



INCOME CLASSIFICATION

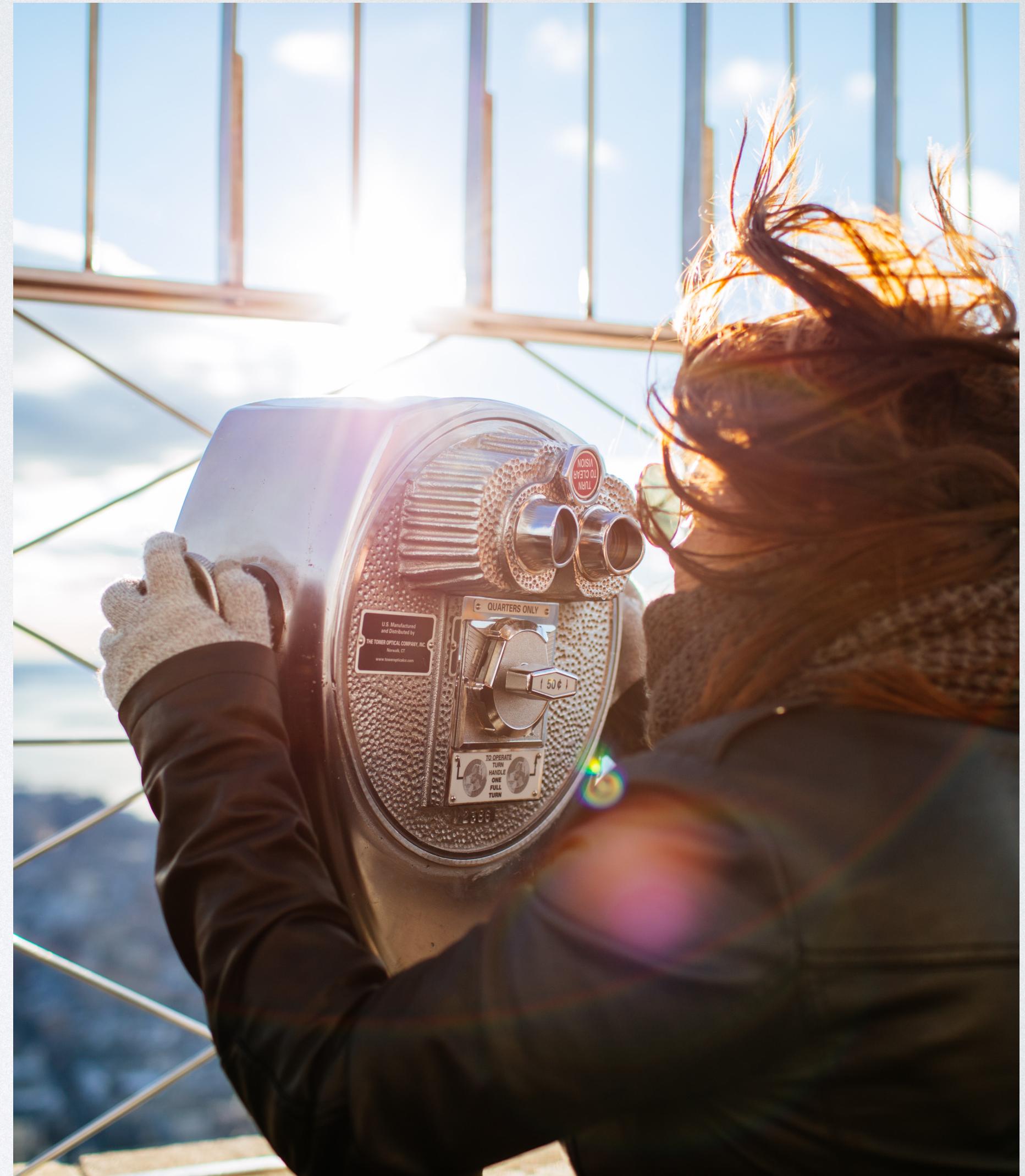
Dream Team Analytics:
Bruno Santos, Jeremy Opacich,
David Trichter, Eric Stokes

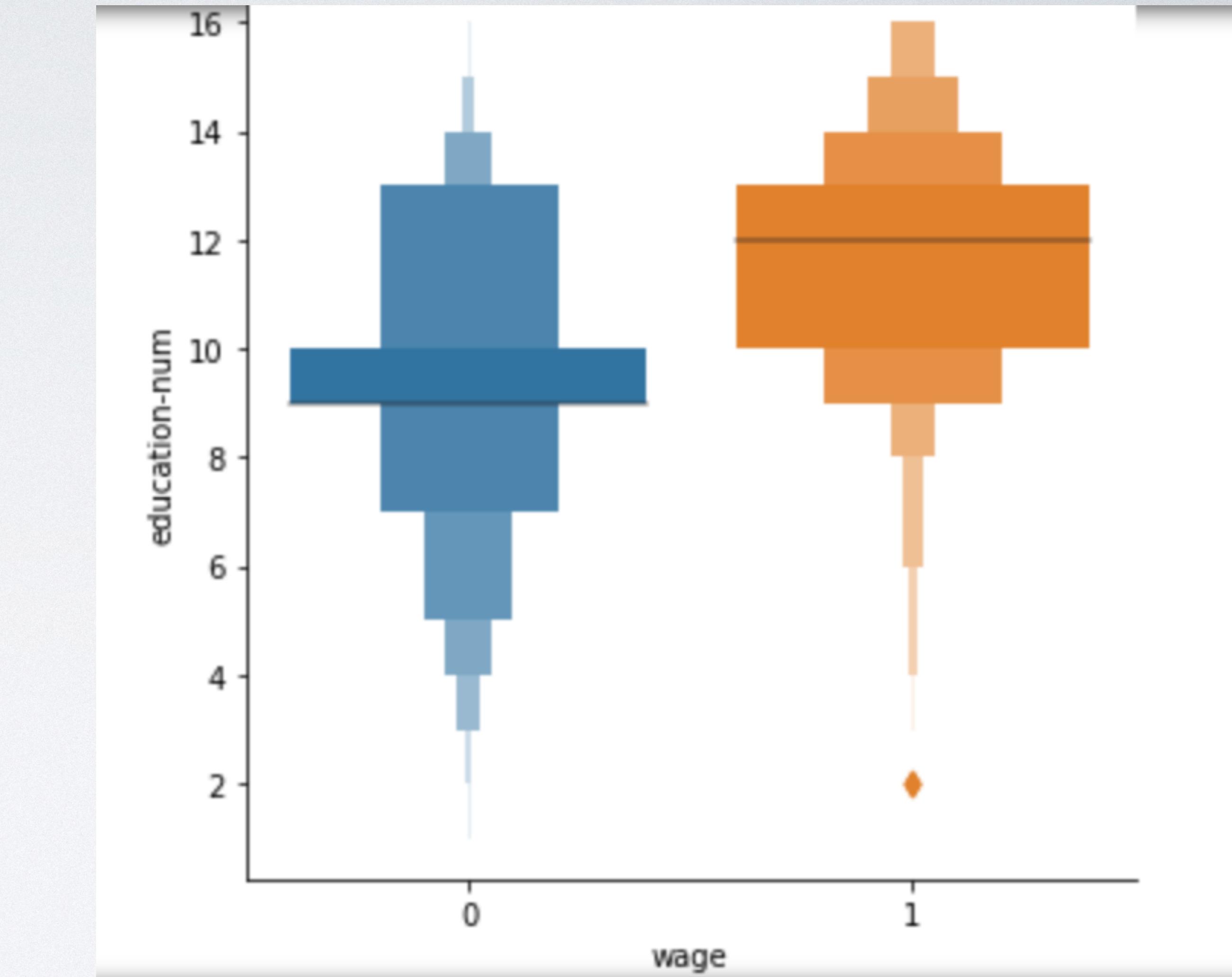
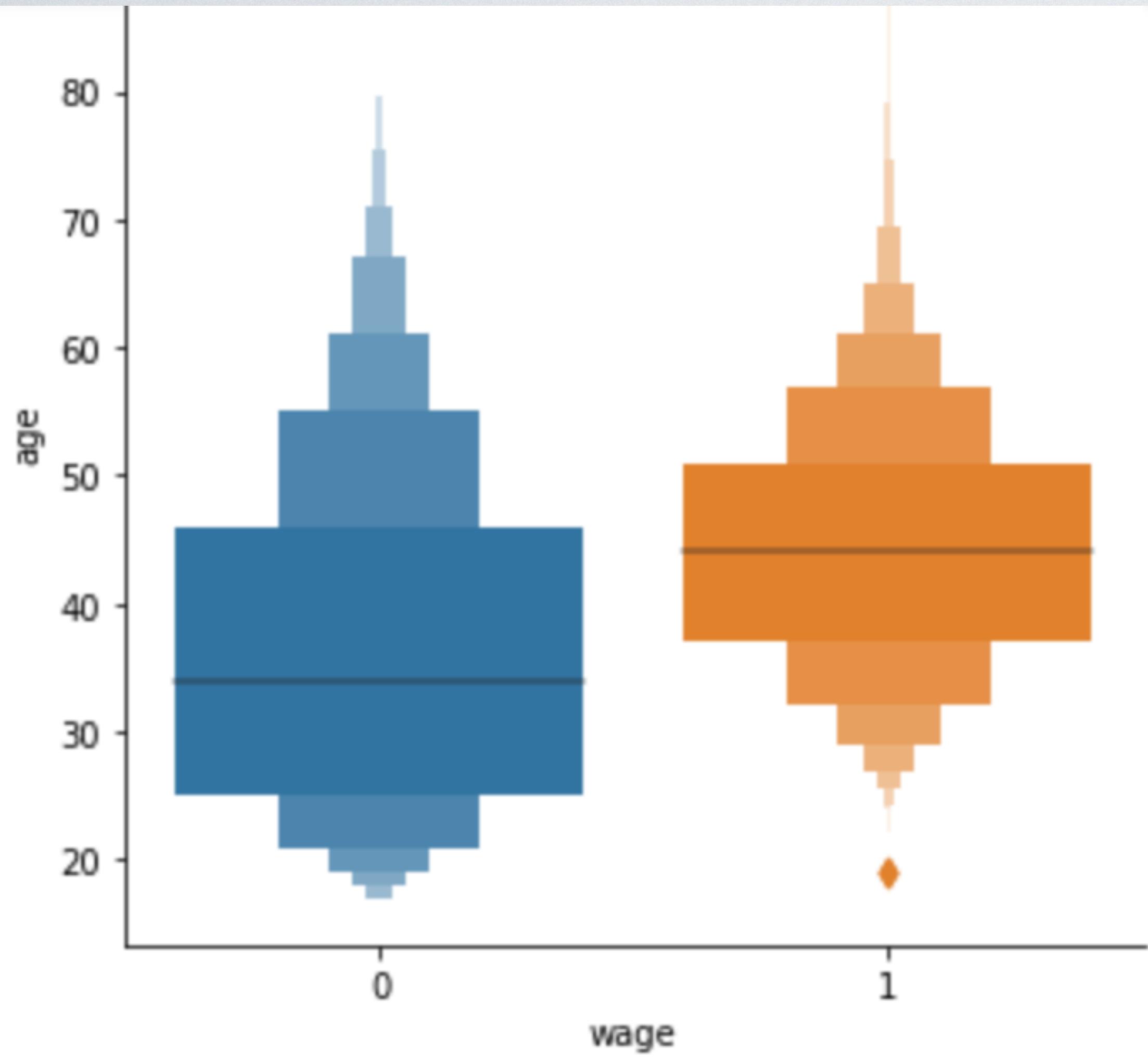
1994 Census

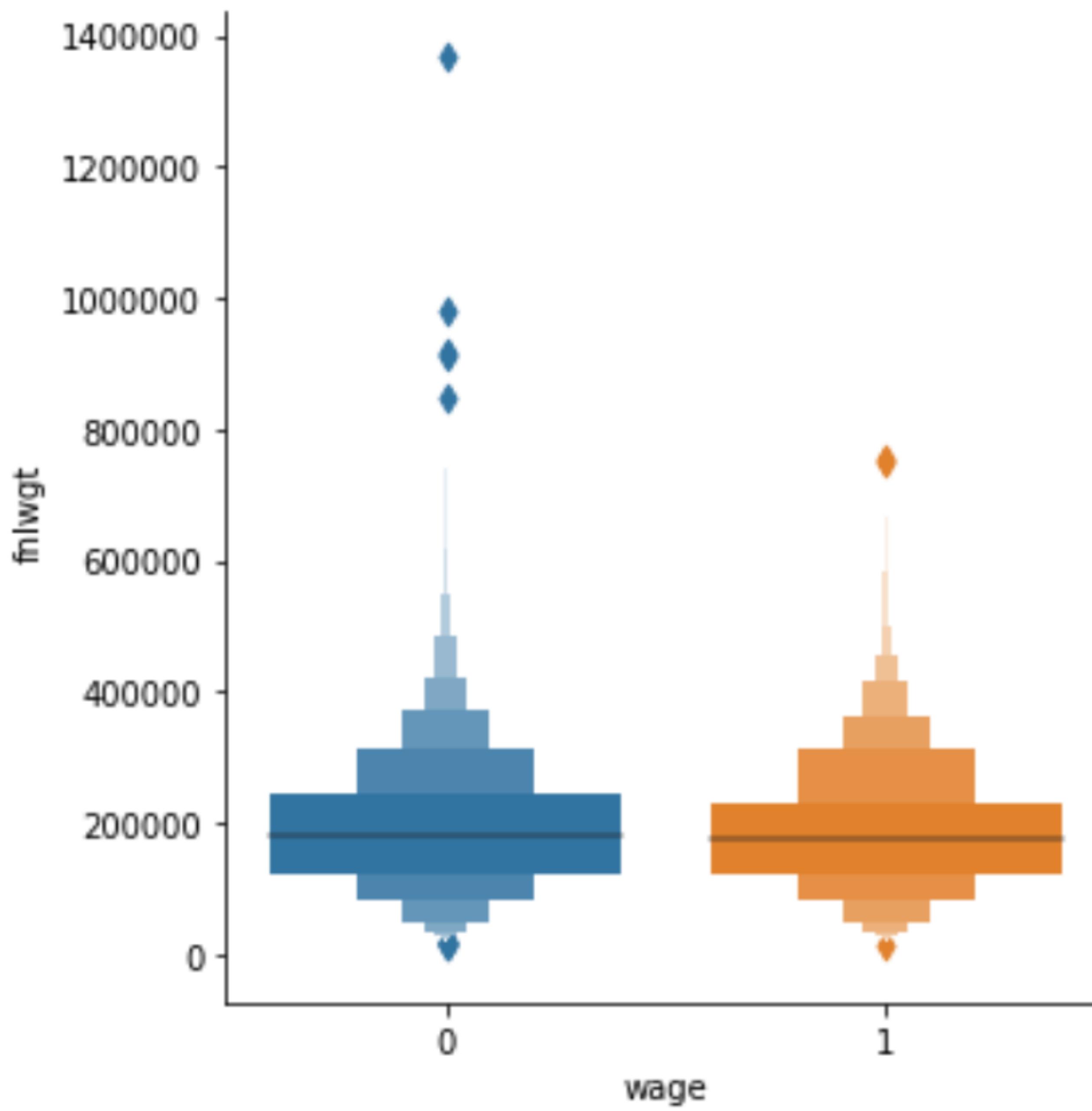
CAN CERTAIN PERSONAL FEATURES
ENABLE US TO PREDICT INCOME
STATUS?

INITIAL OBSERVATIONS

- Removed Financial Weight over 800k
- Majority of population were United States Citizens







FEATURE ENGINEERING

- Created interaction items:
 - Education * Being Married
 - Age * Education
 - Education * Not Married
- Created dummy columns for all Categorical data
- Made ordinal values for Education
- Bootstrap our data

MODEL EVALUATION

Baseline: 74.2%

- Models attempted:
 - AdaBoost
 - Logistic Regression
 - SVC
 - K-Nearest Neighbors
 - GradientBoost
 - DecisionTree

Best Performing, AdaBoost: 85.8%

PRODUCTION MODEL

- Random Forest:
 - Training score: 99.7%
 - Testing score: 92.5%
- Best Features:
 - Married/ Financial Weight
 - Married
 - Age
 - Financial Weight
 - Education

NEXT STEPS

- Continue to collect data to train from
- Address occupation and workplace values
- Build additional interaction items
- Hypertune our parameters

