# Machine Learning Case Study - Binary Classification

You are a Data Scientist working in a Public Policy team. Your team needs you to come up with a prediction model to know if a person, based on his/her demographic data will earn $50,000 or more. This prediction will help the team in making policy decisions for providing financial assistance for the low-income group. You are given a sample data of the population along with their annual income. You can use that data to train your machine learning model. You can build your model in your own hardware / pc / laptop and just upload the prediction as shown in the below format.

Instructions for the case study are provided below –

* Build a ML Model, which is capable of predicting if an individual's income is greater than 50k or not.
* Use 'TrainData' file provided below for building the model.
* Use 'TestData' file provided below for testing your predictions.
* The prediction must be done based on various data attributes provided below.

## Data Attributes description

• age: continuous.

• workclass: Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.

• fnlwgt: continuous.

• education: 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: continuous.

• marital-status: Married-civ-spouse, Divorced, Never- married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.

• occupation: 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: Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.

• race: White, Asian-Pac-Islander, Amer-lndian-Eskimo, Other, Black.

• sex: Female, Male.

• capital-gain: continuous.

• capital-loss: continuous.

• hours-per-week: continuous.

• native-country: United-States, Cambodia, England,

Puerto-Rico, Canada, Germany, Outlying-US(Guam USVl-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.

• income >50K: binary (Target that needs to be predicted)