

Bovine Spongiform Encephalopathy Surveillance Statistics in United Kingdom and a Shiny App for Raman Spectra Analysis

Shaowei Ding

April 24, 2014

1 Background

1.1 What is BSE?

Bovine Spongiform Encephalopathy(BSE), or the so called "Mad Cow" disease, is a kind of prion related disease, which is fatal and trigger neurological dysfunction in Cattles. Prions are neither bacteria nor viruses, with no genetic information, but still causes disease. It has long incubation period and progress inexorably once clinical conditions occur.(Araujo,2013) BSE can be transmitted by eating food containing brain,spinal cord or other nervous system tissue from an infected animal, and prions begin to slowly transform normal protein into the abnormal prion shape, which eventually leads to fatal damage to the nervous system.The human form of BSE is called Creutzfeldt-Jakob Disease(CJD).BSE can cause death usually within one year of onset of illness.

1.2 Current condition

Currently disease diagnosis is still limited, and there isn't mature technique that can detect BSE at a earlier stage. Diseased animals have protein transformation in the central nervous system. Since retina is a derivative of forebrain, it is the most exposed portion of the central nervous system. Raman spectroscopy is a kind of non-invasive and non-destructive technique and can be used to measure unique fingerprint for different disease. And by applying this spectroscopy in retina, the most exposed nervous tissue; it is highly possible that BSE can be diagnosed accurately and at earlier stage.

2 Motivation

Bovine Spongiform Encephalopathy disease was first found in the mid-1980s from 16 cattle, and that number dramatically increased to over 190,000 cases worldwide, with the majority of them in Europe.(Lee et al., 2013) ,Which led to a disaster to human, the agricultural industry and the food industry. The worst case of BSE was happened in the United Kingdom, and in this analysis, we will take a close look at the development of BSE in the United Kingdom.Since the research I am doing is about BSE diagnosis, and I think it is interesting to look at the history of this disease. Moreover, by investigating this case, it makes people have a better awareness of the importance of food safety. This project will first discuss the history and the trend of BSE disease happened in the United Kingdom, then I will build a shiny application to analyze raman spectra, which can be used to do a brief analysis and give a preview of the measured Raman spectra.

3 Data Collecting

These data of United Kingdom is obtained from Animal Health and Veterinary Laboratories Agency,"<http://www.defra.gov.uk/publication/tse-stats-cattle/>".BSE cases in every year/every month is available, and it is also

categorized into active and passive surveillance. Also the birth date and age of all suspected animals are recorded. Four pdfs have data from before year 1987 to 2014, and then I converted them to csv format. Additionally, I downloaded the UK shape file, which can be used to plot the distribution of BSE. Since the name of every city in two files don't match, I need to do some modification to plot the map.

4 Possible questions that could be answered

1. What is the overall trending of BSE from 1987 to 2014? What is the trending of passive Surveillance and active surveillance?
2. What are the locations of BSE?
3. What is the percent reduction year on year?
4. How many confirmed cases of BSE per million head of cattle population over 24 months of age?
5. What is the trending of BSE by confirmed cases of BSE in the UK by year of birth and age?
6. On August 1st, 1996, extra control measures on animal feed containing mammalian meat and bone meal are considered to have been fully implemented, and what is trending of BSE after this date?

5 Analysis of Mad Cow Disease from 1997 to 2014

5.1 Overall BSE cases happened in great britain and northern Ireland

First let's take a look at the general trend of BSE cases from 1987 to 2014 based on passive and active surveillance.

```
library(reshape)

## Loading required package: plyr
##
## Attaching package: 'reshape'
##
## The following objects are masked from 'package:plyr':
##   rename, round_any

library(ggplot2)

pa <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/passive and active.csv")
pa$Year <- as.character(pa$Year)
pa$Year[pa$Year == "Pre 1988"] <- "1986-1988"
a.melt <- melt(pa[1:28, c(1:3, 5:6)], id = c("Year"))
pd <- position_dodge(0.1)
r <- ggplot(data = a.melt, aes(x = Year, y = value, colour = variable)) + geom_line(position = pd,
  aes(group = variable))
plot(r)

## ymax not defined: adjusting position using y instead
```

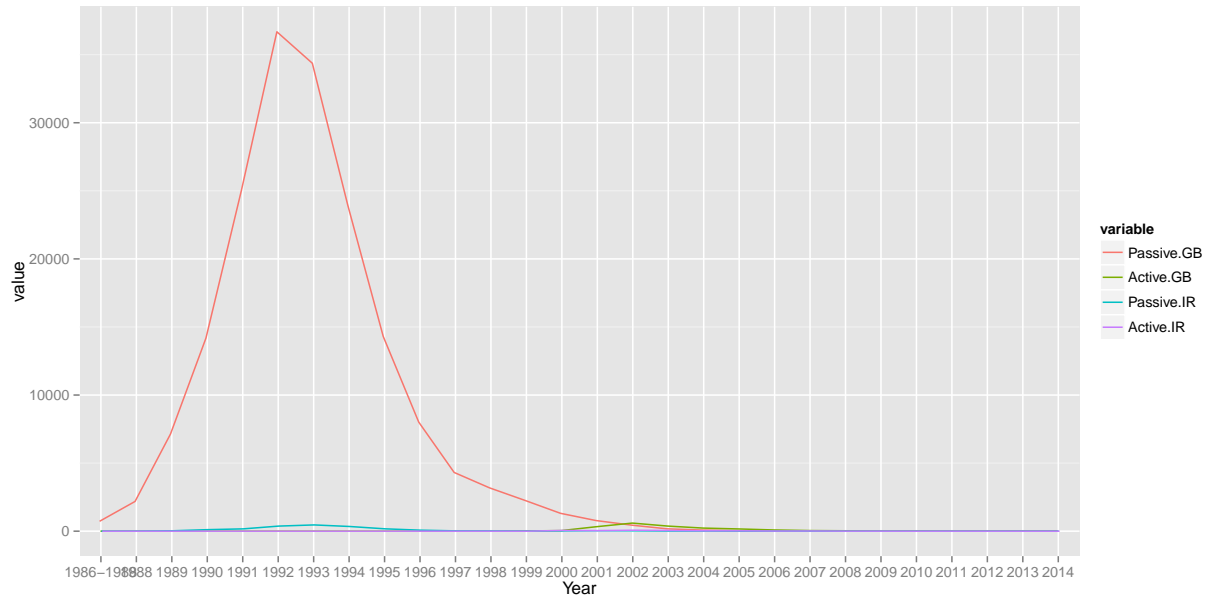


Figure 1: the general trend of BSE cases from 1987 to 2014 based on passive and active surveillance.

This plot shows the majority of data is from passive surveillance in great britain, and BSE cases were the most from 1989-1996. Now we can take a look at condition in great britain.

5.2 Trend of passive surveillance in great britain

```
overall <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/OVERALL.csv")
trend <- overall[1:28, ]
trend.det <- melt(trend)

## Using YEAR as id variables

pd <- position_dodge(0.1)
r <- ggplot(data = trend.det, aes(x = YEAR, y = value, colour = variable)) +
  geom_line(position = pd, aes(group = variable))
plot(r)

## ymax not defined: adjusting position using y instead
```

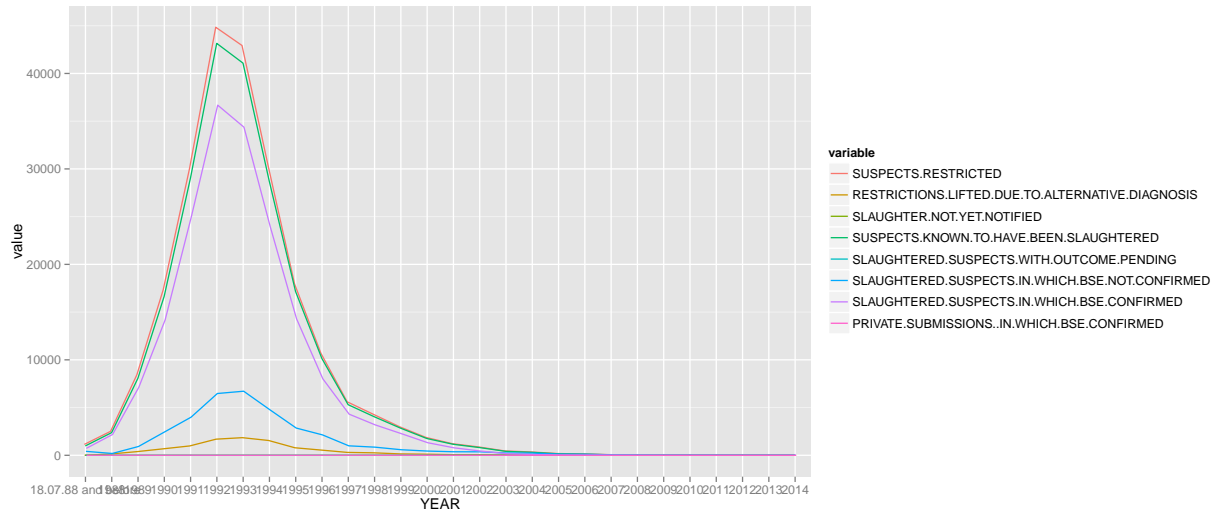


Figure 2: the trend of passive surveillance in great britain.

This plot shows the trend of restricted case and confirmed case. They all follow the similar curve, but the amplitude is different. The amount of suspected cases is a lot and almost all of those animals were slaughtered. Among them, the purple curve represents the actual confirmed cases.

5.3 Cattle population over 24 months of ages and number of confirmed cases per million

Since the population of cattles were not constant during the time when people recorded the data, it is meaningful to take a look at the percentage of suspected of animals based on a time interval. Also, because of a big mortality rate, the analysis can be focused on cattle with age older than 24 months. The x axis is the period when people took measurement, and the y axis is the amount of cattles confirmed with BSE disease.

```
library(reshape)
library(ggplot2)
library(Mcow)

##
## Attaching package: 'Mcow'
##
## The following object is masked from 'package:base':
##
##   date

age24 <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/millioncattle.csv")
age24 <- age24[c(1:75), -c(5, 8)]
colnames(age24) <- c("start", "end", "million of cattles", "# of case by date of confirmation",
  "# of case by date of restriction", "# per million by data of confirmation",
  "# per million by date of restriction")
age.sub <- age24[-1, ]
age.sub$name <- 1:74
age.sub[, 1] <- date(age.sub[, 1])
age.sub[, 2] <- date(age.sub[, 2])
age.sub$"# per million by data of confirmation" <- as.numeric(as.character(age.sub$"# per million by data of confirmation"))
mdfr <- melt(age.sub, measure.vars = c("start", "end"))
a <- ggplot(mdfr, aes(mdfr$value, as.factor(mdfr$"# million of cattles"))) + geom_line(size = 1) +
  xlab("year") + ylab("million of cattles") + theme_bw()
plot(a)
```

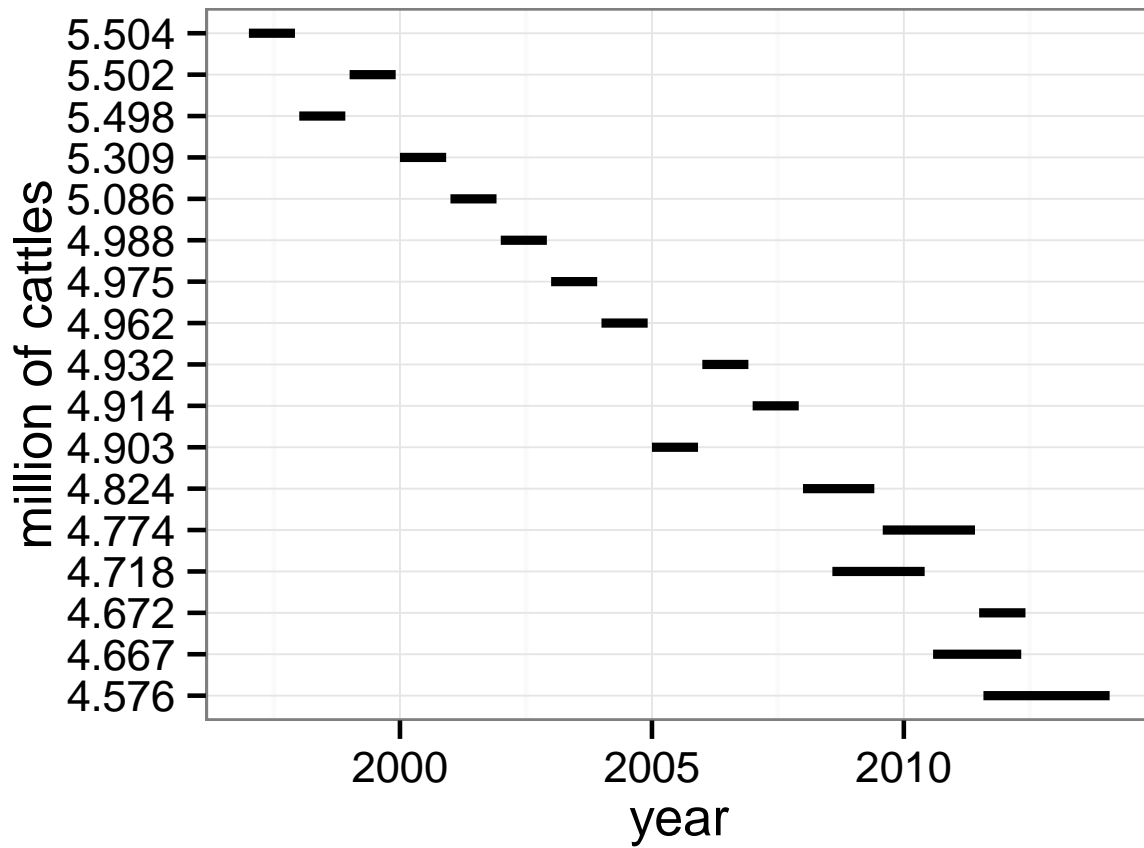


Figure 3: Population of cattles.

This plot indicates the population from 1987 to 2014 has been decreasing and the lowest level is in 2014, which might be under the influence of BSE disease. BSE caused the decrease of the demand of beef and this corresponds to the trend as shown in the plot.

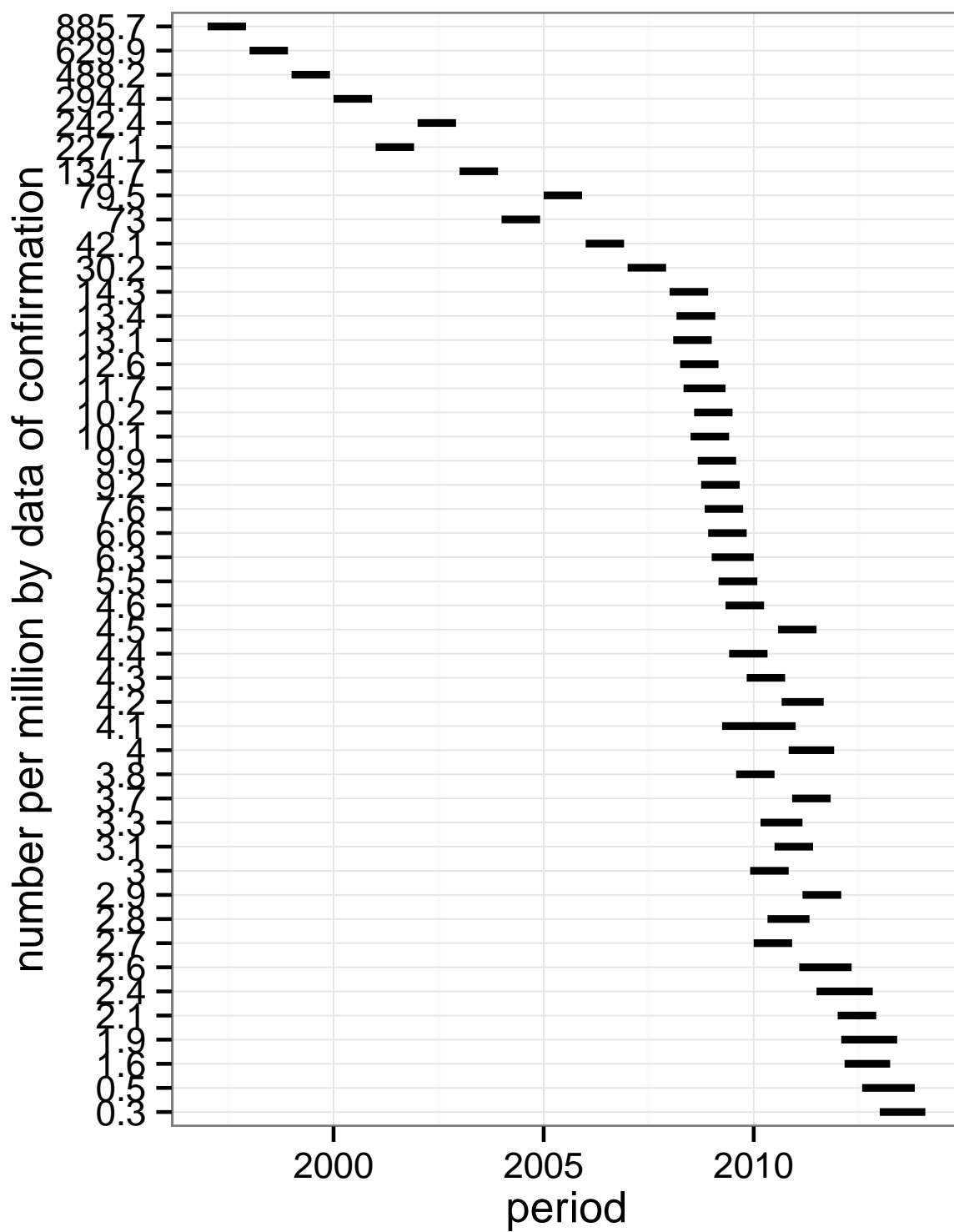


Figure 4: Number of confirmed BSE per million by date of confirmation

With adequate control, from 2008 to 2014 the percentage of BSE cases has been decreasing dramatically. Compared to this, the period from 1987 to 2008, when the BSE was relatively more serious, shows slower

decreasing rate.

5.4 Distribution of BSE in Great Britain

```
library(maptools)

## Loading required package: sp
## Checking rgeos availability: FALSE
## Note: when rgeos is not available, polygon geometry computations in maptools depend on gpclib,
## which has a restricted licence. It is disabled by default;
## to enable gpclib, type gpclibPermit()

library(Mcow)
city <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/city.csv")
uk <- readShapeSpatial("C:/Users/sding/Documents/GitHub/585xproject/uk/map.shp")
xxx <- thinnedSpatialPoly(as(uk, "SpatialPolygons"), tolerance = 0.1, minarea = 0.001,
  topologyPreserve = TRUE)
oz <- extractPolygons(xxx)
# ggplot(oz, aes(x = x, y = y, group = group)) +
#   geom_polygon(colour='black', fill='white')
uk@data$MM_UID <- 1:192
oz.new <- merge(oz, uk@data, by.x = "region", by.y = "MM_UID")
# Clean the data
city$COUNTY <- gsub("&", "and", city$COUNTY)
city$COUNTY[city$COUNTY == "Ayrshire"] <- "South Ayrshire"
city$COUNTY[city$COUNTY == "Cornwall and Isles of Scilly"] <- "Cornwall"
city$COUNTY[city$COUNTY == "East Riding and Northern Lincolnshire"] <- "North Lincolnshire"
city$COUNTY[city$COUNTY == "Eileanan an Iar"] <- "Eilean Siar"
city$COUNTY[city$COUNTY == "Greater London"] <- "London"
city$COUNTY[city$COUNTY == "Greater Manchester"] <- "Manchester"
city$COUNTY[city$COUNTY == "Gloucestershire excl South"] <- "South Gloucestershire"
city$COUNTY[city$COUNTY == "Leicestershire and Rutland"] <- "Leicestershire"
city$COUNTY[city$COUNTY == "Lincolnshire excl North"] <- "North Lincolnshire"
city$COUNTY[city$COUNTY == "North-East Scotland"] <- "Aberdeen"
city$COUNTY[city$COUNTY == "North-East Wales"] <- "Herefordshire"
city$COUNTY[city$COUNTY == "North-West Wales"] <- "Ceredigion"
city$COUNTY[city$COUNTY == "Northern Somerset and South Gloucester"] <- "Northern Somerset"
city$COUNTY[city$COUNTY == "Somerset excl North"] <- "North Somerset"
city$COUNTY[city$COUNTY == "South Wales"] <- "Cardiff"
colnames(city) <- c("region", "farms", "cases")
oz.subset <- oz.new[, c(2:5, 7, 8, 9)]
colnames(oz.subset) <- c("x", "y", "order", "group", "area", "region", "division")
# merge two files
distribution <- merge(oz.subset, city, by = "region")
# distribution of farms and BSE cases
distribution$cases <- as.numeric(distribution$cases)
a <- ggplot() + geom_polygon(data = oz.new, aes(x = x, y = y, order = order,
  group = group)) + geom_polygon(data = distribution, aes(x = x, y = y, order = order,
  group = group, fill = distribution$cases))
plot(a)
```

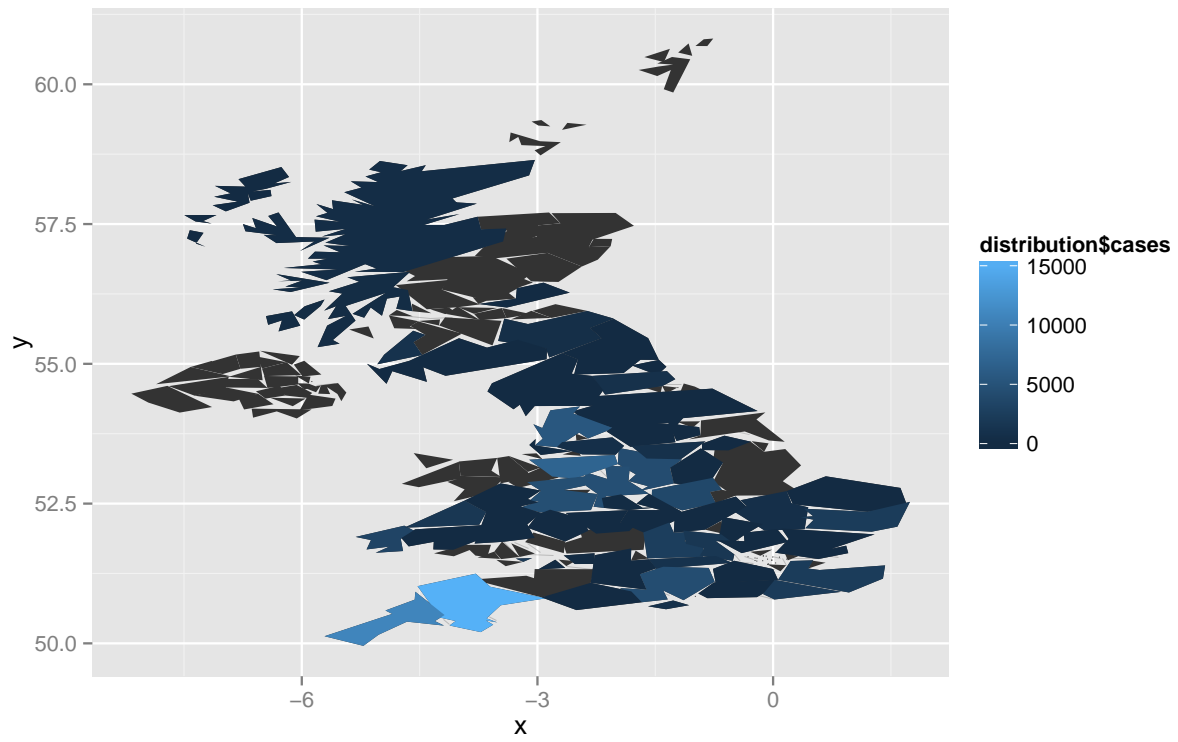


Figure 5: Distribution of BSE cases in Great Britain



Figure 6: Distribution of farms where BSE have been found in Great Britain

The color gradient indicates the distribution of BSE cases and farms where BSE cases have been found.

5.5 Percent reduction of BSE cases in each year

```
red <- trend[, c(1, 2, 8)]
red$"percent reduction year on year(suspected)" <- NA
red$"percent reduction year on year(confirmed)" <- NA
i <- NA
j <- NA
per <- NA
red[1:28, 2:3] <- as.matrix(sapply(red[1:28, 2:3], as.numeric))
percent <- function(x, digits = 2, format = "f", ...) {
  paste(formatC(100 * x, format = format, digits = digits, ...), "%", sep = "")
}

for (i in 3:(length(red$SUSPECTS.RESTRICTED))) {
  red$"percent reduction year on year(suspected)"[i] <- as.numeric((-red$SUSPECTS.RESTRICTED[i -
1] + red$SUSPECTS.RESTRICTED[i])/red$SUSPECTS.RESTRICTED[i - 1])
  red$"percent reduction year on year(confirmed)"[i] <- as.numeric((-red$SLAUGHTERED.SUSPECTS.IN.WHICH.BSE.CONFIRMED[i -
1] + red$SLAUGHTERED.SUSPECTS.IN.WHICH.BSE.CONFIRMED[i])/red$SLAUGHTERED.SUSPECTS.IN.WHICH.BSE.CONFIRMED[i -
1])
  if (red$SLAUGHTERED.SUSPECTS.IN.WHICH.BSE.CONFIRMED[i - 1] == 0) {
    red$"percent reduction year on year(confirmed)"[i] <- 0
  }
}

red.melt <- melt(red[, c(1, 4, 5)], id = c("YEAR"))
red.melt$value <- as.numeric(red.melt$value)
red.melt$value[is.na(red.melt$value)] <- 0
pd <- position_dodge(0.1)
a <- ggplot(data = red.melt, aes(x = YEAR, y = value, colour = variable)) +
```

```
geom_line(position = pd, aes(group = variable))
plot(a)

## ymax not defined: adjusting position using y instead
```

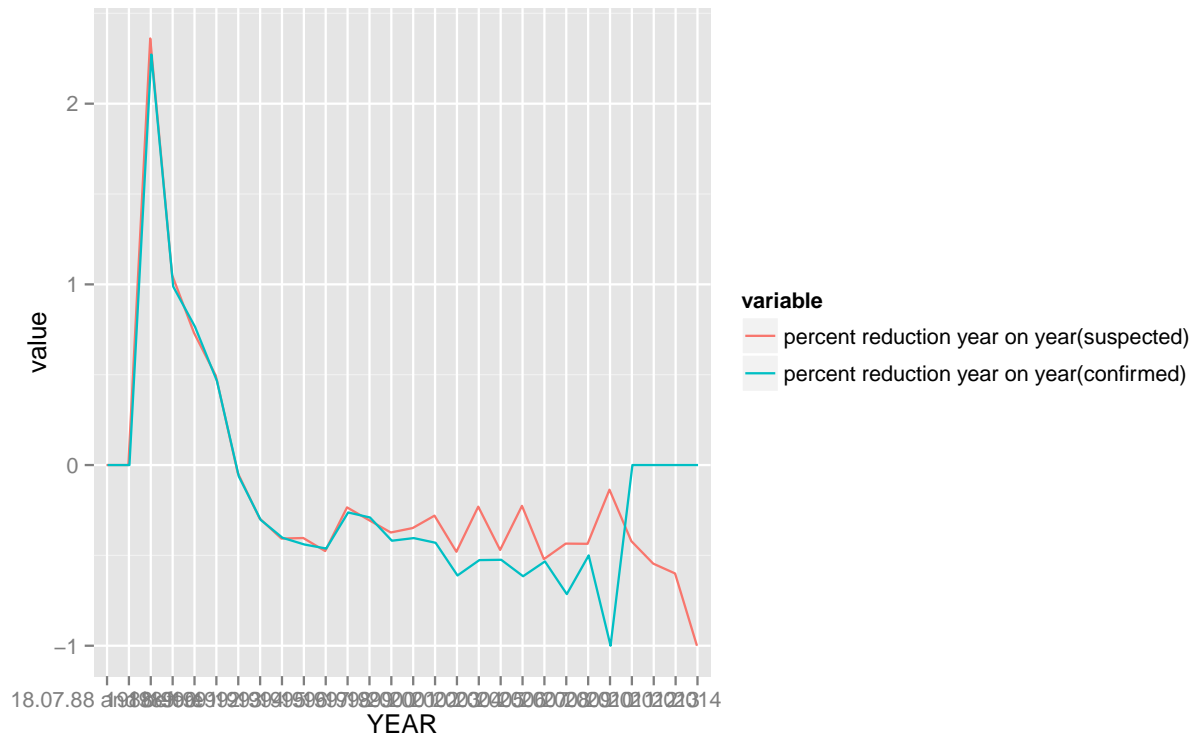


Figure 7: Percentage reduction

This plot shows the percentage reduction of confirmed and suspected cases year on year.

5.6 The trend of confirmed BSE cases of different age in great britain

```
age <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/age.gb.csv")
age <- age[-1, -2]
age$"Birth Period" <- 1982:2006

## Error: replacement has 25 rows, data has 27

colnames(age) <- c("Birth Period", "1 year old", "2 year old", "3 year old",
  "4 year old", "5 year old", "6 year old", "7 year old", "8 year old", "9 year old",
  "10 year old")
age.melt <- melt(age, id = "Birth Period")
age.melt$value <- as.numeric(age.melt$value)
age.melt$"Birth Period" <- as.numeric(age.melt$"Birth Period")
age.melt$variable <- as.factor(age.melt$variable)

pd <- position_dodge(0.1)
a <- ggplot(data = age.melt, aes(x = age.melt$"Birth Period", y = value, colour = variable)) +
  geom_line(position = pd, aes(group = variable))
plot(a)

## ymax not defined: adjusting position using y instead
```

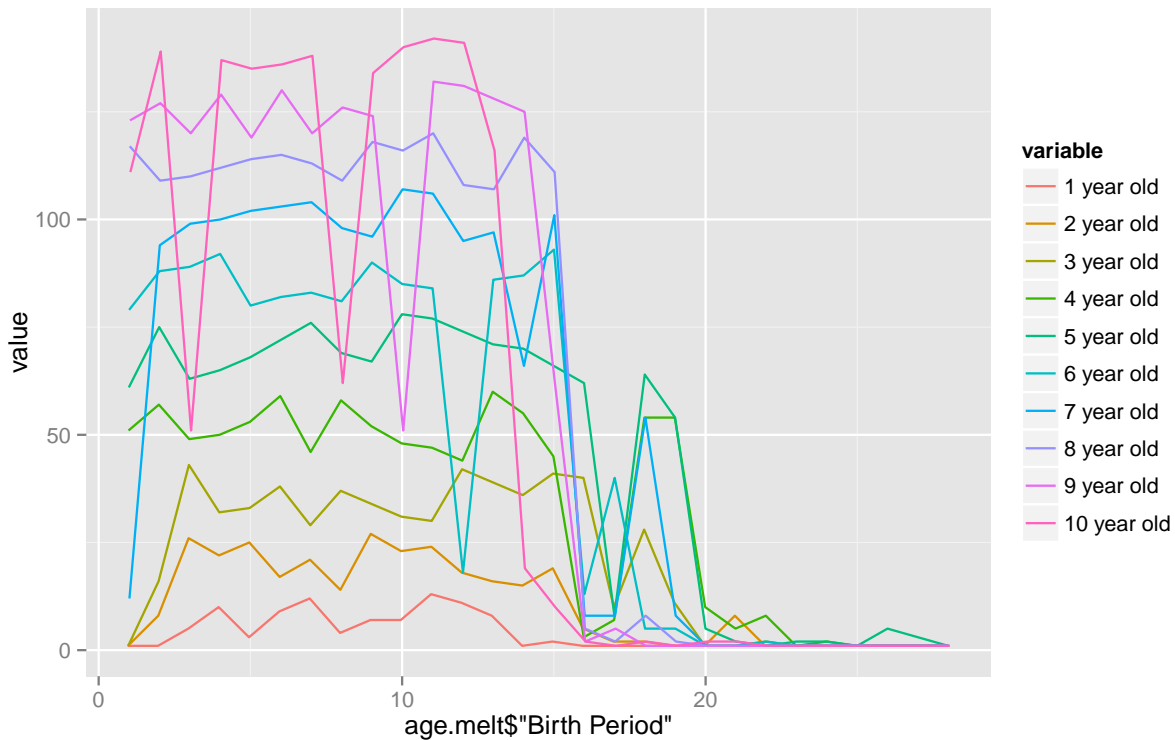


Figure 8: confirmed BSE cases of different age

If we set age as a factor, age of the animals seems to be proportional to number of cases.

5.7 The distribution of BSE cases in Great Britain from 2013 to 2014

```
recent <- read.csv("C:/Users/sding/Documents/GitHub/585xproject/recent.csv")
a <- recent[-1, c(1, 2, 3, 4)]
recent[, c(2, 3, 4)] <- as.matrix(sapply(recent[, c(2, 3, 4)], as.numeric))
a$cases <- rowSums(recent[-1, c(2, 3, 4)])
a$County <- gsub("&", "and", a$County)
a <- a[, -c(2, 3, 4)]
a$County[a$County == "North-West Wales"] <- "Ceredigion"
colnames(a) <- c("region", "cases")

rec.dis <- merge(oz.subset, a, by = "region")
rec.dis$cases <- as.numeric(rec.dis$cases)
a <- ggplot() + geom_polygon(data = oz.subset, aes(x = x, y = y, order = order,
  group = group)) + geom_polygon(data = rec.dis, aes(x = x, y = y, order = order,
  group = group, fill = rec.dis$cases))
plot(a)
```

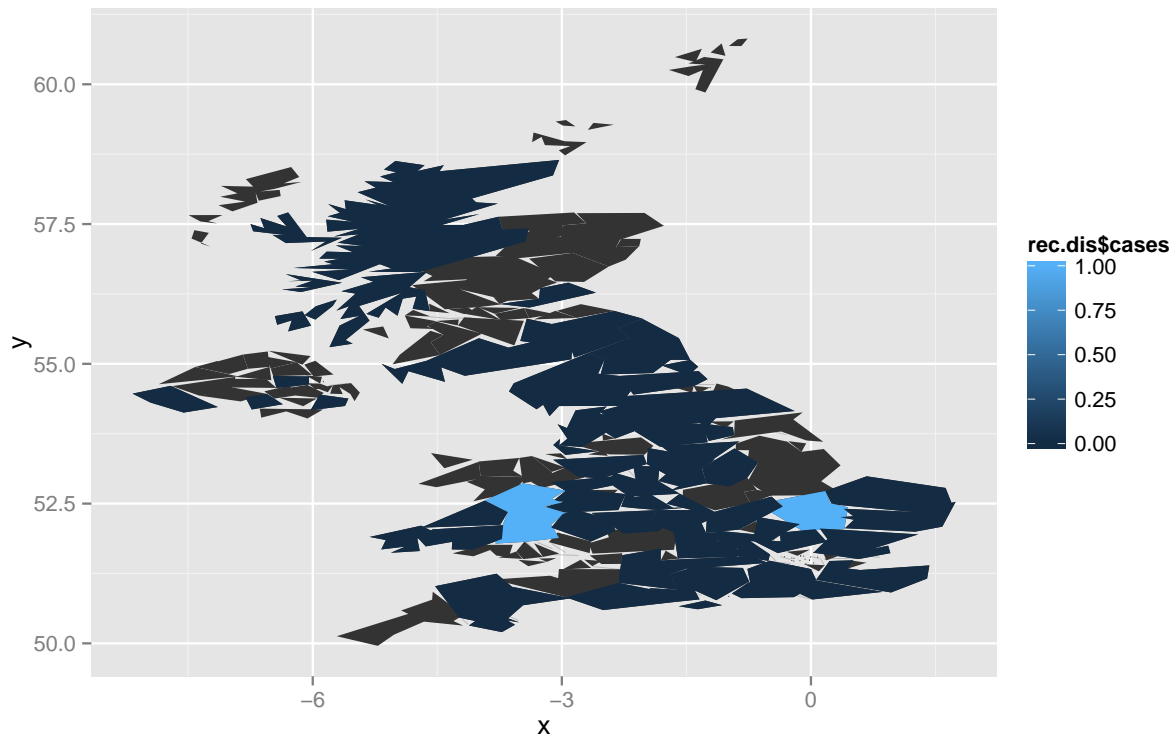
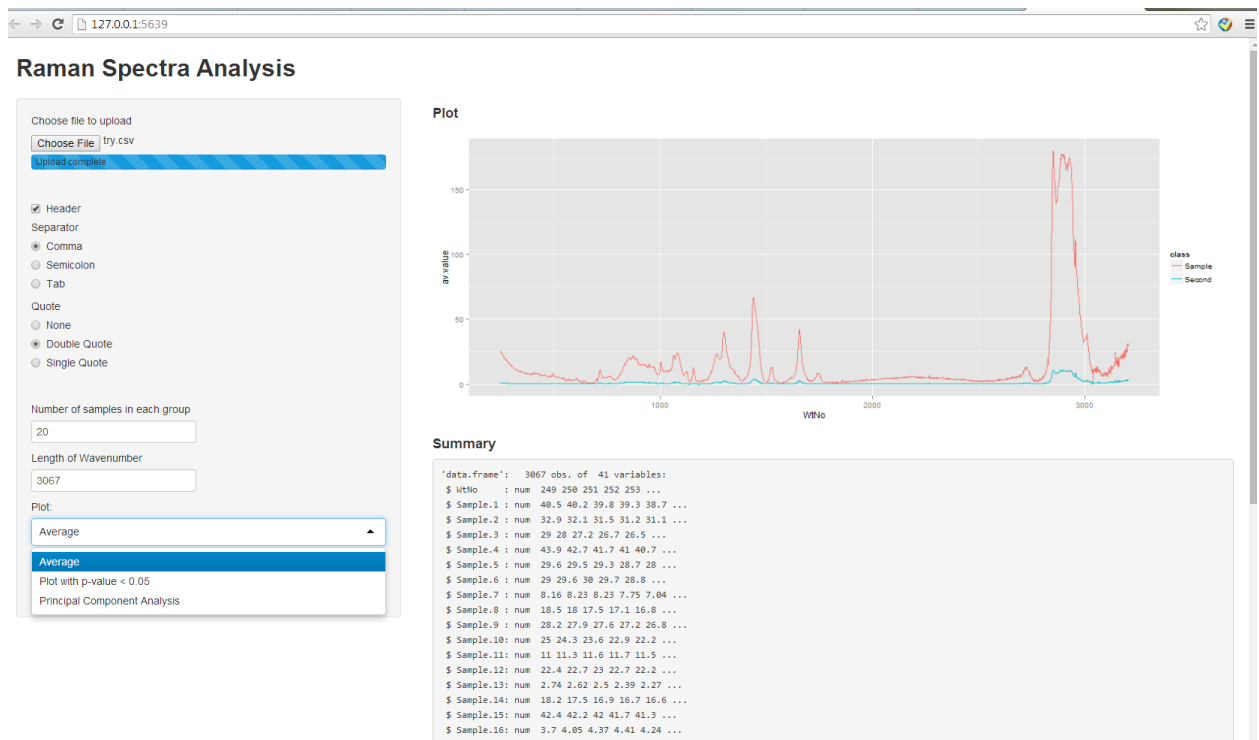


Figure 9: The distribution of BSE cases in Great Britain from 2013 to 2014

Even though BSE is no longer active in the UK, there are still cases appeared in recent years. Since prion disease is hard to diagnose and it spread easily through food contamination, attention is still required to monitor the development of BSE.

6 Shiny App for Raman analysis

```
library(shiny) runGist(11195851)
```



The screenshot above is the shiny app I built for Raman spectra analysis. Data is uploaded as required format and based on user's preference, the desired plot will be displayed in the main panel. Additionally, the result can be downloaded from the button on the left. However, uploaded data need to be corrected by polynomial baseline correction, smoothing and normalization.

7 Mcow package

The packages I built for this project is consisted of 4 functions: `data`, `extractPolygons` (the function in class), `percent` and `percentage_reduction`. *This is not quite complete and some modifications need to be made.*

`date`: This function to reformat the date in the given dataset
`extractPolygons`: The function to extract polygons
`percent`: The function to format data into percentage format
`percentage_reduction`: *This function to calculate percentage reduction*

8 Discussion and Conclusion

From this analysis, BSE cases have been decreasing dramatically during the past decade, however, there are still BSE existing and more prion diseases like scrapie and CJDs appear, which threaten human and animal health. Since prion caused disease is very hard to diagnose, with the application of Raman Spectroscopy, people can detect the protein transformation at a earlier stage, which is beneficial for disease control, as well leading to a better medical treatment.