

Do workplace smoking bans effect to reduce smoking?

Description

SmokeBans is a cross-sectional data set with observations on 10,000 workers. The data set contains information on whether individuals were, or were not, subject to a workplace smoking ban.

Format

A data frame containing 10,000 observations on 7 variables.

Variables:

smoker: Factor. Is the individual a current smoker?

Ban: Factor. Is there a work area smoking ban?

Age: Age in years.

Education: Factor indicating highest education level attained: high school (hs) drop out, high school graduate, some college, college graduate, master's degree (or higher).

Afam: Factor. Is the individual African, American?

Hispanic: Factor. Is the individual Hispanic?

Gender: Factor : gender male/female.

Our target is smoker, because the smoke bans one of the effects

Reference

Reference of Dataset: <https://vincentarelbundock.github.io/Rdatasets/datasets.html>

Tools I use to analysis and clinging data and modelling

Data process: Pandas, Numpy

Modelling: ScikitLearn

Visualization: Matplotlib, Seaborn

For modelling

I try to build by ScikitLearn library.