Statistical overview of the “Ethnicity” column from the dataset. Here’s what each value means:

1. **Count (2392)**: This is the number of entries in the dataset for ethnicity. There are 2392 records.
2. **Mean (0.877508361)**: The mean value of the Ethnicity variable is approximately 0.88. This indicates the average value across all records. In the context of ethnicity, this value suggests that the data are not uniformly distributed and are skewed towards lower values.
3. **Standard Deviation (1.028475776)**: This measures the amount of variation or dispersion in the Ethnicity values from the mean. A standard deviation of approximately 1.03 indicates that there is considerable variability in the Ethnicity values around the average. The fact that this value is close to 1 implies significant spread in the ethnicity data.
4. **Min (0)**: The minimum value of the Ethnicity variable is 0. This is the lowest value recorded, which likely represents one of the categories of ethnicity.
5. **25th Percentile (0)**: Also known as the first quartile. This means that 25% of the ethnicity values are 0 or less. Since the value is 0, this indicates that at least 25% of the data falls into the lowest category of ethnicity (0).
6. **50th Percentile (0)**: Also known as the median. This is the middle value when all entries are sorted. For the Ethnicity variable, a median of 0 means that 50% of the records are 0 or less, showing that the lower categories are more common.
7. **75th Percentile (2)**: Also known as the third quartile. This means that 75% of the ethnicity values are 2 or less. It indicates that a significant portion of the data falls into the categories up to 2, but 25% of the records have a value higher than 2.
8. **Max (3)**: The maximum value of the Ethnicity variable is 3. This is the highest value recorded, which represents one of the higher categories of ethnicity.

**Summary**

* **Range**: Ethnicity values range from 0 to 3.
* **Mean**: The average value of 0.88 indicates that the distribution leans towards lower values.
* **Spread**: The standard deviation of about 1.03 shows that there is a considerable spread in the ethnicity values.
* **Percentiles**:
  + 25% of the entries are 0 or lower.
  + 50% (the median) of the entries are 0 or lower, indicating that the lower categories are prevalent.
  + 75% of the entries are 2 or lower.

In summary, the Ethnicity data is categorized into several groups with values from 0 to 3. Most records fall into the lower categories, with a significant spread in the values. The distribution is skewed towards the lower end, with only a smaller proportion falling into higher categories.