

Exploratory Data Analysis and Visualization - MSc AIDA UoM

Professor George Evaggelidis

Academic Year: 2021-2022

The purpose of this assignment is to apply data visualization and analysis techniques on the Birth_weight dataset using R. Study the structure of the Birth_weight dataset

```
head(Birth_weight) # Display the first few entries of the Birth_weight dataset
str(Birth_weight) # Display the structure of the Birth_weight dataset
```

Provide a diagram for each of the following:

1. Study the distribution of the 'income' variable values
2. Study the distribution of the 'baby_wt' (baby weight) variable values
3. Study the distribution of the 'mother_age' variable values
4. Study the relationship between the variables 'mother_age' and 'income'
5. Study the relationship between the variables 'baby_wt' and 'smoke'
6. Study the relationship between the variables 'mother_age' and 'mother_wt' (mother weight)
7. Study the relationship between the variables 'mother_wt', 'baby_wt' and 'smoke'
8. Study the distribution (boxplot) of the 'baby_wt' variable in relation to the 'smoke' variable
9. Study the distribution (boxplot) of the 'baby_wt' variable in relation to the 'smoke' and 'income' variables