Median Income results

The output you provided shows summary statistics for two groups based on the "removed" variable:

| removed | mean\_income | sd\_income | n | se\_income | lower\_ci | upper\_ci |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | 111075 | 47100 | 577 | 1961 | 107232 | 114919 |
| 1 | 94942 | 47331 | 60 | 6110 | 82966 | 106918 |

Here's a breakdown of what each column represents:

* **removed:** The group or category. In this case, it's either "0" (not removed) or "1" (removed).
* **mean\_income:** The average income for that group.
* **sd\_income:** The standard deviation of income for that group.
* **n:** The sample size (number of data points) in that group.
* **se\_income:** The standard error of the mean income for that group.
* **lower\_ci:** The lower bound of the 95% confidence interval for the mean income.
* **upper\_ci:** The upper bound of the 95% confidence interval for the mean income.

**Interpretation:**

* **Mean Income:** The "not removed" group has a higher mean income (111075) compared to the "removed" group (94942).
* **Standard Deviation:** The standard deviation is similar for both groups, indicating similar variability in income within each group.
* **Sample Size:** The "not removed" group has a much larger sample size (577) than the "removed" group (60).
* **Standard Error:** The standard error is smaller for the "not removed" group due to the larger sample size. This means the estimate of the mean income for the "not removed" group is more precise.
* **Confidence Intervals:** The confidence intervals overlap for the two groups. This suggests that the difference in mean income between the groups might not be statistically significant at a 95% confidence level.

**Reliability:**

The reliability of the data can be assessed based on:

* **Sample Size:** Larger sample sizes generally lead to more reliable estimates. The "not removed" group has a larger sample size, making its estimates more reliable.
* **Standard Error:** Smaller standard errors indicate more precise estimates. The "not removed" group has a smaller standard error, suggesting a more reliable estimate of the mean income.
* **Confidence Intervals:** If the confidence intervals for two groups overlap, it suggests that the difference in means might not be statistically significant.

**Overall:**

While the "not removed" group has a higher mean income, the overlapping confidence intervals indicate that the difference might not be statistically significant. The larger sample size and smaller standard error for the "not removed" group suggest that its mean income estimate is more reliable.

**Additional Considerations:**

* **Data Quality:** The reliability of the data also depends on its quality, such as the accuracy of income reporting and the representativeness of the sample.
* **External Factors:** Factors like economic conditions, policy changes, or other external influences can affect the income distribution and the reliability of the data.

By considering these factors, you can assess the overall reliability of the data and the conclusions drawn from the analysis.