This feature indicates if an individual is a smoker or not.

Unique Values – 0 (non-smokers), 1 (smokers)

**Mean**: 0.204783

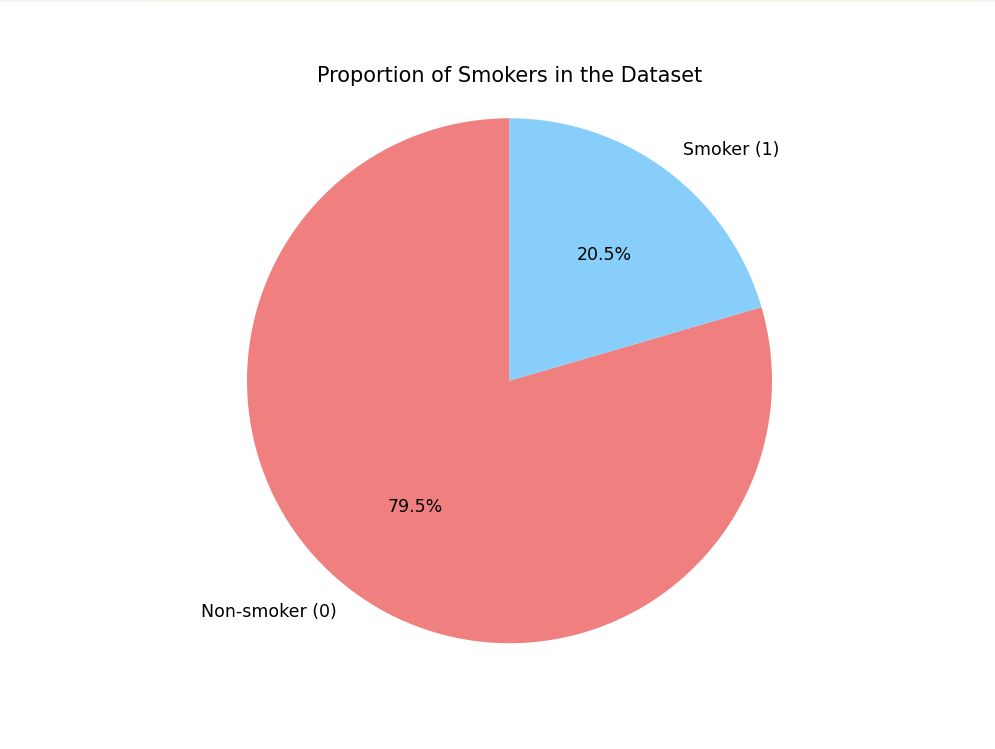
* + roughly 20.5% of the observations are smokers. Non-smokers form most of the population

1. **Median**: 0.0
   * more than half of the observations are non-smokers.
2. **Mode**: 0
   * non-smokers are more common on the dataset than smokers.
3. **Standard Deviation**: 0.403694

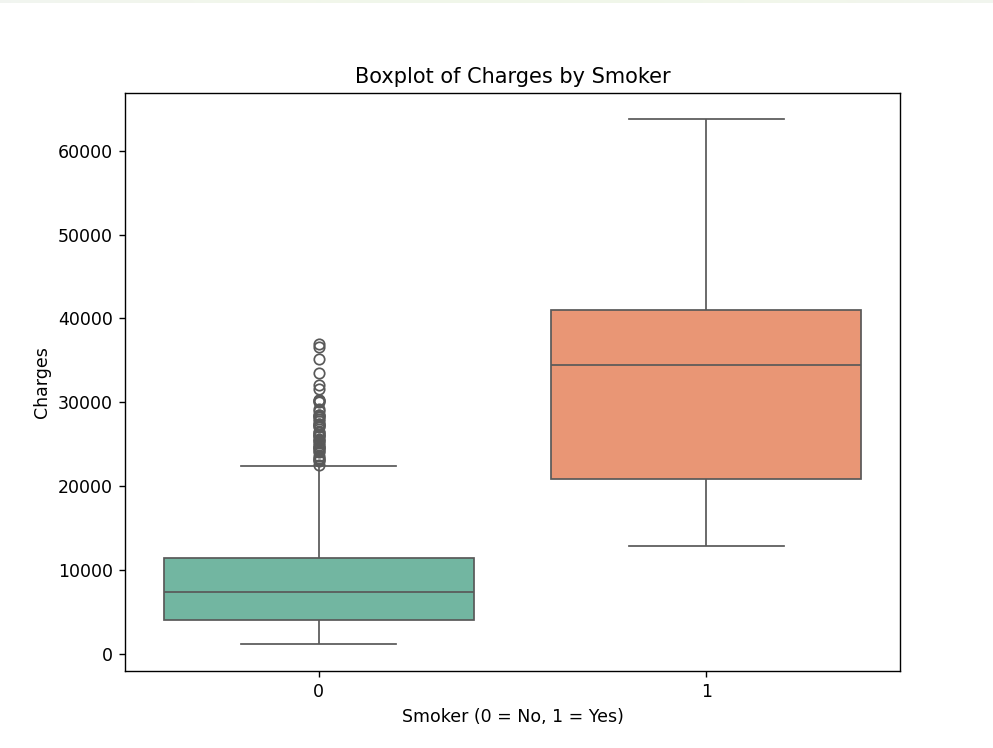
This shows some level of imbalance, with the majority being non-smokers.

1. **Variance**: 0.162969
   * shows that most values are close to the mean
   * indicates that most individuals are non-smokers.
2. **Percentiles (25th, 50th, 75th)**:
   * **25th Percentile**: 0.0 - At least 25% of the observations are non-smokers.
   * **50th Percentile (Median)**: 0.0 - At least half of the observations are non-smokers.
   * **75th Percentile**: 0.0 - 75% of the observations are non-smokers

Plotting



-The pie chart indicates a large majority of non-smokers in the dataset.



* Indicates that smokers have higher medical charges compared to non-smokers.
* Smokers have higher variability in their medical expenses which is apparent from its wide range of charges.
* This suggests a strong association between smoking and high medical charges.