Analysis of LDSQ/ CAIDS-Q data-set. Report 5-11th Feb

Pablo M. Rodriguez

8th February 2018

1. Summary statistics for categorical variables

Adults

```
##     Gender        Area.of.residence           Ethnicity         Religion
##  Female:681   GM-Manchester: 87     White British  :308   None      :470
##  Male   : 63   Bolton       : 32     White - British:300   Christian :171
##  NA's   :  1   GM-Bolton    : 32     Pakistani      : 23   Not stated: 42
##                GM-Salford   : 26     White - other  : 15   Muslim    : 32
##                GM-Rochdale  : 24     White Other    :  9   Other     :  3
##                (Other)      :537     (Other)        : 88   (Other)   :  5
##                NA's         :  7     NA's           :  2   NA's      : 22
##     DV.history  DASH.Score         Relationship.to.alleged.perp.
##  No      :192   <14 : 46   Stranger                :140
##  Yes     :156   >=14: 24   Acquaintance < 24 hours: 84
##  Current : 98   >14 : 46   Partner                 : 81
##  Historic: 48   NA's:629   Ex partner              : 73
##  None    :221              Friend                  : 60
##  Unknown :  2              (Other)                 :294
##  NA's    : 28              NA's                    : 13
##         No..of.perps. Met.on.internet             Assault.type.1
##  Multiple perps: 47   No  :662        Penile Vaginal     :220
##  Single perp   :647   Yes : 57        Vaginal rape       :200
##  Unknown       : 35   NA's: 26        Unknown            :123
##  NA's          : 16                   Digital penetration: 43
##                                       Anal rape          : 31
##                                       (Other)            : 99
##                                       NA's               : 29
##              Assault.type.2            Assault.type.3
##  Penile Oral        : 41    Digital penetration: 11
##  Penile Anal        : 31    Oral rape          :  9
##  Oral rape          : 30    Penile Anal        :  7
##  Vaginal rape       : 27    Digital Vaginal    :  6
##  Digital penetration: 26    Cunnilingus        :  4
##  (Other)            : 74    (Other)            : 22
##  NA's               :516    NA's               :686
##          Assault.type.4 Strangulation   Sex.Worker
##  Penile Oral    : 41    No    :601    No     :658
##  Penile Anal    : 31    Yes   : 69    Yes    :  8
##  Digital Vaginal: 25    Unkown: 24    UnkNown:  2
##  Digital Anal   :  3    NA's  : 51    NA's   : 77
##  Cunnilingus    :  2
##  (Other)        : 11
##  NA's           :632
##  Emergency.contraception         HIV.PEP             Hep.B
```

```
##  Declined    : 29        Declined    : 19   Declined    : 82
##  No          : 13        No          : 48   No          : 26
##  Yes         :288        Yes         : 98   Yes         :255
##  Not required:186        Not required:298   Not required:180
##  NA's        :229        NA's        :282   NA's        :202
##
##
##                       Referrer        LDSQ..
##  Police                   :594   100     :353
##  Self                     :122   86      :167
##  GP/other health referrals:  8   71      : 51
##  3rd Sector referrals     :  3   Not done: 39
##  Social Services          :  2   57      : 27
##  (Other)                  :  2   (Other) : 61
##  NA's                     : 14   NA's    :  1
```

Children

```
##              FME.context     Gender        Area.of.residence
##  Acute             :253   Female:383   GM-Manchester: 49
##  Historic          :175   Male  : 60   Cheshire East: 22
##  Forensic follow-up:  7   NA's  :  3   GM-Bolton    : 20
##  NA's              : 11                GM-Salford   : 20
##                                        GM-Stockport : 19
##                                        (Other)      :311
##                                        NA's         :  5
##           Ethnicity                    Religion   DV.history CSE.CSE.risk
##  White - British:194   Christian           : 98   No  :288   No  :297
##  White British  :158   Muslim              : 38   Yes : 70   Yes : 91
##  Black - African: 17   None                :263   NA's: 88   NA's: 58
##  Pakistani      : 13   Not stated          : 22
##  African        :  8   Other (please state):  2
##  (Other)        : 50   Other               :  5
##  NA's           :  6   NA's                : 18
##         Relationship.to.alleged.perp.        No..of.perps.
##  Father                  : 41           Multiple perps: 35
##  Unknown                 : 41           Single perp   :364
##  Acquaintance > 24 hours: 37            Unknown       : 29
##  Stranger                : 34           NA's          : 18
##  Friend                  : 31
##  (Other)                 :243
##  NA's                    : 19
##  Met.on.internet             Assault.type.1             Assault.type.2
##  No  :403        Penile Vaginal     : 82    Penile Oral        : 15
##  Yes : 25        Vaginal rape       : 82    Digital penetration: 14
##  NA's: 18        Unknown            : 70    Vaginal rape       : 14
##                  Digital penetration: 41    Digital Vaginal    : 11
##                  Digital Vaginal    : 28    Oral rape          : 11
##                  (Other)            :120    (Other)            : 40
##                  NA's               : 23    NA's               :341
##                  Assault.type.3 Emergency.contraception
##  Anal rape              :  6    Declined    :  5
##  Digital Vaginal        :  4    No          :  6
##  Non penetrative assault:  3    Yes         : 79
```

```
##  Fellatio               :  2    Not required:208
##  Cunnilingus            :  1    NA's        :148
##  (Other)                :  7
##  NA's                   :423
##             HIV.PEP                Hep.B         U.16.DVD
##  Declined       :  8   Declined       : 44   Declined:  3
##  No             : 20   No             : 16   No      :148
##  Yes            :  8   Yes            : 71   Yes     :192
##  not appropriate:  2   No required    :185   NA's    :103
##  Not required   :238   not appropriate:  1
##  NA's           :170   NA's           :129
##
##                       Referrer       CAIDSQ..
##  Police                   :359   100     :158
##  Social Services          : 32   86      : 62
##  Police/Social Services   : 16   Not done: 30
##  Self                     : 11   57      : 21
##  GP/other health referrals:  5   71      : 19
##  (Other)                  : 10   (Other) : 15
##  NA's                     : 13   NA's    :141
```

2. Summary statistics for numeric variables

Adults

```
##       Age
##  Min.   :15.00
##  1st Qu.:21.00
##  Median :26.00
##  Mean   :28.95
##  3rd Qu.:35.00
##  Max.   :96.00
##  NA's   :5
```

Children

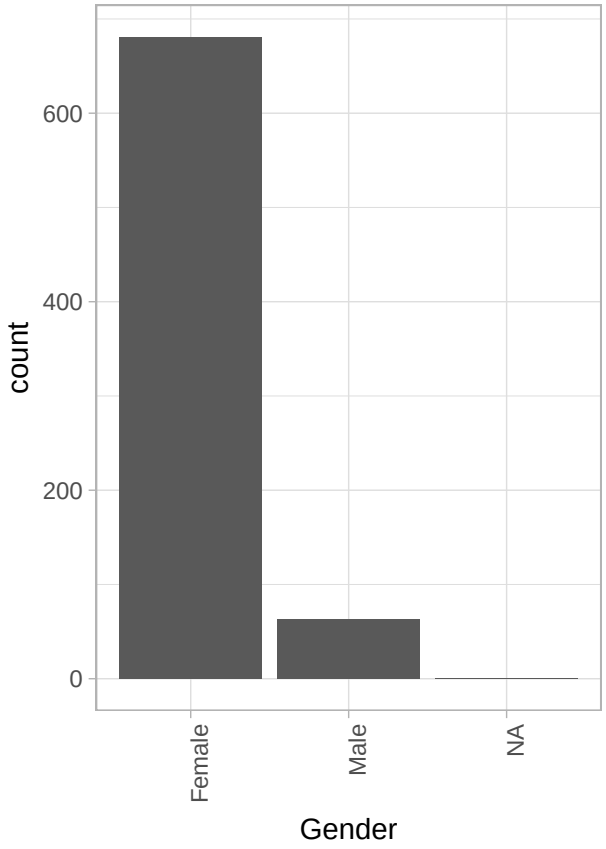
```
##       Age
##  Min.   : 0.00
##  1st Qu.: 7.00
##  Median :13.00
##  Mean   :11.11
##  3rd Qu.:15.00
##  Max.   :17.00
##  NA's   :2
```

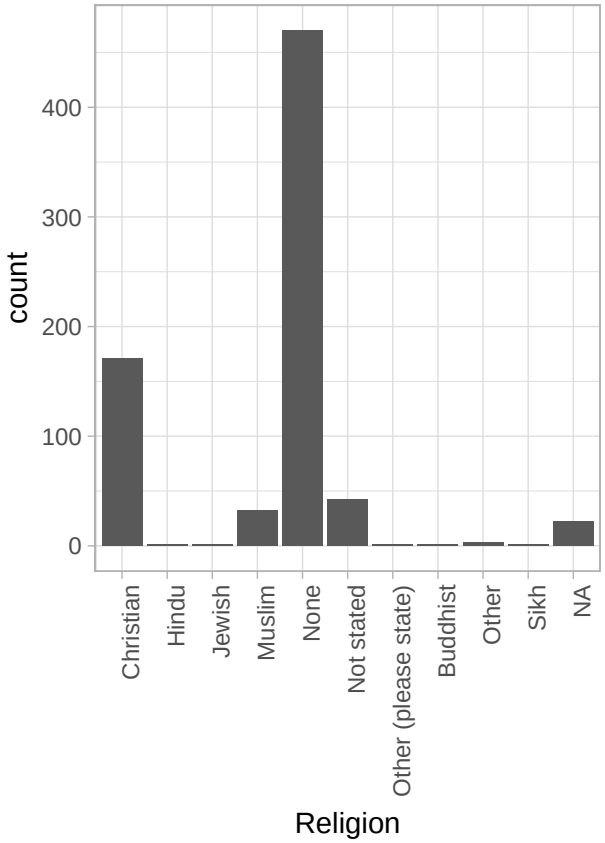
Age is the only numeric variable. FME also is, left aside because getting the difference in mins is troubling.

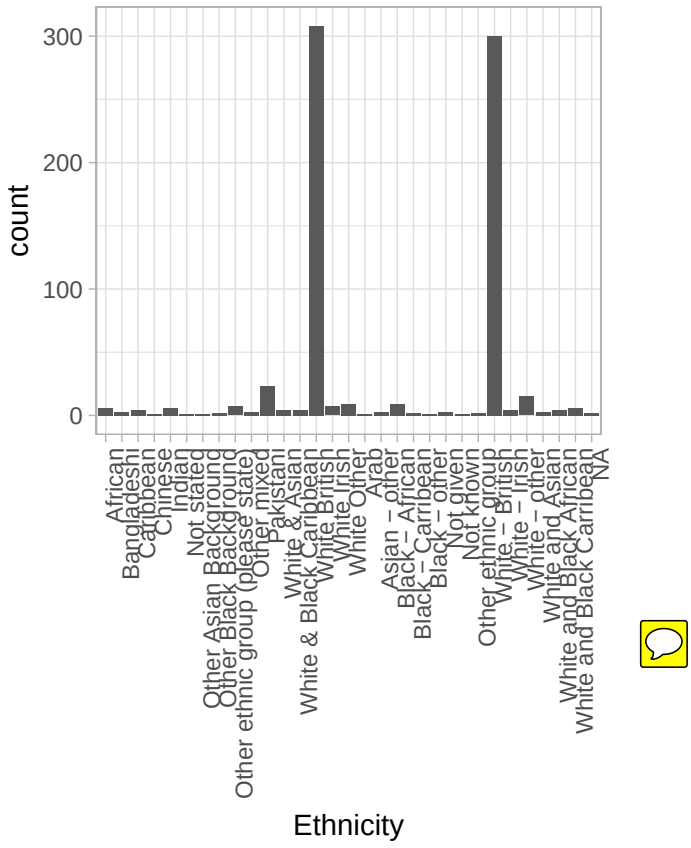
2. Histograms for categorical variables

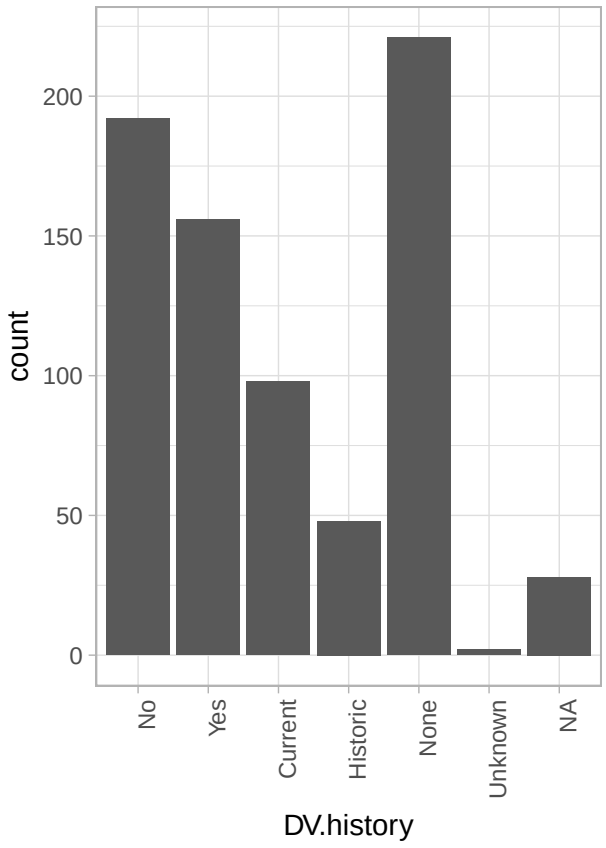
NAs are included in each. Representing this can be meaningful for some variables. To avoid overplotting, area of residence have been omitted as it is the column with more distinct values, which takes a lot of space to cover properly. Basic information about this variable can be found above

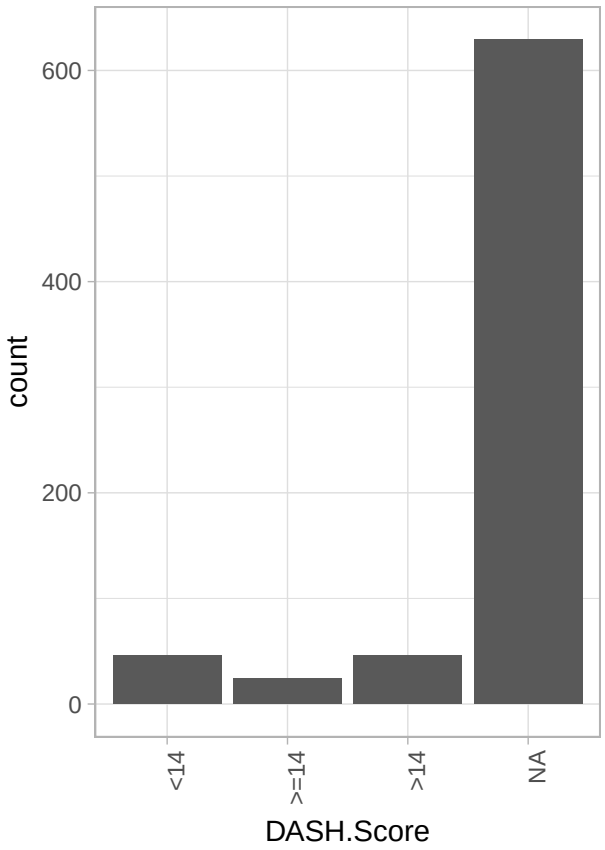
Adults

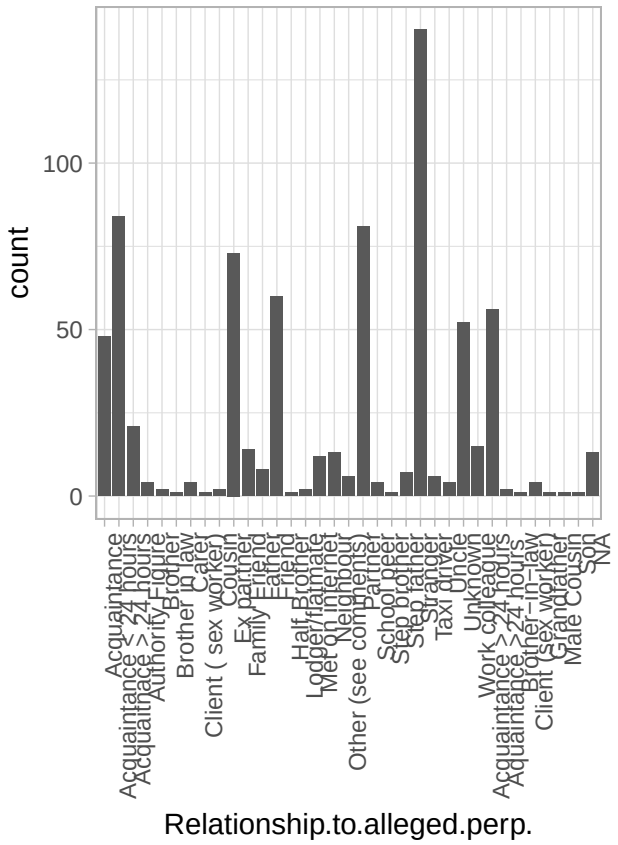

600
400
200
0
count
Female
Male
NA
Gender

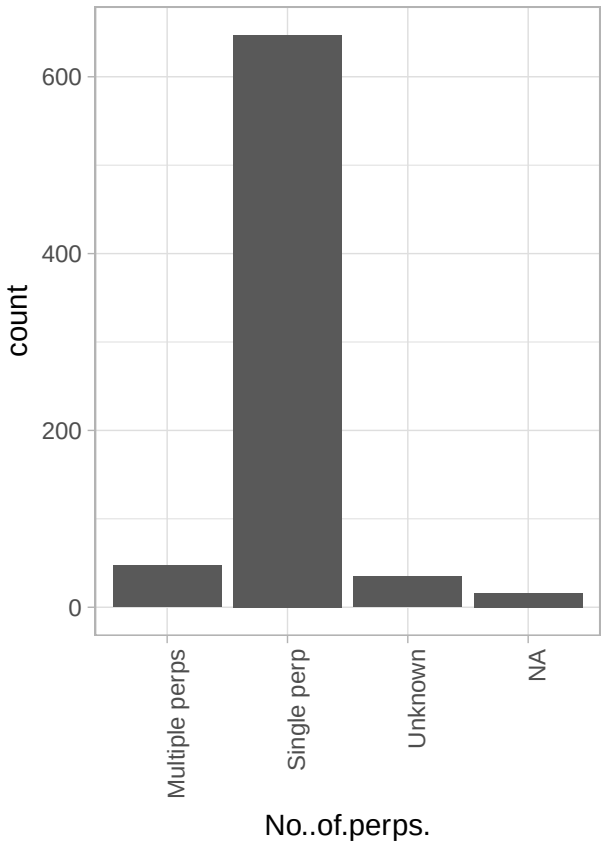

400
300
200
100
0
count
Christian
Hindu
Jewish
Muslim
None
Not stated
Other (please state)
Buddhist
Other
Sikh
NA
Religion

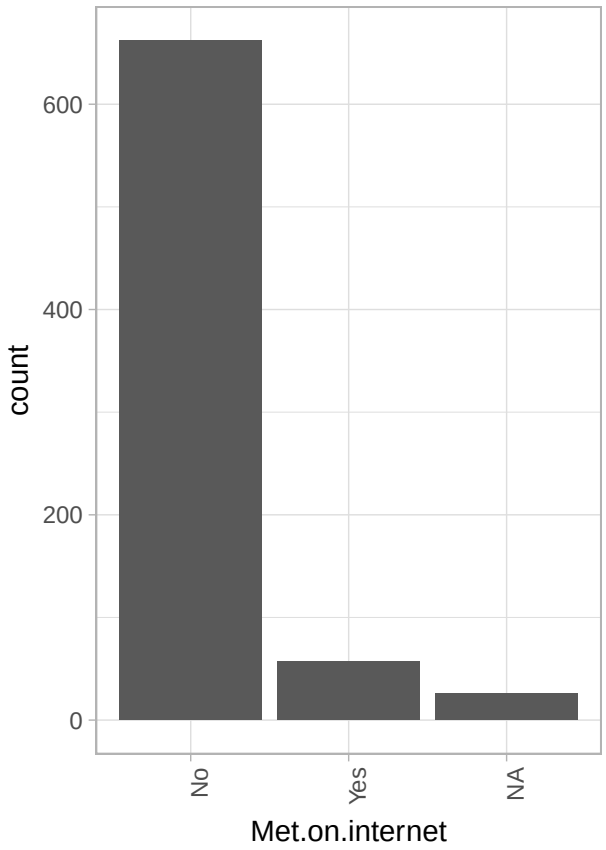

count
300
200
100
0
African
Bangladeshi
Caribbean
Chinese
Indian
Not stated
Other Asian Background
Other Black Background
Other ethnic group (please state)
Other mixed
Pakistani
White & Asian
White & Black Caribbean
White British
White Irish
White Other
Arab
Asian – other
Black – African
Black – Carribean
Black – other
Not given
Not known
Other ethnic group
White – British
White – Irish
White – other
White and Asian
White and Black African
White and Black Carribean
NA
Ethnicity

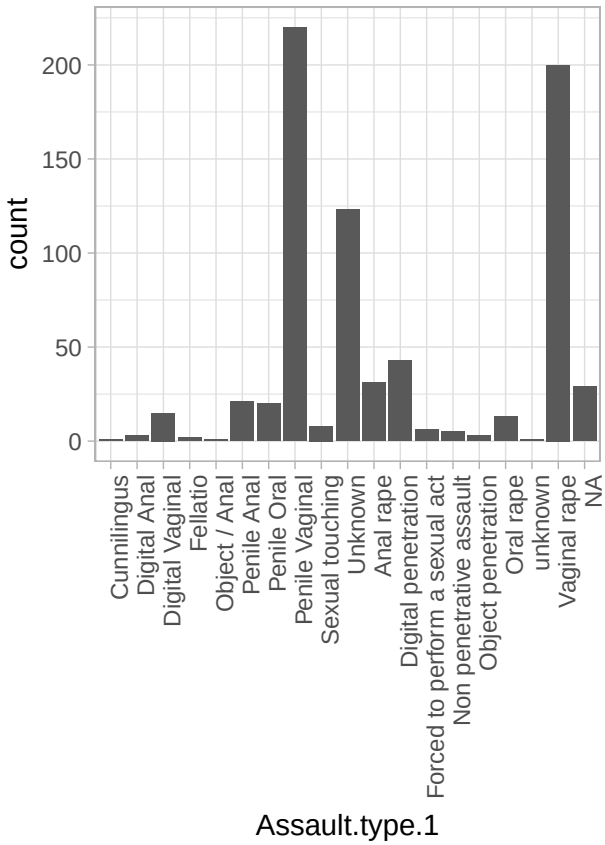

count
200
150
100
50
0
No
Yes
Current
Historic
None
Unknown
NA
DV.history

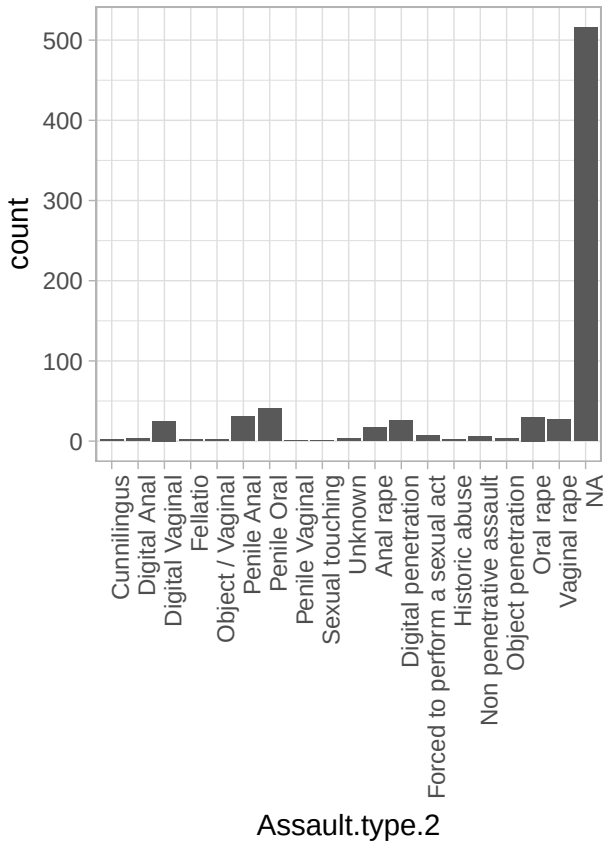

count
600
400
200
0
<14
>=14
>14
NA
DASH.Score

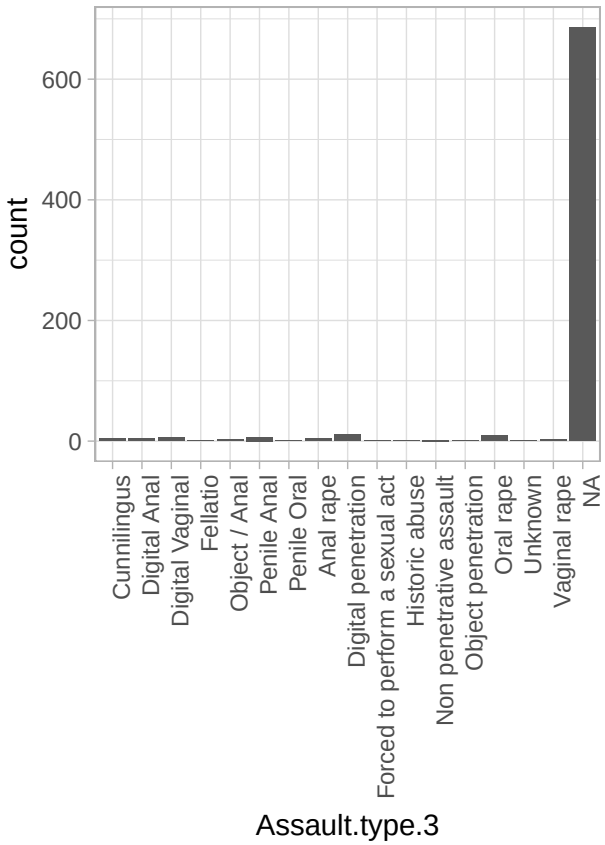

count
100
50
0
Acquaintance
Acquaintance < 24 hours
Acquaintance > 24 hours
Authority figure
Brother
Brother in law
Carer
Client (sex worker)
Cousin
Ex partner
Family member
Friend
Half Brother
Lodger/flatmate
Met on internet
Neighbour
Other (see comments)
Partner
School peer
Step brother
Stranger
Taxi driver
Uncle
Unknown
Work colleague
Acquaintance < 24 hours
Acquaintance > 24 hours
Brother-in-law
Client (sex worker)
Grandfather
Male Cousin
Son
NA
Relationship.to.alleged.perp.

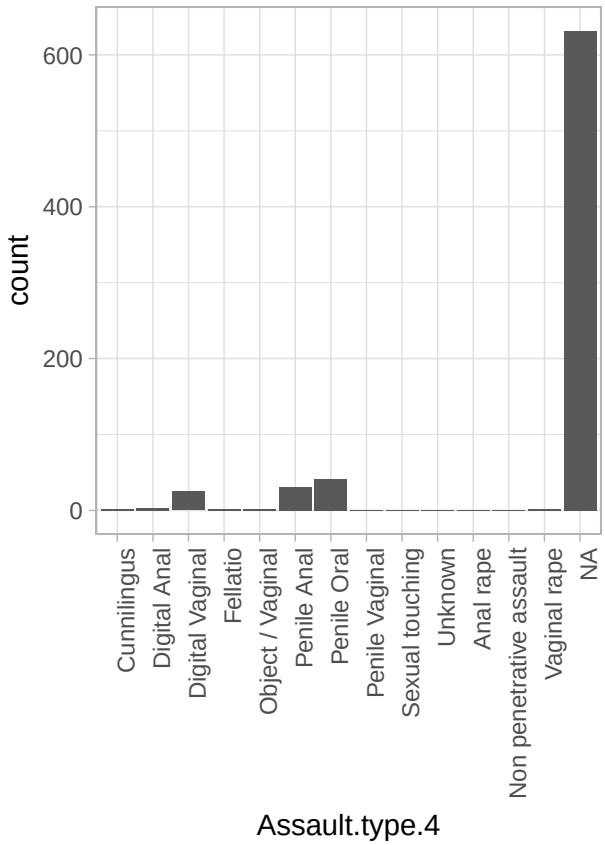

count
600
400
200
0
Multiple perps
Single perp
Unknown
NA
No..of.perps.

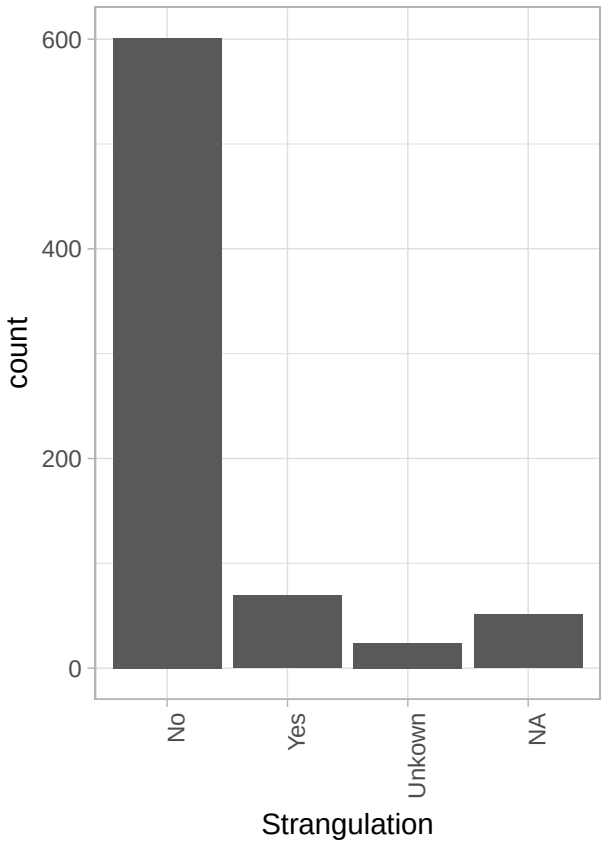

count
600
400
200
0
No
Yes
NA
Met.on.internet

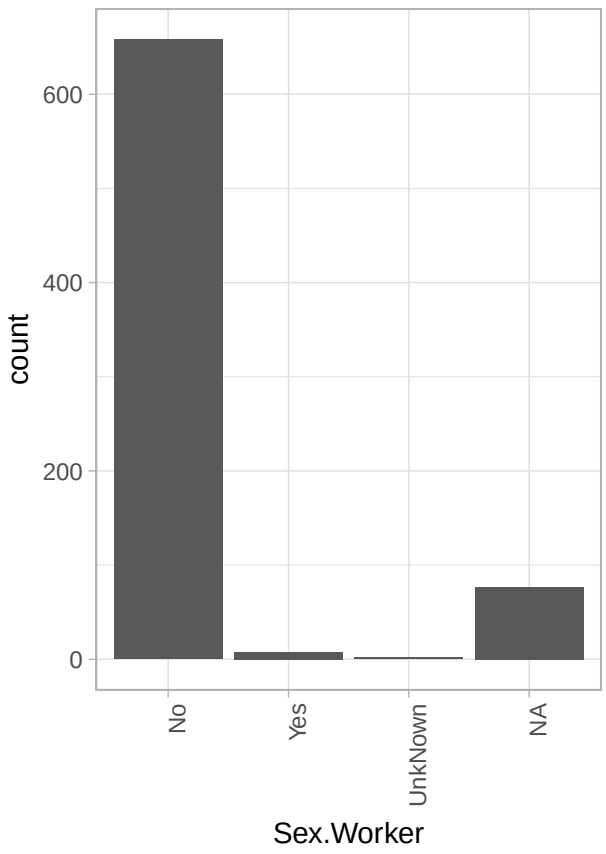

count
200
150
100
50
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Object / Anal
Penile Anal
Penile Oral
Penile Vaginal
Sexual touching
Unknown
Anal rape
Digital penetration
Forced to perform a sexual act
Non penetrative assault
Object penetration
Oral rape
unknown
Vaginal rape
NA
Assault.type.1

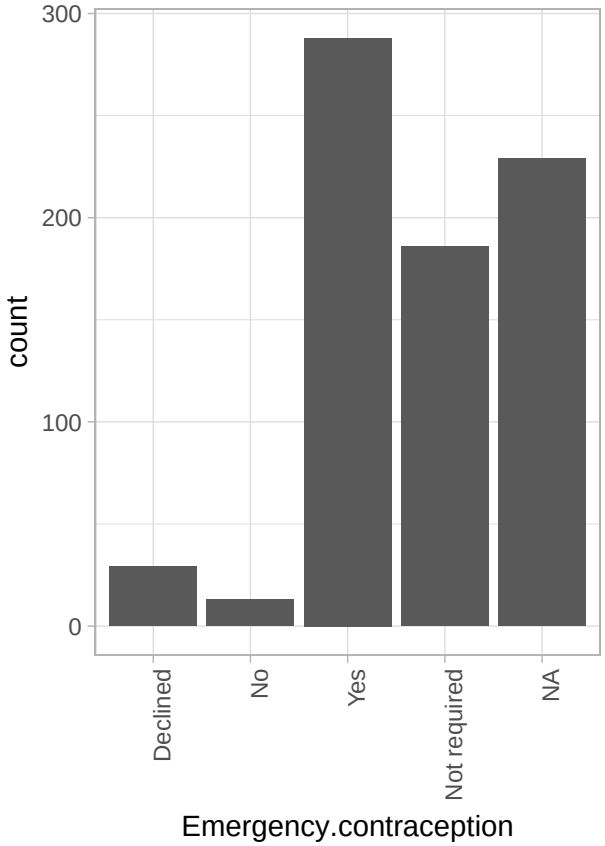

count
500
400
300
200
100
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Object / Vaginal
Penile Anal
Penile Oral
Penile Vaginal
Sexual touching
Unknown
Anal rape
Digital penetration
Forced to perform a sexual act
Historic abuse
Non penetrative assault
Object penetration
Oral rape
Vaginal rape
NA
Assault.type.2

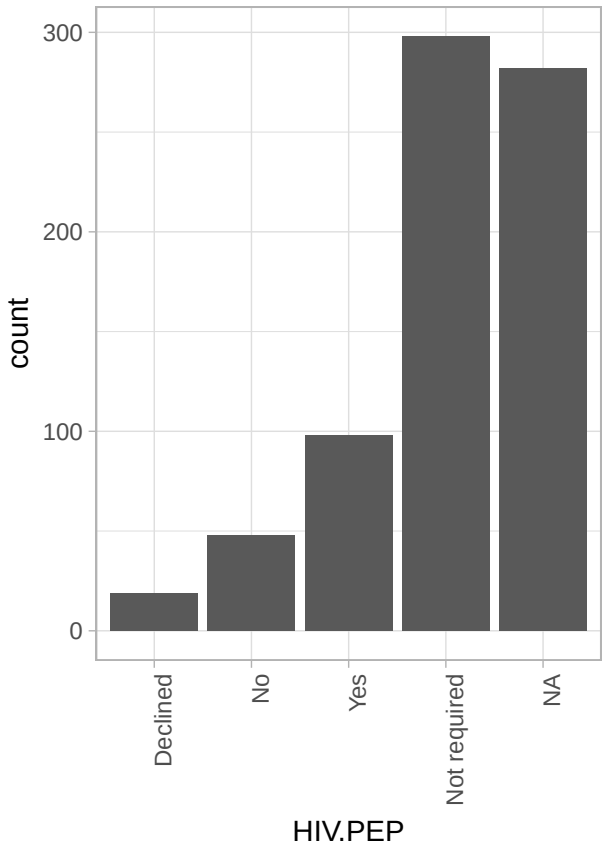

count
600
400
200
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Object / Anal
Penile Anal
Penile Oral
Anal rape
Digital penetration
Forced to perform a sexual act
Historic abuse
Non penetrative assault
Object penetration
Oral rape
Unknown
Vaginal rape
NA
Assault.type.3

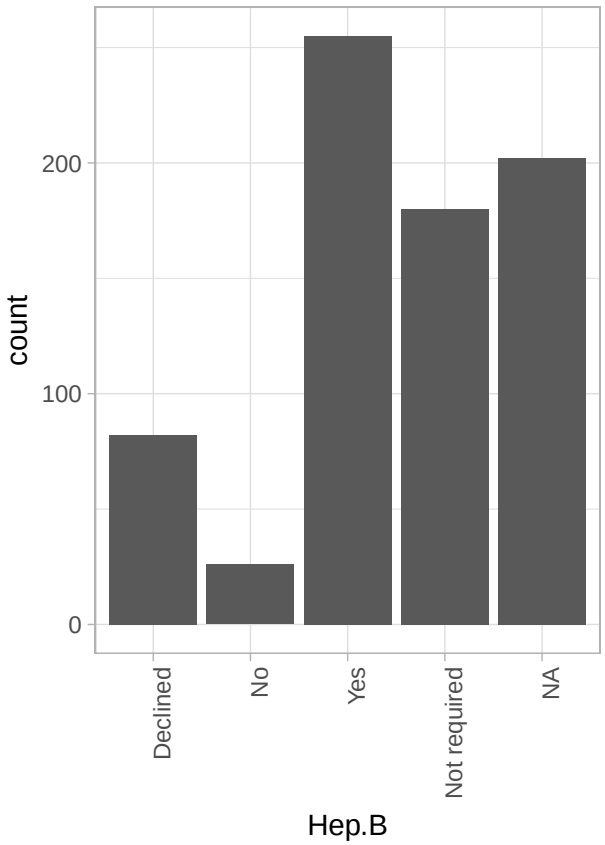

count
600
400
200
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Object / Vaginal
Penile Anal
Penile Oral
Penile Vaginal
Sexual touching
Unknown
Anal rape
Non penetrative assault
Vaginal rape
NA
Assault.type.4

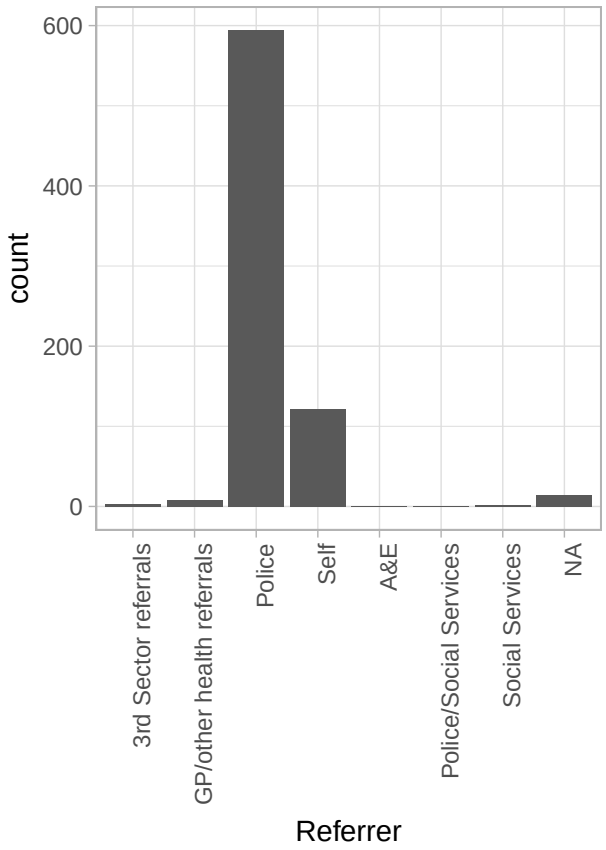

count
600
400
200
0
No
Yes
Unkown
NA
Strangulation

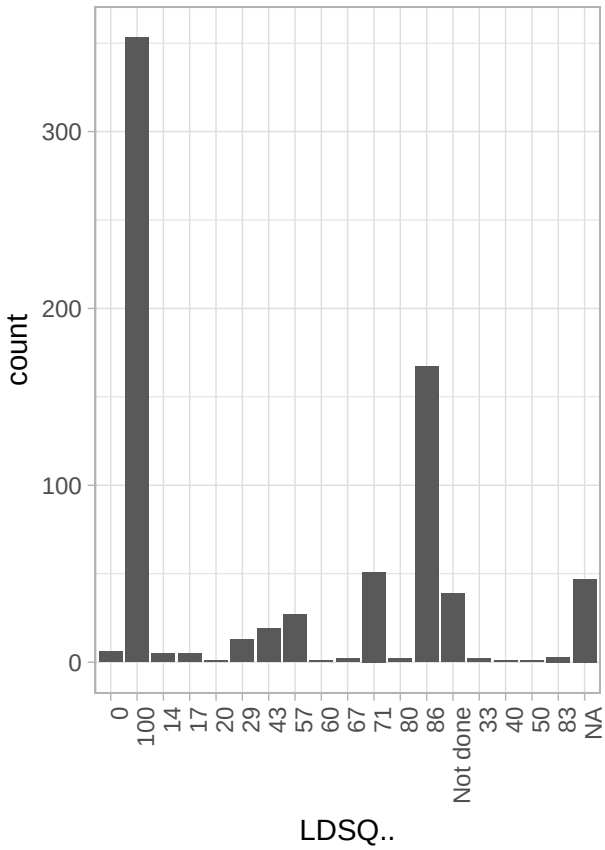

600
400
200
0
count
No
Yes
UnkNown
NA
Sex.Worker


300
200
100
0
count
Declined
No
Yes
Not required
NA
Emergency.contraception

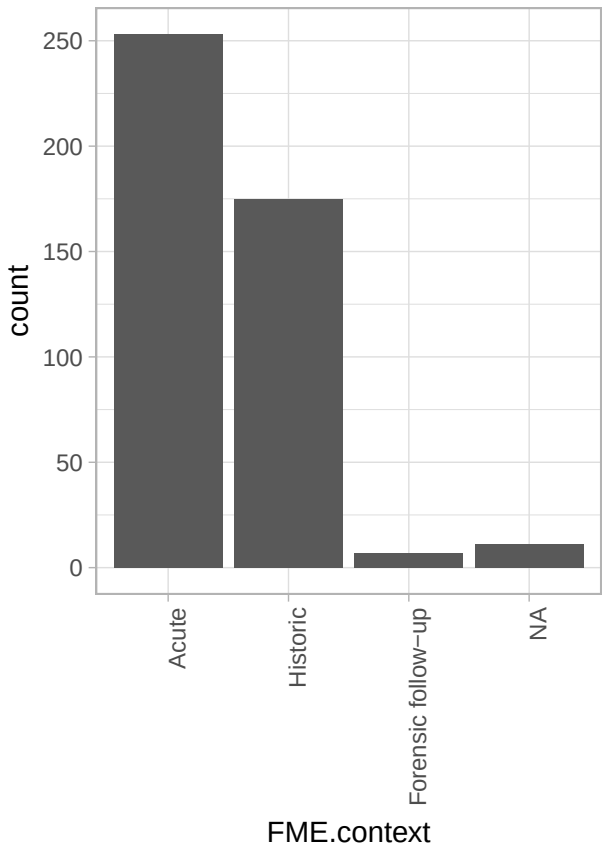

300
200
100
0
count
Declined
No
Yes
Not required
NA
HIV.PEP

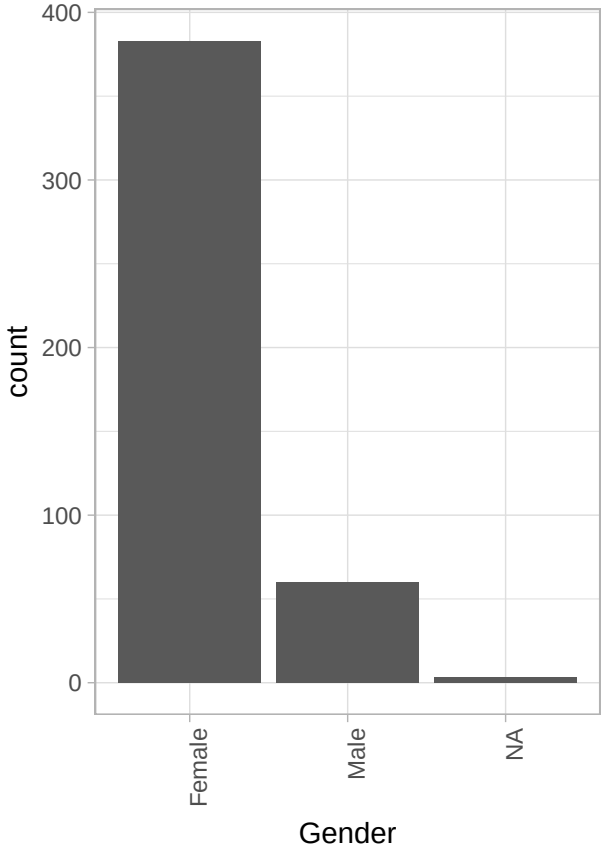

200
100
0
count
Declined
No
Yes
Not required
NA
Hep.B

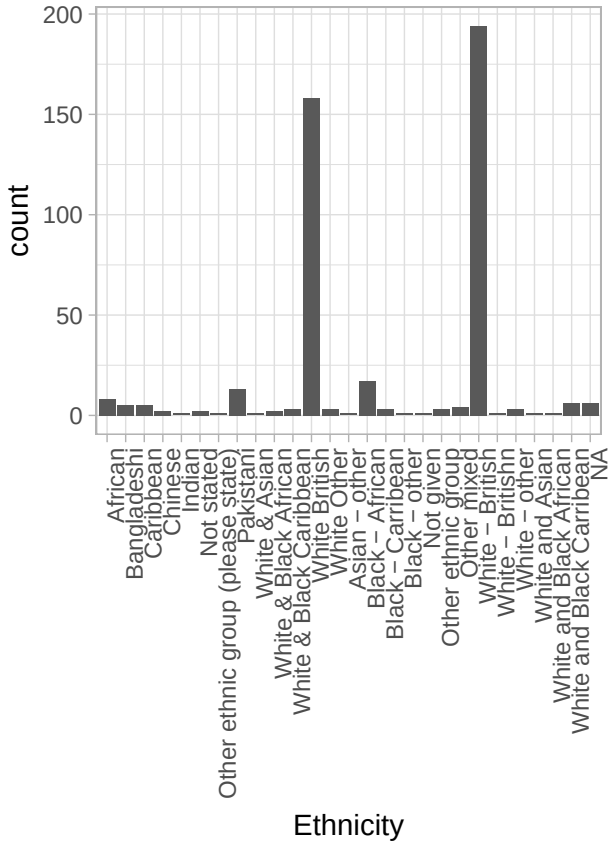

600
400
200
0
count
3rd Sector referrals
GP/other health referrals
Police
Self
A&E
Police/Social Services
Social Services
NA
Referrer

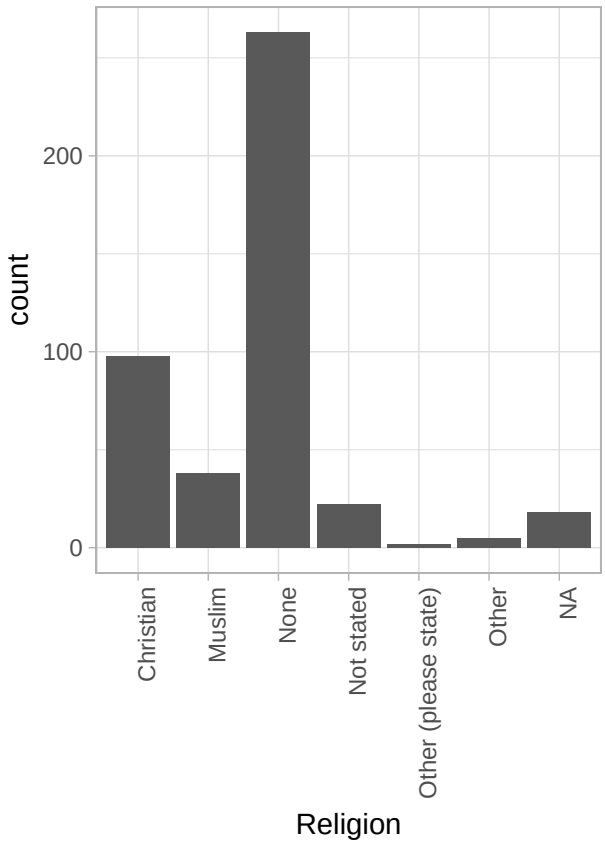

300
200
100
0
count
0
100
14
17
20
29
43
57
60
67
71
80
86
Not done
33
40
50
83
NA
LDSQ..

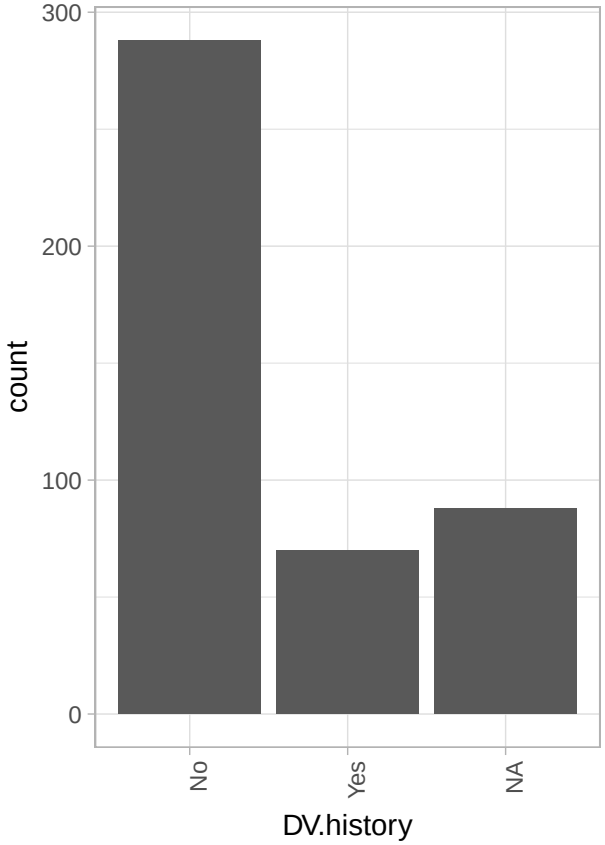
Children

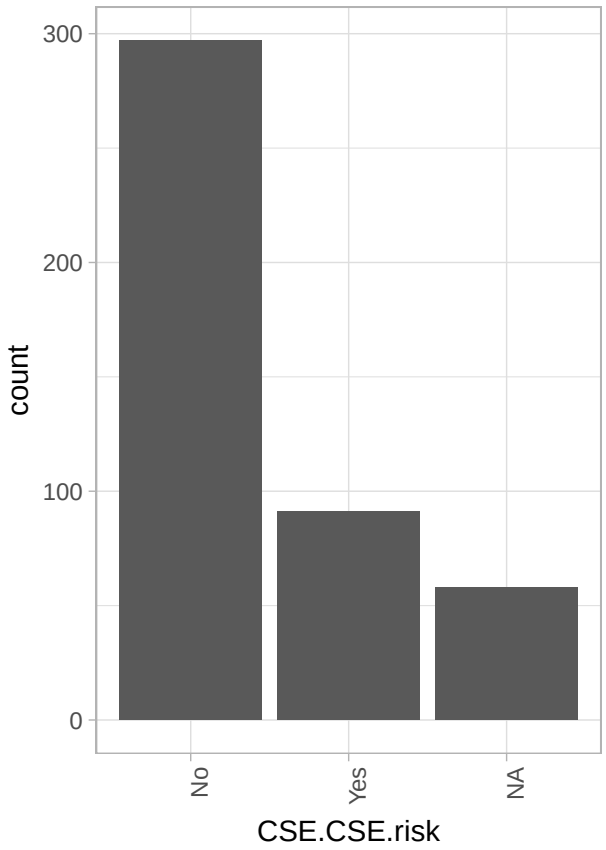

250
200
150
100
50
0
count
Acute
Historic
Forensic follow−up
NA
FME.context

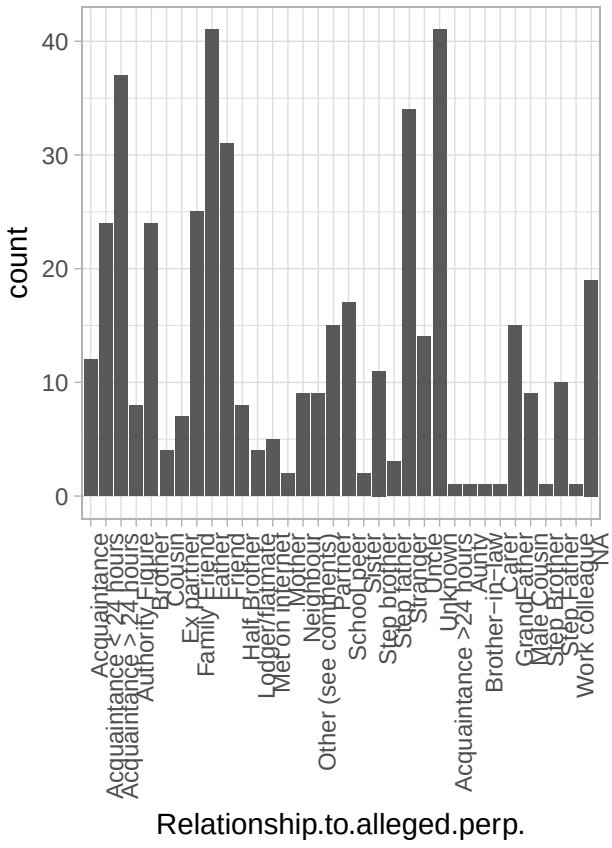

400
300
200
100
0
count
Female
Male
NA
Gender

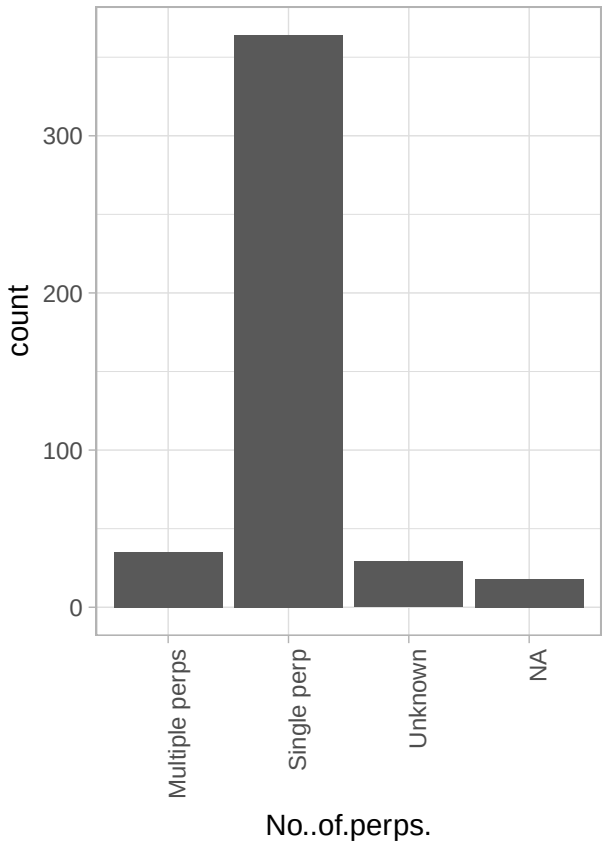

count
200
150
100
50
0
African
Bangladeshi
Caribbean
Chinese
Indian
Not stated
Other ethnic group (please state)
Pakistani
White & Asian
White & Black African
White & Black Caribbean
White British
White Other
Asian – other
Black – African
Black – Carribean
Black – other
Not given
Other ethnic group
Other mixed
White – British
White – Britishn
White – other
White and Asian
White and Black African
White and Black Carribean
NA
Ethnicity

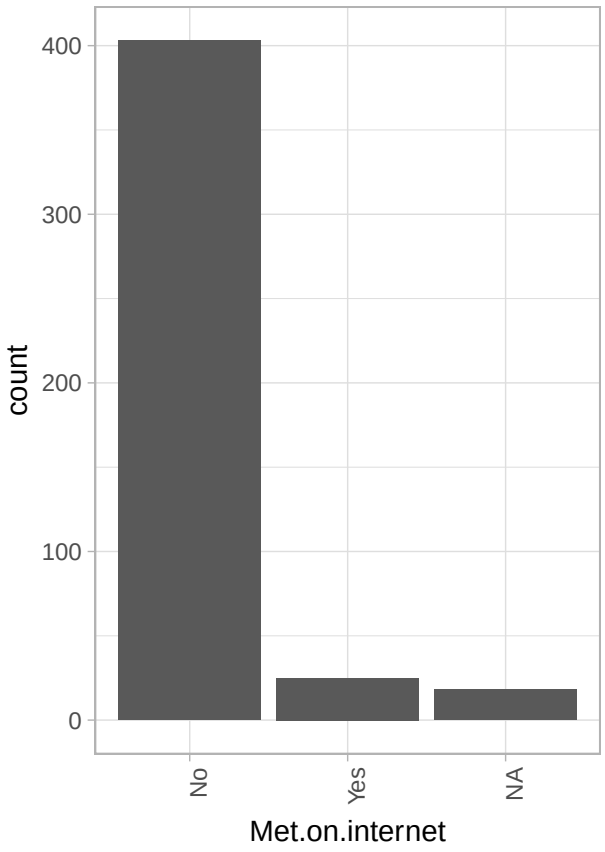

count
200
100
0
Christian
Muslim
None
Not stated
Other (please state)
Other
NA
Religion

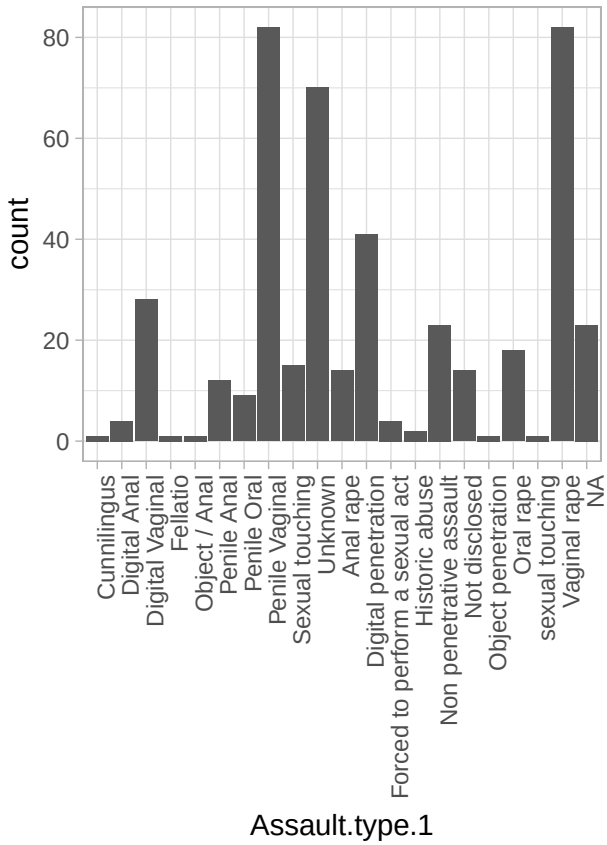

count
300
200
100
0
No
Yes
NA
DV.history

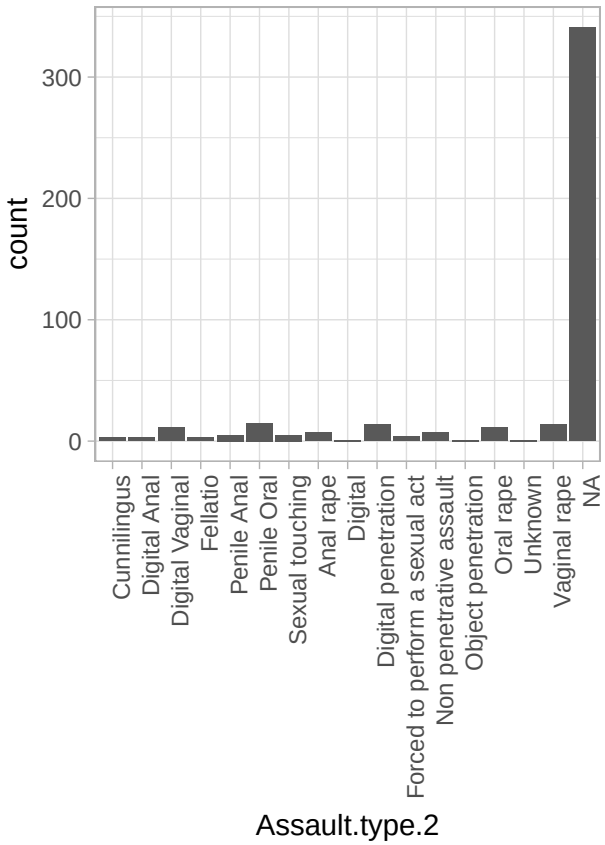

300
200
100
0
count
No
Yes
NA
CSE.CSE.risk

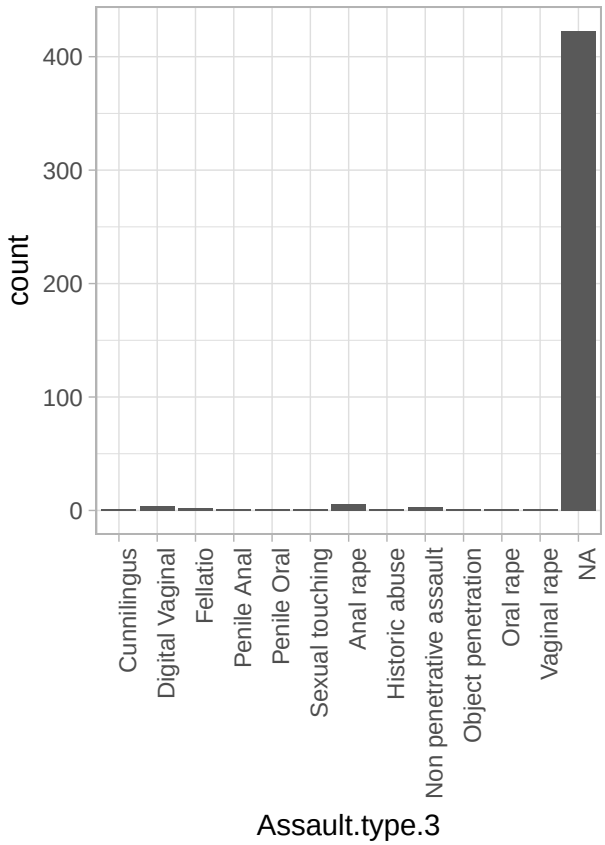

40
30
20
10
0
count
Acquaintance
Acquaintance < 24 hours
Acquaintance > 24 hours
Authority figure
Brother
Cousin
Ex partner
Family friend
Father
Friend
Half Brother
Lodger/flatmate
Mother in law/inlaw
Mother
Neighbour
Other (see comments)
Partner
School peer
Sister
Step brother
Step father
Stranger
Uncle
Unknown
Acquaintance >24 hours
Aunty
Brother−in−law
Carer
Grandfather
Male Cousin
Step Brother
Step Father
Work colleague
NA
Relationship.to.alleged.perp.

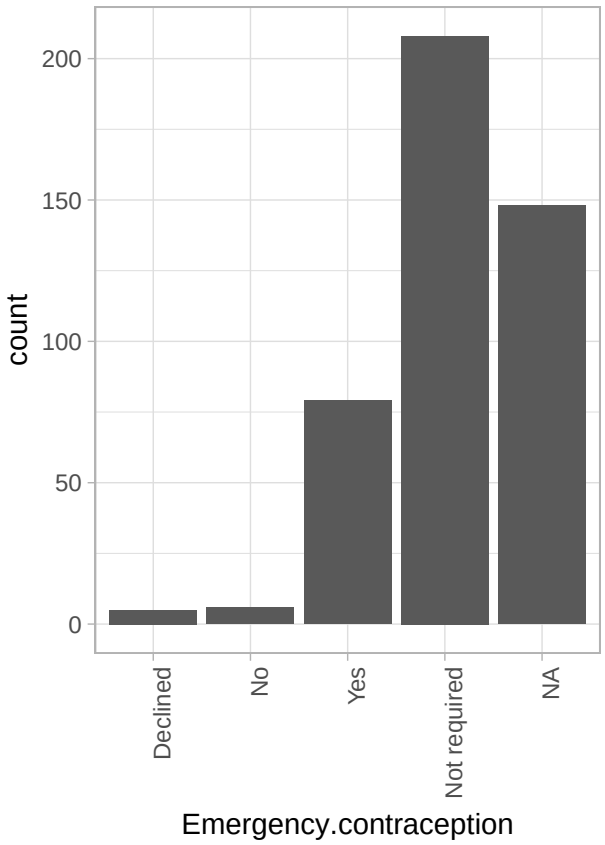

count
300
200
100
0
Multiple perps
Single perp
Unknown
NA
No..of.perps.

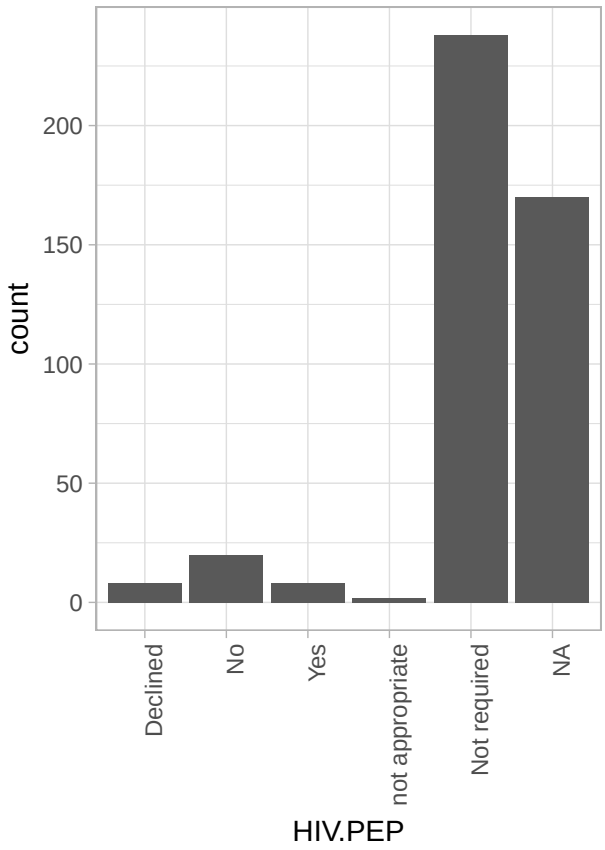

count
400
300
200
100
0
No
Yes
NA
Met.on.internet

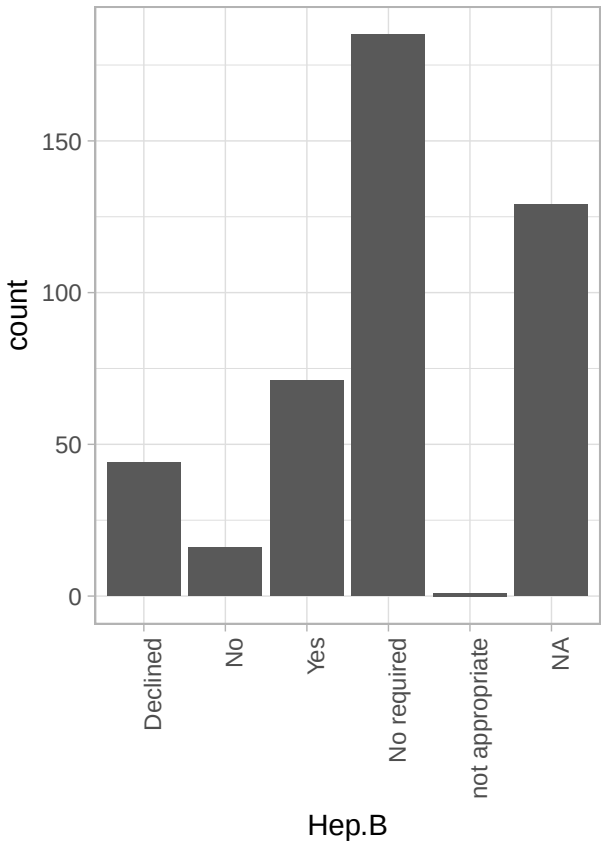

count
80
60
40
20
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Object / Anal
Penile Anal
Penile Oral
Penile Vaginal
Sexual touching
Unknown
Anal rape
Digital penetration
Forced to perform a sexual act
Historic abuse
Non penetrative assault
Not disclosed
Object penetration
Oral rape
sexual touching
Vaginal rape
NA
Assault.type.1

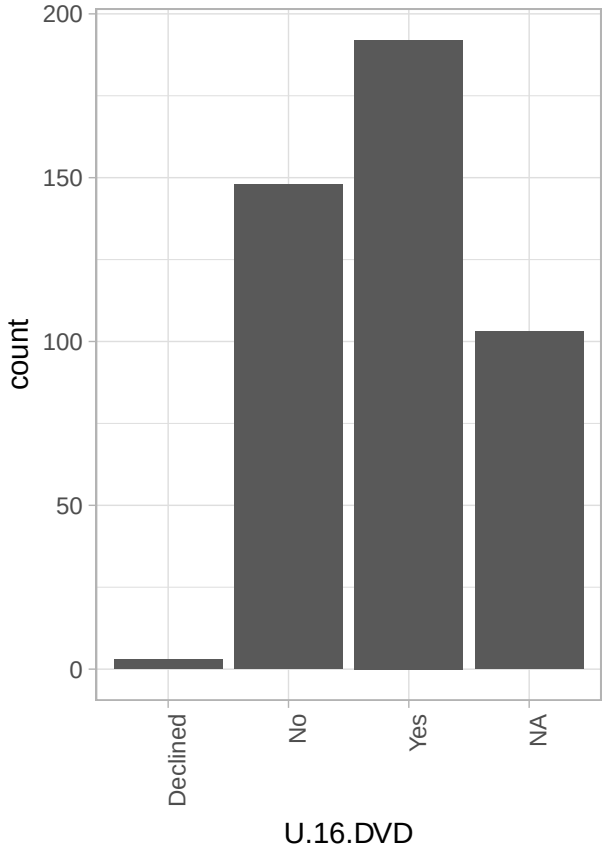

count
300
200
100
0
Cunnilingus
Digital Anal
Digital Vaginal
Fellatio
Penile Anal
Penile Oral
Sexual touching
Anal rape
Digital
Digital penetration
Forced to perform a sexual act
Non penetrative assault
Object penetration
Oral rape
Unknown
Vaginal rape
NA
Assault.type.2

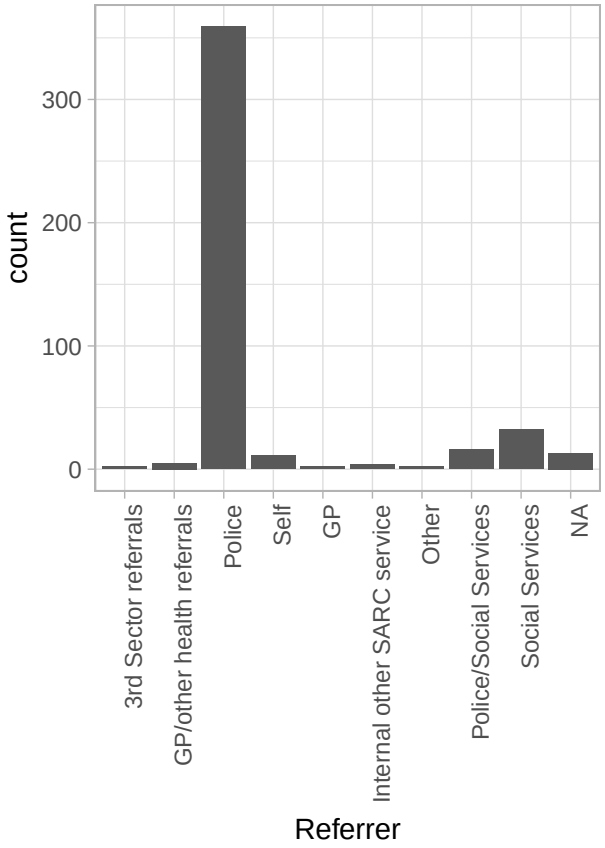

count
400
300
200
100
0
Cunnilingus
Digital Vaginal
Fellatio
Penile Anal
Penile Oral
Sexual touching
Anal rape
Historic abuse
Non penetrative assault
Object penetration
Oral rape
Vaginal rape
NA
Assault.type.3

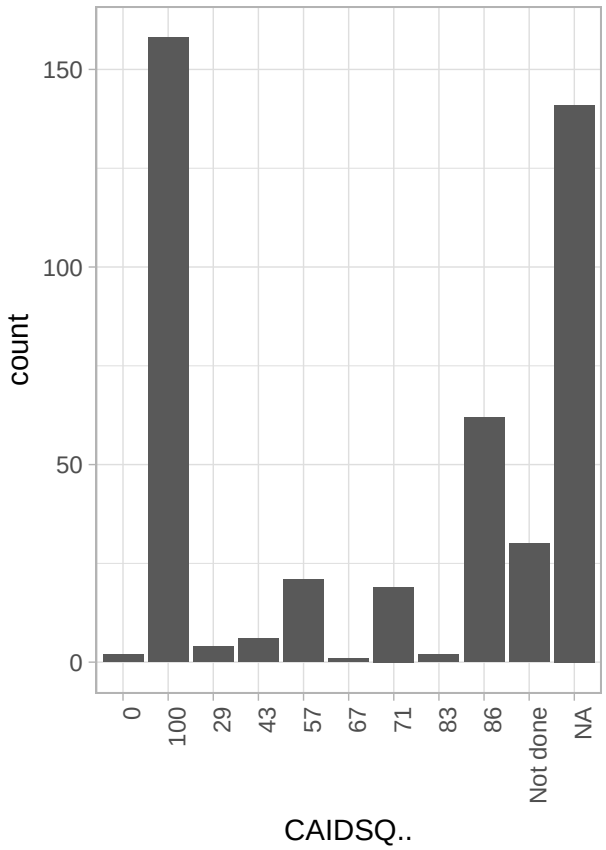

count
200
150
100
50
0
Declined
No
Yes
Not required
NA
Emergency.contraception


count
200
150
100
50
0
Declined
No
Yes
not appropriate
Not required
NA
HIV.PEP


count
150
100
50
0
Declined
No
Yes
No required
not appropriate
NA
Hep.B


200
150
100
50
0
count
Declined
No
Yes
NA
U.16.DVD

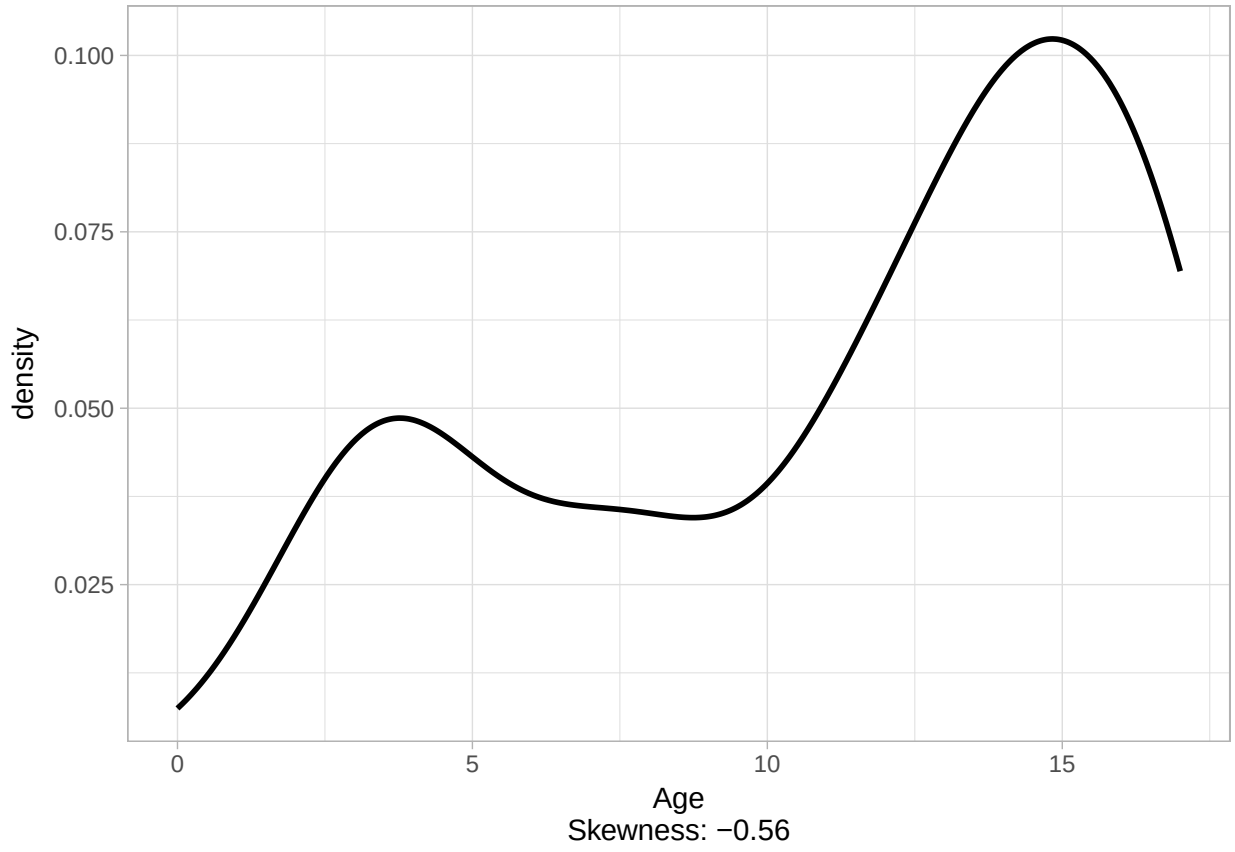

300
200
100
0
count
3rd Sector referrals
GP/other health referrals
Police
Self
GP
Internal other SARC service
Other
Police/Social Services
Social Services
NA
Referrer


150
100
50
0
count
0
100
29
43
57
67
71
83
86
Not done
NA
CAIDSQ..

3. Age distribution among clients

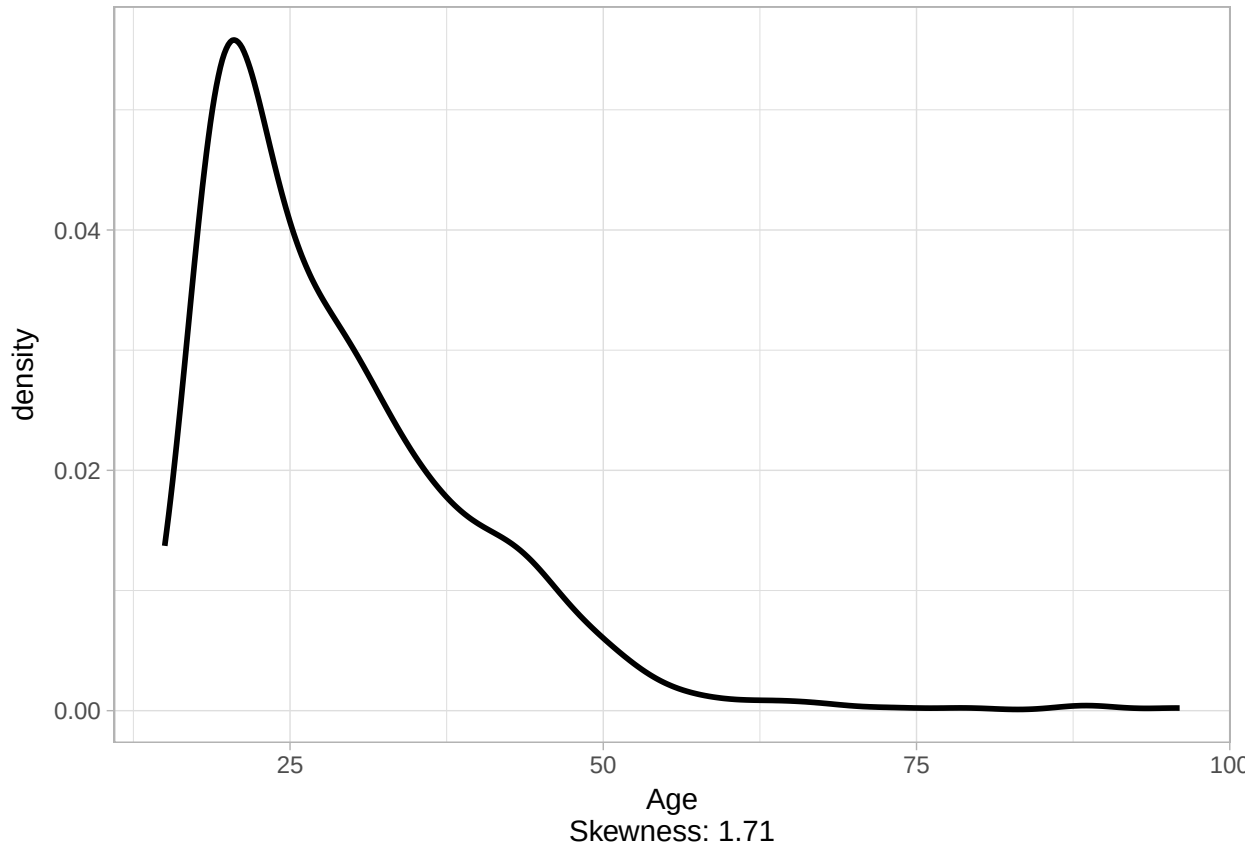
Children

```
## Warning: Removed 2 rows containing non-finite values (stat_density).
```



Adults

```
## Warning: Removed 5 rows containing non-finite values (stat_density).
```



Comments

- The histograms were made to make visualization easy. In some of them the labels in the X-axis overlap. To make them clearer, they can be generated again using more space.
- Some of the variables are quite meaningful themselves. Interpretaton is clear in most of them. This helps to look for relationships.
- I have not been able to solve the problem with computing the FME for the Jun-Nov set in a non-mechanical way. Because "FME start time" and "FME end time" are times, computing the difference gets difficult as there is no date associated to the second. This way the program does not know whether it should return the difference in mins from 8pm to 3am of the next day or of the same day. I am gonna try to find a solution which does not involve individually typing in the difference in mins. But I will if that is ther is no apparent solution to it. As we discussed, these variables can be used to get info about the followed protocol.
- The "Area of residence" column has so many distinct rows that makes it cumbersome to visualize, and has not been included. The good part is that it will be used for the maps of density
- I mentioned the variables I omitted when I merged both datasets, but it is important to keep in mind that "Learning disability" in Jun-Nov set, even though only collected in this period, is relevant to the study as It can validate the LDSQ/ CAIDS-Q score.

- Next step will be, firstly, elaborate a draft of the density maps and the temporal variances throughout the year. Secondly, it will involve looking at correlations between LDSQ/CAIDS-Q scored and other independent variables and presenting graphs with high correlation (and take a look at other correlations that can show up). Finally, before moving on, I need to review the code in order to check for any mistakes and to simplify it.
- The graphs shown come only from the unique adult and children data sets. I will need to take another look at those not-common variables that left aside for consistency.