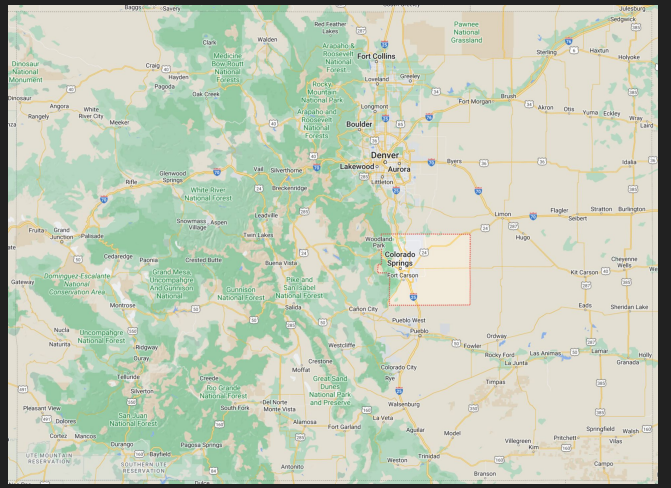
The Impact of Covid on Crimes in the El Paso County

By Eli Corpron

- The most populous county in Colorado
- Population of 730,395 in 2,130 square miles

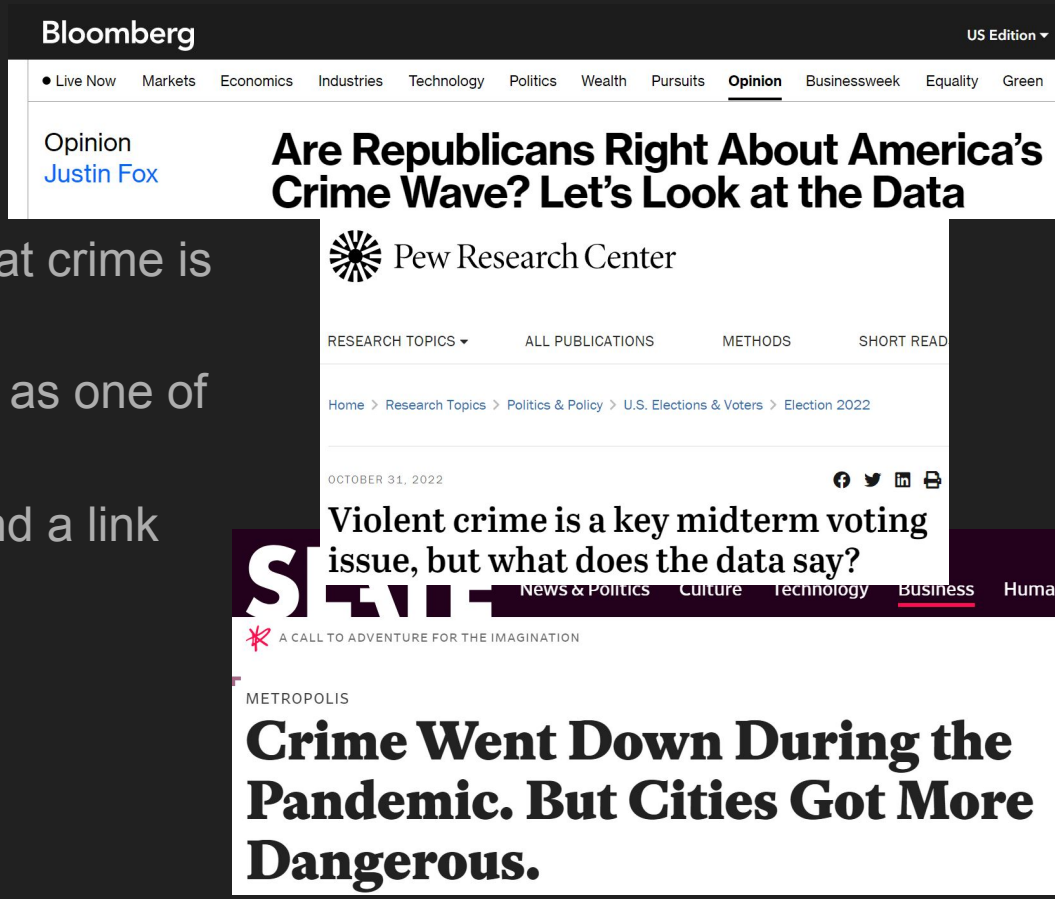
El Paso County Information Continued

- Voted Republican in every presidential election since 1920 except for 1936 and 1964
- Last Democrat governor was elected in 1982



Why Crime?

- Constantly see news articles that crime is on the rise
- The recent midterms had crime as one of the central themes
- Interested in seeing if I could find a link between Covid and Crime

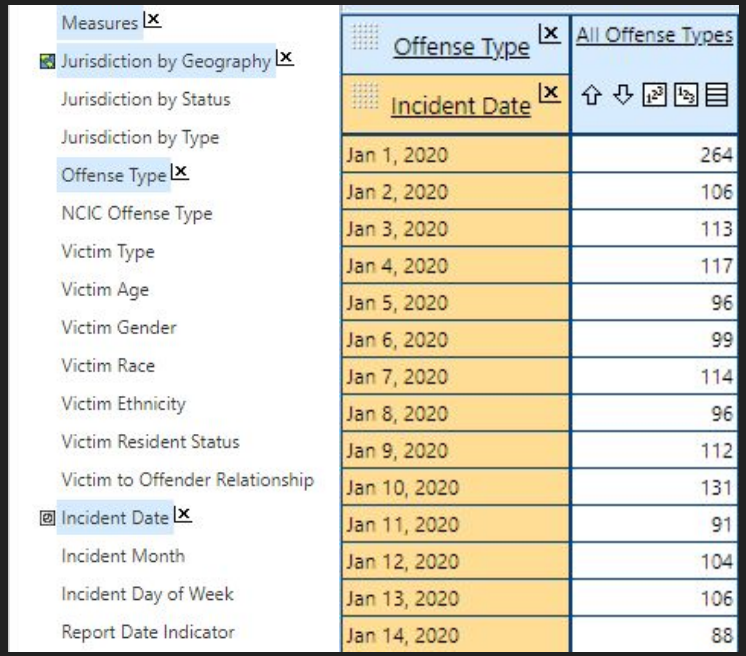


Keeping it Human Centered

- The goal is to have a model that is understandable and explainable.
  - I am not looking for one that predicts crime well without being easily understood
- The data I am looking at is collective data about a broad geographic areas without specific data about people
- The results of this and other similar research can be used to better improve our response to future pandemics

The Data

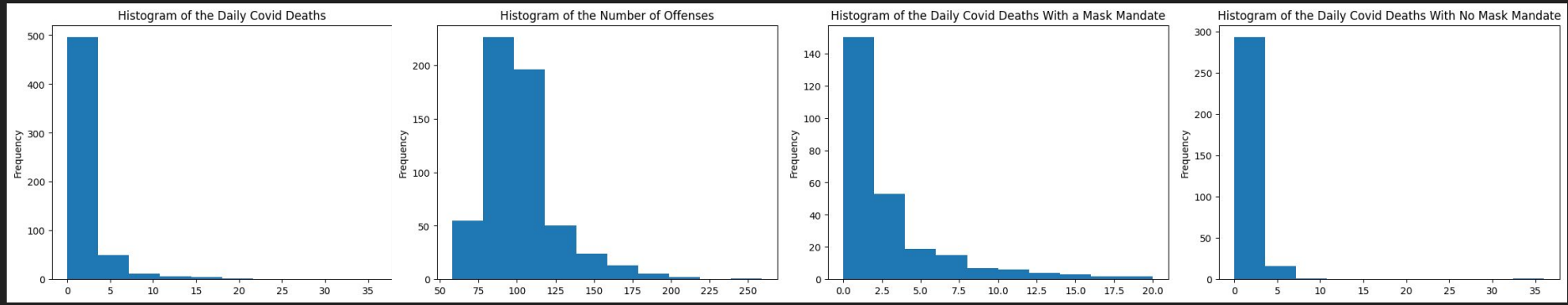
- Covid death data comes from John Hopkins University, Covid masking data comes from the CDC
- The El Paso County Crime Data comes from Colorado Crime Statistics
  - Funded by the Colorado Automobile Theft Prevention (CATPA), which is in the Colorado State Patrol



Incident Date	All Offense Types
Jan 1, 2020	264
Jan 2, 2020	106
Jan 3, 2020	113
Jan 4, 2020	117
Jan 5, 2020	96
Jan 6, 2020	99
Jan 7, 2020	114
Jan 8, 2020	96
Jan 9, 2020	112
Jan 10, 2020	131
Jan 11, 2020	91
Jan 12, 2020	104
Jan 13, 2020	106
Jan 14, 2020	88

The Data Continued - Data Exploration

- Significance Skewness in each histogram
- To keep the model easily understandable I didn't transform any of the variables.

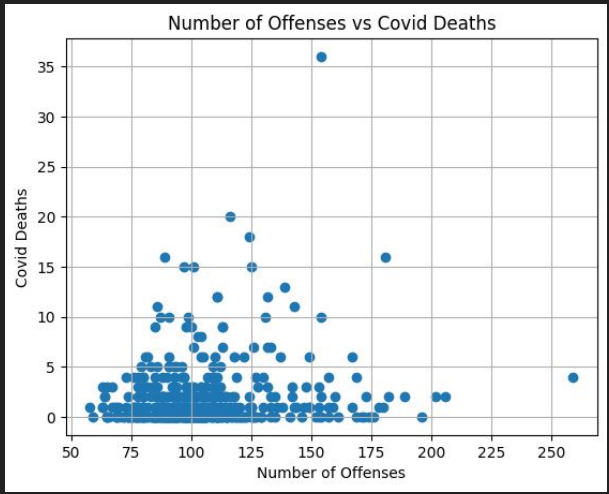


The Research Question

- Initially wanted to know how covid influenced the number of crimes committed in El Paso County
- Specific research question is:
  - Does the number of Covid related deaths and whether a masking mandate was in effect impact the number of offenses that occured in the El Paso County
  - The null hypothesis is that there is not a relation between covid deaths and masking mandates and the number of incidents that occur, with a significance value of .05.

Analyzation Method

- With predicting a numeric response variable with numeric and categorical variable I decided to use multiple linear regression
- Then checked the linear relation assumption between offenses and Covid deaths



The Results

Results were not ideal

- Significant p-values below .05 for both covid deaths and the mask mandate
- R-squared value very low ~ .136
- Skew is extreme
- Other model measures were not good

OLS Regression Results

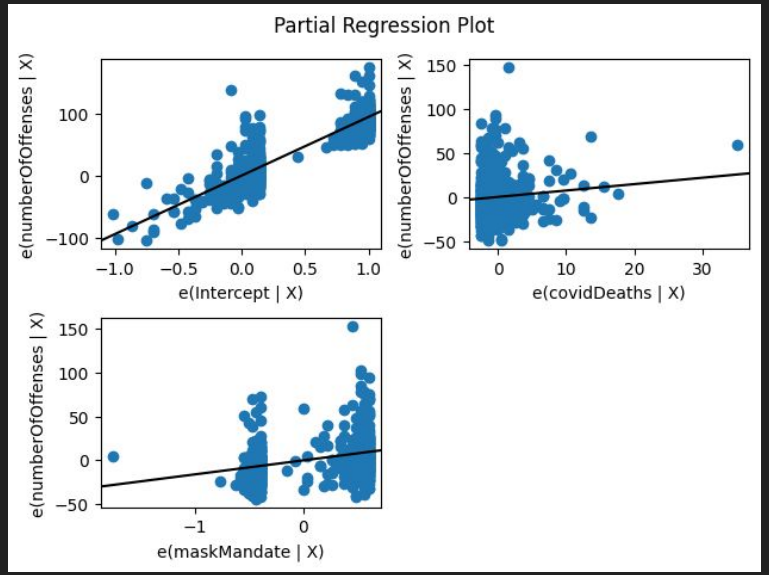
Dep. Variable:	numberOfOffenses	**R-squared:**	0.136
Model:	OLS	**Adj. R-squared:**	0.133
Method:	Least Squares	**F-statistic:**	44.80

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	94.6371	1.291	73.282	0.000	92.101	97.174
covidDeaths	0.7260	0.307	2.362	0.019	0.122	1.330
maskMandate	15.9753	1.911	8.357	0.000	12.221	19.730

Omnibus:	203.810	**Durbin-Watson:**	1.432
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	889.703
Skew:	1.564	**Prob(JB):**	6.36e-194
Kurtosis:	8.248	**Cond. No.**	8.08

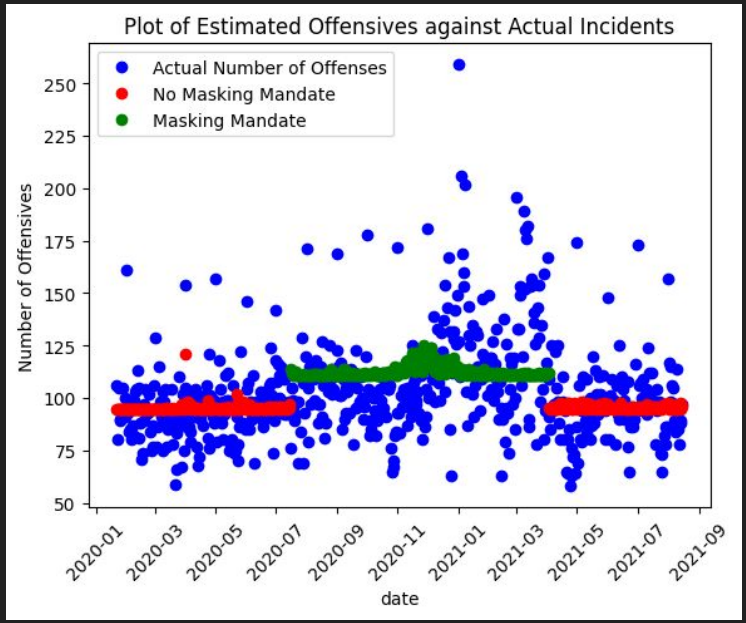
The Results - Continued

- Residuals are clearly not normal
- Residuals do not have constant variance
- Model could be improved by transforming the variables



What the Results Mean

- Results were already very constrained
  - Applied to the El Paso county, from the specific date ranges of 1/22/2020 to 8/15/2021.
- The conclusions are not reliable.
  - Significant p-values undermined by not normal and not constant residuals
  - R-squared value of .136 means not much variance of the data is explained.



Further Work

- Two avenues forward:
  - Keep linear regression as the model, change the data used
    - I could use different data from Colorado Crime Statistics
    - Perform data transformations to attempt to make the variables more normally distributed
  - Change the model
    - Try and find a different predictive model that performs better with the given data