

## Art of Statistics: Figure 9-2 (page 235) Funnel plot of bowel cancer rates

The data are discussed in a blog by Paul Barden on the Understanding Uncertainty site and on his own blog. Note that data for Wales is not included, as it was reported for the whole of Wales rather than by region.

```
library(ggplot2)
```

```
bowel.data<-read.csv("../data/09-2-bowel-cancer-data.csv",header=T)
summary(bowel.data)
```

```
##      Country      District      n      d
## Length:379      Length:379      Min.   : 6.0      Min.   : 31332
## Class :character Class :character 1st Qu.: 24.0      1st Qu.: 140110
## Mode  :character Mode  :character Median : 32.0      Median : 189202
##                                     Mean  : 39.6      Mean  : 224700
##                                     3rd Qu.: 45.0      3rd Qu.: 267794
##                                     Max.   :251.0      Max.   :1268959
```

```
attach(bowel.data)
mean.prop=sum(n)/sum(d)
props=n/d
max.props=max(props)
```

```
# try funnelR package
library(funnelR)
```

```
# Numerator must be called n, denominator d
funnel_limits <- fundata(input=bowel.data,benchmark=mean.prop, alpha=0.95, alpha2=0.998, method='approx')
```

```
funnel_plot <- funplot(input=bowel.data, fundata=funnel_limits)
```

```
funnel_plot = funnel_plot + coord_cartesian(ylim = c(0,max.props) )
```

```
## Coordinate system already present. Adding new coordinate system, which will
## replace the existing one.
```

```
#funnel_plot = funnel_plot + geom_hline(yintercept=mean.prop, colour="darkred", linetype=6, size=1)
funnel_plot = funnel_plot + scale_x_continuous(name="Population (100,000's)", breaks=100000*(0:14), labels=100000*(0:14))
funnel_plot = funnel_plot + scale_y_continuous(name="Annual bowel cancer mortality rate per 100,000", breaks=c(0,10,20,30,40,50,60,70,80,90,100,110,120,130,140,150,160,170,180,190,200,210,220,230,240,250))
```

```
## Scale for y is already present.
```

```
## Adding another scale for y, which will replace the existing scale.
```

```
glasgow <- subset(bowel.data, District == "Glasgow City") # identify Glasgow City in data frame
```

```
#funnel_plot = funnel_plot + geom_text(data=glasgow, label="Glasgow City", vjust=1)
```

```
funnel_plot = funnel_plot + annotate("text", x=glasgow$d,y=glasgow$n/glasgow$d,label="Glasgow City",hjust=0.5)
funnel_plot
```

```
## Warning: Use of `fundata$d` is discouraged.
## i Use `d` instead.

## Warning: Use of `fundata$up` is discouraged.
## i Use `up` instead.

## Warning: Use of `fundata$d` is discouraged.
## i Use `d` instead.

## Warning: Use of `fundata$lo` is discouraged.
## i Use `lo` instead.

## Warning: Use of `fundata$d` is discouraged.
## i Use `d` instead.

## Warning: Use of `fundata$up2` is discouraged.
## i Use `up2` instead.

## Warning: Use of `fundata$d` is discouraged.
## i Use `d` instead.

## Warning: Use of `fundata$lo2` is discouraged.
## i Use `lo2` instead.

## Warning: Use of `fundata$benchmark` is discouraged.
## i Use `benchmark` instead.

## Warning: Use of `input$d` is discouraged.
## i Use `d` instead.

## Warning: Use of `input$n` is discouraged.
## i Use `n` instead.

## Warning: Use of `input$d` is discouraged.
## i Use `d` instead.

## Warning: Removed 374 rows containing missing values or values outside the scale range
## (`geom_line()`).

## Warning: Removed 218 rows containing missing values or values outside the scale range
## (`geom_line()`).

## Warning: Removed 928 rows containing missing values or values outside the scale range
## (`geom_line()`).

## Warning: Removed 542 rows containing missing values or values outside the scale range
## (`geom_line()`).
```

