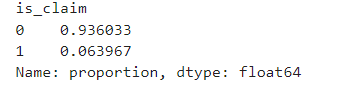
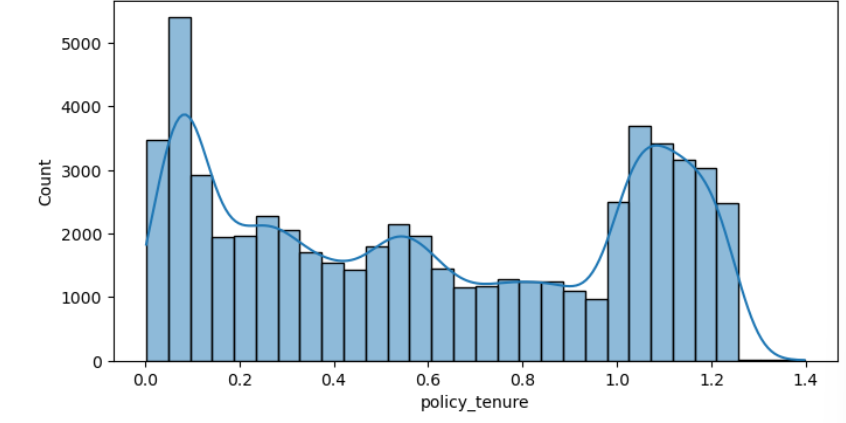
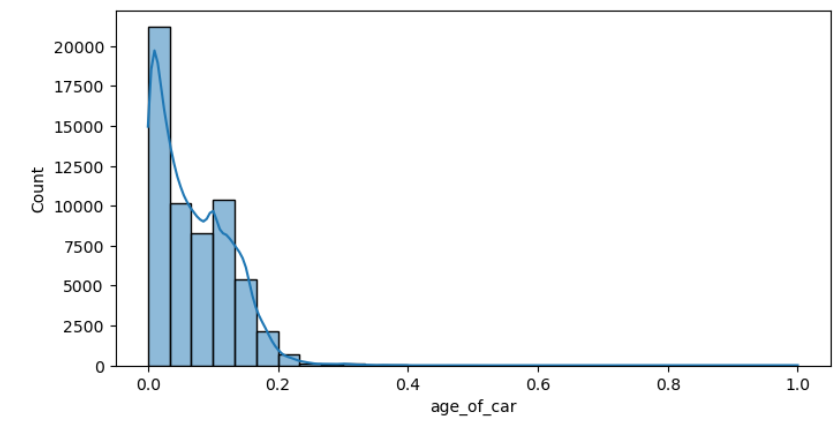
General Observations:-



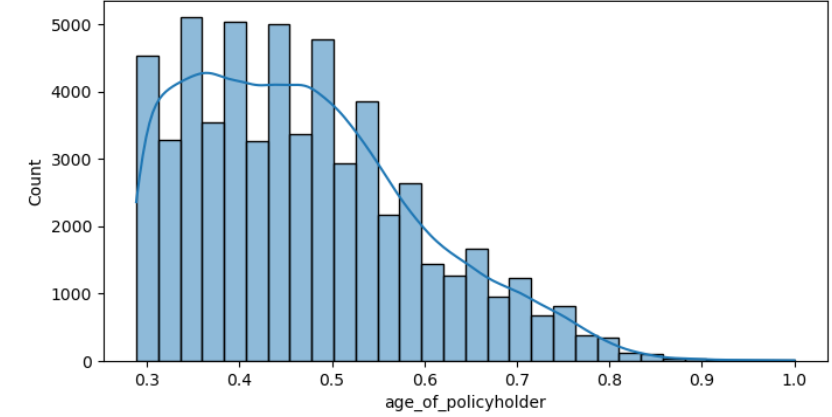
1. According to the provided Dataset, **93% of the people don't claim insurance**, while **7% of the people claim insurance.**



1. There is a significant peak at the beginning of the policy tenure range (0.0467 - 0.0933). **This indicates that a substantial number of new policies have been added, suggesting a considerable influx of new customers**. (histogram-policy\_tenure)[Number of Bins=30]
2. **The presence of two significant peaks suggests a bimodal distribution, indicating two distinct groups of policyholders based on policy tenure.** One group prefers shorter-term insurance policies, while the other opts for longer-term commitments. The **notable peak around the policy tenure of 1.0 to 1.2.**



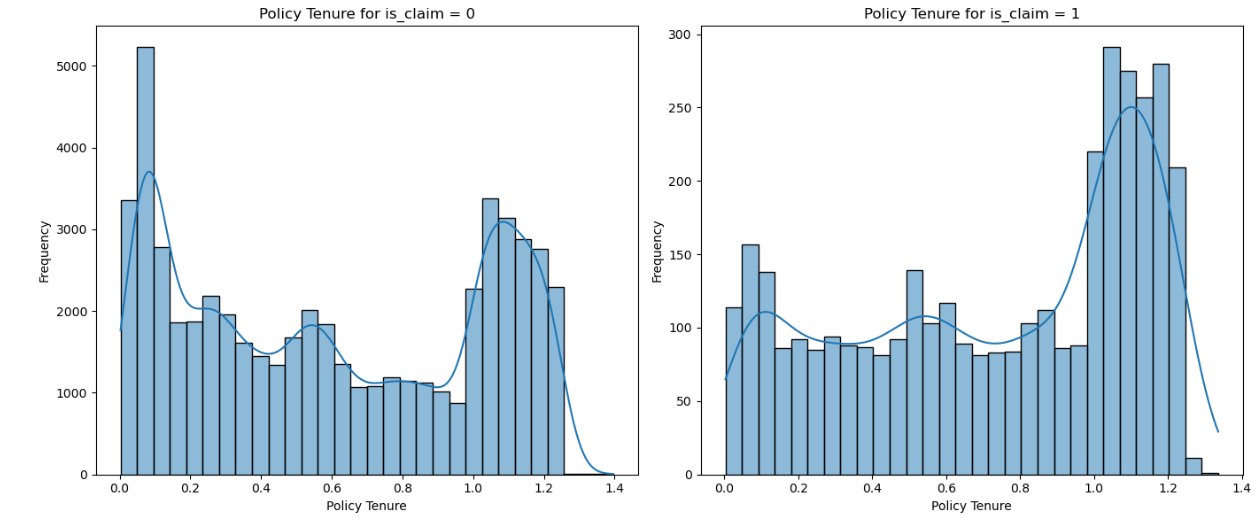
1. The **peak of the distribution is near zero,** indicating that a significant number of new insurance policies are **purchased for newer cars**.
2. **As cars age, fewer people tend to opt for insurance policies**. Generally, individuals are more likely to **forgo insurance for older vehicles.**



**Individuals in their late 20s to late 50s or early 60s as we can see in the histogram are purchasing insurance in this class because they represent the working population**, possessing both the purchasing power **and the need for a vehicle.**

In C8, there are most number of people purchasing insurance policy, it has a lower population density but its really

In C2 to C3 its moderately popular there is a gradual decrease in popularity



Insights from Policy Tenure Distribution by Claim Occurrence

From the provided histograms of policy tenure distributions separated by claim occurrence (is\_claim = 0 and is\_claim = 1), we can derive several insights:

Different Distributions for Claims vs. No Claims:

The distribution of policy tenure for is\_claim = 0 (left plot) and is\_claim = 1 (right plot) shows distinct patterns. This indicates that the tenure distribution is different for those who filed claims compared to those who did not.

High Frequency of Short Tenures Without Claims:

For is\_claim = 0, there is a significant peak at the beginning of the policy tenure range (around 0.0). This suggests that a large number of policyholders do not file claims within a short duration of their policy. The frequency decreases as the policy tenure increases.

Longer Tenures with Claims:

For is\_claim = 1, the distribution shows a noticeable increase in the frequency of claims for policy tenures around 1.0 to 1.2. This suggests that policyholders are more likely to file claims after having the policy for a longer duration.

Multiple Peaks:

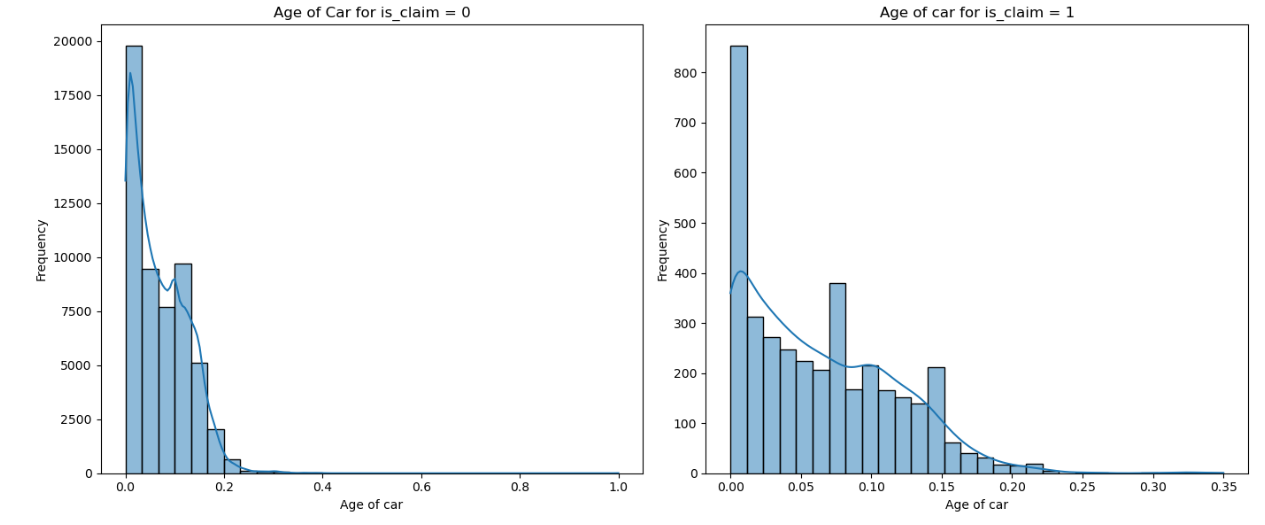
Both plots exhibit multiple peaks. In the is\_claim = 0 plot, there are peaks around 0.0, 0.2, and 1.0, while in the is\_claim = 1 plot, peaks appear around 0.1, 0.8, and 1.1. These peaks might indicate specific renewal periods or other factors influencing policy tenure.

Policy Tenure and Claim Occurrence Correlation:

The comparison of both plots suggests a potential correlation between policy tenure and the likelihood of filing a claim. Longer policy tenures seem to have a higher frequency of claims.

Lower Frequencies for Mid-Range Tenures:

Both plots show lower frequencies for mid-range policy tenures (around 0.4 to 0.8), which could indicate a transitional period where fewer claims occur.



Insights from Age of Car Distribution by Claim Occurrence

From the provided histograms of car age distributions separated by claim occurrence (is\_claim = 0 and is\_claim = 1), we can derive several insights:

Newer Cars Have Higher Frequencies Without Claims:

For is\_claim = 0 (left plot), the distribution shows a significant peak at the beginning of the car age range (around 0.0). This indicates that newer cars have a higher frequency of not filing claims. The frequency decreases rapidly as the car age increases.

Claims Occur More Frequently with Slightly Older Cars:

For is\_claim = 1 (right plot), the distribution also shows a high frequency for newer cars (around 0.0), but the decline in frequency is less steep compared to is\_claim = 0. This suggests that slightly older cars (up to 0.1 normalized years) have a relatively higher frequency of filing claims.

Higher Frequency of Very New Cars with Claims:

Both plots exhibit a high frequency for very new cars (close to 0.0), indicating that very new cars have a significant proportion of both claim and non-claim instances. However, the non-claim frequency is much higher.

Long Tail in Claim Distribution:

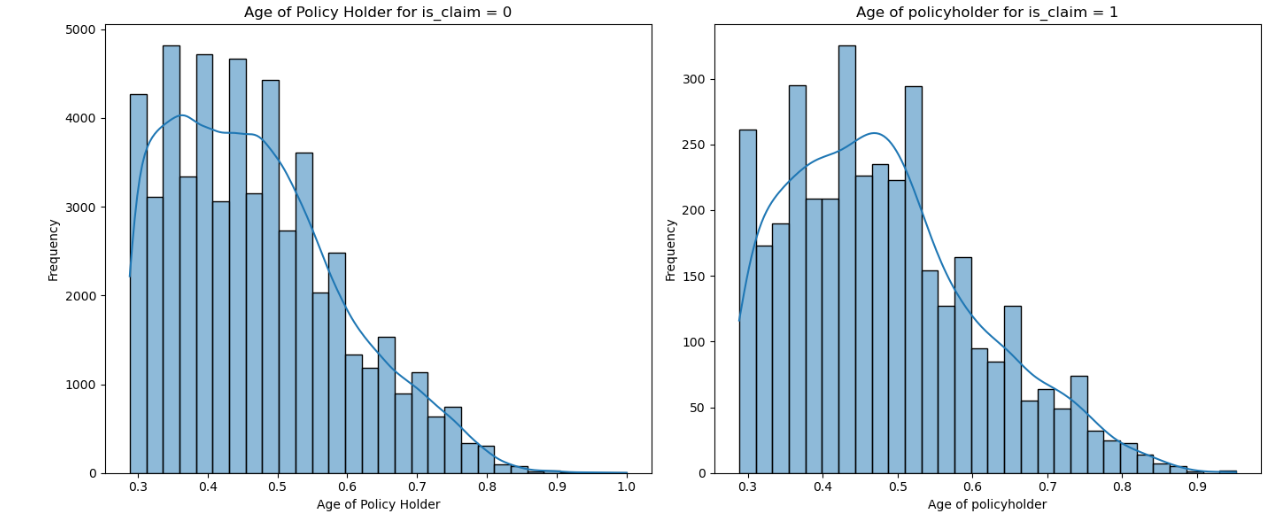
The distribution for is\_claim = 1 shows a longer tail, indicating that claims are spread out more evenly across a wider range of car ages, whereas the non-claim distribution is more concentrated towards newer cars.

Multiple Peaks in Claim Distribution:

The is\_claim = 1 plot shows multiple peaks, with secondary peaks around 0.05, 0.1, and 0.15. This could suggest specific age intervals where cars are more prone to filing claims.

Potential Correlation:

The comparison of both plots suggests that newer cars have a higher likelihood of not filing claims, but there is a significant proportion of claims in slightly older cars. This could imply that car age is a relevant factor in predicting claim occurrence.



Observations:

Non-claim Policyholders (Left Plot):

The distribution is right-skewed.

Most policyholders who have not filed a claim fall within the 0.3 to 0.6 range of normalized age.

The frequency decreases significantly after 0.6.

Claim Policyholders (Right Plot):

The distribution is also right-skewed but has fewer overall counts compared to non-claim policyholders.

The majority of policyholders who filed a claim are within the 0.3 to 0.5 range of normalized age.

There is a noticeable drop in frequency after 0.6.

Insights:

Age of Policyholder:

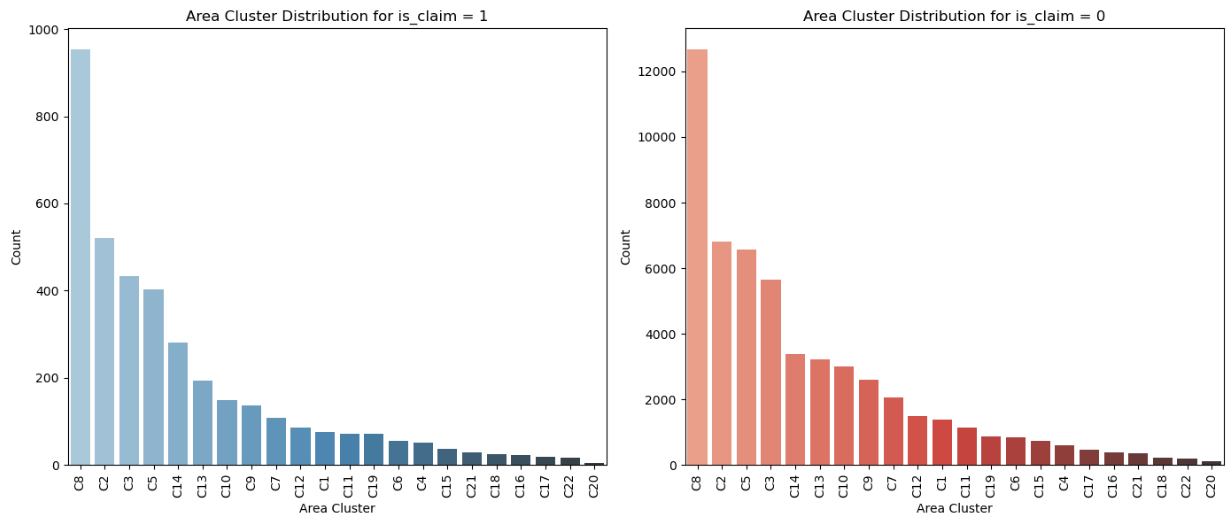
Younger policyholders (normalized age 0.3 to 0.5) are more common among both claim and non-claim groups.

There is a slightly higher proportion of policyholders in the younger age brackets among those who have filed a claim compared to those who haven't.

Potential Risk Factor:

Younger policyholders might have a higher risk of filing a claim, although further analysis is needed to confirm this trend.

Heatmap drawwwwwwwwwww



Visually analyze