A. Mail-Order Sales

(a) x_1 = number of purchases from a home gift catalog within 24 months
x_2 = proportion of single people in the zip code area
x_3 = number of credit cards

$$x_4 = \begin{cases} 1 & \text{apartment dweller} \\ 0 & \text{single-family home} \end{cases}$$

$$x_5 = \begin{cases} 1 & \text{has purchased from a similar catalog} \\ 0 & \text{has not purchased from a similar catalog} \end{cases}$$

$$y = \begin{cases} 1 & \text{customer makes purchase} \\ 0 & \text{customer does not make purchase} \end{cases}$$

(b) x_1, x_2, x_3 are numerical; x_4, x_5, y are binary.

B. College Student Drinking

(a) 0.70

(b) The odds that a college woman drinks frequently are 0.70 times the odds that a college man drinks frequently.

(c) The two odds ratios are reciprocals of each other:

$$(1.44) \times (0.70) = 1.008 \approx 1.0$$

These are just two alternative ways of saying the same thing.

C. Successful Franchises

- 17.1
  The proportion of exclusive-territory firms which are successful is

$$\widehat{p} = \frac{\#\text{ in exclusive territory group which are successful}}{\text{total number in exclusive-territory group}} = \frac{108}{142} = \boxed{0.7606}$$

  The odds are

$$\text{Odds} = \frac{\widehat{p}}{1-\widehat{p}} = \frac{0.7606}{0.2394} = \boxed{3.1771}$$

  For firms without exclusive territories,

$$\widehat{p} = \frac{15}{28} = \boxed{0.5357}$$

$$\text{Odds} = \frac{0.5357}{0.4643} = \boxed{1.1538}$$

- 17.3
  For exclusive territories, $\log(\text{Odds}) = \log(3.1771) = \boxed{1.1560}$
  For nonexclusive territories, $\log(\text{Odds}) = \log(1.1538) = \boxed{0.1431}$

Followup Exercise:

(a) • Response:

$$y = \begin{cases} 1 & \text{if successful} \\ 0 & \text{if not successful} \end{cases}$$

• Predictor:

$$x = \text{ Territory } = \begin{cases} 1 & \text{if exclusive} \\ 0 & \text{if nonexclusive} \end{cases}$$

(b)

$$\text{log odds(success)} = \beta_0 + \beta_1 x$$

(c)

$$\begin{aligned} \text{log odds(success)} &= \widehat{\beta}_0 + \widehat{\beta}_1 x \\ &= 0.1431 + 1.0127\, x \end{aligned}$$

(d) Nonexclusive $\Longrightarrow x = 0.$

$$\begin{aligned} \text{log odds} &= \widehat{\beta}_0 + \widehat{\beta}_1 x \\ &= \widehat{\beta}_0 + \widehat{\beta}_1(0) = \widehat{\beta}_0 = \boxed{0.1431} \end{aligned}$$

$$\Longrightarrow \text{ odds of success} = e^{0.1431} = \boxed{1.1538}$$

$$\Longrightarrow \text{ probability of success} = \frac{\text{odds}}{\text{odds} + 1} = \frac{1.1538}{2.1538} = \boxed{0.5357}$$

(e) Exclusive $\Longrightarrow x = 1.$

$$\begin{aligned} \text{log odds} &= \widehat{\beta}_0 + \widehat{\beta}_1 x \\ &= \widehat{\beta}_0 + \widehat{\beta}_1(1) = \widehat{\beta}_0 + \widehat{\beta}_1 = 0.1431 + 1.0127 = \boxed{1.1558} \end{aligned}$$

$$\Longrightarrow \text{ odds of success} = e^{1.1558} = \boxed{3.1766}$$

$$\Longrightarrow \text{ probability of success} = \frac{\text{odds}}{\text{odds} + 1} = \frac{3.1766}{4.1766} = \boxed{0.7606}$$

(f) Yes, except for round-off error.

D. Stock Options

- 17.19

  (a) 0.8022,   4.0556

  (b) 0.6881,   2.2059

  (c) 1.8385

  The odds of offering stock options for high-tech companies are 1.8385 times the odds for non-high-tech companies.

- 17.20

  (a) * High-tech firms: log odds = log(4.056) = $\boxed{1.4002}$

  * Non-high-tech firms: log odds = log(2.206) = $\boxed{0.7912}$

  (b)

  Logistic model: log odds (stock options) $= \beta_0 + \beta_1 x$

  * $x = 0 \implies$ Non-high tech firms.

$$\begin{aligned} \text{log odds} &= \beta_0 + \beta_1(0) \\ 0.7912 &= \beta_0 \\ \implies b_0 = \widehat{\beta}_0 &= \boxed{0.7912} \end{aligned}$$

  * $x = 1 \implies$ High-tech firms.

$$\begin{aligned} \text{log odds} &= \beta_0 + \beta_1(1) \\ 1.4002 &= \beta_0 + \beta_1 \\ 1.4002 &= 0.7912 + \beta_1 \\ \beta_1 &= 1.4002 - 0.7912 = 0.6090 \\ \implies b_1 = \widehat{\beta}_1 &= \boxed{0.6090} \end{aligned}$$

  (c)

$$e^{b_1} = e^{0.6090} = \boxed{1.8386}$$

  This is (almost) the same answer for the odds ratio that we calculated directly from the data in Exercise 17.19 part (c).

- **Followup Exercise:** Here are MINITAB's answers:

  (a) $b_0 = \widehat{\beta}_0 = \boxed{0.791128}$ $b_1 = \widehat{\beta}_1 = \boxed{0.608960}$

  (b) the odds ratio = $\boxed{1.84}$

  (c) A 95% CI for the odds ratio is $\boxed{(0.95,\ 3.54)}$

  (d) No, since 1 is contained in the 95% CI for the odds ratio.

  (e) No. If the odds are not significantly different then the probabilities aren't either (and vice versa.)

E. Blood Pressure and Cardiovascular Disease

- 17.25

  (a)

  - Response:

$$y = \begin{cases} 1 & \text{if died} \\ 0 & \text{if did not die} \end{cases}$$

  - Predictor:

$$x = \text{ Blood Pressure } = \begin{cases} 1 & \text{if high} \\ 0 & \text{if low} \end{cases}$$

  (b) 2.12

  (c)

  1.

$$H_0\colon\ \beta_1 = 0$$
$$H_A\colon\ \beta_1 \neq 0$$

  2. $Z = \boxed{2.91}$ P-value $= \boxed{0.004}$
  3. Reject H_0 since P-value $= 0.004 < .05 = \alpha$.
  4. We conclude that blood pressure (high or low) is related to death from heart disease.

  (d) From MINITAB, a 95% CI for the odds ratio is $\boxed{1.28 \text{ to } 3.51}$

  We are 95% confident that the odds that a man with high blood pressure dies from heart disease are between 1.28 and 3.51 times the odds that a man with low blood pressure dies from heart disease.

F. Healthy Versus Failed Companies

(a) 4.0572

(b) 59.5609

(c)

$$(4.0572)(14.68) = 55.5597 \quad \text{(This answer contains some roundoff error.)}$$

(d)

$$\text{Odds} = (59.5609)(14.68) = 874.3540$$

$$\text{Probability} = \frac{874.3540}{875.3540} = 0.9989 = \boxed{99.89\%}$$

(e)

$$\text{Log Odds} = -2.6293 + 2.6865x$$

$$p = 1/2 \implies \text{Odds} = \frac{1/2}{1/2} = 1:1 = 1 \implies \text{Log Odds} = 0$$

$$0 = -2.6293 + 2.6865x$$

$$(2.6865)x = 2.6293$$

$$x = \frac{2.6293}{2.6865} = \boxed{0.9787}$$

G. Pizza Hut

(a) Variables:

$$y = \text{Choice of pizza shop} = \begin{cases} 1 & \text{if Pizza Hut} \\ 0 & \text{if competitor} \end{cases}$$

$$x_1 = \text{Price of Pizza Hut pizza, in dollars} \quad \text{(quantitative)}$$

$$x_2 = \text{Gender} = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases}$$

(b)

1. log odds $= \beta_0 + \beta_2\, x_2$
2. log odds $= -1.5712 + 0.0623\,(\text{Gender})$
3. From MINITAB, the odds ratio is $\boxed{1.06}$

   The odds that a male student chooses Pizza Hut are approximately 1.06 times the odds that a female student chooses Pizza Hut.

(c)

1. log odds $= \beta_0 + \beta_1\, x_1$
2. log odds $= 1.2433 - 0.2503\,(\text{Price})$
3. From MINITAB, the odds ratio is $\boxed{0.78}$

   The odds of choosing Pizza Hut change (decrease) by a factor of 0.78 for every one-dollar increase in the price of Pizza Hut pizza.

(d)

1. log odds $= \beta_0 + \beta_1\, x_1 + \beta_2\, x_2$
2. log odds $= 1.2195 - 0.2502\,(\text{Price}) + 0.0377\,(\text{Gender})$
3.
   - From MINITAB, the odds ratio for Price is $\boxed{0.78}$

     The odds of choosing Pizza Hut change (decrease) by a factor of 0.78 for every one-dollar increase in the price of Pizza Hut pizza, if gender is held constant.
   - From MINITAB, the odds ratio for Gender is $\boxed{1.04}$

     The odds that a male student chooses Pizza Hut are approximately 1.04 times the odds that a female student chooses Pizza Hut, if the price of Pizza Hut pizza is held constant.

(e) The full model (with both predictors) should not be chosen since Gender is not significant in that model ($P = 0.918$).

The model which predicts from Gender alone should also not be chosen since Gender is not significant in that model ($P = 0.862$).

The model which predicts from Price alone is best conservative. Price is a significant predictor in that model ($P = 0.007$).

(continued)

(f) Our chosen model from (e) is

$$\text{log odds } = 1.2433 - 0.2503\,(\text{Price})$$

so we'll make interpretations and predictions from this model.

- Odds ratio for Price:

  The odds in favor of choosing Pizza Hut change (decrease) by a factor of 0.78 for every one-dollar increase in the price of Pizza Hut pizza.

- Odds ratio for Gender: Gender is not significant ($\beta_2 = 0$) so the odds ratio for Gender is

  $$e^{\beta_2} = e^0 = 1$$

  The odds for choosing Pizza Hut pizza are not related to gender, after accounting for price.

(g) Gender is not relevant. $x_1 = \$8.99$:

$$\begin{aligned}\text{log odds} &= 1.2433 - (0.2503)\,x_1 \\ &= 1.2433 - (0.2503)(8.99) \\ &= 1.2433 - 2.2502 = -1.0069 \\ \text{odds} &= e^{-1.0069} = .3653 \\ \widehat{p} &= \frac{\text{odds}}{1+\text{odds}} = \frac{.3653}{1.3653} = \boxed{.2676}\end{aligned}$$

(h) Gender is not relevant. $x_1 = \$11.49$:

$$\begin{aligned}\text{log odds} &= 1.2433 - (0.2503)\,x_1 \\ &= 1.2433 - (0.2503)(11.49) \\ &= 1.2433 - 2.8759 = -1.6326 \\ \text{odds} &= e^{-1.6326} = .1954 \\ \widehat{p} &= \frac{\text{odds}}{1+\text{odds}} = \frac{.1954}{1.1954} = \boxed{.1635}\end{aligned}$$

(i) \$13.75

H. Acceptable Cheese

(a) Variables:

$$y = \text{acceptability of cheese} = \begin{cases} 1 & \text{if acceptable (success)} \\ 0 & \text{if not acceptable (failure)} \end{cases}$$

$$x_1 = \text{percent of acetic acid in cheese mixture}$$

$$x_2 = \text{percent of hydrogen sulfide (H2S) in cheese mixture}$$

$$x_3 = \text{percent of lactic acid in cheese mixture}$$

All three predictor variables are quantitative.

Full model: $\log \text{odds} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

(b) Fitted final model:

$$\begin{aligned} \log \text{odds} &= \widehat{\beta}_0 + \widehat{\beta}_2 x_2 \\ &= -7.2789 + 0.9399\,(\text{H2S}) \end{aligned}$$

(c) Interpret odds ratios:

- Odds ratio for H2S = $\boxed{2.56}$:

  The odds in favor of making acceptable cheese change (increase) by a factor of 2.56 for every one percent increase in the quantity of hydrogen sulfide.

- Odds ratios for acetic and lactic acid: $\beta_1 = \beta_3 = 0 \Longrightarrow$ odds ratio $= e^0 = \boxed{1.0}$

  The odds in favor of making acceptable cheese do not depend on the percentages of acetic acid and lactic acid, after accounting for the percentage of hydrogen sulfide.

(d) 95% confidence intervals:

- H2S:

  We are 95% certain that the odds for acceptable cheese increase by a factor of between 1.30 and 5.03 for every one percent increase in the quantity of hydrogen sulfide.

- Acetic acid: Not applicable since this predictor is not used in the chosen model.
- Lactic acid: Not applicable since this predictor is not used in the chosen model.

(e) Estimate the probability that the cheese mixture is acceptable:

$$\begin{aligned} \log \text{odds} &= -7.2789 + (0.9399)\,(\text{H2S}) \\ &= -7.2789 + (0.9399)(7) \\ &= -7.2789 + 6.5793 = -0.6996 \end{aligned}$$

$$\text{odds} = e^{-0.6966} = .4968$$

$$\widehat{p} = \frac{\text{odds}}{1 + \text{odds}} = \frac{.4968}{1.4968} = \boxed{.3319} \quad \text{or} \quad \boxed{33.19\%}$$

I. Lawn Service

(a) response $y = \begin{cases} 1 & \text{if home uses lawn service} \\ 0 & \text{if home does not use lawn service} \end{cases}$

$$x_1 = \text{Attitude toward outdoor work} = \begin{cases} 1 & \text{if favorable} \\ 0 & \text{if unfavorable} \end{cases}$$

x_2 = Income (thousands of dollars)
x_3 = Lawn Size (thousand of square feet)
x_4 = Number of Teenagers
x_5 = Age of Head of Household

The variables y and x_1 are binary; the other variables x_2, x_3, x_4, x_5 are quantitative.

(b) Use the *full* model for this purpose. (See notes page 14.)

The odds of using lawn service change (decrease) by a factor of 0.82 for each additional teenager, when the other four variables are all held constant.

(c)

The odds of using lawn service change (increase) by a factor of 2.94 for each additional year of age, when the other four variables are all held constant.

(d)

$$\text{log odds} = -13.7385 + 0.1547\,(\text{Income}) - 2.0962\,(\text{Attitude})$$

(e)

- Estimated slope for Income is $\widehat{\beta}_2 = 0.1547 > 0$ so the odds of using a lawn service increase with greater income.
- Estimated slope for Attitude is $\widehat{\beta}_1 = -2.0962 < 0$ so the odds of lawn service decrease if attitude toward yard work is favorable rather than unfavorable.

(f) Odds ratios

- For Income: odds ratio = $\boxed{1.17}$

  The odds of using a lawn service increase by a factor of 1.17 for every additional $1000 in Income, when Attitude is held constant.

- For Attitude: odds ratio = $\boxed{0.12}$

  The odds of using a lawn service decrease by a factor of 0.12 if Attitude is favorable rather than unfavorable, when Income is held constant.

- All other variables: x_3 = lawn size, x_4 = number of teenagers, x_5 = age of head of household:

  The slopes $\beta_i = 0$ so the odds ratios are $e^{\beta_i} = e^0 = 1$: The odds for using a lawn service are not related to these other factors, after accounting for Income and Attitude.

(continued)

(g) Estimate the probability of using lawn service for this particular household:

$$\text{log odds} = -13.7385 + 0.1547\,(\text{Income}) - 2.0962\,(\text{Attitude})$$

$$\begin{aligned}\text{log odds} &= -13.7385 + (0.1547)(100) - (2.0962)(0)\\ &= -13.7385 + 15.47\\ &= 1.7315\end{aligned}$$

$$\text{odds} = e^{1.7315} = 5.6491$$

$$\widehat{p} = \frac{\text{odds}}{1+\text{odds}} = \frac{5.6491}{6.6491} = \boxed{.8496} \quad \text{or} \quad \boxed{84.96\%}$$