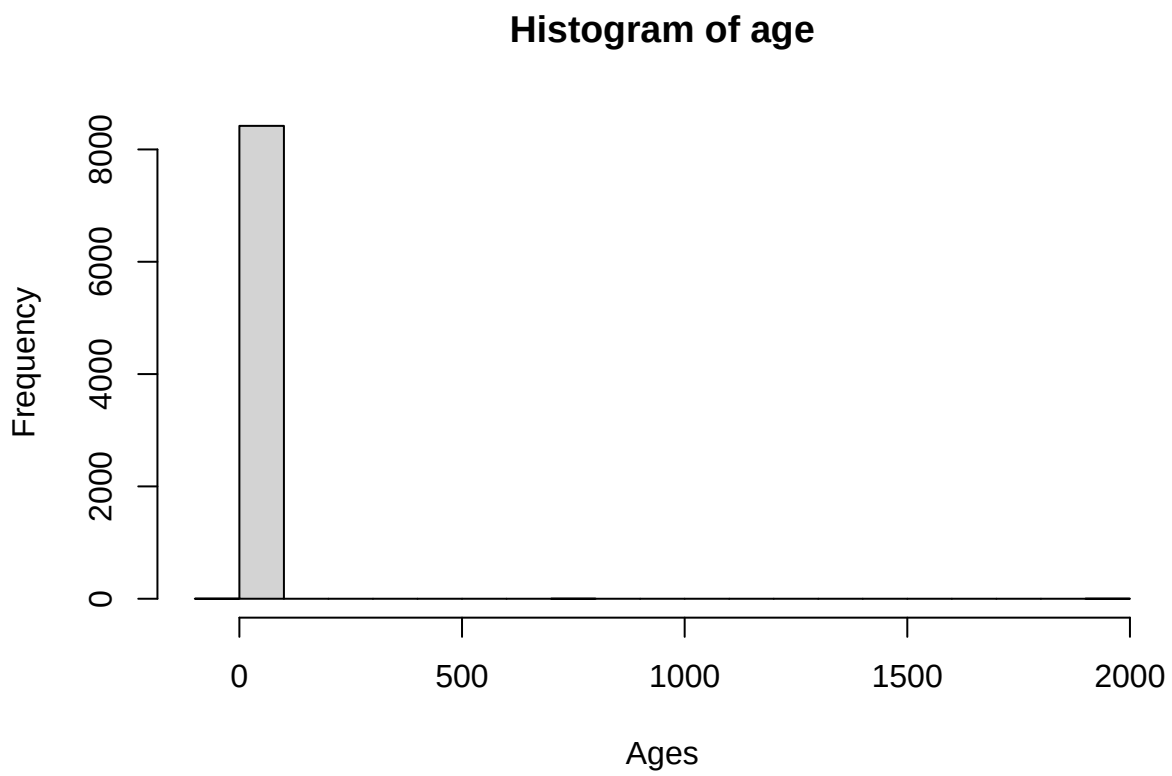
Cleaning HCWS

Check the gender

```
table(hcws$d_gender, useNA = "always") %>% pander()
```

Female	Male	NA
3858	3319	1594

```
hist(hcws$d_age_i, main = "Histogram of age", xlab = "Ages")
```



Display the age in descending order and identify the extreme and improbable values

```
hcws %>%
  filter(d_age_i < 18 | d_age_i > 75) %>%
  select(ID = ctry_id, Country = country, Age = d_age_i) %>%
```

```
  arrange(Age) %>%
  pander()
```

ID	Country	Age
2321	Ghana	-37
4528	Ghana	2
1837	Ghana	2
1782	Ghana	3
1214	Ghana	3
1215	Ghana	3
3662	Ghana	4
1053	Ghana	4
376	Ghana	6
4045	Ghana	6
2212	Ghana	12
81	Cameroon	15
469	Cameroon	16
529	Uganda	17
1005	Uganda	17
4110	Ghana	760
1781	Ghana	1990

What should we do to rectify this?

Marital status

Use the new marital status grouping

```
hcws %$% table(marital) %>% pander()
```

Married/Partnership	Prefer not to say	Single
4323	590	3499

```
table(hcws$education) %>% pander()
```

Table 4: Table continues below

Certificate	Diploma	High School	High School/Middle School
48	25	833	1

Post-Graduate	Remove?	Undergraduate
1204	2	3260

The levels targeted for removal. Their entries are suspicious

```
hcws %>%
  filter(education == "Remove?") %>%
  select(ID = ctry_id, Country = country, Education = d_edu_c_o, Other = educ_other, Profession = d_job
```

```
## # A tibble: 2 x 6
##      ID Country Education Other            Profession     Type
##   <dbl> <chr>   <chr>     <chr>            <chr>          <chr>
## 1  1455 Ghana   Other     Doctorate degree Pharmacist     Prefer not to say
## 2  3524 Ghana   Other     Kindergarten     Medical Doctor Prefer not to say
```

Suggestion

Rename `Undergraduate` correctly as graduate, then all others under it as undergraduate.

Number of children

Different countries used different scales. As such, it has to be combined as one. The new groupings are seen below

```
hcws %$% table(children) %>%
  pander()
```

0	1	2 to 4	More than 4
2291	1653	3890	582

Hospitals

```
hcws %$%
  table(hospital, country) %>%
  pander()
```

	Cameroon	Ghana	Senegal	Uganda
Denomination	142	674	0	0
Hospital	0	1025	294	0
Level 1 Hospital	0	0	34	0
Level 2 Hospital	0	391	31	0
Level 3 hospital	0	0	125	0
Level 4 Hospital	0	277	0	0
Private hospital	2405	753	27	363
Teaching hospital	0	1078	0	0
University hospital	0	156	0	0

Senegal classified the hospital as Level 1 to 3. It would be easier (and best in my opinion) to arrange the hospitals as such.

Job roles

```
hcws %$%
  table(job_role) %>%
  pander()
```

Table 8: Table continues below

Administrative staff	Dentist	Diagnostic staff	Health Assistant
66	20	1	406

Table 9: Table continues below

Laboratory staff	Non-Clinical staff	Nurse Assistant	Nurse/Midwife
9	179	301	2265

Table 10: Table continues below

Nutritionist/Dietician	Pharmacy	Pharmacy Assistant	Physician/Surgeon
1	7	4	1991

Public Health Worker	Sanitation	Student
148	1	13

The grouping above closely matches the stated job role or as indicated in the "other". We cal combine all others as 1, possible removing responses from students.