

```

1  -- Table: daily_activity
2
3  CREATE TABLE daily_activity_clean (
4      Id BIGINT,
5      ActivityDate DATE,
6      TotalSteps INT,
7      TotalDistance FLOAT,
8      TrackerDistance FLOAT,
9      LoggedActivitiesDistance FLOAT,
10     VeryActiveDistance FLOAT,
11     ModeratelyActiveDistance FLOAT,
12     LightActiveDistance FLOAT,
13     SedentaryActiveDistance FLOAT,
14     VeryActiveMinutes INT,
15     FairlyActiveMinutes INT,
16     LightlyActiveMinutes INT,
17     SedentaryMinutes INT,
18     Calories INT
19 );
20
21 -- Table: sleep_data
22 CREATE TABLE sleep_data_clean (
23     Id BIGINT,
24     SleepDay DATE,
25     TotalSleepRecords INT,
26     TotalMinutesAsleep INT,
27     TotalTimeInBed INT
28 );
29
30
31 INSERT INTO daily_activity_clean (
32     Id,
33     ActivityDate,
34     TotalSteps,
35     TotalDistance,
36     TrackerDistance,
37     LoggedActivitiesDistance,
38     VeryActiveDistance,
39     ModeratelyActiveDistance,
40     LightActiveDistance,
41     SedentaryActiveDistance,
42     VeryActiveMinutes,
43     FairlyActiveMinutes,
44     LightlyActiveMinutes,
45     SedentaryMinutes,
46     Calories
47 )
48 SELECT
49     field1,
50     field2,
51     field3,
52     field4,
53     field5,
54     field6,
55     field7,
56     field8,
57     field9,
58     field10,
59     field11,
60     field12,
61     field13,
62     field14,
63     field15
64 FROM daily_activity;
65
66
67 INSERT INTO sleep_data_clean(
68     Id,
69     SleepDay,
70     TotalSleepRecords,
71     TotalMinutesAsleep,
72     TotalTimeInBed
73 )
74 SELECT
75     field1,
76     field2,

```

```
77     field3,
78     field4,
79     field5
80 FROM sleep_data;
81
82 DROP TABLE daily_activity;
83
84 ALTER TABLE daily_activity_clean RENAME TO daily_activity;
85
86
87 DROP TABLE sleep_data;
88
89 ALTER TABLE sleep_data_clean RENAME TO sleep_data;
90
91
92
93
94 -- View sample data
95 SELECT * FROM daily_activity LIMIT 10;
96 SELECT * FROM sleep_data LIMIT 10;
97
98 -- Count unique users
99 SELECT COUNT(DISTINCT Id) AS unique_users
100 FROM daily_activity;
101
102 -- Null or zero checks
103 SELECT COUNT(*) AS zero_step_days
104 FROM daily_activity
105 WHERE TotalSteps = 0;
106
107 SELECT COUNT(*) AS zero_sleep_days
108 FROM sleep_data
109 WHERE TotalMinutesAsleep = 0;
110
111 -- Data Cleaning
112 DELETE FROM daily_activity
113 WHERE TotalSteps = 0;
114
115 DELETE FROM sleep_data
116 WHERE TotalMinutesAsleep = 0;
117
118 -- Feature Engineering: Add weekday (SQLite)
119 SELECT
120     ActivityDate,
121     strftime('%w', ActivityDate) AS weekday_number,
122     CASE strftime('%w', ActivityDate)
123         WHEN '0' THEN 'Sunday'
124         WHEN '1' THEN 'Monday'
125         WHEN '2' THEN 'Tuesday'
126         WHEN '3' THEN 'Wednesday'
127         WHEN '4' THEN 'Thursday'
128         WHEN '5' THEN 'Friday'
129         WHEN '6' THEN 'Saturday'
130     END AS weekday_name,
131     TotalSteps
132 FROM daily_activity;
133
134 -- Join daily_activity and sleep_data
135 SELECT
136     a.Id,
137     a.ActivityDate,
138     a.TotalSteps,
139     a.Calories,
140     a.VeryActiveMinutes,
141     s.TotalMinutesAsleep,
142     s.TotalTimeInBed
143 FROM daily_activity a
144 JOIN sleep_data s
145     ON a.Id = s.Id AND a.ActivityDate = s.SleepDay;
146
147 -- Summary statistics
148 SELECT
149     Id,
150     ROUND(AVG(TotalSteps), 2) AS avg_steps,
151     ROUND(AVG(Calories), 2) AS avg_calories
152 FROM daily_activity
```

```

153 GROUP BY Id;
154
155 SELECT
156     Id,
157     ROUND(AVG(TotalMinutesAsleep), 2) AS avg_sleep
158 FROM sleep_data
159 GROUP BY Id;
160
161 -- Average steps per weekday
162 SELECT
163     CASE strftime('%w', ActivityDate)
164         WHEN '0' THEN 'Sunday'
165         WHEN '1' THEN 'Monday'
166         WHEN '2' THEN 'Tuesday'
167         WHEN '3' THEN 'Wednesday'
168         WHEN '4' THEN 'Thursday'
169         WHEN '5' THEN 'Friday'
170         WHEN '6' THEN 'Saturday'
171     END AS weekday_name,
172     ROUND(AVG(TotalSteps), 2) AS avg_steps
173 FROM daily_activity
174 GROUP BY weekday_name;
175
176 -- Identify high-activity users
177 SELECT
178     Id,
179     SUM(VeryActiveMinutes) AS total_active_minutes,
180     ROUND(AVG(Calories), 2) AS avg_calories
181 FROM daily_activity
182 GROUP BY Id
183 ORDER BY total_active_minutes DESC
184 LIMIT 5;
185
186 SELECT Id, TotalSteps, Calories
187 FROM daily_activity
188 LIMIT 20;
189

```