Joel_IVRS_Script

Harun

June 6, 2022

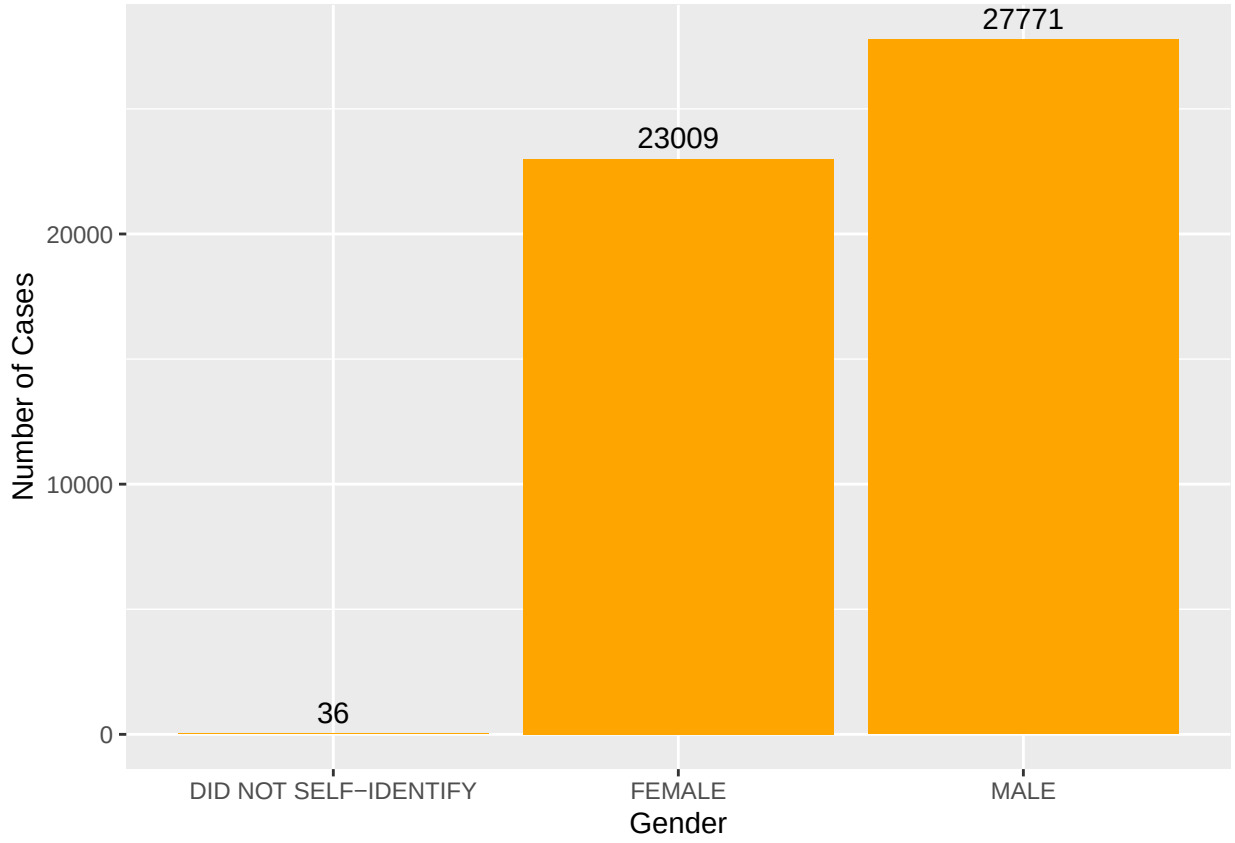
Exploration of the Iowa Vocational Rehabilitation Services (IVRS) Data

This data comes from the **data.iowa.gov** website and provides information on closed cases where the individual received services from IVRS. This data includes cases closed after October 1, 2008.

Number of Cases by Gender

This section looks at the number of cases for each gender.

From this graph we can see that there were more cases for men although it can be noted that the number of cases for women isn't to far off from the total of men's cases.

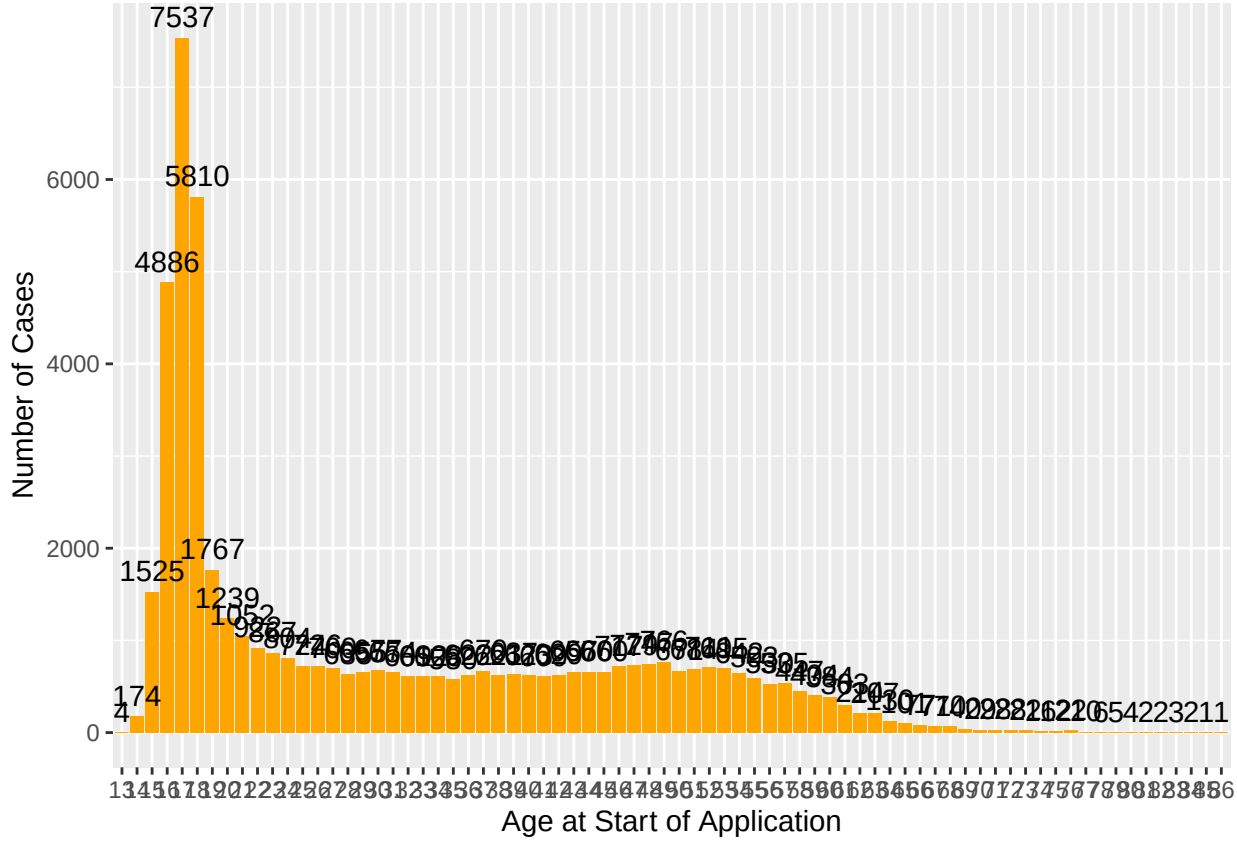


Number of Cases by Age

Age at Start of Application

This section looks at the number of cases by the age of the client when the case was started.

Initially I put the data into a bar graph but it is difficult to read the various ages and the bar labels are crowded.



I then looked at a table of the data sorted by the number of cases which was more helpful.

From this table we can see that the majority of cases were from those in their mid teen to mid twenties (as could be seen in the bar graph).

```
## # A tibble: 74 x 2
## # Groups:   Age.at.Application [74]
##    Age.at.Application Total.Cases
##                 <int>       <int>
##  1                 17        7537
##  2                 18        5810
##  3                 16        4886
##  4                 19        1767
##  5                 15        1525
##  6                 20        1239
##  7                 21        1052
##  8                 22         922
##  9                 23         867
## 10                 24         804
```

```
## # ... with 64 more rows
```

Age at Closure of Application

This section looks at the number of cases by the age of the client when the case was closed.

As before I put the data into a bar graph but again it is difficult to read the various ages and the bar labels are crowded. We do however see the same trend as before of the majority of those with closed cases are in their mid teen to mid twenties.

I then looked at a table of the data sorted by the number of cases.

From this table we can see that the majority of cases were actually from those in their late teen to late twenties.

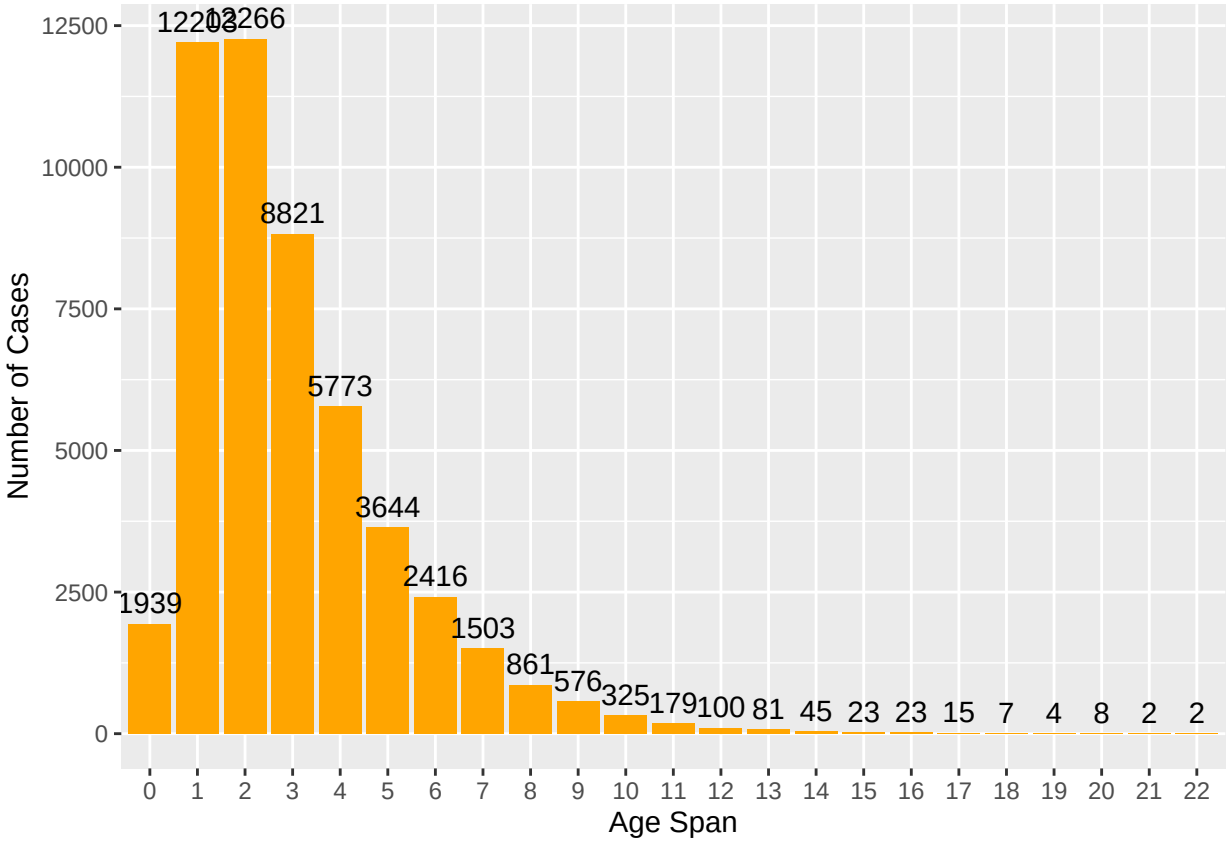
```
## # A tibble: 74 x 2
## # Groups:   Age.at.Closure [74]
##    Age.at.Closure Total.Cases
##             <int>       <int>
##  1             20        4910
##  2             19        4716
##  3             21        3891
##  4             22        2990
##  5             23        2427
##  6             18        1962
##  7             24        1907
##  8             25        1472
```

```
##  9             26        1140
## 10             27         930
## # ... with 64 more rows
```

Age Span of Clients

I then looked at the age span of the cases by subtracting the persons age at the end of their case from their age at the start of the case.

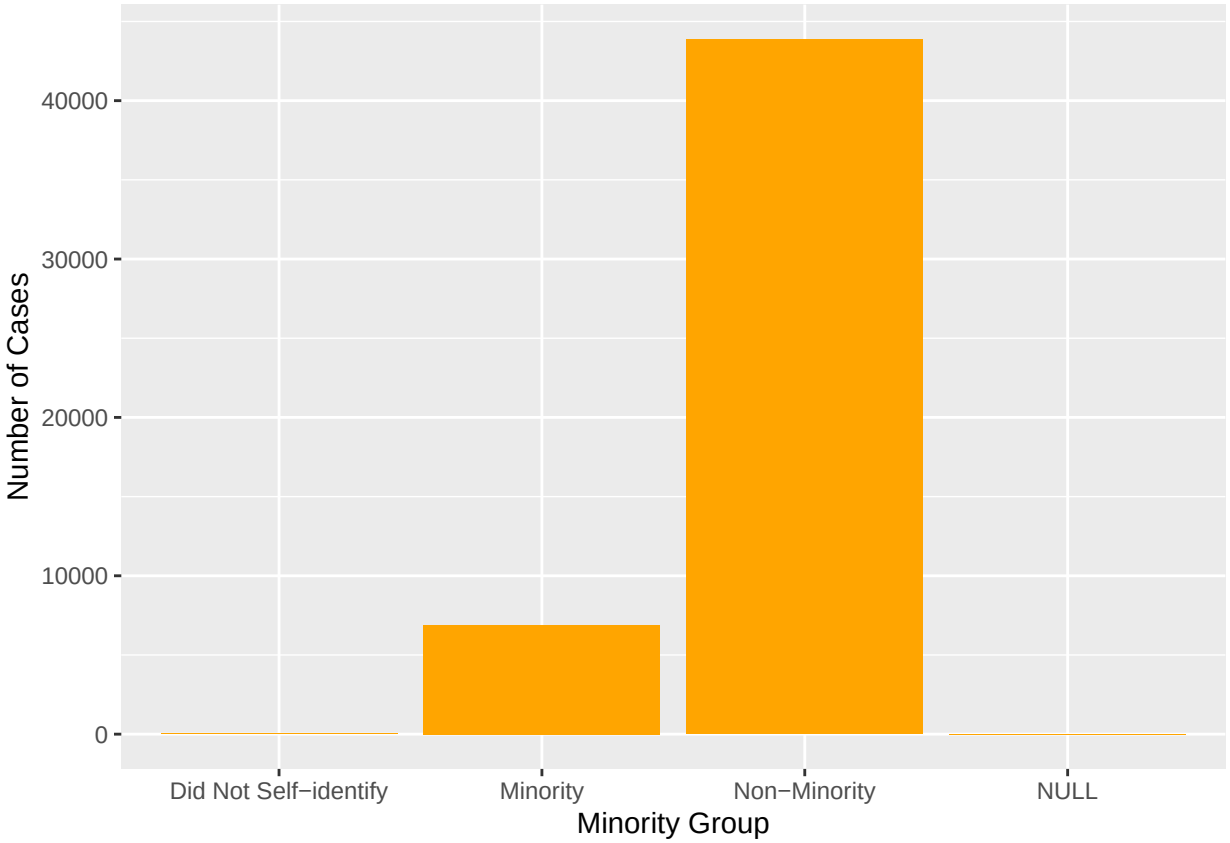
We can see that the majority of the cases were only open for 1 - 3 years. This makes sense since from our previous graphs it appeared that those in the majority age range for their case start age were in the majority for their case closed age.



Minority Groups

This section looks at the minority groups.

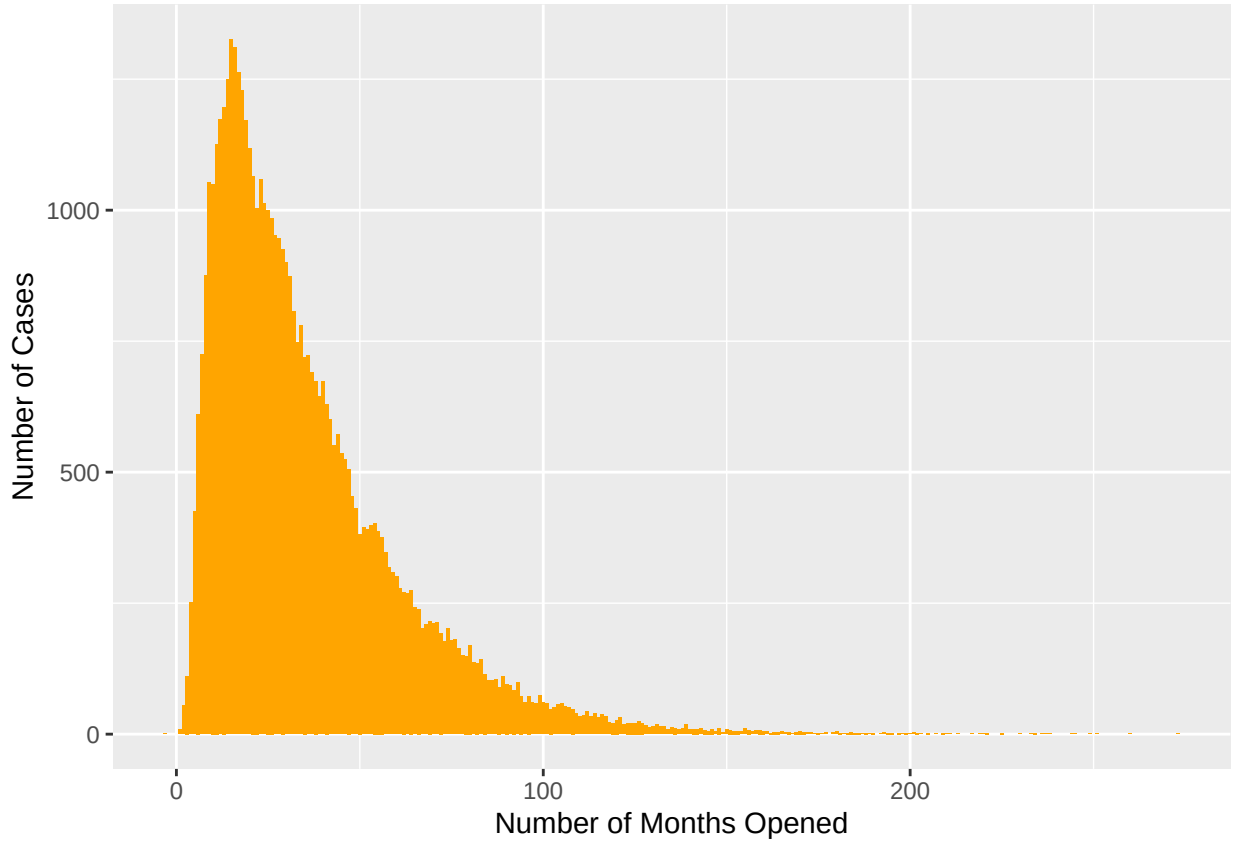
We can see that most of the cases were non-minority clients.



Monthly Cases

This section looks at how long the closed cases were opened.

I created a bar chart to view this data, however there are a large number of data points and it is difficult to see specifics. We can see that the majority of cases were opened between about 10 - 50 months (about 1 - 5 years).



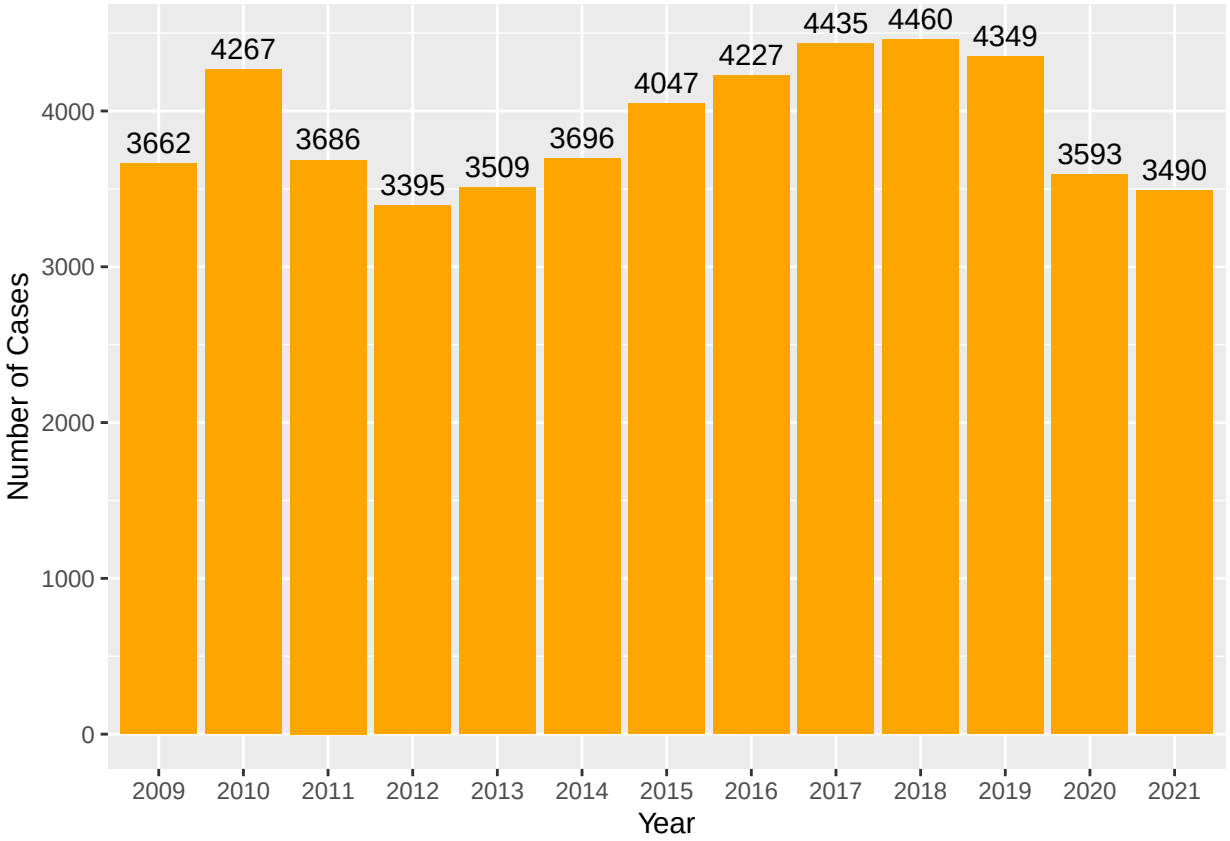
Creating a sorted table shows that most of the cases were opened for about 1 $\frac{1}{2}$ - 2 years.

```
## # A tibble: 224 x 2
## # Groups:   Months.Case.Open [224]
##    Months.Case.Open Total.Cases
##               <int>       <int>
##  1               15        1326
##  2               16        1310
##  3               17        1263
##  4               14        1249
##  5               18        1228
##  6               13        1197
##  7               12        1173
##  8               19        1171
##  9               11        1126
## 10               20        1118
## # ... with 214 more rows
```

Yearly Cases

This section looks at how many cases were closed each year.

We can see that the number of closed cases has remained relatively the same over the course of the data sets history.

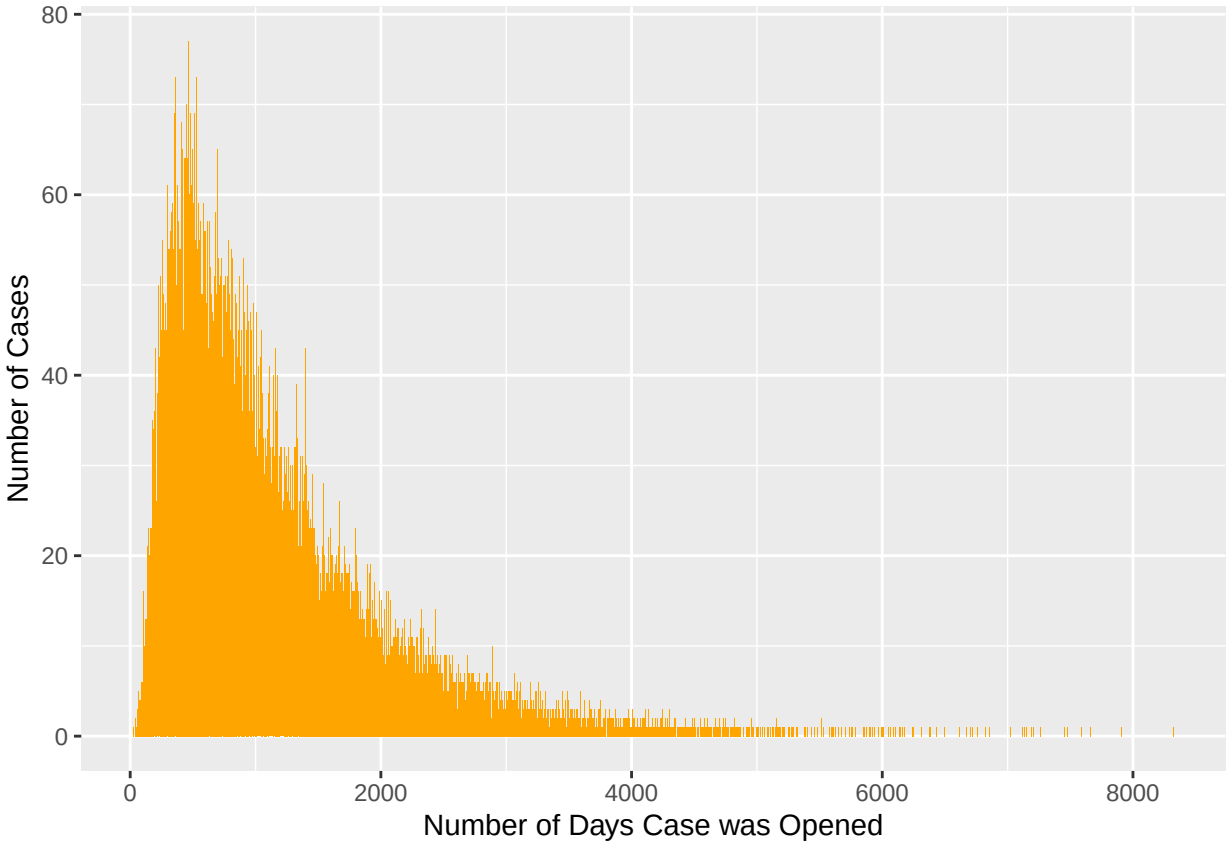


Date Span

This section looks at the span of the cases by start date of application to date of closure.

A bar graph is somewhat helpful however it is difficult to see specifics. It appears that the majority of the cases are opened between 0 - 1000 days (about 0 - 32 months or about 0 - 3 years)

```
## Don't know how to automatically pick scale for object of type difftime. Defaulting to continuous.
```



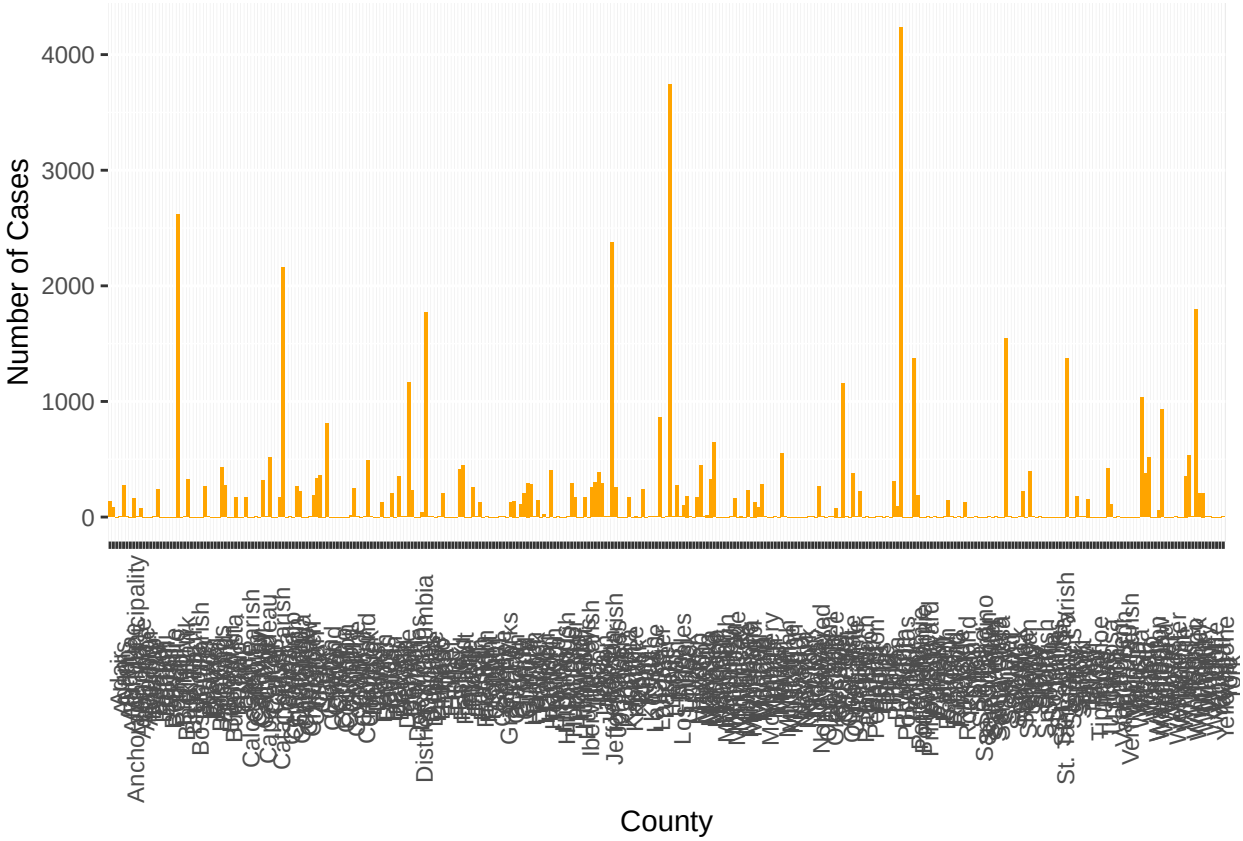
Looking at a table we can see the highest number of total cases is 462 days which is about 1 $\frac{1}{4}$ years which is close to the range we found from looking at the months a case was open.

```
## # A tibble: 3,927 x 2
## # Groups:   date.span [3,927]
##    date.span Total.Cases
##    <drtn>          <int>
##  1 462 days           77
##  2 364 days           73
##  3 525 days           73
##  4 448 days           70
##  5 350 days           69
##  6 483 days           69
##  7 511 days           69
##  8 406 days           68
##  9 532 days           67
## 10 420 days           65
## # ... with 3,917 more rows
```

Cases by Client County

This section looks at the clients for each case by county they are from.

Initially I attempted to make a bar chart hoping it would hold all of Iowa's 99 counties. However I learned that there are many counties in the data set outside of Iowa which made it almost impossible to view the data in a bar graph format.



I then opted to make a table showing the number of closed cases for each county sorted by those with the highest number of cases.

From this table we can see that Polk and Linn counties have the most closed cases. A quick **Google search** showed from the 2020 American Community Survey (ACS) data that the counties with the highest population were Polk and Linn so this data aligns with what we would expect.

```
## # A tibble: 329 x 2
## # Groups:   Client.County [329]
##    Client.County Total.Cases
##    <chr>               <int>
##  1 Polk                 4238
##  2 Linn                 3744
##  3 Black Hawk           2622
##  4 Johnson              2374
##  5 Cerro Gordo          2158
##  6 Woodbury             1794
##  7 Dubuque              1772
##  8 Scott                1542
##  9 Story                1374
## 10 Pottawattamie        1369
## # ... with 319 more rows
```

Cases by Office Area

This section looks at the number of closed cases by the office area.

We can see that the Des Moines (026) office has the most number of closed cases.