

## Caractéristiques du jeu de données ACSIncome

Le jeu de données est décrit dans l'article Ding, Frances, et al. "Retiring adult: New datasets for fair machine learning." *Advances in neural information processing systems* 34 (2021): 6478-6490. Cet article est accessible avec le lien <https://arxiv.org/pdf/2108.04884.pdf>.

Le jeu de données ACSIncome est détaillé dans l'Annexe B1 extraite ci-dessous. Les valeurs des features et du labels sont fournies dans deux fichiers séparés.

Predict whether US working adults' yearly income is above \$50,000.

**Target:** PINCP (Total person's income): an individual's label is 1 if PINCP > 50000, otherwise 0. Note that with our software package, this chosen income threshold can be toggled easily to label the ACS PUMS data differently, and construct a new prediction task.

### Features:

- AGE (Age): Range of values:
  - 0 - 99 (integers)
  - 0 indicates less than 1 year old.
- COW (Class of worker): Range of values:
  - N/A (not in universe)
  - 1: Employee of a private for-profit company or business, or of an individual, for wages, salary, or commissions
  - 2: Employee of a private not-for-profit, tax-exempt, or charitable organization
  - 3: Local government employee (city, county, etc.)
  - 4: State government employee
  - 5: Federal government employee
  - 6: Self-employed in own not incorporated business, professional practice, or farm
  - 7: Self-employed in own incorporated business, professional practice or farm
  - 8: Working without pay in family business or farm
  - 9: Unemployed and last worked 5 years ago or earlier or never worked

- SCHL (Educational attainment): Range of values:
  - N/A (less than 3 years old)
  - 1: No schooling completed
  - 2: Nursery school/preschool
  - 3: Kindergarten
  - 4: Grade 1
  - 5: Grade 2
  - 6: Grade 3
  - 7: Grade 4
  - 8: Grade 5
  - 9: Grade 6
  - 10: Grade 7
  - 11: Grade 8
  - 12: Grade 9
  - 13: Grade 10
  - 14: Grade 11
  - 15: 12th Grade - no diploma
  - 16: Regular high school diploma
  - 17: GED or alternative credential
  - 18: Some college but less than 1 year
  - 19: 1 or more years of college credit but no degree
  - 20: Associate's degree
  - 21: Bachelor's degree
  - 22: Master's degree
  - 23: Professional degree beyond a bachelor's degree
  - 24: Doctorate degree
- MAR (Marital status): Range of values:
  - 1: Married
  - 2: Widowed
  - 3: Divorced
  - 4: Separated
  - 5: Never married or under 15 years old
- OCCP (Occupation): Please see ACS PUMS documentation for the full list of occupation codes
- POBP (Place of birth): Range of values includes most countries and individual US states; please see ACS PUMS documentation for the full list.

- RELP (Relationship): Range of values:
  - 0: Reference person
  - 1: Husband/wife
  - 2: Biological son or daughter
  - 3: Adopted son or daughter
  - 4: Stepson or stepdaughter
  - 5: Brother or sister
  - 6: Father or mother
  - 7: Grandchild
  - 8: Parent-in-law
  - 9: Son-in-law or daughter-in-law
  - 10: Other relative
  - 11: Roomer or boarder
  - 12: Housemate or roommate
  - 13: Unmarried partner
  - 14: Foster child
  - 15: Other nonrelative
  - 16: Institutionalized group quarters population
  - 17: Noninstitutionalized group quarters population
- SEX (Sex): Range of values:
  - 1: Male
  - 2: Female
- RAC1P (Recoded detailed race code): Range of values:
  - 1: White alone
  - 2: Black or African American alone
  - 3: American Indian alone
  - 4: Alaska Native alone
  - 5: American Indian and Alaska Native tribes specified, or American Indian or Alaska Native, not specified and no other races
  - 6: Asian alone
  - 7: Native Hawaiian and Other Pacific Islander alone
  - 8: Some Other Race alone
  - 9: Two or More Races