

**Data Analytics**

**Assignment - 2**

Dharmesh Harsha (ESD15I018)

Akshay Kumar (CED15I031)

Vivek Khatua (MFD15I014)

### **Descriptive Statistics:**

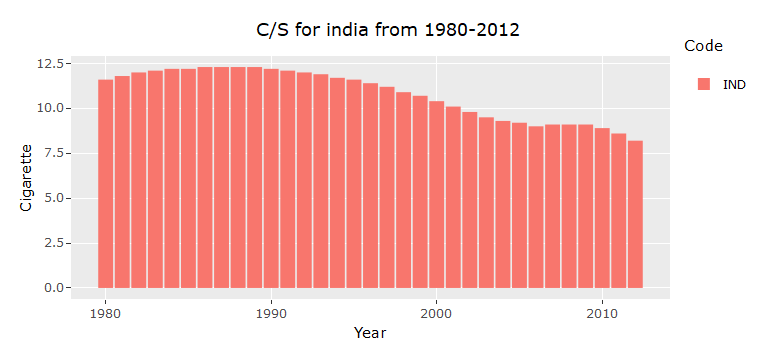
Descriptive Statistics is about quantitatively describing or summarizing the aspects of a collection of information. Distinguished from inferential statistics, there in no prediction/learning in Descriptive Statistics rather just visualisation and summarizing of a sample.

The descriptive statistics techniques we have used of our analysis are :

* Bar Chart
* Box-Whisker Plot
* Pie Chart
* Histogram
* Scatter Plot
* Violin Plot

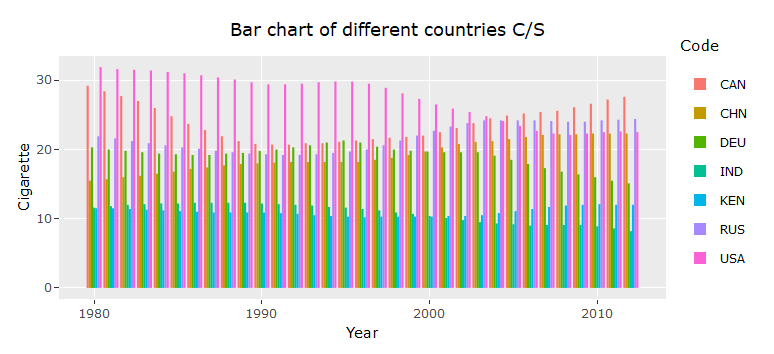
### **Analysed Features :**

1. **Cigarettes per Smoker in India for last 3 decades :**

****

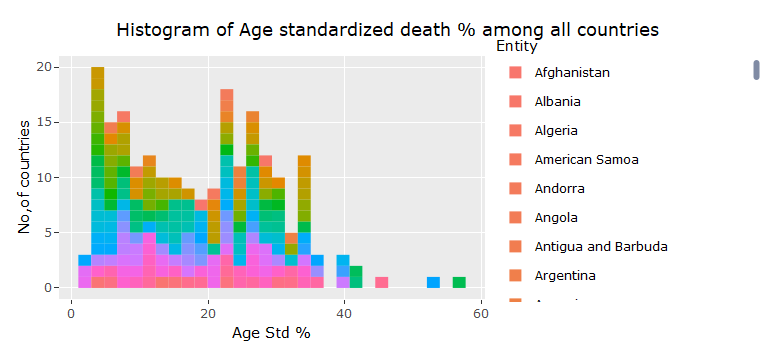
* 1. The above histogram summarizes that the number of cigarettes per smoker has gradually decreased after 1990s.
  2. After 2008, the chart shows a subtle increase in the smoking rate which again has been in control after 2010.
  3. In 2012 the smoking rate has decreased by 25% as compared to the peak observed around 1986.

1. **Comparison of Cigarettes per Smoker over years in various countries :**

****

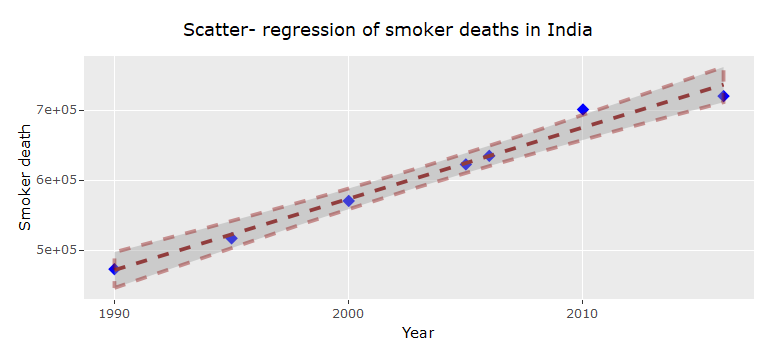
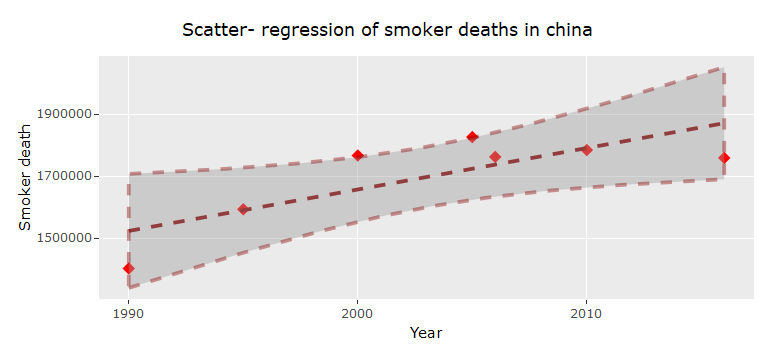
* 1. The Cigarettes per smoker has started to decrease at a double rate in USA as compared to the trend in 1980s.
  2. India and Kenya hasn’t seen much change as compared to other countries in this aspect.
  3. The smoking rate in Canada has increased after 1990 while for other nations its been decreasing.
  4. China hasn’t seen any decrement in the C/S ratio in the last 3 decades.

1. **Age Standardized Death % among all countries in 2010 :**

****

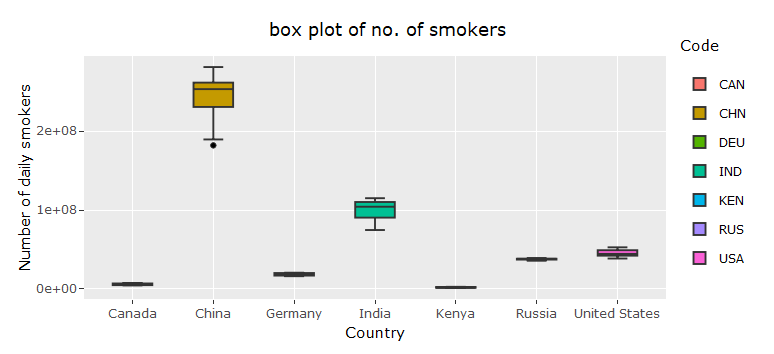
* 1. We can see that in many African Nations the age standardised death is 2-6 years which says there is huge impact of passive smoking.
  2. The Age group of 22-26 which is where most death happen in nations due to youth getting into chain smoking.
  3. There are few outlier countries with age standardized deaths above 40.

1. **Death due to smoking in China and India over the last 3 decades :**

****

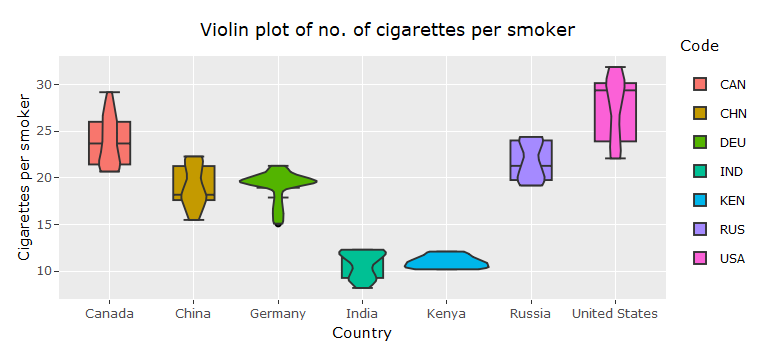
* 1. With every year the deaths in India due to smoking is rising steeply.
  2. As compared to India the deaths in China have been high while the increase rate of this factor has been less.
  3. Taking in concern the population of both nations, we observe that larger population% is getting affected in India.

1. **No. of daily Smokers in various countries :**

****

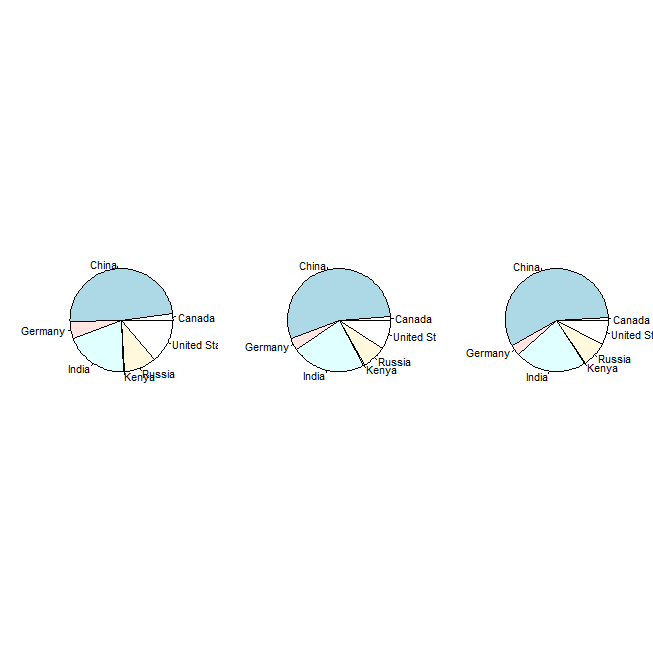
* 1. By the width of Boxes of China and India, we can observe that the no. of smokers had a huge variation in these nations over the last 3 decades.
  2. While the other nation’s saw less change in the number of daily smokers, China saw a vast change with also a higher minimum and maximum which might be due to improper law enforcement and lack of awareness.
  3. The very small span of Box for United States says they have had proper awareness that the increase was very less.

1. **Cigarettes per smoker in various countries over 1980-2012 :**

****

* 1. These violin plot of various countries over the last 3 decades show the change in Cigarette per Smokers that they have observed.
  2. Where we can see that compared to all counties considered here, India has the lowest C/S factor.
  3. The violin of Germany and Kenya having sudden wide region shows that they have seen a no change in number of cigarettes per smoker. This might be due to various reasons.

1. **Overall contribution of Smokers to the world smokers count (1998 ) :**

****

* 1. We can see that the nation China has been been contributing highly to the global smokers count.
  2. On the other hand we see countries that have decreased their share in world smoker count.
  3. India seems to be having the constant share towards world smoker count.

### DataSet Source :

* https://ourworldindata.org/smoking