**Executive summery**

The Panel Study of Income Dynamics (PSID) is a nationally representative, longitudinal study of families in the USA that began in 1968 - about 50 years ago. It is the longest-running household panel survey in the world, and has been the model for similar studies launched in a number of other countries

**A dataframe containing :**

intnum - 1968 interview number

persnum - person number

age - age of individual

educatn - highest grade completed

earnings - total labor income

hours - annual work hours

kids - live births to this individual

married - last known marital status (married, never married, windowed, divorced, separated, NA/DF, no histories)

**Actions performed to analyse the data set.**

Identify the field with data value. Seqnum, Intnum and persnum are identification numbers that is not needed in analyzing.

Visualize the dataset. Since small number of variable exists, box plot is constructed for each variable.

Make assumptions and Remove outliers

Identify the relationship between variables.

Build a hypothesis.

Analyze the hypothesis.