NAME : Ameya
Dabholkar
UID : 2021300023
CLASS : BE COMPS A
BATCH : ADV BATCH F

ADV EXPERIMENT 4

DATASET :

E-commerce sales dataset -
https://www.kaggle.com/datasets/tunguz/us-estimated-crimes

DATASET DESCRIPTION :

This dataset contains estimated data at the state and national level and was derived from the Summary Reporting System (SRS). These data reflect the estimates the FBI has traditionally included in its annual publications. Download this dataset to see the FBI's estimated crime totals for the nation and all 50 states

REPORT :

1. What is the purpose of the density plot for property crime?

Answer:

The density plot for property crime provides a smooth estimate of the distribution of property crime rates across all states in 2019. It helps us understand how property crime is spread among the states and highlights whether the distribution is skewed or symmetric.

2. What insights can we gain from the density plot of violent crime?

Answer:

The density plot of violent crime shows the distribution of violent crime rates across states. If the plot has a peak or cluster, it indicates that most states have similar crime rates. If the distribution is spread out, it shows variability among states. Peaks at both ends might indicate a few states have unusually high or low violent crime rates.

3. Why do we use the IQR method for outlier detection?

Answer:

The IQR method is used to detect outliers by identifying values that are significantly higher or lower than the majority of the data. It is a robust method that isn’t heavily influenced by extreme values, ensuring that the density plots represent the majority of the data accurately, without skew from outliers.

4. What does a wider density curve in the property crime plot indicate?
Answer:
A wider density curve in the property crime plot indicates a larger variation in property crime rates across states. This means the rates are more spread out, with some states experiencing much higher or much lower property crime rates than others.
5. What could cause peaks in the density plot for violent crime?
Answer:
Peaks in the density plot for violent crime could indicate that a large number of states have similar violent crime rates. For example, if there is a single peak, it suggests that most states fall into a particular crime rate range. Multiple peaks could indicate clusters where certain states share similar crime rates at different levels.
6. How do the property crime and violent crime distributions compare?
Answer:
By comparing the two density plots, we can observe whether property crime or violent crime is more consistent across states. For instance, if the property crime plot has a more narrow peak while the violent crime plot is more spread out, it suggests property crime rates are more consistent across states, while violent crime rates vary more.
7. Why are density plots better for this data than histograms?
Answer:
Density plots provide a smoother view of the distribution, which can be more visually intuitive compared to histograms, especially when there is a lot of data. They allow us to see trends without being restricted by the bins used in histograms. Additionally, density plots are helpful in detecting the overall shape of the data distribution more clearly.
8. How does removing outliers affect the interpretation of the data?
Answer:
Removing outliers ensures that the density plots represent the majority of the data without being skewed by extreme values. This provides a clearer view of the typical distribution of crime rates across states, allowing for more meaningful analysis of trends and patterns.
9. What could cause an outlier in property or violent crime data?
Answer:
Outliers in property or violent crime data could be caused by several factors, such as a state with an unusually high or low population, significant local events (e.g., a major natural disaster or social unrest), or reporting inconsistencies. Such outliers need to be removed to avoid skewing the overall trends.
10. Why are violent crime and property crime analyzed separately?
Answer:
Violent crime and property crime represent different types of offenses with different societal and economic implications. Analyzing them separately helps in understanding the unique factors driving each type of crime. For example, violent crime typically involves more direct harm to individuals, while property crime affects material possessions, and these two types of crime might exhibit different trends across states.

CONCLUSION :

- Hence using the given plots we can conclude that the crime rate are relatively high in the states of Kentucky and Louisianna
- Furthermore using the line chart we can conclude that historically the state of California had the highest amount of crime this can be due to variety of reasons such as immigration, unemployment etc.
- Using the area chart we can also induce that the number of violent crimes have reduced due to increase in police routing, better government crackdown and regulation of drugs and guns etc