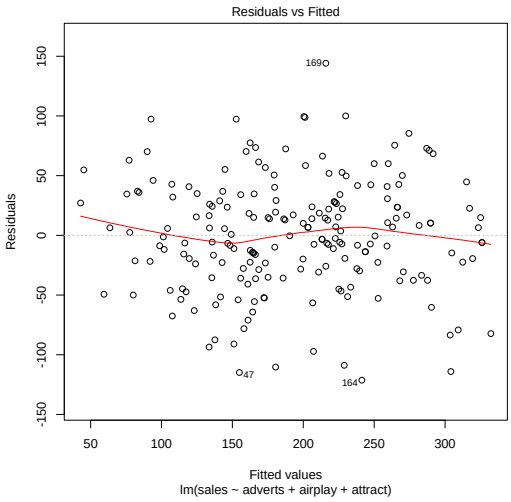
Chapter 7

1. R^2 is
    a. The percentage of variance in the predictor accounted for by the outcome variable.
    b. The proportion of variance in the outcome accounted for by the predictor variable or variables.*
    c. The proportion of variance in the predictor accounted for by the outcome variable.
    d. The percentage of variance in the outcome accounted for by the predictor variable or variables.
2. Which of the following statements about the *t*-statistic in regression is not true?
    a. The *t*-statistic tests whether the regression coefficient, *b*, is equal to 0.
    b. The *t*-statistic provides some idea of how well a predictor predicts the outcome variable.
    c. The *t*-statistic can be used to see whether a predictor variables makes a statistically significant contribution to the regression model.
    d. The *t*-statistic is equal to the regression coefficient divided by its standard deviation.*
3. Which of the following statements about outliers is not true?
    a. Outliers are values very different from the rest of the data.
    b. Outliers have an effect on the mean.
    c. Outliers have an effect on regression parameters.
    d. Outliers are influential cases.*
4. For which regression assumption does the Durbin–Watson statistic test?
    a. Linearity.
    b. Independence of errors.*
    c. Homoscedasticity.
    d. Multicollinearity.
5. Using the model in Chapter 5, how many records would be predicted to be sold if £29,000 was spent on advertising, the record was played 19 times on radio and the band were rated 7 on the attractiveness scale?
    a. 2,461,660 records.
    b. 2435 records.
    c. 2488 records.
    d. 2,435,050 records.*

6. The following graph shows:



a. Non-linearity.

b. Heteroscedasticity and non-linearity.

c. Regression assumptions that have been met.*

d. Heterscedasticity

Recent research has shown that lecturers are among the most stressed workers. A researcher wanted to know exactly what it was about being a lecturer that created this stress and subsequent burnout. She recruited 75 lecturers and administered several questionnaires that measured: **Burnout** (high score = burnt out), **Perceived Control** (high score = *low* perceived control), **Coping Ability** (high score = *low* ability to cope with stress), **Stress from Teaching** (high score = teaching creates a lot of stress for the person), **Stress from Research** (high score = research creates a lot of stress for the person), and **Stress from Providing Pastoral Care** (high score = providing pastoral care creates a lot of stress for the person). The outcome of interest was burnout, and Cooper's (1988) model of stress indicates that perceived control and coping style are important predictors of this variable. The remaining predictors were measured to see the unique contribution of different aspects of a lecturer's work to their burnout. The **R** output is below and the remaining questions relate to this output.

```
summary(burnoutModel.1)

Call:
glm(formula = burnout ~ loc + cope, family = binomial(), data = burnoutData)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9217  -0.5163  -0.3730   0.1273   2.0848

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept) -4.484493   0.379458 -11.818  < 2e-16 ***
loc          0.061080   0.010915   5.596 2.19e-08 ***
cope         0.082714   0.009369   8.829  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 530.11  on 466  degrees of freedom
Residual deviance: 364.18  on 464  degrees of freedom
AIC: 370.18

Number of Fisher Scoring iterations: 5
```

```
logisticPseudoR2s(burnoutModel.1)
```

```
Pseudo R^2 for logistic regression
Hosmer and Lemeshow R^2   0.313
Cox and Snell R^2         0.299
Nagelkerke R^2            0.441
```

```
exp(burnoutModel.1$coefficients)
```

```
(Intercept)         loc        cope
 0.01128261  1.06298389  1.08623164
```

```
exp(confint(burnoutModel.1))
```

```
                 2.5 %     97.5 %
(Intercept) 0.005160721 0.02292526
loc         1.041229885 1.08691181
cope        1.067210914 1.10722003
```

```
summary(burnoutModel.2)
```

```
Call:
glm(formula = burnout ~ loc + cope + teaching + research + pastoral,
    family = binomial(), data = burnoutData)

Deviance Residuals:
     Min        1Q    Median        3Q       Max
-2.41592  -0.48290  -0.28690   0.02966   2.63636

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.43993    1.08565  -4.090 4.32e-05 ***
loc          0.11079    0.01494   7.414 1.23e-13 ***
cope         0.14234    0.01639   8.684  < 2e-16 ***
teaching    -0.11216    0.01977  -5.673 1.40e-08 ***
research     0.01931    0.01036   1.863 0.062421 .
pastoral     0.04517    0.01310   3.449 0.000563 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 530.11  on 466  degrees of freedom
Residual deviance: 321.20  on 461  degrees of freedom
AIC: 333.2

Number of Fisher Scoring iterations: 6
```

```
modelChi; chidf; chisq.prob
```

```
[1] 208.9086
[1] 5
[1] 0
```

```
logisticPseudoR2s(burnoutModel.2)
```

```
Pseudo R^2 for logistic regression
Hosmer and Lemeshow R^2   0.394
Cox and Snell R^2         0.361
```

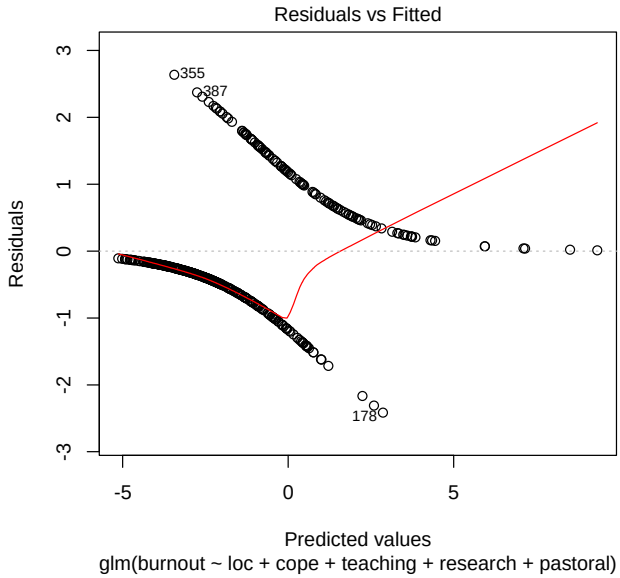
```
Nagelkerke R^2            0.531
```

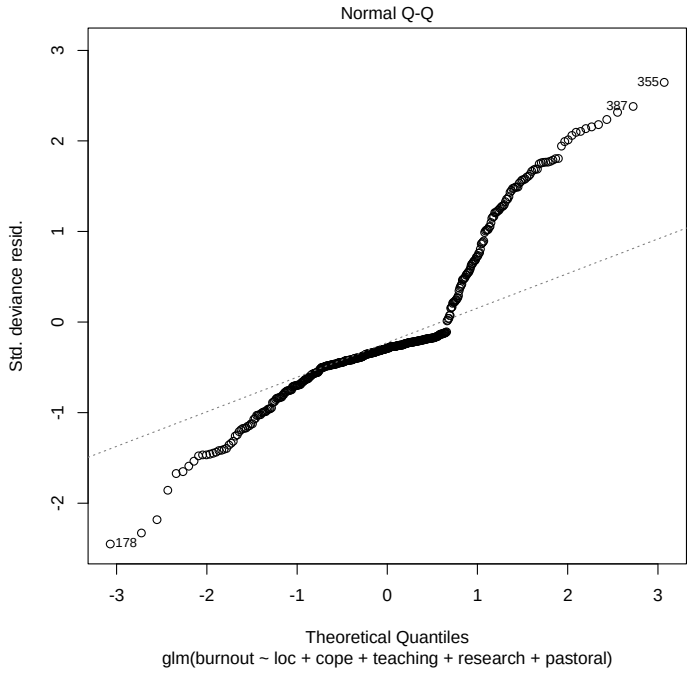
```
exp(burnoutModel.2$coefficients)
```

```
(Intercept)    loc          cope       teaching    research    pastoral
 0.01179680  1.11715594  1.15296414  0.89389904  1.01949919  1.04620942
```

```
exp(confint(burnoutModel.2))
```

```
                2.5 %       97.5 %
(Intercept) 0.001317788 0.09419003
loc         1.086274965 1.15212014
cope        1.118430575 1.19286786
teaching    0.858532732 0.92793154
research    0.999115252 1.04068582
pastoral    1.020119629 1.07403586
```





7. What analysis has been carried out?
    a. Simple regression.
    b. Hierarchical multiple regression.*
    c. Factor analysis.
    d. Reliability analysis.
8. How much variance in burnout does the final model explain?
    a. 3.61–5.31%
    b. 36.1–53.1% *
    c. 0.361–0.531%
    d. 31.3%
9. What does the normal Q-Q plot show?
    a. That the data are not normally distributed.*
    b. Homoscedasticity of errors only.
    c. Independence of errors and homoscedasticity.
    d. Heteroscedasticity and independence of errors.