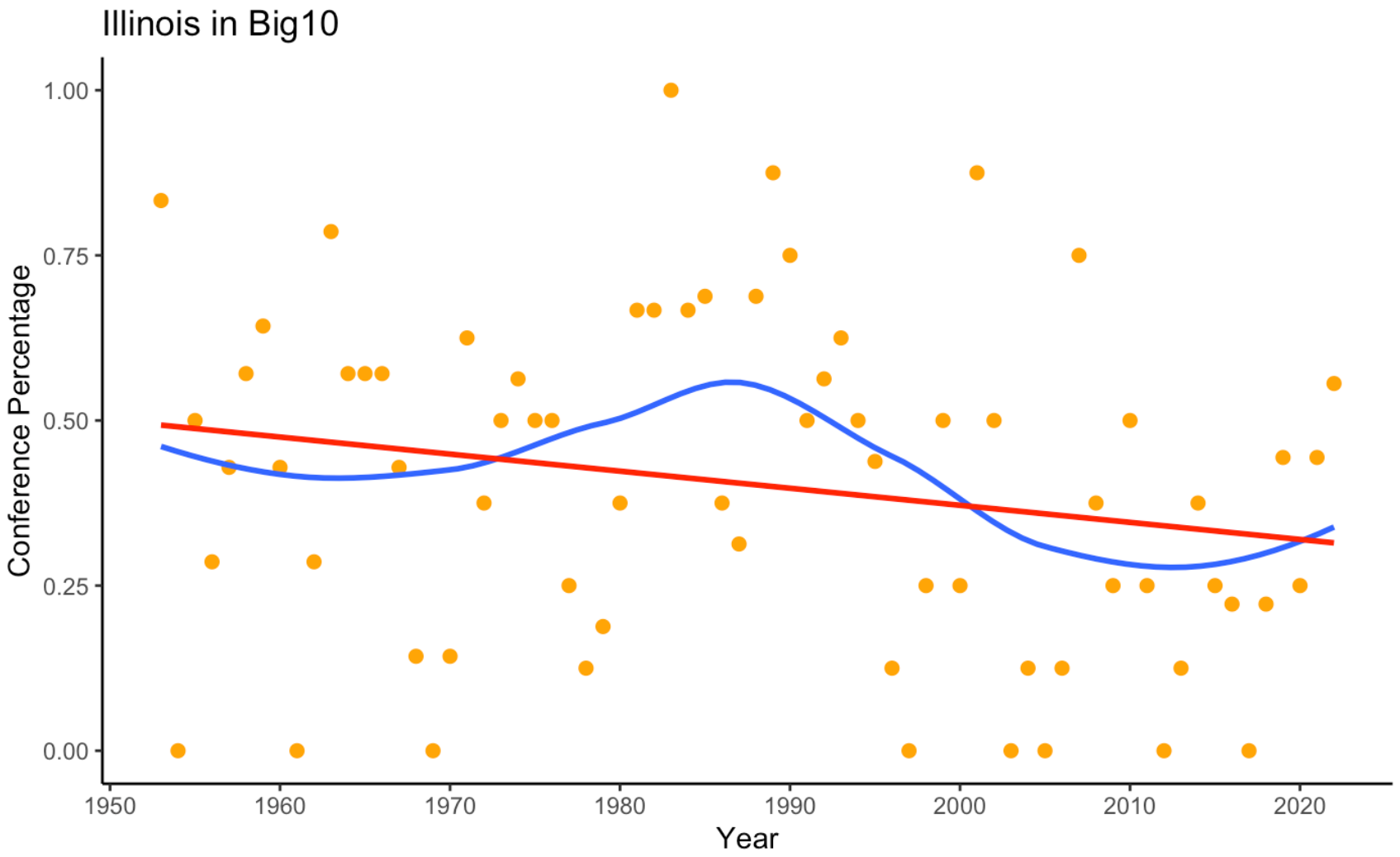
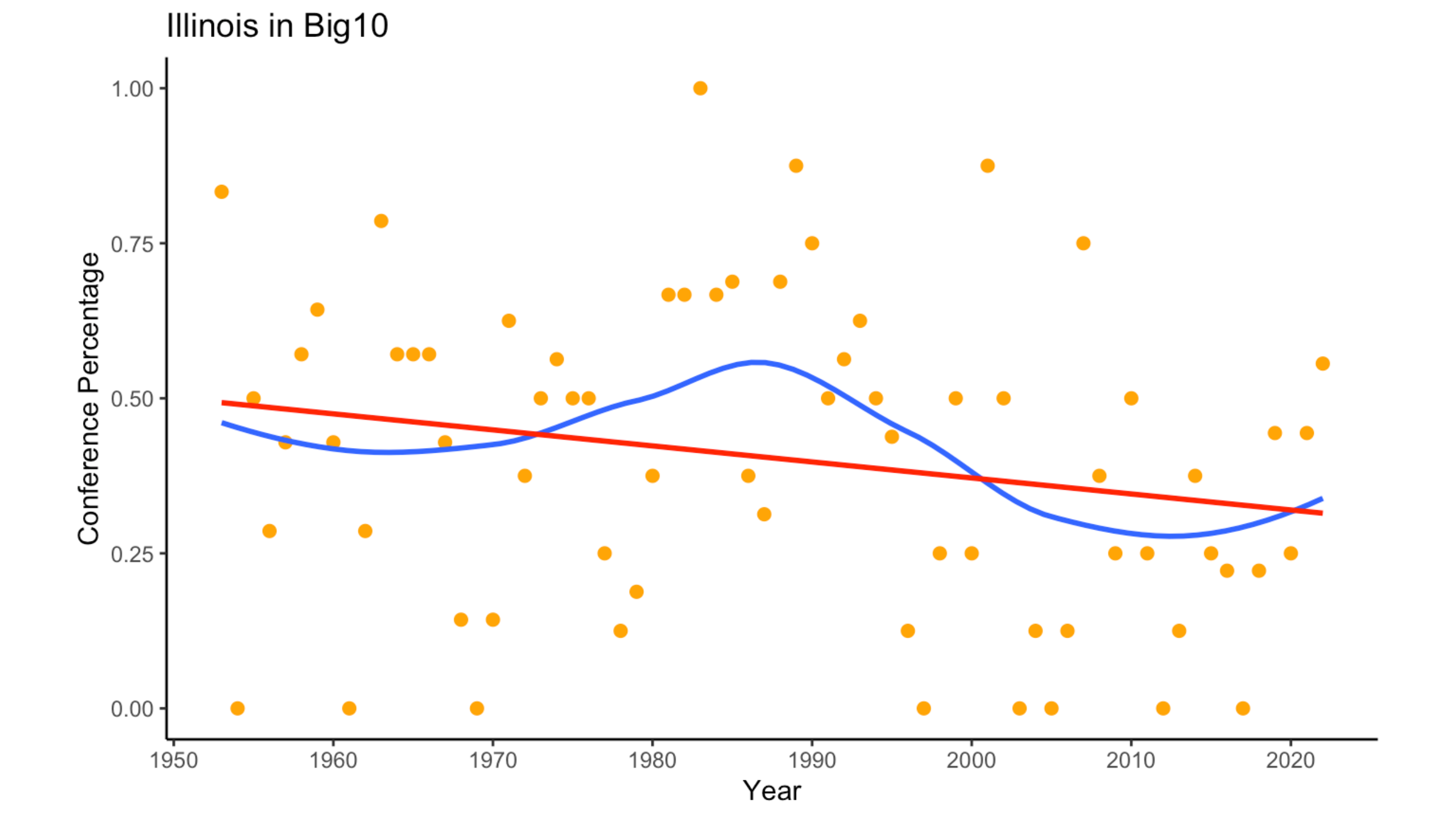
Tips to Improve Your **Data Visualizations**

Christopher Kinson, PhD

What jumps out and what do you notice first?

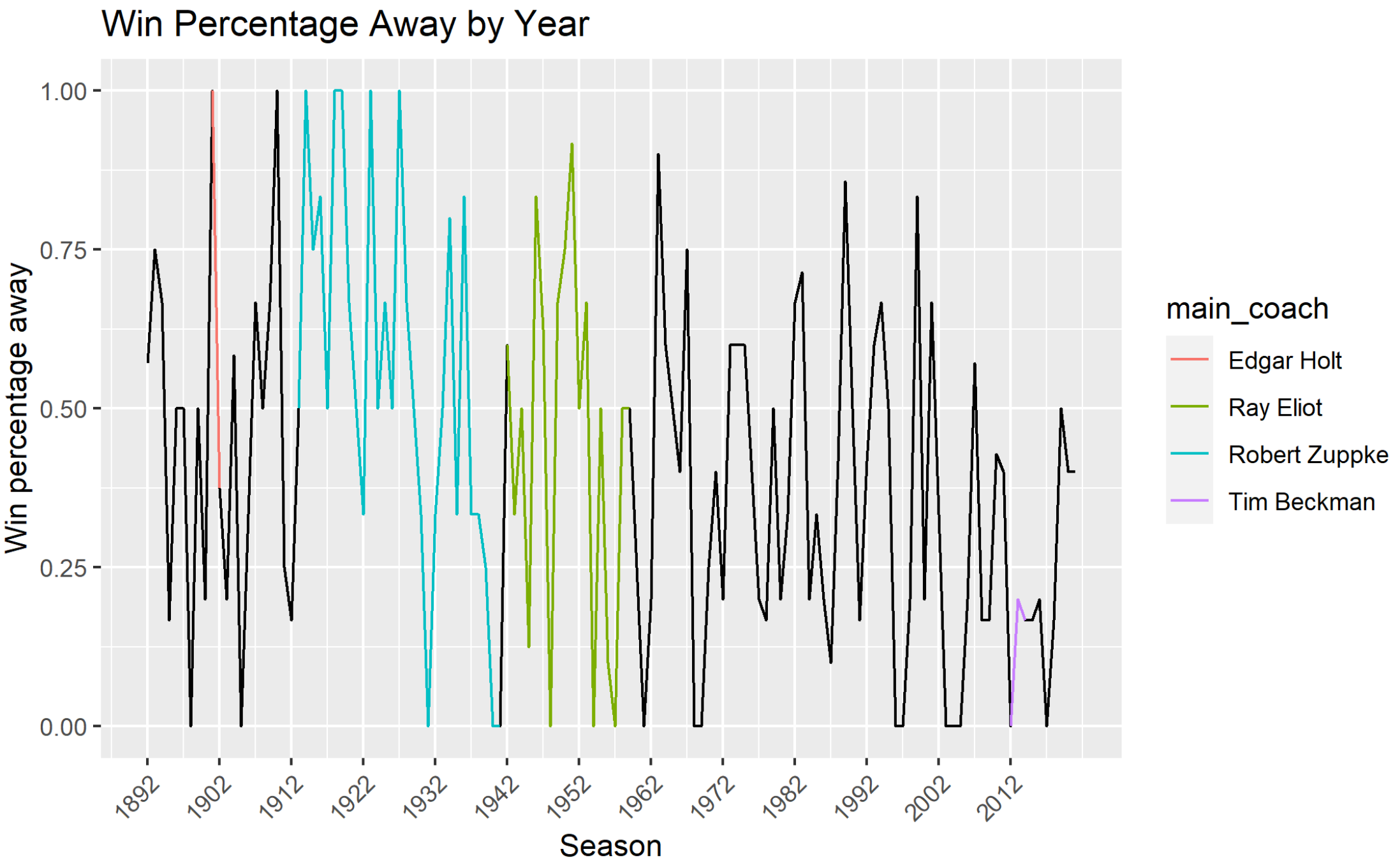


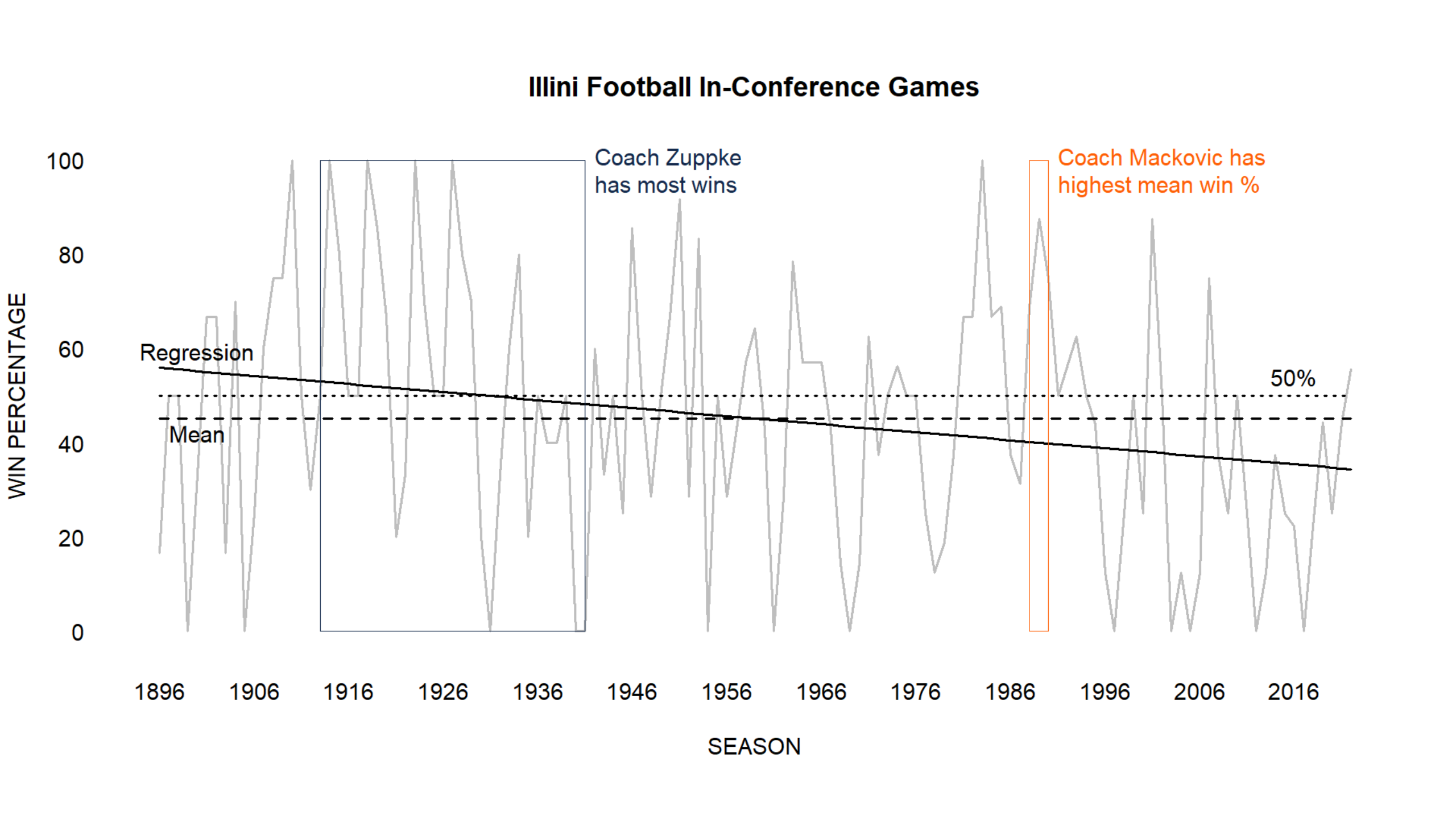
Tip: Use lines for displaying time information.


Illinois in Big10
Conference Percentage
1.00
0.75
0.50
0.25
0.00
1950
1960
1970
1980
1990
2000
2010
2020
Year

1.000
.804
.713
.667
.659
.649
.625
.612
.571
.563
.556
.533
.530
.520
.477
.448
.447
.417
.400
.380
.324
.304
.300
.227
.200
.000
Pct
Arthur Hall
Bill Cubit
Bob Blackman
Bret Bielema
E.K. Hall
Edgar Holt
Fred Lowenthal
Fred Smith
Gary Moeller
George Huff
George Washington Woodruff
James Valek
John Mackovic
Justa Lindgren
Lou Tepper
Louis Vail
Lovie Smith
Mike White
Pete Elliott
Ray Eliot
Robert Zuppke
Rod Smith
Ron Turner
Ron Zook
Tim Beckman
Vic Koenning
Coach
fill
percentage

Tip: Leverage data variables when they are related to time.


Win Percentage Away by Year
Win percentage away
1.00
0.75
0.50
0.25
0.00
1892
1902
1912
1922
1932
1942
1952
1962
1972
1982
1992
2002
2012
Season
main_coach
Edgar Holt
Ray Eliot
Robert Zuppke
Tim Beckman


Illini Football In-Conference Games
Coach Zuppke has most wins
Coach Mackovic has highest mean win %
Regression
Mean
50%
WIN PERCENTAGE
100
80
60
40
20
0
1896
1906
1916
1926
1936
1946
1956
1966
1976
1986
1996
2006
2016
SEASON

Base R

• Create basic line plot with no axes

`plot(…,type="l",axes=FALSE)`

• To add custom axis

`axis()`

• To add lines beyond base plot:

`lines() ; abline()`

• To add enclosures

`rect()`

• To add text

`text()`

Tidyverse

• Create basic line plot with no axes

`geom_line(mapping=aes())`

• To add custom axis

`scale_x_* ; scale_y_*`

• To add lines beyond base plot

`geom_line() ; geom_smooth()`

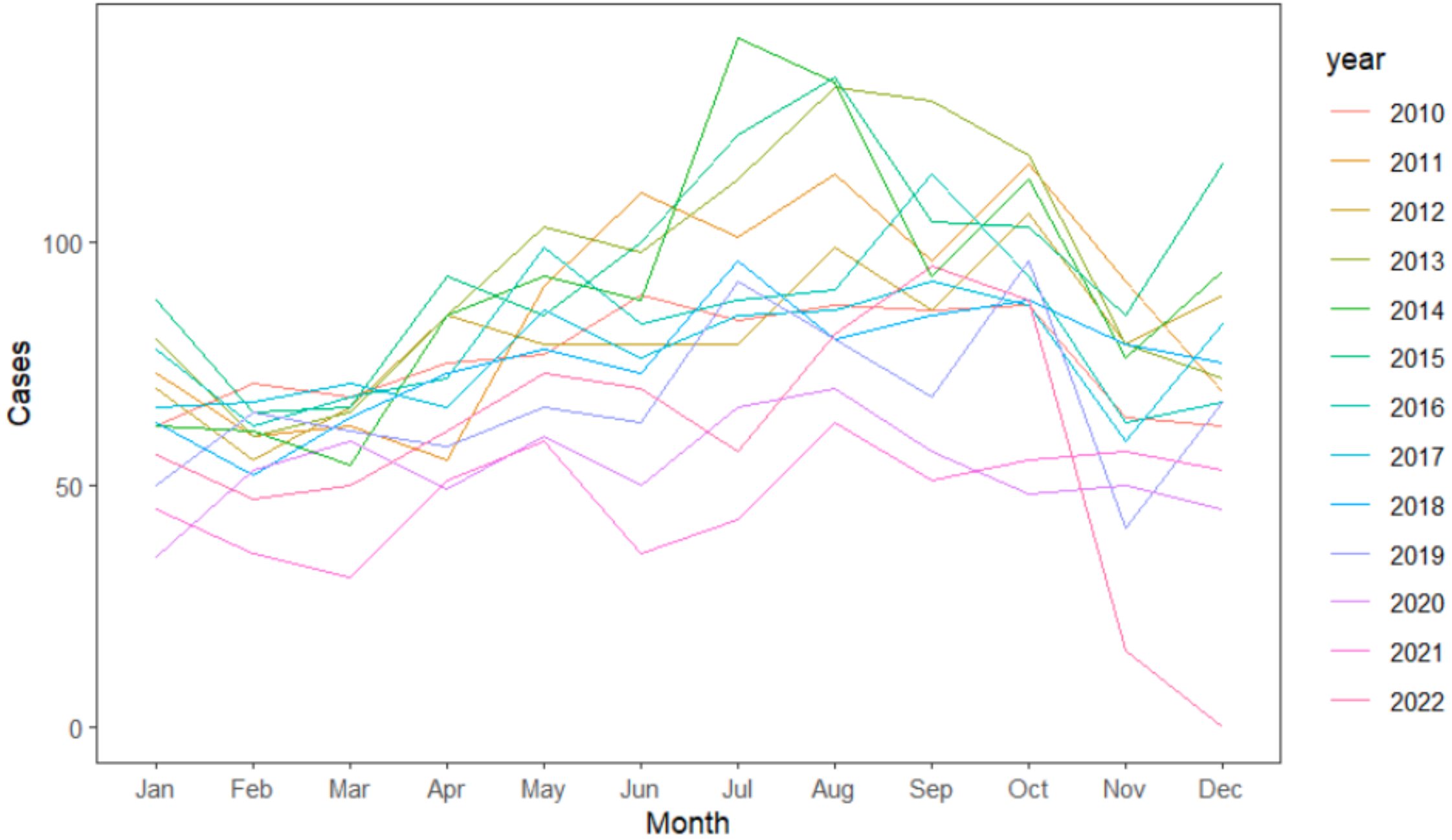
• To add enclosures

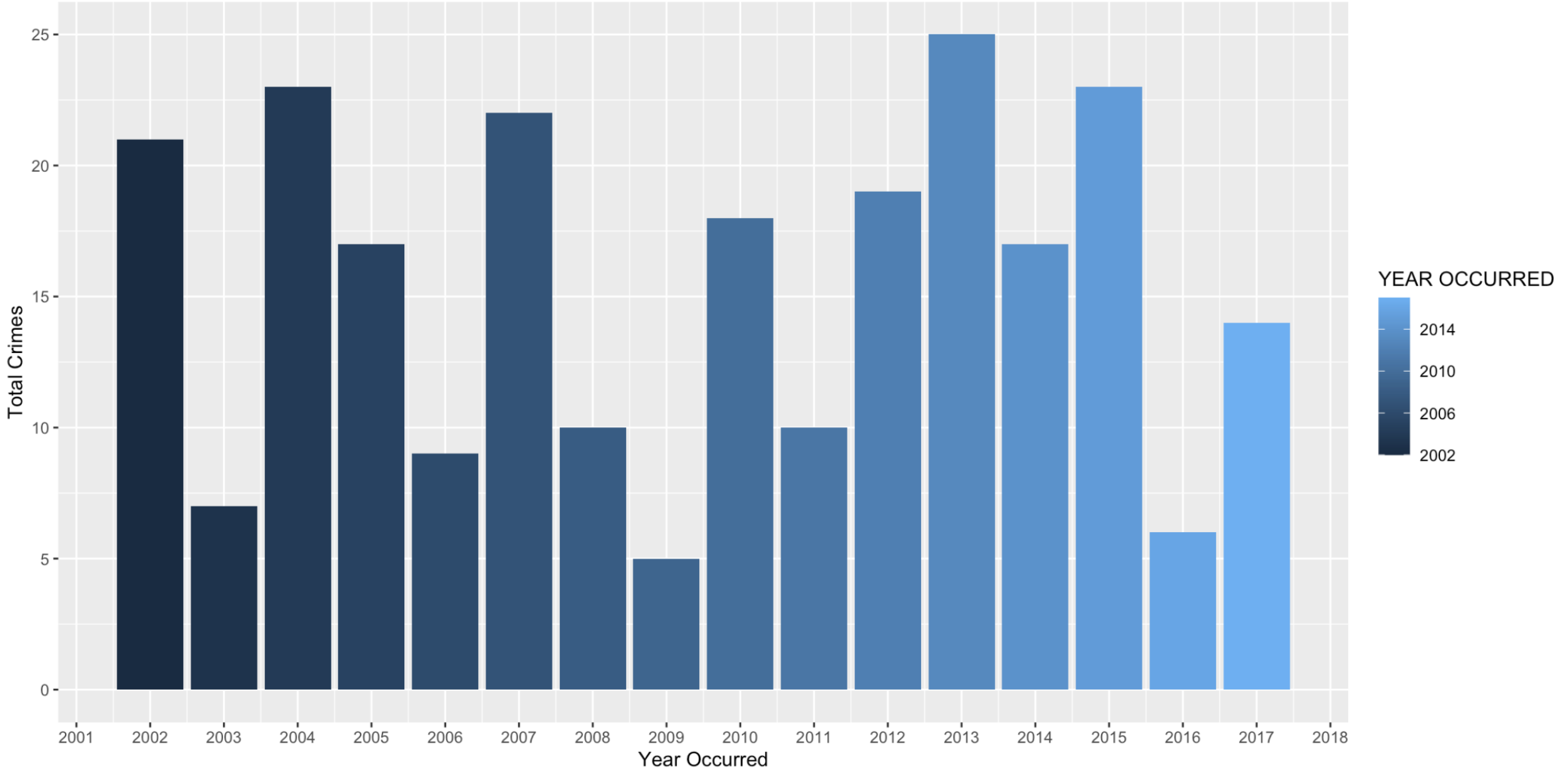
`annotate("rect")`

• To add text

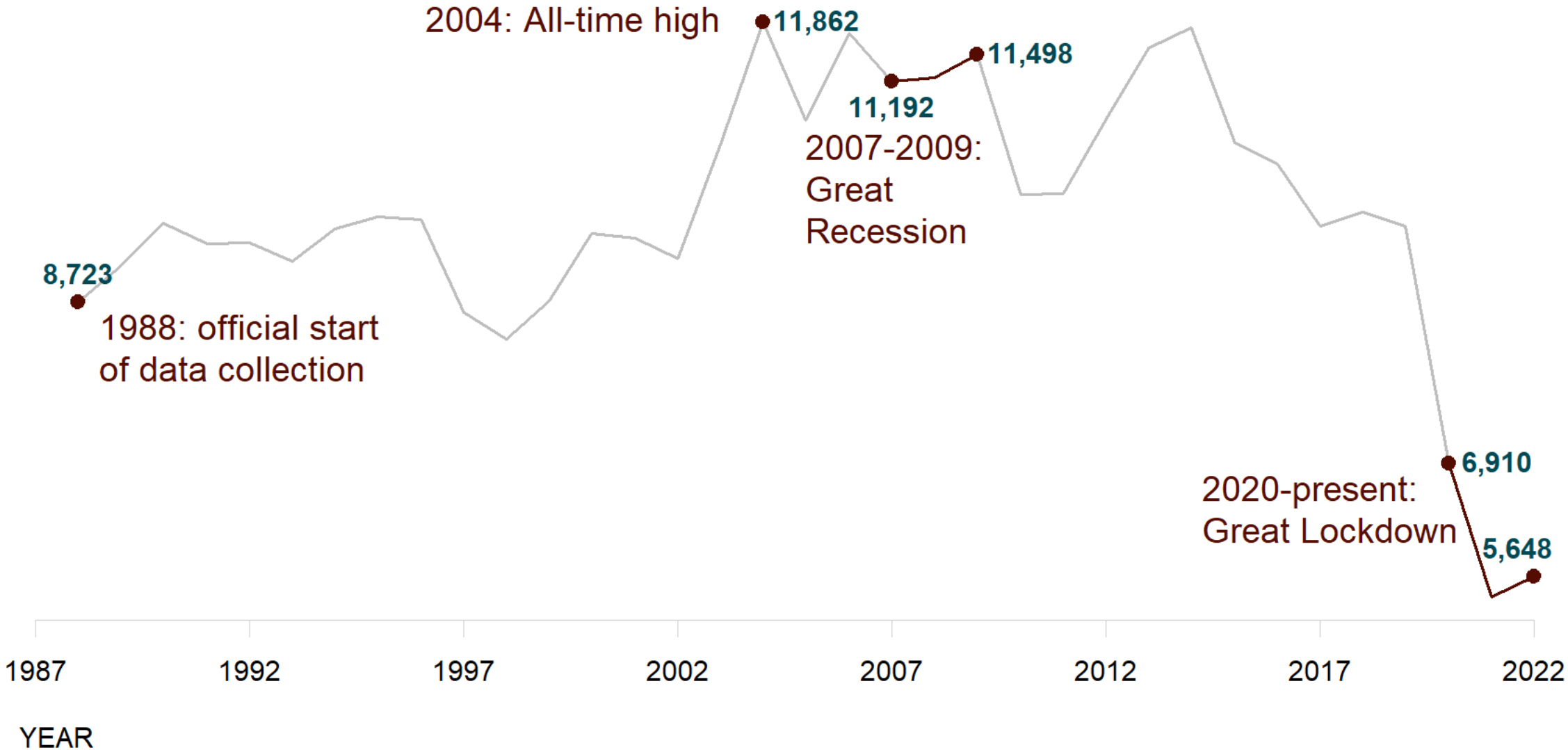
`annotate("text")`

Tip: Use color for emphasis.


Cases
100
50
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Month
year
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022


25
20
15
10
5
0
Total Crimes
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
Year Occurred
YEAR OCCURRED
2014
2010
2006
2002

Number of Incidents Involving the Urbana Police



Base R

- To begin plot with blank canvas

```
plot(…,axes=FALSE,type="n",xlab="",ylab="",main="")
```

- To add RGB-specific color to main plotting characters

```
plot(…,col=rgb())
```

- To add custom horizontal axis label in left-most position

```
title(xlab="",adj=0)
```

Tidyverse

- To begin plot with blank canvas

```
theme(panel.background=
element_blank(),
legend.position="none")
```

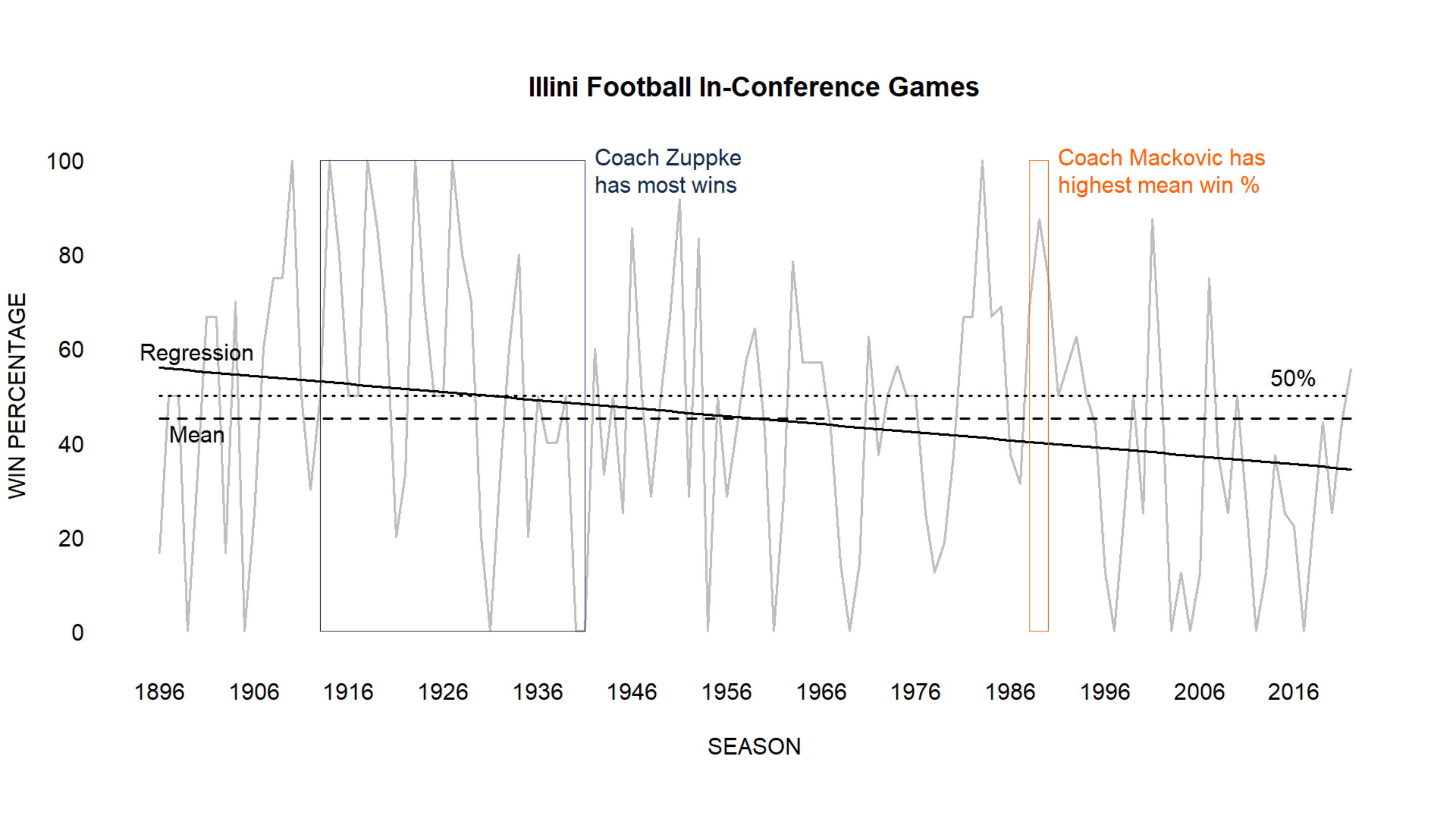
- To add RGB-specific color to main plotting characters

```
ggplot(…,color=rgb())
```

- To add custom horizontal axis label in left-most position

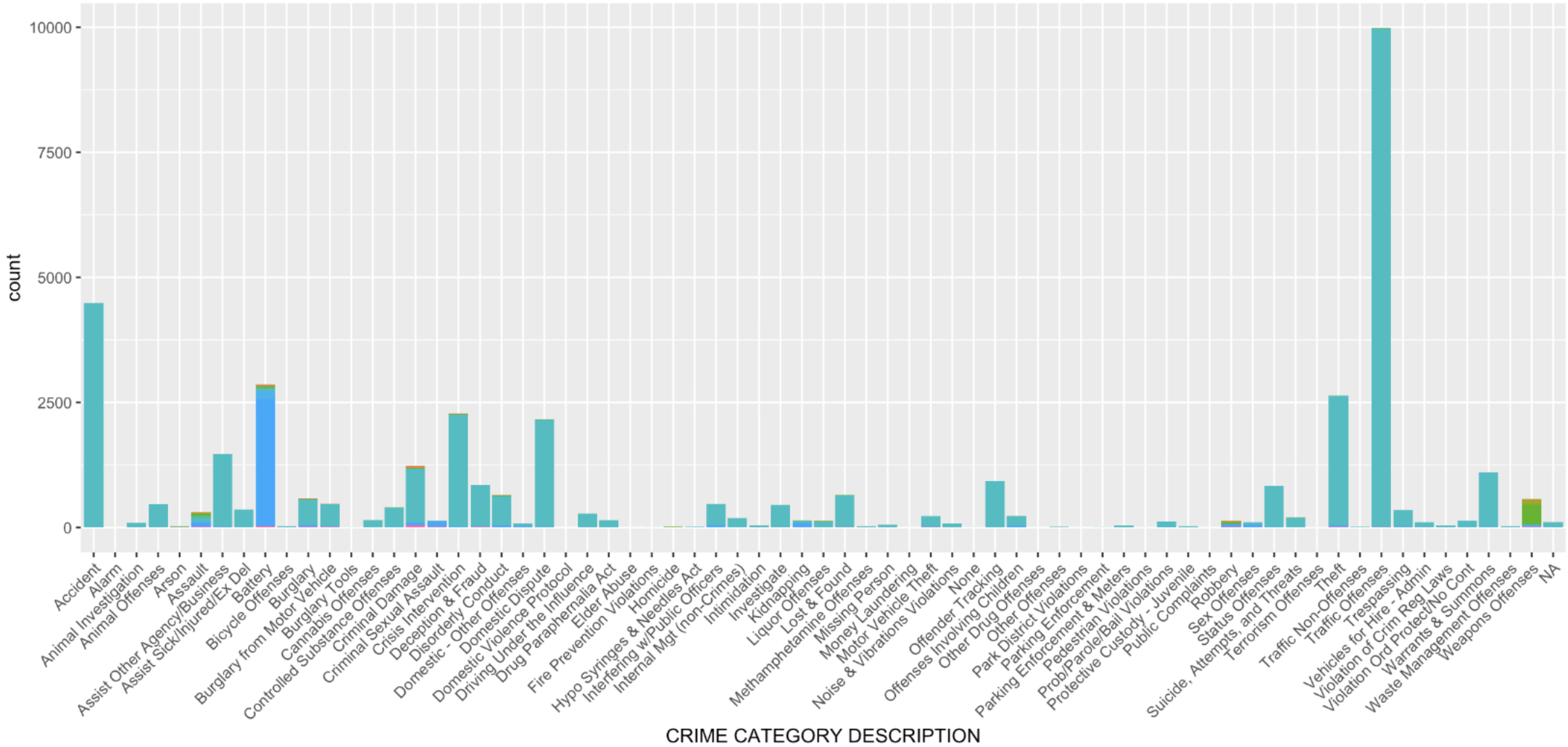
```
theme(axis.title.x=
element_text(margin=
margin(l=))
```

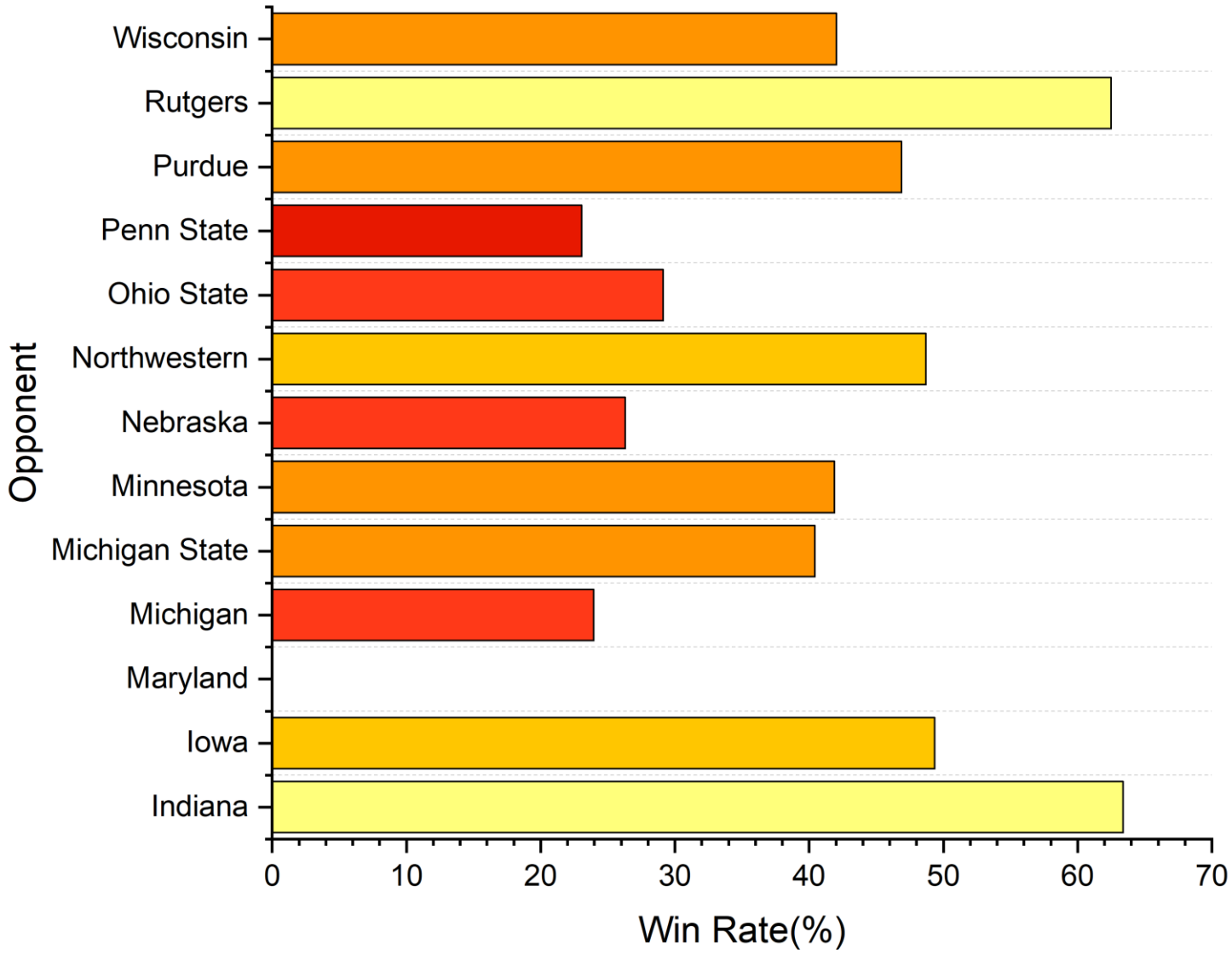
Tips: Check color wheels, such as www.canva.com/colors/color-wheel/, for complementary colors. When color is in the branding, use it.


Illini Football In-Conference Games
Coach Zuppke has most wins
Coach Mackovic has highest mean win %
Regression
Mean
50%
WIN PERCENTAGE
100
80
60
40
20
0
1896
1906
1916
1926
1936
1946
1956
1966
1976
1986
1996
2006
2016
SEASON

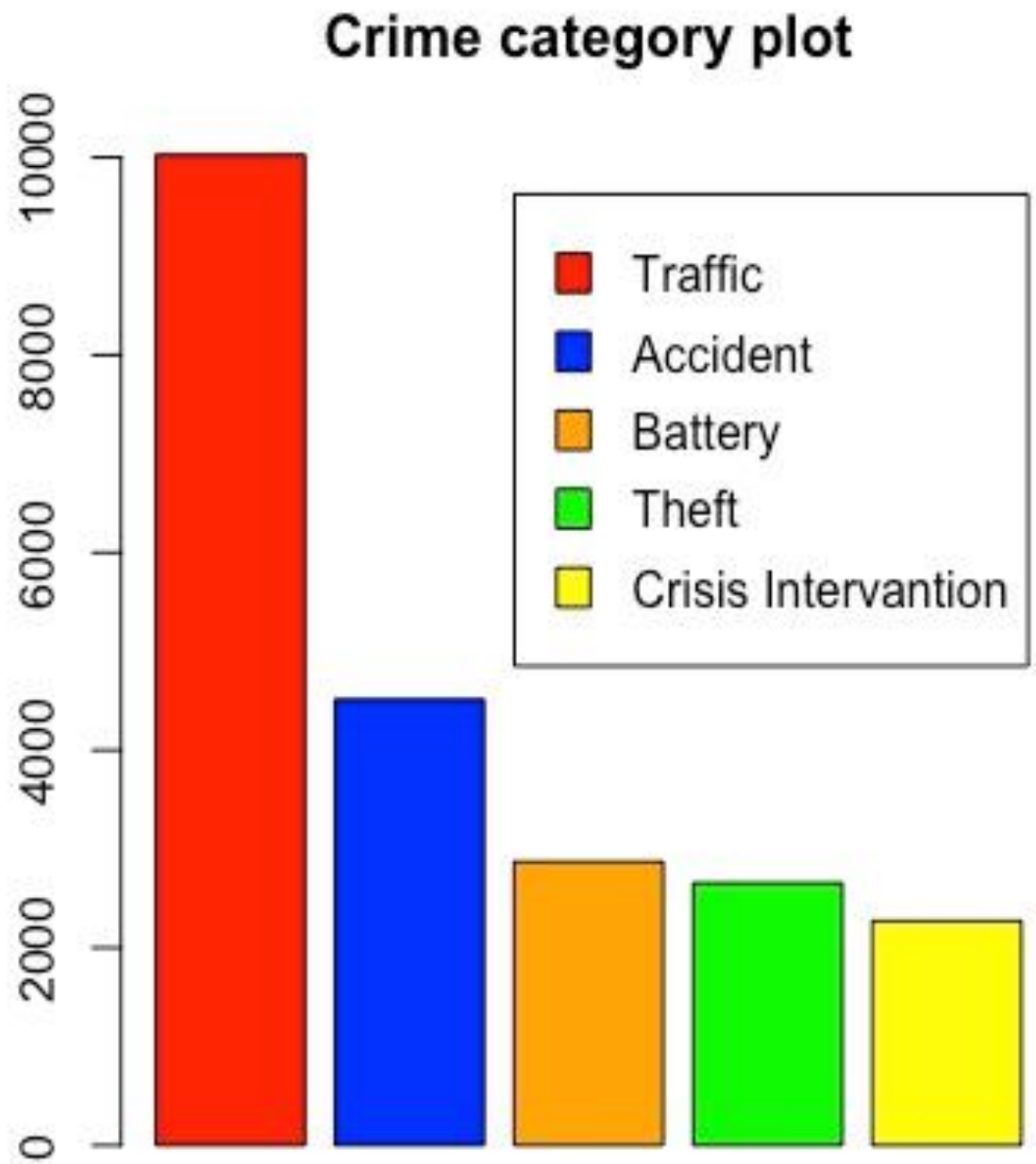
https://marketing.illinois.edu/visual-identity/color

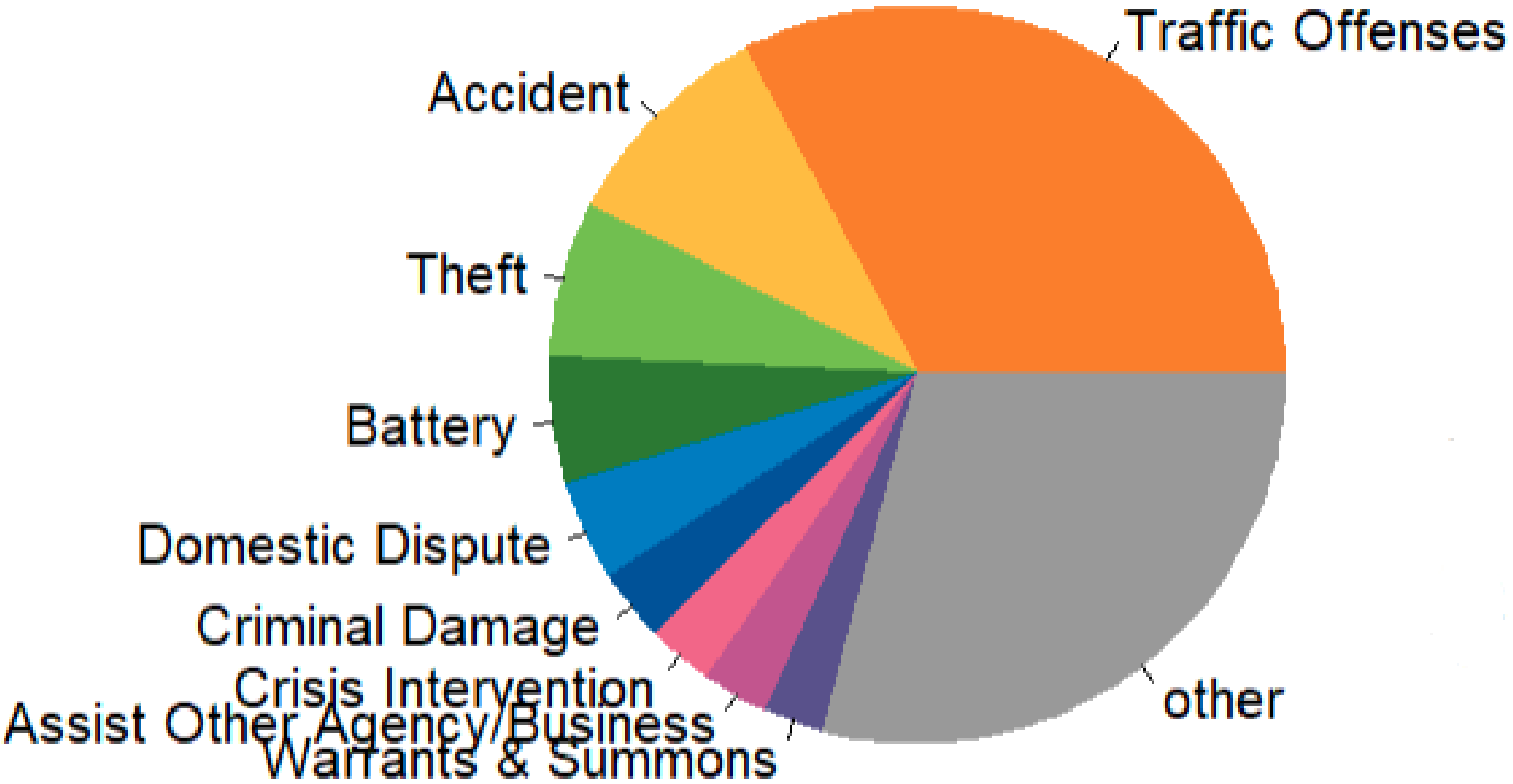
Tips: Don’t overload the audience. Your data viz should be self-sufficient with a clear context.


count
10000
7500
5000
2500
0
Accident
Alarm
Animal Investigation
Animal Offenses
Arson
Assault
Assist Other Agency/Business
Assist Sick/Injured/Ex Del
Bicycle Offenses
Battery
Burglary
Burglary from Motor Vehicle
Burglary Tools
Cannabis Offenses
Controlled Substance Offenses
Criminal Damage
Criminal Sexual Assault
Crisis Intervention
Deception & Fraud
Disorderly Conduct
Domestic - Other Offenses
Domestic Dispute
Domestic Violence Protocol
Driving Under the Influence
Drug Paraphernalia Act
Elder Abuse
Fire Prevention Violations
Homicide
Hypo Syringes & Needles Act
Interfering w/Public Officers
Internal Mgt (non-Crimes)
Intimidation
Investigate
Kidnapping
Liquor Offenses
Lost & Found
Methamphetamine Offenses
Missing Person
Money Laundering
Motor Vehicle Theft
Noise & Vibrations Violations
None
Offender Tracking
Offenses Involving Children
Other Drug Offenses
Other Offenses
Park District Violations
Parking Enforcement
Parking Enforcement & Meters
Pedestrian Violations
Prob/Parole/Bail Violations
Protective Custody - Juvenile
Public Complaints
Robbery
Sex Offenses
Status Offenses
Suicide, Attempts, and Threats
Terrorism Offenses
Theft
Traffic Non-Offenses
Traffic Offenses
Trespassing
Vehicles for Hire - Admin
Violation of Crim Reg Laws
Violation Ord Protect/No Cont
Warrants & Summons
Waste Management Offenses
Weapons Offenses
NA
CRIME CATEGORY DESCRIPTION

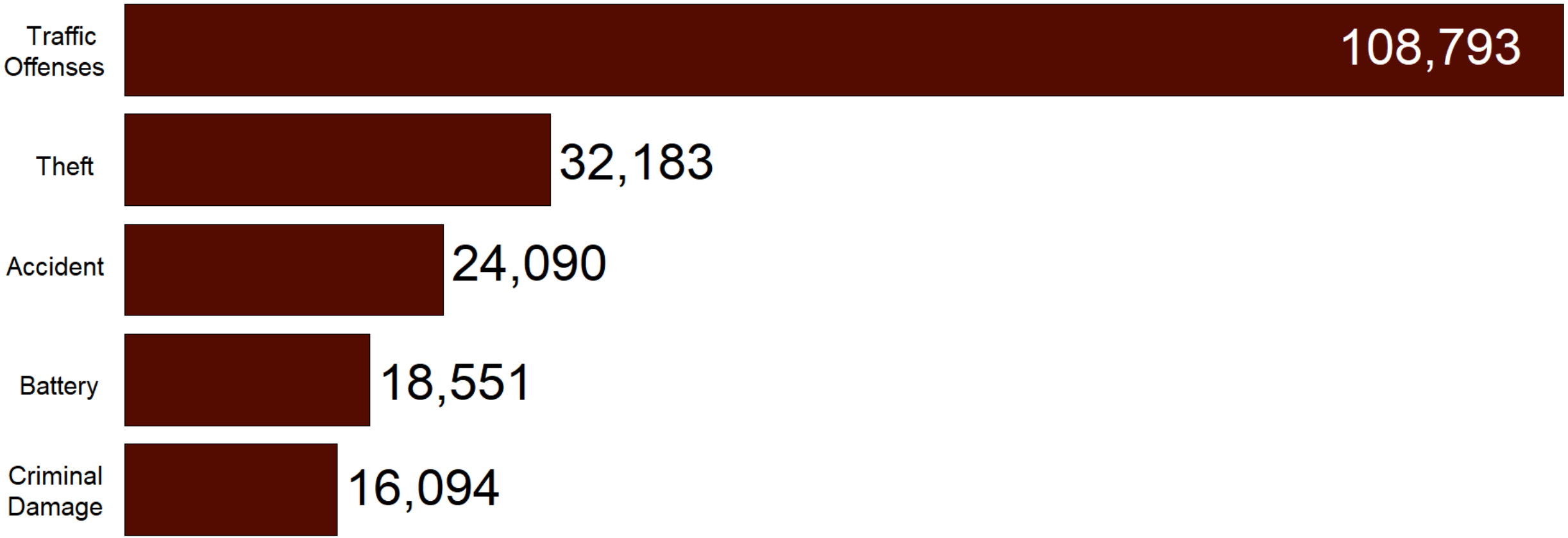

Opponent
Wisconsin
Rutgers
Purdue
Penn State
Ohio State
Northwestern
Nebraska
Minnesota
Michigan State
Michigan
Maryland
Iowa
Indiana
0
10
20
30
40
50
60
70
Win Rate(%)

Tip: Use bars to represent counts.


Crime category plot
0
2000
4000
6000
8000
10000
Traffic
Accident
Battery
Theft
Crisis Intervantion


Traffic Offenses
Accident
Theft
Battery
Domestic Dispute
Criminal Damage
Crisis Intervention
Assist Other Agency/Business
Warrants & Summons
other

Top 5 Most Common Crime Categories for Urbana Police during 1988-2022



Base R

• To begin with bar plot

`barplot()`

• To remove names of bars

`barplot(…,names.arg="")`

• To create horizontal bars

`barplot(…,horiz=TRUE)`

Tidyverse

• To begin with bar plot

`geom_col() ; geom_bar()`

• To remove names of bars

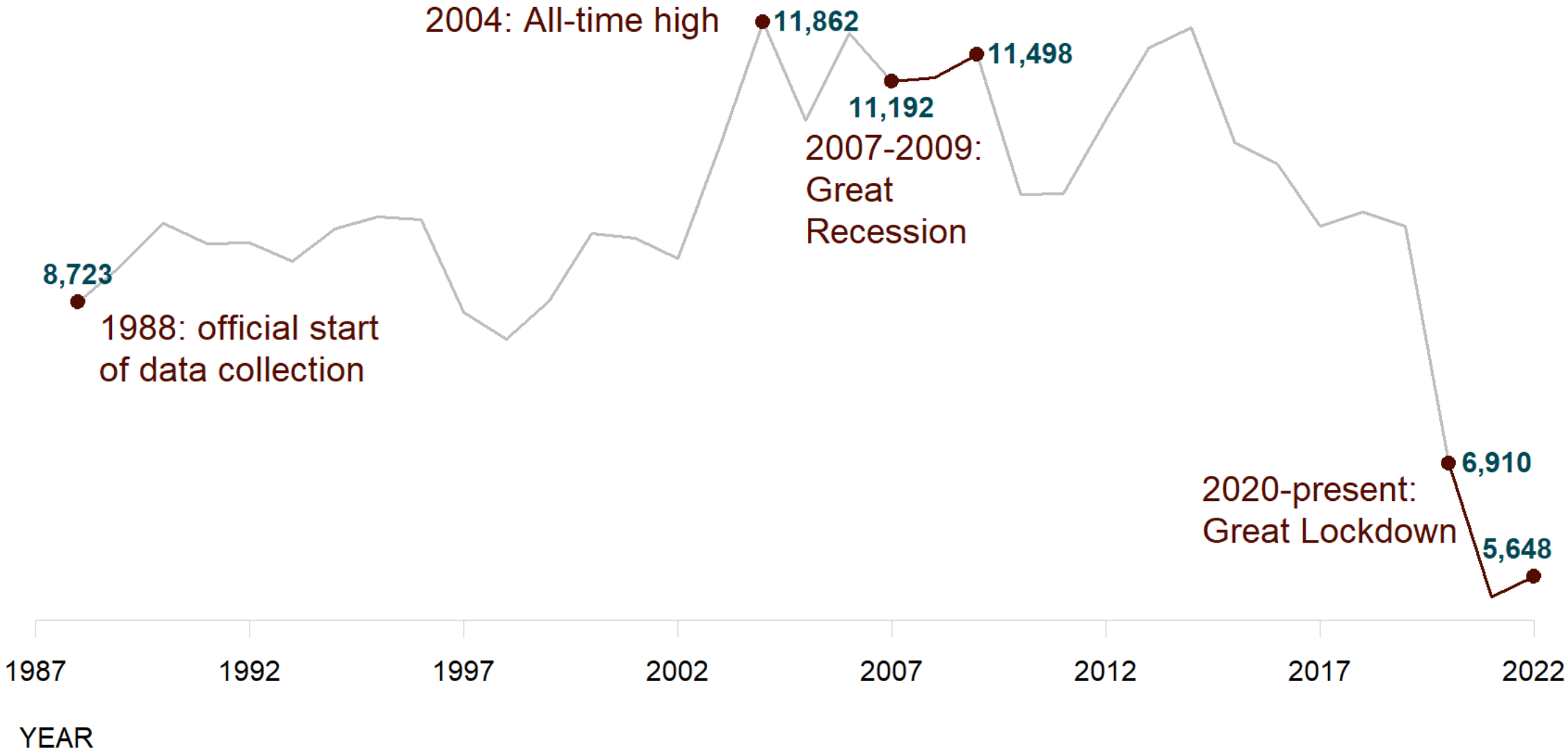
```
theme(axis.text.x=
element_blank(),
axis.ticks.x=
element_blank())
```

• To create horizontal bars

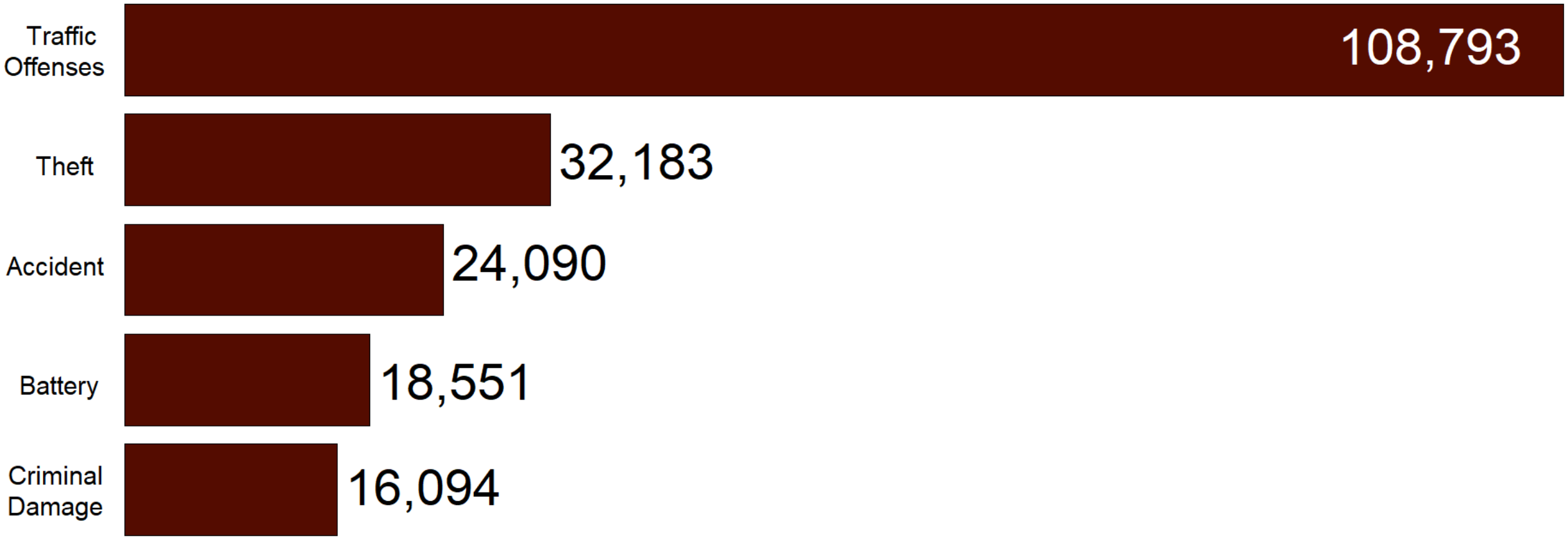
`coord_flip()`

Tip: Space, size, and shape can be used to emphasize and focus the audience's attention.

Number of Incidents Involving the Urbana Police



Top 5 Most Common Crime Categories for Urbana Police during 1988-2022



Base R

The following are arguments within either the `plot()` or `text()` functions.

• To change plotting character shape

`pch=`

• To change plotting character size

`lwd=`

• To change text size

`cex=`

• To change text orientation

`las=`

Tidyverse

The following are arguments within either `ggplot()` or `theme()` functions.

• To change plotting character shape

`shape=`

• To change plotting character size

`size=`

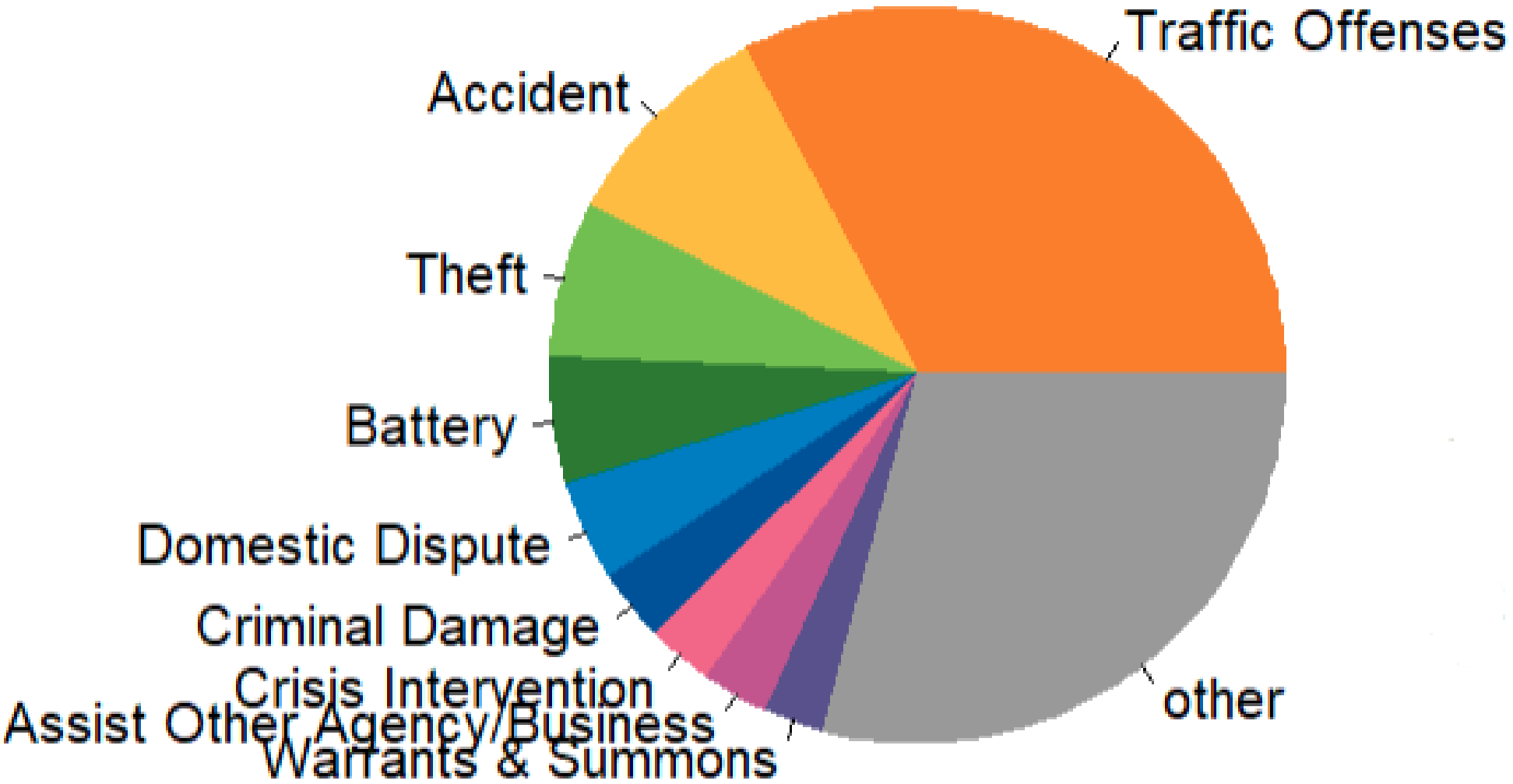
• To change text size

`element_text(size=))`

• To change text orientation

`angle=`

Tip: Do not display screen shots.


Traffic Offenses
Accident
Theft
Battery
Domestic Dispute
Criminal Damage
Crisis Intervention
Assist Other Agency/Business
Warrants & Summons
other

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins | Project: (None)

data-viz-improvements.R

Source on Save | Run | Source

```
1  library(tidyverse)
2  ###############
3  # art theater #
4  ###############
5  film_titles <- read_csv("C:/Users/kinson2/Documents/data_the_a
6  film_titles$EventName
7
8  #using old omdb from RStudio shiny gallery "movie explorer"
9  library(RSQLite)
10 db <- src_sqlite("C:/Users/kinson2/Documents/data_the_art_thea
11 omdb <- tbl(db, "omdb")
12 tomatoes <- tbl(db, "tomatoes")
13
```

1:1 (Top Level) R Script

Console | Terminal | Background Jobs

R 4.2.2 · ~/

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Environment | History | Connections | Tutorial

Import Dataset | 205 MiB | List

R | Global Environment

Environment is empty

Files | Plots | Packages | Help | Viewer | Presentation

Zoom | Export

- Save as Image...
- Save as PDF...
- Copy to Clipboard...

29°F Temps rise

11:40 AM 3/19/2023

Base R

- To save a plot as an image with .png extension

```
png("plot_name.png")
```

- To save a plot as an image with .pdf extension

```
pdf("plot_name.pdf")
```

- To save a plot as an image with .jpeg extension

```
jpeg("plot_name.jpg")
```

- To close the graphics device

```
dev.off()
```

Tidyverse

- To save a plot as an image with .png extension

```
ggsave("plot_name.png")
```

- To save a plot as an image with .pdf extension

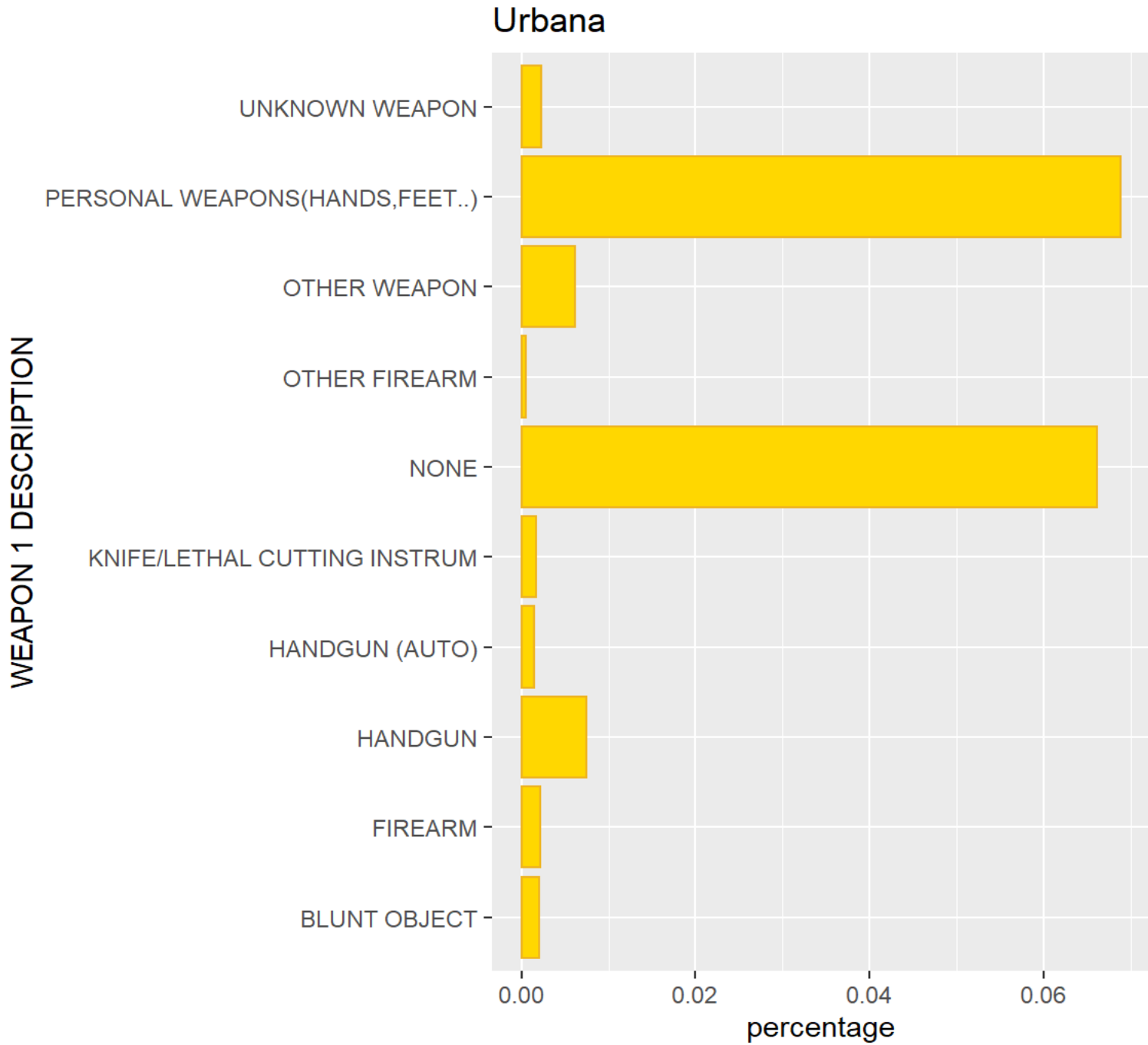
```
ggsave("plot_name.pdf")
```

- To save a plot as an image with .jpeg extension

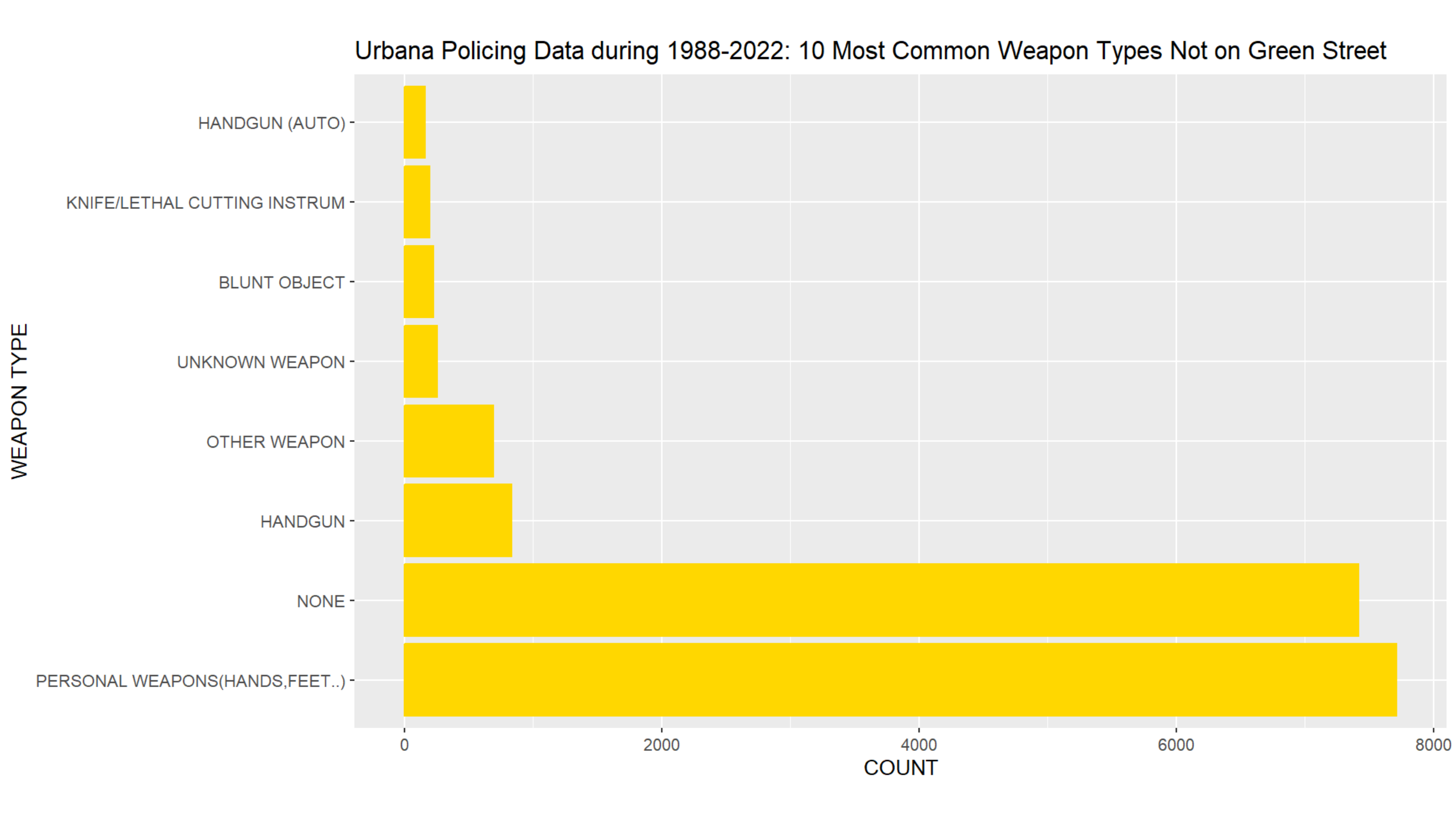
```
ggsave("plot_name.jpg")
```

Tip: Employ visual hierarchy, which includes removing clutter and using pre-attentive attributes such as size, shape, and color.

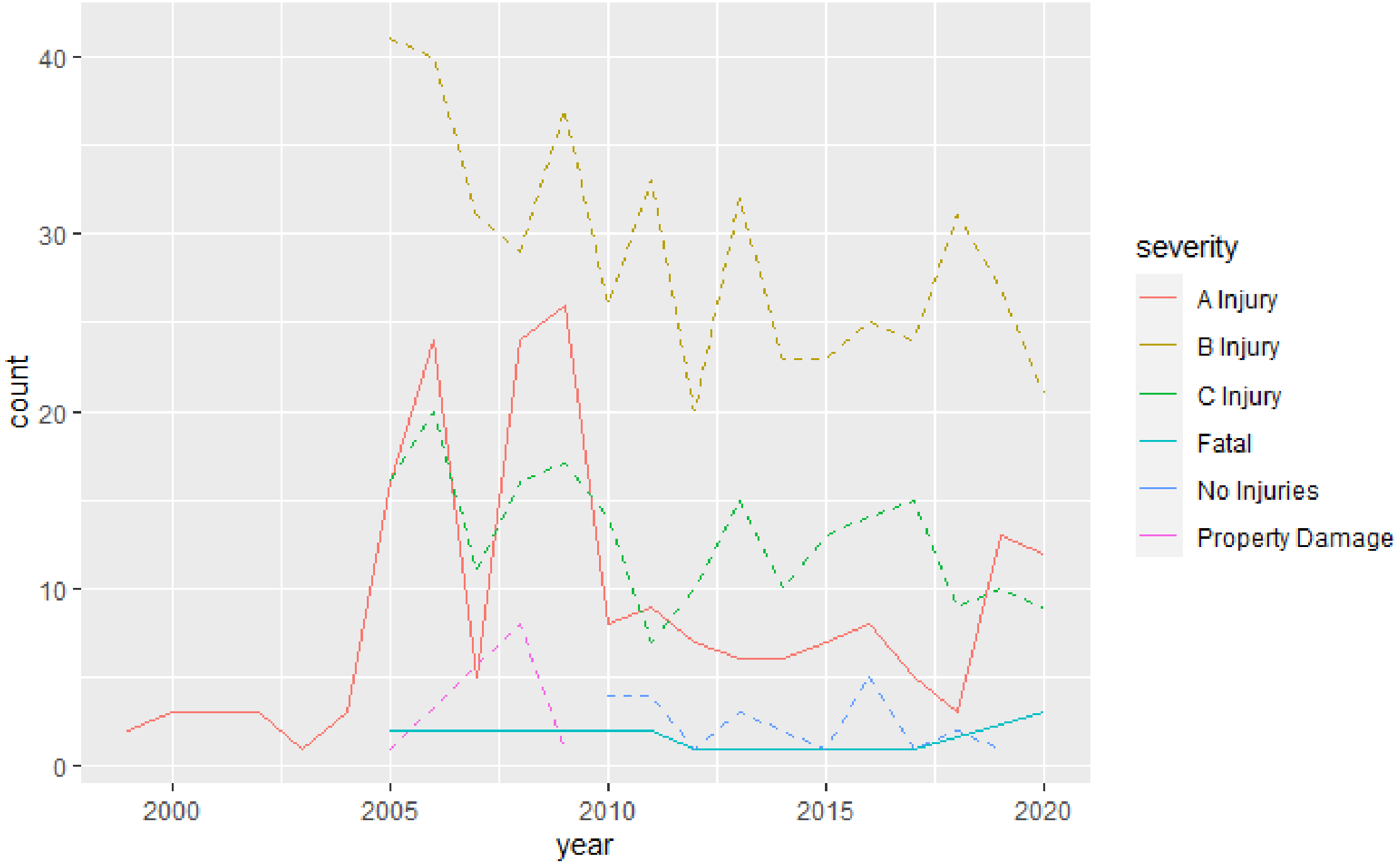
Original


Urbana
WEAPON 1 DESCRIPTION
UNKNOWN WEAPON
PERSONAL WEAPONS(HANDS,FEET..)
OTHER WEAPON
OTHER FIREARM
NONE
KNIFE/LETHAL CUTTING INSTRUM
HANDGUN (AUTO)
HANDGUN
FIREARM
BLUNT OBJECT
0.00
0.02
0.04
0.06
percentage

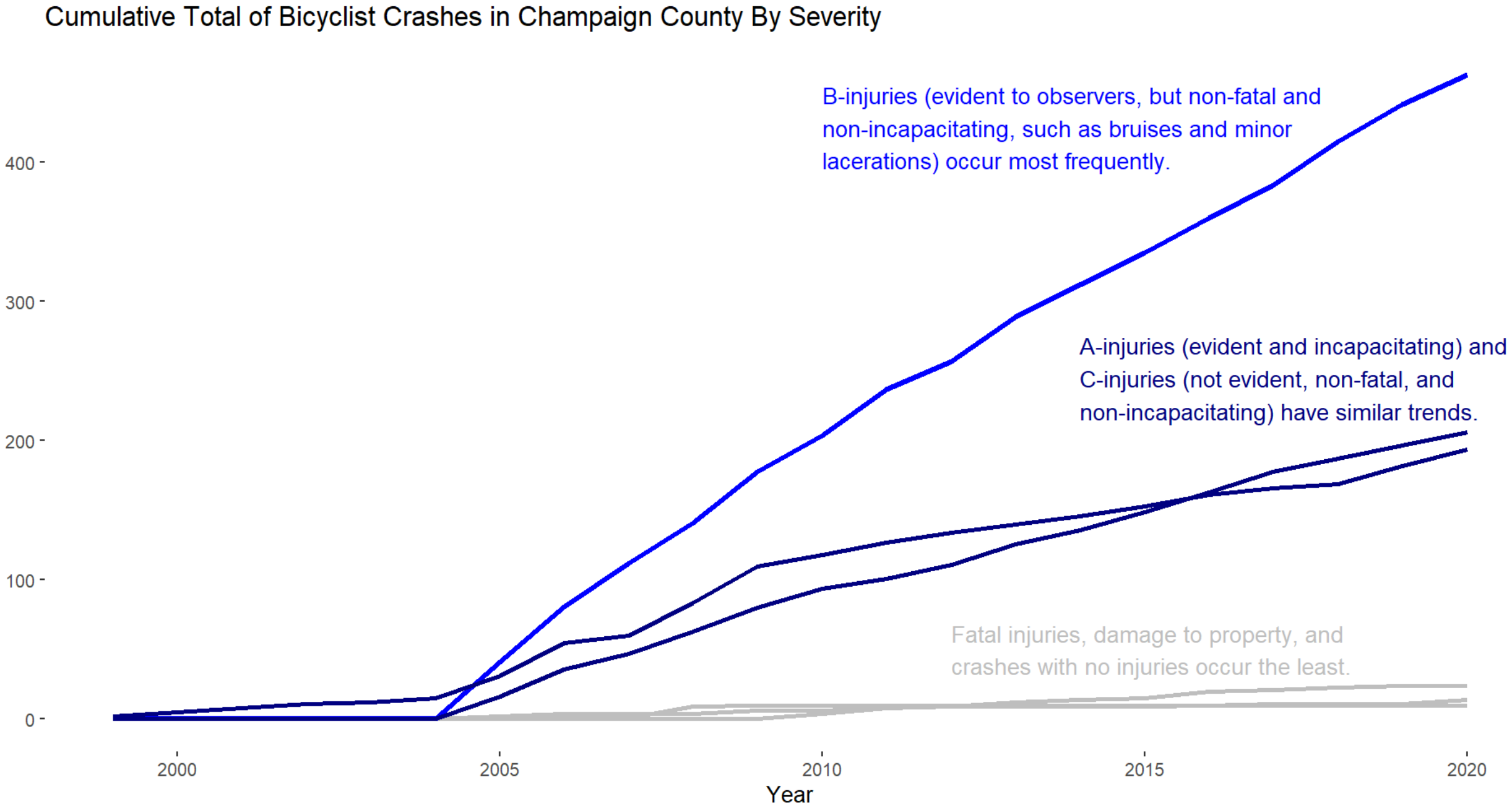
Improvement


Urbana Policing Data during 1988-2022: 10 Most Common Weapon Types Not on Green Street
WEAPON TYPE
HANDGUN (AUTO)
KNIFE/LETHAL CUTTING INSTRUM
BLUNT OBJECT
UNKNOWN WEAPON
OTHER WEAPON
HANDGUN
NONE
PERSONAL WEAPONS(HANDS,FEET..)
0
2000
4000
6000
8000
COUNT

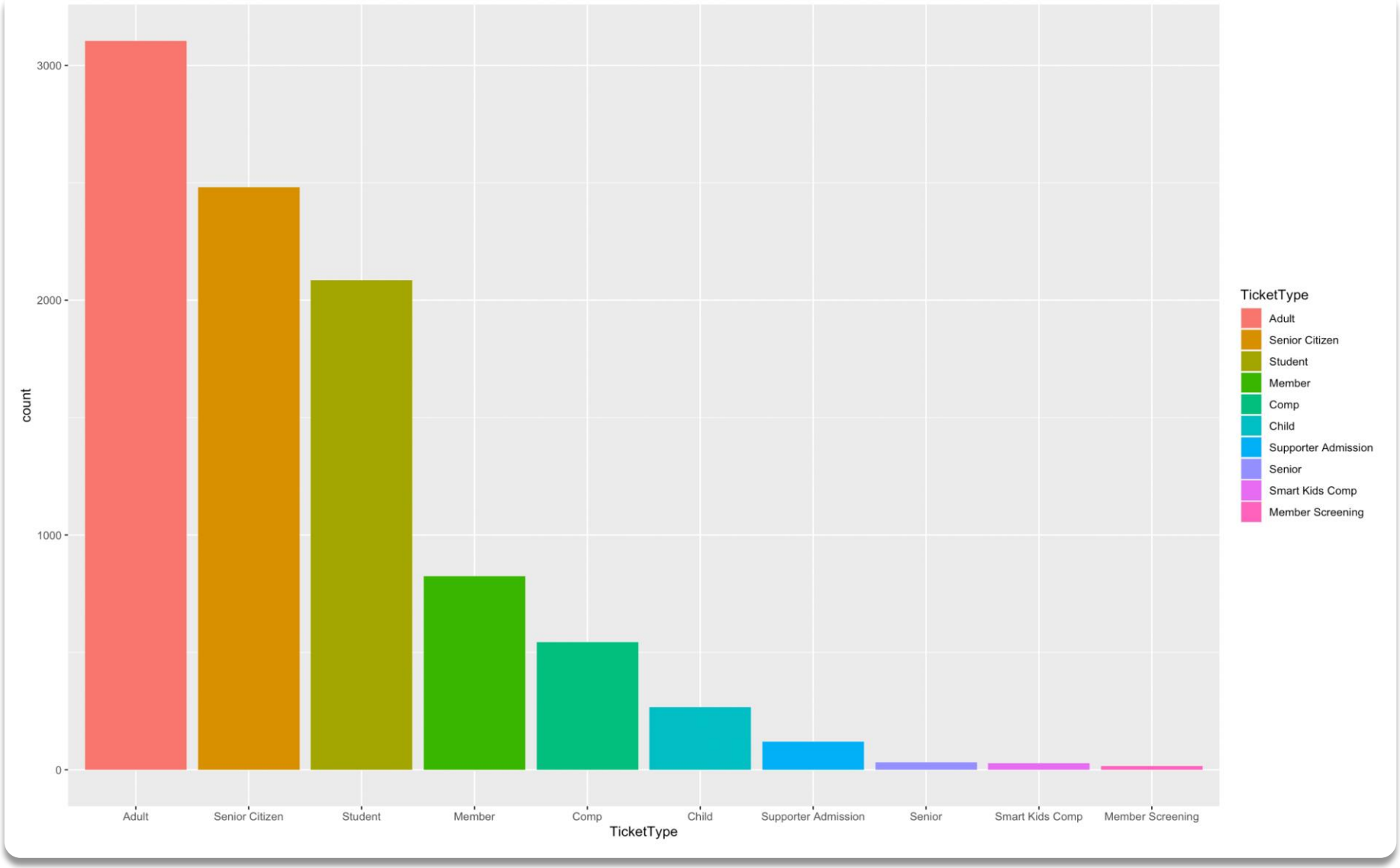
Original


40
30
20
10
0
count
2000
2005
2010
2015
2020
year
severity
A Injury
B Injury
C Injury
Fatal
No Injuries
Property Damage

Improvement

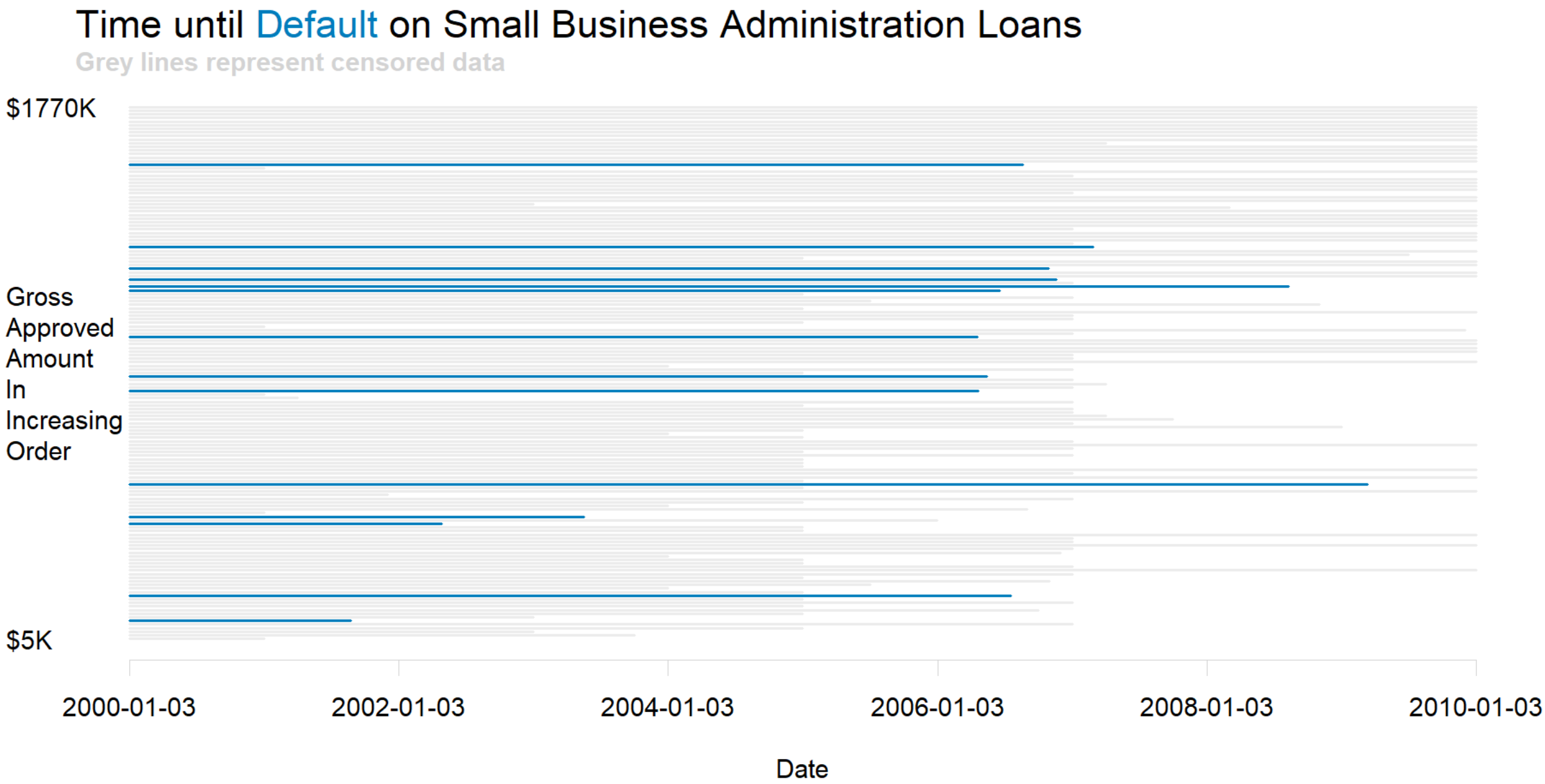

Cumulative Total of Bicyclist Crashes in Champaign County By Severity
B-injuries (evident to observers, but non-fatal and non-incapacitating, such as bruises and minor lacerations) occur most frequently.
A-injuries (evident and incapacitating) and C-injuries (not evident, non-fatal, and non-incapacitating) have similar trends.
Fatal injuries, damage to property, and crashes with no injuries occur the least.
400
300
200
100
0
2000
2005
2010
2015
2020
Year

Based on previous tips, what would you change?



Exercise: Using the SBA Loans Data and starter code in the GitHub repo, create a data visualization that addresses the question, **“How long are businesses taking to default on their loan?”**

Intentionally left blank


Time until Default on Small Business Administration Loans
Grey lines represent censored data
$1770K
Gross Approved Amount In Increasing Order
$5K
2000-01-03
2002-01-03
2004-01-03
2006-01-03
2008-01-03
2010-01-03
Date