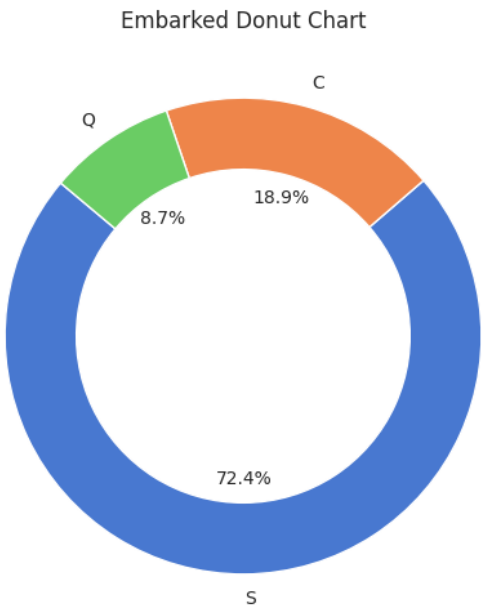
How to interpret Donut chart?



A. Interpretation of the Donut Chart Components:

- **The Outer Ring:** The entire ring represents the whole dataset or 100% of the observations for the "Embarked" variable.
- **Arcs (Slices):** The ring is divided into different colored arcs, each representing a category of the "Embarked" variable ('S', 'C', and 'Q').
- **Size of the Arcs:** The length (and area) of each arc is proportional to the proportion or percentage of observations belonging to that category relative to the whole dataset.
- **Labels:** Each arc is associated with a label indicating the category ('S', 'C', 'Q') and the corresponding percentage of the total within the arc itself.
    - 'S' (blue) occupies the largest arc and represents 72.4% of the embarked passengers.
    - 'C' (orange) occupies the second largest arc and represents 18.9% of the embarked passengers.
    - 'Q' (green) occupies the smallest arc and represents 8.7% of the embarked passengers.

- **The Hole in the Center:** The central empty circle is the defining characteristic of a donut chart. While it doesn't inherently represent data in this simple univariate case, it can be used to display additional information in more complex donut charts (e.g., a total number or another related metric). In this univariate context, it primarily serves an aesthetic purpose.

B. Interpreting the "Embarked" Distribution:

The donut chart visually shows the relative proportions of passengers embarking from each port, similar to a pie chart:

- **'S' (72.4%):** The vast majority of passengers (almost three-quarters) embarked from the port represented by 'S'.
- **'C' (18.9%):** A smaller but still significant proportion of passengers embarked from the port represented by 'C'.
- **'Q' (8.7%):** The smallest proportion of passengers embarked from the port represented by 'Q'.

Like a pie chart, the donut chart provides a quick visual understanding of the relative contribution of each category to the total.

The use cases for donut charts in univariate analysis are very similar to those of pie charts, with a few potential nuances:

- **Showing Parts of a Whole:** Like pie charts, donut charts effectively illustrate how a whole is divided into its constituent parts, emphasizing the proportion of each category relative to the total.
- **Simple Datasets with Few Categories:** Donut charts work best with a limited number of distinct categories (typically 2-5) for readability.
- **Highlighting Large Proportions:** They can quickly convey which category represents a significant portion of the data.
- **Aesthetic Preference:** Some people find donut charts visually more appealing or less visually overwhelming than pie charts due to the central hole.
- **Potential for Additional Information in the Center:** The central hole can be utilized in more complex scenarios (e.g., multivariate donut charts or dashboards) to display a summary statistic or another

relevant piece of information related to the overall dataset or the categories. However, this isn't a factor in a simple univariate donut chart.

Similar Limitations to Pie Charts:

Like pie charts, donut charts share some limitations compared to bar plots:

- **Difficulty in Comparing Sizes:** Accurately comparing the sizes of arcs (especially by area) can be challenging, especially for arcs that are close in size. It's generally easier to compare the lengths of bars.
- **Handling Many Categories:** Donut charts become cluttered and difficult to read with a large number of categories.
- **Difficulty in Judging Small Differences:** Small differences in proportions can be hard to discern visually.
- **Lack of Direct Count Representation:** The chart primarily focuses on proportions, and the exact counts are not as immediately apparent as in a bar plot.

In summary, donut charts are a visual variation of pie charts used to show the proportions of a categorical variable. They share similar strengths and weaknesses with pie charts and are best suited for simple datasets with a few categories where the focus is on the parts of a whole. While the central hole offers a potential space for additional information in more complex visualizations, for basic univariate analysis, the choice between a pie chart and a donut chart often comes down to aesthetic preference. Bar plots generally offer a more robust and easier-to-interpret way to visualize categorical data.