Assignment 1
Yuzhe Yang

Unemployment dataset
What information is this dataset showing? (Based on what you can deduct from the dataset itself)

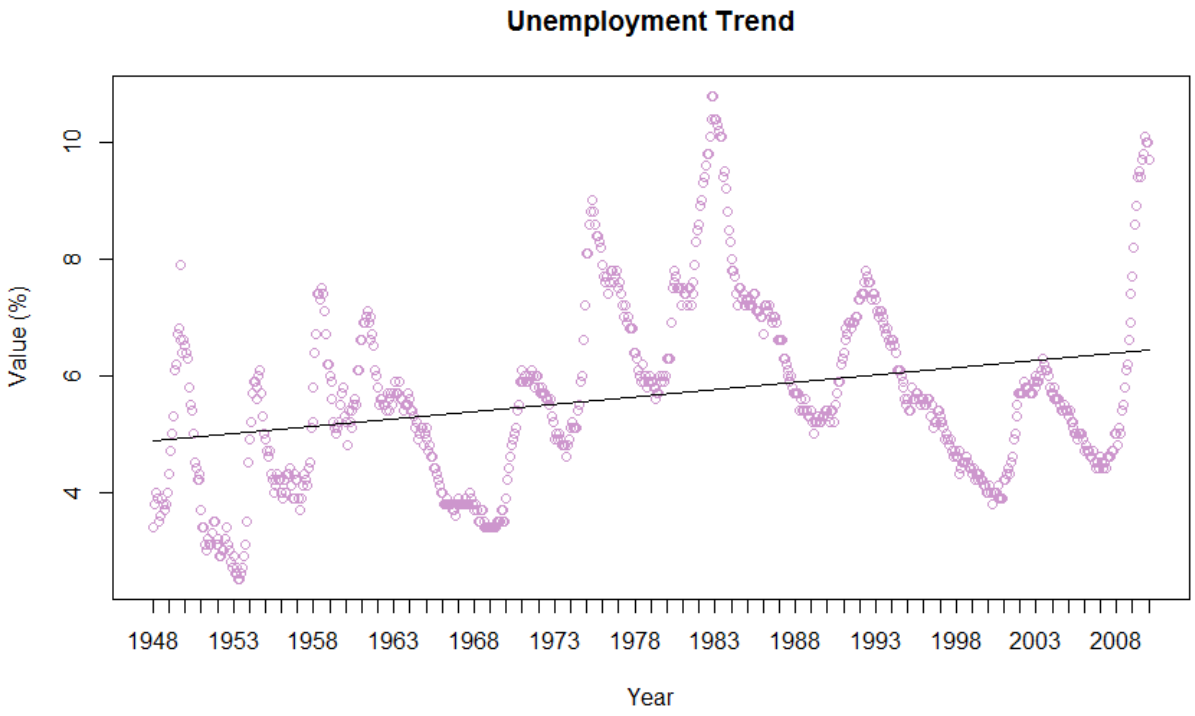
The dataset shows the unemployment rate of each month from 1948 to 2010.

What is the number and type [categorical,ordinal,quantitative] of attributes is shows?

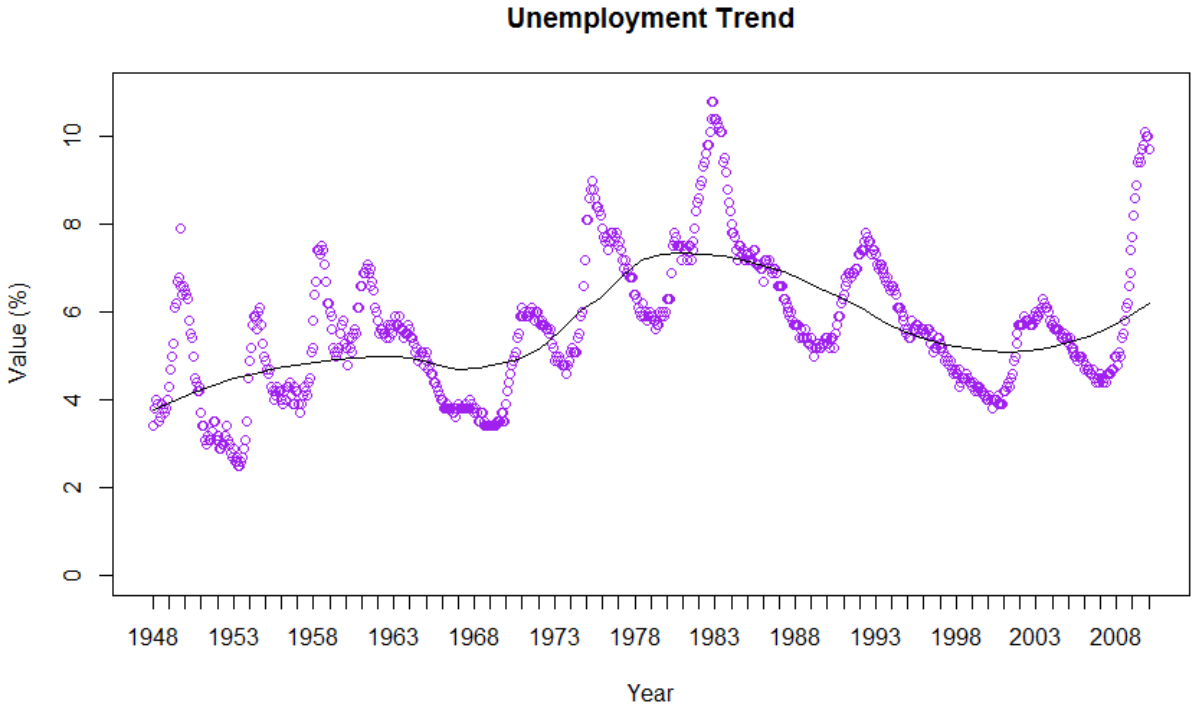
Serial ID is a categorical attribute
Year and Period are ordinal attributes
Value is a quantitative attribute

Viz idiom choice – we would like to show the number of people unemployed and how this number changed over time.

The unemployment rate of each month from 1948 to 2010 is plotted as a scatter plot. The scatter plot can be fitted with a straight line.



We can also fit the scatter plot with a LOESS curve which represents the data better.



Please refer to Rplots "Unemployment_Trend_linefit" and "Unemployment_Trend_smoothfit" in the "Folder Unemployment" folder.

Birth rate dataset

What information is this dataset showing? (Based on what you can deduct from the dataset itself)

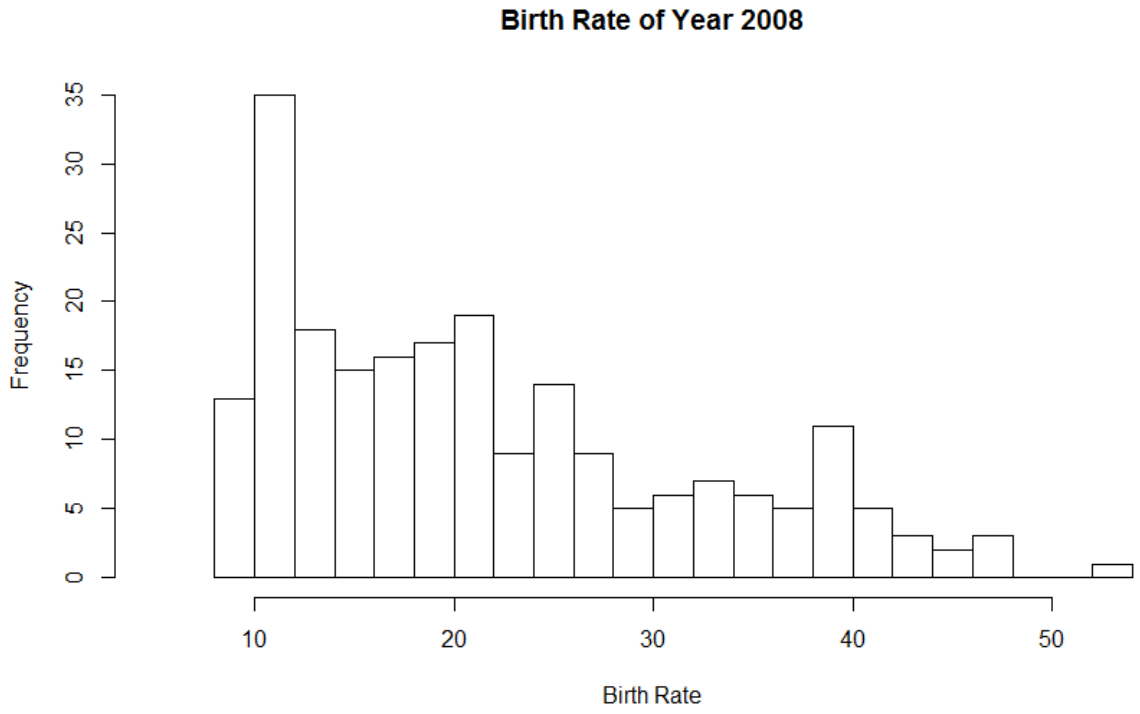
The dataset shows annual birthrates of different countries from 1960 to 2008.

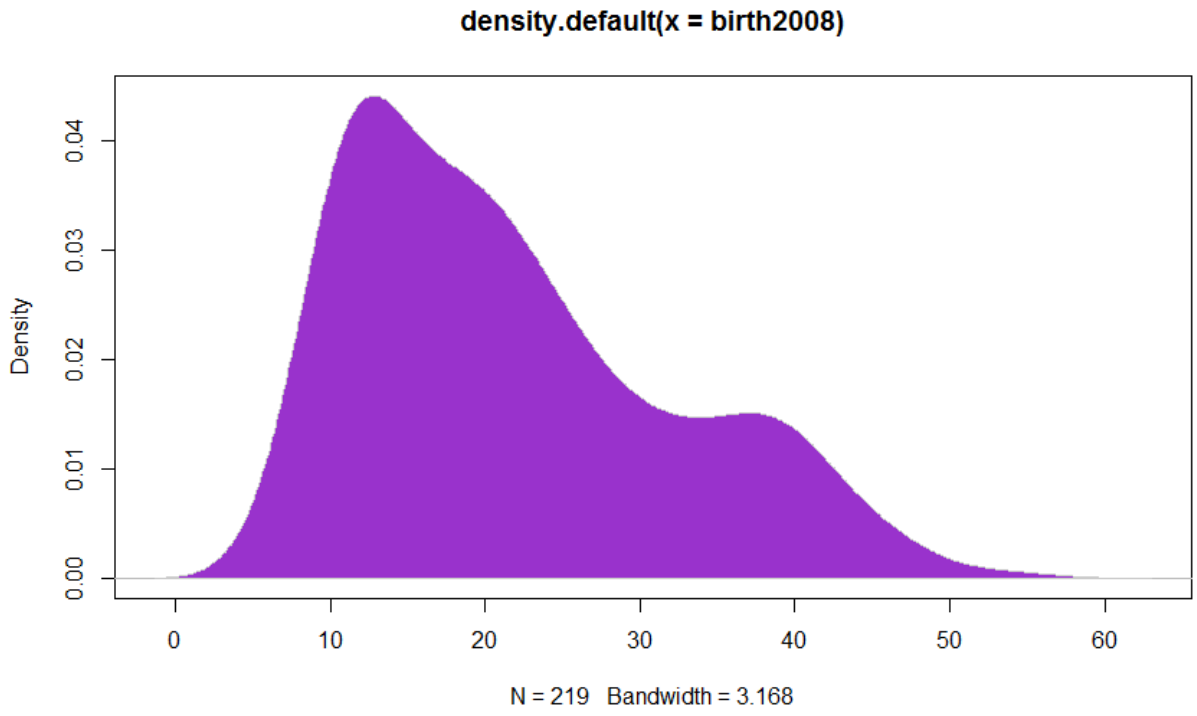
What is the number and type [categorical,ordinal,quantitative] of attributes is shows?

Country is a categorical attribute
Year is an ordinal attribute
Birthrate is a quantitative attribute

Viz idiom choice – we would like to show the distribution of newborns over the year of 2008

The distribution of newborns over the year of 2008 can be plotted into a histogram of birth rates or a density plot of birth rates as well.





Please refer to Rplots “Distribution_Newborn_2008_hist” and “Distribution_Newborn_2008_density” in “Birth Rate” folder.

Crime rate dataset

What information is this dataset showing? (Based on what you can deduct from the dataset itself)

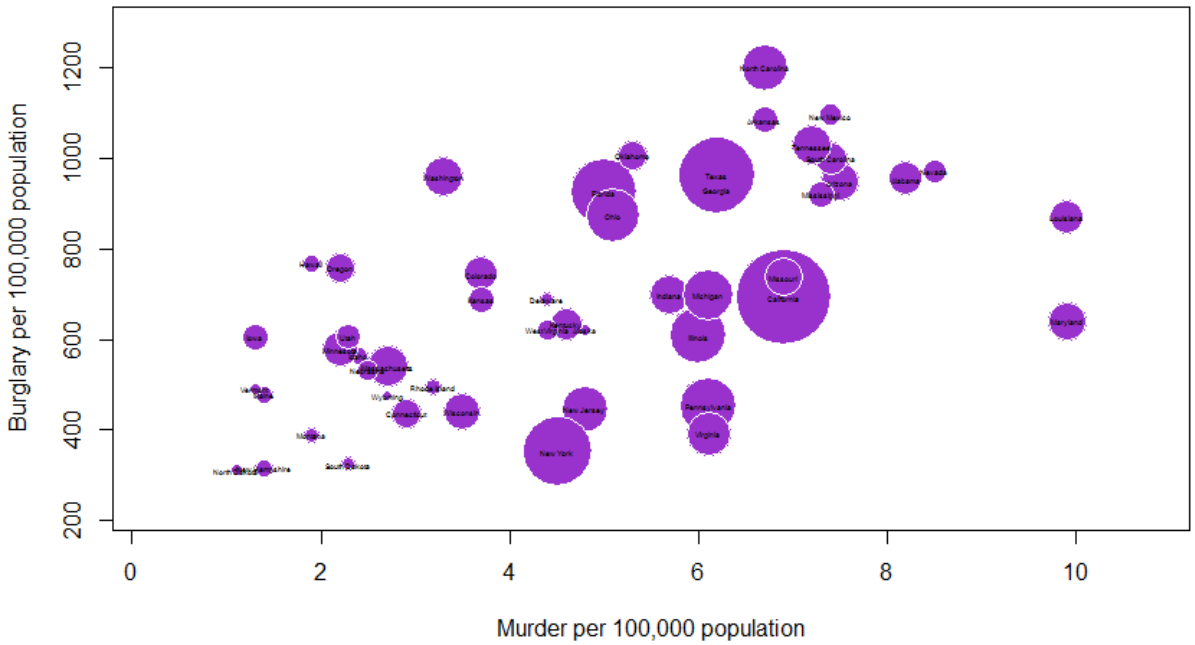
The dataset shows the number of different types of crime cases by U.S. states in the year of 2005 as well as the United States averages in that year. The dataset also provides the population information of each states and the country to help the readers to get incidences of crime.

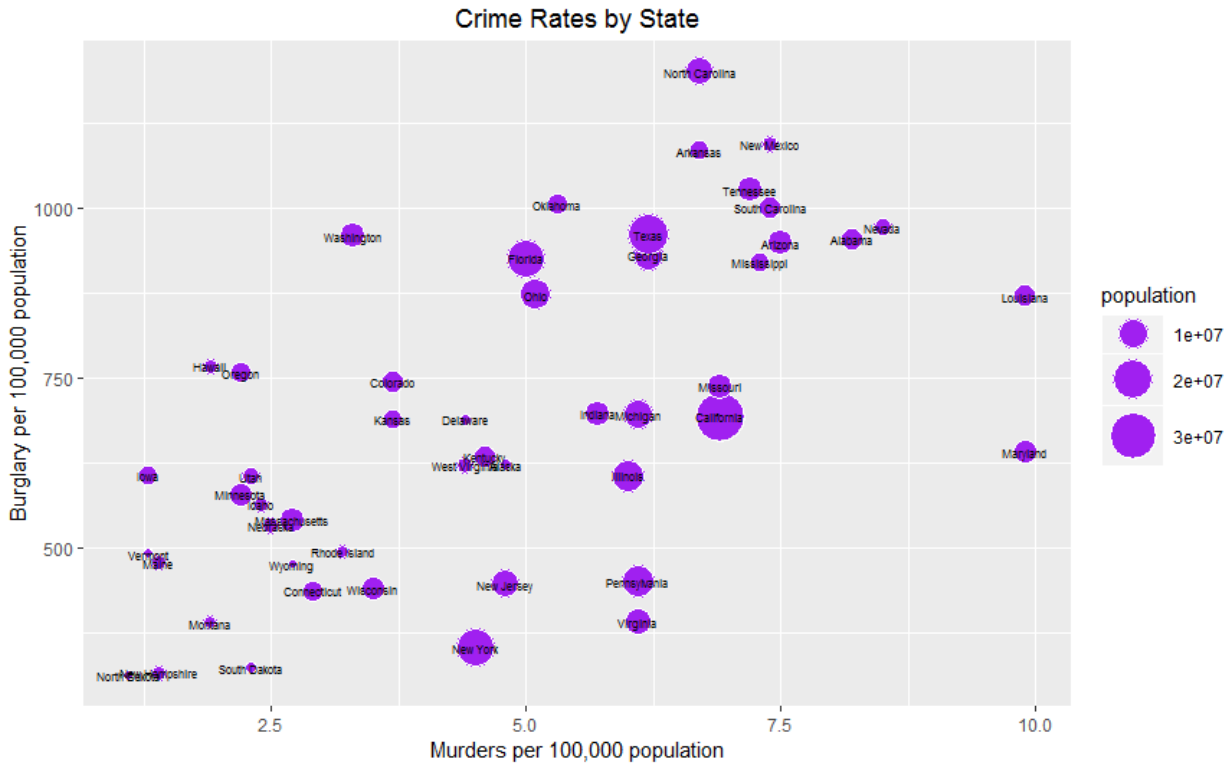
What is the number and type [categorical,ordinal,quantitative] of attributes is shows?

State is a categorical attribute
Types of crime are categorical attributes
Number of cases are quantitative attributes
Population is a quantitative attribute

Viz idiom choice – we would like to show, in a single graph, the relationships between the murder and burglary rate. We would like to show this in relation to the population of each place.

Since there are three attributes including murder and burglary rate and population of states, it is preferred to use bubble chart to visualize the relationships. The chart shows a positive correlation between murder rate and burglary rate. States that have higher murder rates tend to have higher burglary rates. The bubble chart can be drawn directly or with the ggplot2 library.





Please refer to Rplot “Crime_Rates_by_State” and “Crime_Rates_by_State_ggplot” in “Crime Rate” folder.