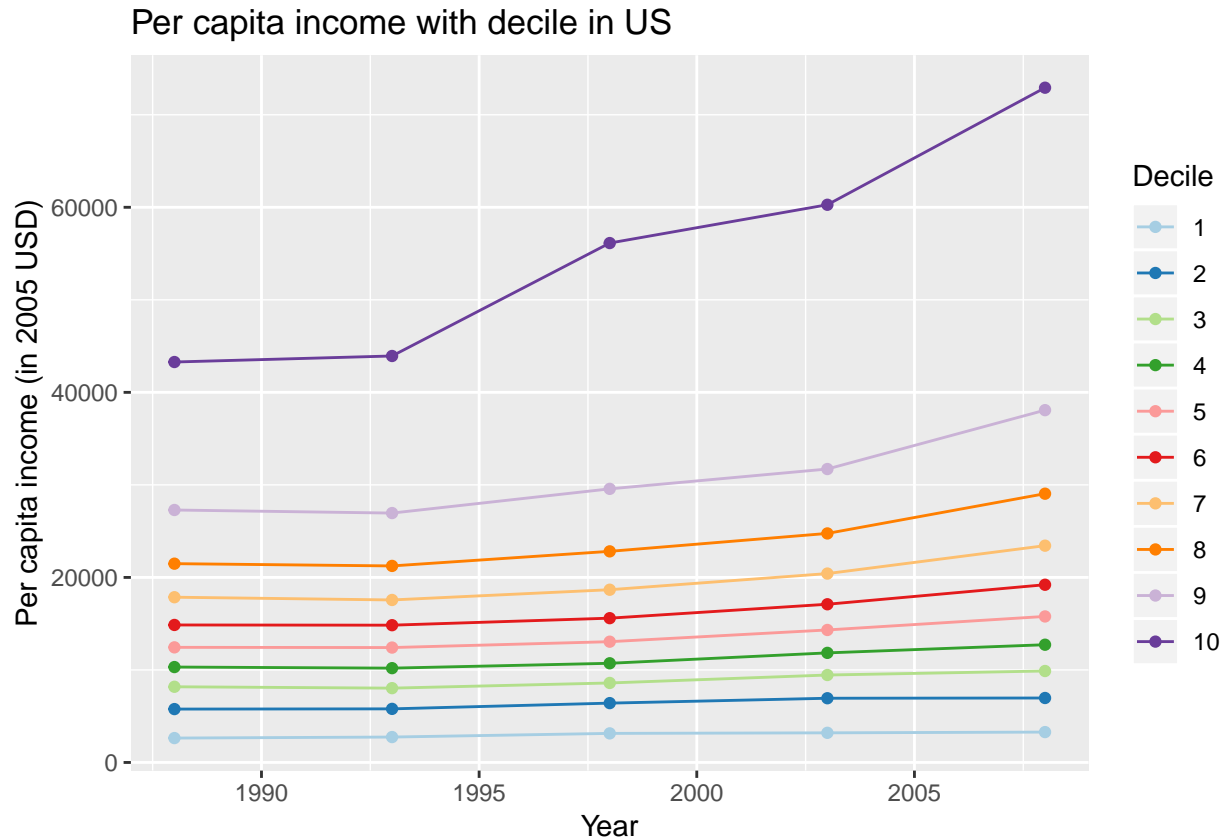# Mini project 1

Abhilash Kuhikar, Raj Thakkar, Darshan Shinde and Dhruuv Agarwal

February 24, 2019

1. How has per capita income for each decile in the U.S. changed since 1988?

Use the WPID data set to answer this. Give both graphical and numerical results, and describe in words both the differences in income level between the deciles and the differences in changes in income for each decile.

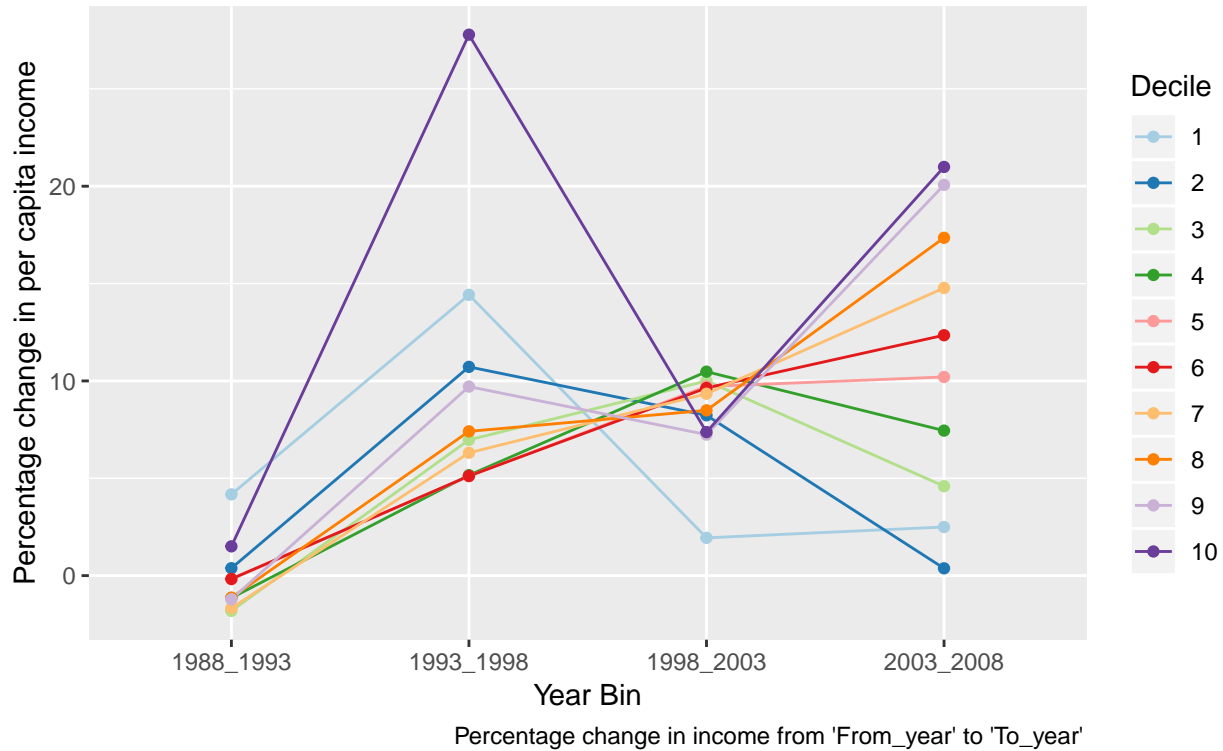Plotting the Income Levels for each decile

- Difference in income level between the deciles:

  We can see from the graph above that the income for the lower deciles have not changed much over the period compared to the higher deciles. For example, let's look at the decile '10' for which the income has increased significantly over the period of 15 years and the income for this decile is far above the other deciles. We can also observe that the amount of increase in the per capita income is directly proportional to the decile number. In other words, the income for the lowest decile has increased the least and the income for the highest decile has increased the most over the entire period. We can see the general trend that the gap in income between the rich and the poor is increasing.

- We see that the growth of income has been maximum for the 10th decile and lowest (almost flat line) for the 1st decile.

Per capita income change with decile in US
Example: 1988_1993 indicates the change in income from 1988 to 1993

Percentage change in income from 'From_year' to 'To_year'

- The differences in changes in income for each decile:

  The percentage change is calculated as:

  $$percentage\_change = \frac{current\_year\_income - previous\_year\_income}{previous\_year\_income} \times 100$$
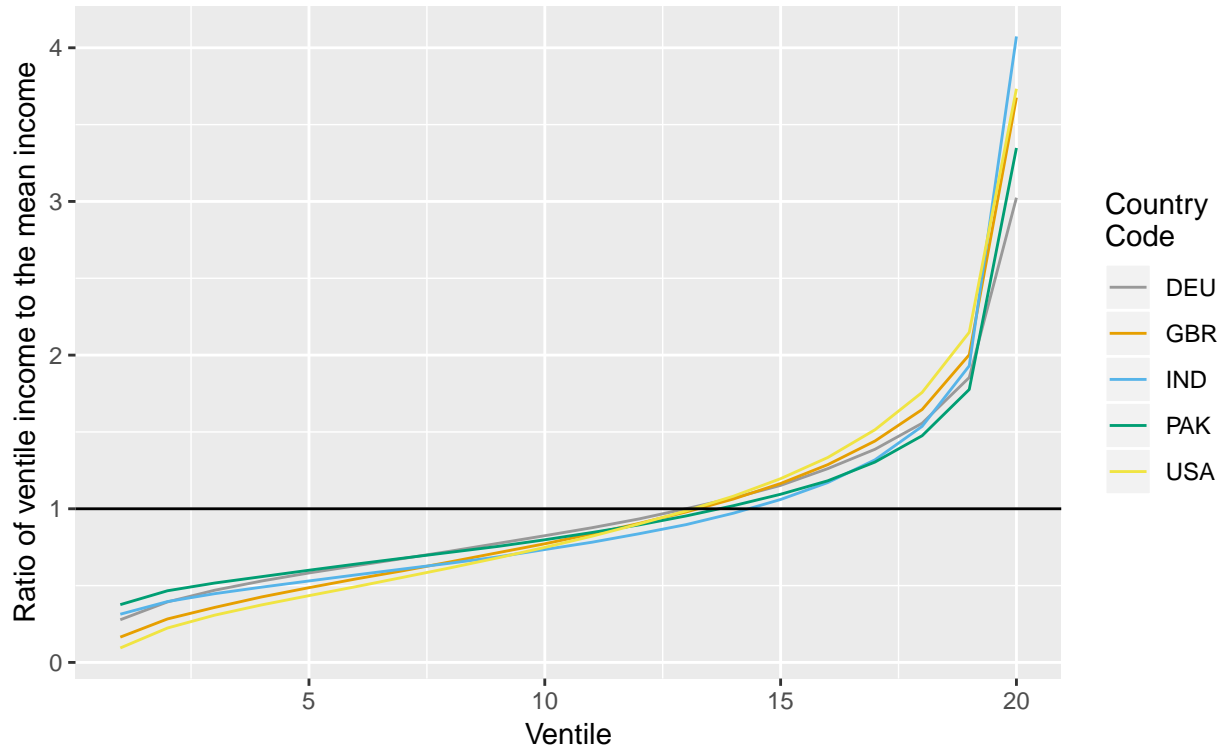
  The above graph precisely points out the percentage change in income for the given year bins. For example, the negative values for certain deciles for the year bin 1988-1993 indicates that the income has been decreased for these deciles decreased in 1993 as compared to 1988.

  We can observe from the graph that the percentage change values after 1993 are always positive for all the deciles. This means that the change in income has always been positive indicating that the per capita income for all the deciles is monotonically increasing after 1993. We can observe the steepest increase in percentage change for the decile 10 from the year 1993 to 1998.

  The rate of change of the income increases with the increase in the decile number.

2. How does the present distribution of income, relative to a country's mean, differ between selected countries?

Distribution of countries' incomes relative to their respective mean incomes
Horizontal line represents the income where ventile income = mean income

- We can observe from the above graph that in general, the distribution of income relative to mean is almost similar for all the five countries over the ventiles. We can also observe that the incomes of ventiles 1 to 12 are below the mean income for all the countries. This means that more than 60 percent of the population has income below the mean income of the country. We will now take a detailed look at the distribution of each country.
  RVM: ratio of the ventile income to the mean income of that country
  Germany(DEU): The 20th ventile income is about 3 times the mean income of the country. This ratio of highest ventile income to the mean income of the country is lowest for Germany suggesting us that the gap between the income of the highest ventile and the mean income is lowest for Germany and hence their economy is rather less skewed compared to other 4 countries.
  United Kingdom(GBR): The distribution of income is almost similar to that of United States. The ventile incomes of about first 13 ventiles is below the mean.
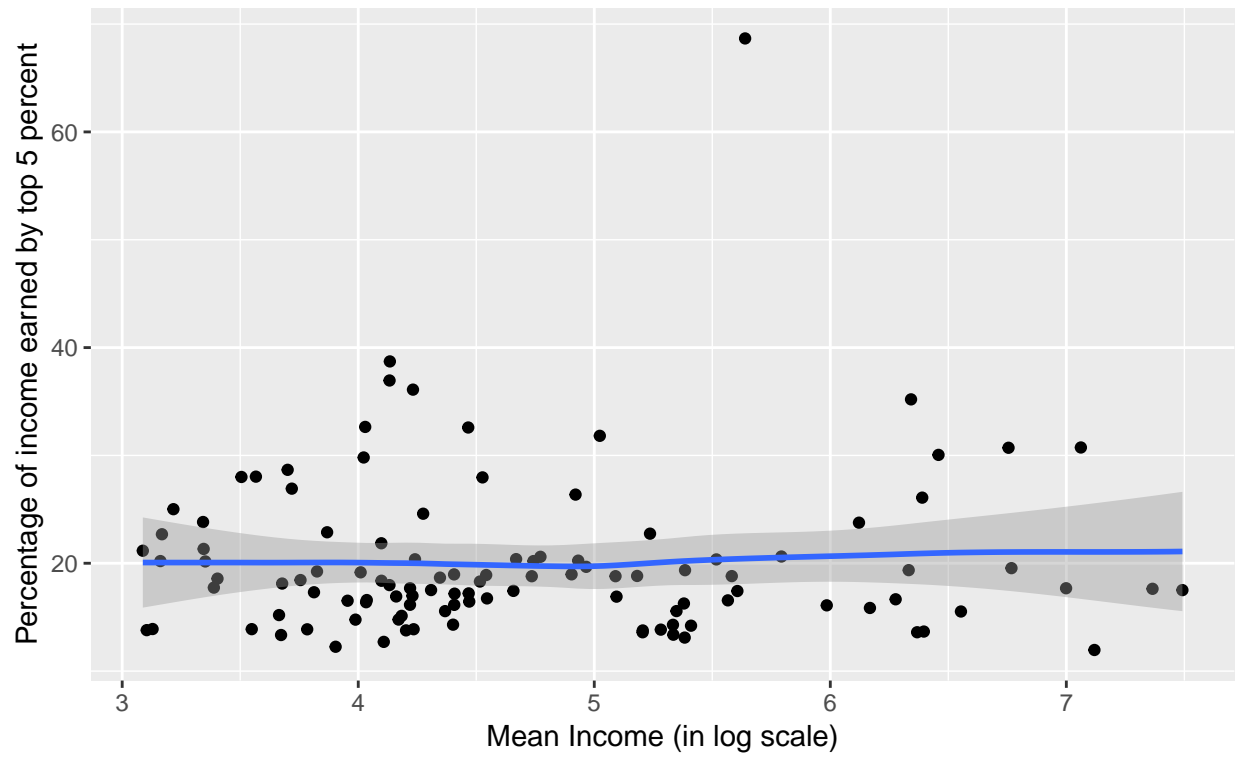 **India(IND):** The income for highest ventile is about 4 times the mean income. This RVM for 20th ventile is the highest amongst all the countries. We can also observe that about 14 ventiles are below the mean income. This means that India has the highest percentage of population amongst all countries below the mean.
  Pakistan(PAK): The RVM for ventile 1 is closest to the mean compared to the other 4 countries. About 13 ventiles are beow the mean income.
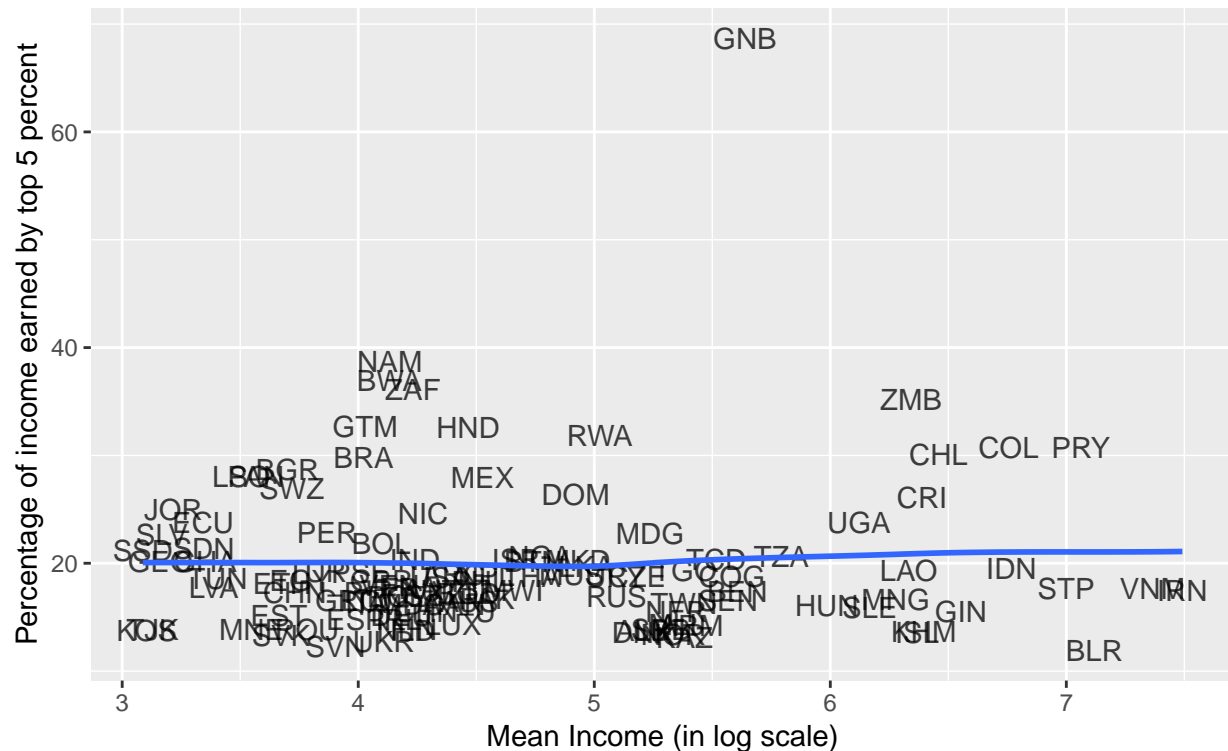 **United States(USA):** The RVM for ventile 1 was found to be the farthest from the mean amongst all the countries.

3. Is the percentage of income earned by the top 5 percent in a country related to mean income in that country? What about the percentage of income earned by the bottom 5 percent?

Relationship between percentage of income earned by top 5 percent and m[e]

Mean income represented in log scale

Percentage of income earned by top 5 percent v/s mean income
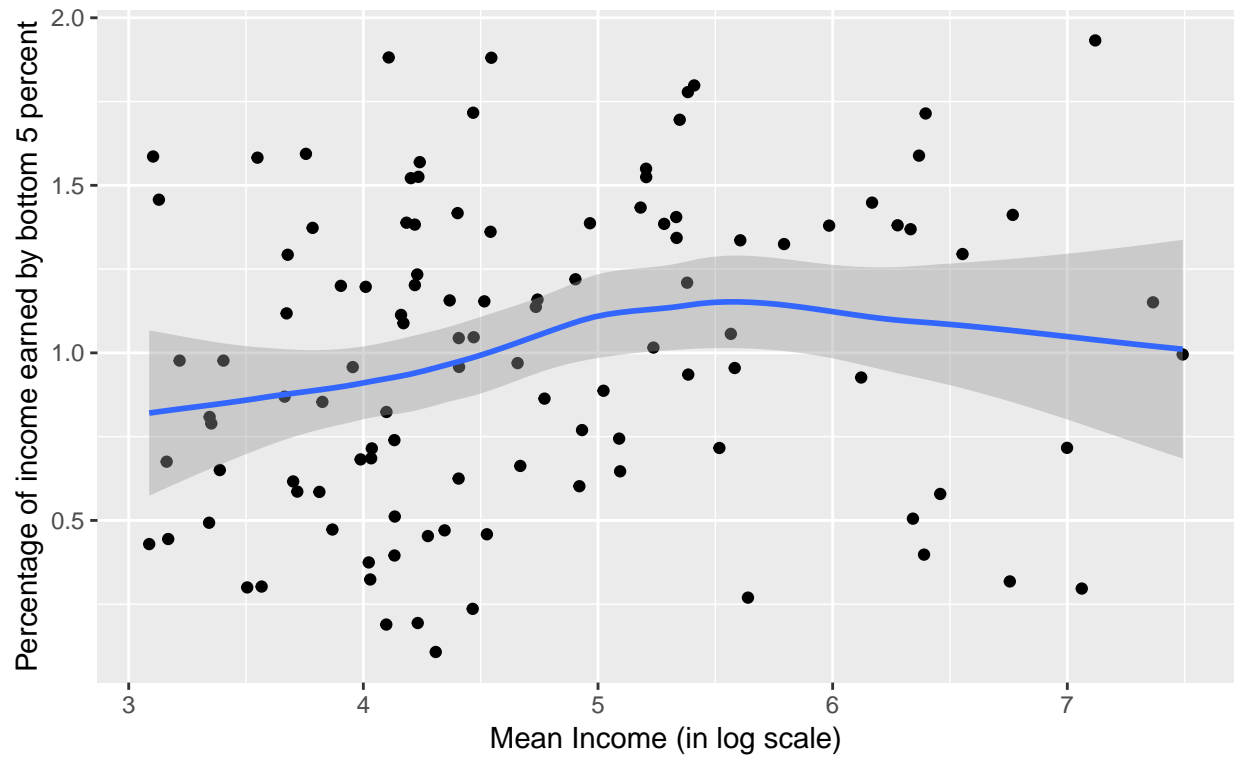### Mean income represented in log scale

Fom the graph above, we can say that there is no relationship between the percentage of income earned by the top 5 percent in a country and mean income in that country. One of the major outliers is country Guinea-Bissau (country code = GNB) which is located in Africa in which the top 5% of the population earns 60% of the total income earned in that country.

Other countries which are farther from the fitted lines are Zambia (country code = ZMB), Namibia (country code = NAM) etc. which are located in the African sub contitnent. And countries such as Chile (country code = CHL), Colombia (country code = COL), Paraguay (country code = PRY) etc whic are located in the South American subcontinent.

Thus we can say that countries from Afrcian and South American sub contitnent have the trend that top 5% of the population earns more than 20% of the country's total income.

Percentage of income earned by bottom 5 percent v/s mean income

Mean income represented in log scale

Percentage of income earned by bottom 5 percent

Mean Income (in log scale)

Percentage of income earned by bottom 5 percent v/s mean income
Mean income represented in log scale

Fom the graph above, we can say that there is a non monotonic relationship between the percentage of income earned by the bottom 5 percent in a country and mean income in that country (It increases and then decreases). There are no major outliers observed in this graph.

Bottom 5 percent in Austria (country code = AUT) have the least contributuon in the total income earned by the population in the country where as the highest contributuon is by the bottom 5 percent population of Belarus (country code = BLR), Pakistan (Country code = PAK) and Ukraine (country code = UKR)