Spokane Population Growth and Crime

FBI UCR Data (1985-2014)

Tyler Hart

April 8, 2019

Someone I was talking with recently about Spokane's economy and growth mentioned their concern about crime increasing as the city grows. It makes sense on the surface - more people could mean more criminals, and therefore more crime. But there are many factors that contribute to an environment where criminality occurs. Communities also experience a diversity of crime - violent and non-violent. Within each type are varying degrees of offense as well. Saying "crime went up" may be accurate, but within that trend violent crime (murder, robbery, etc.) could go down while non-violent crime (theft, malicious mischief) goes up. I decided to do a quick exploration of what data is available and what it says about changes in population and crime. We'll compare different types of crime, both violent and non-violent, with population data to see if correlations exist.

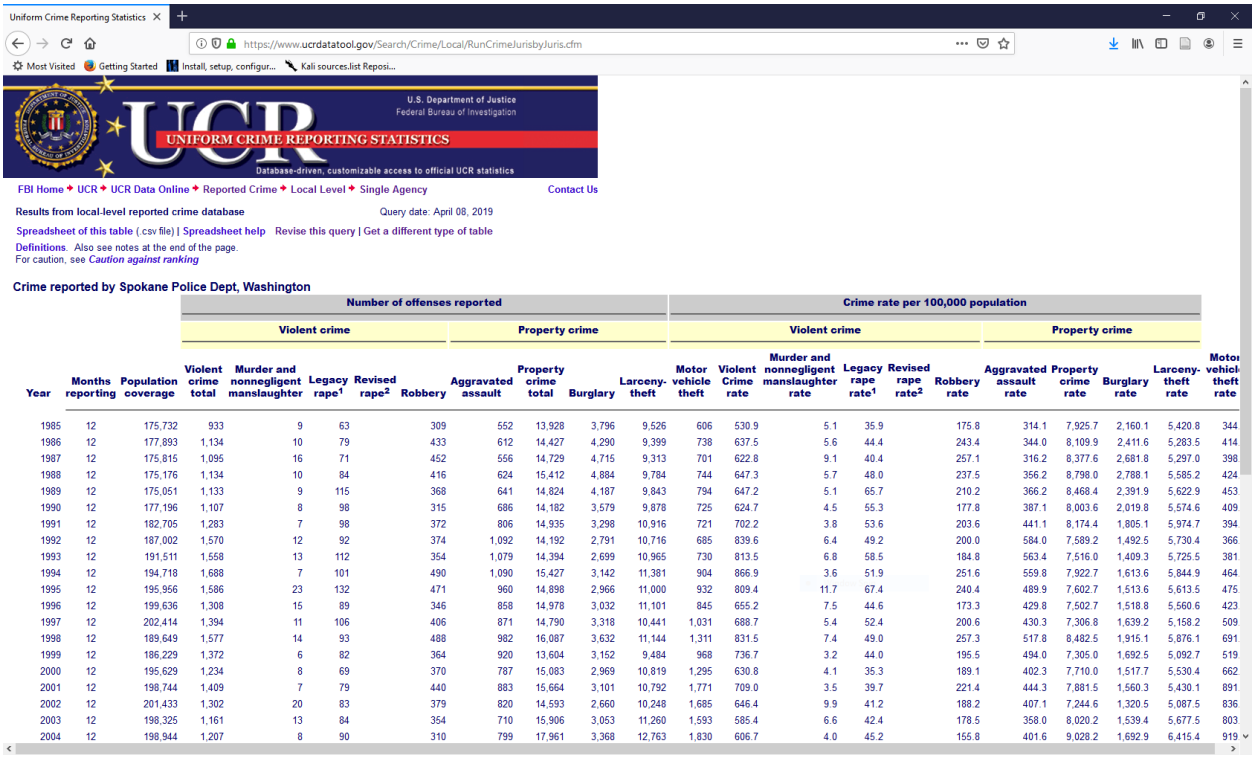
Data

The Federal Bureau of Investigation (FBI) makes crime data available for many localities through the Uniform Crime Reporting (UCR) webpage. In the case of Spokane, WA there is data available from 1985 through 2014. All the data used in this report is sourced from the FBI UCR webpage at this URL:

https://www.ucrdatatool.gov/Search/Crime/Local/JurisbyJuris.cfm

We can select from a number of crime types. For this report we chose all available types. The output of a search for Spokane data results in a simple HTML table output to the screen:

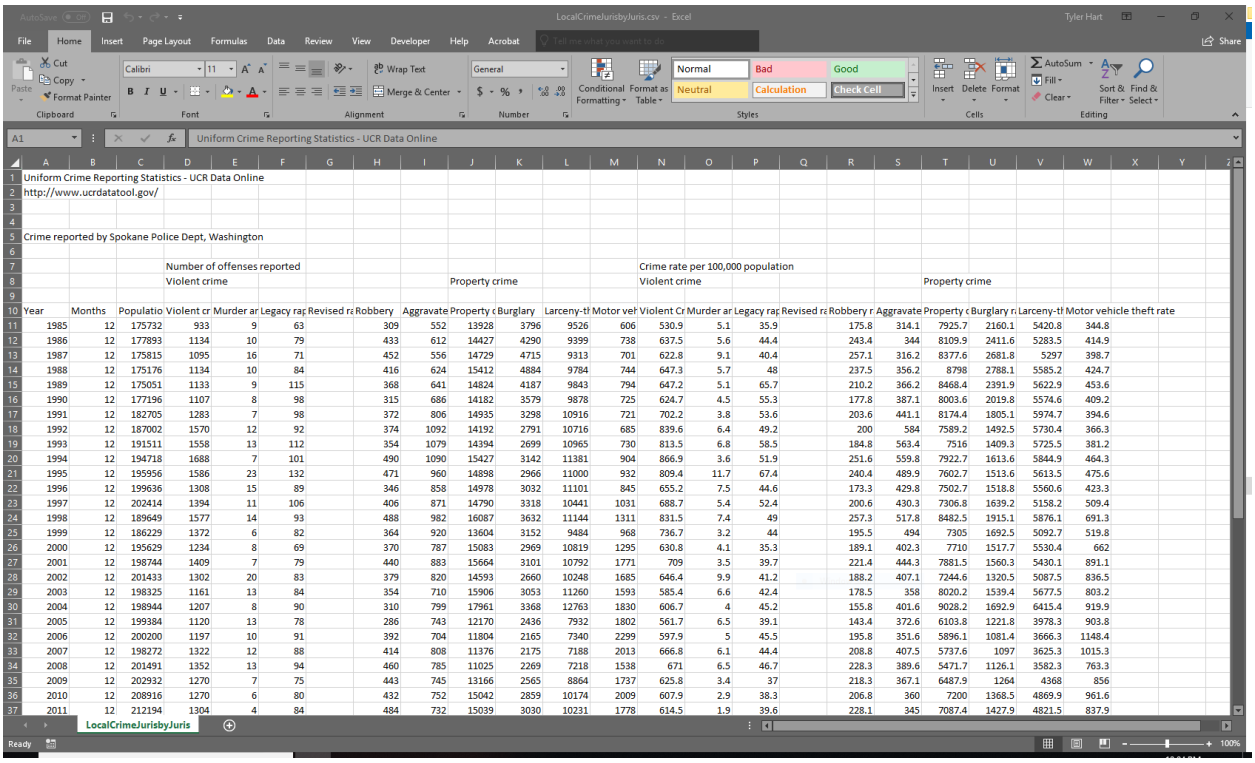
```
knitr::include_graphics('./ucr_data_table_snip.png')
```



			Number of offenses reported									Crime rate per 100,000 population									
			Violent crime					Property crime				Violent crime					Property crime				
Year	Months reporting	Population coverage	Violent crime total	Murder and nonnegligent manslaughter	Legacy rape[1]	Revised rape[2]	Robbery	Aggravated assault	Property crime total	Burglary	Larceny-theft	Motor vehicle theft	Violent Crime rate	Murder and nonnegligent manslaughter rate	Legacy rape rate[1]	Revised rape rate[2]	Robbery rate	Aggravated assault rate	Property crime rate	Burglary rate	Larceny-theft rate	Motor vehicle theft rate
1985	12	175,732	933	9	63		309	552	13,928	3,796	9,526	606	530.9	5.1	35.9		175.8	314.1	7,925.7	2,160.1	5,420.8	344
1986	12	177,893	1,134	10	79		433	612	14,427	4,290	9,399	738	637.5	5.6	44.4		243.4	344.0	8,109.9	2,411.6	5,283.5	414
1987	12	175,815	1,095	16	71		452	556	14,729	4,715	9,313	701	622.8	9.1	40.4		257.1	316.2	8,377.6	2,681.8	5,297.0	398
1988	12	175,176	1,134	10	84		416	624	15,412	4,884	9,784	744	647.3	5.7	48.0		237.5	356.2	8,798.0	2,788.1	5,585.2	424
1989	12	175,051	1,133	9	115		368	641	14,824	4,187	9,843	794	647.2	5.1	65.7		210.2	366.2	8,468.4	2,391.9	5,622.9	453
1990	12	177,196	1,107	8	98		315	686	14,182	3,579	9,878	725	624.7	4.5	55.3		177.8	387.1	8,003.6	2,019.8	5,574.6	409
1991	12	182,705	1,283	7	98		372	806	14,935	3,298	10,916	721	702.2	3.8	53.6		203.6	441.1	8,174.4	1,805.1	5,974.7	394
1992	12	187,002	1,570	12	92		374	1,092	14,192	2,791	10,716	685	839.6	6.4	49.2		200.0	584.0	7,589.2	1,492.5	5,730.4	366
1993	12	191,511	1,558	13	112		354	1,079	14,394	2,699	10,965	730	813.5	6.8	58.5		184.8	563.4	7,516.0	1,409.3	5,725.5	381
1994	12	194,718	1,688	7	101		490	1,090	15,427	3,142	11,381	904	866.9	3.6	51.9		251.6	559.8	7,922.7	1,613.6	5,844.9	464
1995	12	195,956	1,586	23	132		471	960	14,898	2,966	11,000	932	809.4	11.7	67.4		240.4	489.9	7,602.7	1,513.6	5,613.5	475
1996	12	199,636	1,308	15	89		346	858	14,978	3,032	11,101	845	655.2	7.5	44.6		173.3	429.8	7,502.7	1,518.8	5,560.6	423
1997	12	202,414	1,394	11	106		406	871	14,790	3,318	10,441	1,031	688.7	5.4	52.4		200.6	430.3	7,306.8	1,639.2	5,158.2	509
1998	12	189,649	1,577	14	93		488	982	16,087	3,632	11,144	1,311	831.5	7.4	49.0		257.3	517.8	8,482.5	1,915.1	5,876.1	691
1999	12	186,229	1,372	6	82		364	920	13,604	3,152	9,484	968	736.7	3.2	44.0		195.5	494.0	7,305.0	1,692.5	5,092.7	519
2000	12	195,629	1,234	8	69		370	787	15,083	2,969	10,819	1,295	630.8	4.1	35.3		189.1	402.3	7,710.0	1,517.7	5,530.4	662
2001	12	198,744	1,409	7	79		440	883	15,664	3,101	10,792	1,771	709.0	3.5	39.7		221.4	444.3	7,881.5	1,560.3	5,430.1	891
2002	12	201,433	1,302	20	83		379	820	14,593	2,660	10,248	1,685	646.4	9.9	41.2		188.2	407.1	7,244.6	1,320.5	5,087.5	836
2003	12	198,325	1,161	13	84		354	710	15,906	3,053	11,260	1,593	585.4	6.6	42.4		178.5	358.0	8,020.2	1,539.4	5,677.5	803
2004	12	198,944	1,207	8	90		310	799	17,961	3,368	12,763	1,830	606.7	4.0	45.2		155.8	401.6	9,028.2	1,692.9	6,415.4	919

Fortunately, there is a handy link to download the data in CSV format so no web scraping is required. The download is named "LocalCrimeJurisbyJuris.csv". Unfortunately, some HTML header information from the webpage is included in the CSV file, along with loads of extraneous spaces. We can see the extra "data" when opening the CSV file in Excel:

```
knitr::include_graphics('./ucr_csv_formatting_trouble.png')
```



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	Uniform Crime Reporting Statistics - UCR Data Online																								
2	http://www.ucrdatatool.gov/																								
3																									
4																									
5	Crime reported by Spokane Police Dept, Washington																								
6																									
7				Number of offenses reported										Crime rate per 100,000 population											
8				Violent crime						Property crime				Violent crime							Property crime				
9																									
10	Year	Months	Populatio	Violent cr	Murder ar	Legacy rap	Revised ra	Robbery	Aggravate	Property c	Burglary	Larceny-th	Motor veh	Violent Cr	Murder ar	Legacy rap	Revised ra	Robbery r	Aggravate	Property c	Burglary r	Larceny-th	Motor vehicle theft rate		
11	1985	12	175732	933	9	63		309	552	13928	3796	9526	606	530.9	5.1	35.9		175.8	314.1	7925.7	2160.1	5420.8	344.8		
12	1986	12	177893	1134	10	79		433	612	14427	4290	9399	738	637.5	5.6	44.4		243.4	344	8109.9	2411.6	5283.5	414.9		
13	1987	12	175815	1095	16	71		452	556	14729	4715	9313	701	622.8	9.1	40.4		257.1	316.2	8377.6	2681.8	5297	398.7		
14	1988	12	175176	1134	10	84		416	624	15412	4884	9784	744	647.3	5.7	48		237.5	356.2	8798	2788.1	5585.2	424.7		
15	1989	12	175051	1133	9	115		368	641	14824	4187	9843	794	647.2	5.1	65.7		210.2	366.2	8468.4	2391.9	5622.9	453.6		
16	1990	12	177196	1107	8	98		315	686	14182	3579	9878	725	624.7	4.5	55.3		177.8	387.1	8003.6	2019.8	5574.6	409.2		
17	1991	12	182705	1283	7	98		372	806	14935	3298	10916	721	702.2	3.8	53.6		203.6	441.1	8174.4	1805.1	5974.7	394.6		
18	1992	12	187002	1570	12	92		374	1092	14192	2791	10716	685	839.6	6.4	49.2		200	584	7589.2	1492.5	5730.4	366.3		
19	1993	12	191511	1558	13	112		354	1079	14394	2699	10965	730	813.5	6.8	58.5		184.8	563.4	7516	1409.3	5725.5	381.2		
20	1994	12	194718	1688	7	101		490	1090	15427	3142	11381	904	866.9	3.6	51.9		251.6	559.8	7922.7	1613.6	5844.9	464.3		
21	1995	12	195956	1586	23	132		471	960	14898	2966	11000	932	809.4	11.7	67.4		240.4	489.9	7602.7	1513.6	5613.5	475.6		
22	1996	12	199636	1308	15	89		346	858	14978	3032	11101	845	655.2	7.5	44.6		173.3	429.8	7502.7	1518.8	5560.6	423.3		
23	1997	12	202414	1394	11	106		406	871	14790	3318	10441	1031	688.7	5.4	52.4		200.6	430.3	7306.8	1639.2	5158.2	509.4		
24	1998	12	189649	1577	14	93		488	982	16087	3632	11144	1311	831.5	7.4	49		257.3	517.8	8482.5	1915.1	5876.1	691.3		
25	1999	12	186229	1372	6	82		364	920	13604	3152	9484	968	736.7	3.2	44		195.5	494	7305	1692.5	5092.7	519.8		
26	2000	12	195629	1234	8	69		370	787	15083	2969	10819	1295	630.8	4.1	35.3		189.1	402.3	7710	1517.7	5530.4	662		
27	2001	12	198744	1409	7	79		440	883	15664	3101	10792	1771	709	3.5	39.7		221.4	444.3	7881.5	1560.3	5430.1	891.1		
28	2002	12	201433	1302	20	83		379	820	14593	2660	10248	1685	646.4	9.9	41.2		188.2	407.1	7244.6	1320.5	5087.5	836.5		
29	2003	12	198325	1161	13	84		354	710	15906	3053	11260	1593	585.4	6.6	42.4		178.5	358	8020.2	1539.4	5677.5	803.2		
30	2004	12	198944	1207	8	90		310	799	17961	3368	12763	1830	606.7	4	45.2		155.8	401.6	9028.2	1692.9	6415.4	919.9		
31	2005	12	199384	1120	13	78		286	743	12170	2436	7932	1802	561.7	6.5	39.1		143.4	372.6	6103.8	1221.8	3978.3	903.8		
32	2006	12	200200	1197	10	91		392	704	11804	2165	7340	2299	597.9	5	45.5		195.8	351.6	5896.1	1081.4	3666.3	1148.4		
33	2007	12	198272	1322	12	88		414	808	11376	2175	7188	2013	666.8	6.1	44.4		208.8	407.5	5737.6	1097	3625.3	1015.3		
34	2008	12	201491	1352	13	94		460	785	11025	2269	7218	1538	671	6.5	46.7		228.3	389.6	5471.7	1126.1	3582.3	763.3		
35	2009	12	202932	1270	7	75		443	745	13166	2565	8864	1737	625.8	3.4	37		218.3	367.1	6487.9	1264	4368	856		
36	2010	12	208916	1270	6	80		432	752	15042	2859	10174	2009	607.9	2.9	38.3		206.8	360	7200	1368.5	4869.9	961.6		
37	2011	12	212194	1304	4	84		484	732	15039	3030	10231	1778	614.5	1.9	39.6		228.1	345	7087.4	1427.9	4821.5	837.9		

We need to strip out data from the first 10 rows in the file and the last 11 rows. Using Excel is the quickest way to drop the extra stuff and get to the data. Next, we import raw data from the CSV and format the *Year* column:

```
df.crime_data <- read_csv(
  "LocalCrimeJurisbyJuris.csv",
  col_types = cols(Year = col_date(format = "%Y"))
  )
```

To make handling columns in the data frame easier we'll switch all the names from mixed-case to lower-case:

```
names(df.crime_data) <- tolower(names(df.crime_data))
```

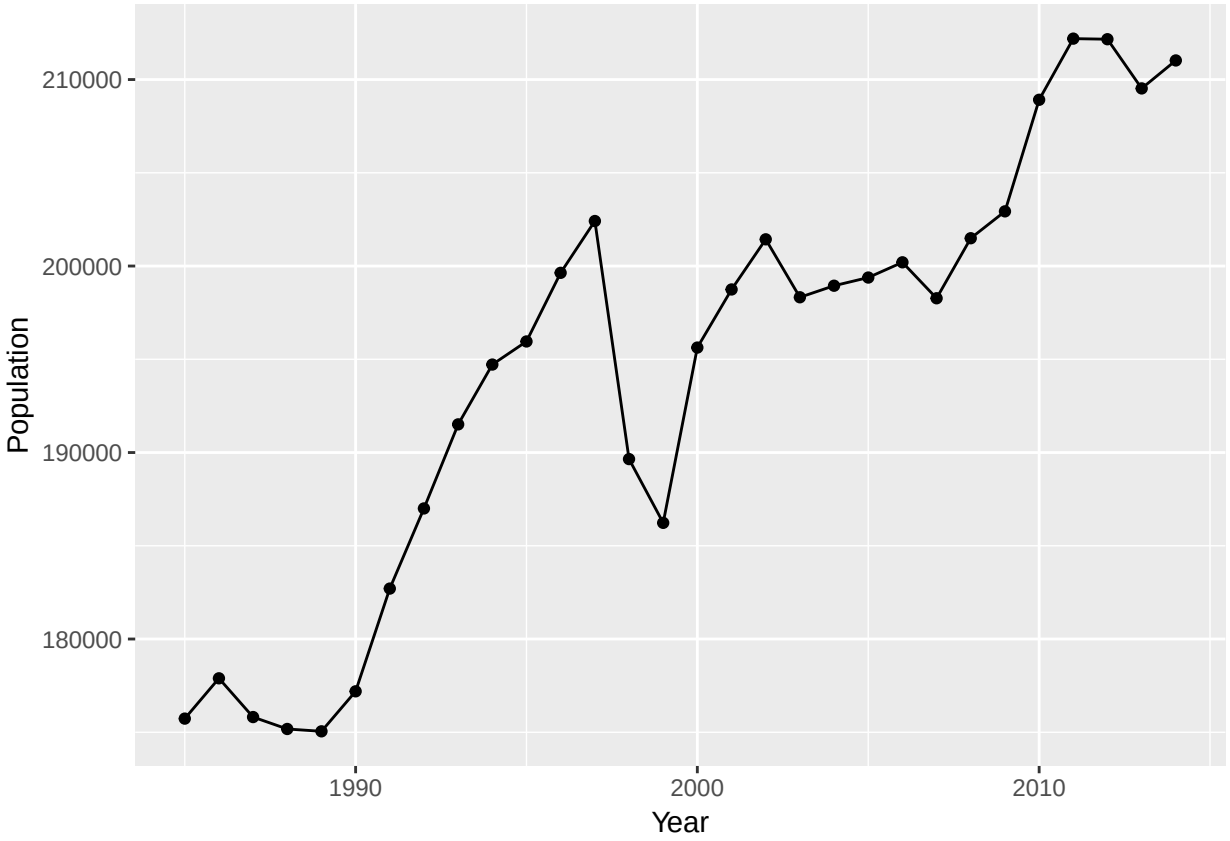
Next, we'll load up the *df.crime_data* data frame into a reusable ggplot object. In the process we'll also set the *year* column as the *X* aesthetic so it won't need to be done repeatedly:

```
plot.annual_crime <- df.crime_data %>%
  ggplot(aes(year)) +
  xlab("Year")
```

Setting the X-axis label in the ggplot object also saves us work later. Let's leverage that reusable plot object and create a basic line plot of overall population. We'll also include the data points to make the line more meaningful. This plot gives us a feel for how Spokane has grown from 1985-2014:

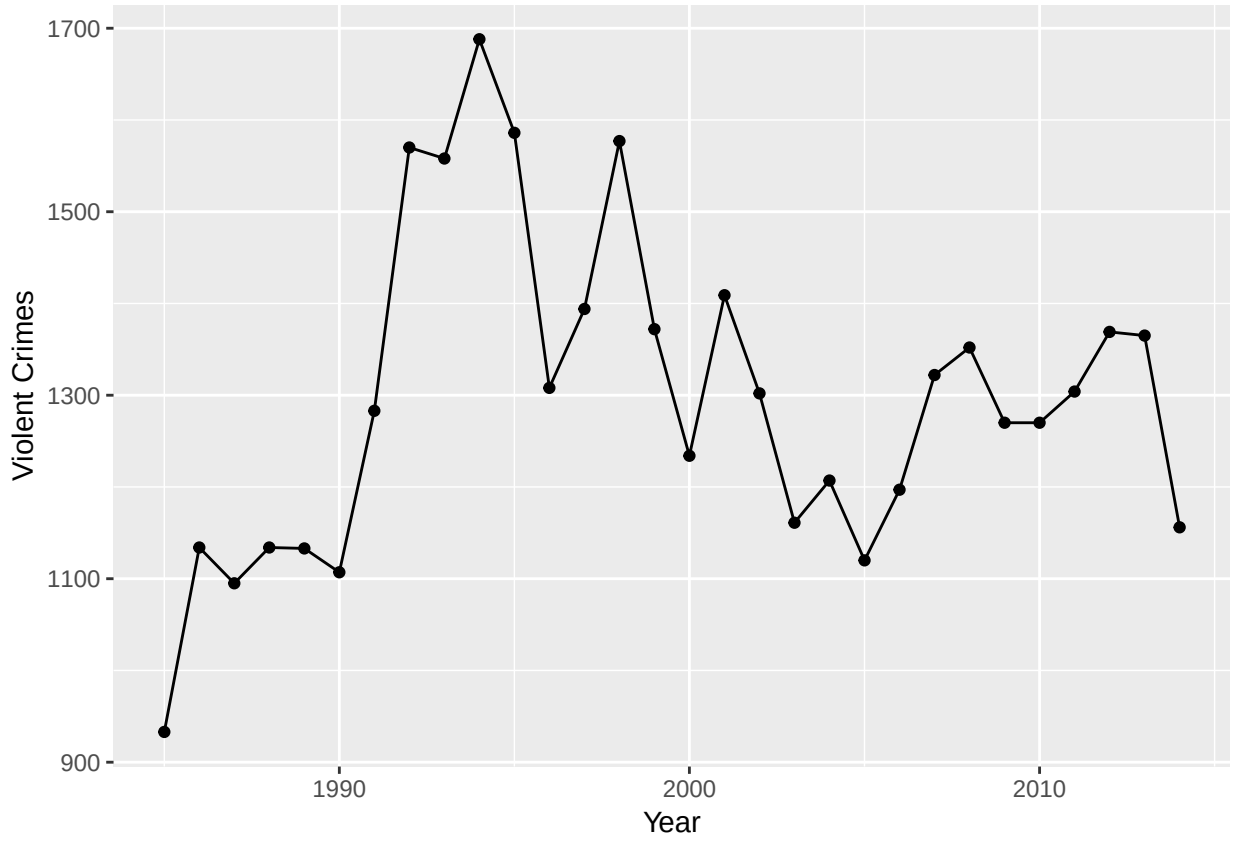
```
plot.annual_crime +
  geom_line(aes(y = population)) +
```

```
  geom_point(aes(y = population)) +
  ylab("Population")
```



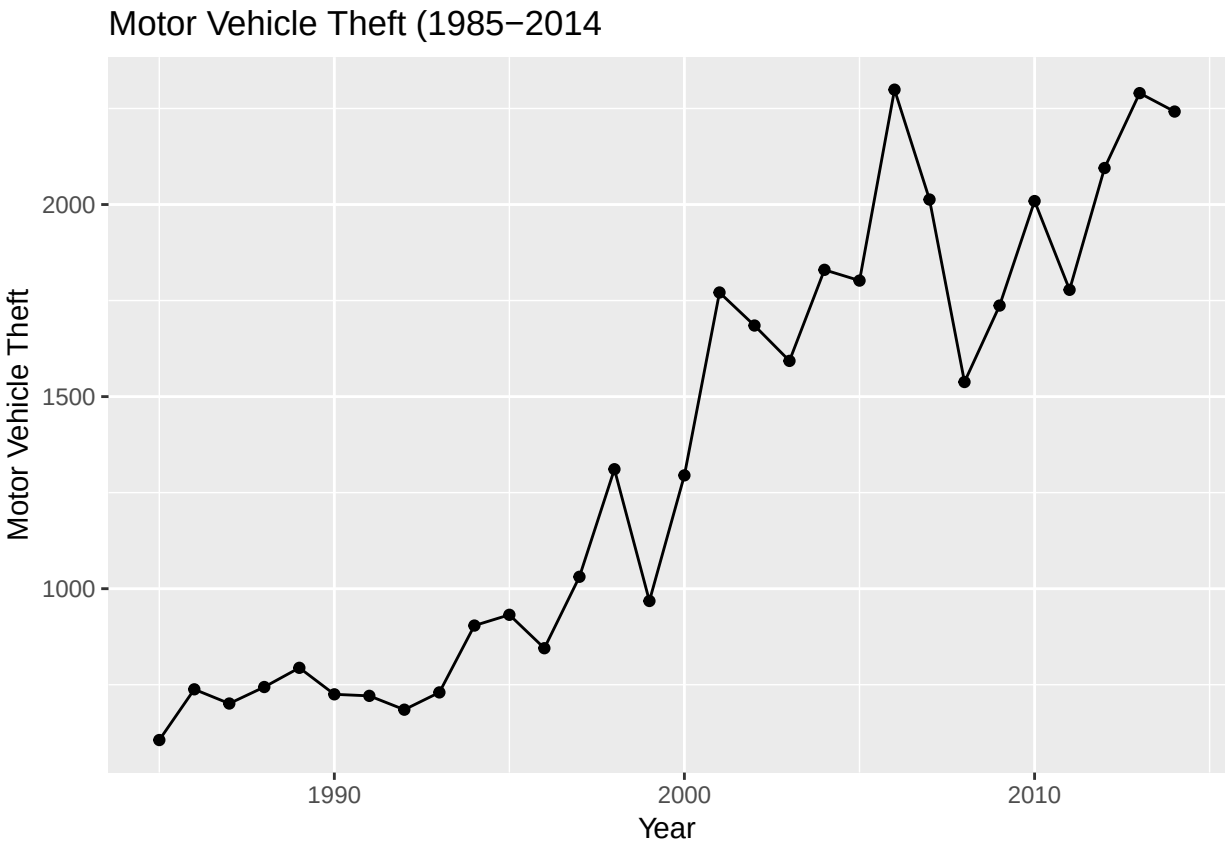
The population has grown quite a bit according to the graph. Let's create the same plot, but with the total number of violent crimes per year instead of population numbers:

```
plot.annual_crime +
  geom_line(aes(y = `violent crime total`)) +
  geom_point(aes(y = `violent crime total`)) +
  ylab("Violent Crimes")
```



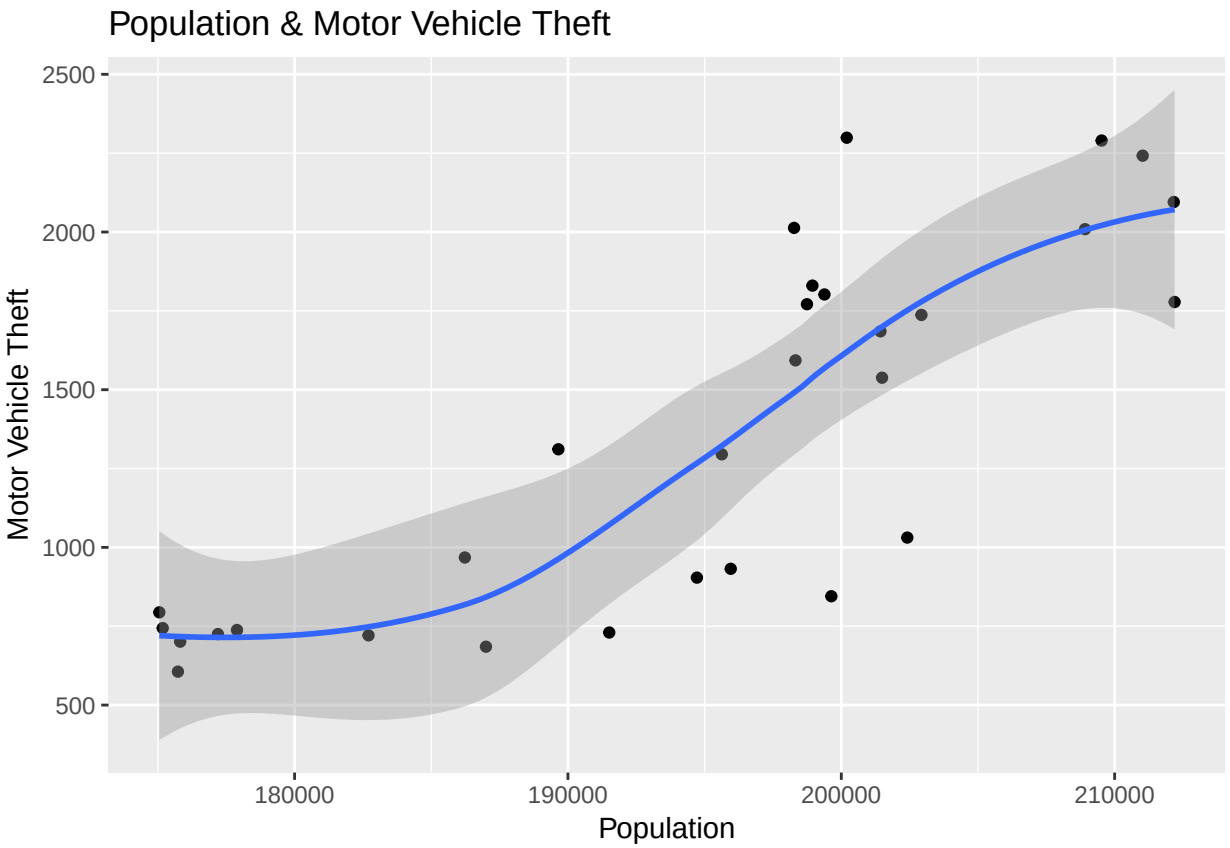
There doesn't appear to be a hard-and-fast link between violent crime numbers and population count. Let's create the same plot for motor vehicle theft and compare again:

```
plot.annual_crime +
  geom_line(aes(y = `motor vehicle theft`)) +
  geom_point(aes(y = `motor vehicle theft`)) +
  ylab("Motor Vehicle Theft") +
  ggtitle("Motor Vehicle Theft (1985-2014")
```



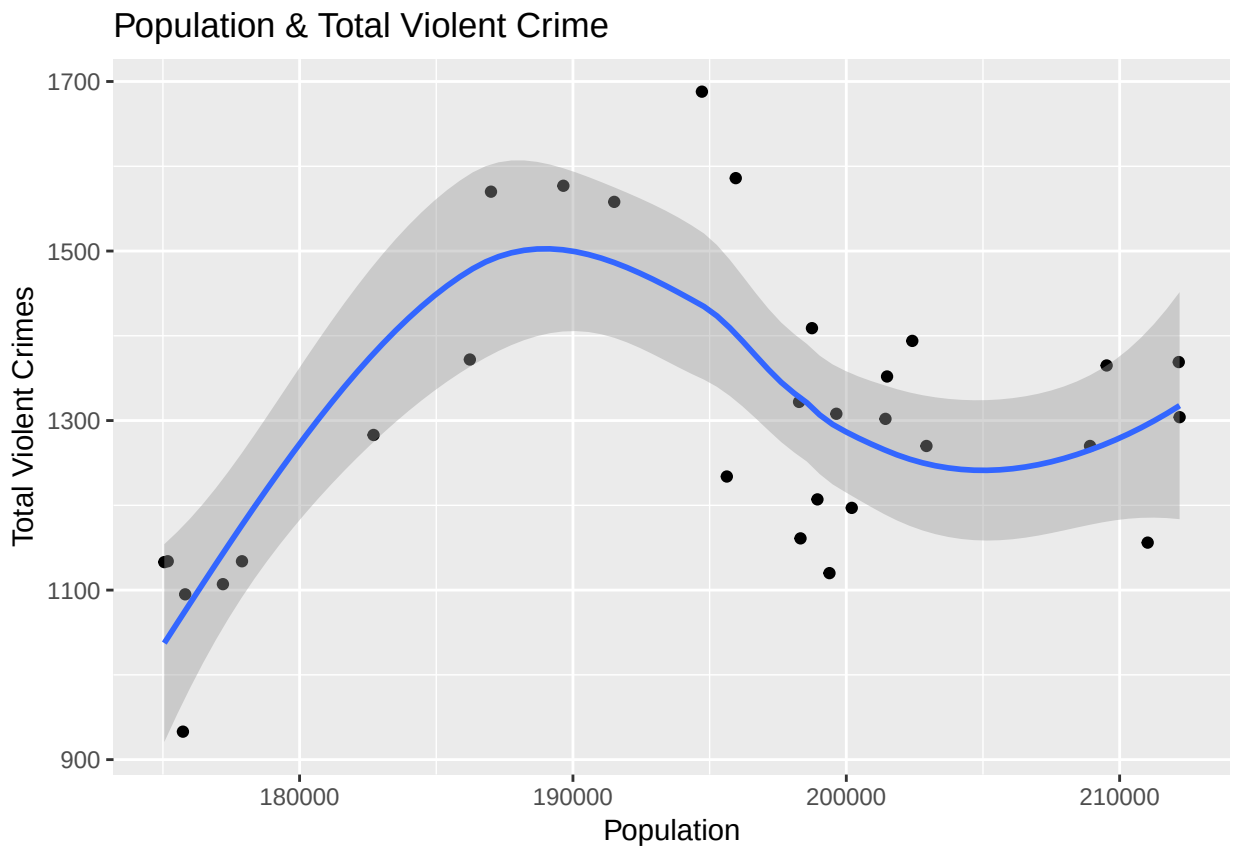
In the case of motor vehicle theft there does appear to be a steady upward trend that resembles population growth. Let's use a scatter plot to visualize the two values together:

```
df.crime_data %>%
  ggplot(aes(population, `motor vehicle theft`)) +
  geom_point() +
  geom_smooth() +
  xlab("Population") +
  ylab("Motor Vehicle Theft") +
  ggtitle("Population & Motor Vehicle Theft")
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

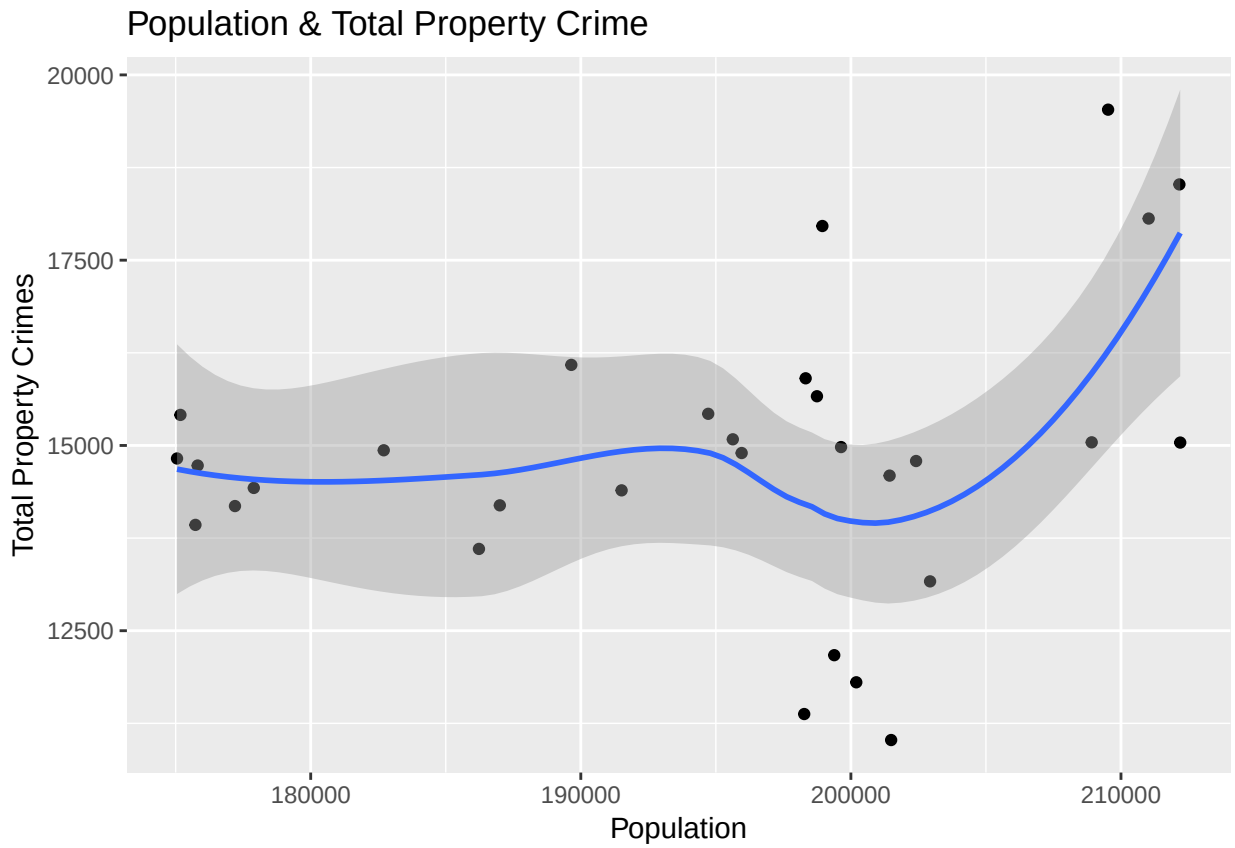


It looks like a relationship may exist between population and vehicle thefts. We'll use scatter plots for the other types of crime as well:

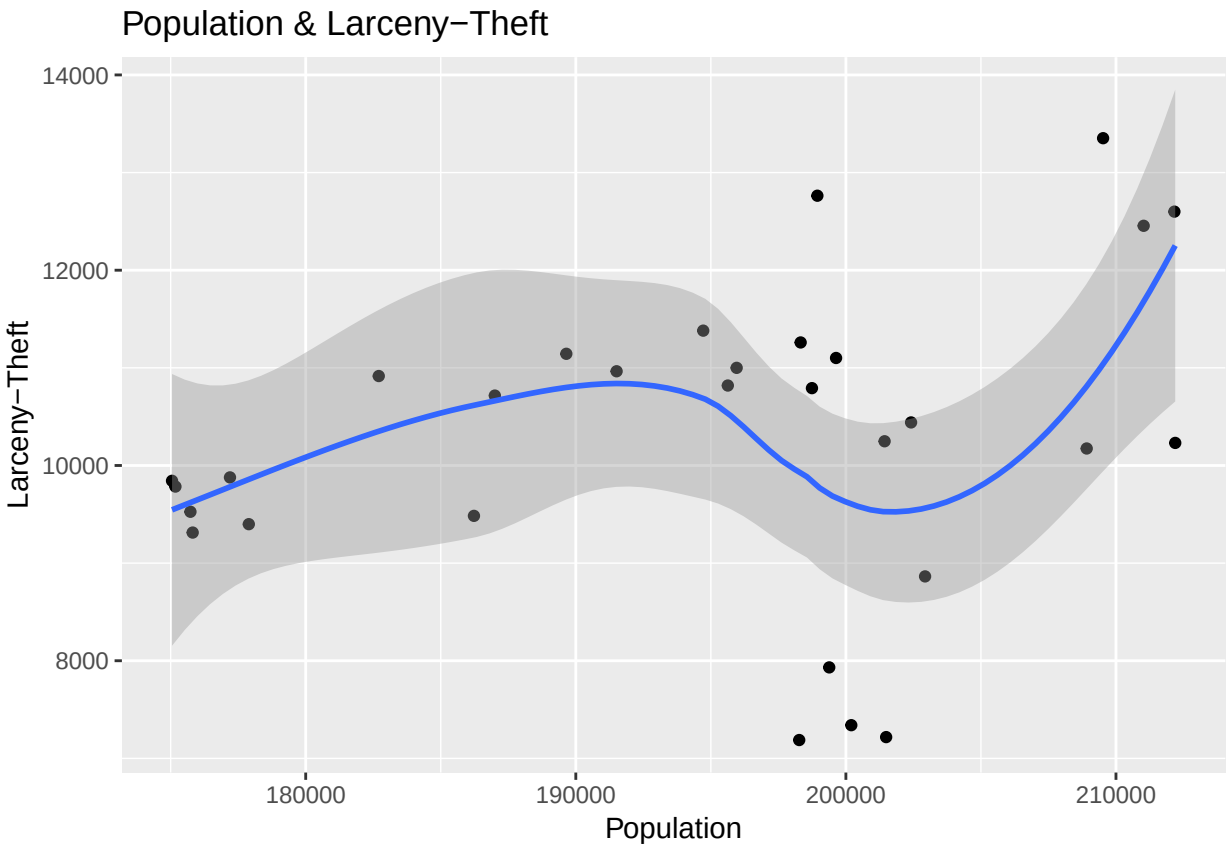
```
df.crime_data %>%
  ggplot(aes(population, `violent crime total`)) +
  geom_point() +
  geom_smooth() +
  xlab("Population") +
  ylab("Total Violent Crimes") +
  ggtitle("Population & Total Violent Crime")
```



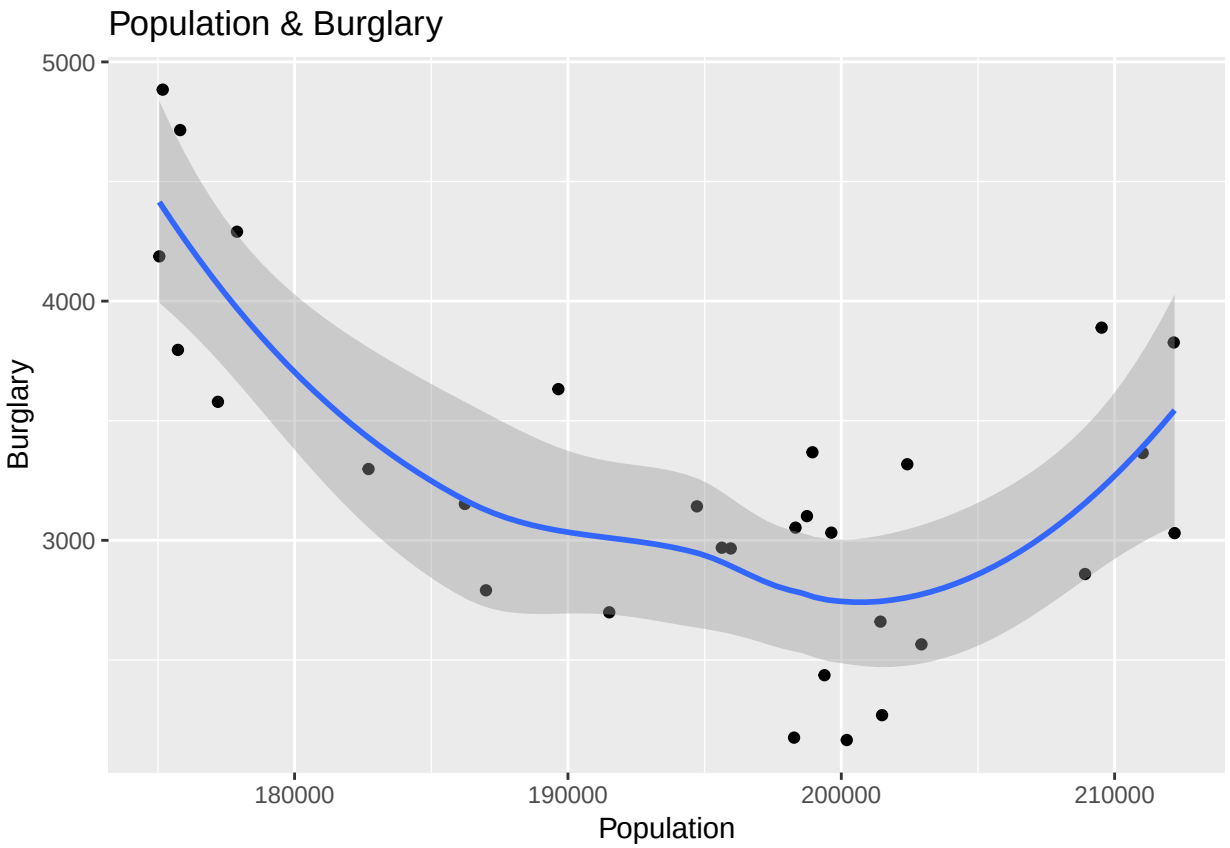

```
df.crime_data %>%
  ggplot(aes(population, `property crime total`)) +
  geom_point() +
  geom_smooth() +
  xlab("Population") +
  ylab("Total Property Crimes") +
  ggtitle("Population & Total Property Crime")
```




```
df.crime_data %>%
  ggplot(aes(population, `larceny-theft`)) +
  geom_point() +
  geom_smooth() +
  xlab("Population") +
  ylab("Larceny-Theft") +
  ggtitle("Population & Larceny-Theft")
```




```
df.crime_data %>%
  ggplot(aes(population, burglary)) +
  geom_point() +
  geom_smooth() +
  xlab("Population") +
  ylab("Burglary") +
  ggtitle("Population & Burglary")
```



While the first scatter plot we looked at seemed clear, I'm not confident there is much positive correlation between crime and population outside of motor vehicle theft. We can use the Pearson correlation coefficient to quantify (possible) relationships between the two sets of data. The coefficient value is always in the range of -1 to 1. Values less than zero mean there is a negative correlation; When one value increases, the other decreases. Values above zero indicate a positive correlation; When one value increases, the other does as well. A zero value indicates no correlation at all. Let's look at the correlation values for total violent and property crimes:

```
corr.crimes <- c(
  cor(
    df.crime_data$population,
    df.crime_data$`violent crime total`
  ),

  cor(
    df.crime_data$population,
    df.crime_data$`property crime total`
    )
)

df.correlation_values <- data.frame(
  crime_type = c("Total Violent Crime", "Total Property Crime"),
  correlation_coef = corr.crimes
)

kable(df.correlation_values, caption = "Total Crime Correlations")
```

Table 1: Total Crime Correlations

crime_type	correlation_coef
Total Violent Crime	0.2959880
Total Property Crime	0.2520579

There is a weak positive correlation between total violent crime and population. An even weaker correlation exists for property crime. Given more recent data and the continuing downturn in crime, it's possible they have become weaker or even turn negative. I wouldn't be confident making the blanket statement, "crime goes up along with population in Spokane". Now let's look at the same correlation values for individual types of crime like robbery and murder:

```
corr.specific_crimes <- c(
  cor(
    df.crime_data$population,
    df.crime_data$`motor vehicle theft`
    ),

  cor(
    df.crime_data$population,
    df.crime_data$`larceny-theft`
    ),

  cor(
    df.crime_data$population,
    df.crime_data$`murder and nonnegligent manslaughter`
    ),

  cor(
    df.crime_data$population,
    df.crime_data$`aggravated assault`
    ),

  cor(
    df.crime_data$population,
    df.crime_data$`burglary`
    ),

  cor(
    df.crime_data$population,
    df.crime_data$`robbery`
    )
)

df.correlation_values <- data.frame(
  crime_type = c(
    "Motor Vehicle Theft",
    "Larceny-Theft",
    "Murder & Non-Negl. Manslaughter",
    "Aggravated Assault",
    "Burglary",
    "Robbery"
    ),
```

```
  correlation_coef = corr.specific_crimes
)

kable(
  df.correlation_values[order(-df.correlation_values$correlation_coef),],
  caption = "Crime-Type Correlations"
  )
```

Table 2: Crime-Type Correlations

	crime_type	correlation_coef
1	Motor Vehicle Theft	0.8206645
6	Robbery	0.3820753
2	Larceny-Theft	0.2500194
4	Aggravated Assault	0.1753144
3	Murder & Non-Negl. Manslaughter	0.0112465
5	Burglary	-0.5400008

There is a strong positive correlation between motor vehicle theft and population. We could argue that more cars are being stolen while the population increased, but it's important to remember that *correlation is not causation*. More research is needed before we could say definitively that more cars are being stolen *because of* population growth. A positive correlation exists for robbery as well, though it's less than half as strong. The coefficient for murder and non-negligent manslaughter is almost zero - we can't really say, "more people are being murdered as the population grows". Neither can we say, "less people are being murdered as the population grows", because it's pretty much static.

One surprising result is the negative correlation between population and burglaries. As the population grew, burglaries went down. Again, we can't say the population change was the *cause* of reduction in burglaries. However, we can be comfortable saying, "burglaries went down as the population increased".

Summary

Is there a correlation between population growth and crime? **Yes**. Can we argue that crime in-general goes up along with the population? **No**. While motor vehicle theft and robbery increased along with the population, murders stayed the same, while burglaries went down.