BÁO CÁO BÀI ASSIGNMENT 2

1. Import dữ liệu:  
   Tạo collection **usAdultIncome** để import dữ liệu từ csv:   
   A computer screen shot of a black screen

   Description automatically generated
2. ERD

A black background with white text

Description automatically generated

* Thiết kế các collection:
  + Finance sẽ lưu thông tin về tài chính như income\_bracket
  + Occupation sẽ lưu thông tin về công việc như: occupation, wokclass, hours\_per\_week
  + Relationship sẽ lưu thông tin về mối quan hệ như: relationship và martial\_status
  + Personal\_info sẽ lưu thông tin các nhân như: race, native\_country, age, gender
  + Eduation sẽ lưu thông tin học vấn như education, education\_num
  + Collection user sẽ lấy khóa ngoài là id từ các bảng trên và lưu thông tin về capital\_loss, gain, total
* Quan hệ giữa user và các collection khác là quan hệ 1 nhiều ví dụ là nhiều user cùng trường thông tin education và education\_num thì sẽ có cùng education\_id

1. Business query
   1. Có bao nhiêu người là nữ lam việc hơn 50 tiếng một tuần  
      Câu truy vấn:   
      A screen shot of a computer program

      Description automatically generated  
      Kết quả:  
      A screen shot of a computer code

      Description automatically generated
   2. Có bao nhiêu người ở Mỹ có mức thu nhập trên 50k  
      A screen shot of a computer program

      Description automatically generated  
      kết quả:  
      A computer screen shot of a code

      Description automatically generated
   3. Tìm số người cử nhân ở Mỹ:  
      A screen shot of a computer program

      Description automatically generated  
      Kết quả:  
      A computer screen with text and numbers

      Description automatically generated
   4. Liệt kê các loại tình trạng hôn nhân:   
      A screen shot of a computer program

      Description automatically generated  
      Kết quả  
      A screen shot of a computer program

      Description automatically generated
   5. Liệt kê các loại quan hệ:  
      A screen shot of a computer code

      Description automatically generated  
      Kết quả:  
      A screen shot of a computer program

      Description automatically generated
   6. Tìm người thạc sỹ chưa có gia đinh:  
      A screen shot of a computer program

      Description automatically generated  
      Kết quả:  
      A screen shot of a computer

      Description automatically generated
2. Index  
   A computer screen shot of text

   Description automatically generated
   1. Tạo index trong collection eduction:  
      Chỉ tạo index cho trường education để dễ truy vấn  
      Không tạo index cho trường education\_num do trường thông tin cập nhật thường xuyên ( hàng năm)
   2. Tạo index cho collection personlInfo:  
      Tạo index cho trwongf native\_country  
      Không tạo index cho trường gender vì chỉ có 2 giá trị nên việc tìm kiếm theo index sẽ không đạt hiệu quả  
      Không tạo index cho trường age vì cập nhật hàng năm
   3. Tạo index cho collection relationship  
      Tạo index cho trường relationship và martial\_status: tạo index có relationship trước martial\_status vì relationship ít giá trị hơn -> hiệu quả cao hơn chi chọn martial\_status trước
   4. Tạo index cho collection occupation:  
      tạo index cho cả 3 trường occupation, workclass, hours\_per\_week