Solutions Quiz Given on last day of class

The dataset presented here contains measurements of weight (g), tar(mg), nicotine content (mg), carbon monoxide content (mg), brand, and flavor for 25 cigarettes.

Write out the true model and the estimated model for each of the following.

Model 1

```
##
## Call:
## lm(formula = weight ~ tar + nicotine + COcontent, data = cig.dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.10234 -0.05625 -0.00326  0.04390  0.16439
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.864079   0.061532   14.04  3.8e-12
## tar          0.001161   0.018006    0.06     0.95
## nicotine     0.110888   0.220102    0.50     0.62
## COcontent   -0.000412   0.012255   -0.03     0.97
##
## Residual standard error: 0.0812 on 21 degrees of freedom
## Multiple R-squared:  0.25,Adjusted R-squared:  0.143
## F-statistic: 2.34 on 3 and 21 DF,  p-value: 0.103
```

TRUE model:

$\mu\{wt|tar, nic, CO\} = \beta_0 + \beta_1 tar + \beta_2 nic + \beta_3 CO$

Estimated model:

$\widehat{\mu\{wt|tar, nic, CO\}} = 0.864 + 0.001 tar + 0.111 nic - 0.0004 CO$

What hypotheses are being tested in the NICOTINE line of the R output above?

$H_0: \beta_2 = 0$

$H_A: \beta_2 \neq 0$

Provide the conclusion, in context, of the test of the hypotheses in the previous question.

There is no evidence that the mean weight depends on nicotine after accounting for tar and CO content (p-value = 0.62 from t-stat = 0.5 on 21 df).

Model 3

```
##
## Call:
## lm(formula = weight ~ nicotine * Flavor, data = cig.dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.1173 -0.0358 -0.0127  0.0231  0.1554
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                  0.80892    0.06881   11.76  5.7e-09
## nicotine                     0.14714    0.07822    1.88     0.08
## Flavorchocolate              0.20173    0.18370    1.10     0.29
## Flavorclove                  0.16767    0.19428    0.86     0.40
## Flavormenthol                0.06397    0.09712    0.66     0.52
## Flavorregular                0.44373    0.27800    1.60     0.13
## nicotine:Flavorchocolate    -0.15398    0.21593   -0.71     0.49
## nicotine:Flavorclove        -0.15735    0.23658   -0.67     0.52
## nicotine:Flavormenthol       0.00303    0.09839    0.03     0.98
## nicotine:Flavorregular      -0.47486    0.28428   -1.67     0.12
##
## Residual standard error: 0.0744 on 15 degrees of freedom
## Multiple R-squared:  0.551,Adjusted R-squared:  0.281
## F-statistic: 2.04 on 9 and 15 DF,  p-value: 0.107
## Analysis of Variance Table
##
## Response: weight
##                  Df Sum Sq Mean Sq F value Pr(>F)
## nicotine          1 0.0462  0.0462    8.35  0.011
## Flavor            4 0.0345  0.0086    1.56  0.236
## nicotine:Flavor   4 0.0210  0.0052    0.95  0.463
## Residuals        15 0.0830  0.0055
```

TRUE model:

$$\mu\{wt|nicotine, flavor\} = \beta_0 + \beta_1 nic + \beta_2 I_{choc} + \beta_3 I_{clove} + \beta_4 I_{menthol} + \beta_5 I_{reg} + \beta_6 nic\, I_{choc} + \beta_7 nic\, I_{clove} + \beta_8 nic\, I_{menthol} + \beta_9 nic\, I_{reg}$$

$$I_{choc} = \begin{cases} 1 & \text{if chocolate} \\ 0 & \text{if else} \end{cases} \qquad I_{clove} = \begin{cases} 1 & \text{if clove flava} \\ 0 & \text{if else} \end{cases}$$

$$I_{menthol} = \begin{cases} 1 & \text{if menthol} \\ 0 & \text{if else} \end{cases} \qquad I_{reg} = \begin{cases} 1 & \text{if reg} \\ 0 & \text{if else} \end{cases}$$

ESTIMATED model:

$$\mu\{wt|nicotine, flavor\} = 0.809 + 0.147 nic + 0.202 I_{choc} + 0.168 I_{clove} + 0.064 I_{menthol} + 0.444 I_{reg}^2 - 0.154 nic\, I_{choc} - 0.157 nic\, I_{clove} + 0.003 nic\, I_{menthol} - 0.0475 nic\, I_{reg}$$

What is the estimated regression line for menthol flavored cigarettes?

$\hat{M}\{wt|nicotine, flavor=menthol\} = \beta_0 + \beta_1 nic + \beta_4 + \beta_8 nic$

$\hat{M}\{wt|nicotine, flavor=menthol\} = (0.809 + 0.064) + (0.147 + 0.003) nic$

What hypotheses are being tested in the nicotine:Flavor row above?

H_0: $\beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$

H_A: Not all β_6, β_7, β_8, and β_9 are 0

Write a conclusion for these hypotheses.

There is no evidence that the linear relationship between nicotine and weight changes across flavor (p-value = 0.463 from F-stat = 0.95 on 4 and 15 df).

The response is the length of odontoblasts (teeth) in each of 10 guinea pigs at each of three dose levels of Vitamin C (0.5, 1, 2, or 2.5 mg) with each of two delivery methods (orange juice or pill). A total of 60 guinea pigs were studied.

Here is a partial ANOVA table for this example. Fill in the blanks.

Analysis of Variance Table

Response: len

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
supp	1	205.35	205.35	15.572	0.0002312 ***
dose	2	2426.43	1213.22	92.4	< 2.2e-16 ***
supp:dose	2	108.32	54.16	4.11	0.0218603 *
Residuals	54	712.11	13.19		

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Write out the hypotheses being tested in the supp:dose row above.

H_0: $w_{OJ0.5} = w_{OJ1} = w_{OJ2.5} = w_{P0.5} = w_{P1} = w_{P2.5} = 0$

H_A: At least one $w_{jk} \neq 0$

Write out a conclusion for these hypotheses.

There is moderate evidence difference in mean tooth length between delivery methods changes across levels of dose (p-value = 0.022 from F-stat = 4.11 on 2 and 54 df).