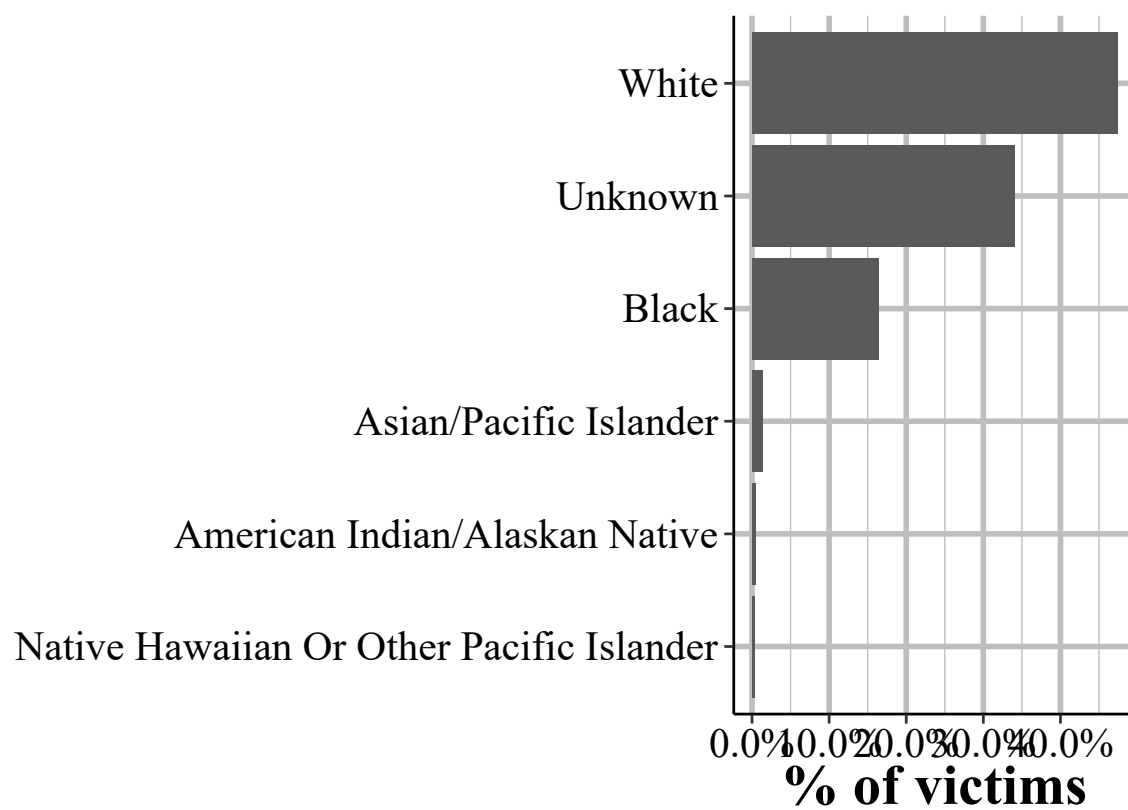
6.1.11   Ethnicity
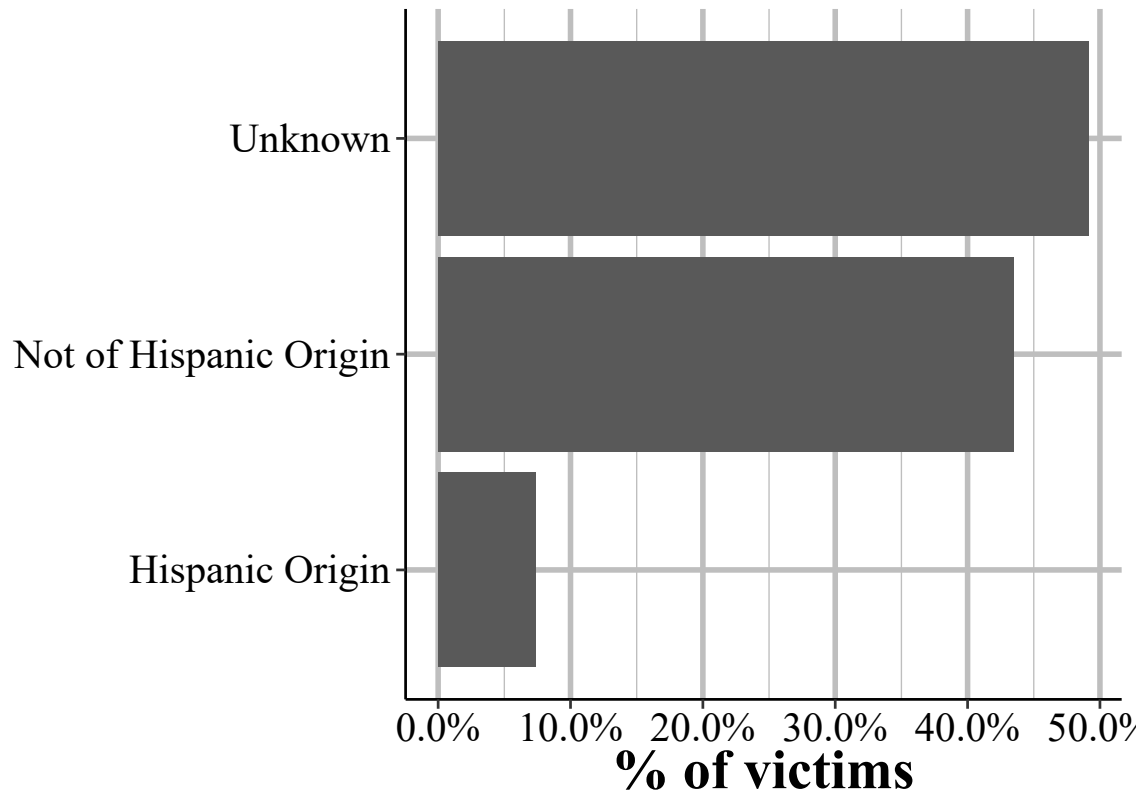
Chapter 7

Arrestee, Group B Arrestee, and Window Arrestee Segment

7.1 Important variables

7.1.1 Crime arrested for (excluding Group B arrests)

\begin{figure}

Show 100 ▾ entries   Search:

	Crime Category	# of Offenses	% of Offenses
1	Drug/Narcotic Violations	523,732	25.82%
2	Simple Assault	385,695	19.02%
3	Shoplifting	228,355	11.26%
4	Aggravated Assault	127,192	6.27%
5	All Other Larceny	104,244	5.14%
6	Drug Equipment Violations	95,730	4.72%
7	Destruction/Damage/Vandalism of Property	69,153	3.41%
8	Burglary/Breaking And Entering	56,613	2.79%
9	Intimidation	52,972	2.61%
10	Weapon Law Violations	51,907	2.56%
11	All Other Offenses	49,288	2.43%
12	Stolen Property Offenses (Receiving, Selling, Etc.)	32,953	1.62%
13	Motor Vehicle Theft	28,489	1.40%
14	Robbery	25,590	1.26%
15	False Pretenses/Swindle/Confidence Game	23,183	1.14%
16	Theft From Motor Vehicle	18,780	0.93%
17	Counterfeiting/Forgery	18,022	0.89%
18	Theft From Building	15,899	0.78%
19	Disorderly Conduct	11,413	0.56%
20	Driving Under The Influence	10,084	0.50%
21	Impersonation	9,429	0.46%
22	Kidnapping/Abduction	8,941	0.44%
23	Credit Card/Atm Fraud	6,846	0.34%
24	Fondling (Incident Liberties/Child Molest)	6,659	0.33%
25	Trespass of Real Property	6,439	0.32%
26	Rape	6,406	0.32%
27	Embezzlement	6,327	0.31%
28	Prostitution	5,492	0.27%
29	Murder/Nonnegligent Manslaughter	4,788	0.24%
30	Liquor Law Violations	4,267	0.21%
31	Identity Theft	3,802	0.19%
32	Drunkenness	3,753	0.19%
33	Pocket-Picking	3,137	0.15%
34	Pornography/Obscene Material	3,055	0.15%
35	Arson	3,016	0.15%
36	Family Offenses, Nonviolent	2,498	0.12%
37	Theft of Motor Vehicle Parts/Accessories	2,026	0.10%
38	Animal Cruelty	1,852	0.09%
39	Assisting Or Promoting Prostitution	1,442	0.07%
40	Sodomy	1,374	0.07%
41	Statutory Rape	1,137	0.06%
42	Purse-Snatching	866	0.04%
43	Curfew/Loitering/Vagrancy Violations	840	0.04%
44	Sexual Assault With An Object	647	0.03%
45	Purchasing Prostitution	609	0.03%
46	Theft From Coin-Operated Machine Or Device	426	0.02%
47	Negligent Manslaughter	327	0.02%
48	Operating/Promoting/Assisting Gambling	262	0.01%
49	Betting/Wagering	262	0.01%
50	Extortion/Blackmail	254	0.01%
51	Welfare Fraud	241	0.01%
52	Human Trafficking - Commercial Sex Acts	232	0.01%
53	Bribery	230	0.01%
54	Bad Checks	210	0.01%
55	Wire Fraud	180	0.01%
56	Incest	150	0.01%
57	Runaway	90	0.00%
58	Gambling Equipment Violations	89	0.00%
59	Hacking/Computer Invasion	66	0.00%
60	Peeping Tom	39	0.00%
61	Human Trafficking - Involuntary Servitude	27	0.00%
62	Sports Tampering	1	0.00%

Showing 1 to 62 of 62 entries   Previous   1   Next

{

7.1.2   Group B Crimes arrested for

\begin{figure}

Show 100 ▾ entries   Search: [   ]

	Crime Category	# of Offenses	% of Offenses
1	All Other Offenses	1,095,755	57.26%
2	Driving Under The Influence	351,926	18.39%
3	Disorderly Conduct	117,707	6.15%
4	Drunkenness	116,343	6.08%
5	Trespass of Real Property	108,546	5.67%
6	Liquor Law Violations	68,862	3.60%
7	Family Offenses, Nonviolent	31,251	1.63%
8	Runaway	9,535	0.50%
9	Curfew/Loitering/Vagrancy Violations	9,360	0.49%
10	Bad Checks	3,911	0.20%
11	Peeping Tom	414	0.02%

Showing 1 to 11 of 11 entries   Previous   1   Next

{

7.1.3 Arrest date
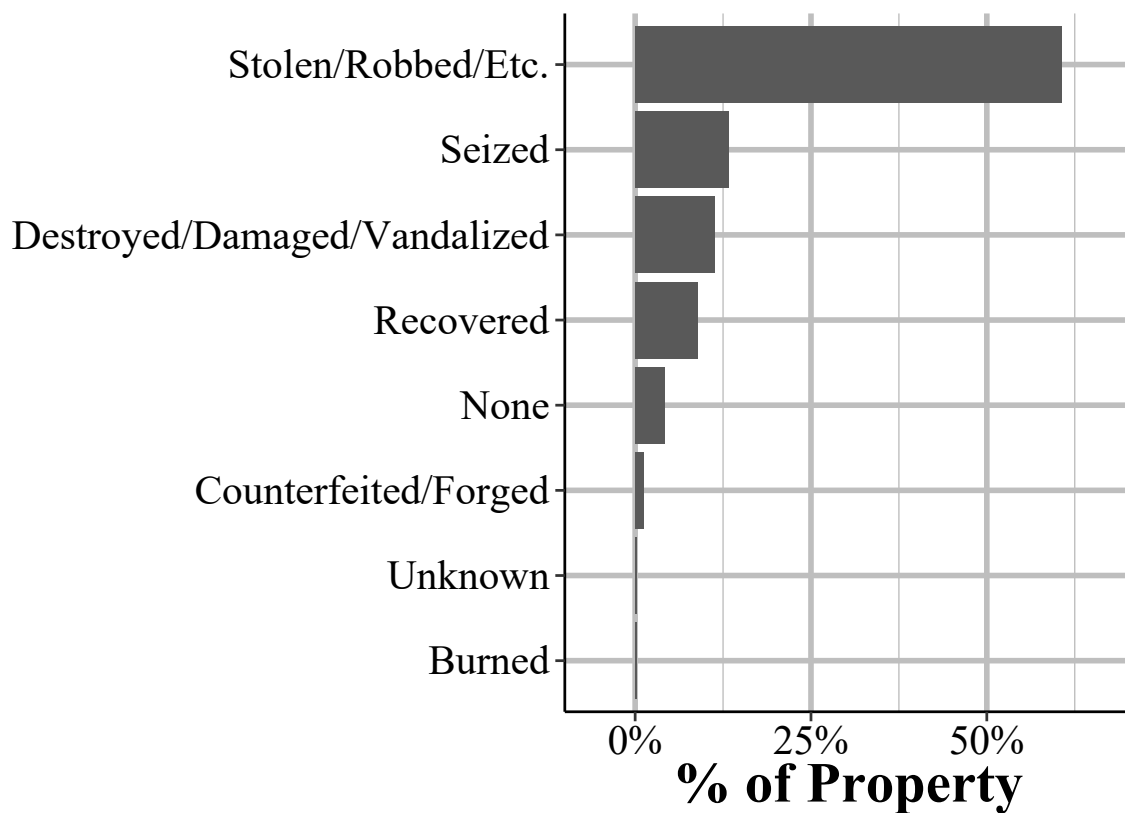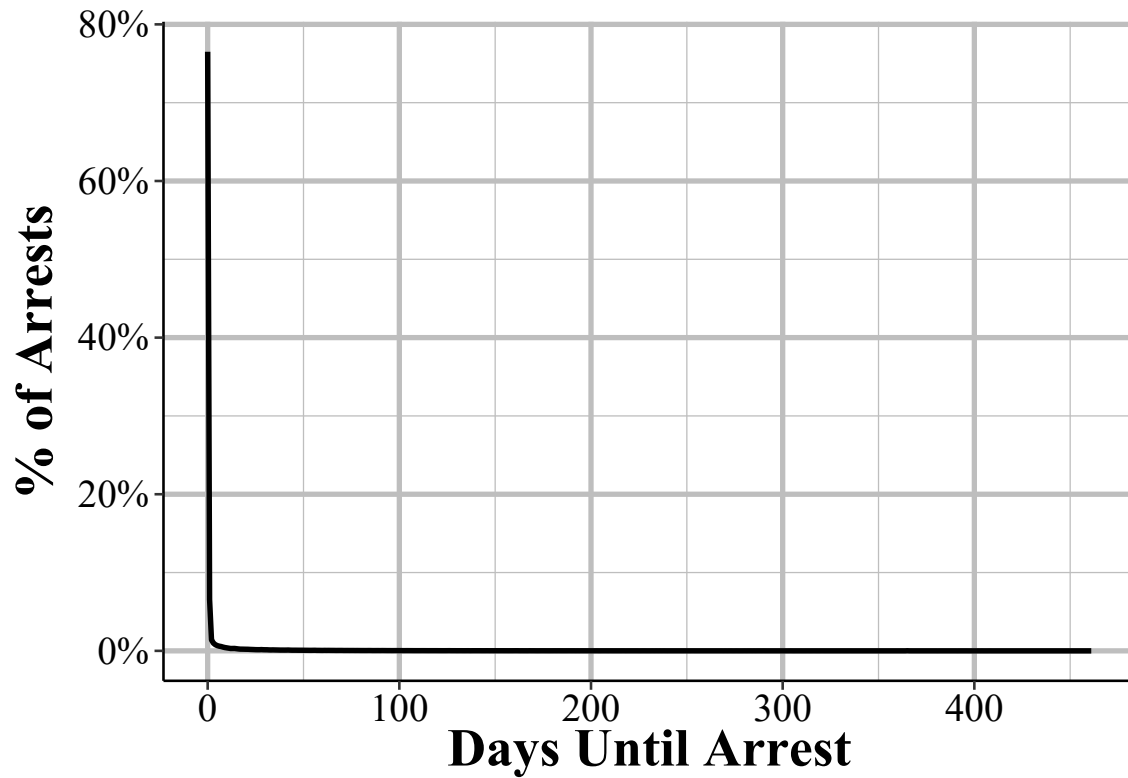
7.1.4 Weapons

All crimes can potentially have a weapon
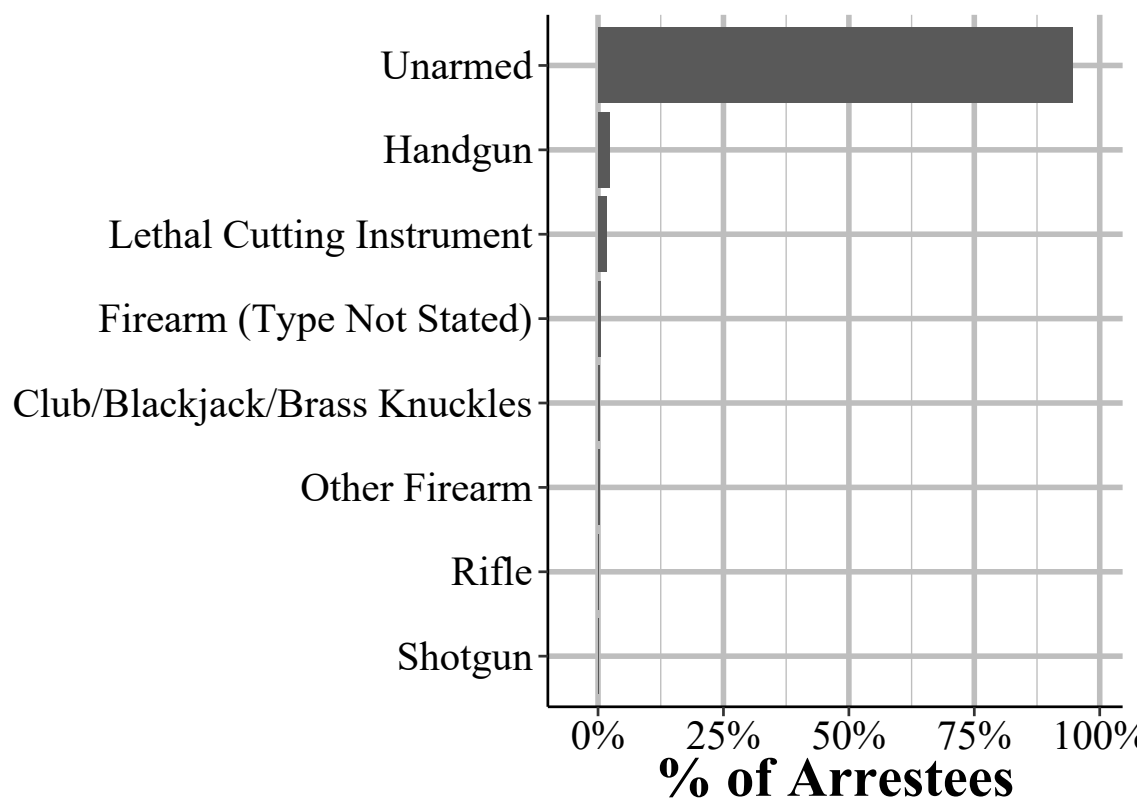
Figure 7.1: The bias motivation (i.e. if hate crime or not) for all arrestees reported in 2019.

7.1.5 Automatic weapons

This variable only tells you if the weapon is automatic
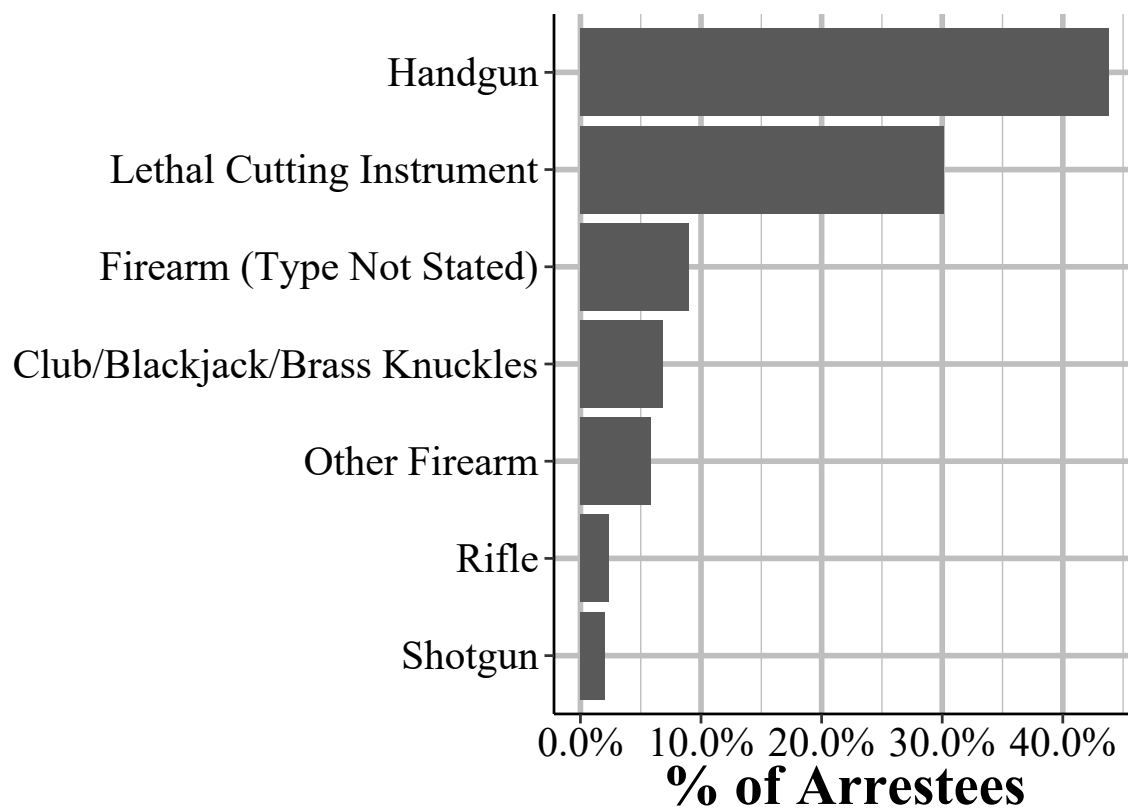
Figure 7.2: The bias motivation (i.e. if hate crime or not) for all arrestees reported in 2019.
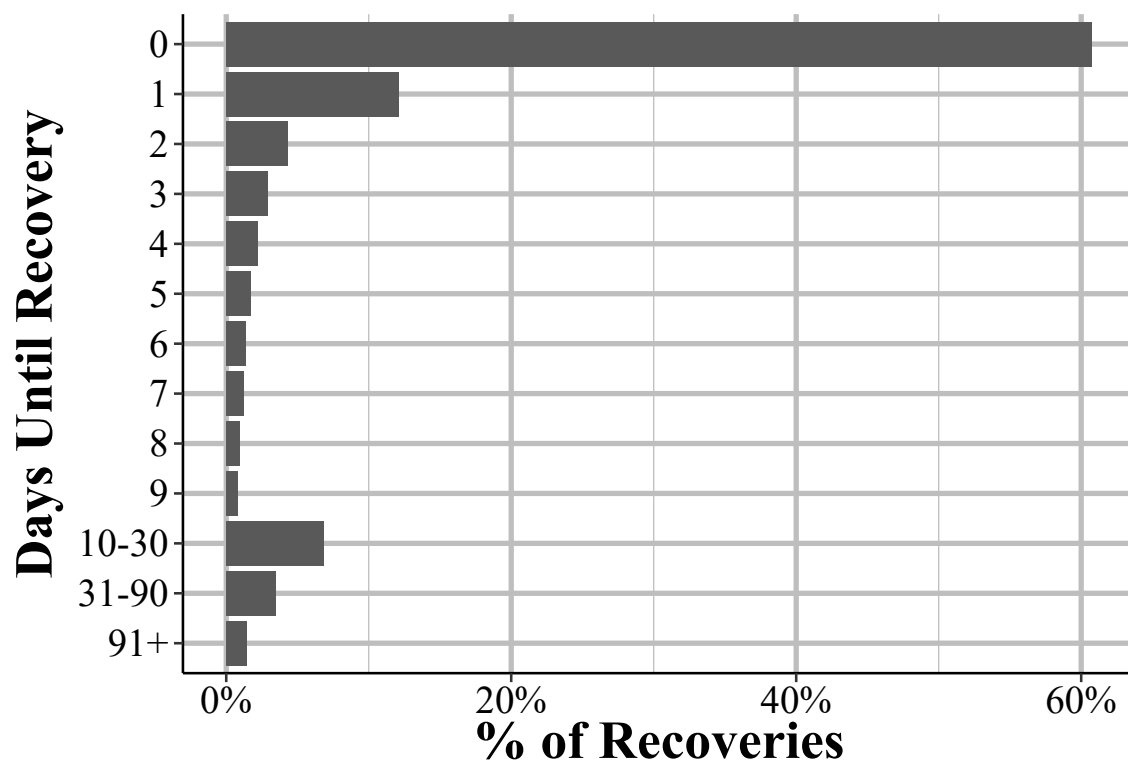
7.1.6 Type of arrest
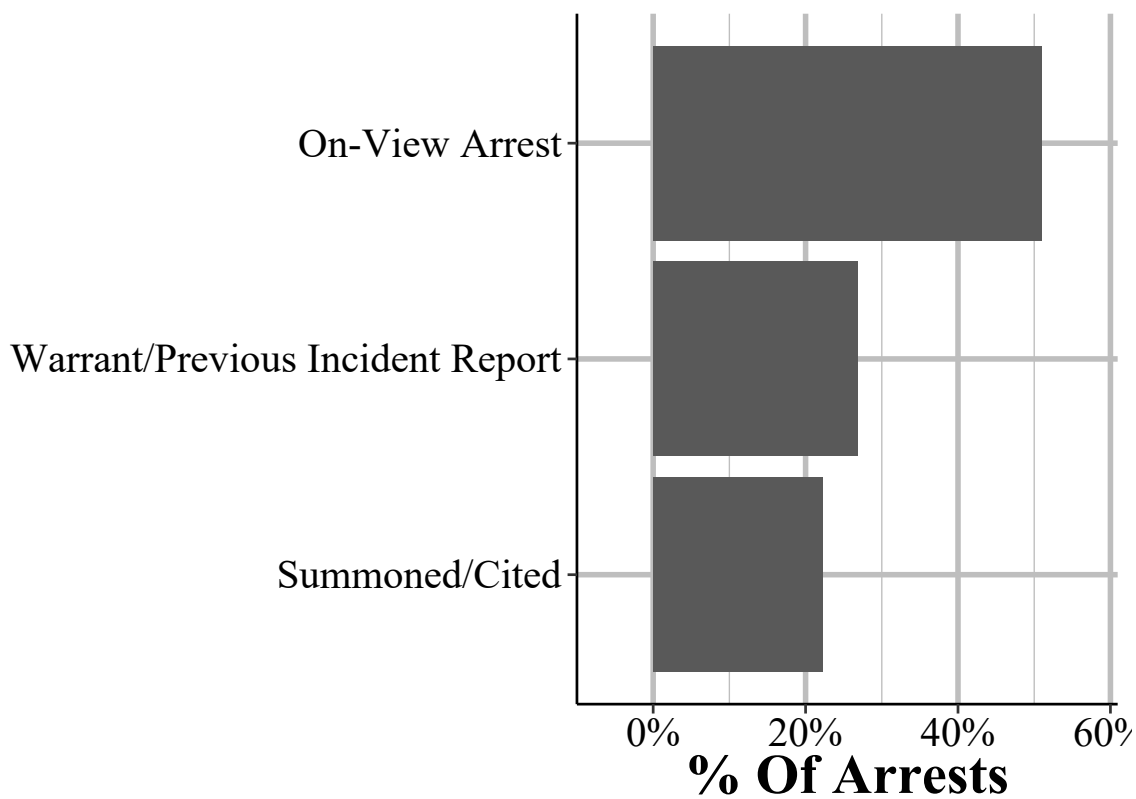
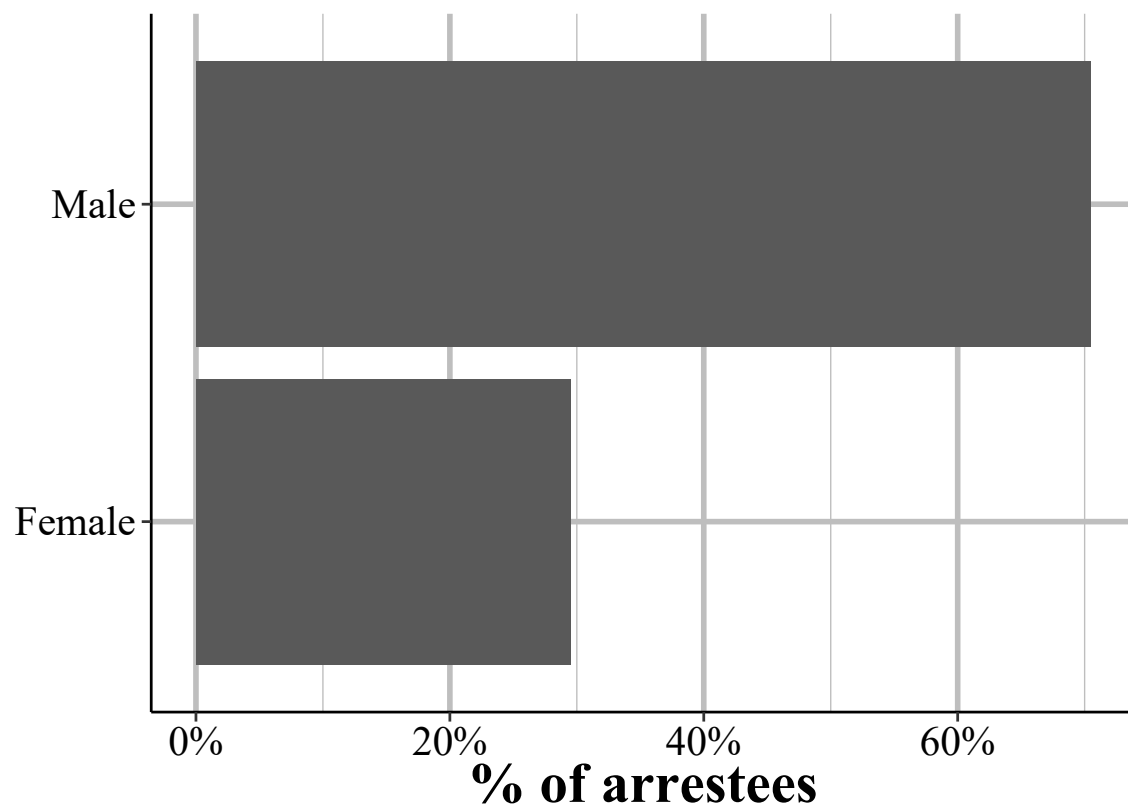### 7.1.7 Disposition for juvenile arrestees
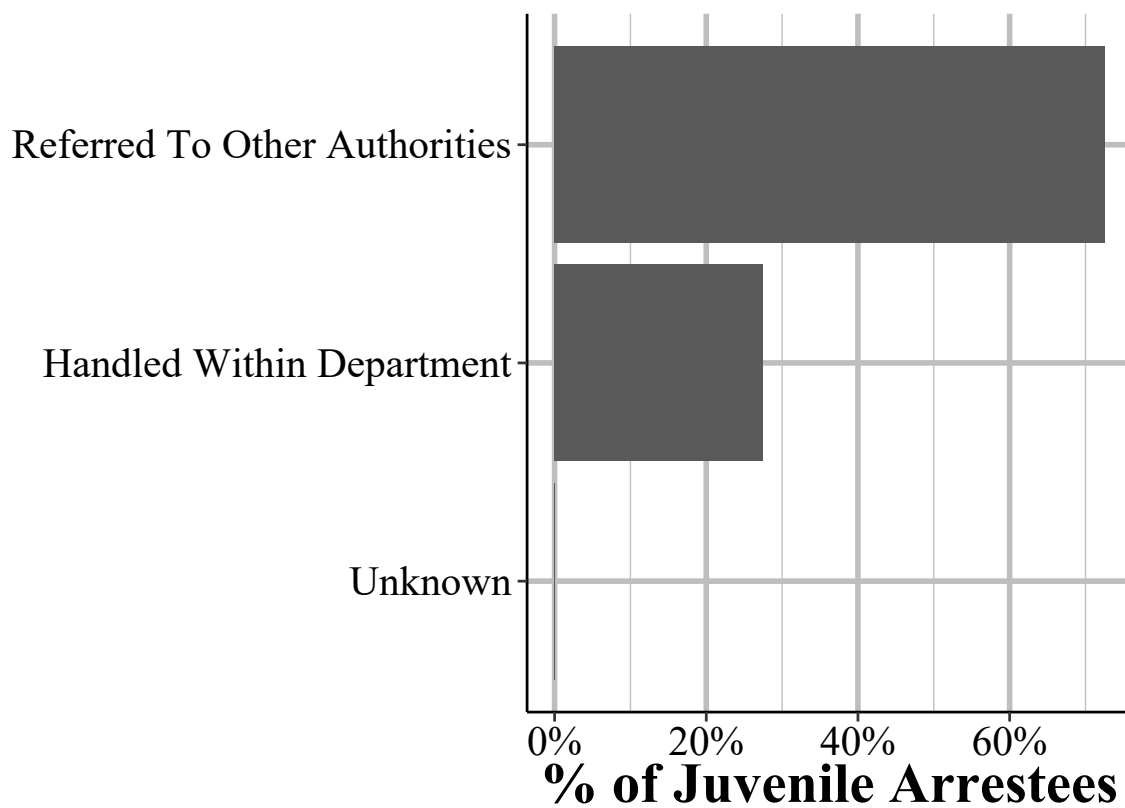
7.1.8 Residence status
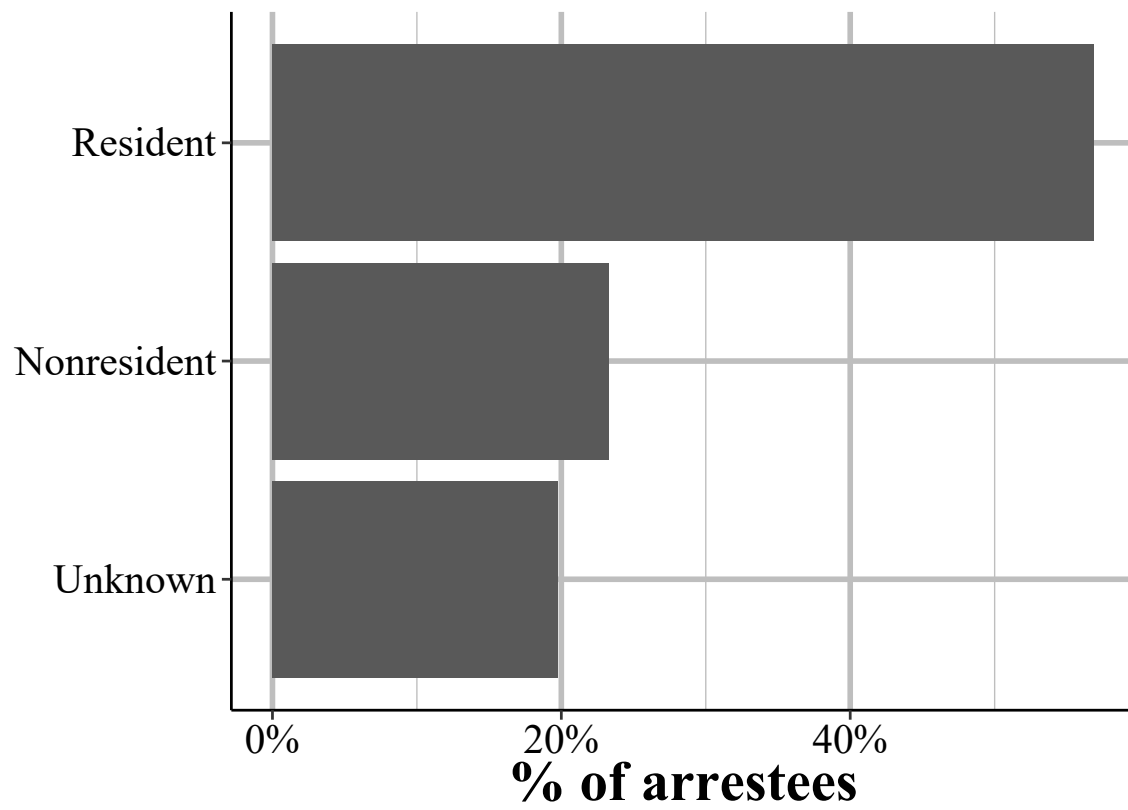
7.1.9 Demographics

7.1.9.1   Age
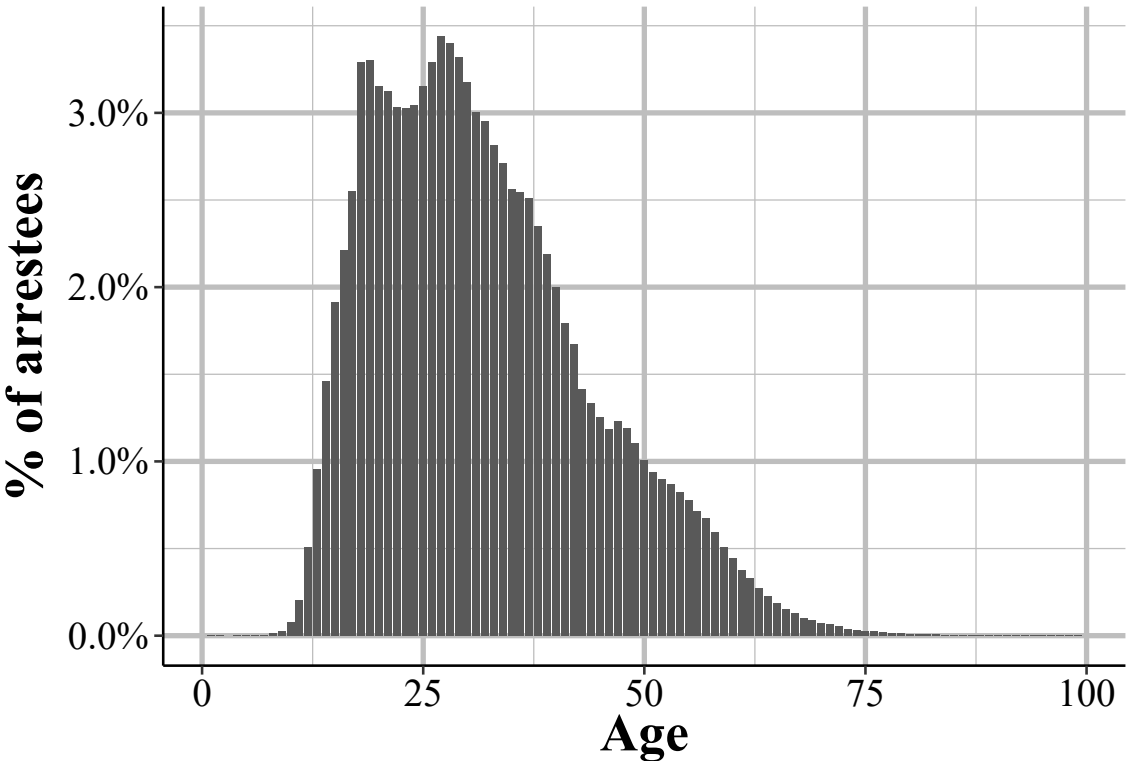
7.1.9.2   Sex

7.1.9.3 Race
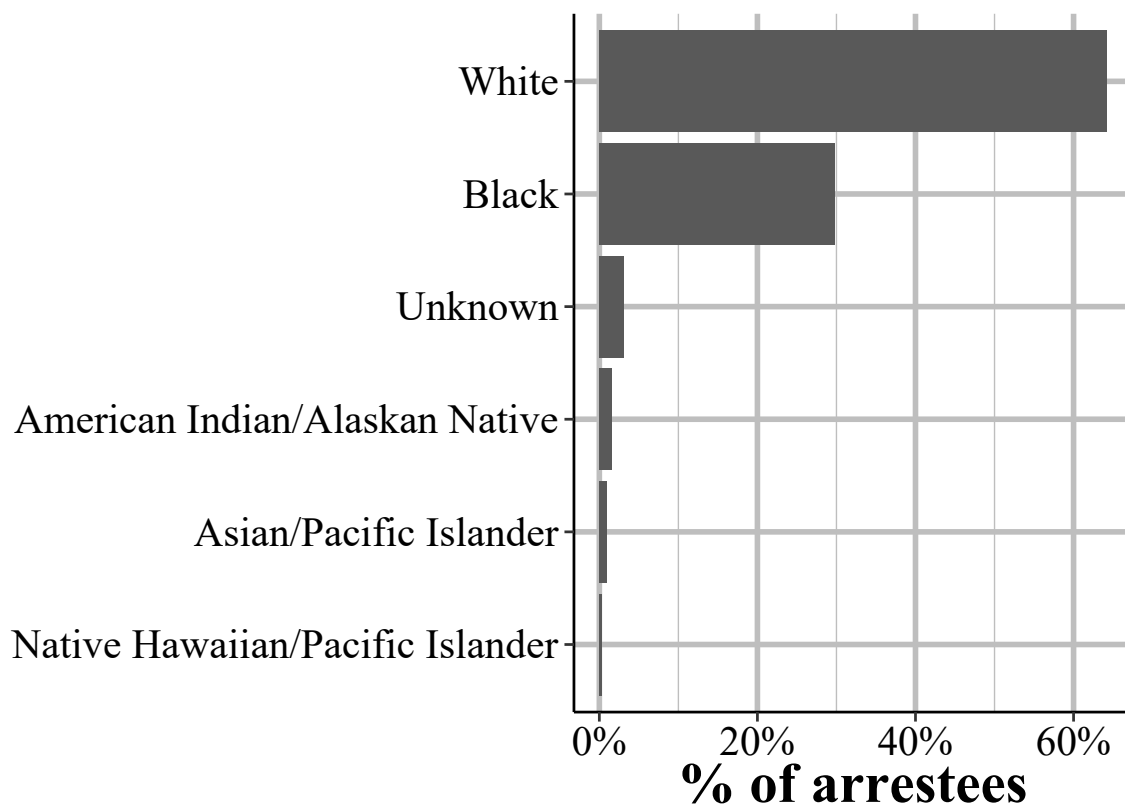
7.1.9.4 Ethnicity
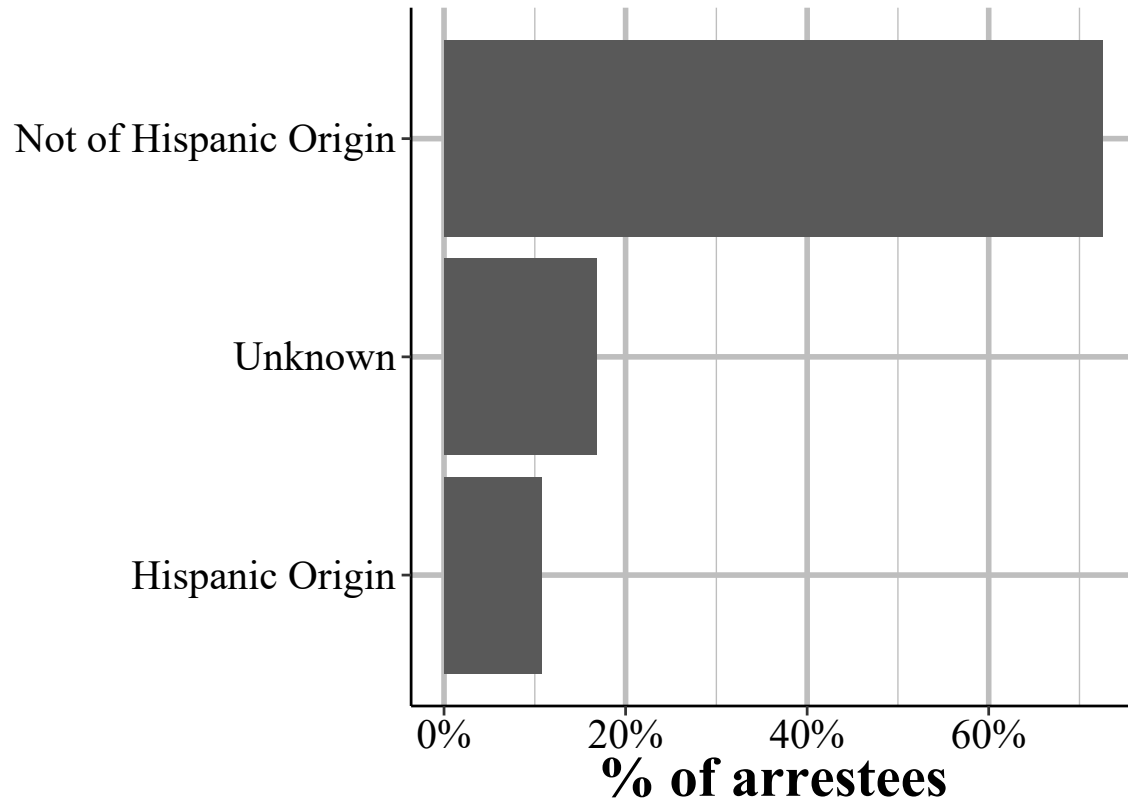
Chapter 8

Property and Window Property Segment

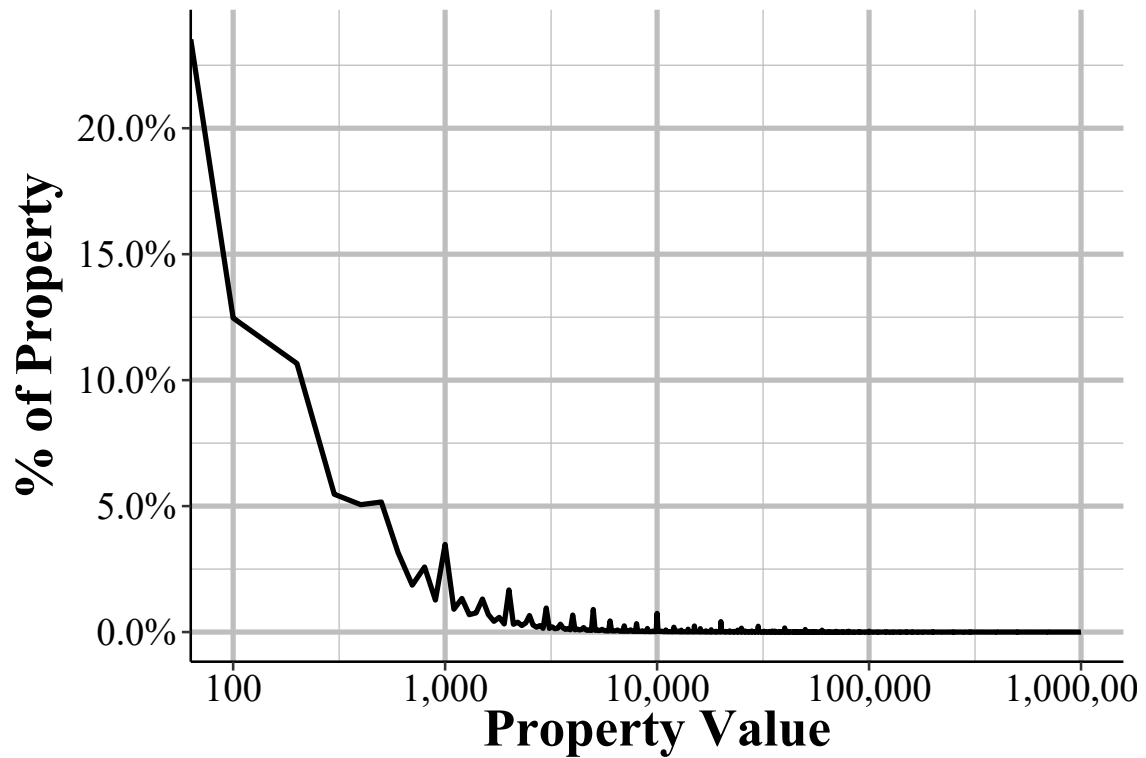## 8.1 Important variables

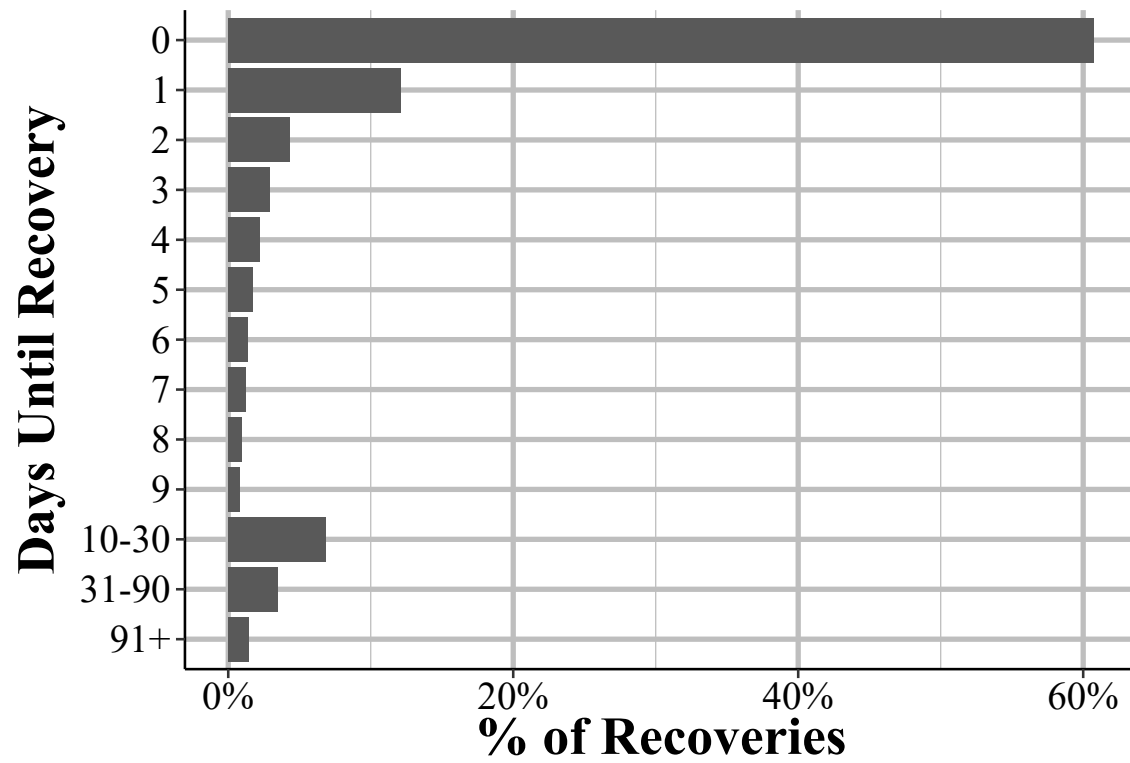### 8.1.1 Description of property
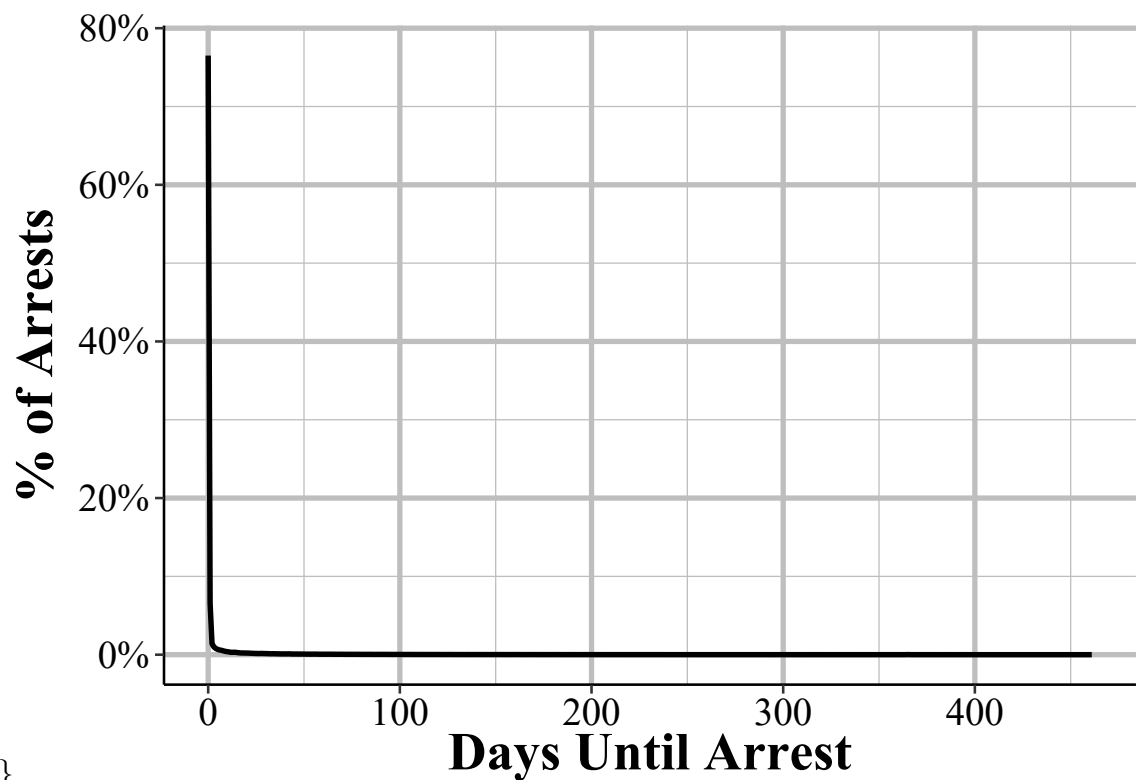
\begin{center}

8.1.2 Type of property loss
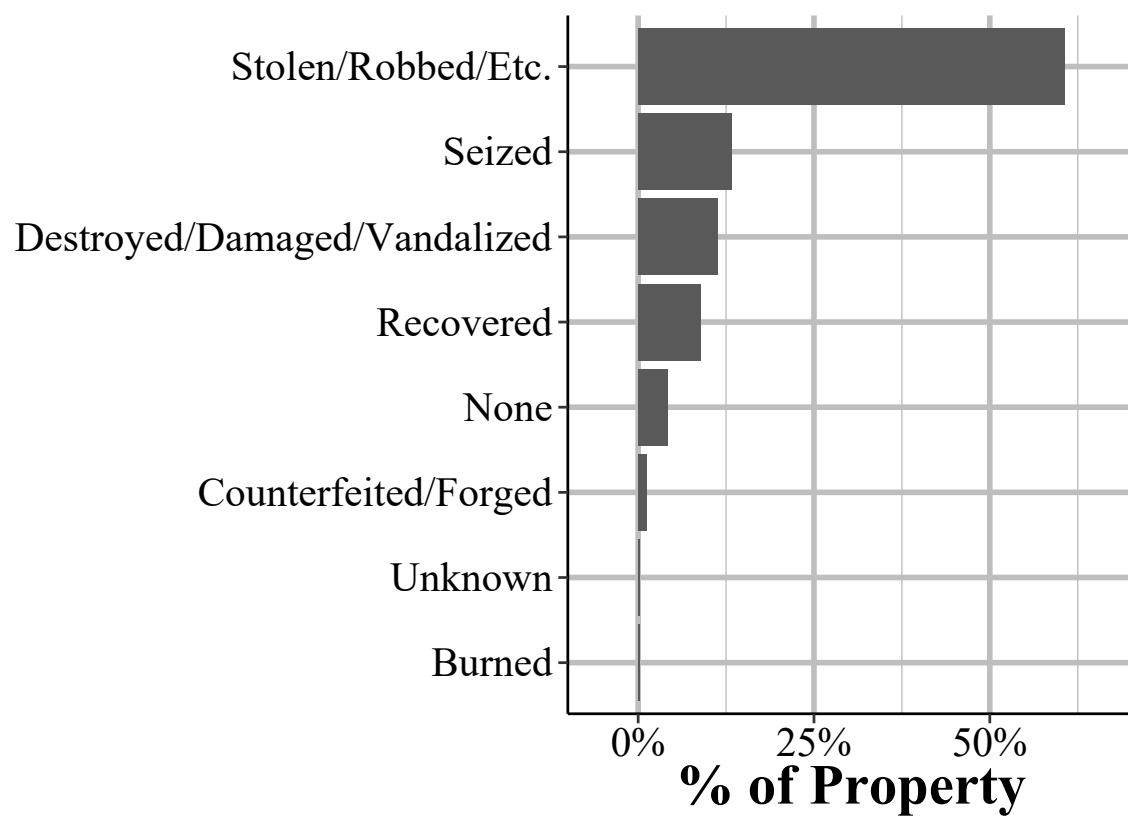
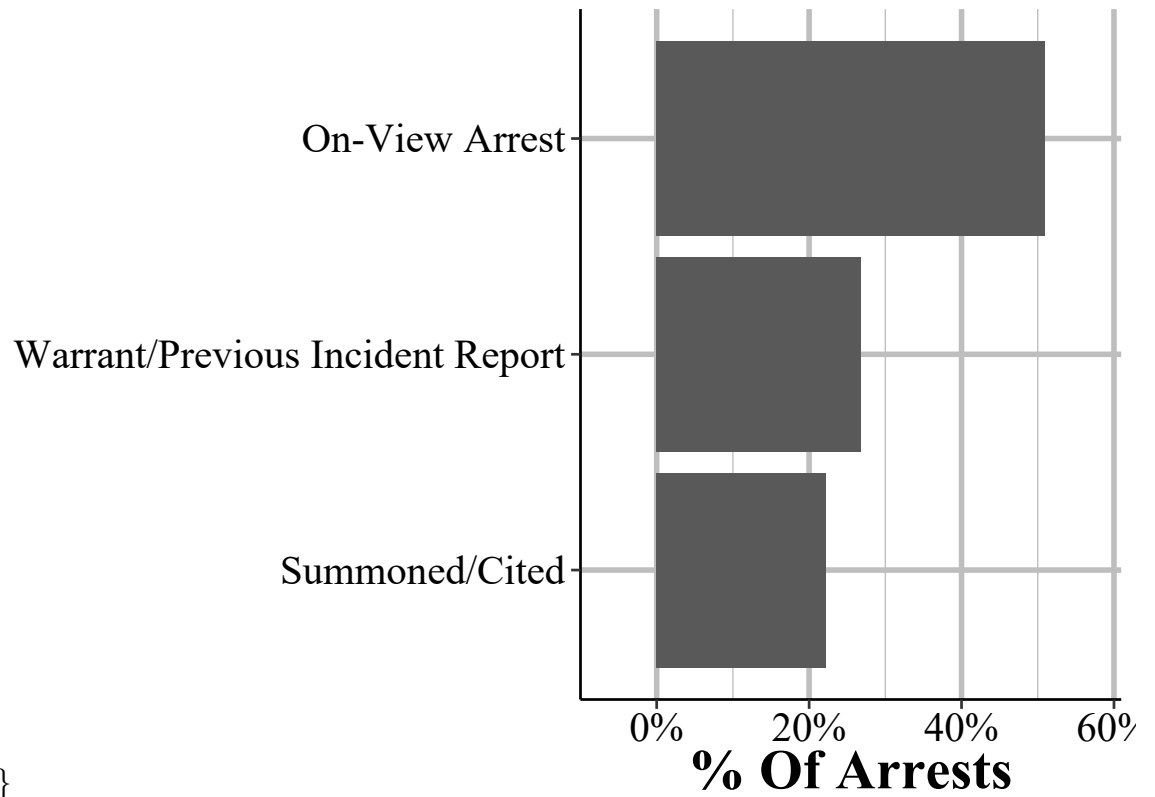### 8.1.3   Value of stolen property
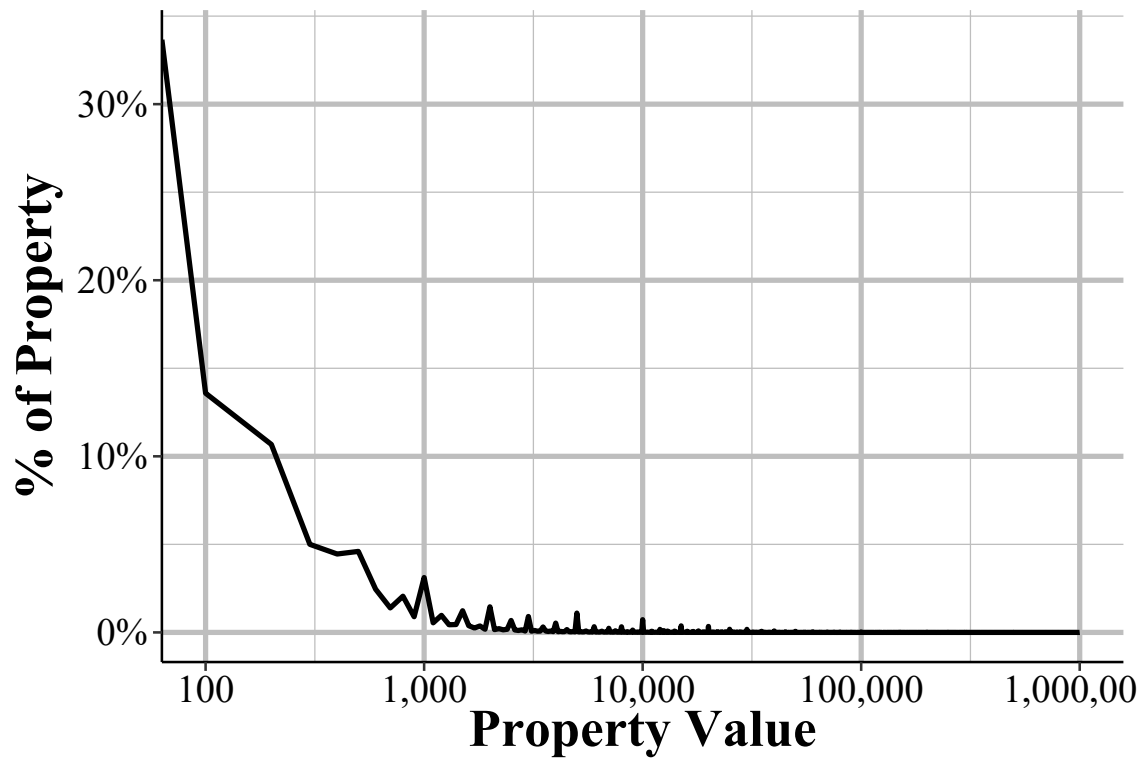
8.1.4   Date property was recovered

8.1.5   Drugs

8.1.5.1   Suspected drug type

\begin{center}

8.1.5.2   Amount of drugs

- Dosage Unit/Items (Pills, Etc.)
- Fluid Ounce
- Gallon
- Gram
- Kilogram
- Liter
- Milliliter
- Not Reported
- Number of Plants
- Ounce
- Pound