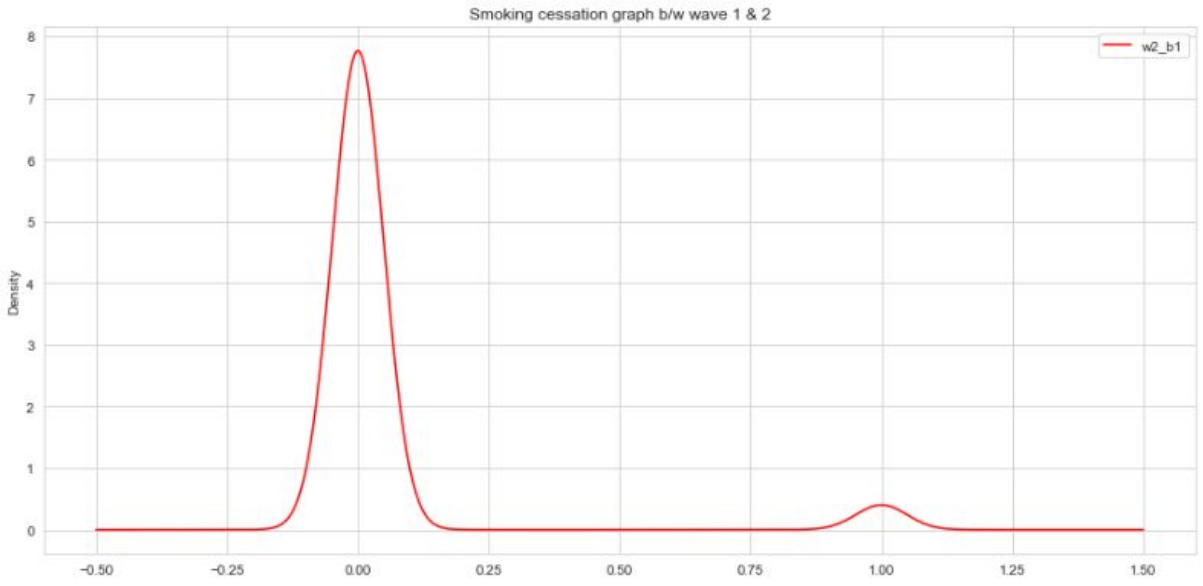
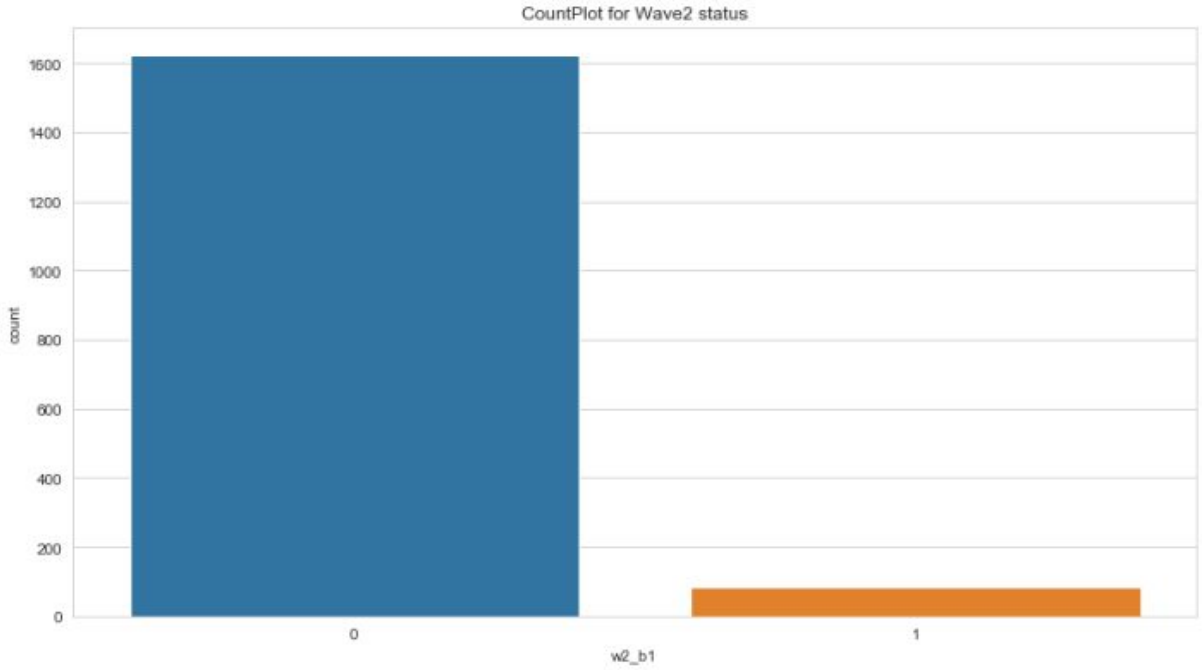
Firdale assessment task

1) **Data of 73 people is absent** in the dataset for wave 3 smoking type(**w3_smk_type**) i.e Data shows lack of consistency and the dataset requires some cleaning for better monitoring and evaluation. Along with this, **there's also discrepancy in age data as for 75 people, their w2_age> w3_age & for 44 people w2_age<w1_age which is possibly wrong** if the waves are carried out in a consecutive fashion i.e w2 after w1 & w3 after w2.

2) a) The smoking cessation rate gives insights on whether a person quit smoking or not between the interval of wave 1 & wave 2 i.e whether he/she showed any significant change. On the basis of data analysis, 84 people out of 1706 quit smoking in between wave 1 & wave 2. It gives a **smoking cessation percentage of 4.9237983 i.e out of every 100 people(irrespective of race,gender,geopolitical location) 4.93 people recovered and quit smoking in wave 2.**

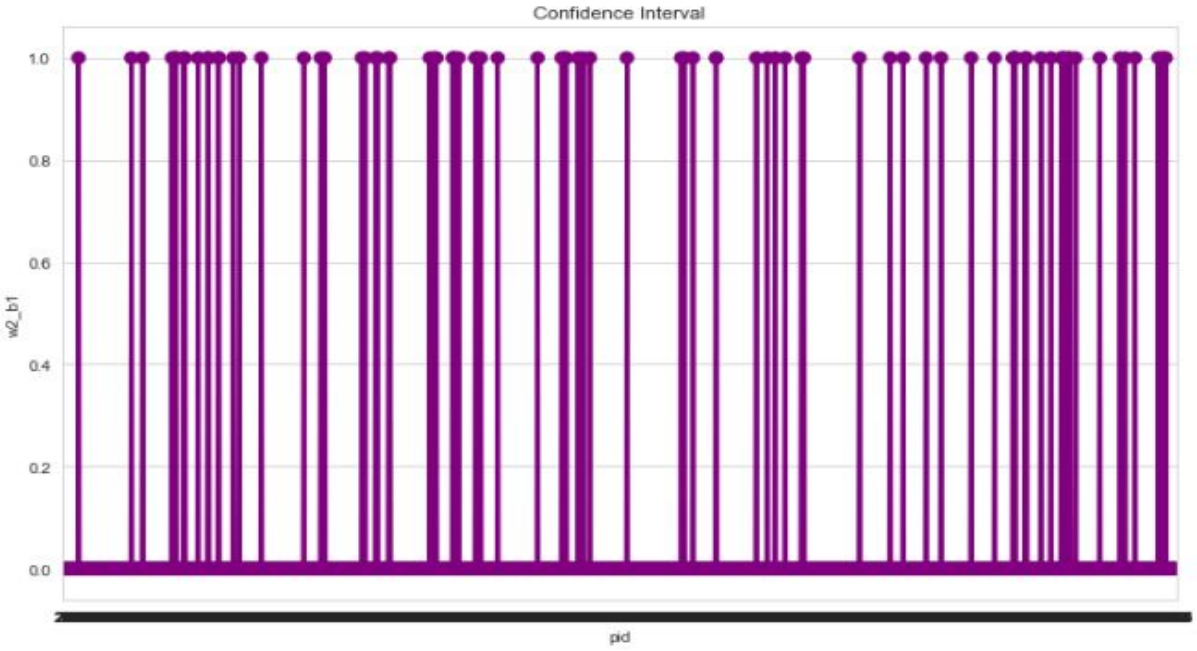


This graph shows the **kernel density** i.e shows the dense regions. Here, **Zero represents not quitting smoking & 1 represents stoppage in smoking.**



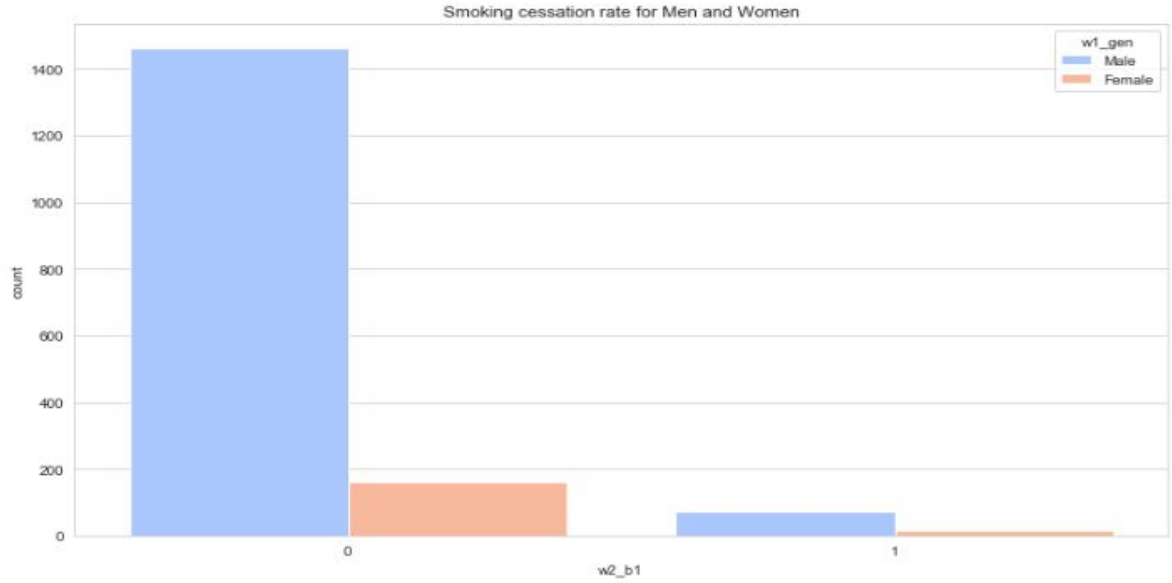
This Countplot is showing the no of people who quit smoking and the ones who don't between **interval(wave1 to wave2)** and it can be clearly observed that no of people who quit are way less than those who don't. It also shows the incompetence of the wave2 campaign in recovering more people.

The Confidence interval graph displays the extent to which a result can be believed.
A confidence interval doesn't quantify variability.



This confidence interval is connecting all possible values for quitting and not quitting smoking i.e 0 & 1.

2) b) On analyzing the data it can be concluded that both Men and Women respond differently to the treatment in waves 1 & 2 but women show a better recovery trend in comparison to men. **Total no of women & men who participated in wave2 were 176 and 1530 and out of it 14 women and 70 men recovered respectively which gives a smoking cessation rate of 7.954% for women and 4.575% for men**. It tells that the **wave2 campaign planning is more successful for women and less for men.**

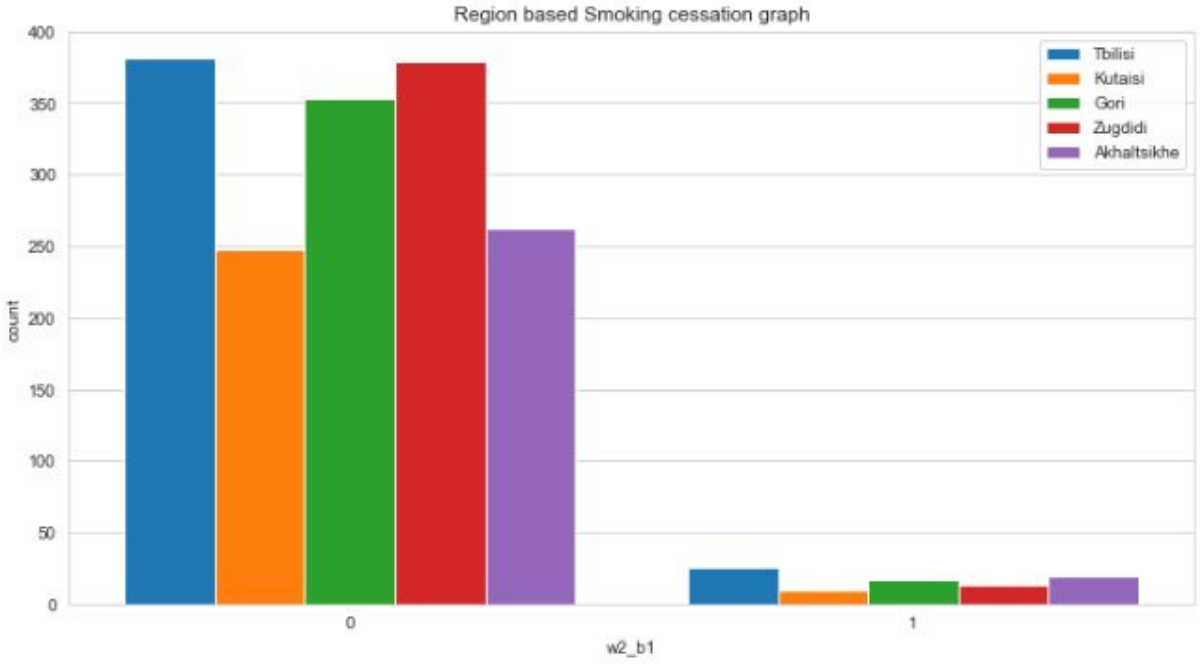


Graph displays the no of cases of both quitting and not quitting smoking for both men and women separately and all concluded points can be verified from the graph.

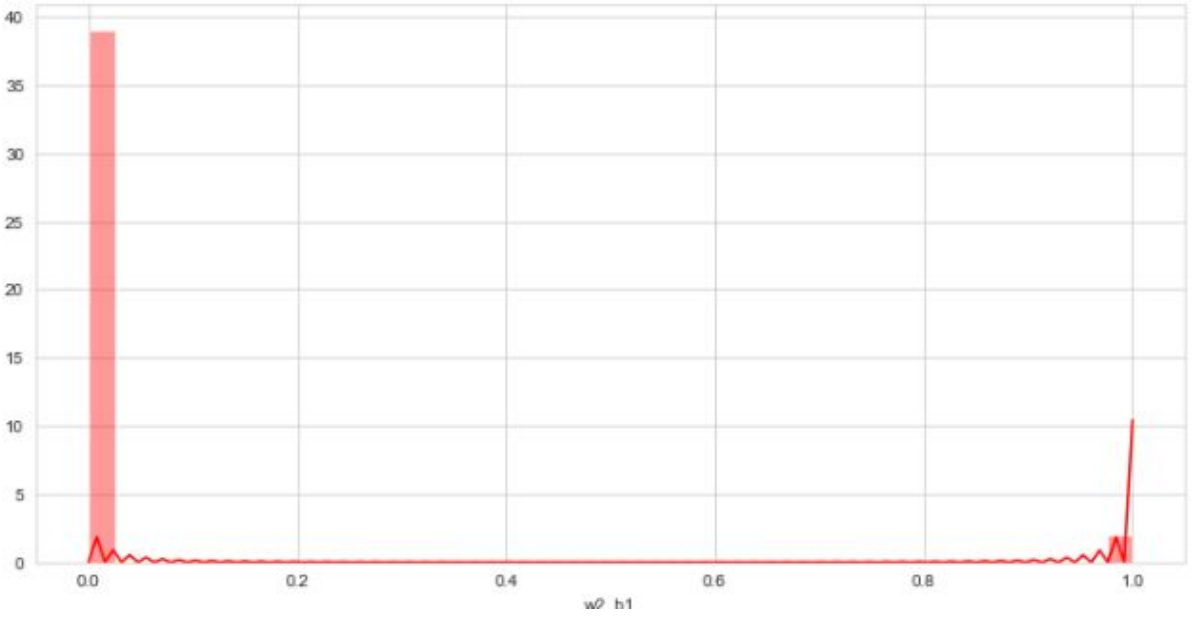
2) c) Different regions respond differently to the wave2 and display different metrics. On analyzing the dataset, different regions show different recovery patterns.

1) ***SC_tbilisi_rate= 6.157635467980295%***
2) ***SC_Kutaisi_rate=3.8910505836575875%***
3) ***SC_gori_rate= 4.594594594594595%***
4) ***SC_Zugdidi_rate=3.3163265306122 45%***
5) ***SC_Akhaltsikhe_rate=6.761565836298933%***

Out of all regions, **Akhaltsikhe region showcase the best response** to smoking with A Rate of 6.761565836298933% & **Zugdidi shows worst response** with a %rate of 3.316326530612245% .



This Graph represents all of the findings after assessment of data and proves the results.

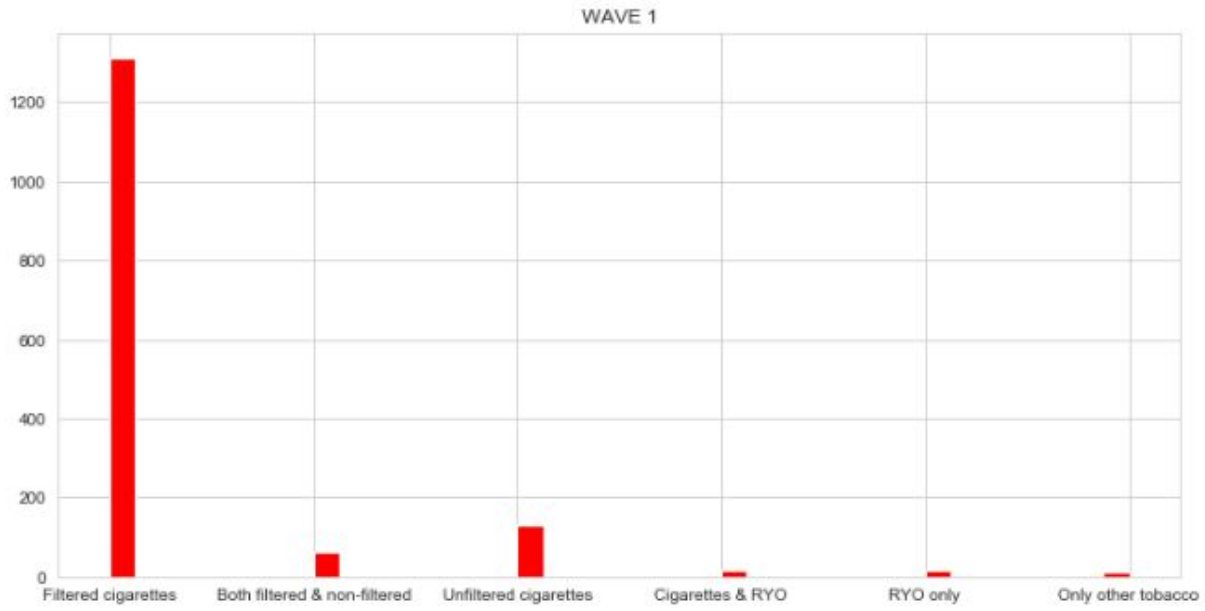


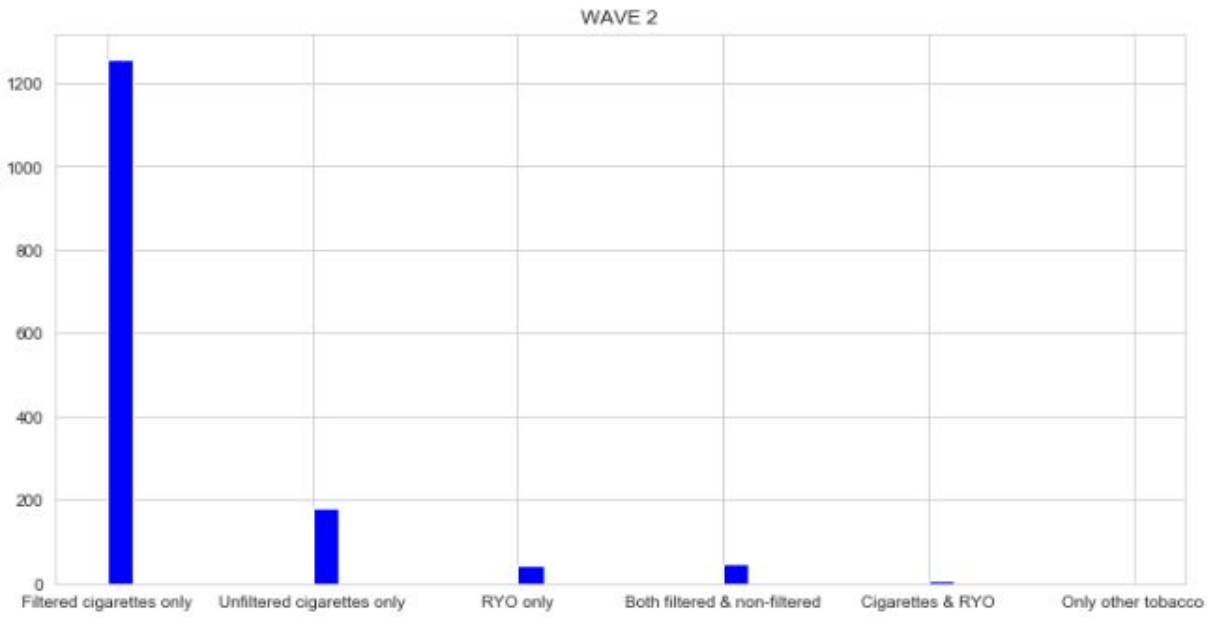
3) On calculation,there are precisely **84 people** that gave up smoking in wave 2 but started smoking again in wave 3.
4) Smoking pattern of people changes over different waves.every wave is displaying some unique trend and gives different insights.

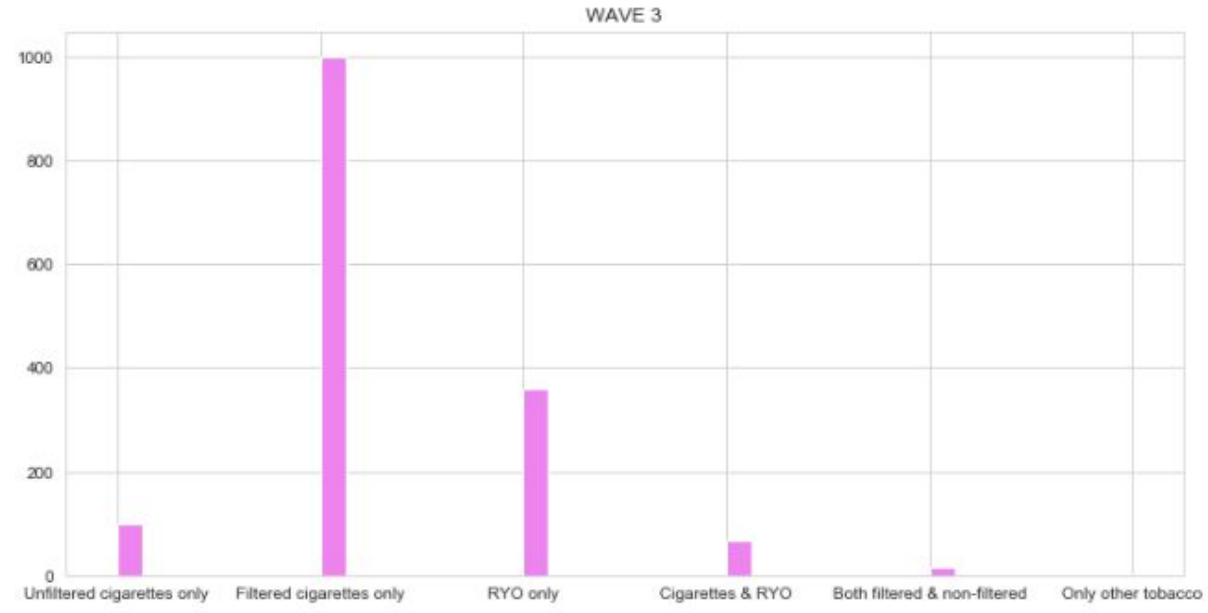
On monitoring,Filtered cigarettes are the most popular choice among people because of it's easy availability and smooth taste. After that, there are mixed choices which are famous among people. A very small portion of people smokes RYO and the likability of "other tobacco" is very less and sounds more like an uncommon idea among people.

The number of cases where **'other tobacco'** gets consumed **reduces significantly** as different waves pass by. It shows a significant proportion in the first wave but becomes almost **zero to nil in the 2nd to 3rd wave**.On the other hand, **unfiltered cigarettes consuming people increases in the 2nd wave but reduces significantly in the 3rd one**. The **market proportion of filtered cigarettes reduces to some extent** which shows the **competence of waves into eliminating easily available drugs from people's lives. RYO consuming people displays a gradual growth in 2nd wave and a drastic growth from 2nd to 3rd wave, increasing the no of RYO preferring people.**

Graphs are added for factual assessment of dataset and to prove the concluded points.







All points are concluded and graphs are plotted only after doing a logical and factual assessment of the dataset through data analytics technique.