Solutions

Marking: **K** knowledge, **App** use of the case or data, **An** analysis (a chain of reasons: cause, then effect), **Ev** evaluation (a judgement with a reason). In calculations, **M1** is for the method and **A1** for the answer; **B1** is an independent mark.

Each answer shows the steps, then the **result**.

2.1

(a) quantitative, ratio

(b) categorical, nominal

(c) categorical, ordinal

(d) quantitative, interval

2.2

(a) non-response bias

(b) survivorship bias

(c) selection bias

2.3

weekday $50 \times \dfrac{720}{1{,}200} = \mathbf{30}$

weekend **20**

2.4

$k = \dfrac{1{,}200}{60} = \mathbf{20}$

orders **27** and **47**

3.1

(a) ratio; ordinal; nominal

(b) $\dfrac{1{,}014.2}{37} = \mathbf{27.4}$ minutes

(c) a delivery never takes 0 minutes, so the mean is pulled down; the missing ones are likely busy-time orders, which are slower, so the bias is larger still

(d) that 3 of 40 are missing, why, and that the mean uses the 37 complete cases

3.2

(a) $\dfrac{300}{2{,}000} = \mathbf{15}\%$

(b) lowest $\dfrac{240}{2{,}000} = \mathbf{12}\%$; highest $\dfrac{240 + 1{,}700}{2{,}000} = \mathbf{97}\%$

(c) people who reply may be happier, or angrier, than those who do not, so 80% may not describe all customers; follow up a random sample of non-replies, e.g. by phone

3.3

(a) weekday $80 \times \dfrac{3{,}000}{4{,}000} = \mathbf{60}$; weekend **20**

(b) the question compares two groups; equal sizes give the most precise comparison, while 20 weekend times would leave that mean uncertain

(c) **4**: no: the free times are a convenience sample of Saturday evenings only (App); they miss weekday times and other weekend hours, so they cannot answer the question (An); a larger sample does not fix a biased design (K); judgement: take the 40 and 40 random checks; the free data can describe Saturday evenings only (Ev). **2–3**: the bias named, partly applied. **1**: one relevant point.

exercises.lol exercises.com.cn