Department of Statistics

COURSE STATS 330

Assignment 1, 2003

Instructions: Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 7th August.

The Federal Trade Commission in the US 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.

The data overleaf give data on 25 brands of cigarettes. The variables are

Brand name
Tar content (mg)
Nicotine content (mg)
Weight (g)
Carbon monoxide content (mg)

The data (in the form of a text file) are available on the course web page under the title cigs.txt.

1. Read the data into R and make a data frame **cigs.df** . Inspect the data for any unusual values and if you find any, make a new data frame with the outliers deleted. Delete no more than 2 brands. Print out the new data frame in R.

2. Suppose we want to construct a linear regression model to explain the carbon monoxide content in terms of the other variables. Using the new data, draw some plots to explore the suitability of the model. Do the plots suggest that fitting a linear regression model might be appropriate? Give reasons.

3. Fit a regression model to the data, using all the variables. Comment on the goodness of fit. Comment on the significance of the variables. Do you think the variable **weight** is required in the model? i.e. does it help explain the amount of carbon monoxide emitted?

4. Fit a regression using carbon monoxide as the response and nicotine as the only explanatory variable. Contrast the regression coefficient of nicotine and its significance in this model with that of nicotine in the model fitted in Q3. Can you explain the difference?

Data are listed overleaf.

Brand	tar	nicotine	weight	CO
Alpine	14.1	0.86	0.9853	13.6
Benson&Hedges	16.0	1.06	1.0938	16.6
BullDurham	29.8	2.03	1.1650	23.5
CamelLights	8.0	0.67	0.9280	10.2
Carlton	4.1	0.40	0.9462	5.4
Chesterfield	15.0	1.04	0.8885	15.0
GoldenLights	8.8	0.76	1.0267	9.0
Kent	12.4	0.95	0.9225	12.3
Kool	16.6	1.12	0.9372	16.3
L&M	14.9	1.02	0.8858	15.4
LarkLights	13.7	1.01	0.9643	13.0
Marlboro	15.1	0.90	0.9316	14.4
Merit	7.8	0.57	0.9705	10.0
MultiFilter	11.4	0.78	1.1240	10.2
NewportLights	9.0	0.74	0.8517	9.5
Now	1.0	0.13	0.7851	1.5
OldGold	17.0	1.26	0.9186	18.5
PallMallLight	12.8	1.08	1.0395	12.6
Raleigh	15.8	0.96	0.9573	17.5
SalemUltra	4.5	0.42	0.9106	4.9
Tareyton	14.5	1.01	1.0070	15.9
True	7.3	0.61	0.9806	8.5
ViceroyRichLight	8.6	0.69	0.9693	10.6
VirginiaSlims	15.2	1.02	0.9496	13.9
WinstonLights	12.0	0.82	1.1184	14.9