Insights:

- There are almost equal data for both male and female, same goes for all four regions
- In the dataset, 80% are non-smokers and 20% are smokers
- 43% data is of patients with severity level 0 followed by level 1
- Average & median hospitalization charges of male are slightly greater than female
- Average & median hospitalization charges of smokers are greater than non-smokers
- Average & median hospitalization charges of southeast region are higher than other regions
- Average & median hospitalization charges of severity level 2,3,4 is greater than other severity levels
- Hospitalization charges of females with severity level 2,3,4 is similar
- Hospitalization charges of smokers are greater than non-smokers across both genders, all regions and all severity levels
- There are very less outliers present in age and viral load. And in hospitalization charges, even though there are few outliers but we will keep those as there can be high hospitalization charges in some cases

- **Statistical Test Results:**
    1. Mean hospitalization charges for smokers are greater than non-smokers. Proved by Right tailed T-test.
    2. Mean viral load for both female and male is equal. Proved by two tailed T-test.
    3. Proportion of smoking is same across different regions. Proved by Chi-square test.
    4. Viral load of women with severity level 0,1,2 is similar. Proved by one-way anova test.
    5. Hospitalization charges of male and female are same. Proved by Kruskal-Wallis test.

Recommendations:

- As we can see that hospitalization of people who do smoking is greater than those who don't. So apollo hospitals can plan for the resources according to the diseases that are caused by smoking

- Viral load for both male and female is equal means amount of virus is equal in both infected males and females, so in case of any infectious decease, apollo can plan such that the resources for both male and female are equal

- Proportion of smoking is same across regions. And, we already know that hospitalization of smokers is greater than non-smokers so apollo can equally divide resources in all 4 regions (northeast, southeast, southwest and northwest) of Delhi

- Viral load of women with severity level 0,1,2 is similar. Viral load is independent of severity level in females so in case of infectious decease, this can be used by apollo to utilise the resources effectively