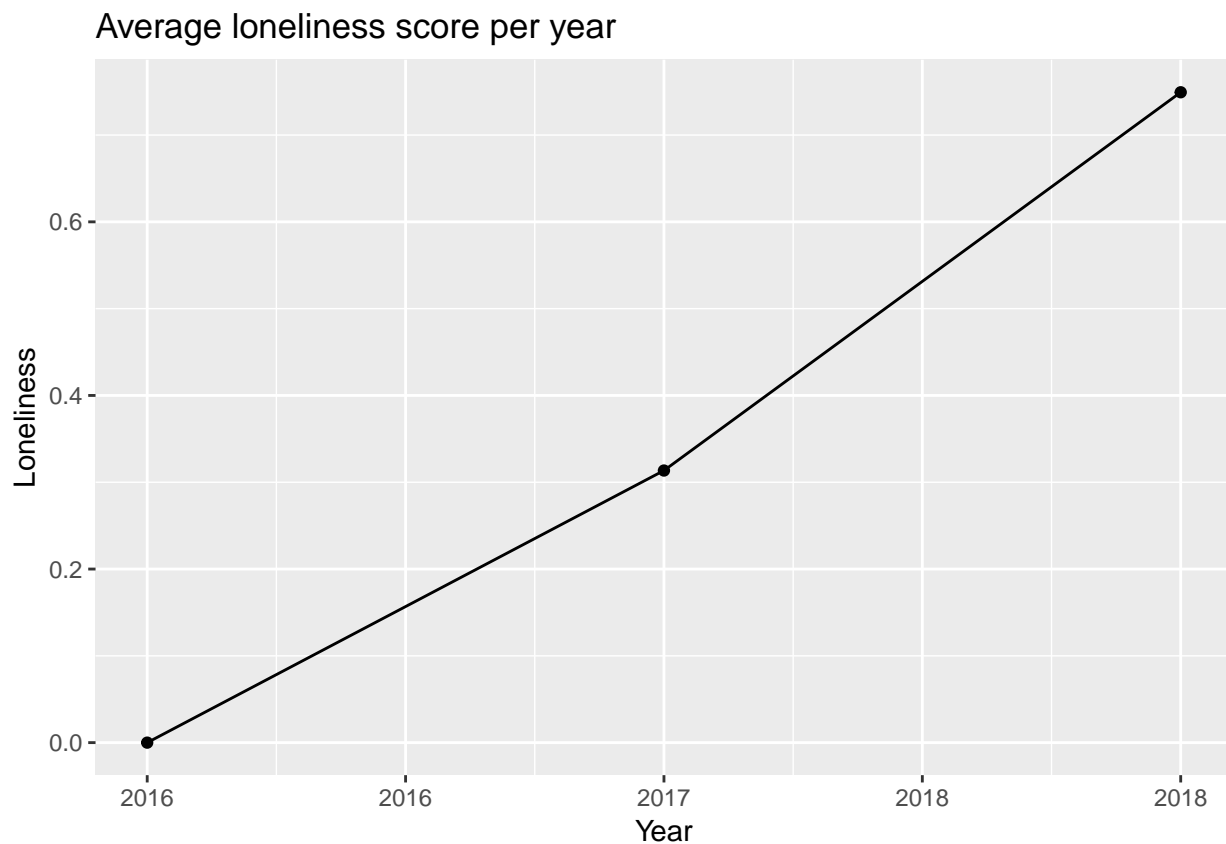


Descriptive analysis based on the `final_data` dataset

Loneliness appears to be increasing over the three years

```
data_lonely <- group_by(data, Year)
data_lonely <- summarise(data_lonely, Loneliness = mean(loneills))

ggplot(data_lonely, aes(y=Loneliness, x=Year))+
  geom_point() +
  geom_line() +
  labs(title = "Average loneliness score per year")
```



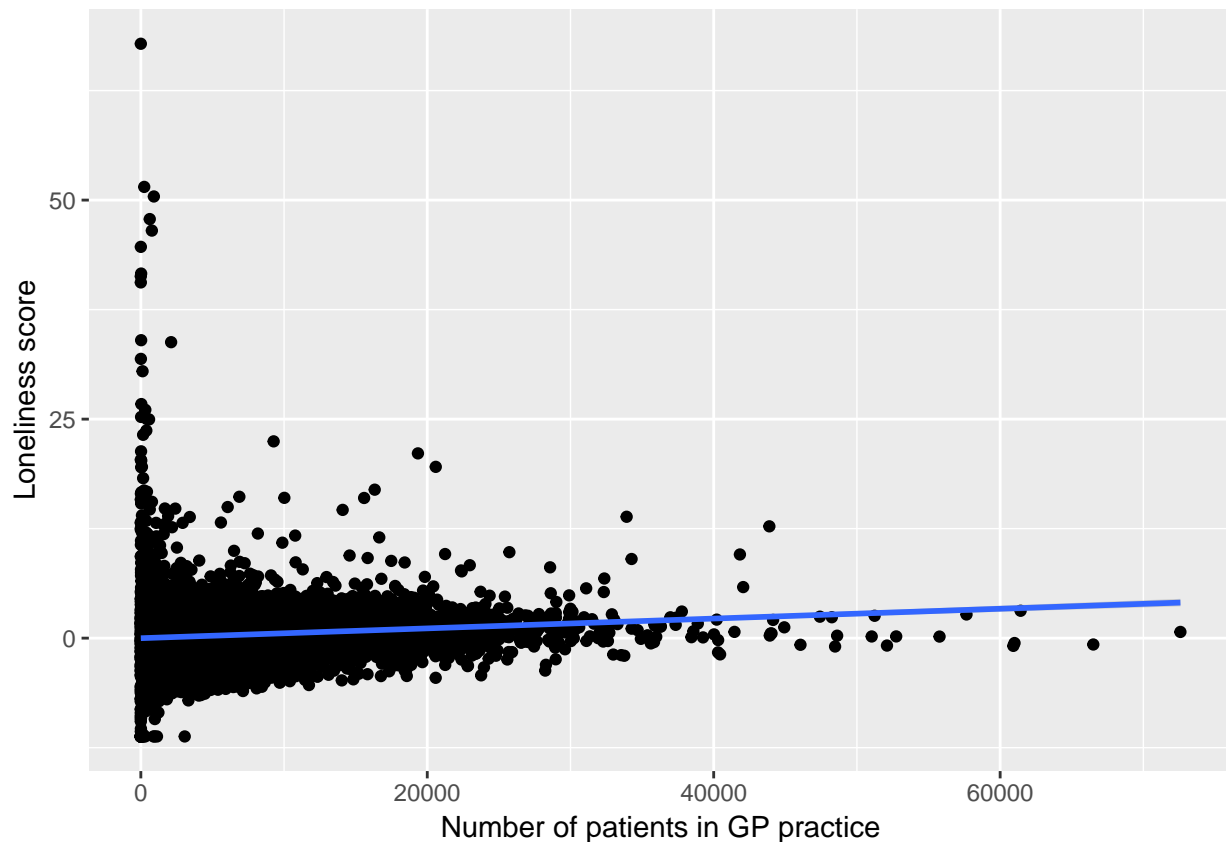
```
kable(data_lonely)
```

Year	Loneliness
2016	0.0000
2017	0.3135
2018	0.7493

Loneliness seems to increase (slightly) based with patient numbers

```
ggplot(data, aes(x=data$NUMBER_OF_PATIENTS, y=data$loneills))+
  geom_point() +
  geom_smooth(method='lm') +
  labs(y="Loneliness score",
```

```
x="Number of patients in GP practice")
```



Could be worth exploring further, as potentially population density (for which no of patients is an indicator) might be a factor? Population estimates for the MSOA's can be found [here](#), and will hopefully be easy enough to begin linking in with the data we currently have.

Loneliness score by urban/rural category (no. of categories = 10)

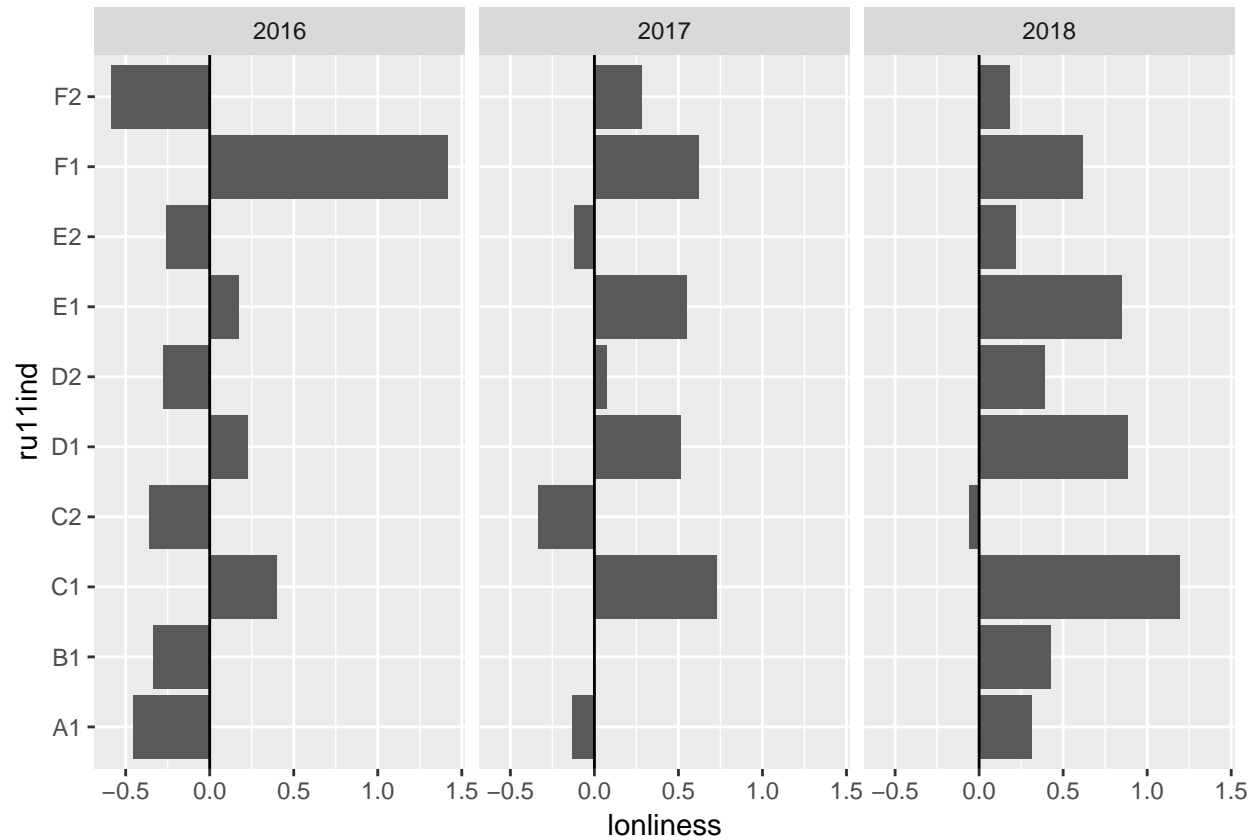
I used the following definitions of urban vs rural to create both of the plots below, sourced from [here](#):

- Urban: Major Conurbation (A1)
- Urban: Minor Conurbation (B1)
- Urban: City and Town (C1)
- Urban: City and Town in a Sparse Setting (C2)
- Rural: Town and Fringe (D1)
- Rural: Town and Fringe in a Sparse Setting (D2)
- Rural: Village (E1)
- Rural: Village in a Sparse Setting (E2)
- Rural: Hamlets and Isolated Dwellings (F1)
- Rural: Hamlets and Isolated Dwellings in a Sparse Setting (F2)

```
data_ru1lind <- group_by(data, ru1lind, Year)
data_ru1lind <- summarise(data_ru1lind, loneliness = mean(loneills))

ggplot(data_ru1lind, aes(x = ru1lind, y=loneliness)) +
  facet_grid(~Year)+
  coord_flip() +
```

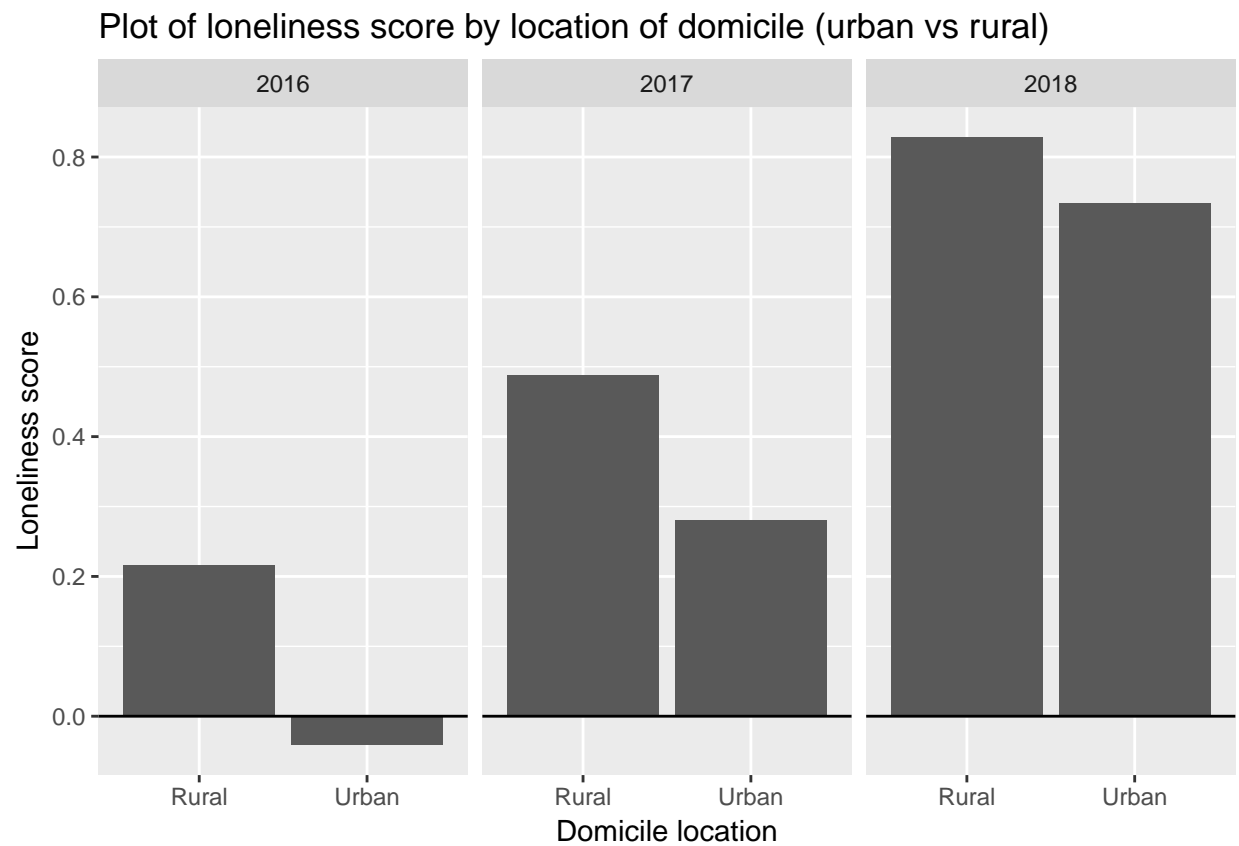
```
geom_bar(stat="identity") +  
geom_hline(yintercept = 0)
```



Loneliness score with urban/rural as binary variable

```
data <- data %>%  
  mutate(  
    ru11ind_group = ifelse(  
      ru11ind == "A1" |  
      ru11ind == "B1" |  
      ru11ind == "C1" | ru11ind == "C2",  
      "Urban",  
      "Rural"  
    )  
  )  
  
data_ru11ind_grp <- group_by(data, ru11ind_group, Year)  
data_ru11ind_grp <- summarise(data_ru11ind_grp, loneliness = mean(loneliness))  
  
ggplot(data_ru11ind_grp, aes(x = ru11ind_group, y=loneliness)) +  
  facet_grid(~Year) +  
  geom_bar(stat="identity") +  
  geom_hline(yintercept = 0) +  
  labs(y="Loneliness score",  
       x="Domicile location",
```

```
title = "Plot of loneliness score by location of domicile (urban vs rural)"
```



It does seem that location of domicile (urban vs rural) has an effect on average loneliness.