Use world bank data.

# Define the World Bank indicators

indicators = {

'EN.ATM.CO2E.PC': 'CO2Emissions',

'NY.GDP.PCAP.PP.KD': 'GDPpercapita',

'GB.XPD.RSDV.GD.ZS': 'Research',

'NE.TRD.GNFS.ZS': 'Trade',

'EG.FEC.RNEW.ZS': 'RenewableEnergy',

'NV.MNF.TECH.ZS.UN': 'MedAndHighTechInd',

'SP.URB.TOTL.IN.ZS': 'Urban',

'EG.USE.PCAP.KG.OE': 'EnergyUse'

}

# Set the date range for filtering after fetching the data

start\_year = 1990

end\_year = 2020 #

add Income levels

your resulting dataset should only countian data for countries, not aggregated units.

1. For each index, plot histograms. Discuss symmetry/tails and whether the distribution is unimodal or multimodal.
2. For indices with skewed distributions, apply log transformations and plot the histograms again. Did the transformation improve the symmetry?
3. Plot boxplots for each index by income level. Are there differences in the distribution across different income levels?
4. Run all pairwise scatter plots, find pairs of indices that have the strongest relationships, and discuss the nature of these relationships: monotonic, linear, or non-linear.
5. Create a bar plot for the income levels.
6. In the scatter plot of log-transformed CO2 vs log-transformed GDP, color the points by income level. Do you observe any patterns?
7. Repeat step 6, but now average both log-transformed CO2 and log-transformed GDP values over years. Do you observe any patterns?
8. Repeat step 6 again, but now average both log-transformed CO2 and log-transformed GDP values over countries. Do you observe any patterns?
9. Do you see any difference in patters from 7.8.9?

For log transformed CO2 emission

10. Calculate the average of it over countries, plot against the years Discuss the trend

11. which countries have the highest CO2 (top3), which ciontirs have the lowest CO2(bottom 3)

12 For each index, plot a scatterplot of that index against log-transformed CO2. If you log-transformed that index in step 2, keep it log-transformed here. Which indices appear to have a relationship with CO2?

13 For each index and each country, subtract the average of that index for the country from each observation (i.e., center the data for each country) and repeat step 12. Does the relationship between the indices and CO2 change after centering the data?

14 Repeat 13 but now remove averages by Year