Response Summary:

Q3. Student Information *

First Name	Claire
Last Name	O'Malley
Major	UX Design
Course ***(e.g. CGT 270-001)***	CGT270
Term ***(e.g. F2019)***	F2022

Q4. Email Address *
(University Email Address is required.)
omalle18@purdue.edu

Q5. Visualization Assignment *
- Final Project

Generate

Q7. Identify appropriate data sources: is the data publicly available? What search methods were used? *

Data source 1	Non-Fatal Overdose Data-publicly available. Google Search.
Data source 2	SUDORS Fatal Overdose Data-publicly available. Google Search.
Data source 3	NCHS Data Brief #428-publicly available. Google Search.

Q8. Data format: what format is the data in? Structured vs instructed? All text, a combination, multiple sources? Is it primary or secondary data? *
1. Excel Spreadsheet. Structured. Combination. Primary data.
2. Excel Spreadsheet. Structured.. Combination. Primary data.
3. PDF. Structured.. Combination. Secondary Data.

Q9. Data types: what types of data are in the data? How are they stored? What is the access to the data (API, JSON, txt, csv, etc.)? What structure holds the data (data base, spreadsheet, etc.)? *
Characters, Strings, Floats, Boolean, Integers. Stored in Excel spreadsheets and PDF. CSV and PDF files. Two Excel spreadsheets and a PDF.

Evaluate

Q11. Variables: list the data variables? What are the parameters? Give them names. What are the dependent variables and independent variables? *

Rate of overdose deaths by state and drug or drug class-Numeric. Independent Data.
Percentages of overdose deaths involving select drugs and drug classes-Numeric; percentage ranging from 0-100. Independent Data.
Percentage of overdose deaths involving the most common opioids and stimulants alone or in combination-Numeric; . percentage ranging from 0-100,Character. Dependent Data.
Distribution of overdose deaths by opioid and stimulant involvement-Numeric; percentage ranging from 0-100 Dependent Variable.
How many drug overdose deaths occurred each month in 2020?-Numeric. Independent Data.
Who died of a drug overdose in 2020? By Sex-Numeric; percentage ranging from 0-100 Dependent Variable.
Who died of a drug overdose in 2020? By Race/Ethnicity-Numeric; percentage ranging from 0-100 Dependent Variable.
Who died of a drug overdose in 2020? By Age (In Years)-Numeric; percentage ranging from 0-100.Dependent Variable.
Who died of a drug overdose in 2020? By Age and Sex-Numeric; percentage ranging from 0-100.Dependent Variable.
Potential opportunities for intervention-Numeric; percentage ranging from 0-100.Dependent Variable.
Additional circumstances surrounding overdose deaths-Numeric; percentage ranging from 0-100.Dependent Variable.
State-Character. Independent Variable.
Geo-Location-Character. Independent Variable.
Start Year-Integer. Independent Variable.
End Year-Integer. Independent Variable.
Start Month-Integer. Independent Variable.
End Month-Integer. Independent Variable.
All Percent Change-Float. Dependent Variable.
Opioid Percent Change-Float. Dependent Variable.
Heroin Percent Change-Float. Dependent Variable.
Stimulant Percent Change-Float. Dependent Variable.
All LS Significant-Binary. Independent Variable.
Opioid LS Significant-Binary. Independent Variable.
Heroin Ls Significant-Binary. Independent Variable.
Stimulant LS Significant-Binary . Independent Variable.
All Significance-String. Independent Variable.
Opioid Significance-String. Independent Variable.
Heroin Significance-String . Independent Variable.
Stimulant Significant-String. Independent Variable.
Gender-String. Independent Variable.
Age Range-String. Independent Variable.
Jurisdiction Count-String. Independent Variable.
Comparison Type-String. Independent Variable.
Year-Integer. Independent Variable.
Number-Integer. Independent Variable.
Deaths per 100,000-Float. Dependent Variable.
Sex-String. Independent Variable.
Age Group-Integer. Independent Variable.
Race and Hispanic Origin-String. Independent Variable.
Opioid Type-String. Independent Variable.
Stimulant Type-String. Independent Variable.

Q12. Audience & Assumptions: list any assumptions you have about the data. Who is your audience? *

The data sources are trustworthy and as for the non-fatal data, legalized drugs were likely not taken into consideration.

Generate

Q14. What real life behavior does the data reflect? Does it show patterns of activity, regularity of events, a timeline, population data, etc? Explain. *

Addiction shows repeated human drug use. Death shows a pattern of activity because it shows the common result of addiction. We can see the statistics changing from year to year.

Q15. What are the weaknesses of the data source? Is it likely that the source will be available in the future? Is the data complete? What is the quality of the data? Is it specific to your needs for. the current project? Is the data in the format you need? Are there missing data? Explain. *

The nonfatal data is solely based off of reported cases and there is some missing data. The fatal data does not take into account people that died from overdoses as well as other causes. Yes it is likely it will be available in the future because the data is from the government. Not all of the data is complete. The quality is good because it comes from a trusted source. It is specific for the needs of our project. We had to change the PDF document into an excel sheet. There was missing data on the non-fatal dataset , but it has no significant impact.

Q16. What information is emphasized? What is the central focus of the data? Explain. *

The information that is emphasized are the percent changes present in the data. The central focus of the data is the deaths and impacts of opioids.

Q17. At what level of granularity is the data provided? Is the data summarized, or do you have access to the raw data? Is the data categorized or is the data in a format that allows you to create your own categories, etc. Explain. *

The data is summarized. Yes we have access to raw data. The data is categorized.

Q18. What is the scope of the data? What topics can be covered using the data? Is there a time range/frame? Is the data for a specific area/discipline/demographic etc.? Explain. *

The topics that are covered are how many non-fatal impacts occurred between 2018-2022 regarding the opioid crisis. As well as the deaths across the country as of 2020.