

Application exercise 7.1: Linear regression

Team name: _____

Lab section: 8:30 10:05 11:45 1:25 3:05

Write your responses in the spaces provided below. WRITE LEGIBLY and SHOW ALL WORK! Only one submission per team is required. One team will be randomly selected and their responses will be discussed and graded. Concise and coherent are best!

Cigarettes and CO

The Federal Trade Commission annually rates varieties of domestic cigarettes according to their tar, nicotine, and carbon monoxide content. The United States Surgeon General considers each of these substances hazardous to a smoker's health. Past studies have shown that increases in the tar and nicotine content of a cigarette are accompanied by an increase in the carbon monoxide emitted from the cigarette smoke.

In this exercise we will work with data from 2007 on cigarettes sold in the US. Each row in the dataset represents a cigarette. There are 11 variables in the dataset:

- BRAND_NAME
- TYPE: Type of cigarette, REGULAR or MENTHOL
- NIC: Nicotine content, in mg
- TAR: Tar content, in mg
- CO: Carbon monoxide, in mg
- LEN: Length of cigarette, in mm
- FLTR: Filter, F or NF
- PACK: Pack type, HARD or SOFT
- STRENGTH: Strength of cigarette, ULTRA LIGHT, LIGHT, MEDIUM, REGULAR FULL, or FLAVOR
- STYLE: Some information of style of cigarette (not available for all cigarettes, and not used in this analysis)
- OTHER: Other relevant information (not available for all cigarettes, and not used in this analysis)

To load the data use the following:

```
download("http://stat.duke.edu/~mc301/data/cig07.RData", destfile = "cig07.RData")  
load("cig07.RData")
```

1. First we try a full model with all explanatory variables in the model. The output of this model is shown below. Interpret the slopes of nicotine and tar variables. Are these results surprising? Why, or why not? Make sure to use appropriate terminology in your answer, and use summary statistics and/or visualizations of the data to support your answer.