**Problem Statement:**

Using the Census Income dataset, classify whether an individual makes more than $50,000 annually.

**Dataset**:

The Census Income dataset has 48,842 entries. Each entry contains the following information about an individual:

* **age**: the age of an individual
  + Integer greater than 0
* **workclass**: a general term to represent the employment status of an individual.
  + Private, Self­emp­not­inc, Self­emp­inc, Federal­gov, Local­gov, Without­pay, Never worked.
* **fnlwgt**: final weight, the number of people the census believes the entry represents.
  + Integer greater than 0
* **education**: the highest level of education achieved by an individual.
* **education­num**: the highest level of education achieved in numerical form.
  + Integer greater than 0
* **marital­status**: marital status of an individual.
  + Married­civ­spouse corresponds to a civilian spouse while Married­AF­spouse is a spouse in the Armed Forces.
  + Married­civ­spouse, Divorced, Never­married, Separated, Widowed, Married­spouse­absent, Married­AF­spouse.
* **occupation**: the general type of occupation of an individual
  + Tech­support, Craft­repair, Other­service, Sales, Exec­managerial, Prof­specialty, Handlers­cleaners, Machine­op­inspct, Adm­clerical, Farming­fishing, Transport­moving, Priv­house­serv, Protective­serv, Armed­Forces.
* **relationship**: represents what this individual is relative to others.
  + For example, an individual could be a Husband. Each entry only has one relationship attribute and is somewhat redundant with marital status. We might not make use of this attribute at all
  + Wife, Own­child, Husband, Not­in­family, Other­relative, Unmarried.
* **gender**: the biological gender of the individual
  + Male, Female
* **capital­gain**: capital gains for an individual
  + Integer greater than or equal to 0
* **capital­loss**: capital loss for an individual
  + Integer greater than or equal to 0
* **hours­per­week**: the hours an individual has reported to work per week
  + continuous.
* **native­country**: country of origin for an individual
* **class**: whether an individual makes more than $50,000 annually.
  + ○ <=50k, >50k

**Steps to solve:**

1. Use the CREATE\_DATASET.sql and create your dataset.
2. Perform Exploratory Data Analysis and understand the data.
3. Divide the data 70%(training) 30%(test).
4. Create and train your model on the training dataset you have just now created. (CREATE\_MODEL\_AND\_TRAIN - Guide.sql). Explain the model and its metrics.
5. Apply the trained model on the test data set and evaluate the performance of the model. (Precision, Recall, and F1 score) (APPLY\_MODEL - Guide.sql)
6. Summarize all the information and present the same.