Summary of Statistical Analysis: "Does Age Impact the Number of Hours Worked per Week?"

Statistical/Hypothetical Question:

The primary question addressed in this analysis was: **"Does age impact the number of hours worked per week?"** This question is crucial for understanding workforce dynamics and has implications for retirement planning and productivity management. It was hypothesized that as individuals age, their work hours might change due to various factors such as career advancement, family responsibilities.

Outcome of Exploratory Data Analysis (EDA):

The **Exploratory Data Analysis (EDA)** focused on five key variables: **age**, **hours worked per week**, **education**, **occupation**, and **income**.

1. **Age**: The age distribution was balanced, with most individuals in their 30s and 40s. This suggests a diverse representation across different life stages.
2. **Hours Worked per Week**: This variable was skewed, with many individuals working 40 hours per week, but also outliers working fewer or more hours, possibly due to part-time or overtime work.
3. **Education**: Most individuals had a high school diploma or some college education. The analysis indicated that education levels might influence work hours, with more educated individuals potentially working in more stable or professional environments.
4. **Occupation**: Occupation types varied widely, with technical and managerial roles generally requiring longer hours compared to manual labor or service jobs.
5. **Income**: The income variable was split into two categories: those earning over $50,000 and those earning less. The analysis suggested that higher income earners tend to work longer hours, possibly reflecting more senior roles or jobs that require extra time.

The EDA showed that **age** does have an impact on **hours worked per week**, but other factors like **education**, **occupation**, and **income** also play significant roles in shaping work hour patterns.

What was missed during analysis and Variables That Could Have Helped:

- **Work Class**: Could explain differences in work hours across different employment types.
- **Marital Status and Family Situation**: Could influence work schedules, especially for older individuals with dependents.
- **Geography**: Urban areas may have different work patterns compared to rural areas due to job availability and commuting.

Assumptions and Incorrect Approaches:

The analysis assumed a **linear relationship** between age and hours worked, which may not capture the full complexity. The impact of age on work hours might differ at various life stages. Additionally, the assumption that **income** directly influences work hours without considering occupation type or industry sector may be overly simplistic.

Challenges Faced:

Handling **outliers** was a challenge, particularly in the **hours worked** variable. Extreme values may represent valid cases, such as part-time or overtime workers, but their impact on the model needed careful consideration. Furthermore, the **linear regression** model used might not fully capture the complexities of how age affects work hours.

Conclusion:

In conclusion, the analysis showed that **age** does have a statistically significant relationship with **hours worked per week**, though the effect was modest. The **R-squared** value suggested that age alone does not explain much of the variance in work hours. To improve the model, incorporating non-linear regression and addressing outliers would enhance the accuracy of the results.