The Adult dataset is from the Census Bureau. Predict whether income exceeds $50K/yr based on census data. Also known as "Census Income" dataset.

Model Goal: Classification task is to determine whether a person makes over 50K | a year.

Prepare the metadata and the data quality report for the dataset and list the variables selected for the analysis

|  |  |
| --- | --- |
| Variable name | Desc |
| age | age |
| workclass | Type of job |
| fnlwgt | Column created a sampling weight to select records from original data. |
| education | Education |
| education-num | Number of years of education |
| marital-status | Marital status |
| occupation | Job |
| relationship | Relationship status |
| race | Race |
| sex | Gender |
| capital-gain | Capital gain from investments |
| capital-loss | Losses from investments |
| hours-per-week | Number of hours worked for a week |
| native-country | Native country |

Output variable (desired target):

Income - income exceeds $50K/yr

Data Source: Adapted from <https://archive.ics.uci.edu/ml/datasets/Adult>