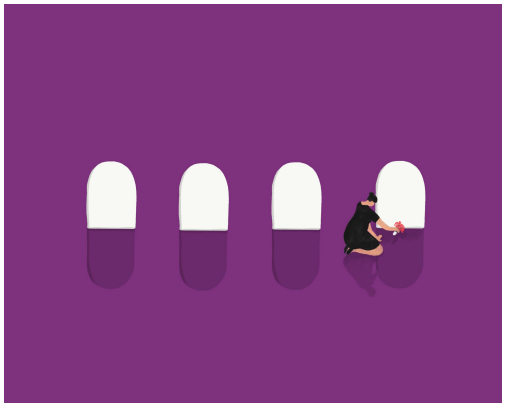
Visualizing the last 30 years drug overdose in the US

Background



As a society, we are much more aware of the threat of drug addiction as a public health threat now than we were thirty years ago. However, the past thirty years have seen an increase across the board in drug overdose deaths, including an alarming increase in the number of opioid-related deaths. Opioids include heroin, synthetic opioids such as fentanyl, and pain relievers available by prescription such as oxycodone, hydrocodone, and morphine. Opioids reduce the sensation of pain by blocking pain receptors in the brain and stimulating the production of dopamine. A side-effect of all opioids is that they can slow breathing to dangerous levels, potentially resulting in death.

Dataset

The Center for Disease Control manages a database with every cause of death for each death reported across the country. With regard to the causes of death related to drug overdose, the CDC has reliable data going back to 1989, which we used for this activity.

Variables

Year: This numeric variable is the year for which drug-overdose death type and number were recorded.

Type of overdose: This categorical variable has the value of either non-opioid or opioid-related. Opioid-related deaths include deaths caused by heroin, fentanyl, and other synthetic opioids, whether illicit or prescription drugs.

Overdose deaths per year: This numeric variable is the number of overdose deaths in a given year according to data collected by the Center for Disease Control.

t (year): This numeric variable is a common way of representing time in a dataset starting from year 1 of the dataset. The value of 1 represents the year 1989 in this study; the value of 31 represents the year 2019.

Activity

1) Make a graph to show the trend in overdose deaths for both opioid and non-opioid drugs in the US from 1989 to 2019.

First, go to the Graph tab (near the top of the screen when viewing the data table in the DataClassroom app). Show Overdose deaths per year on the Y axis and Year on the X axis. To visually separate the opioid deaths from non-opioid deaths, click Show for the Type of overdose variable and make it show on your graph by clicking Z on the right (it will turn red when you do).

2) One of the most important aspects of data science is being able to make inferences from data. If you were to begin researching the cause of the accelerated growth in opioid-related overdoses, approximately what time frame would you begin your search? Why?

3) Configure your graph to only display the opioid-related deaths.

To do this, go to the Type of overdose variable above the graph. Click on the Values button. In the Label, the values window that pops up, exclude values for non-opioid deaths by clicking in the check box. This should leave only opioid deaths represented on your graph.

4) Fit a line of best fit to these data by checking the Regression line box to the right of the graph.

What is the overall trend highlighted by the linear regression model? What is the r squared value?

5. Now try an exponential regression model instead of a linear one. Choose Exponential from the dropdown menu that currently says Linear. You will need to make a small change to the data on the graph. Replace Year with t on the X axis.

Note: Even computers have a tough time handling certain numbers that are really large, and so it's typically best to use small values for X when using exponential models. We can change our independent variable this way as long as it doesn't change what the data represents. In this case, representing 1989 as year 1, 1990 as year 2, and so on, is a simple way to do so.

What is the r-squared value for the exponential regression? Which do you think is the better model, linear or exponential, for these particular data? What are the implications of this for opioid addiction treatment programs in the US?