

## Solution to Exercise 2: Appending and making new REC files EpiData Analysis

### Key Point(s)

- You can append several files with exactly the same fields and field definitions. The several identical REC and CHK files should have been created from one QES or respectively CHK file.

### Task:

- Start with the program *B\_EX01.PGM*, save it as *B\_EX02.PGM*, and edit it as needed to make the new dataset *ABCD.REC*.
- In the same program, create a new dataset *B\_EX02.REC* file that has a new variable *CASE* which can be either positive or negative and one for WHO age groups. Define as positive any examinee who has at least one bacillus in at least one result and make a tabulation by case (column) and age group (row), stratified by a new variable that groups reason into diagnosis, follow-up (irrespective of month) and unknown reason. Calculate row percentages.

### Solution:

The solution output table:

The first table is the unstratified table:

<b>Bacteriologic case definition</b>				
WHO age groups	Non-case	% Case	% Total	%
0 to 14 years	5 (83.3)	1 (16.7)	6 (100.0)	
15 to 24 years	50 (87.7)	7 (12.3)	57 (100.0)	
25 to 34 years	73 (86.9)	11 (13.1)	84 (100.0)	
35 to 44 years	59 (88.1)	8 (11.9)	67 (100.0)	
45 to 54 years	24 (85.7)	4 (14.3)	28 (100.0)	
55 to 64 years	22 (88.0)	3 (12.0)	25 (100.0)	
65 years and older	26 (86.7)	4 (13.3)	30 (100.0)	
Unknown age	3 (100.0)	0 (0.0)	3 (100.0)	
<b>Total</b>	262 (87.3)	38 (12.7)	300	
Unstratified table				

Then follow the stratified tables:

**Reason grouped: Diagnosis**  
**Bacteriologic case definition**

WHO age groups	Non-case	%	Case	%	Total	%
0 to 14 years	1	(50.0)	1	(50.0)	2	(100.0)
15 to 24 years	20	(76.9)	6	(23.1)	26	(100.0)
25 to 34 years	29	(76.3)	9	(23.7)	38	(100.0)
35 to 44 years	17	(70.8)	7	(29.2)	24	(100.0)
45 to 54 years	9	(69.2)	4	(30.8)	13	(100.0)
55 to 64 years	12	(100.0)	0	(0.0)	12	(100.0)
65 years and older	14	(82.4)	3	(17.6)	17	(100.0)
Unknown age	2	(100.0)	0	(0.0)	2	(100.0)
<b>Total</b>	<b>104</b>	<b>(77.6)</b>	<b>30</b>	<b>(22.4)</b>	<b>134</b>	

Percents: (Row)

**Reason grouped: Follow-up**  
**Bacteriologic case definition**

WHO age groups	Non-case	%	Case	%	Total	%
0 to 14 years	4	(100.0)	0	(0.0)	4	(100.0)
15 to 24 years	30	(96.8)	1	(3.2)	31	(100.0)
25 to 34 years	44	(97.8)	1	(2.2)	45	(100.0)
35 to 44 years	42	(97.7)	1	(2.3)	43	(100.0)
45 to 54 years	15	(100.0)	0	(0.0)	15	(100.0)
55 to 64 years	9	(75.0)	3	(25.0)	12	(100.0)
65 years and older	12	(92.3)	1	(7.7)	13	(100.0)
Unknown age	1	(100.0)	0	(0.0)	1	(100.0)
<b>Total</b>	<b>157</b>	<b>(95.7)</b>	<b>7</b>	<b>(4.3)</b>	<b>164</b>	

Percents: (Row)

**Reason grouped: Unknown**  
**Bacteriologic case definition**

WHO age groups	Non-case	%	Case	%	Total	%
0 to 14 years	0	(0.0)	0	(0.0)	0	(100.0)
15 to 24 years	0	(0.0)	0	(0.0)	0	(100.0)
25 to 34 years	0	(0.0)	1	(100.0)	1	(100.0)
35 to 44 years	0	(0.0)	0	(0.0)	0	(100.0)
45 to 54 years	0	(0.0)	0	(0.0)	0	(100.0)
55 to 64 years	1	(100.0)	0	(0.0)	1	(100.0)
65 years and older	0	(0.0)	0	(0.0)	0	(100.0)
Unknown age	0	(0.0)	0	(0.0)	0	(100.0)
<b>Total</b>	<b>1</b>	<b>(50.0)</b>	<b>1</b>	<b>(50.0)</b>	<b>2</b>	

Percents: (Row)

The program B\_EX02.PGM might look as follows:

```
* This is b_ex02 EpiData Analysis program

cls
close

read "a.rec"
append /file="b.rec"
append /file="c.rec"
append /file="d.rec"
drop serno id
savedata "abcd.rec" /replace

close
read "abcd.rec"

* Definition positive: any AFB in any of three cases
* Values: "pos" or "neg" or "unk"

cls
                                define case #
                                let case=0
if (res1>0.0) and (res1<9) then case=1
if (res2>0.0) and (res2<9) then case=1
if (res3>0.0) and (res3<9) then case=1

labelvalue case /0="Non-case" /1="Case"
label case "Bacteriologic case definition"

* define WHO age groups
                                define agegrp #
                                agegrp=8
if age>=00 and age<15 then agegrp=1
if age>=15 and age<25 then agegrp=2
if age>=25 and age<35 then agegrp=3
if age>=35 and age<45 then agegrp=4
if age>=45 and age<55 then agegrp=5
if age>=55 and age<65 then agegrp=6
if age>=65 and age<99 then agegrp=7
if age>=99
                                then agegrp=9
label agegrp "WHO age groups"
labelvalue agegrp /1="0 to 14 years"
labelvalue agegrp /2="15 to 24 years"
labelvalue agegrp /3="25 to 34 years"
labelvalue agegrp /4="35 to 44 years"
labelvalue agegrp /5="45 to 54 years"
labelvalue agegrp /6="55 to 64 years"
labelvalue agegrp /7="65 years and older"
labelvalue agegrp /8="Other" //Note this should not exist - control
labelvalue agegrp /9="Unknown age"

define reason2 #
reason2=2
if reason=0 then reason2=1
if reason=9 then reason2=9
label reason2 "Reason grouped"
labelvalue reason2 /1="Diagnosis"
labelvalue reason2 /2="Follow-up"
labelvalue reason2 /9="Unknown"
```

```
savedata "b_ex02.rec" /replace
```

```
cls
```

```
close
```

```
read "b_ex02.rec"
```

```
cls
```

```
tables case agegrp reason2 /r
```