Name: Mateusz Chetkowski ID: 33654709

C) The problem I decided to study in the income dataset was that of higher education leads to higher paying jobs. I plotted this bar graph using R Studio in order to show this

A graph of different colored bars

Description automatically generated

As you can see, in the <=50K there is a much higher count of people however another thing to notice is that in the >50k bracket there isn’t much of the lower education classes but a much higher distribution of the higher education individuals

D) A computer screen shot of a blue screen

Description automatically generated

E)

A screen shot of a computer

Description automatically generated

F)

* **age**: Age of the individual.
* **workclass**: The type of work the individual is engaged in, such as Private, Self-emp-not-inc, Self-emp-inc, Federal-gov, Local-gov, State-gov, Without-pay, Never-worked.
* **fnlwgt**: Final weight. This is the number of units in the target population that the responding unit represents.
* **education**: The highest level of education completed by the individual, ranging from preschool to a doctorate.
* **education.num**: Numeric representation of education, usually corresponding to the years of education.
* **marital.status**: Marital status of the individual, such as Married-civ-spouse, Divorced, Never-married, Separated, Widowed, Married-spouse-absent, Married-AF-spouse.
* **occupation**: The type of occupation of the individual, such as Tech-support, Craft-repair, Other-service, Sales, Exec-managerial, Prof-specialty, Handlers-cleaners, Machine-op-inspct, Adm-clerical, Farming-fishing, Transport-moving, Priv-house-serv, Protective-serv, Armed-Forces.
* **relationship**: Relationship status of the individual, such as Wife, Own-child, Husband, Not-in-family, Other-relative, Unmarried.
* **race**: Race of the individual, typically categorized as White, Asian-Pac-Islander, Amer-Indian-Eskimo, Other, Black.
* **sex**: Gender of the individual, Male or Female.
* **capital.gain**: Capital gains earned by the individual.
* **capital.loss**: Capital losses incurred by the individual.
* **hours.per.week**: Number of hours worked per week.
* **native.country**: Country of origin of the individual, usually represented by the country name.
* **income**: The target variable indicating whether the individual's income exceeds $50,000 per year. It's typically categorized as <=50K or >50K

G)

A graph of a graph showing the amount of hours

Description automatically generated

This plot presents the density distribution of hours worked per week among individuals with different income levels. It helps visualize whether there's a difference in work hours between different income groups. Adjusting the alpha value controls the transparency of the density curves.

A graph of income by work class

Description automatically generated

This plot illustrates the distribution of income levels across different work classes.

A graph of a graph

Description automatically generated with medium confidence

This plot shows the distribution of ages among individuals with different income levels.