Tax Data Prediction

Christopher Hyek

7/21/2019

1994 Census Data Set
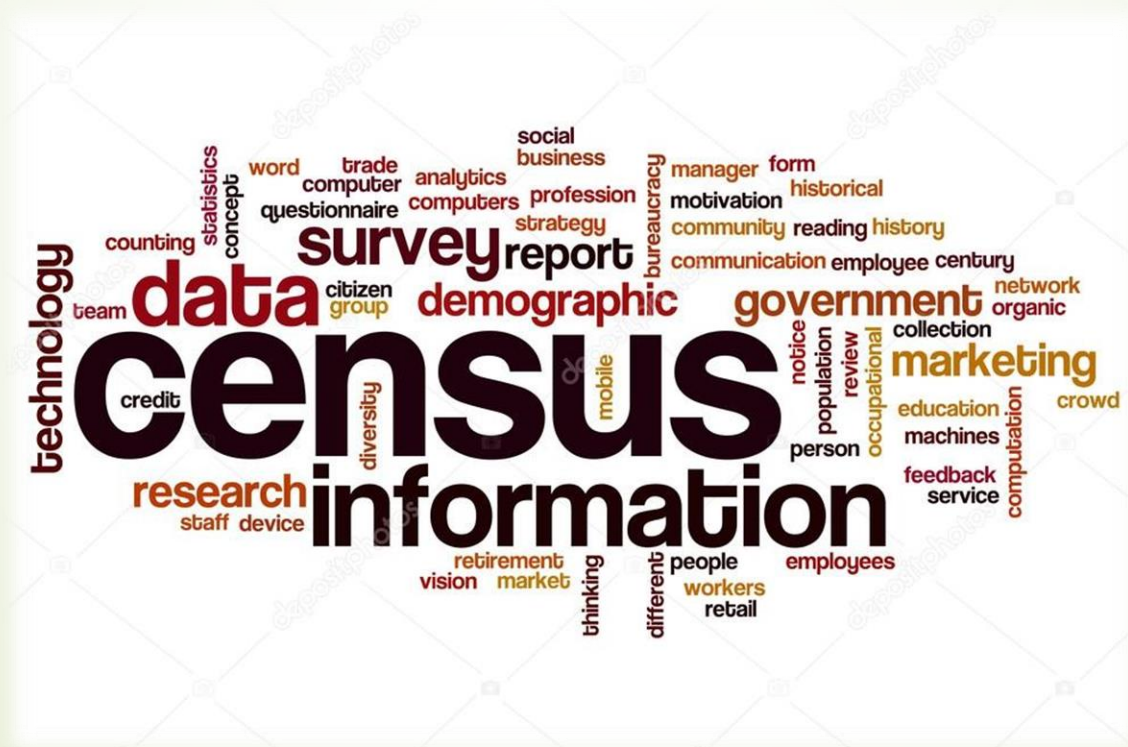
Goal

Predict annual salary based on data

Data Acquisition

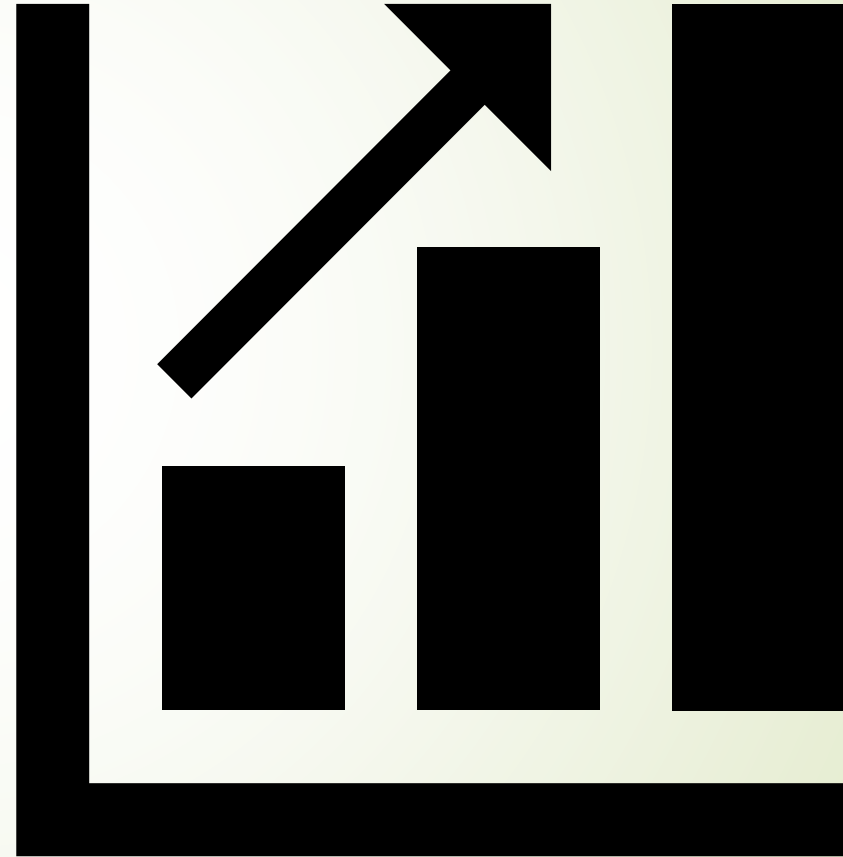To predict tax data we obtain

census bureau information

Scrubbing

Removing serous outliers

Using proper labels

Removing typos

Exploring

Little data is numerical

Most is categorical

Focus on how to use categorical

Modeling

Logistic Regression

Focus on precision

Validate Data

Interpreting

85% correct

Okay precision

Okay accuracy

Recommendations

We should:

    Attempt a new model to verify

    Try different constraints

    Refine the survey data and try again

Thank you