* **Data Set Information**

Extraction was done by Barry Becker from the 1994 Census database. A set of reasonably clean records was extracted using the following conditions: ((AAGE>16) && (AGI>100) && (AFNLWGT>1)&& (HRSWK>0))   
Prediction task is to determine whether a person makes over 50K a year.

* **Dependent Variable**

The dependent variable is **“incomelevel”**, representing the level of income. A value of “<=**50K” indicates “less than or equal to USD 50000”** and “**>50K” indicates “greater than USD 50000”.**

* **Independent Variable**

Below are the independent variables (features or predictors) from the Census Data.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable Name** | **Description** | **Type** | **Possible Values** |
| Age | Age of the individual | Continuous | Numeric |
| Workclass | Class of Work | Categorical | Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked |
| fnlwgt | Final Weight Determined by Census Org | Continuous | Numeric |
| Education | Education of the individual | Ordered Factor | Bachelors, Some-college, 11th, HS-grad, Prof-school, Assoc-acdm, Assoc-voc, 9th, 7th-8th, 12th, Masters, 1st-4th, 10th, Doctorate, 5th-6th, Preschool |
| Education-num | Number of years of education | Continuous | Numeric |
| Marital-status | Marital status of the individual | Categorical | Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse |
| Occupation | Occupation of the individual | Categorical | Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces. |
| Relationship | Present relationship | Categorical | Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried |
| Race | Race of the individual | Categorical | White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black |
| Sex | Sex of the individual | Categorical | Female, Male |
| Capital-gain | Capital gain made by the individual | Continuous | Numeric |
| Capital-loss | Capital loss made by the individual | Continuous | Numeric |
| Hours-per-week | Average number of hours spent by the individual on work | Continuous | Numeric |
| Native-country | Native country of orgin | Categorical | United-States, Cambodia, England, Puerto-Rico, Canada, Germany, Outlying-US(Guam-USVI-etc), India, Japan, Greece, South, China, Cuba, Iran, Honduras, Philippines, Italy, Poland, Jamaica, Vietnam, Mexico, Portugal, Ireland, France, Dominican-Republic, Laos, Ecuador, Taiwan, Haiti, Columbia, Hungary, Guatemala, Nicaragua, Scotland, Thailand, Yugoslavia, El-Salvador, Trinadad&Tobago, Peru, Hong, Holand-Netherlands. |