Oregon Open Crime Data Analysis

Russell_Corey

October, 19 2022

Intro

We will explore crime reports in Oregon. We will look at the number of reported offenses as aggregated by both month and county along with a close look at violent offenses in particular. Rates of offense aggregated along counties will be normalized using 2020 Census data. Source code will be posted on Github

Data

The two data sources used are the "Uniform Crime Reporting" Offenses data and the Oregon 2020 Census data from Portland University.

Crime Data The crime reporting data set is accissable directly from

https://www.oregon.gov/osp/Pages/Uniform-Crime-Reporting-Data.aspx

Population Data The Oregon census data can be found at the website

https://www.pdx.edu/population-research/census-data-oregon

with the direct link being

https://drive.google.com/uc?export=download&id=1JrrmYiQUBPux8nnJ88epAAk9U5rbRDBD

Data Exploration and Validation

Overview of number of records collected per month and the total number of distinct offenses for the data set.

Table 1: Offenses Reports

Agency Name	County	IncidentDate	NIBRS Crime Description	NIBRS Report Title	Distinct Offenses
Washington SO	Washington	9/30/2022	Simple Assault	Simple Assault	4
Washington SO	Washington	9/30/2022	All Other Larceny	Larceny/Theft Offenses	1
Washington SO	Washington	9/30/2022	Robbery	Robbery	1
Washington SO	Washington	9/30/2022	Destruction/Damage/Vandalisim of Property	Vandalism	1
Washington SO	Washington	9/30/2022	All Other Offenses	All Other Offenses	13
Washington SO	Washington	9/30/2022	Driving Under the Influence	Driving Under the Influence	3

Each record has a count of distinct offenses. We can compare the number of reports per month and the total number of distinct offenses.

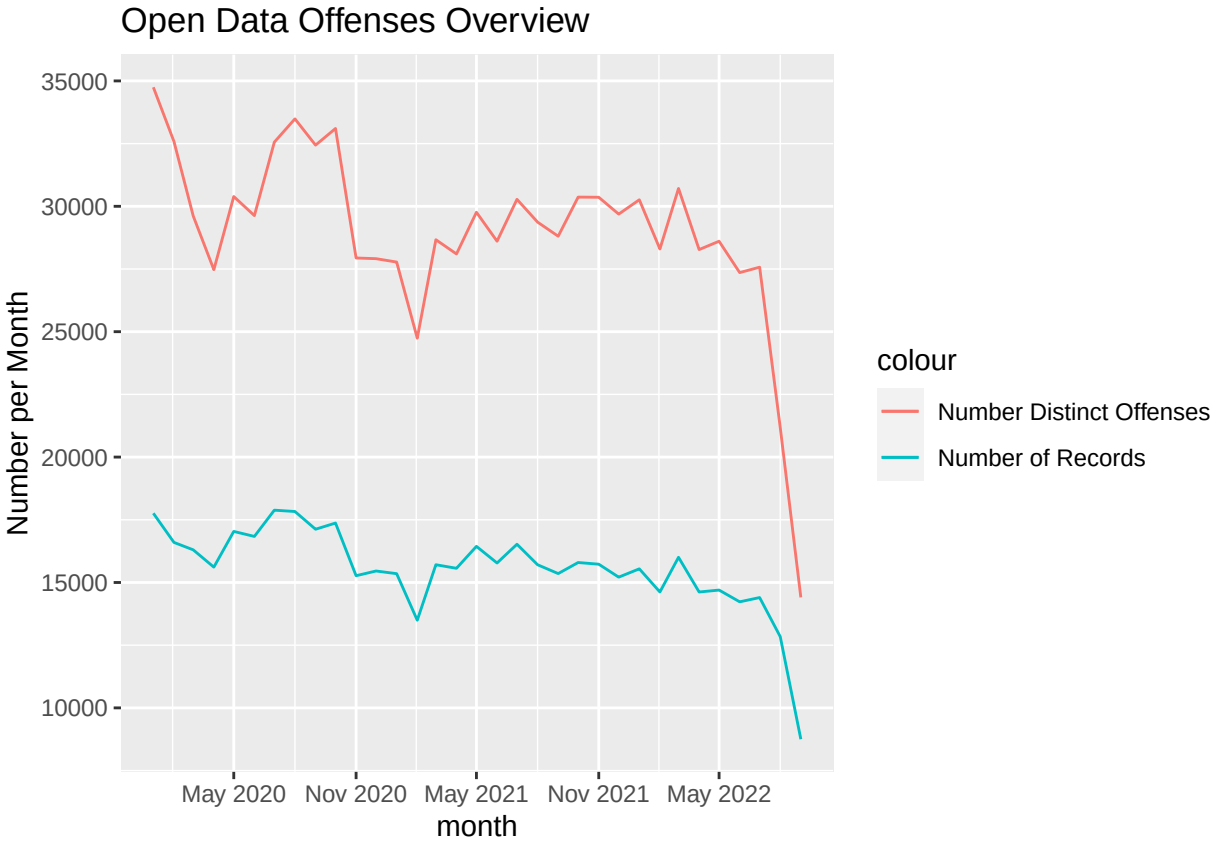


Table 2: Recent Offenses

	month	Total Records	Total Distinct Offenses
28	2022-04-01	14620	28274
29	2022-05-01	14698	28607
30	2022-06-01	14229	27357
31	2022-07-01	14399	27572
32	2022-08-01	12840	21175
33	2022-09-01	8750	14405

There is a shape decline in the last month. We will ignore that month, assuming that it is due to a lag in reporting and assimilating.

Outlier Overview

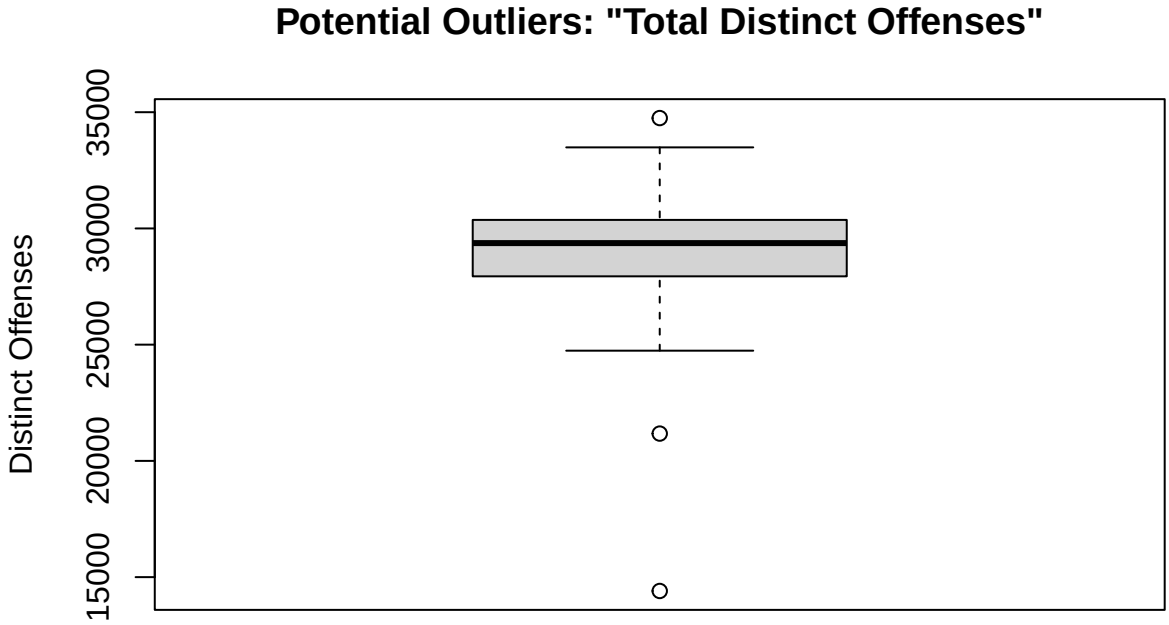


Table 3: Potential Outliers: “Total Distinct Offenses”

	month	Total Records	Total Distinct Offenses	z_score
1	2020-01-01	17760	34747	1.59
32	2022-08-01	12840	21175	-2.13
33	2022-09-01	8750	14405	-3.98

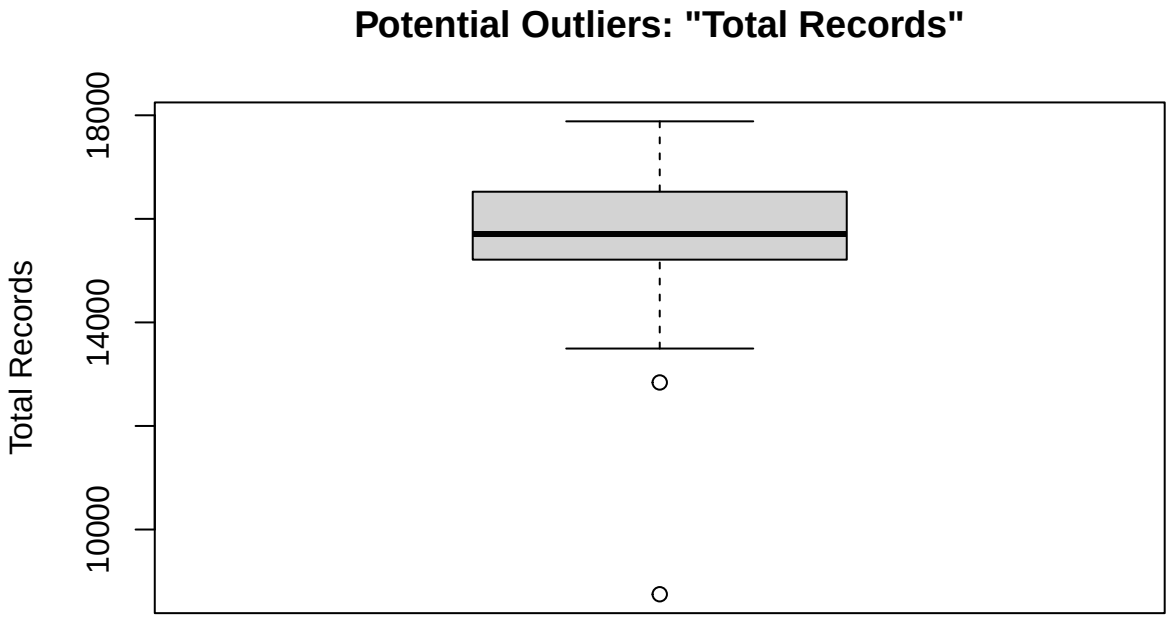


Table 4: Potential Outliers: “Total Records”

	month	Total Records	Total Distinct Offenses	z_score
32	2022-08-01	12840	21175	-1.6
33	2022-09-01	8750	14405	-4.0

We will then be excluding month 33, ‘2022-09-01’.

Population

For per ca-pita calculations we need to know the population broken down by counties. This data is in the Census data set. The data gives us the population in 2010 and 2020 along with the percentage change.

Table 5: County Population

county	2010	2020	percent_change	pop_rate_year
baker	16134	16668	0.0330978	0.0033098
clackamas	375992	421401	0.1207712	0.0120771
lane	351715	382971	0.0888674	0.0088867
multnomah	735334	815428	0.1089219	0.0108922
washington	529710	600372	0.1333975	0.0133398

Violent Offenses

Violent offenses are of particular concern so let’s look at those offenses. To determine which offenses are violent we use the FBI’s definition of violent crimes to categorize the ‘NIBRS Report Title’ column in the data set.

Table 6: Violent Offense Titles

x
Aggravated Assault
Forcible Rape
Forcible Sodomy
Robbery
Willful Murder

We can look at the break down of these violent offenses for the whole state.

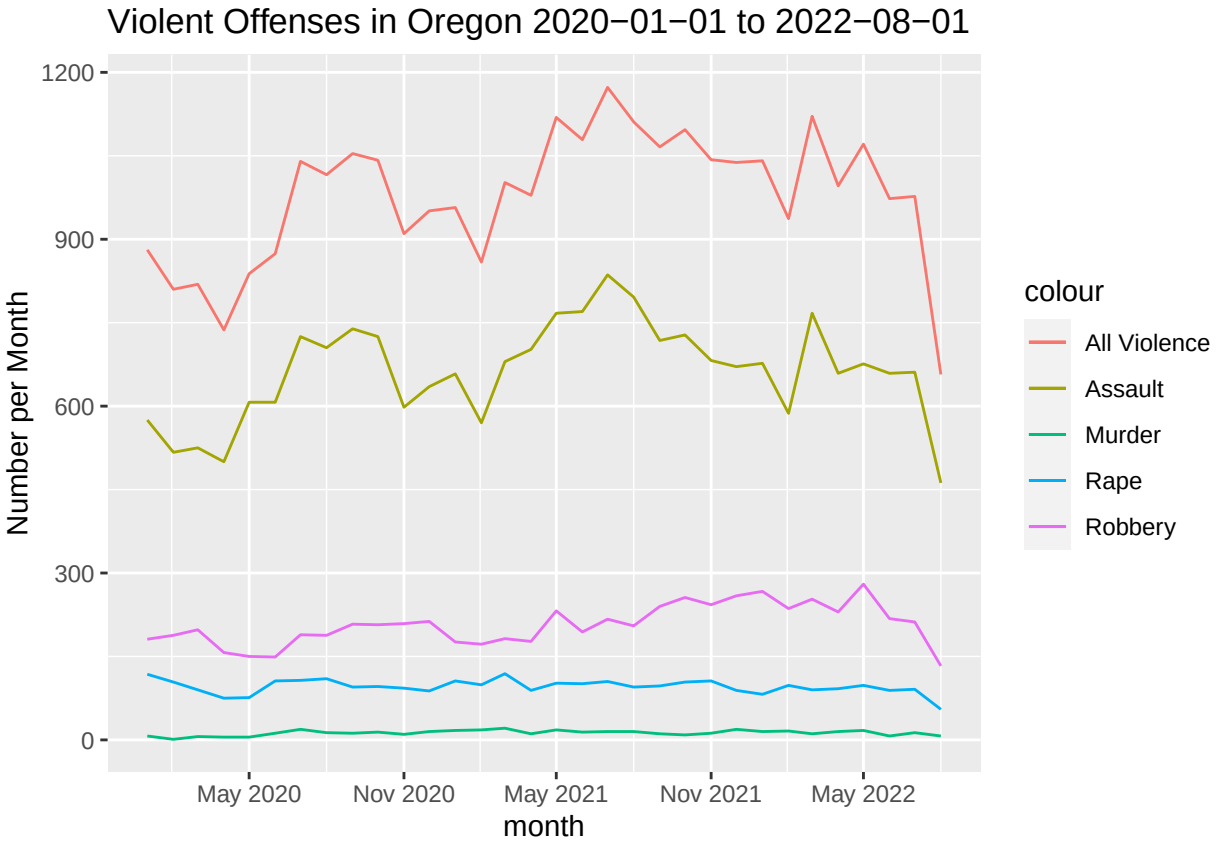
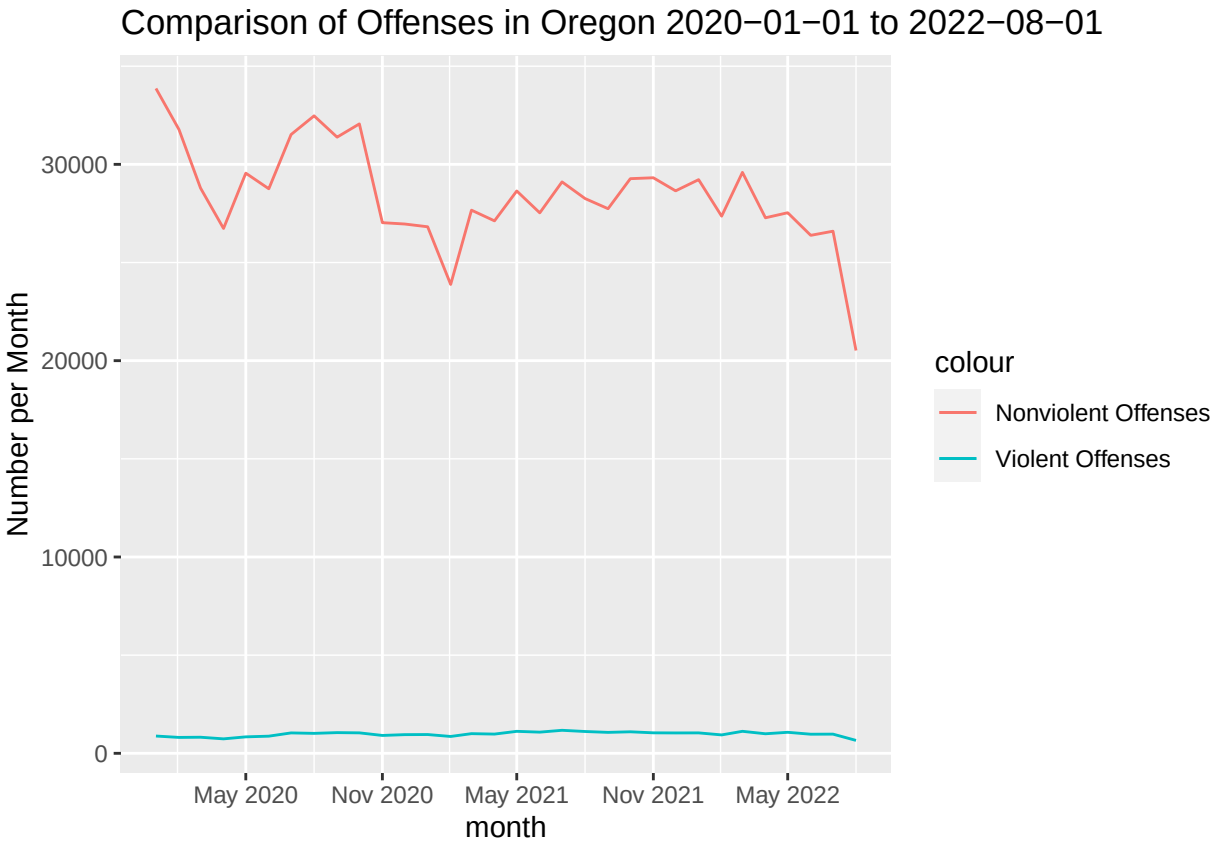


Table 7: Monthly Violent Offenses 2020-01-01 to 2022-08-01

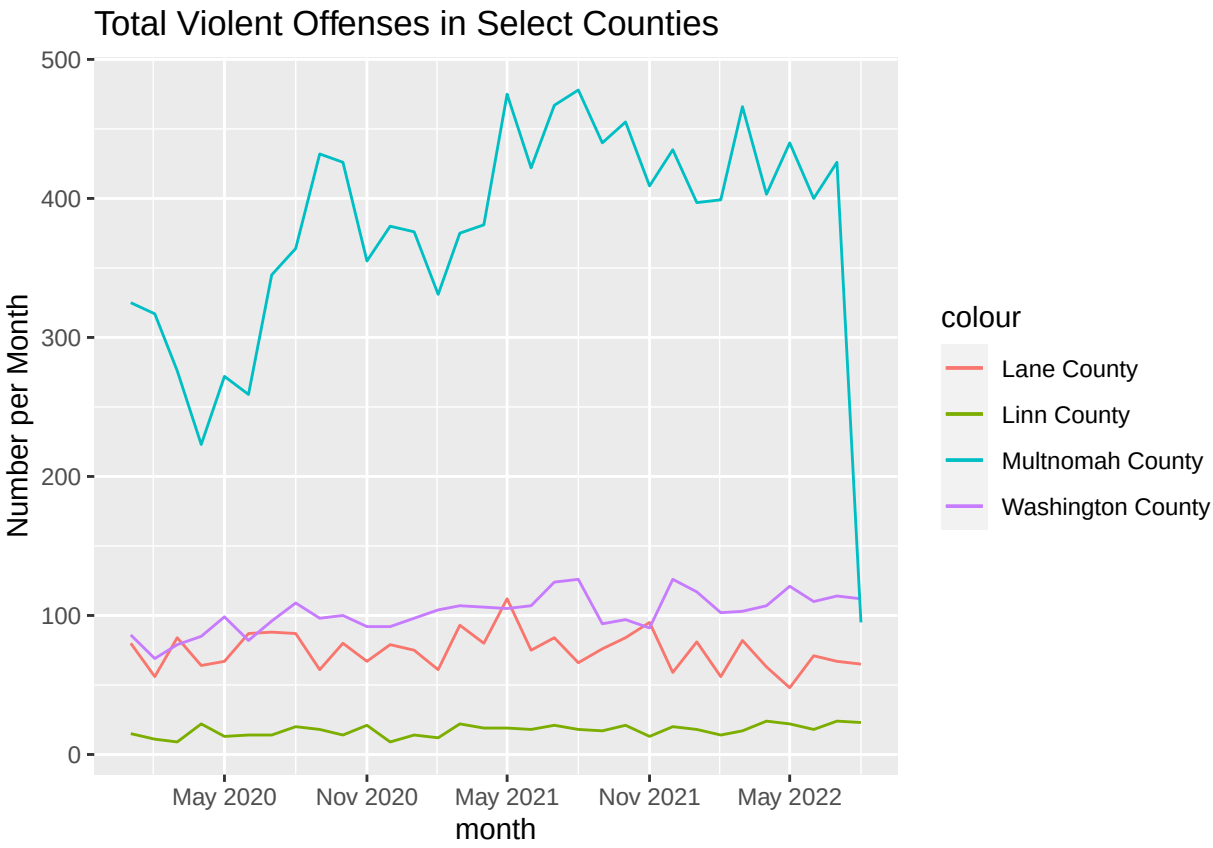
	month	Aggravated Assault	Forcible Rape	Robbery	Willful Murder	All Violent Offenses	Nonviolent Offenses
27	2022-03-01	767	90	253	11	1121	29591
28	2022-04-01	659	92	230	15	996	27278
29	2022-05-01	676	98	280	17	1071	27536
30	2022-06-01	659	89	218	7	973	26384
31	2022-07-01	661	91	212	13	977	26595

	month	Aggravated Assault	Forcible Rape	Robbery	Willful Murder	All Violent Offenses	Nonviolent Offenses
32	2022-08-01	462	55	133	7	657	20518

We can also compare the number of violent offenses verses non-violent offenses for the state over the study period.



Next we can look at the total number of distinct violent offenses by county.



Recall from the population table that these counties vary greatly by population. So, using the population data we can plot the above graph adjusted to the county's population.

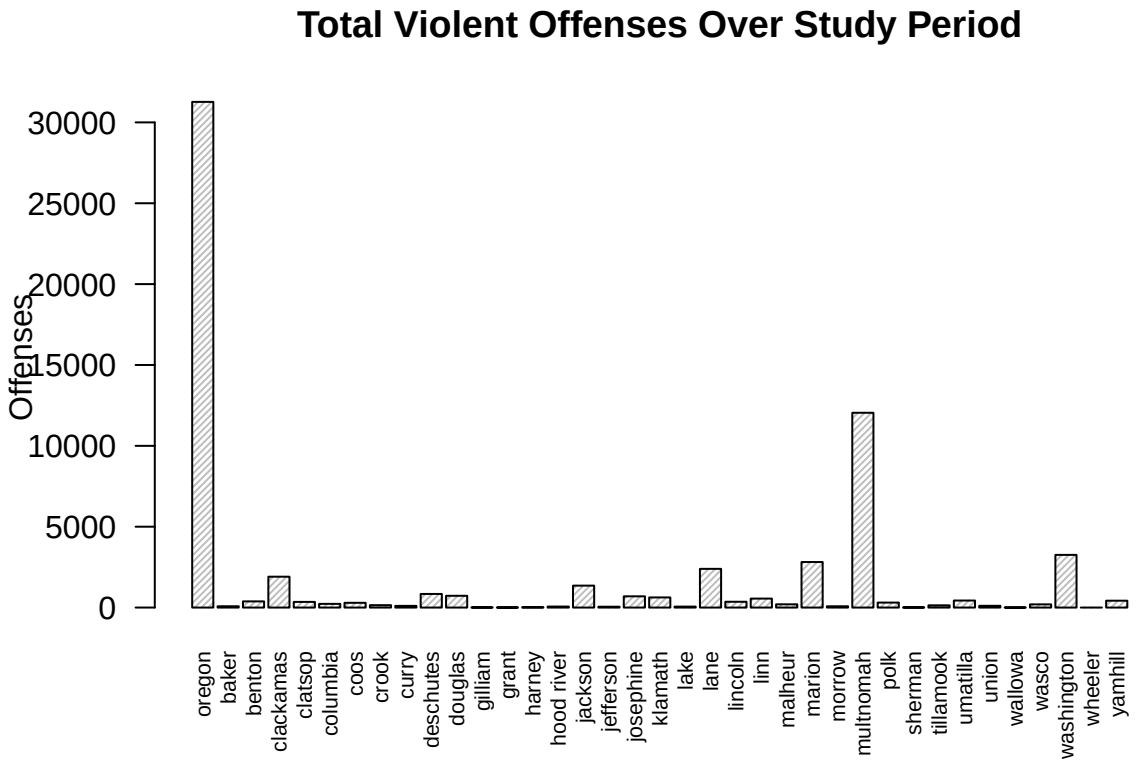
Violent Offenses by County

We can look at the aggregates of violent offenses by county.

Table 8: County Total Violent Offenses

county	offenses	vo_percap
oregon	31268	0.74
baker	79	0.47
benton	382	0.40
clackamas	1908	0.45
clatsop	350	0.85
columbia	228	0.43
coos	294	0.45
crook	148	0.60
curry	105	0.45
deschutes	841	0.42
douglas	728	0.65
gilliam	10	0.50
grant	6	0.08
harney	20	0.27
hood river	61	0.25
jackson	1356	0.61

county	offenses	vo_percap
jefferson	52	0.21
josephine	695	0.79
klamath	626	0.90
lake	58	0.71
lane	2393	0.62
lincoln	359	0.71
linn	554	0.43
malheur	200	0.63
marion	2815	0.81
morrow	78	0.64
multnomah	12044	1.48
polk	307	0.35
sherman	11	0.59
tillamook	138	0.50
umatilla	434	0.54
union	107	0.41
wallowa	2	0.03
wasco	197	0.74
washington	3258	0.54
wheeler	0	0.00
yamhill	424	0.39



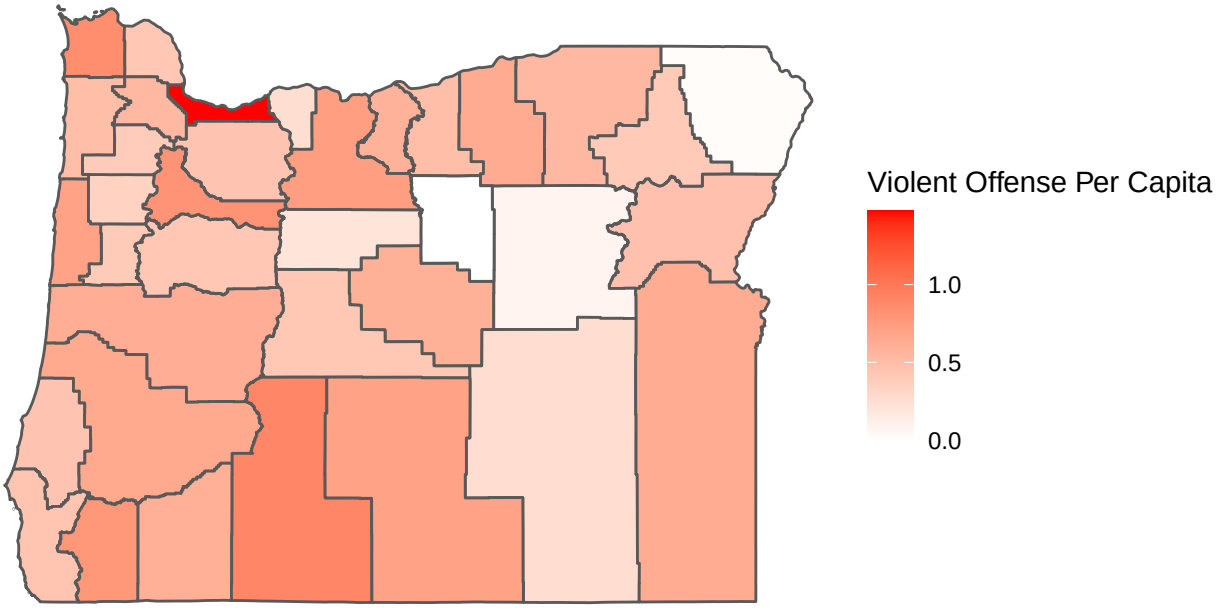
From the table it appears that the rate of violent offenses is far higher in Multnomah county than others.

Table 9: Counties with Highest Per Capita Violent Offenses

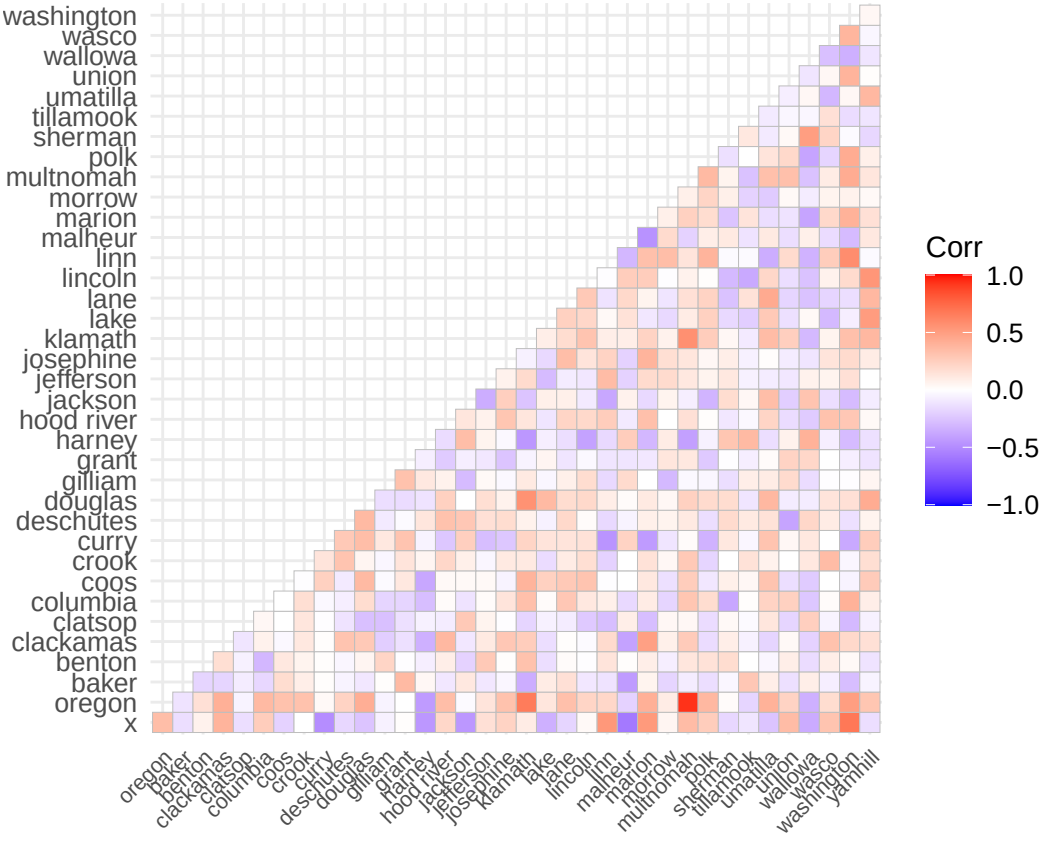
	county	offenses	2020	vo_percap
26	multnomah	12044	815428	1.48
18	klamath	626	69413	0.90
4	clatsop	350	41072	0.85
24	marion	2815	345920	0.81
17	josephine	695	88090	0.79
34	wasco	197	26670	0.74
27	oregon	31268	4237256	0.74
21	lincoln	359	50395	0.71
19	lake	58	8160	0.71
10	douglas	728	111201	0.65
25	morrow	78	12186	0.64
23	malheur	200	31571	0.63
20	lane	2393	382971	0.62
15	jackson	1356	223259	0.61
7	crook	148	24738	0.60
29	sherman	11	1870	0.59
35	washington	3258	600372	0.54
31	umatilla	434	80075	0.54
30	tillamook	138	27390	0.50
11	gilliam	10	1995	0.50
1	baker	79	16668	0.47
6	coos	294	64929	0.45
3	clackamas	1908	421401	0.45
8	curry	105	23446	0.45
5	columbia	228	52589	0.43
22	linn	554	128610	0.43
9	deschutes	841	198253	0.42
32	union	107	26196	0.41
2	benton	382	95184	0.40
37	yamhill	424	107722	0.39
28	polk	307	87433	0.35
13	harney	20	7495	0.27
14	hood river	61	23977	0.25
16	jefferson	52	24502	0.21
12	grant	6	7233	0.08
33	wallowa	2	7391	0.03
36	wheeler	0	1451	0.00

```
## Linking to GEOS 3.10.2, GDAL 3.4.1, PROJ 8.2.1; sf_use_s2() is TRUE
```

Per Capita Violent Offenses



Corrlation



Total monthly violent offenses by county

Table 10: Counties with positiv correlation >0.32

	pearson_cor
x	1.00
oregon	0.34
clackamas	0.39
linn	0.53
marion	0.52
multnomah	0.35
union	0.35
washington	0.68

Table 11: Counties with negative correlation < -0.32

	pearson_cor
curry	-0.49
harney	-0.45
jackson	-0.45
lake	-0.34
malheur	-0.58
wallowa	-0.36

Linear Regression

```
## `geom_smooth()` using formula 'y ~ x'
```

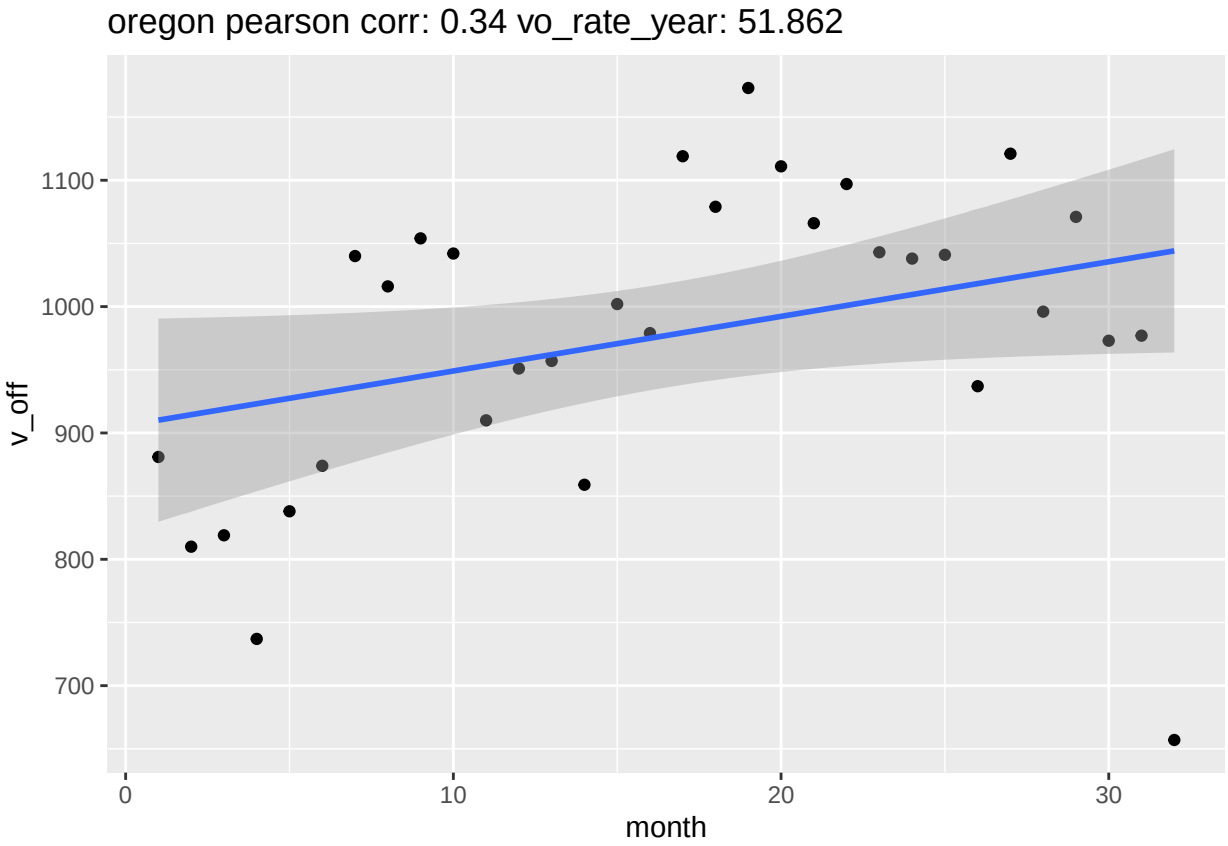


Table 12: pearson correlation and linear reg rate

	county	pearson_cor	2020	vo_rate_year	pop_rate_year	adjusted_rate
37	x	1.00	NA	NA	NA	NA
35	washington	0.68	600372	11.86	0.01	11.85
22	linn	0.53	128610	2.86	0.01	2.85
24	marion	0.52	345920	8.62	0.01	8.61
3	clackamas	0.39	421401	5.53	0.01	5.52
26	multnomah	0.35	815428	37.05	0.01	37.04

Table 13: pearson correlation and linear reg rate

	county	pearson_cor	2020	vo_rate_year	pop_rate_year	adjusted_rate
33	wallowa	-0.36	7391	-0.11	0.01	-0.12
13	harney	-0.45	7495	-0.52	0.00	-0.52
15	jackson	-0.45	223259	-5.19	0.01	-5.20
8	curry	-0.49	23446	-1.69	0.00	-1.69
23	malheur	-0.58	31571	-3.02	0.00	-3.02
36	wheeler	NA	1451	0.00	0.00	0.00

Sources

Oregon Crime Data "Oregon State Police: Uniform Crime Reporting"

https://www.oregon.gov/osp/Pages/Uniform-Crime-Reporting-Data.aspx

Oregon Census Data "Census Data for Oregon" Population Research Center

https://www.pdx.edu/population-research/census-data-oregon

Direct link for data set:

https://drive.google.com/uc?export=download&id=1JrrmYiQUBPux8nnJ88epAAk9U5rbRDBD

FBI "FBI-Violent Crime" FBI: UCR

https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/violent-crime

Oregon Spatial Data Library "Oregon Counties - 2015"

https://spatialdata.oregonexplorer.info/geoportal/details;id=361c06fee9de4e24a72e280fb386a771

APPENDIX A

Appendix description: Raw data tables.

A1 Total Violent Offenses Per Capita per County

Total counts for study period:

Table 14: County Total Violent Offenses

county	offenses	vo_percap
oregon	31268	0.7379304
baker	79	0.4739621
benton	382	0.4013280
clackamas	1908	0.4527754
clatsop	350	0.8521621
columbia	228	0.4335507
coos	294	0.4528023
crook	148	0.5982699
curry	105	0.4478376
deschutes	841	0.4242054
douglas	728	0.6546704
gilliam	10	0.5012531
grant	6	0.0829531
harney	20	0.2668446
hood river	61	0.2544105
jackson	1356	0.6073663
jefferson	52	0.2122276
josephine	695	0.7889658
klamath	626	0.9018484
lake	58	0.7107843
lane	2393	0.6248515
lincoln	359	0.7123723
linn	554	0.4307597
malheur	200	0.6334928
marion	2815	0.8137720
morrow	78	0.6400788
multnomah	12044	1.4770158
polk	307	0.3511260
sherman	11	0.5882353
tillamook	138	0.5038335
umatilla	434	0.5419919
union	107	0.4084593
wallowa	2	0.0270599
wasco	197	0.7386577
washington	3258	0.5426635
wheeler	0	0.0000000
yamhill	424	0.3936058

A2 Population

Table 15: County Population 2010 and 2020

county	2010	2020	percent_change
oregon	3831074	4237256	0.1060230
baker	16134	16668	0.0330978
benton	85579	95184	0.1122355
clackamas	375992	421401	0.1207712
clatsop	37039	41072	0.1088852
columbia	49351	52589	0.0656116
coos	63043	64929	0.0299161
crook	20978	24738	0.1792354
curry	22364	23446	0.0483813
deschutes	157733	198253	0.2568898
douglas	107667	111201	0.0328234
gilliam	1871	1995	0.0662747
grant	7445	7233	-0.0284755
harney	7422	7495	0.0098356
hood river	22346	23977	0.0729885
jackson	203206	223259	0.0986831
jefferson	21720	24502	0.1280847
josephine	82713	88090	0.0650079
klamath	66380	69413	0.0456915
lake	7895	8160	0.0335655
lane	351715	382971	0.0888674
lincoln	46034	50395	0.0947343
linn	116672	128610	0.1023210
malheur	31313	31571	0.0082394
marion	315335	345920	0.0969921
morrow	11173	12186	0.0906650
multnomah	735334	815428	0.1089219
polk	75403	87433	0.1595427
sherman	1765	1870	0.0594901
tillamook	25250	27390	0.0847525
umatilla	75889	80075	0.0551595
union	25748	26196	0.0173994
wallowa	7008	7391	0.0546518
wasco	25213	26670	0.0577876
washington	529710	600372	0.1333975
wheeler	1441	1451	0.0069396
yamhill	99193	107722	0.0859839

A4 Sums of offenses by type by date

Table 16: Aggregate Monthly Counts

month	Total Records	Total Distinct Offenses	Part 3	Traffic Crime	Simple Assault	Fraud Offenses	Larceny/Theft Offenses	Motor Vehicle Theft	Robbery	Vandalism	All Other Offenses	Disorderly Conduct
2020-01-01	17760	34747	4263	1447	1704	1690	7636	1299	181	3032	5207	967
2020-02-01	16598	32584	4113	1409	1609	1601	6980	1118	188	2814	4870	945
2020-03-01	16301	29613	3274	1163	1565	1422	6328	1193	198	2897	4354	903
2020-04-01	15612	27472	2518	1128	1473	1498	6202	1307	157	2924	3661	812
2020-05-01	17036	30390	2766	1381	1667	1416	6294	1154	150	3253	4227	1081
2020-06-01	16836	29629	3118	1382	1690	1459	6335	1170	149	3324	3614	1120

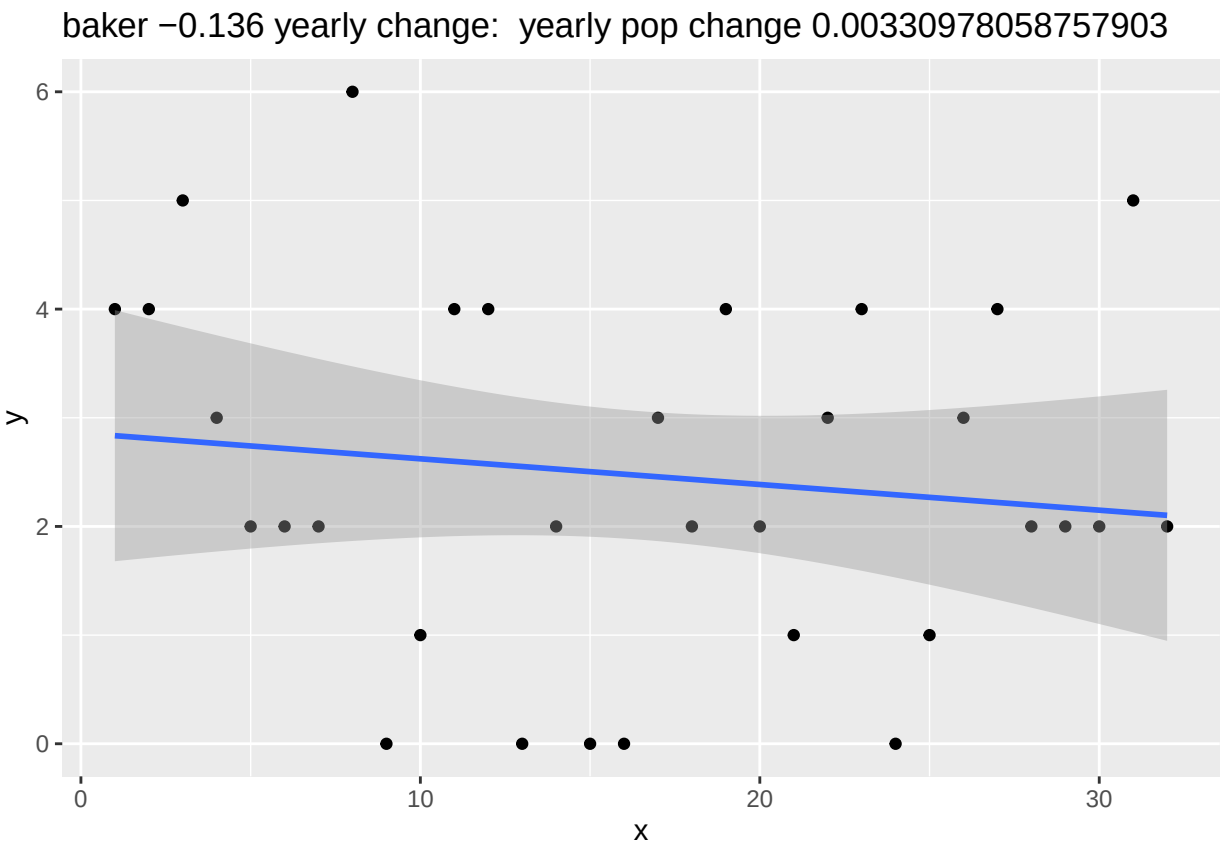
A5 Correlation Matrix of Total Monthly Violent Offenses

Table 17: pearson correlation and linear reg rate

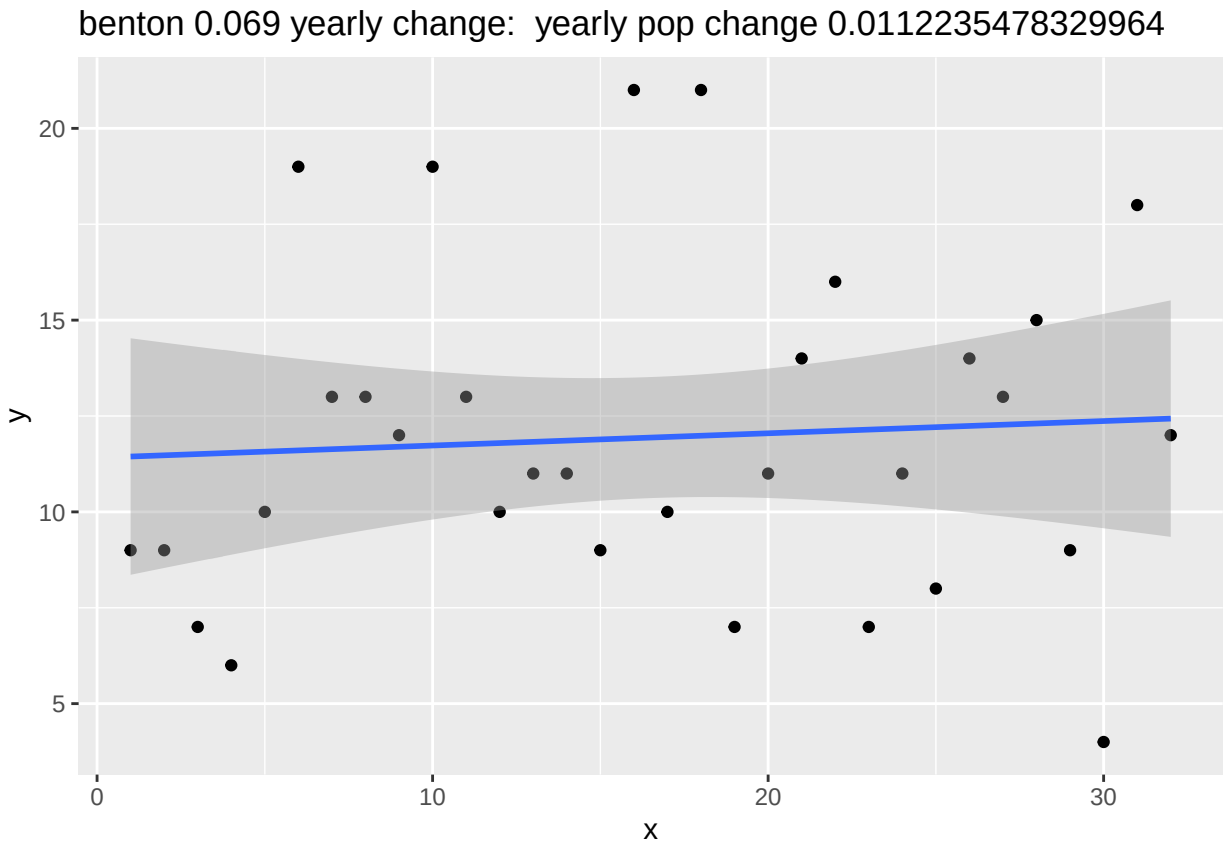
county	pearson_cor	2020	vo_rate_year	pop_rate_year	adjusted_rate
baker	-0.136	16668	-0.284	0.0033098	-0.287
benton	0.069	95184	0.383	0.0112235	0.372
clackamas	0.388	421401	5.534	0.0120771	5.522
clatsop	-0.141	41072	-0.871	0.0108885	-0.882
columbia	0.263	52589	1.157	0.0065612	1.150
coos	-0.192	64929	-0.638	0.0029916	-0.641
crook	0.000	24738	0.000	0.0179235	-0.018
curry	-0.488	23446	-1.687	0.0048381	-1.692
deschutes	-0.170	198253	-1.304	0.0256890	-1.330
douglas	-0.252	111201	-1.786	0.0032823	-1.789
gilliam	-0.064	1995	-0.044	0.0066275	-0.051
grant	0.009	7233	0.004	-0.0028475	0.007
harney	-0.447	7495	-0.519	0.0009836	-0.520
hood river	0.207	23977	0.438	0.0072988	0.431
jackson	-0.453	223259	-5.186	0.0098683	-5.196
jefferson	0.156	24502	0.282	0.0128085	0.269
josephine	0.232	88090	1.880	0.0065008	1.873
klamath	0.097	69413	0.946	0.0045691	0.941
lake	-0.341	8160	-0.739	0.0033566	-0.742
lane	-0.183	382971	-3.187	0.0088867	-3.196
lincoln	0.026	50395	0.121	0.0094734	0.112
linn	0.528	128610	2.859	0.0102321	2.849
malheur	-0.584	31571	-3.018	0.0008239	-3.019
marion	0.525	345920	8.615	0.0096992	8.605
morrow	0.050	12186	0.097	0.0090665	0.088
multnomah	0.348	815428	37.047	0.0108922	37.036
oregon	0.340	4237256	51.862	0.0106023	51.851
polk	0.265	87433	0.922	0.0159543	0.906
sherman	-0.174	1870	-0.134	0.0059490	-0.140
tillamook	-0.110	27390	-0.286	0.0084752	-0.294
umatilla	-0.252	80075	-1.883	0.0055160	-1.889
union	0.348	26196	0.922	0.0017399	0.920
wallowa	-0.364	7391	-0.114	0.0054652	-0.119
wasco	0.310	26670	1.212	0.0057788	1.206
washington	0.682	600372	11.864	0.0133398	11.851
wheeler	NA	1451	0.000	0.0006940	-0.001
x	1.000	NA	NA	NA	NA
yamhill	-0.136	107722	-0.739	0.0085984	-0.748

A6 Violent offenses per county

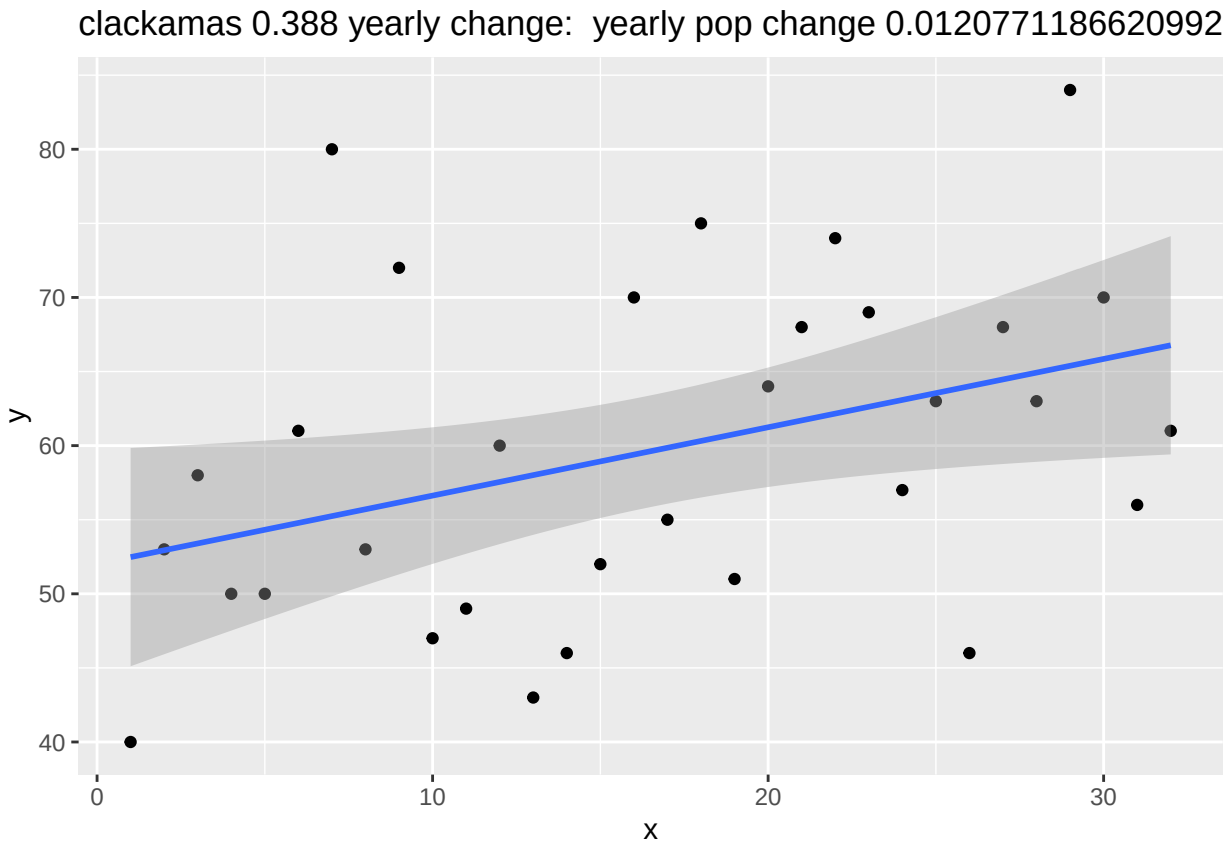
```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(countyname)` instead of `countyname` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
## `geom_smooth()` using formula 'y ~ x'
```



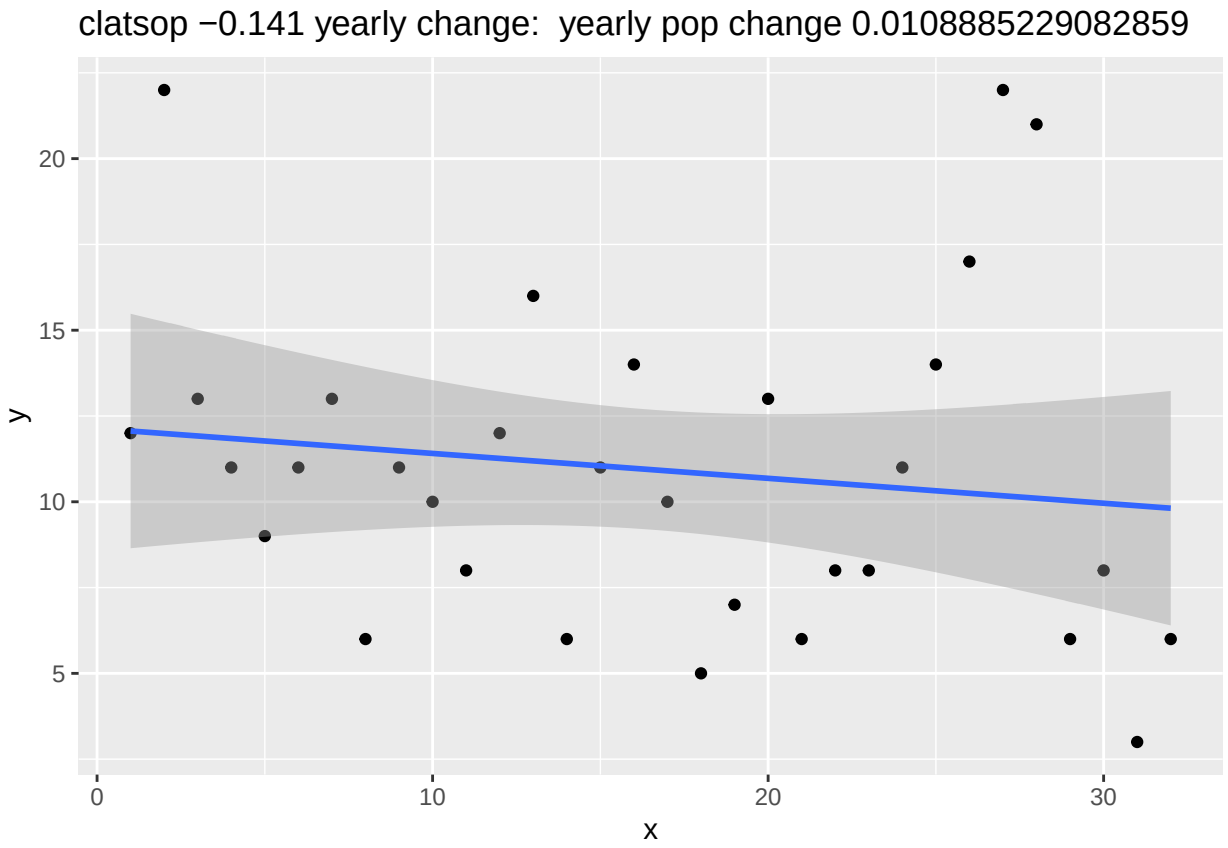

```
## `geom_smooth()` using formula 'y ~ x'
```



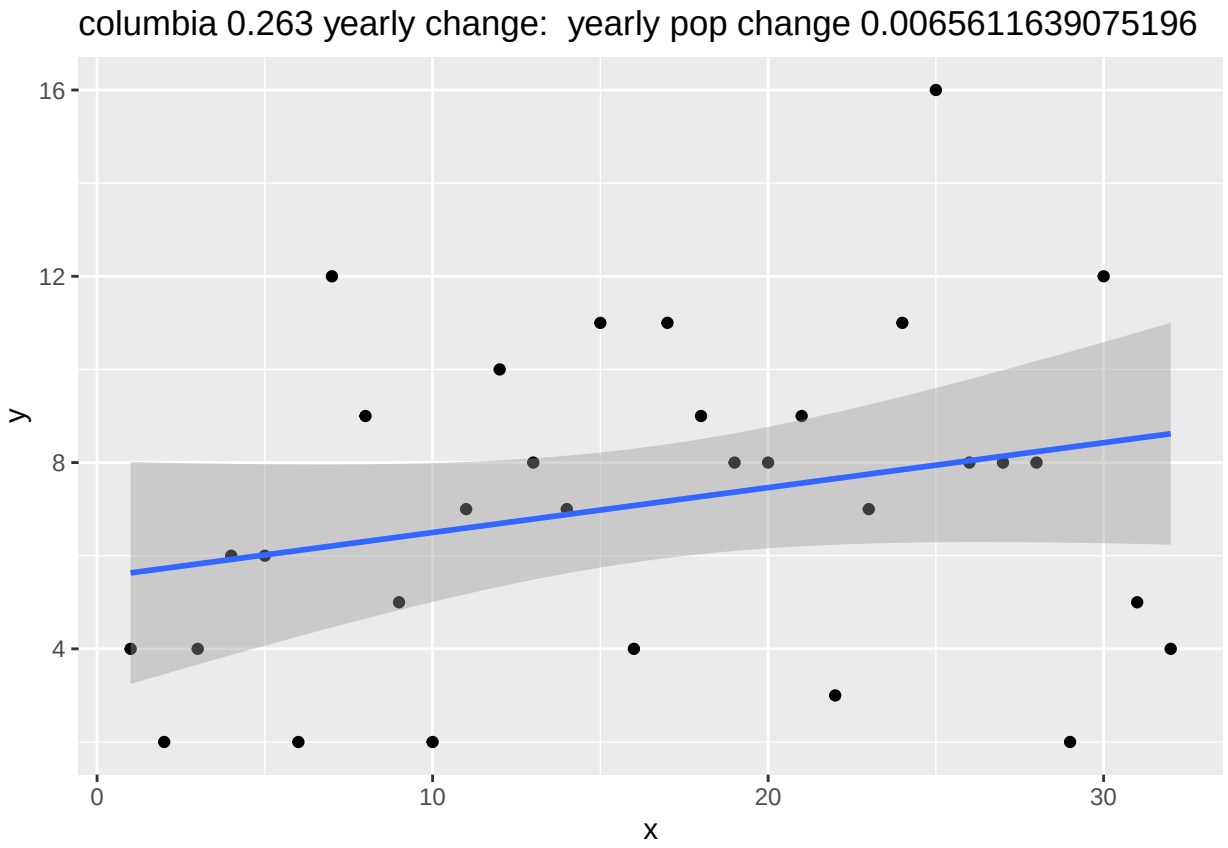

```
## `geom_smooth()` using formula 'y ~ x'
```



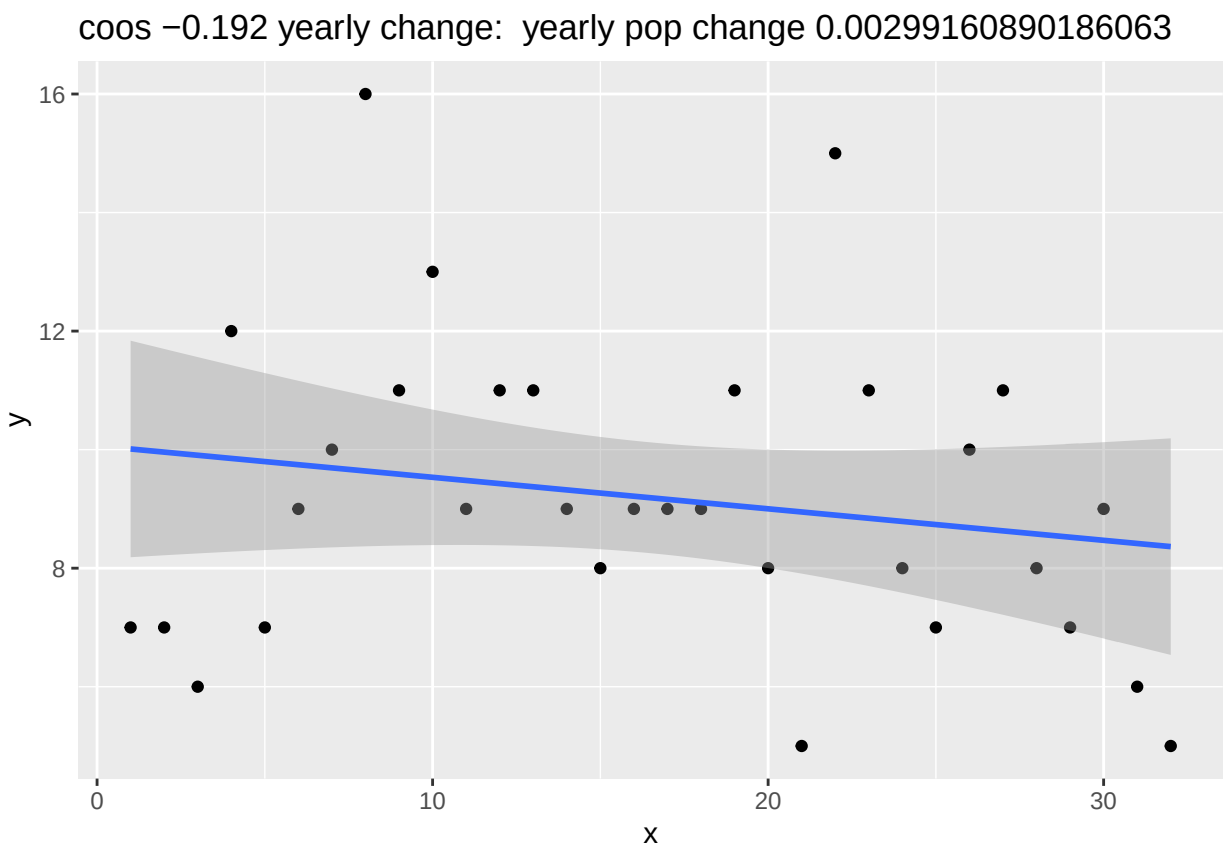

```
## `geom_smooth()` using formula 'y ~ x'
```



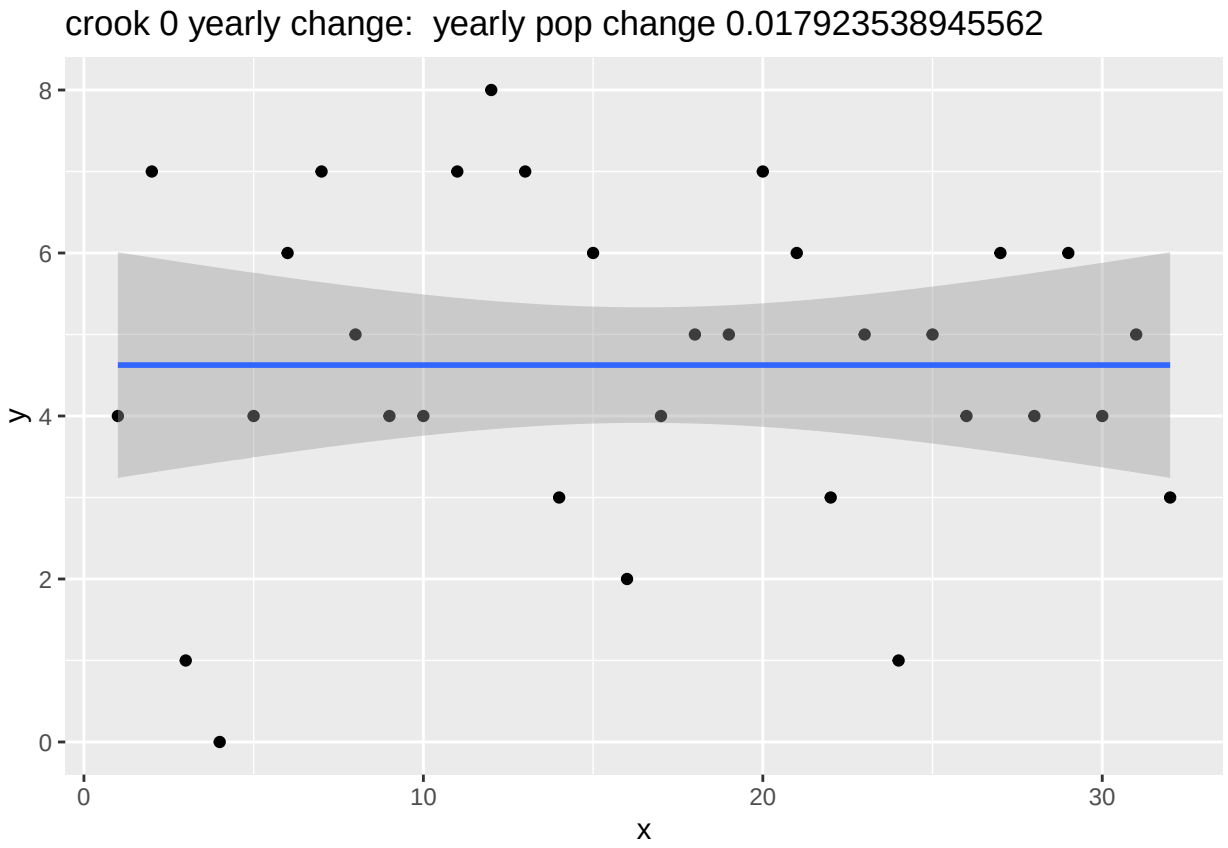

```
## `geom_smooth()` using formula 'y ~ x'
```



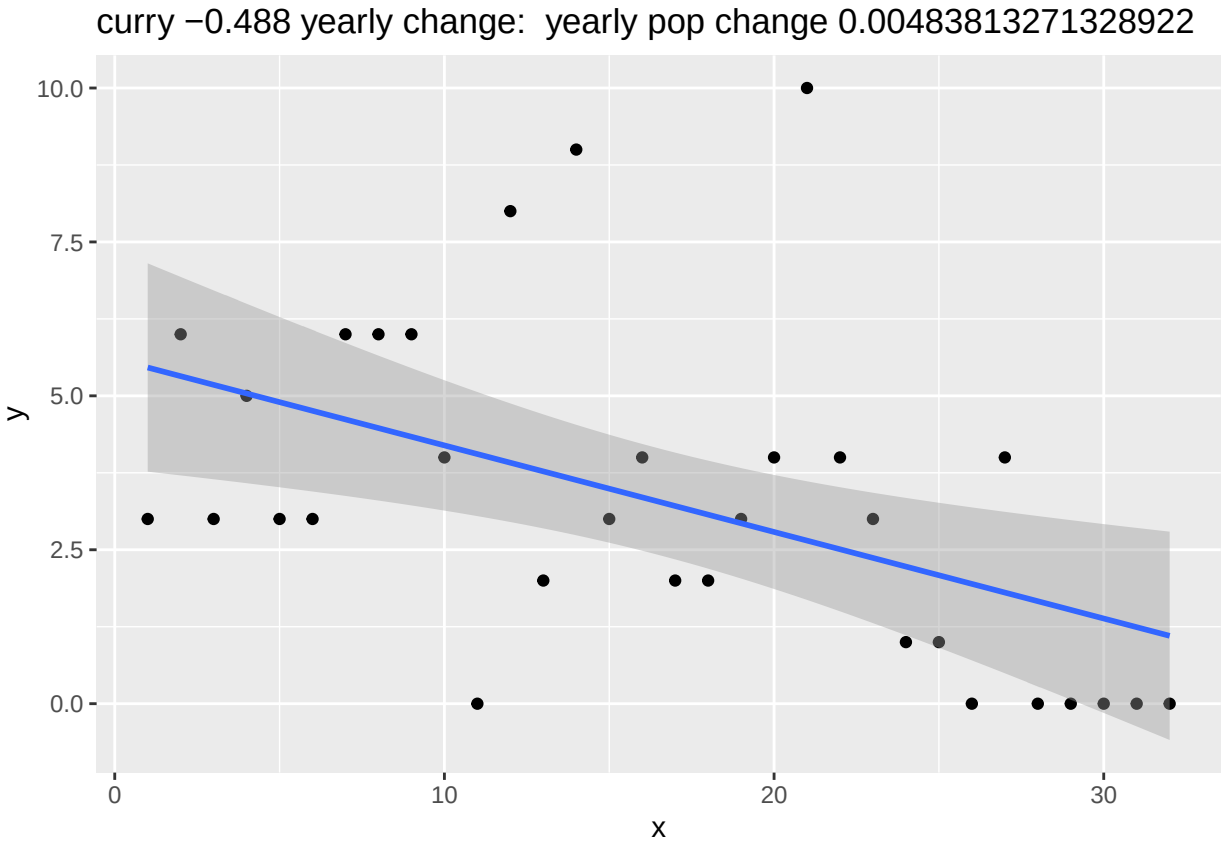

```
## `geom_smooth()` using formula 'y ~ x'
```



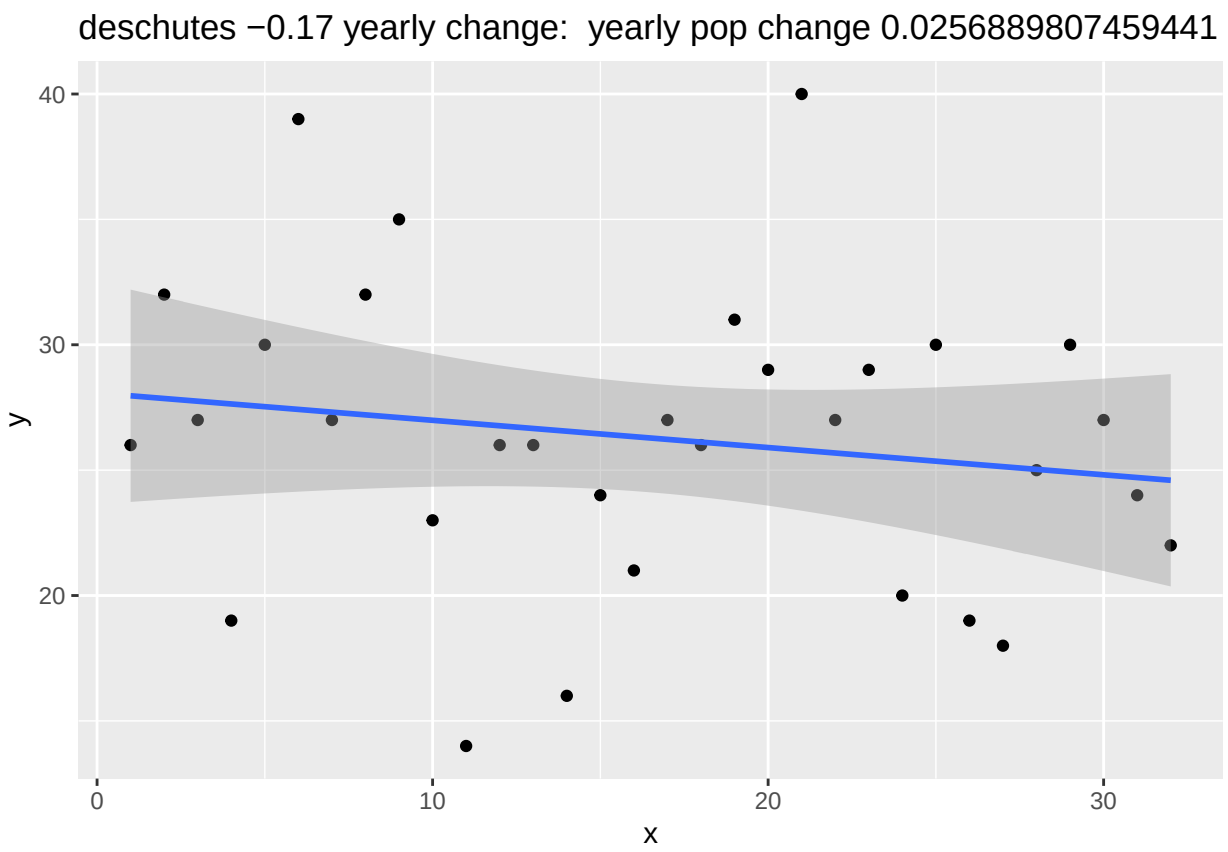

```
## `geom_smooth()` using formula 'y ~ x'
```



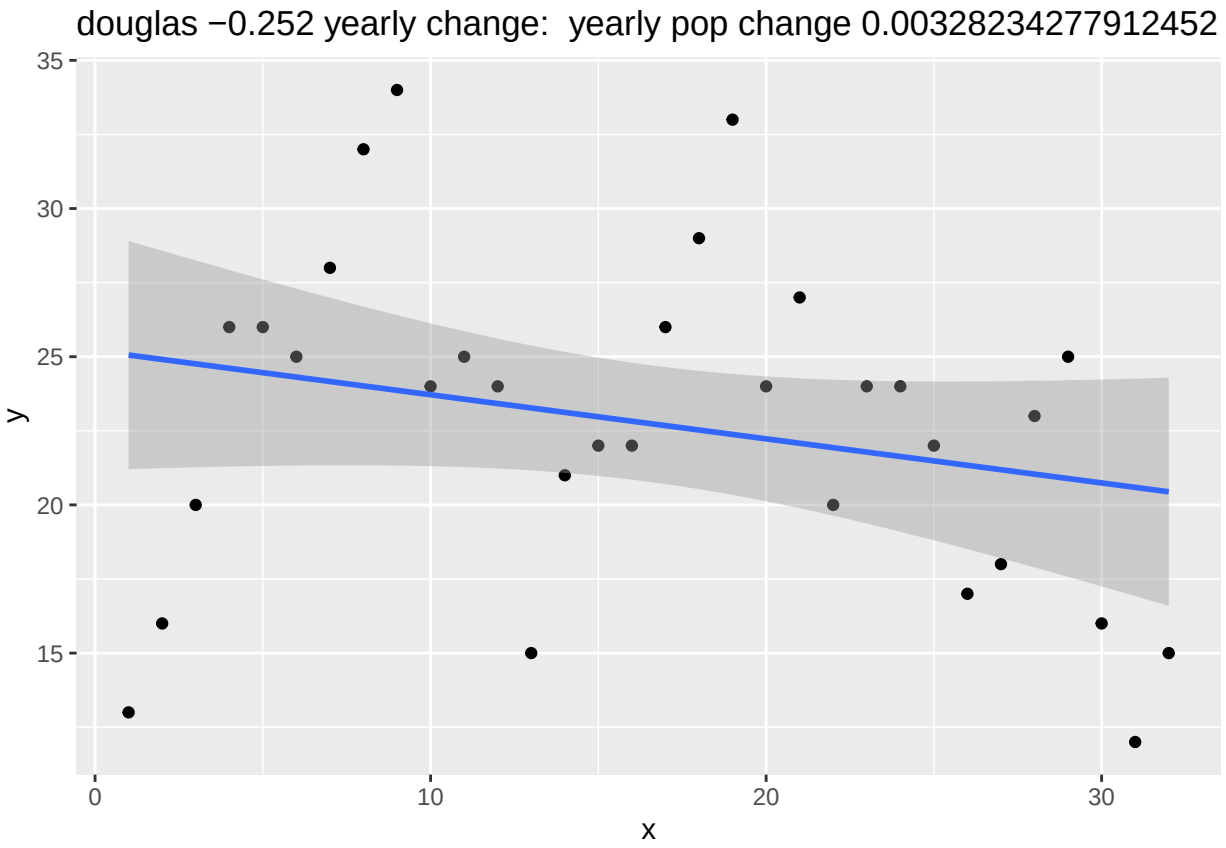

```
## `geom_smooth()` using formula 'y ~ x'
```



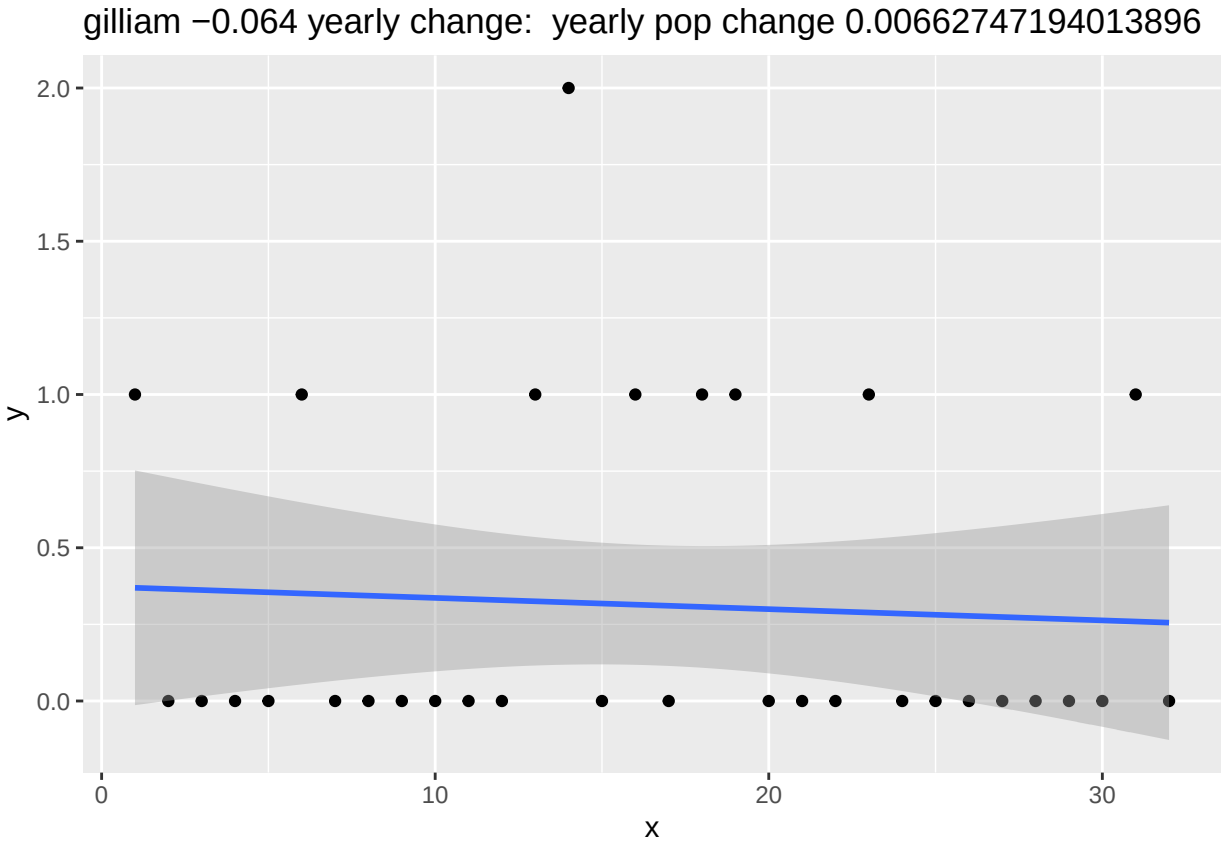

```
## `geom_smooth()` using formula 'y ~ x'
```



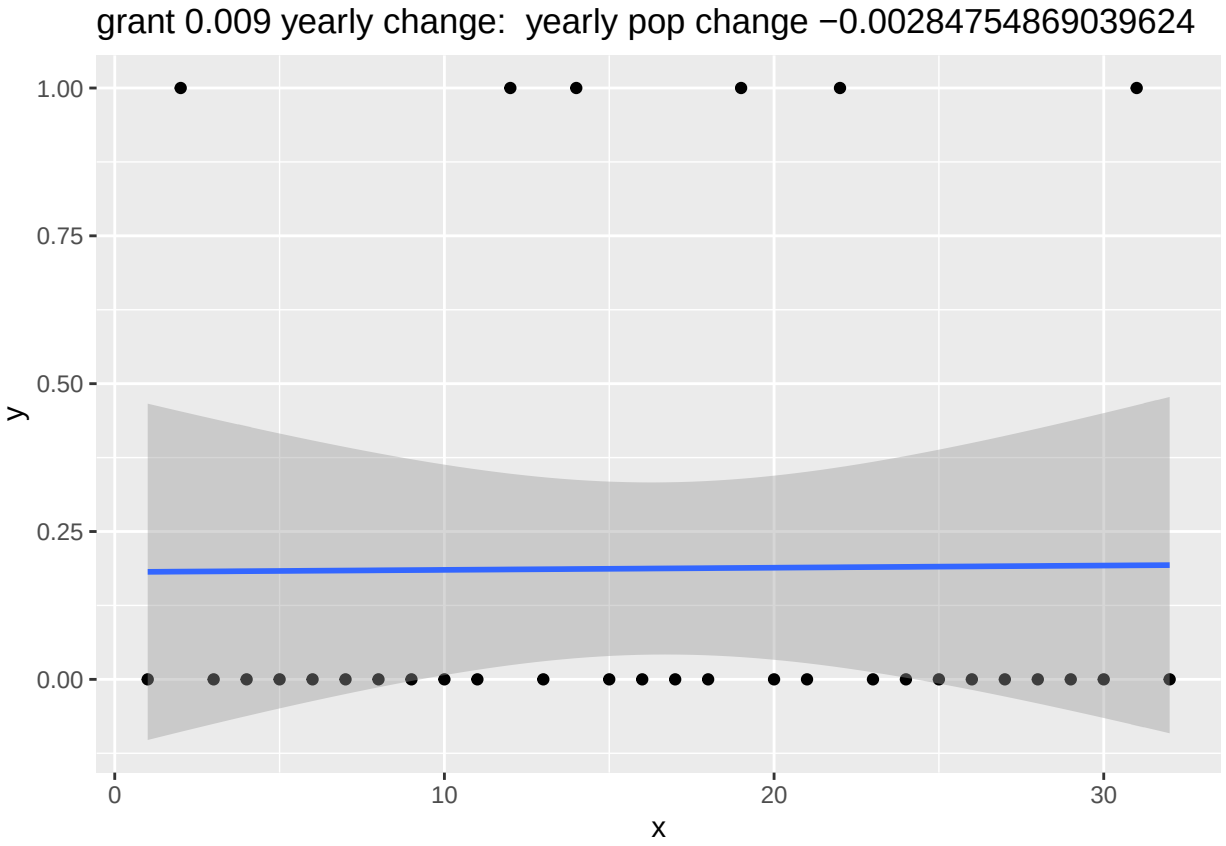

```
## `geom_smooth()` using formula 'y ~ x'
```



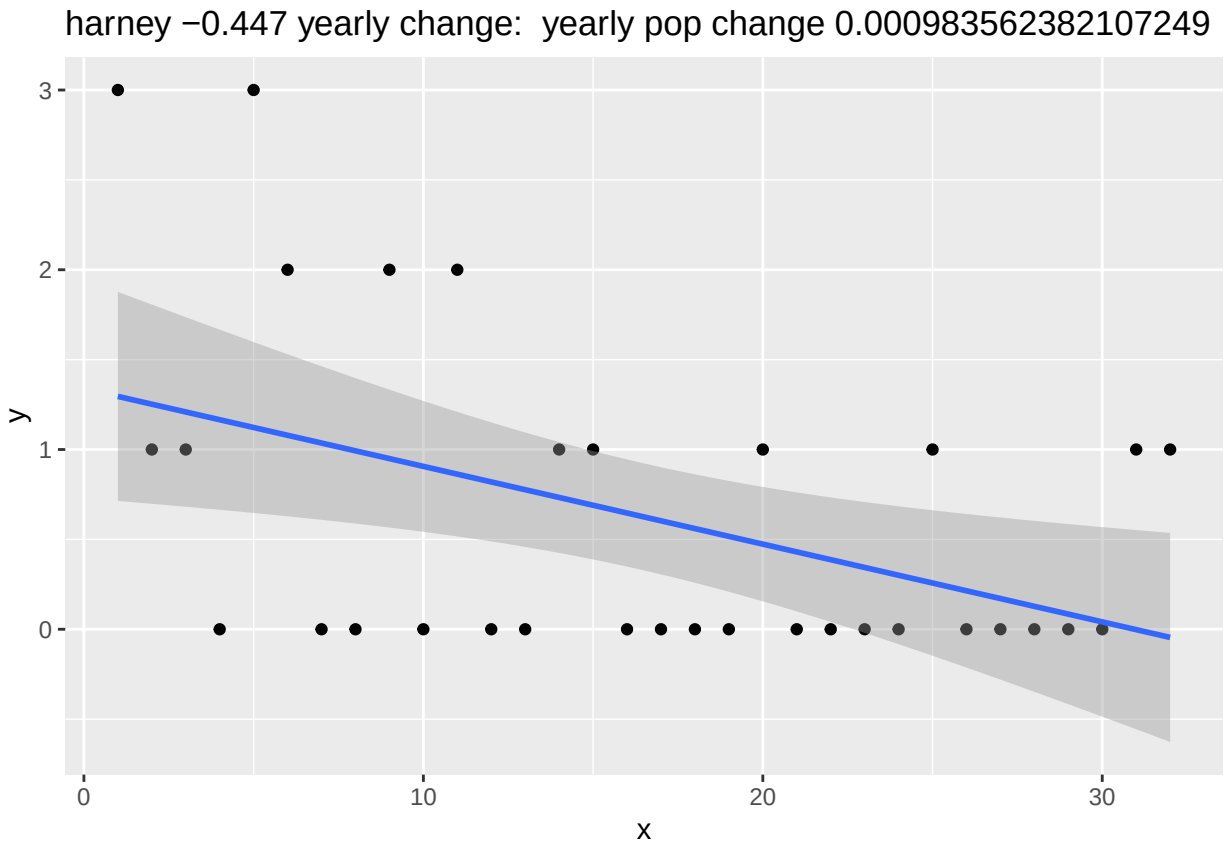

```
## `geom_smooth()` using formula 'y ~ x'
```



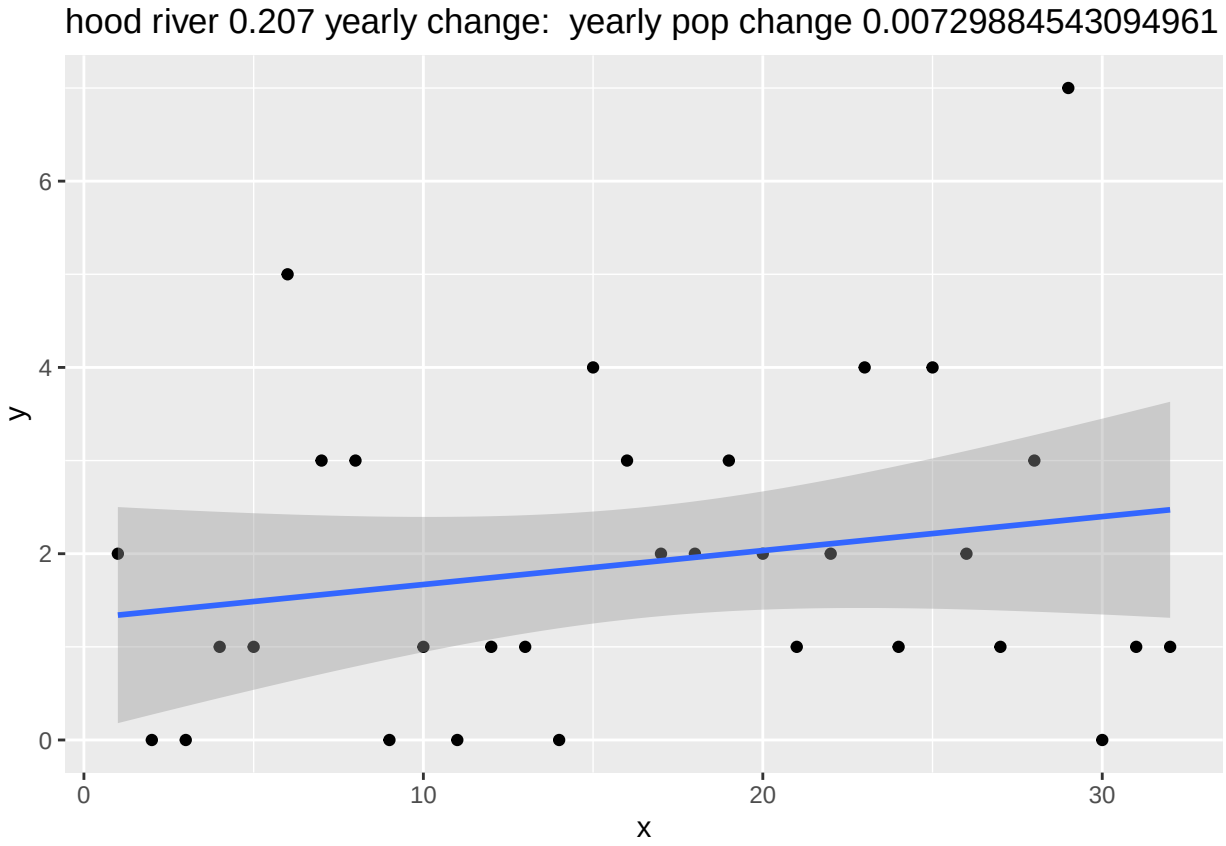

```
## `geom_smooth()` using formula 'y ~ x'
```



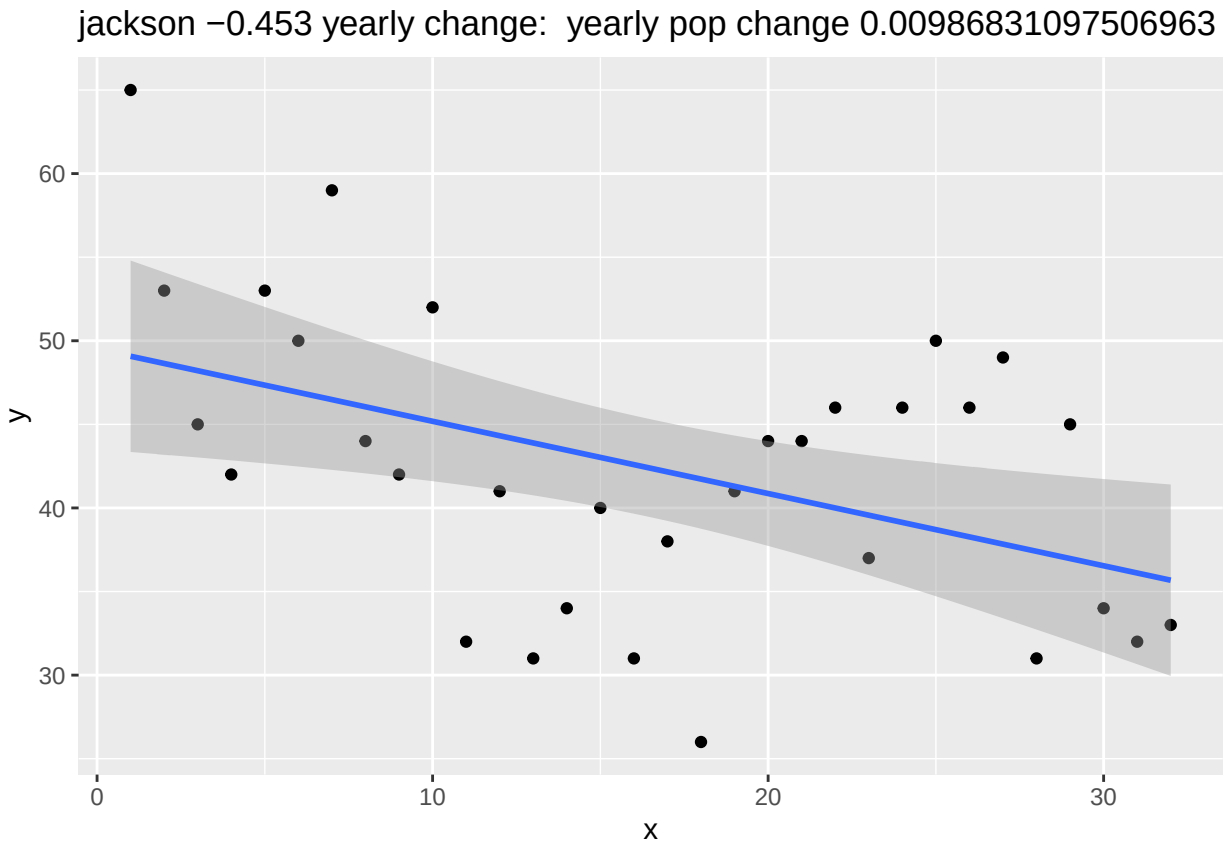

```
## `geom_smooth()` using formula 'y ~ x'
```



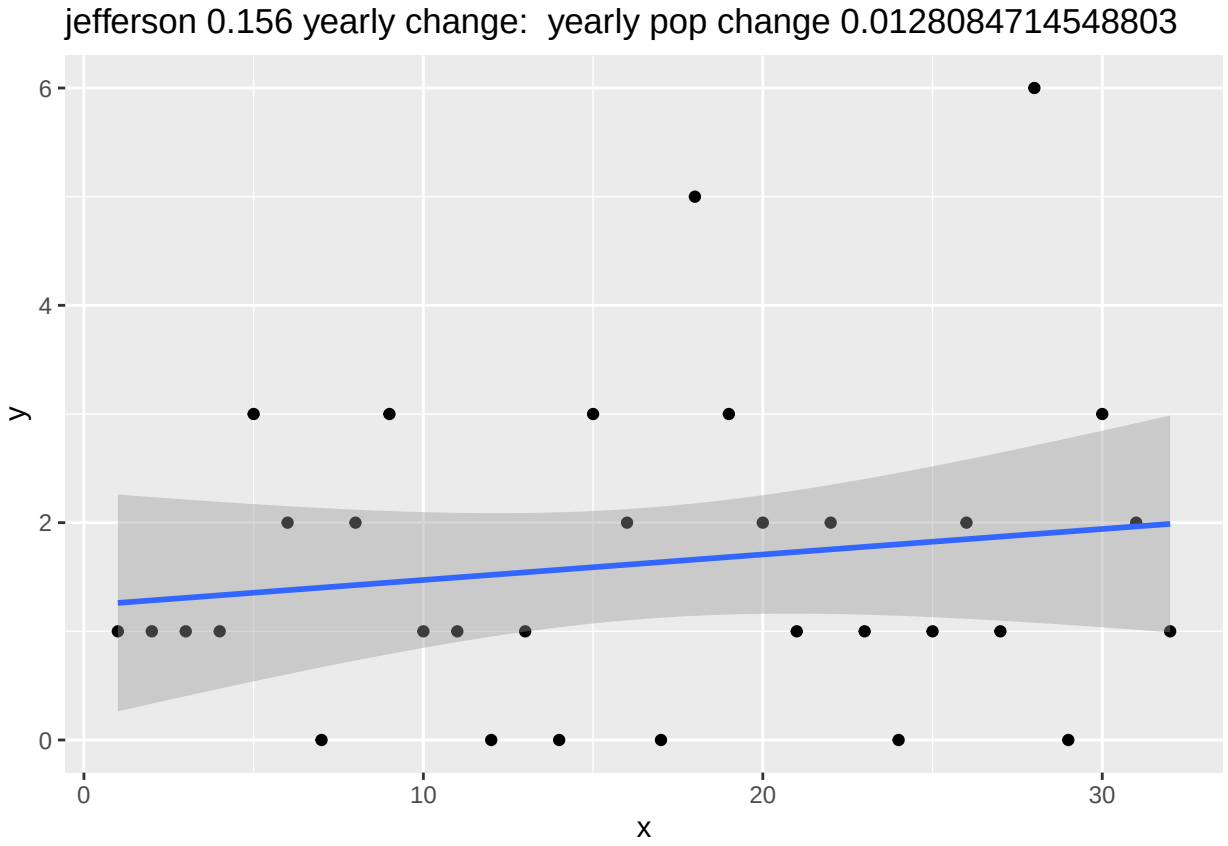

```
## `geom_smooth()` using formula 'y ~ x'
```



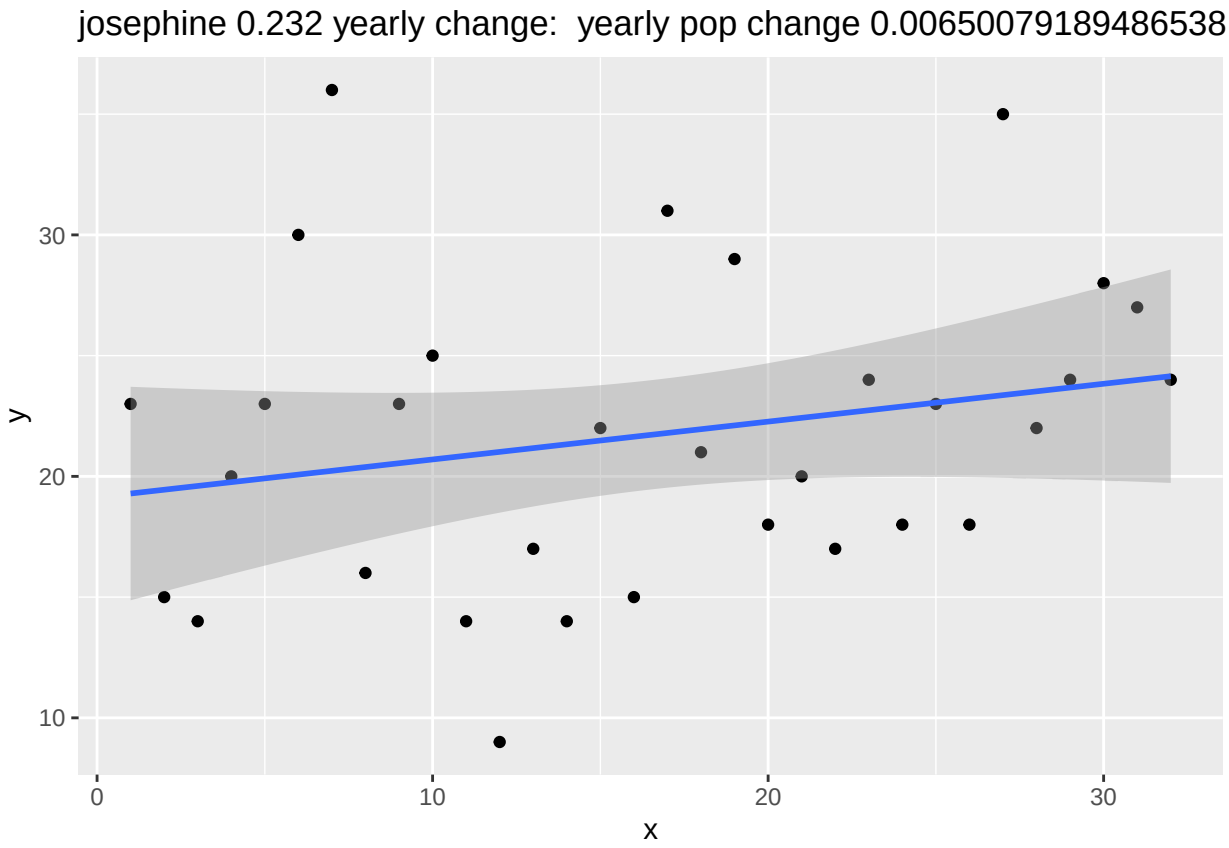

```
## `geom_smooth()` using formula 'y ~ x'
```



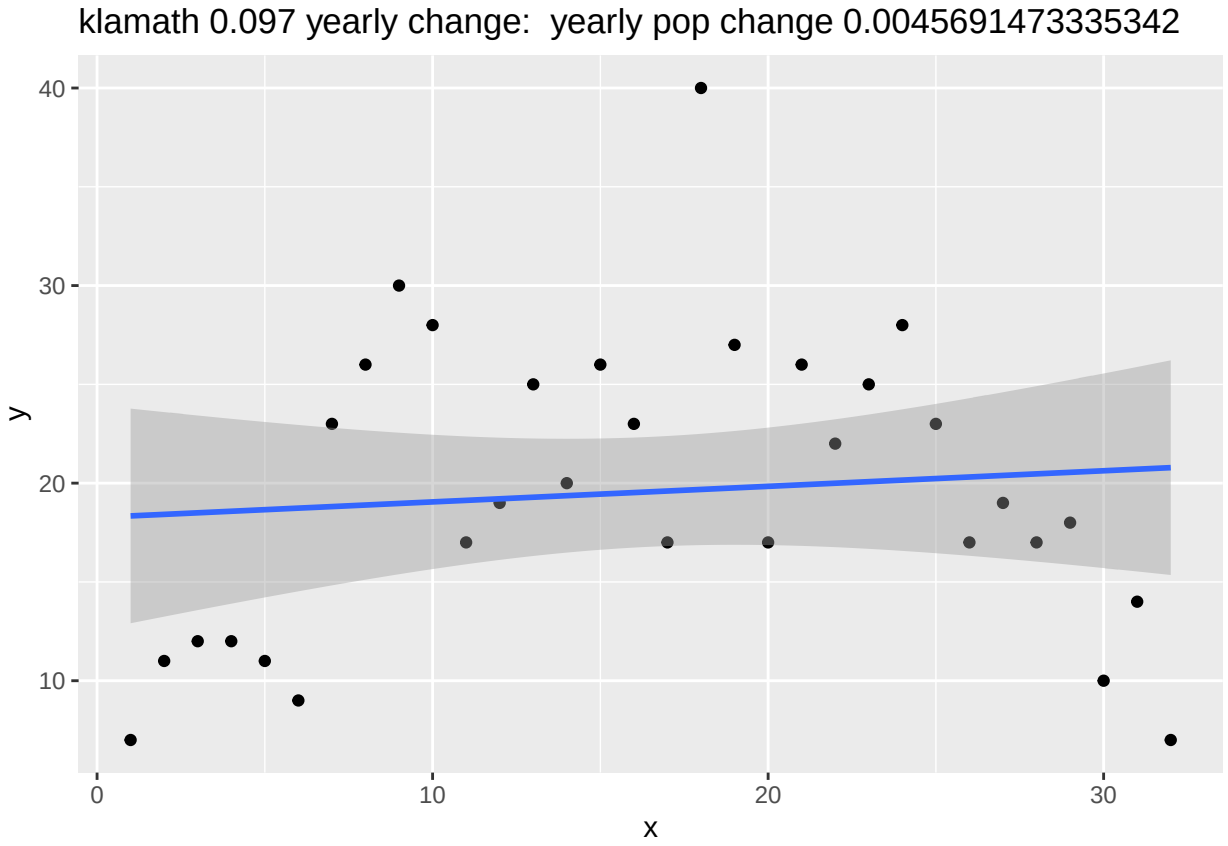

```
## `geom_smooth()` using formula 'y ~ x'
```



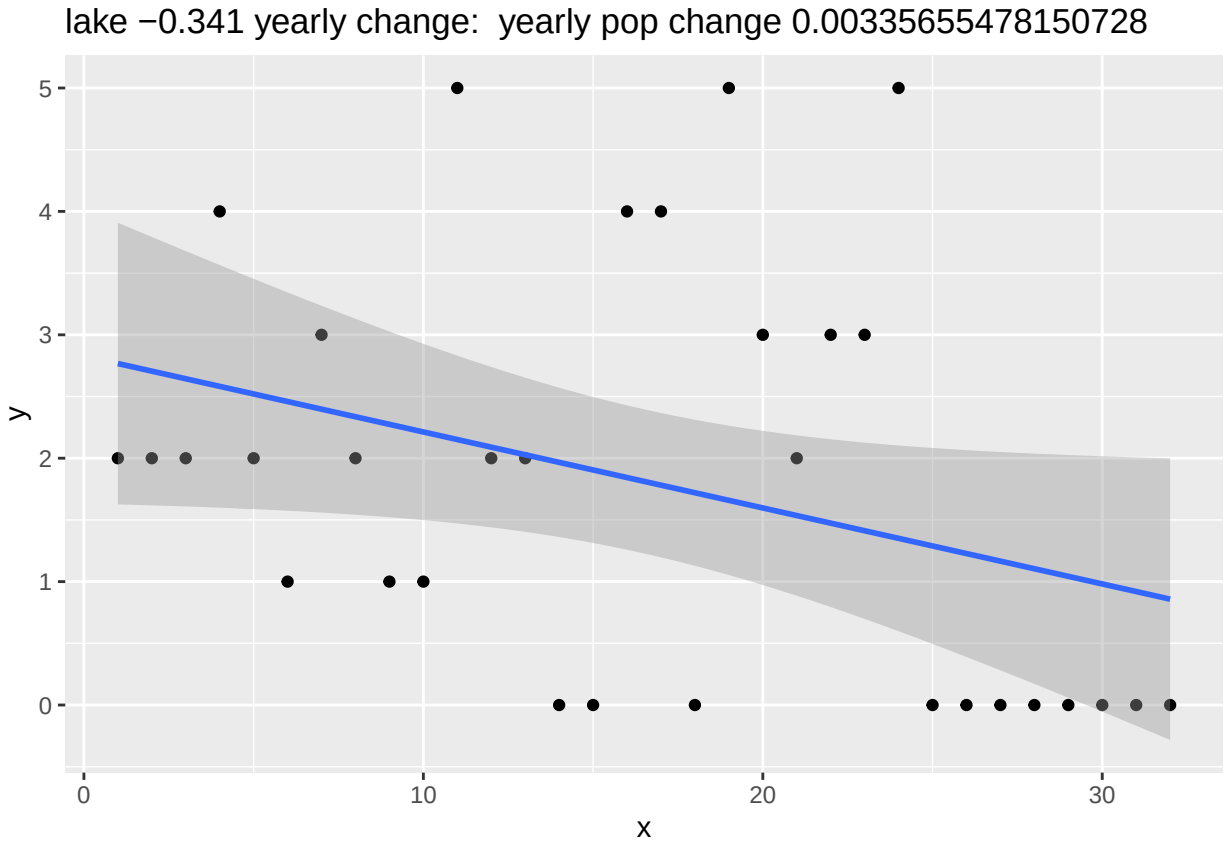

```
## `geom_smooth()` using formula 'y ~ x'
```



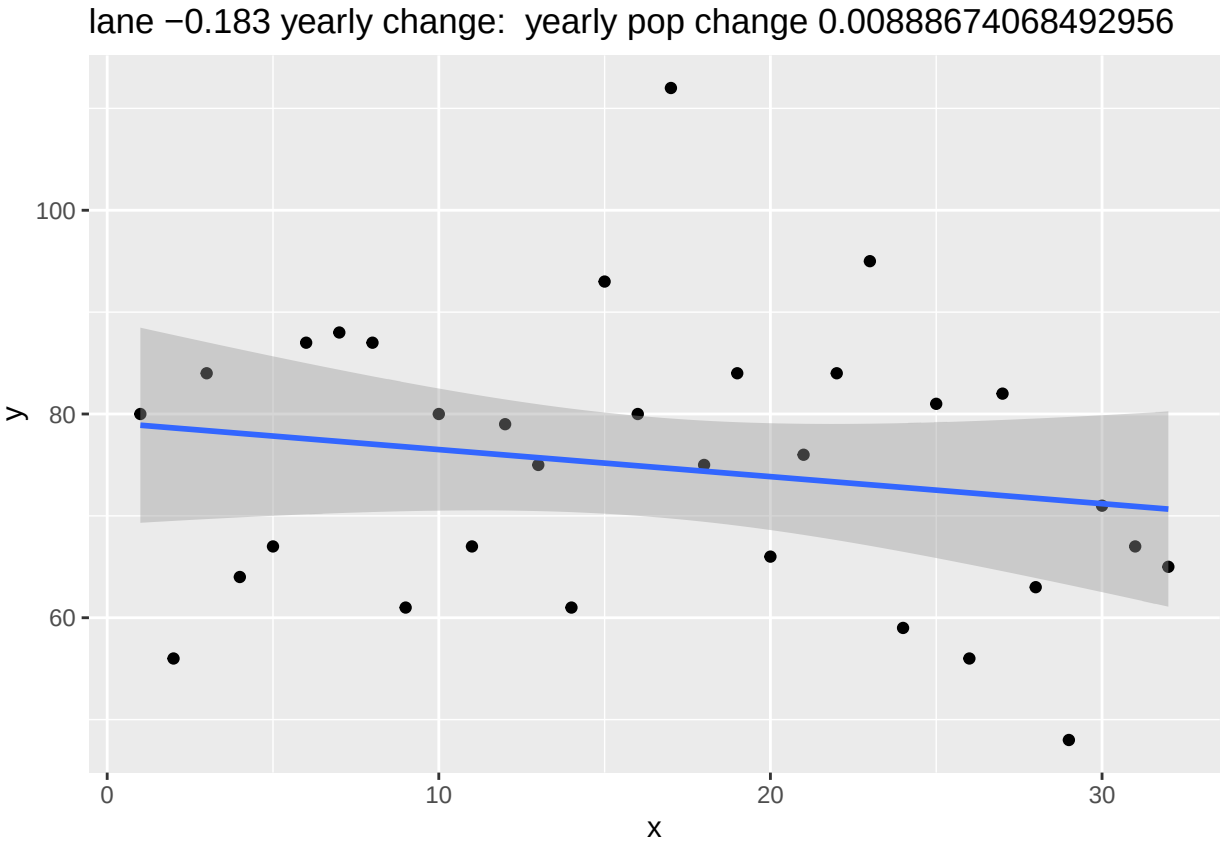

```
## `geom_smooth()` using formula 'y ~ x'
```



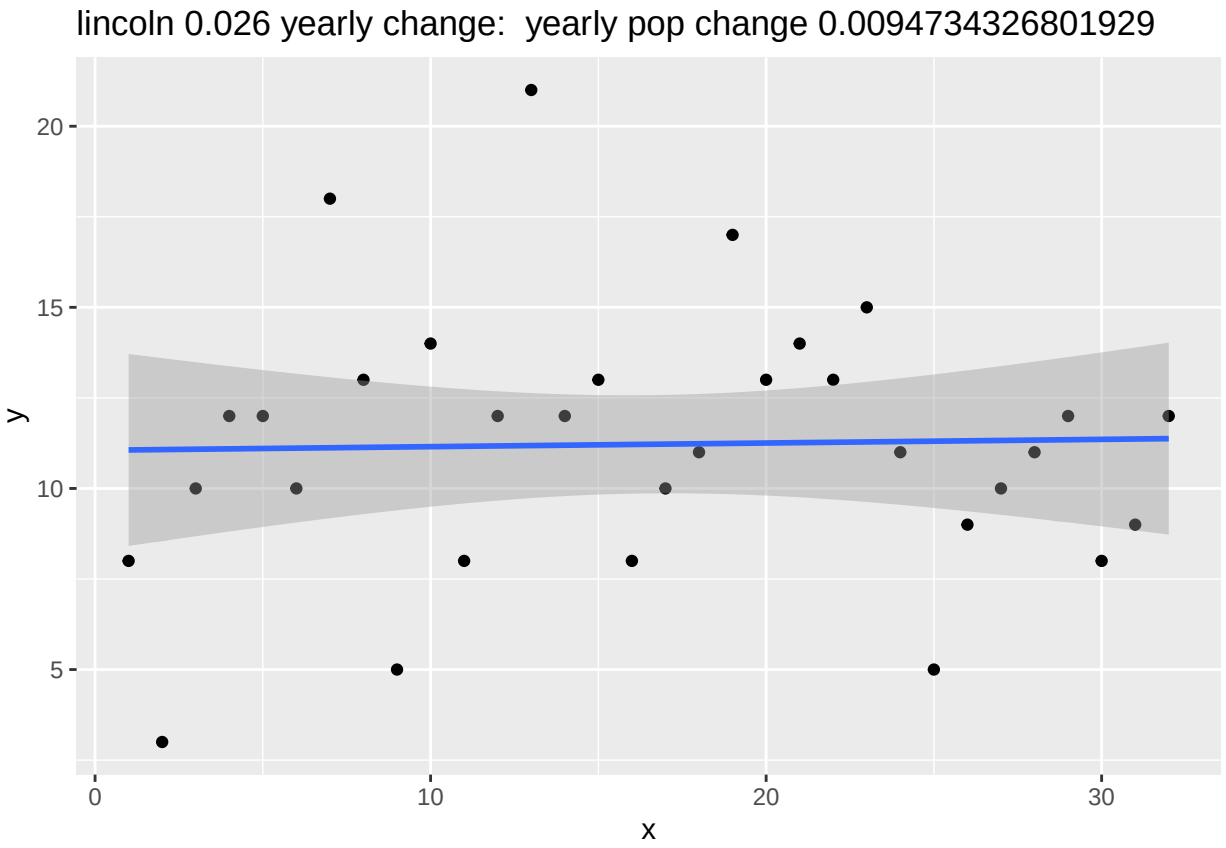

```
## `geom_smooth()` using formula 'y ~ x'
```



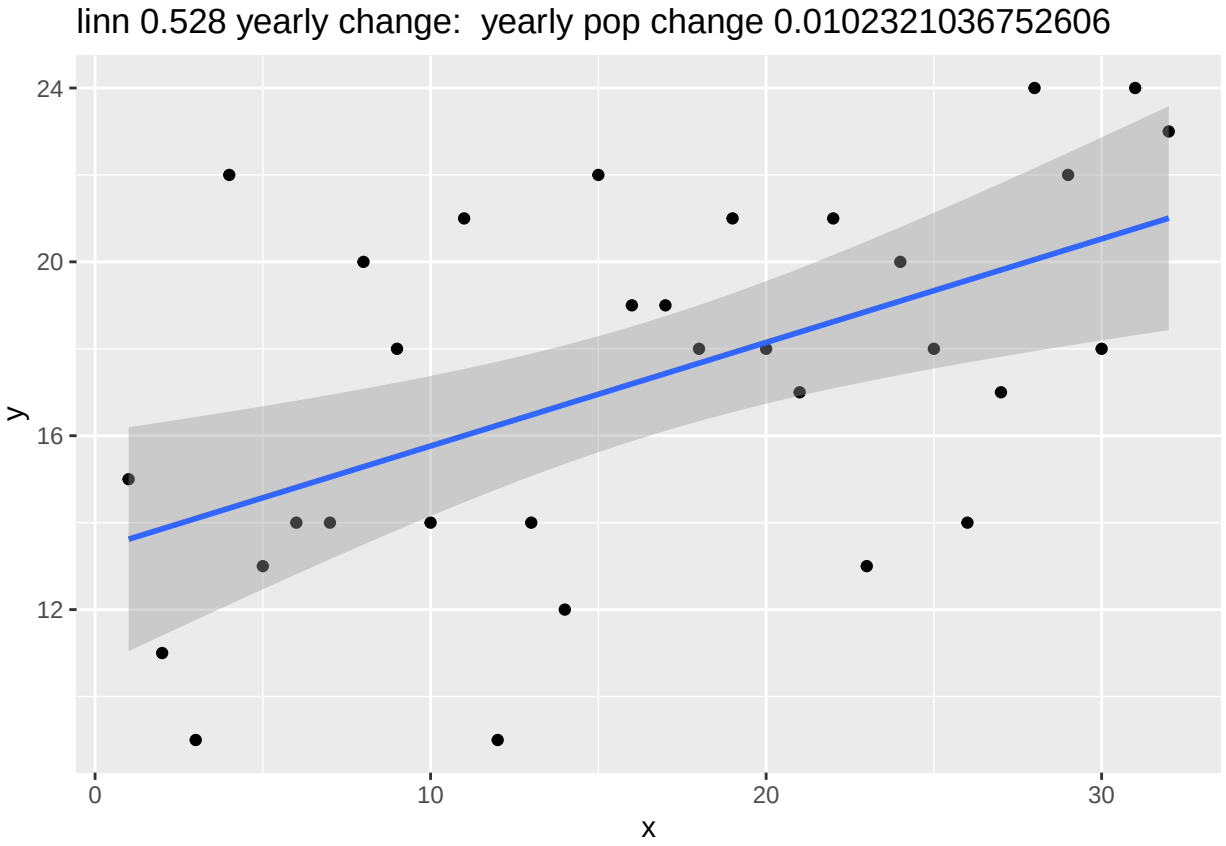

```
## `geom_smooth()` using formula 'y ~ x'
```



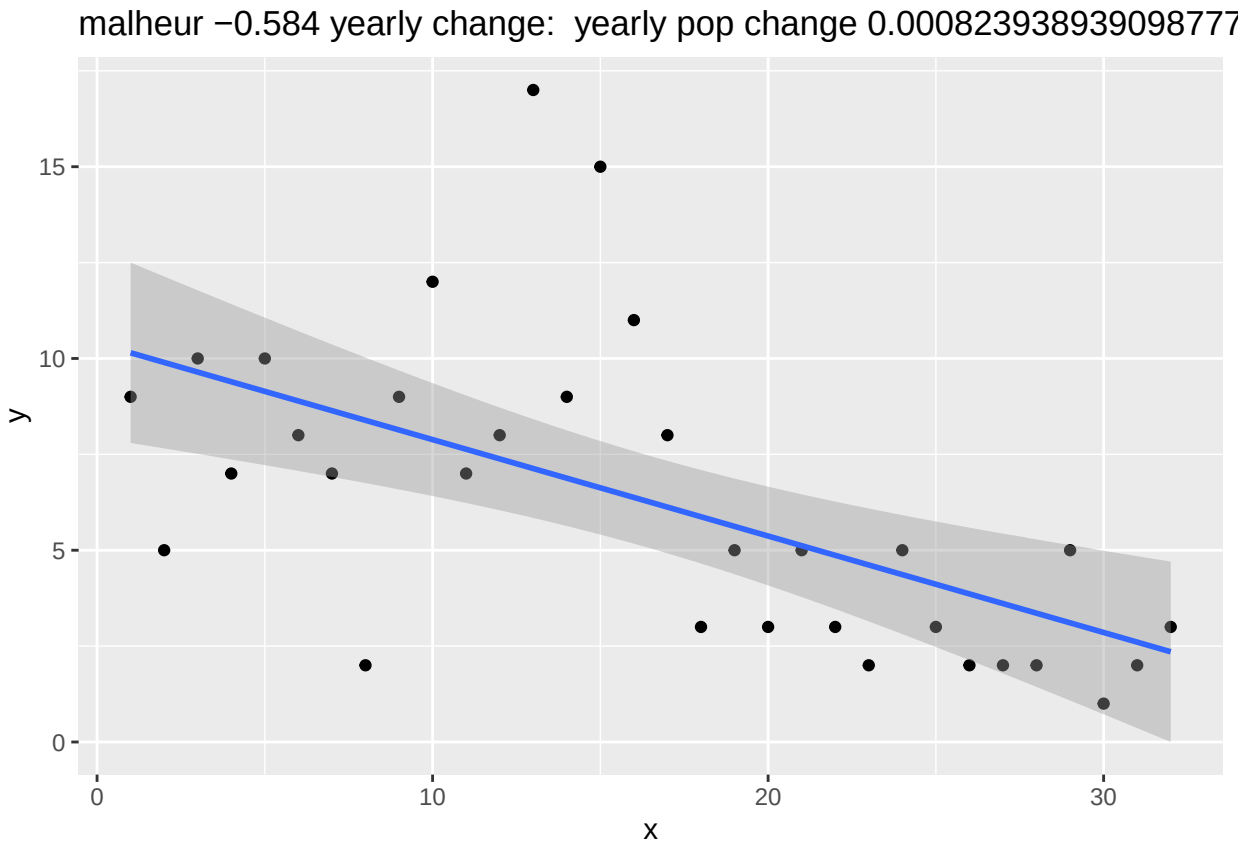

```
## `geom_smooth()` using formula 'y ~ x'
```



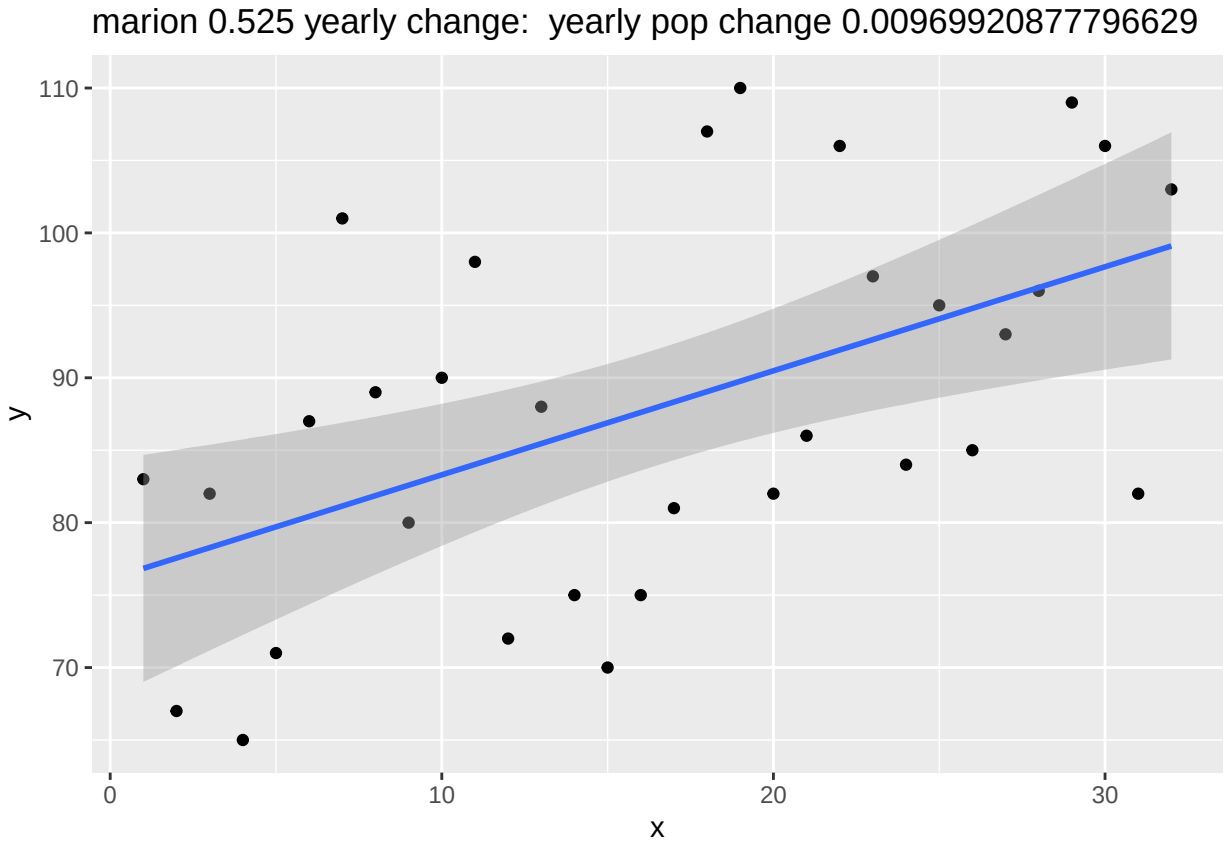

```
## `geom_smooth()` using formula 'y ~ x'
```



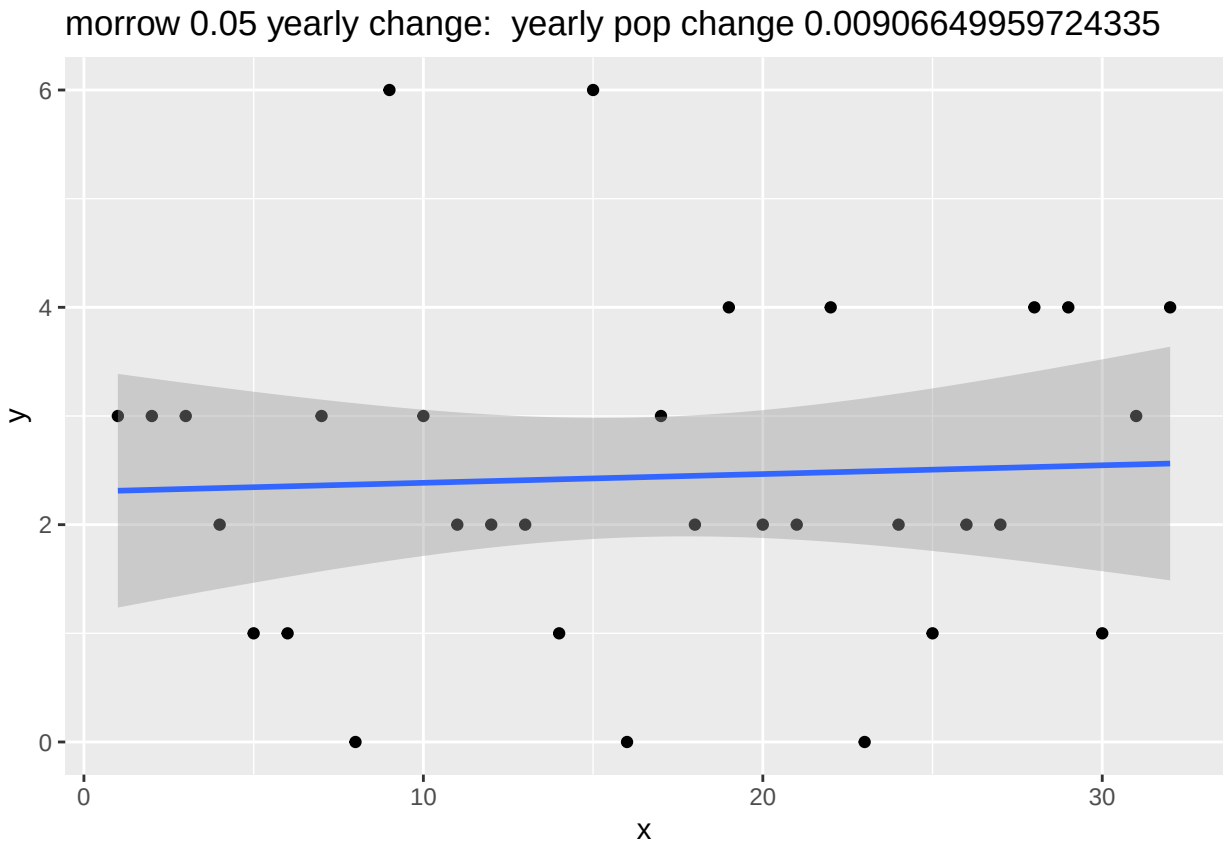

```
## `geom_smooth()` using formula 'y ~ x'
```



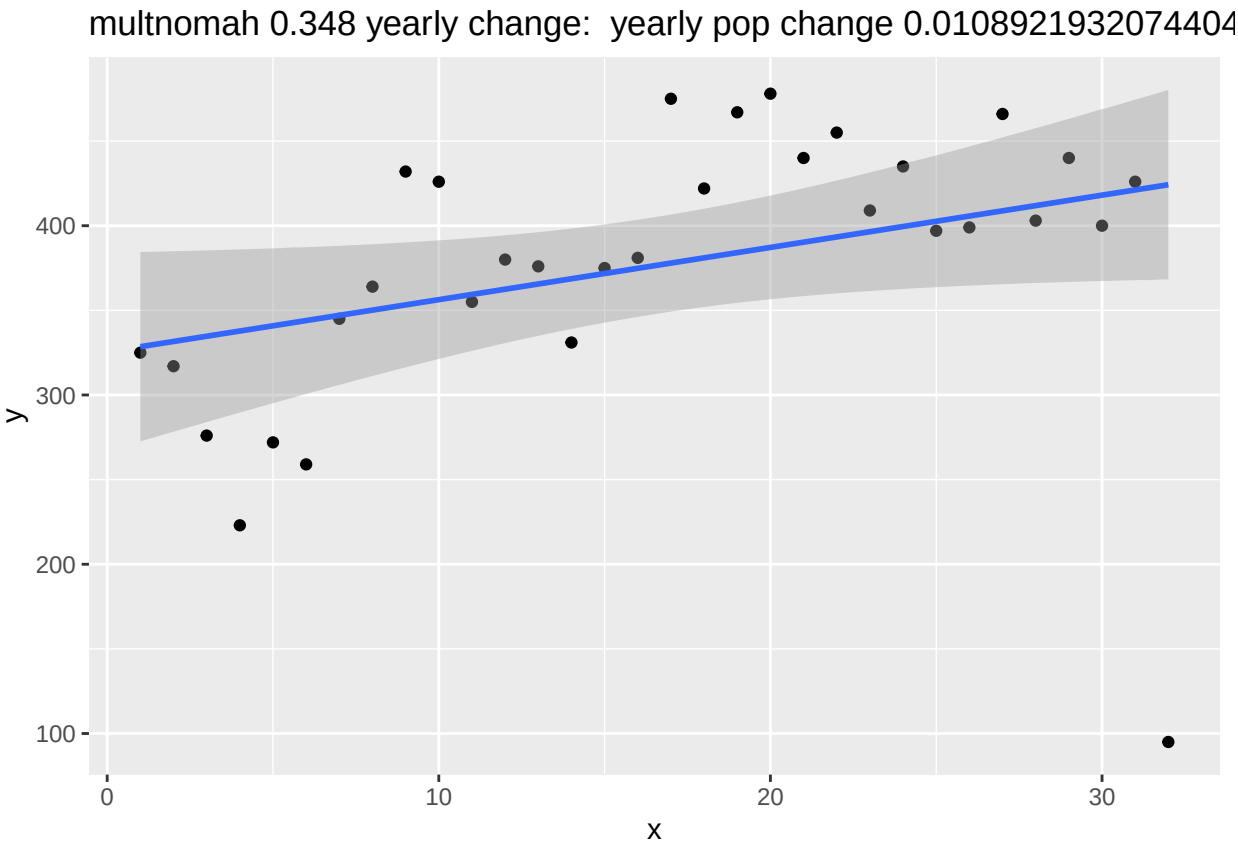

```
## `geom_smooth()` using formula 'y ~ x'
```



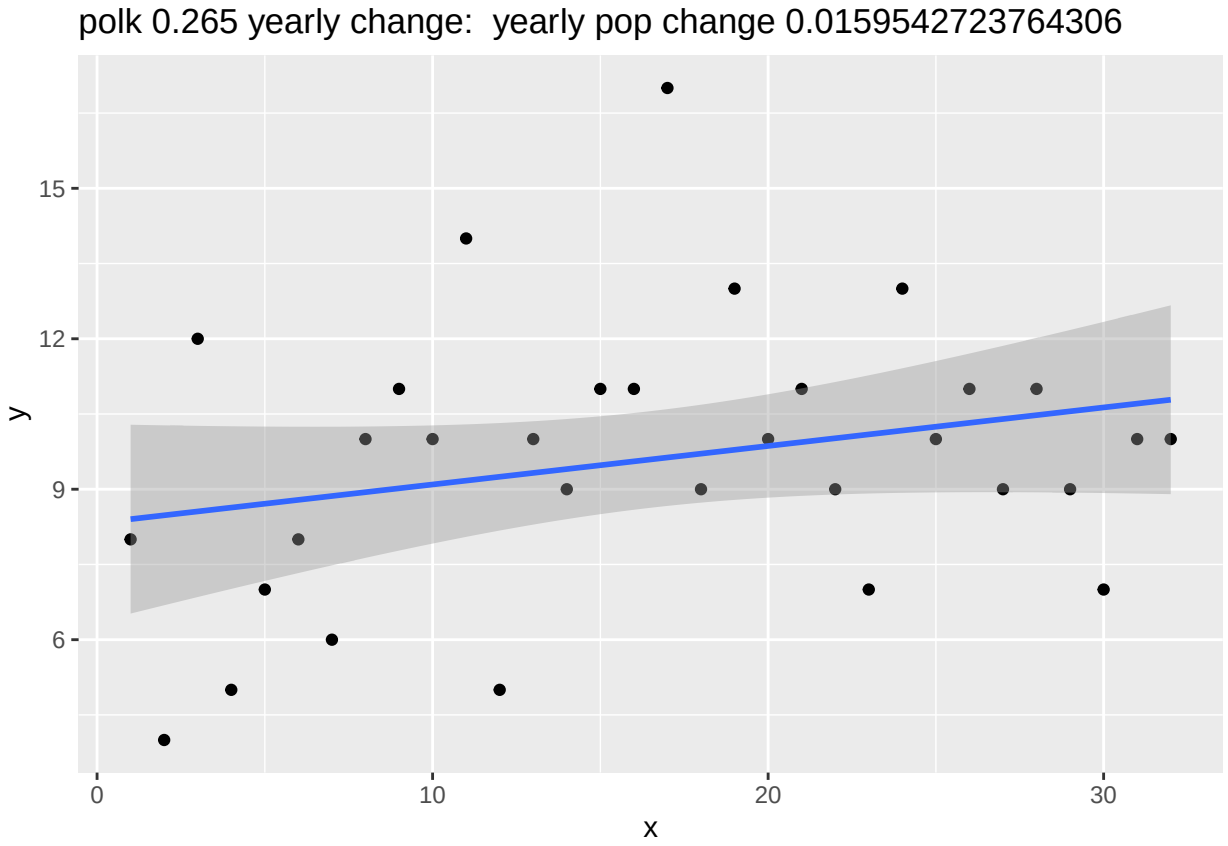

```
## `geom_smooth()` using formula 'y ~ x'
```



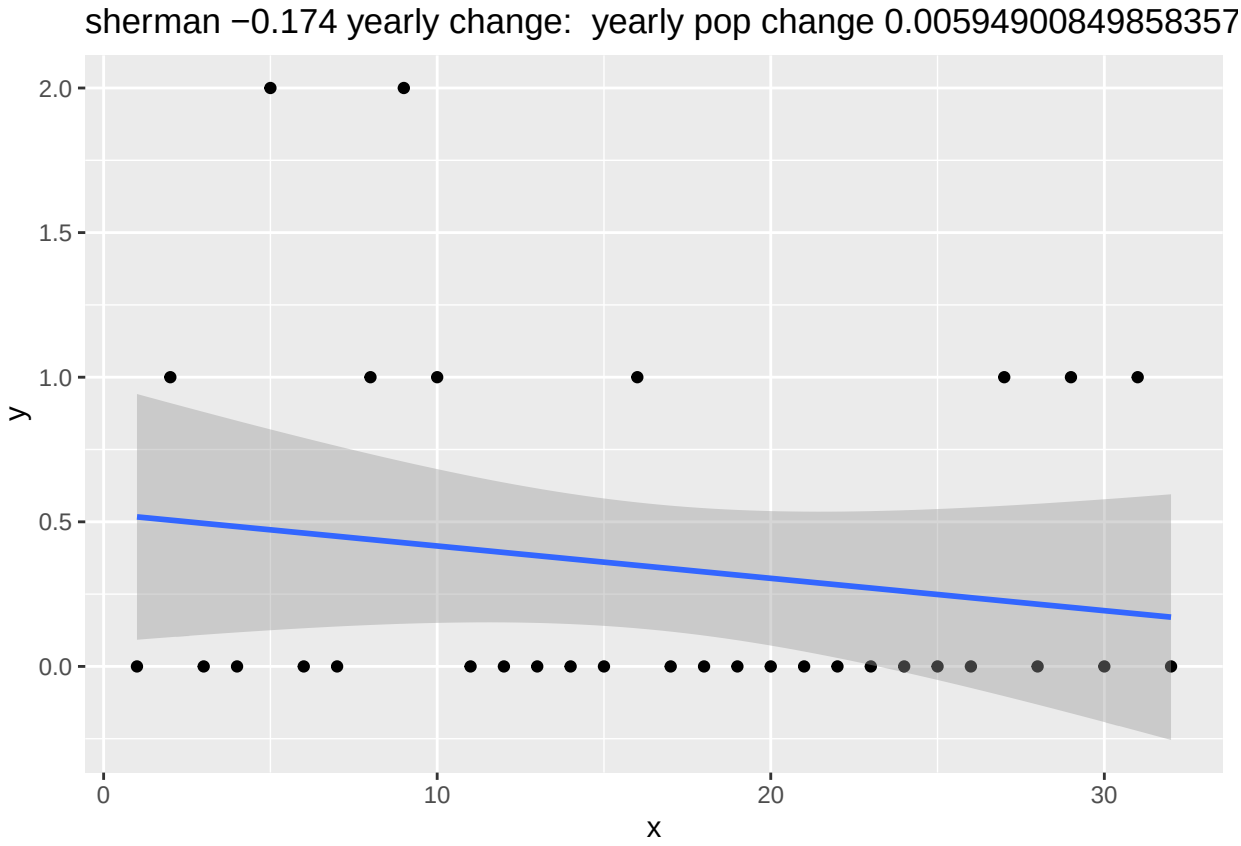

```
## `geom_smooth()` using formula 'y ~ x'
```



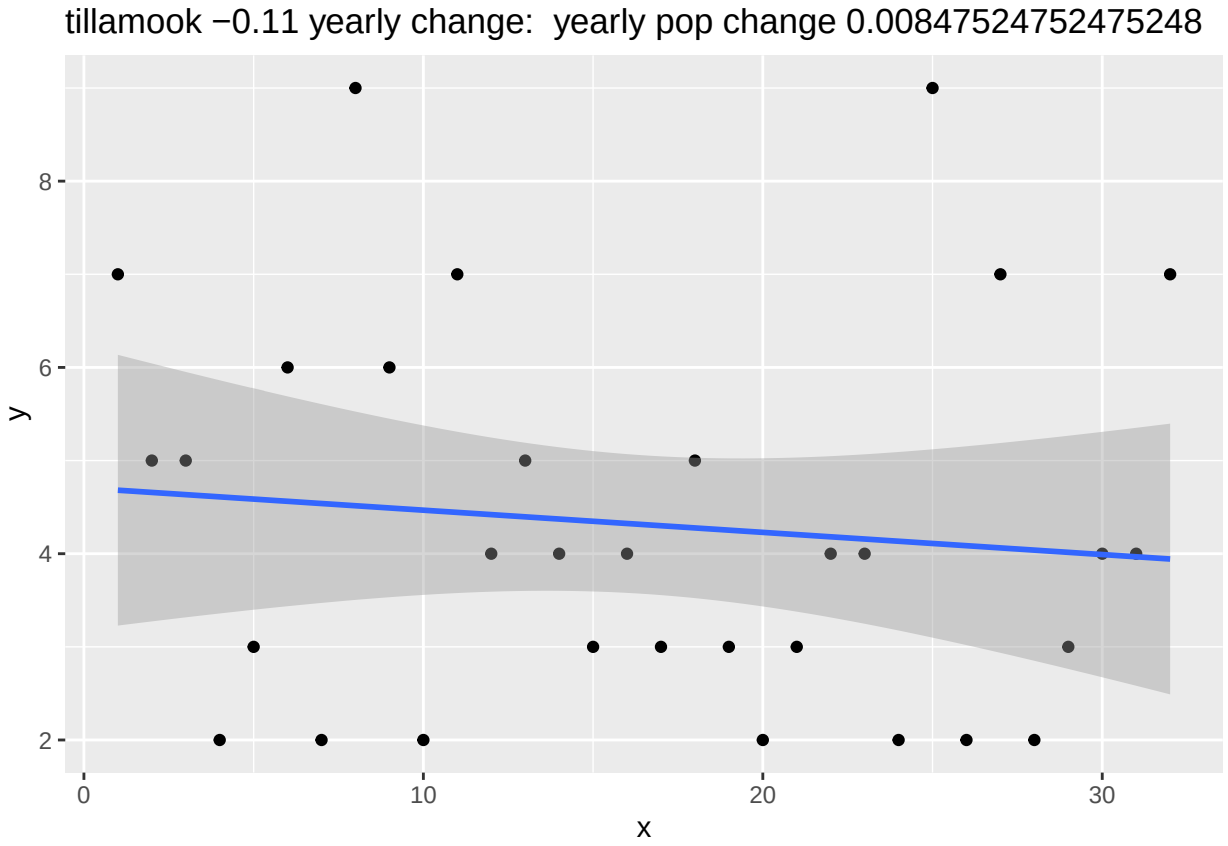

```
## `geom_smooth()` using formula 'y ~ x'
```



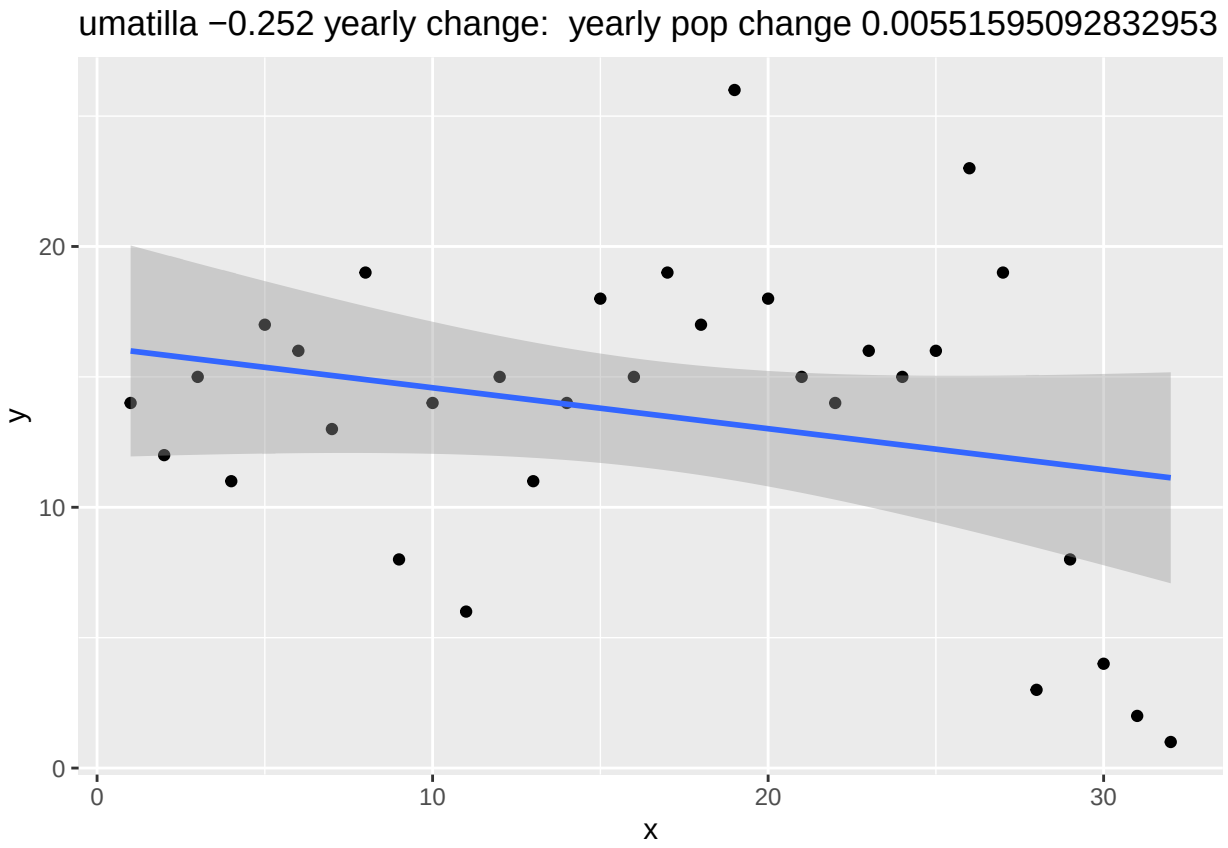

```
## `geom_smooth()` using formula 'y ~ x'
```



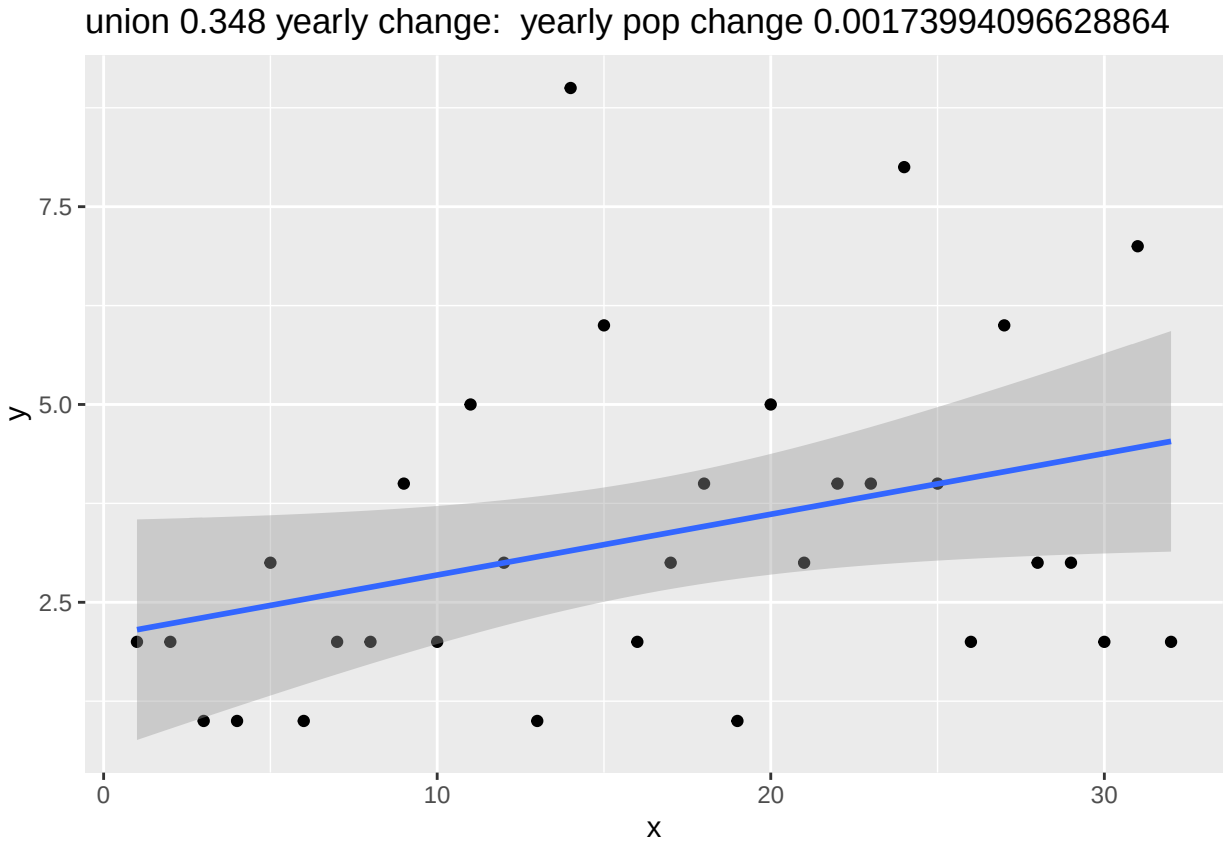

```
## `geom_smooth()` using formula 'y ~ x'
```



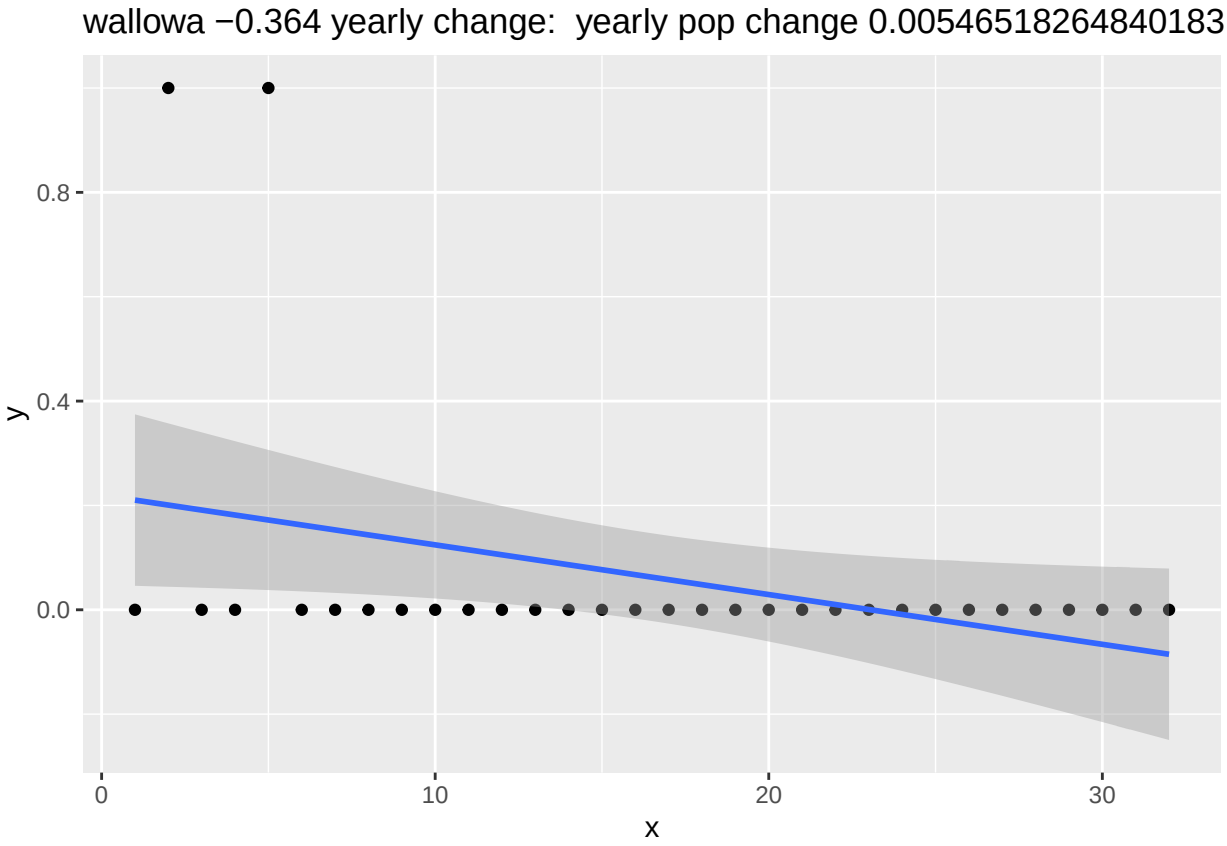

```
## `geom_smooth()` using formula 'y ~ x'
```



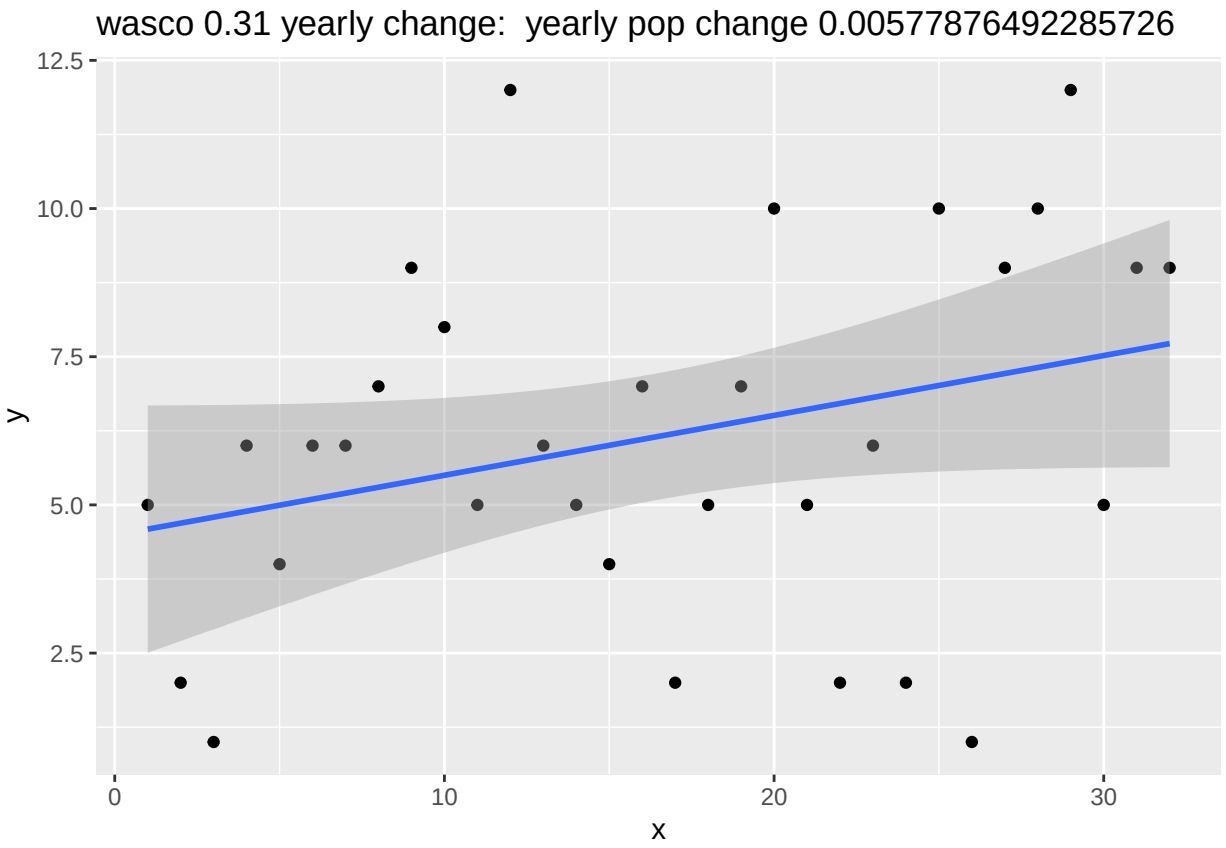

```
## `geom_smooth()` using formula 'y ~ x'
```



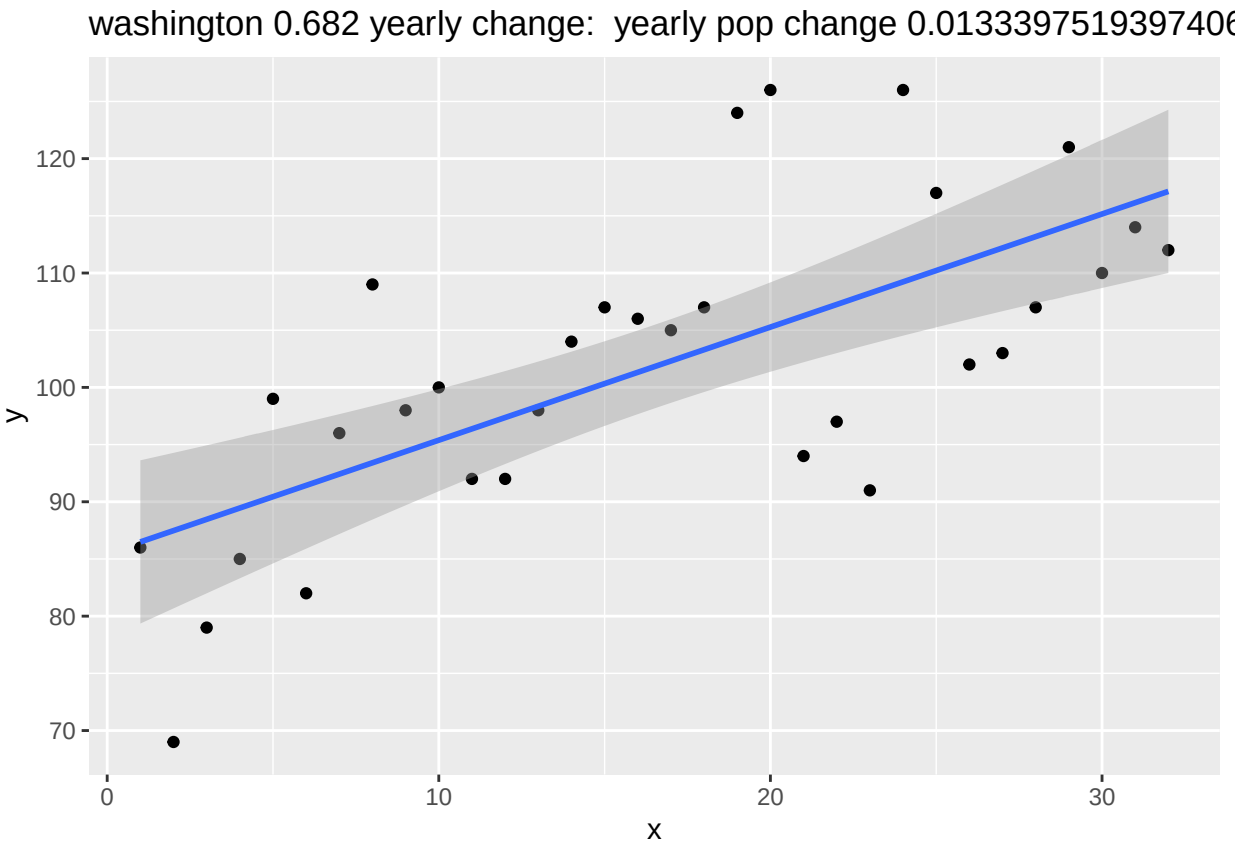

```
## `geom_smooth()` using formula 'y ~ x'
```



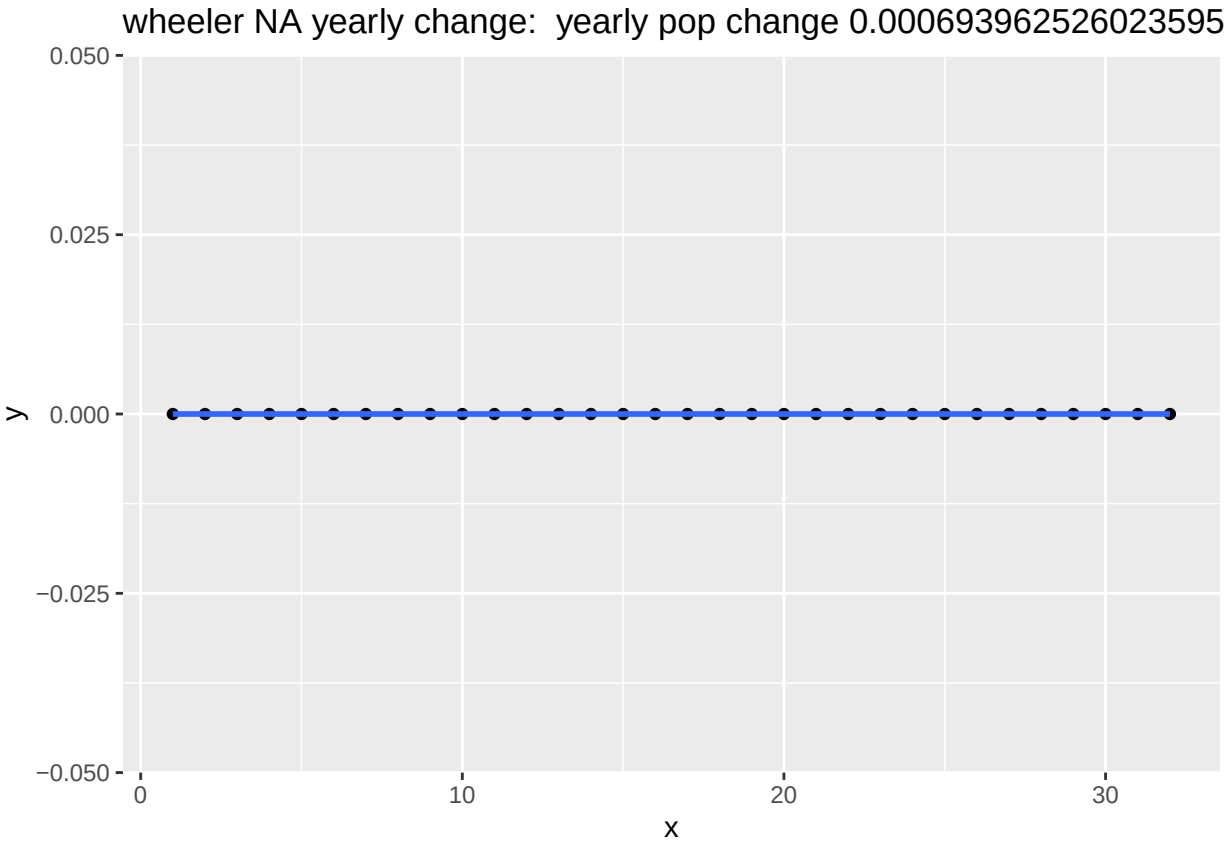

```
## `geom_smooth()` using formula 'y ~ x'
```



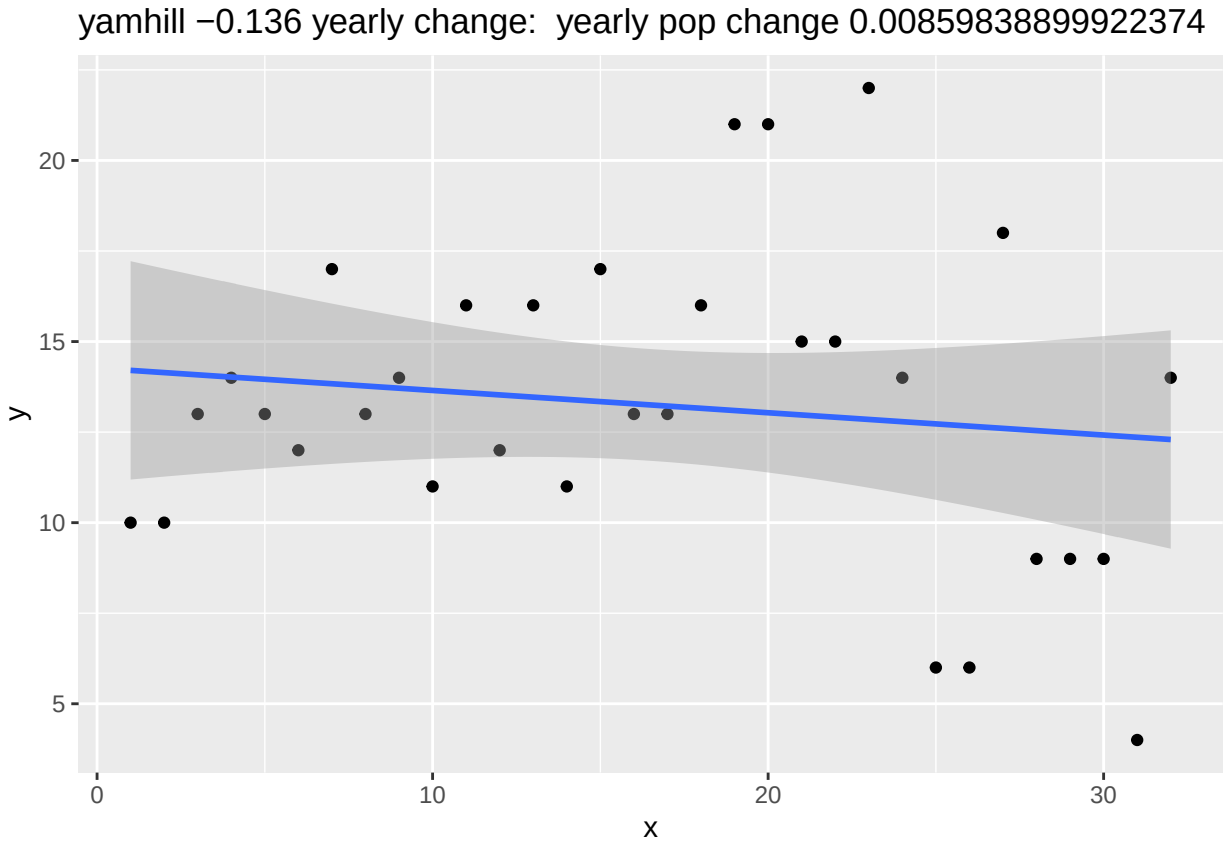

```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```




```
## `geom_smooth()` using formula 'y ~ x'
```


yamhill −0.136 yearly change: yearly pop change 0.00859838899922374
20
15
10
5
y
0
10
20
30
x

Appendix B

source code available at https://github.com/russl-corey/Oregon_Crime