project_a

maria

9/20/2017

Analysing crime in London

```
##  Crime rate - 2013
##  Min.   : 13.93
##  1st Qu.: 62.23
##  Median : 78.35
##  Mean   : 82.40
##  3rd Qu.: 97.72
##  Max.   :222.92
```

Introduction

Modern policy regarding crime rate has transitioned to a focus on preventative measures as opposed to consequential punishment. As outlined by Crawford and Evans (2017) these policies focus on early intervention and improving well-being to secure community safety. A comprehensive understanding of demographic variables that are highly correlated and plausibly causing crime rates is necessary for developing robust policies. Crime rate in London varies drastically by area. In 2013, crime rates ranged from 13.93 to 222.92, with a mean of 82.40 and a standard deviation of 30. With such drastically different levels of crime and significant outliers, there must be an underlying explanation.

The dataset used in this report is from London datastore, an open data-sharing platform with statistics provided from the London mayoral office. The data is from 2013 and its original purpose was to assign a well-being score to different wards in London through measurements of 12 different indicators such as childhood obesity and deliberate fires. However, this report will focus on crime rates in London and explore the relationships between these rates and other demographic features of wards. It will attempt to model crime based on these additional indicators of well-being. Any significant results may prove to be insightful for policy makers considering local crime rates.

Description

Information on the 12 key indicators of well-being are given for 625 wards in London over a four-year time frame from 2009 to 2013. Additional borough-wide information is provided for the 32 London boroughs. The variables are split into 8 different categories, detailed below, aiming to give a holistic view of the standard of well-being in the different areas of the capital.

Variables

Description

Information on the 12 key indicators of well-being are given for 625 wards in London over a four-year time frame from 2009 to 2013. Additional borough-wide information is provided for the 32 London boroughs. The variables are split into 8 different categories, detailed below, aiming to give a holistic view of the standard of well-being in the different areas of the capital.

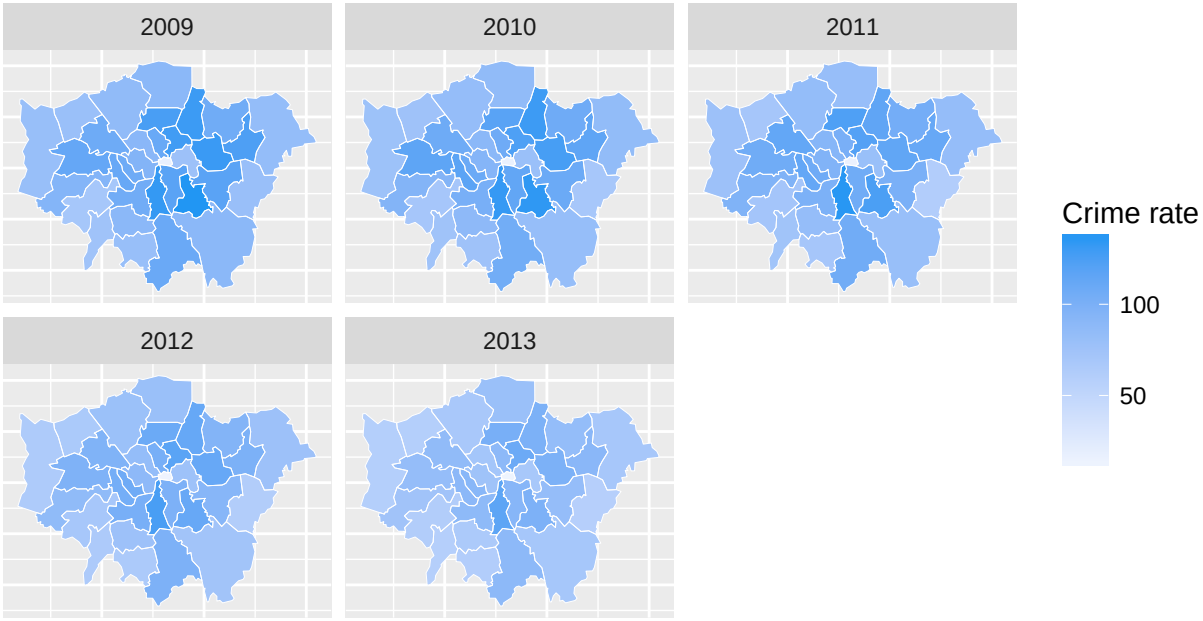


Figure 1: Crime rate by borough from 2009 to 2013

Variables

Category	Variable	X___1
Health	Life Expectancy	Reversed index of life-expectancies for a rolling 5-year perio
	Childhood Obesity	Prevalence of obesity by area of child residence
	Incapacity Benefits claimant rate	Claimant rate of Incapacity Benefit or Severe Disablement A
Economic Security	Unemployment rate	Percent of working-age population claiming Jobseeker's Allo
Safety	Crime rate	Notifiable offences per 1,000 daytime population
	Deliberate fires	Deliberate Fires per 1000 population
Education:	GCSE point scores	Average GCSE point scores
Children:	Unauthorised Pupil Absence	Absence in all maintained schools without permission from a
Families:	Children in out-of-work households	% of children dependent on a parent of guardian claiming ou
Access	Public Transport Accessibility Scores	Measure of the accessibility of a point to the public transpor
Environment:	Access to public open space and nature	Measure of proportion of area that is greenspace and accessi
Happiness:	Subjective well-being average score	Composite Score of Happiness, Life Satisfaction, Worthwhile

Strength and limitations

Overview analysis

Crime by borough

From 2009 to 2013, the crime rates of all boroughs in London have decreased. As seen in the maps of figure x, the boroughs in the outskirts of London seem to have a lower crime rate than the boroughs in the center of London with exception to the City of London.

City of London, remains with the lowest rate of crime across all boroughs with only 14 offences per 1000 daytime inhabitants in 2013. In contrast, its neighbouring boroughs as Lambeth, Hackney, Southwark and Islington continue to present a crime rate of 100. The map below shows the quartile with the highest number

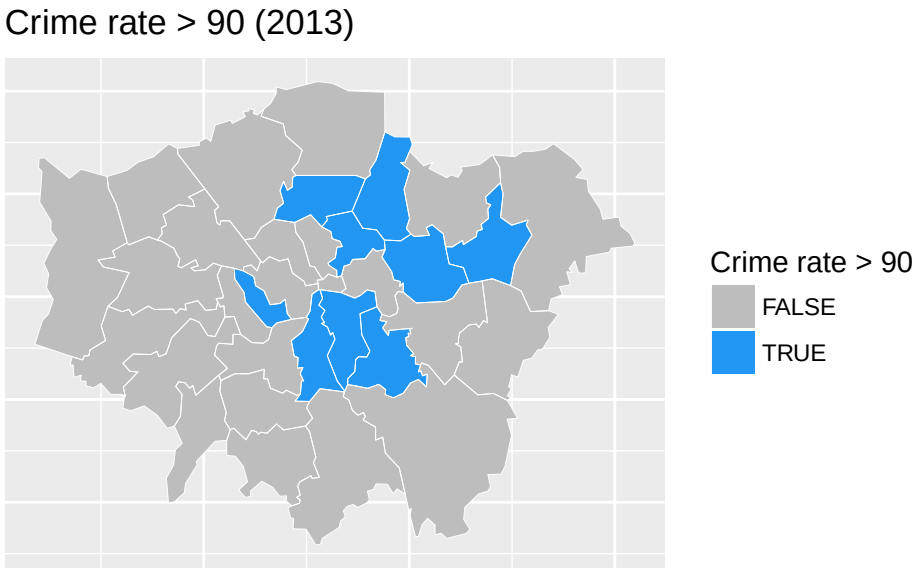


Figure 2: Crime rate > 90 (2013)

of crime rates in London (crime rate over 90). These boroughs are mainly located to the upper and lower side from the City of London.

On average, boroughs have lowered their crime rates to 19% from 2009 to 2013. The boroughs that seem to be showing the highest decrease in crime rates are on the further southeast of London, as for instance in Greenwich, Barking and Dagenham and Lewisham. The boroughs with the least progress in decreasing crime are Tower Hamlets, Enfield and Kensington and Chelsea.

Crime historically

Theory

Correlations

Correlations with crime rate

Theory

Correlations

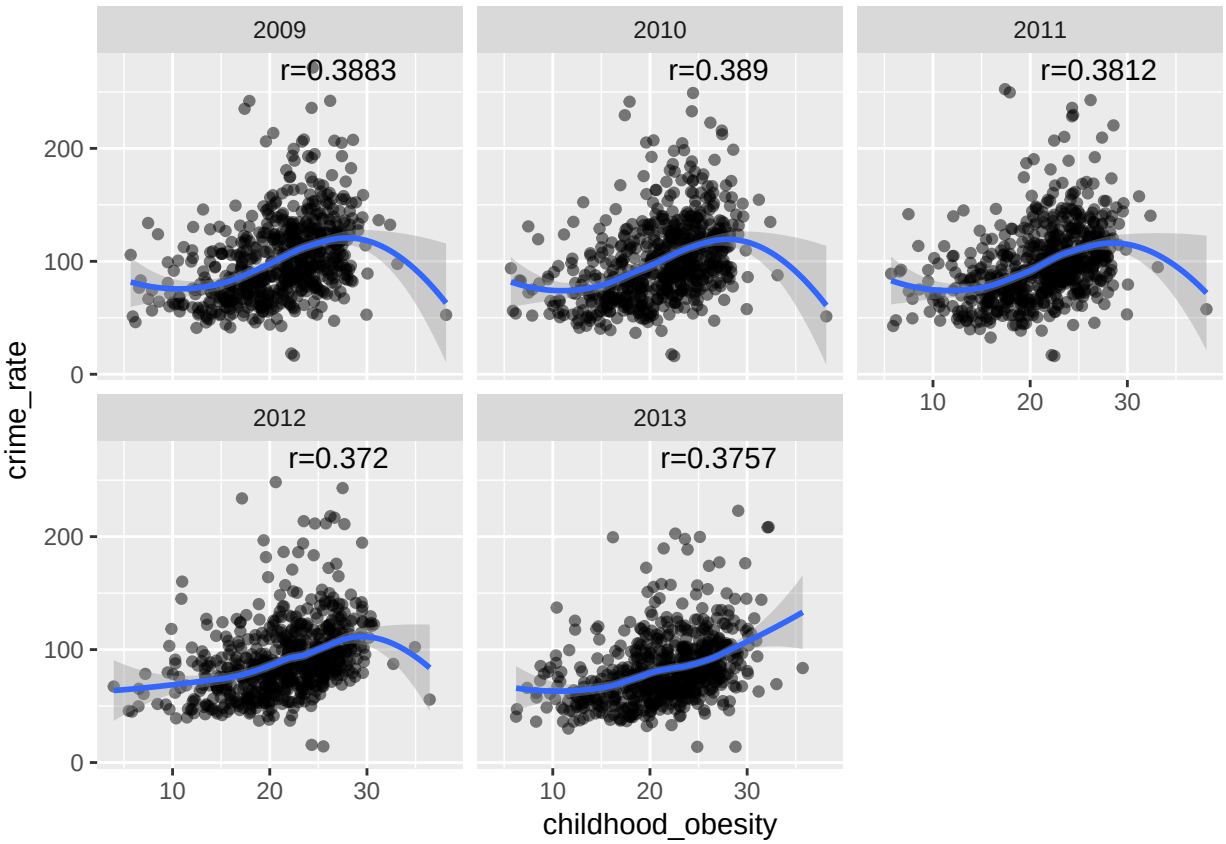
```
#the following codes can produce long table and plots with correlations on them
# correlation part not finished
var_list = c("childhood_obesity","incapacity_benefit","unemployment_rate","crime_rate","deliberate_fires

crime_ward_long = data.frame()
for (i in 1:length(var_list)) {
  temp <- select(crime_ward, Ward, Borough, (i*5+5):(i*5+9))
  colnames(temp)[3:7] <- c(2009:2013)
  tidytemp <- gather_(temp, "year", var_list[i], colnames(temp)[3:7])
  if (i==1) {
```

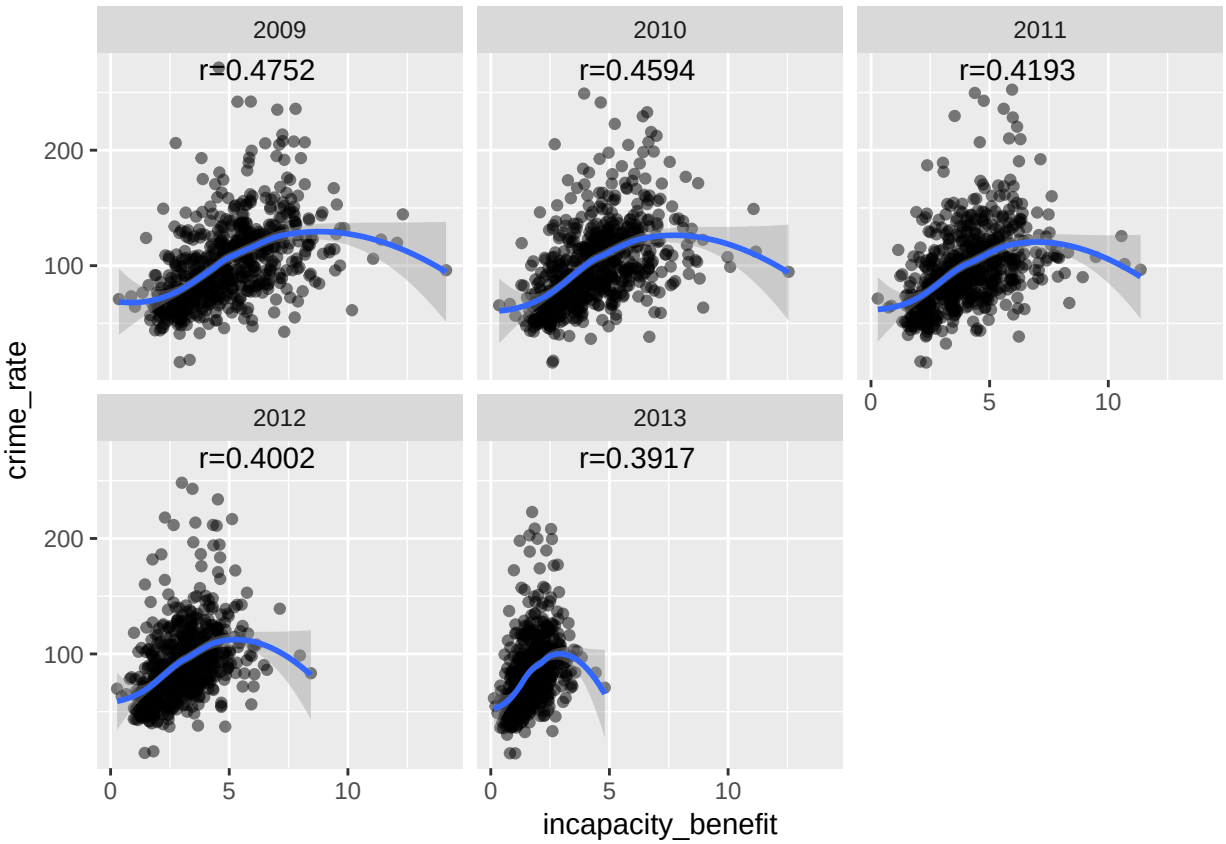
Figure 3: Crime rate growth from 2009 to 2013

```
    crime_ward_long <- tidytemp
  } else {
    crime_ward_long <- cbind(crime_ward_long, tidytemp[,4])
  }
}
gb <- crime_ward_long %>% dplyr::group_by(year)
for (item in var_list) {
  if (item!="crime_rate") {
    corr <- dplyr::summarise(gb,r=round(cor(!!sym(item),crime_rate),4))
    print(ggplot(crime_ward_long,aes_string(x=item,y="crime_rate"))+geom_point(alpha=0.5)+
      facet_wrap(~ year, nrow=2) +
      geom_smooth() +
      geom_text(data=corr, aes(label=paste("r=",r,sep="")),x=quantile(gb[[item]],0.9),y=270) )
  }
}
```

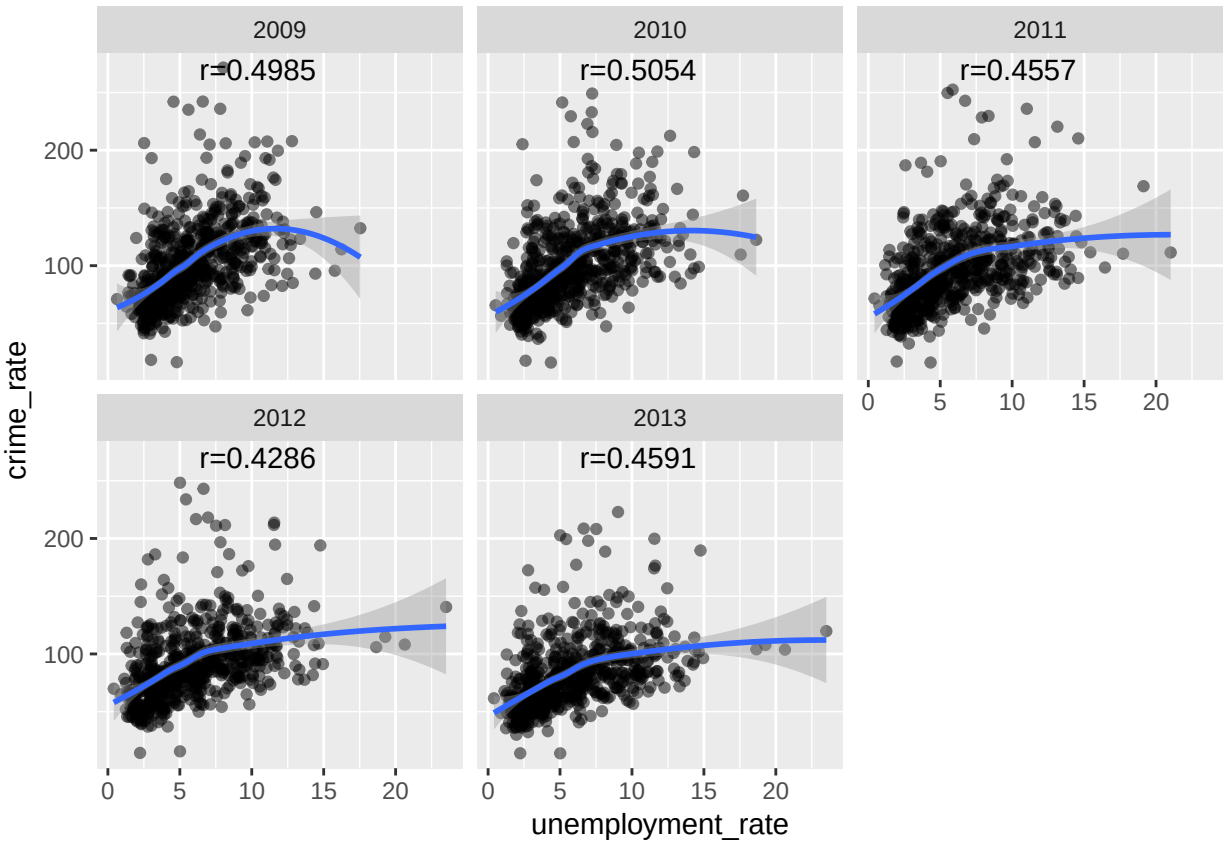
```
## `geom_smooth()` using method = 'loess'
```



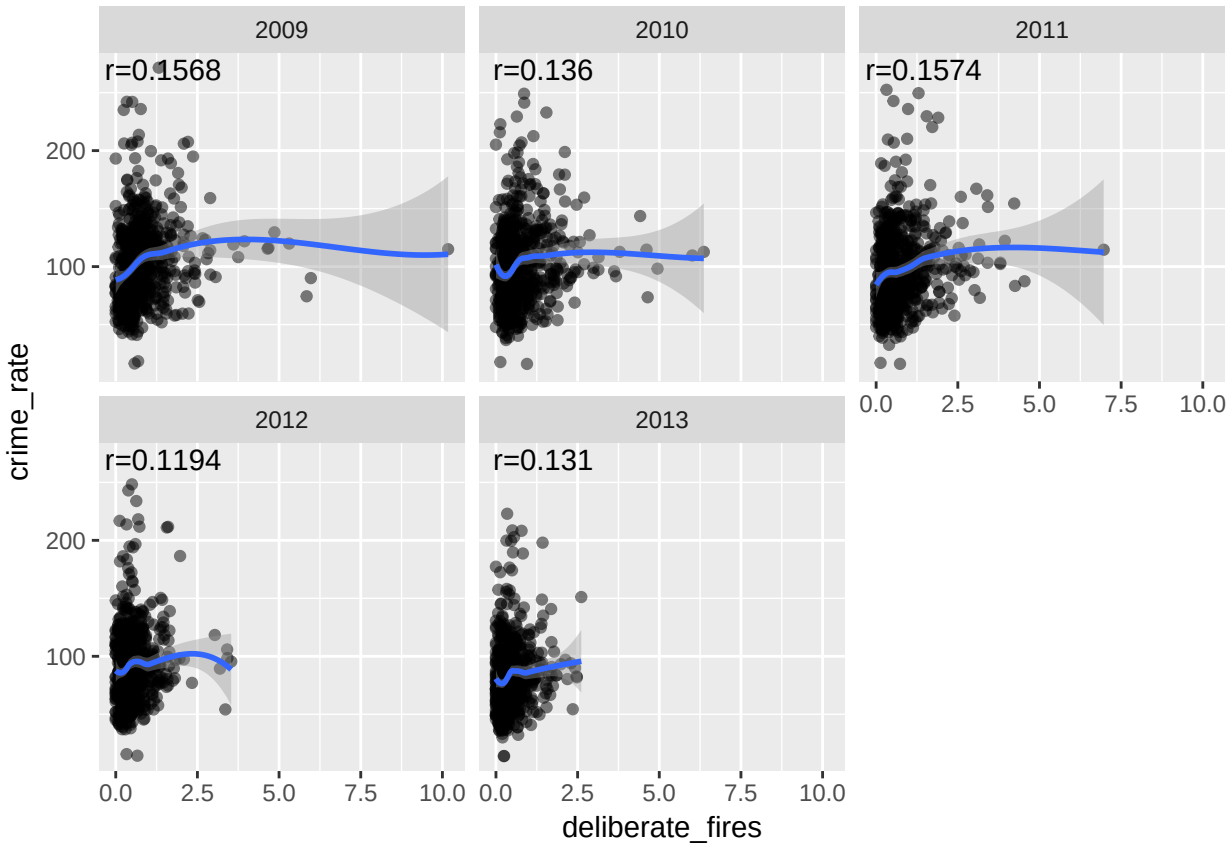

```
## `geom_smooth()` using method = 'loess'
```



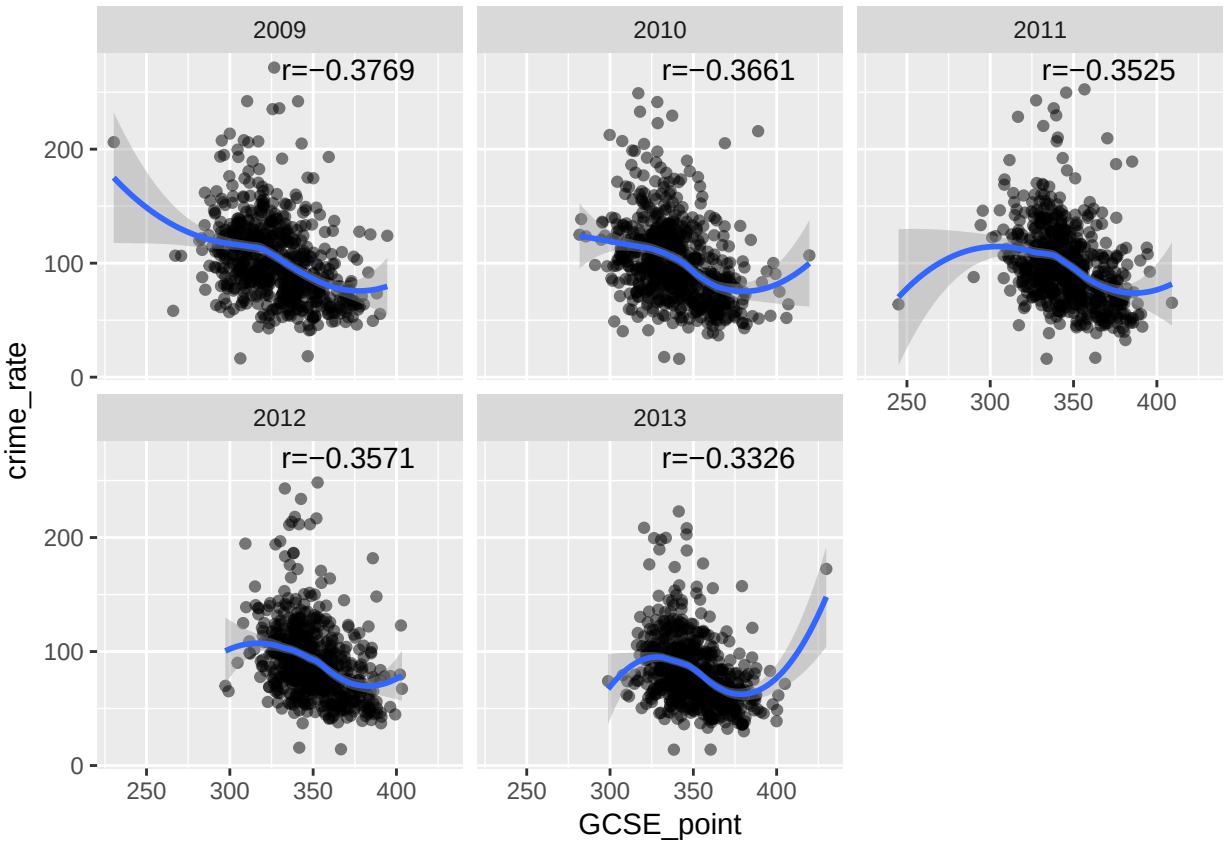

```
## `geom_smooth()` using method = 'loess'
```



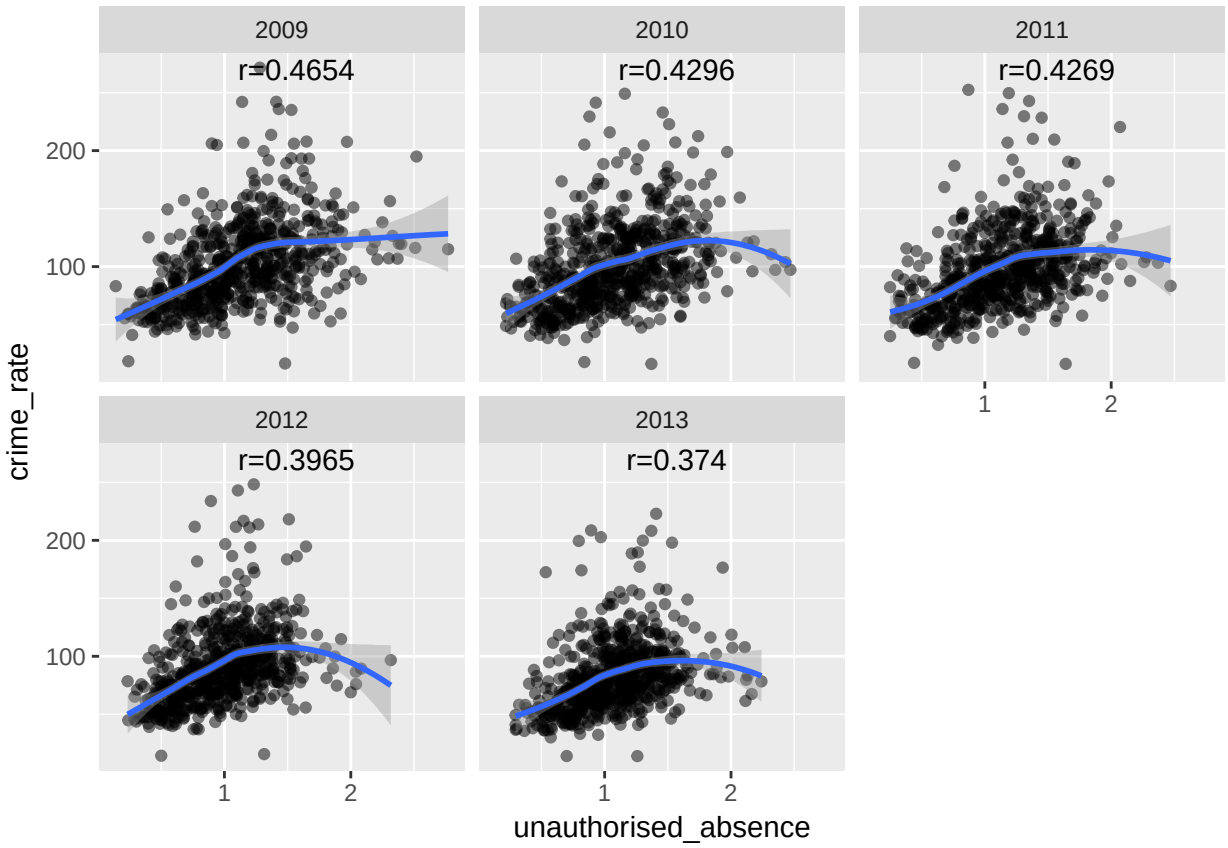

```
## `geom_smooth()` using method = 'loess'
```



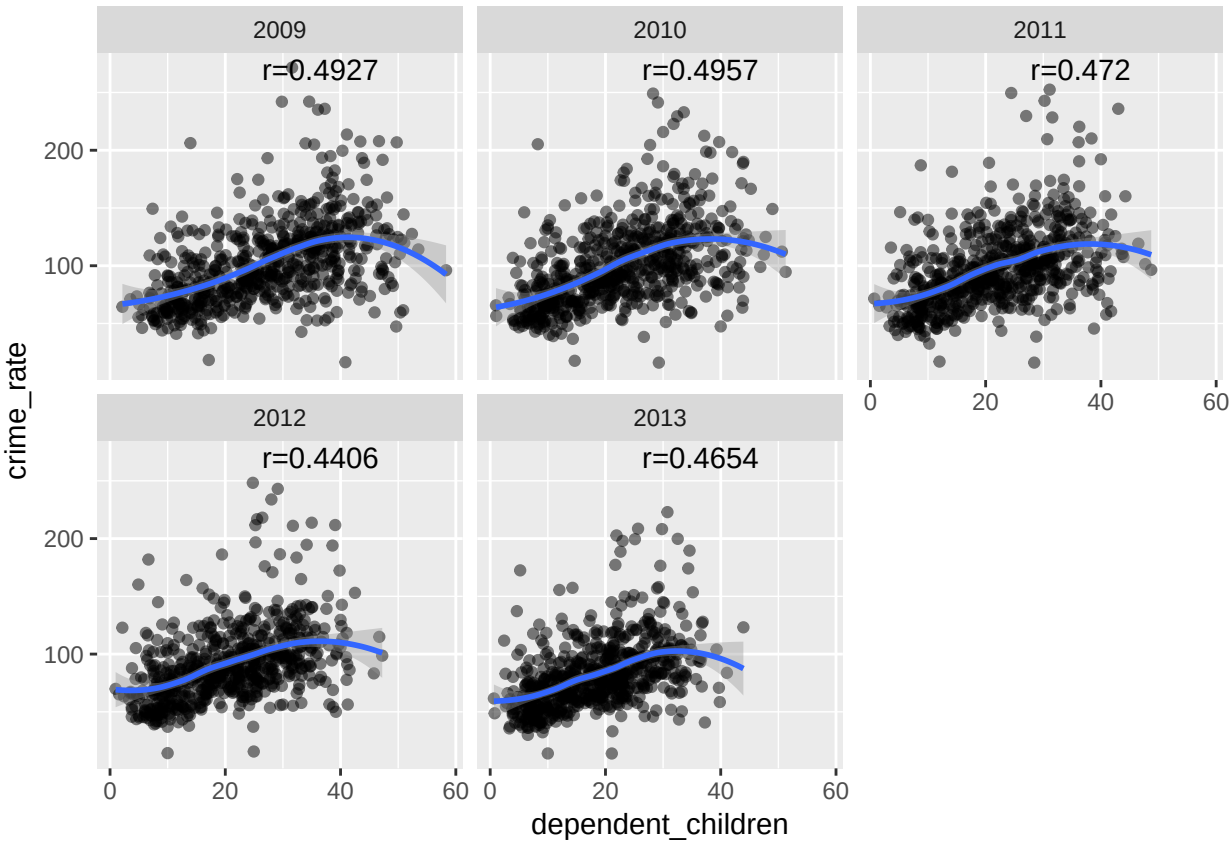

```
## `geom_smooth()` using method = 'loess'
```



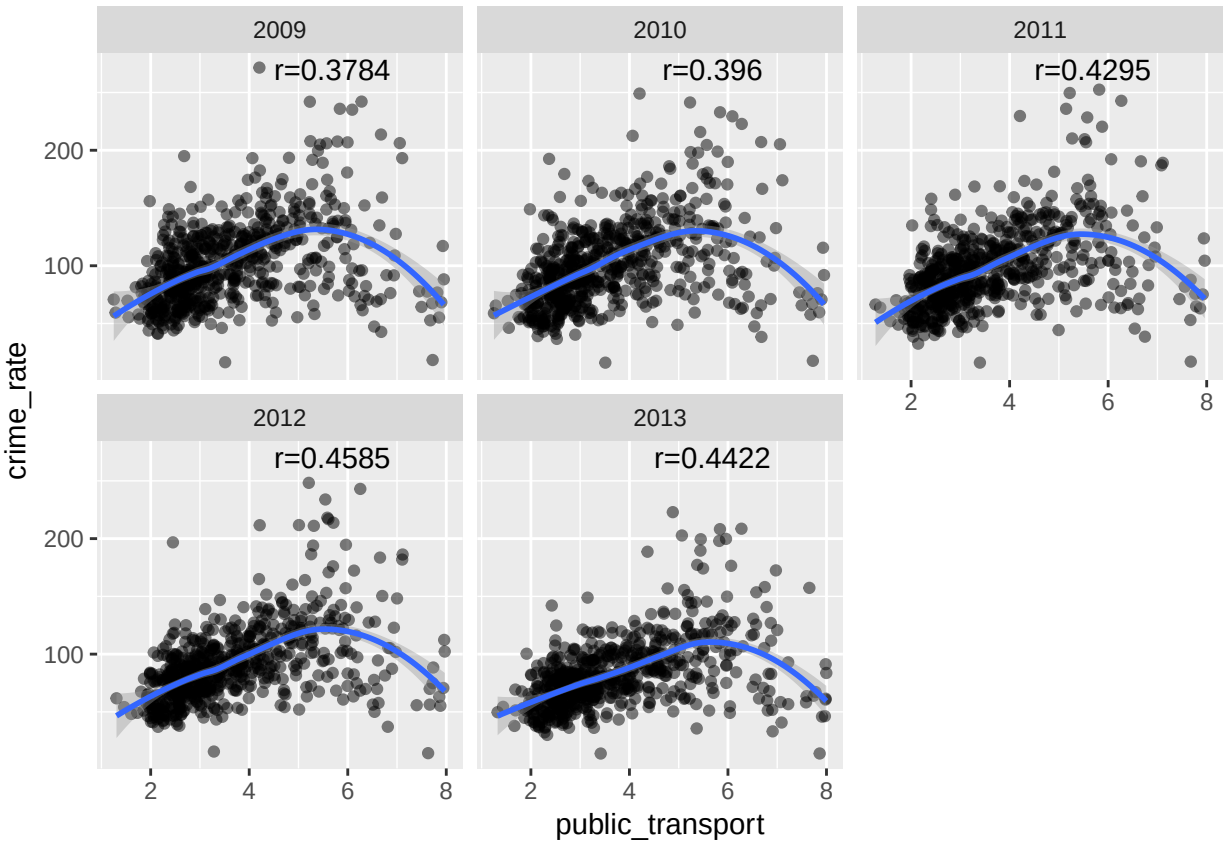

```
## `geom_smooth()` using method = 'loess'
```



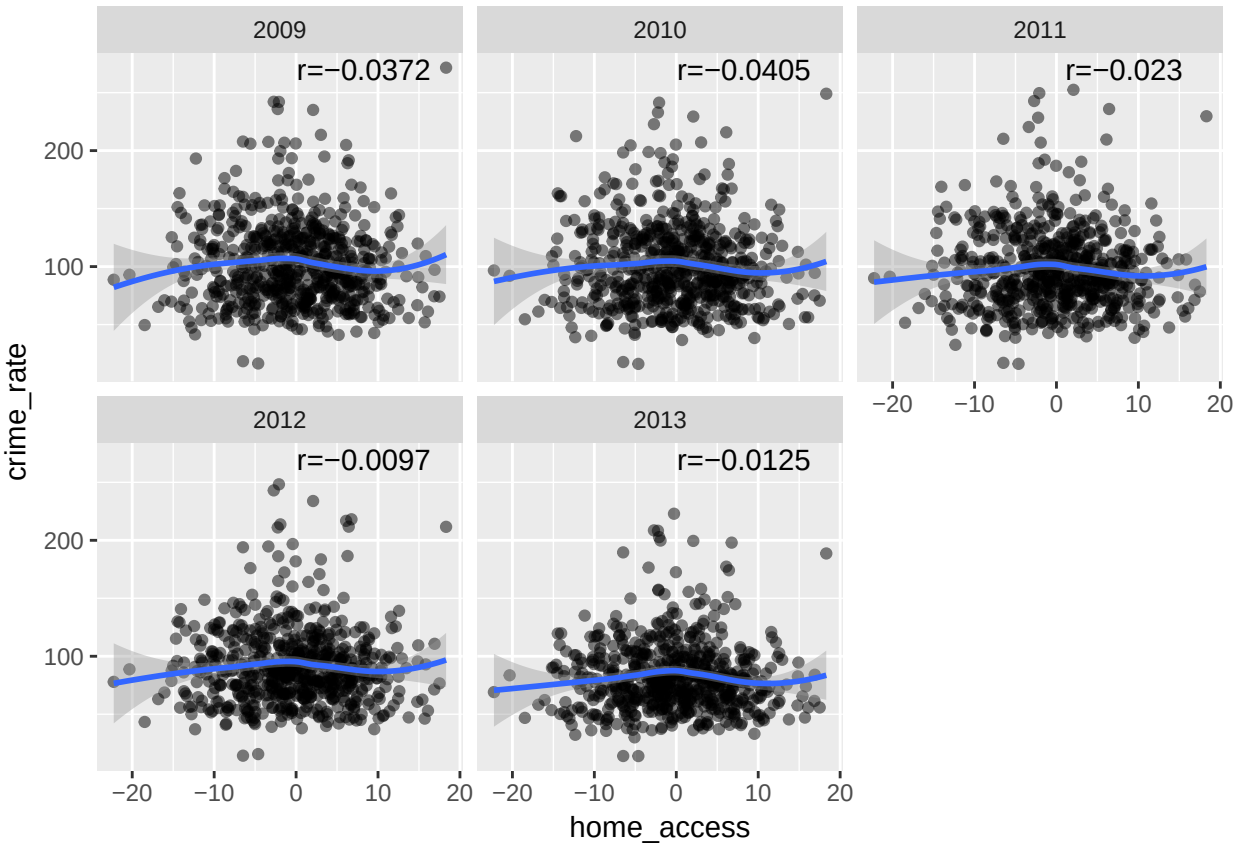

```
## `geom_smooth()` using method = 'loess'
```




```
## `geom_smooth()` using method = 'loess'
```




```
## `geom_smooth()` using method = 'loess'
```



Correlations between crime rate and 3 health-related indexes

Description

These three health-realated indexes including rolling 5-year combined life expectancies, prevalence of obsity by area of child residence, and incapcaity benefit or servere disablement allowance claimant rate.

Rolling 5-year combined life expectancies are used for wards to reduce the efffects of the variability in number of deaths in each year. Index scores were reversed so higher life expectancy equals better well-being. Source: ONS mortality data and GLA population projections, GLA Calculations.

"Prevalence of obesity by area of child residence. The estimates use the latest three years of NCMP data combined eg '2013' covers 2010/11 to 2012/13. Children with a BMI greater than or equal to the 95th centile of the British 1990 growth reference (UK90) BMI distribution have been classified as obese. Earliest data available is 2008/09 to 2010/11, which has been used for 2009-2011. Source: National Obesity Observatory"

"Incapacity Benefit or Severe Disablement Allowance claimant rate. Incapacity Benefit (IB) is paid to people who are incapable of work and who meet certain contribution conditions. Severe Disablement Allowance (SDA) is paid to those unable to work for 28 weeks in a row or more because of illness or disability. SDA was removed for new claims in April 2001. Time period used is a snapshot of May from each year. Demominator is population aged 16-64. Source: IB/SDA from DWP, Population from GLA projections." ###basic data selecting

```
#read all the names of columns
crime_health <- select(crime_ward, Ward:`Incapacity Benefit rate - 2013`, -(`Life Expectancy 2005-2009`
crime_health[,3:18] <- sapply(crime_health[,3:18], as.numeric)
crime_health <- mutate(crime_health, `Childhood Obesity 2009-2013`=(`Childhood Obesity 2009`+`Childhood
crime_health <- mutate(crime_health, `Incapacity Benefit rate - 2009-2013`=(`Incapacity Benefit rate - 
```

```
crime_health <- mutate(crime_health, `Crime rate - 2009-2013`=(`Crime rate - 2009`+`Crime rate - 2010`+

crime_health1 <- select(crime_health, -(`Childhood Obesity 2009`:`Crime rate - 2013`))
```

Correlation Analysis

```
names(crime_health1)
```

```
## [1] "Ward"
## [2] "Borough"
## [3] "Life Expectancy 2009-13"
## [4] "Childhood Obesity 2009-2013"
## [5] "Incapacity Benefit rate - 2009-2013"
## [6] "Crime rate - 2009-2013"
```

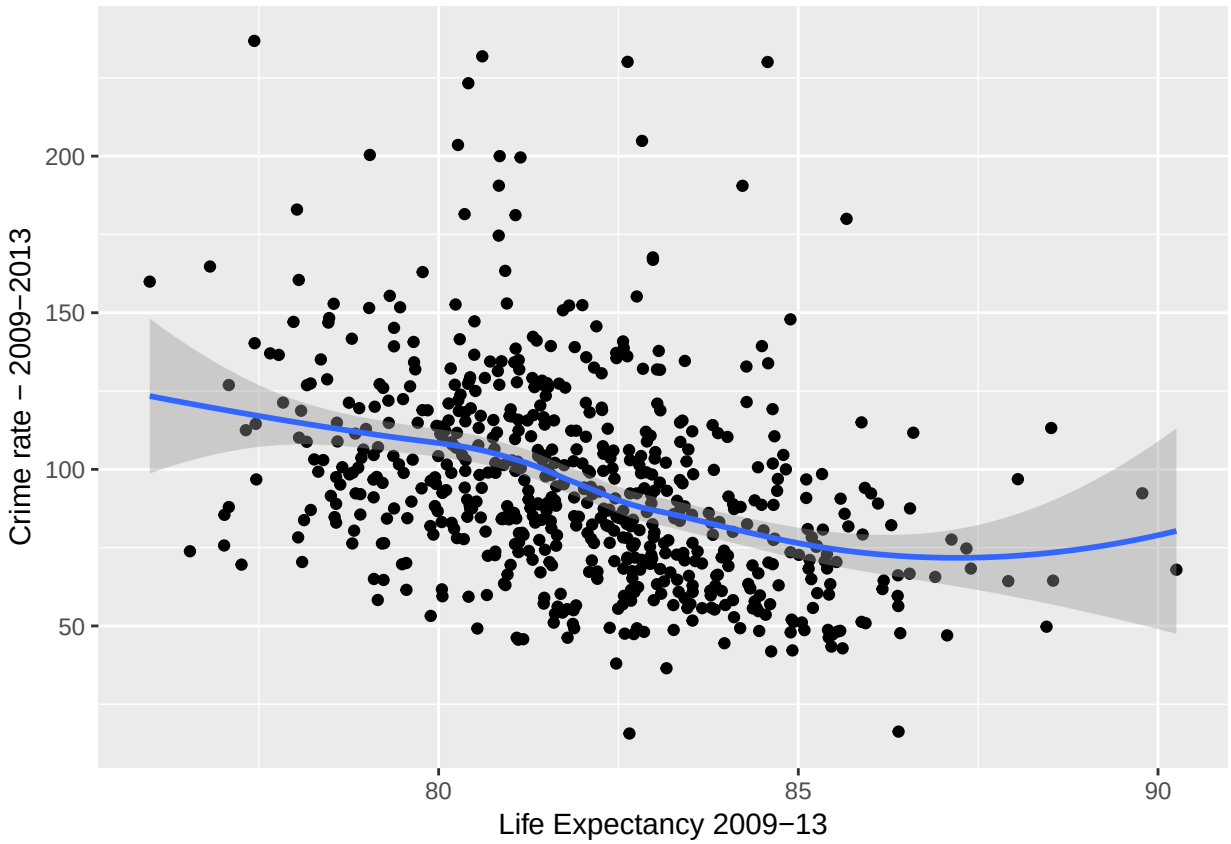
```
crime_health2 <- crime_health1[, c(6,3,4,5)]
cor(crime_health2)
```

```
##                                     Crime rate - 2009-2013
## Crime rate - 2009-2013                           1.0000000
## Life Expectancy 2009-13                         -0.3755371
## Childhood Obesity 2009-2013                      0.4081346
## Incapacity Benefit rate - 2009-2013              0.4472021
##                                     Life Expectancy 2009-13
## Crime rate - 2009-2013                           -0.3755371
## Life Expectancy 2009-13                           1.0000000
## Childhood Obesity 2009-2013                      -0.4892902
## Incapacity Benefit rate - 2009-2013              -0.5907696
##                                     Childhood Obesity 2009-2013
## Crime rate - 2009-2013                                 0.4081346
## Life Expectancy 2009-13                               -0.4892902
## Childhood Obesity 2009-2013                            1.0000000
## Incapacity Benefit rate - 2009-2013                    0.6100361
##                                     Incapacity Benefit rate - 2009-2013
## Crime rate - 2009-2013                                        0.4472021
## Life Expectancy 2009-13                                      -0.5907696
## Childhood Obesity 2009-2013                                   0.6100361
## Incapacity Benefit rate - 2009-2013                           1.0000000
```

Plots

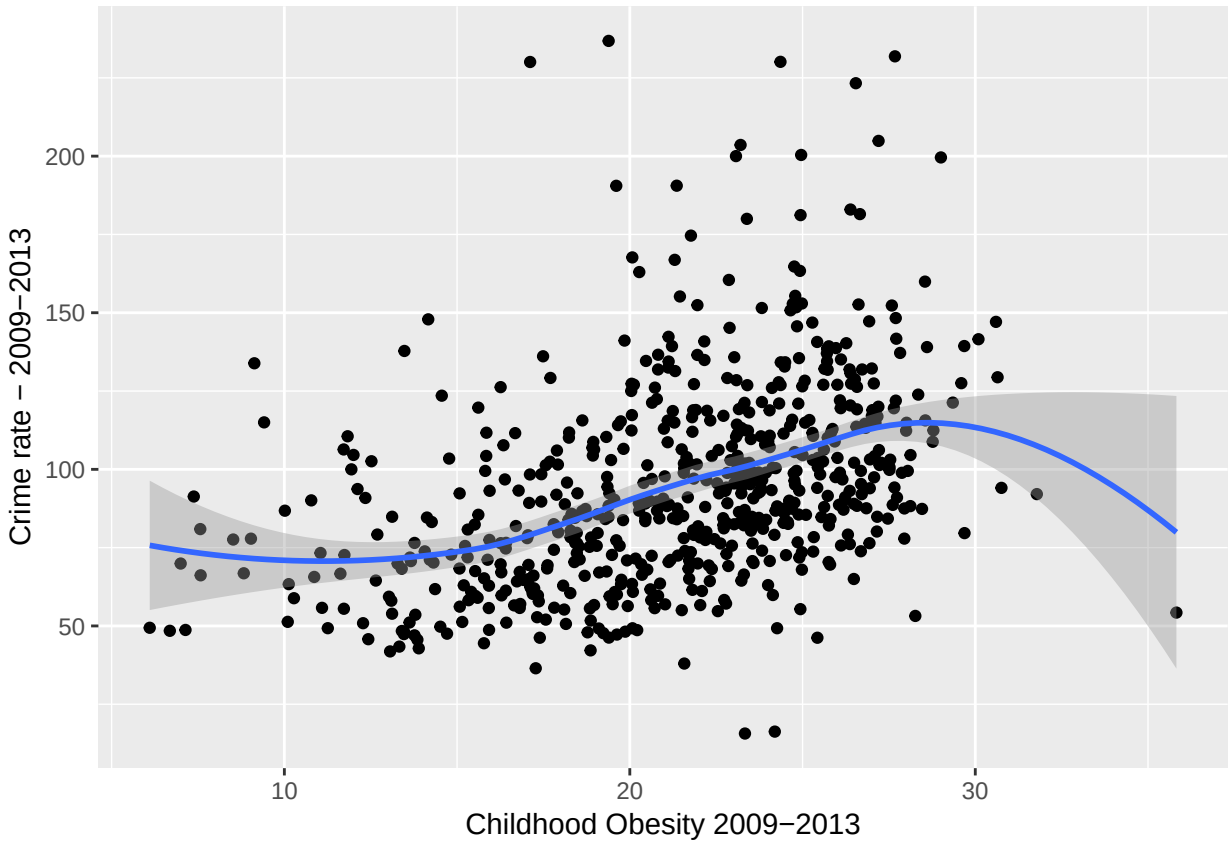
```
ggplot(crime_health2, aes(x=`Life Expectancy 2009-13`, y=`Crime rate - 2009-2013`))+geom_point()+geom_sm
```

```
## `geom_smooth()` using method = 'loess'
```



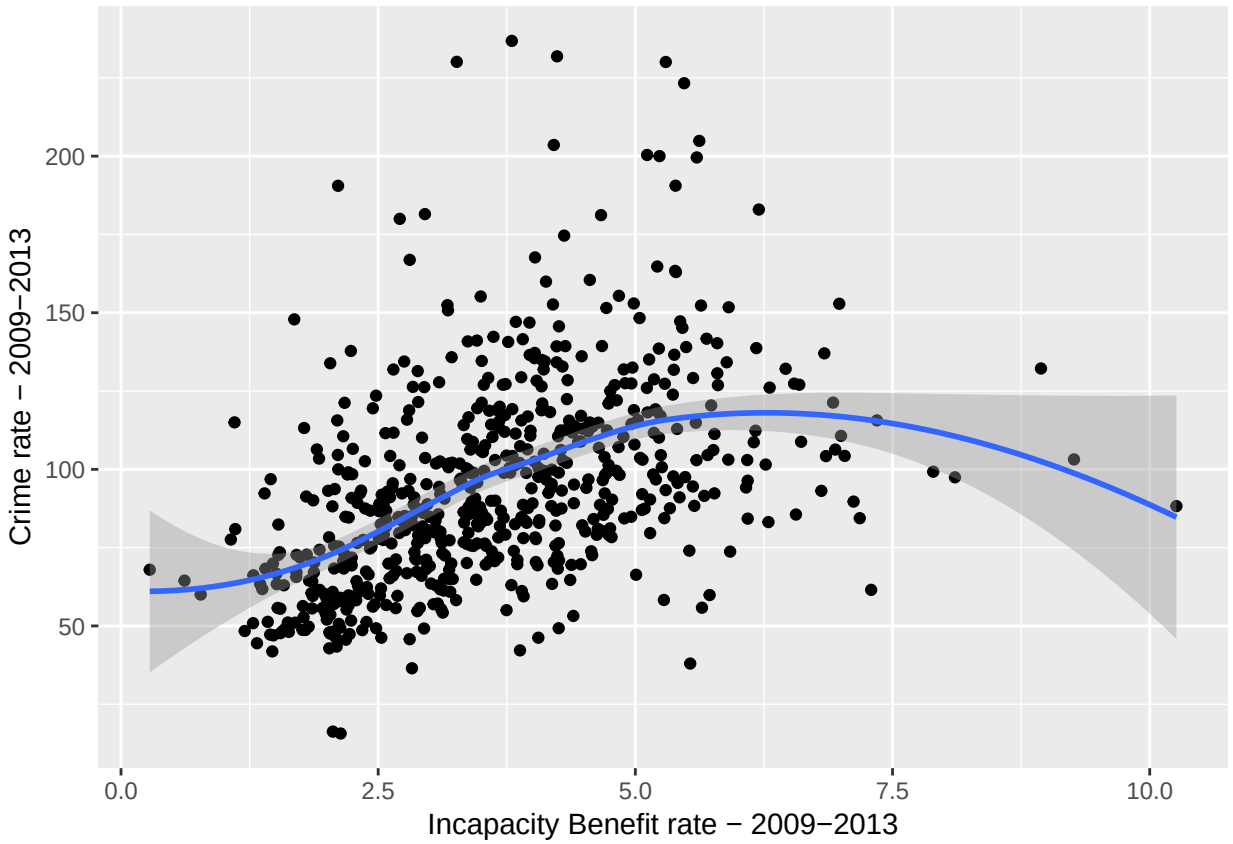

```
ggplot(crime_health2, aes(x=`Childhood Obesity 2009-2013`, y=`Crime rate - 2009-2013`))+geom_point()+ge
```

```
## `geom_smooth()` using method = 'loess'
```

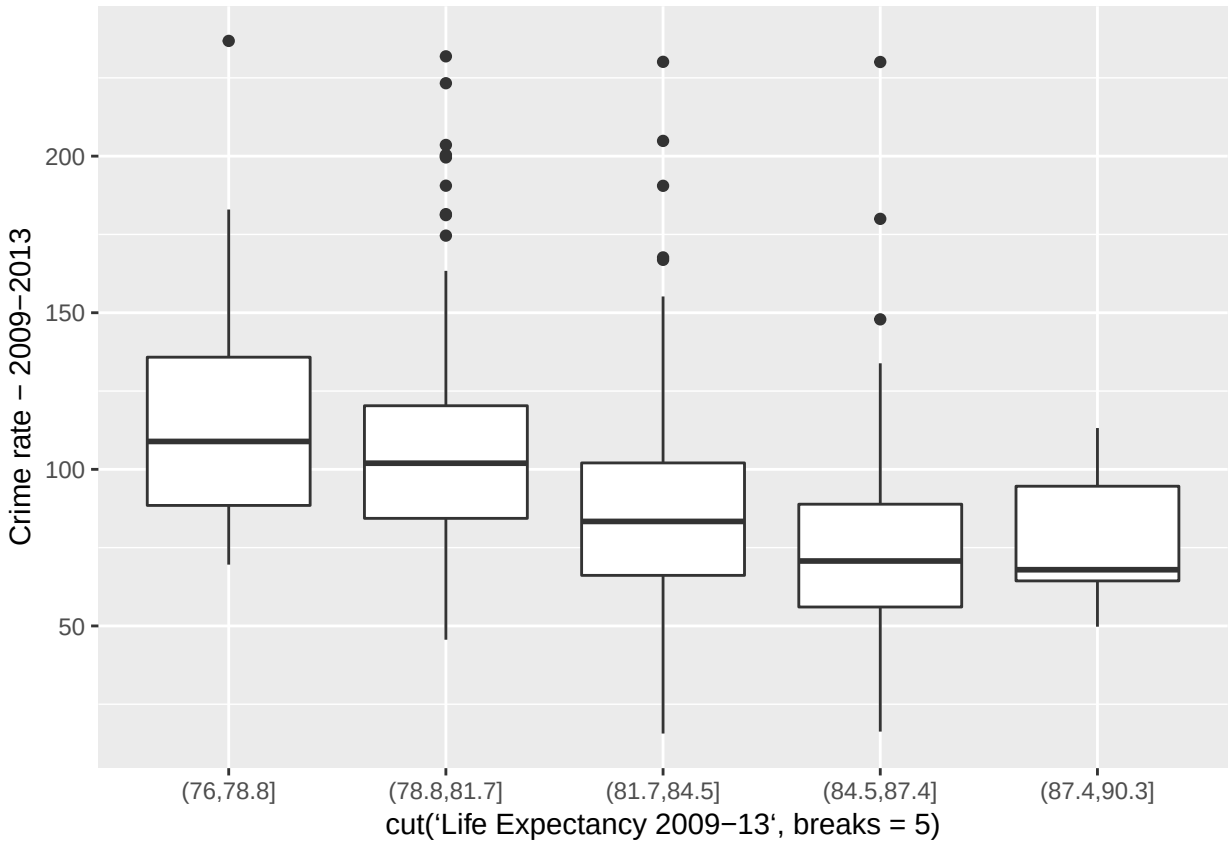



```
ggplot(crime_health2, aes(x=`Incapacity Benefit rate - 2009-2013`, y=`Crime rate - 2009-2013`))+geom_po
```

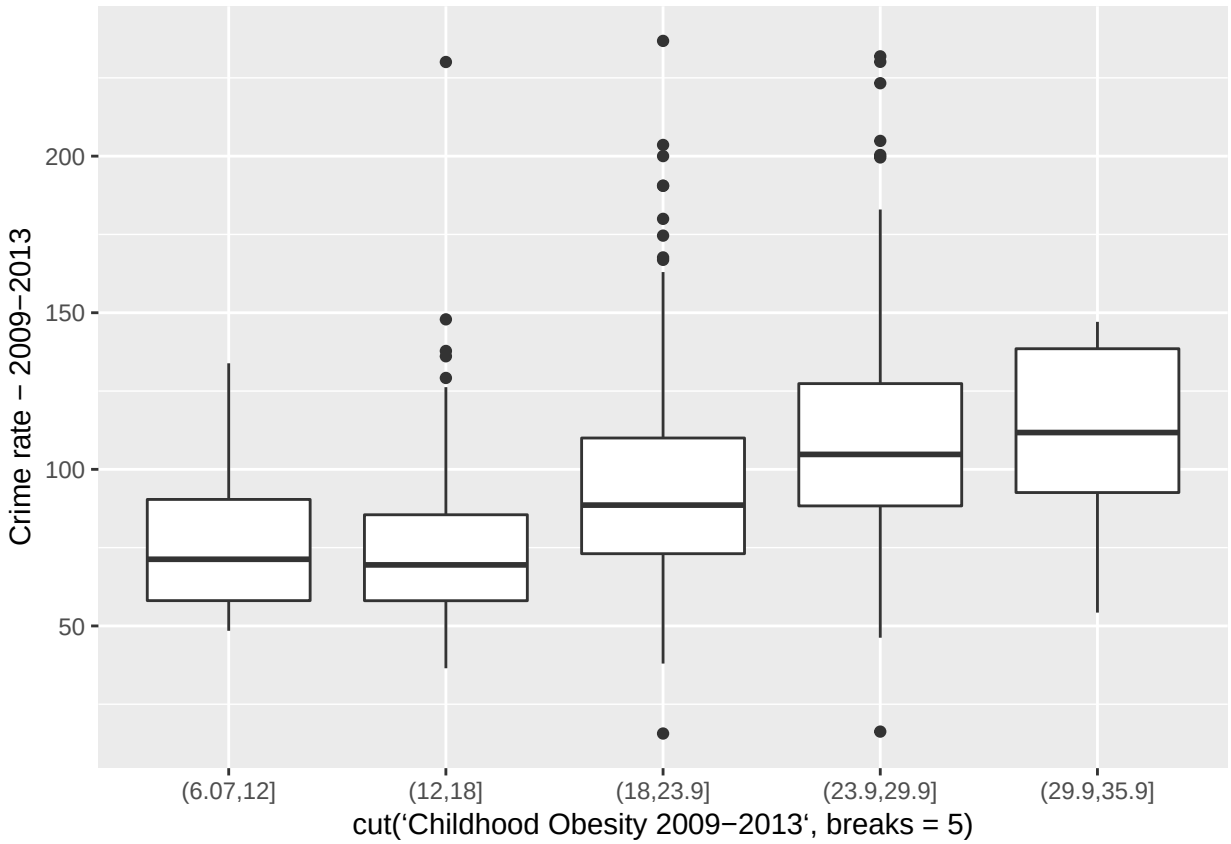
```
## `geom_smooth()` using method = 'loess'
```



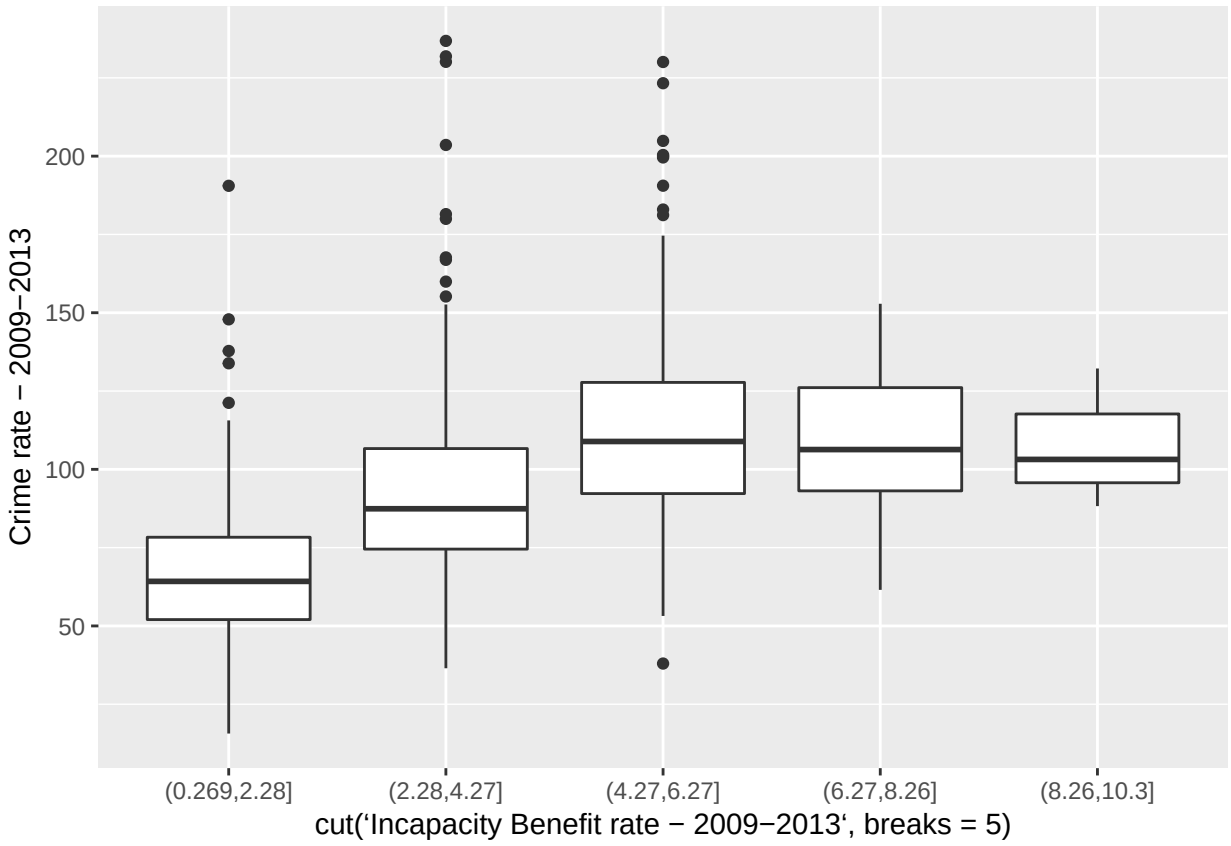

```
ggplot(crime_health2, aes(x=cut(`Life Expectancy 2009-13`, breaks =5), y=`Crime rate - 2009-2013`))+geo
```




```
ggplot(crime_health2, aes(x=cut(`Childhood Obesity 2009-2013`, breaks =5), y=`Crime rate - 2009-2013`))
```




```
ggplot(crime_health2, aes(x=cut(`Incapacity Benefit rate - 2009-2013`, breaks =5), y=`Crime rate - 2009
```



Linear regression

Conclusion and analysis