

Dull, damned dull, and statistics. Discuss.

David Spiegelhalter
Statslab and MRC Biostatistics Unit
Cambridge

IMA, June 20th 2008

With thanks to Mike Pearson, Arciris Garay, Hauke Riesch, Owen Smith, Hayley Jones etc

So what's the point of probability and statistics?

- Spotting patterns in data
- Trying to explain those patterns
- Trying to separate chance from knowledge (distinguish *noise* from *signal*)

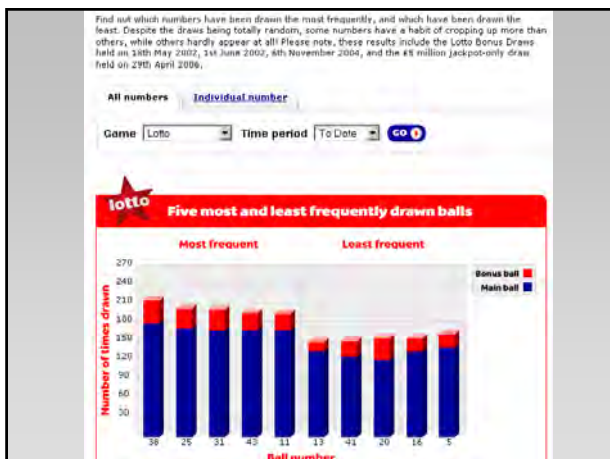
$$f_\delta(x) = \begin{cases} 1 - \gamma_\delta(\delta), & \text{if } x = 0 \\ \frac{x - x_0(\delta)}{\sqrt{r^2(\delta) - (x - x_0(\delta))^2}} \exp\left[\gamma_\delta(\delta) - \sqrt{r^2(\delta) - (x - x_0(\delta))^2}\right], & \text{if } 0 < x \leq x'(\delta) \\ \gamma_\delta(\delta)e^{-x^2}, & \text{otherwise,} \end{cases} \quad (3)$$

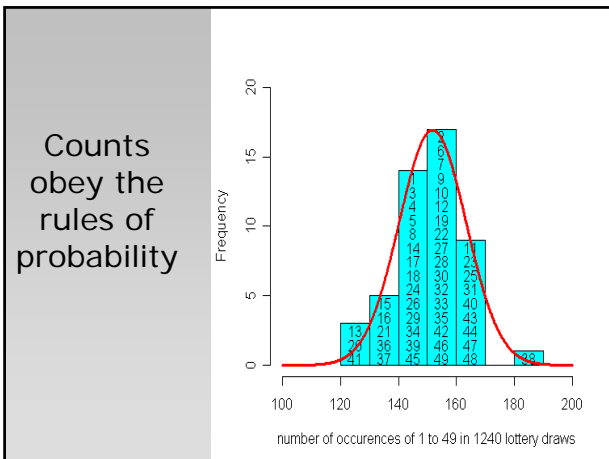
where x' is suitably large. Approximate functional forms for γ_δ , γ and x' are empirically derived and can be described by

$$\begin{aligned} \tilde{\gamma}_0(\delta) &= e^{-0.651(\delta-0.277) + 0.031\delta} - 0.189 \\ \tilde{\gamma}(\delta) &= e^{-0.578(\delta+0.024) + 0.006\delta} \\ \tilde{x}'(\delta) &= 0.170\delta^2 + 1.052\delta - 0.02. \end{aligned} \quad (4)$$

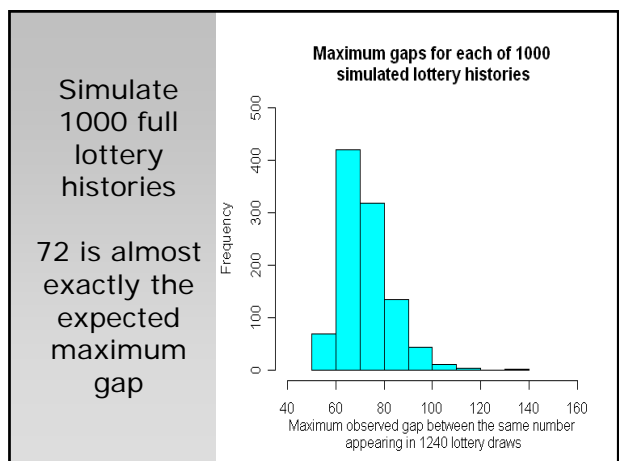
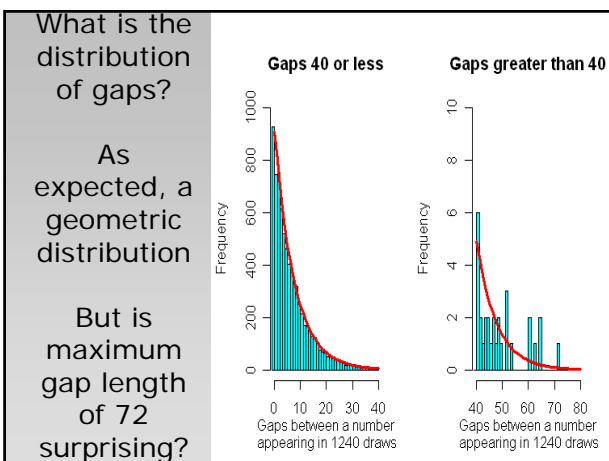
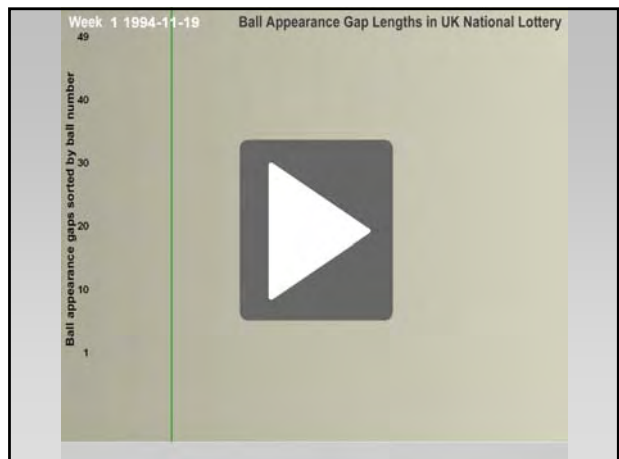
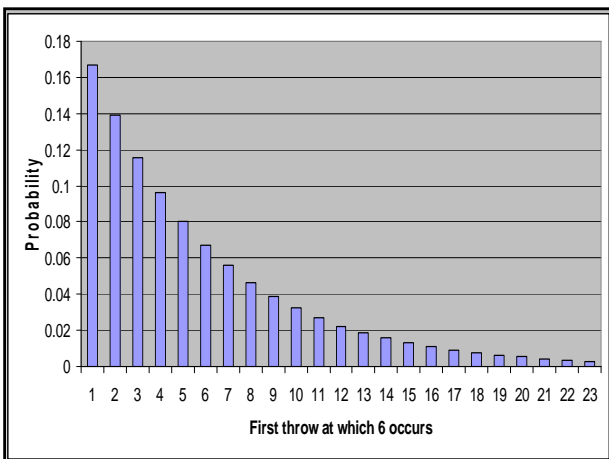
The other functions, r , y_0 and x_0 , which are respectively the radius and origin coordinates of a circle segment (see Figure 2), can be calculated from γ_0 , γ and x' as follows. Let $\rho_0 = \log(\gamma_0)$, $\rho = \log(\gamma) - \rho_0$, $u = \sqrt{x'^2 + (\rho - x'^2)^2}$, $v = \sqrt{2x' - \rho}/\sqrt{2}$ and $\theta = \sin^{-1}(v/u)$, then we have that

$$\begin{aligned} r &= \frac{u}{2\sin\left(\frac{\pi}{2} - \theta\right)} \\ x_0 &= -r \sin\left(2\theta - \frac{3\pi}{4}\right) \\ y_0 &= \rho_0 - r \cos\left(2\theta - \frac{3\pi}{4}\right) \end{aligned} \quad (5)$$





- How long before you throw your first six?

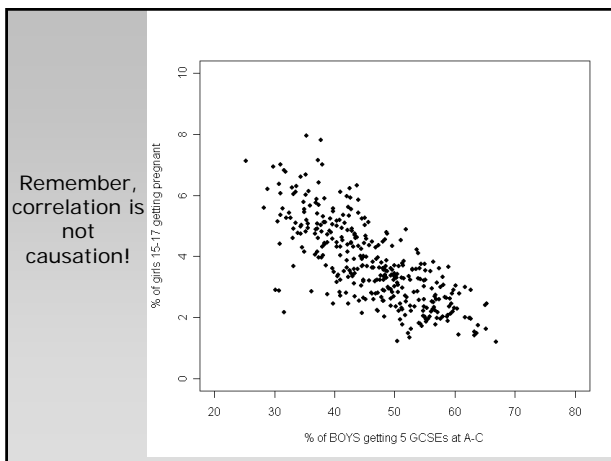
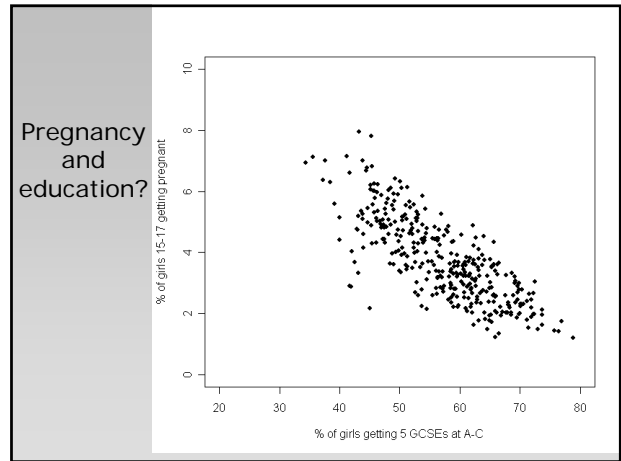
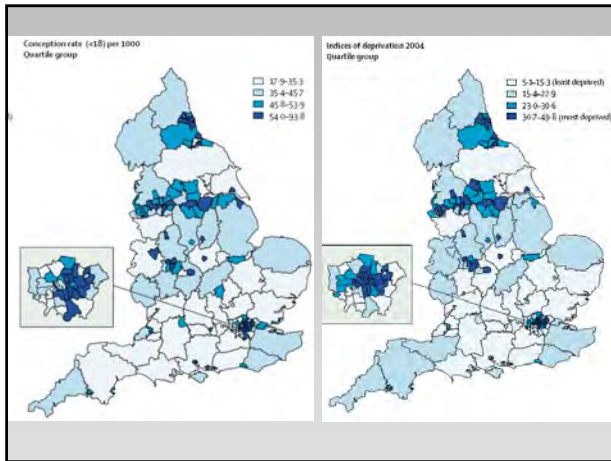


Teenage pregnancy in the East of England

Region	Local authority	Number of pregnancies: females 15-17 2005
East	Peterborough	184
East	Luton	161
East	Basildon	155
East	Norwich	150
East	Southend-on-Sea	136
East	Thurrock	129
East	Ipswich	116
East	Colchester	113
.....		
East	Babergh	35
East	Mid Suffolk	30
East	Maldon	29
East	Forest Heath	27
East	Three Rivers	26
East	Brentwood	20
East	Uttlesford	20

Express as rate per 1000 females aged 15-17 in area

Region	LA_Name	Pregnancies per 1000 females aged 15-17
East	Norwich	72
East	Peterborough	57
East	Great Yarmouth	56
East	Ipswich	49
East	Southend-on-Sea	47
East	Basildon	47
East	Thurrock	44
East	Fenland	43
.....		
East	St Albans	20
East	East Hertfordshire	19
East	South Norfolk	18
East	Mid Suffolk	18
East	Brentwood	15
East	Three Rivers	15
East	Uttlesford	14



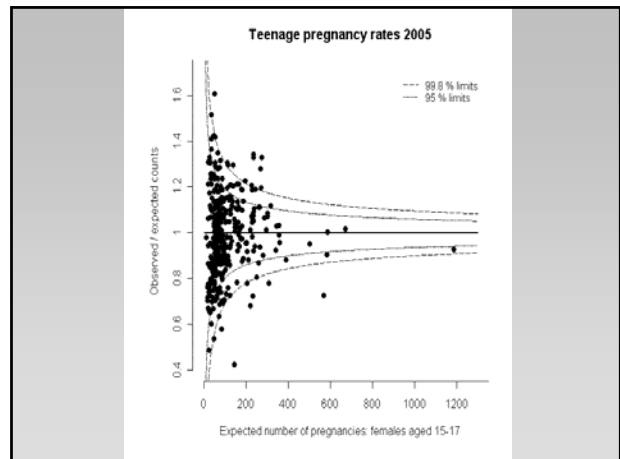
How to fairly compare pregnancy rates?

- Allow for differing environment
- Can calculate **Expected** number of pregnancies, given deprivation, education attainment, urban/rural etc of area
- Then compare **Observed** with **Expected** rates

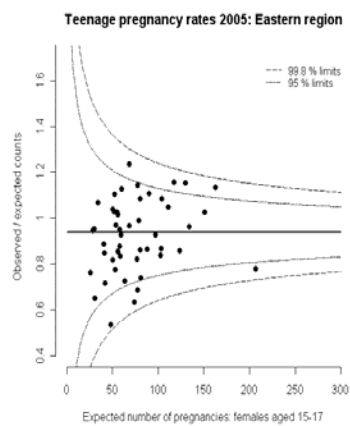
Teenage pregnancy in the East of England, ratio of Observed to Expected

Breckland	1.23
Southend-on-Sea	1.15
Norwich	1.15
Huntingdonshire	1.14
Peterborough	1.13
Fenland	1.13
Great Yarmouth	1.10
St Edmundsbury	1.10
.....	
Cambridge	0.74
North Hertfordshire	0.72
Mid Suffolk	0.71
Epping Forest	0.68
Brentwood	0.65
Stevenage	0.63
Three Rivers	0.53

But are these 'real' differences ?



When we adjust for 'risk factors', and allow for chance, strong differences between areas disappear

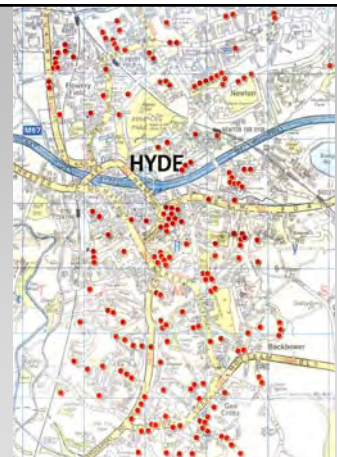


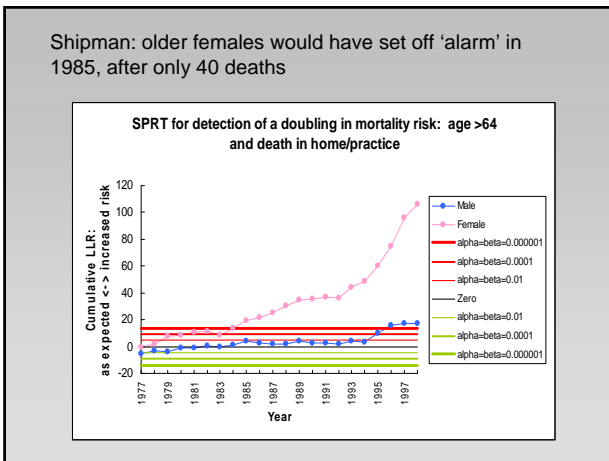
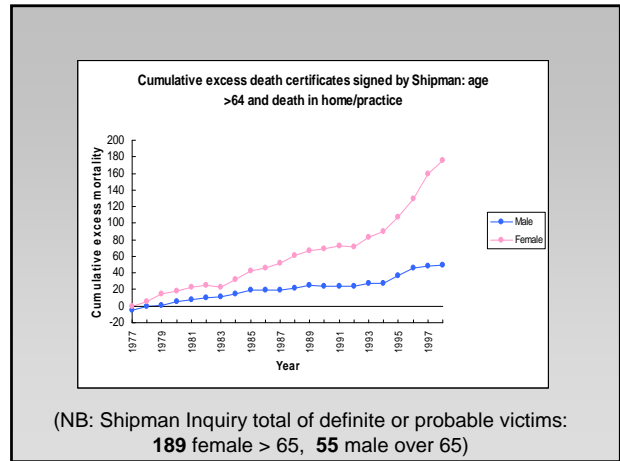
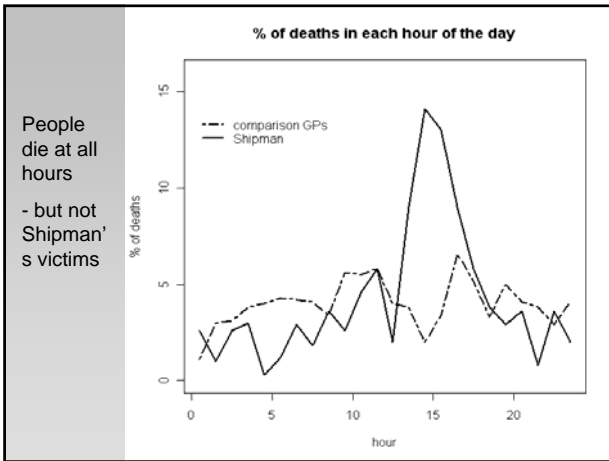
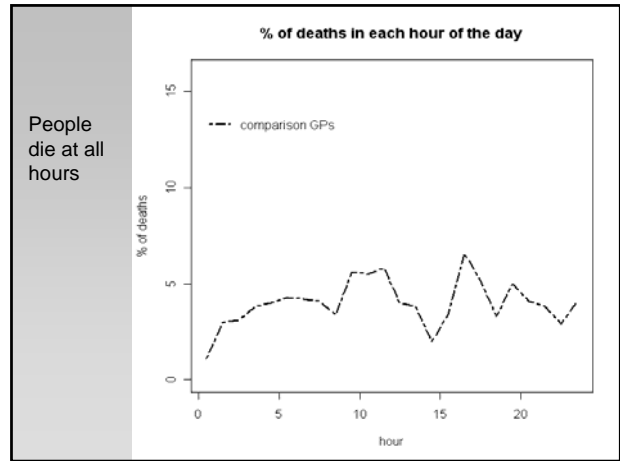
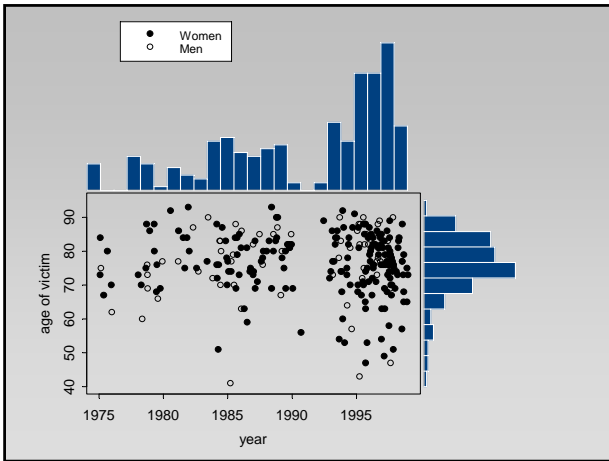
'I have nothing to hide'

Dr Harold Shipman, general practitioner, on his arrest in September 1998

Shipman Inquiry
July 2002:

215 definite victims,
45 probable





What is our programme trying to do?

- Educational lectures and workshops
- Website
- Engagement with media
- Working with people who want to communicate risk
- Inter-disciplinary research

www.understandinguncertainty.org

The screenshot shows the homepage of 'Understanding Uncertainty'. At the top, there is a navigation menu with links for Home, Blog, For Journalists, For Educators, Flat Stuff, and About Us. Below the navigation, there is a search bar and a 'Digg' button. The main content area features a section titled 'Premier League' with a sub-header 'View | Edit | Outline | Revisions | Workflow | Clone'. The text below discusses the Premier League's structure, mentioning 20 teams, 38 matches per team, and the league's ranking system. A sidebar on the right contains 'Featured Content' and 'Pending Content' sections with various links.

The great bacon sandwich debate



The great bacon sandwich debate

- Estimated 20% excess risk of bowel cancer per 50g of processed meat per day
- WCRF Website:
"Try to avoid processed meats such as bacon, ham, salami, corned beef and some sausages"



Press coverage

- Sun: *"Careless pork costs lives"*
- Daily Mail: *"Is anything safe to eat?"*
- Sun Health columnist Dr Keith Hopcroft:
"I'd rather shave a few years off my life with the occasional bacon sarnie than be 110 and dribbling into my All Bran"

The image shows a newspaper clipping from 'Sun Health'. The headline reads 'A bit of sausage won't kill you'. The article is titled 'CANCER SCARE: THE FACTS' and discusses the risks of processed meat. A large image of a sausage on a fork is featured prominently. The text on the right side of the clipping reads: 'warns is dangerous. Prof Spiegelhalter says if you eat a lot of processed meat, or there is a family history of bowel cancer, you may want to cut back. But you may decide the risk is relatively small, and continue to tuck in — as the Professor says he will.'

How long are you going to live?

Microsoft Excel - 13.11.2014 10:49:24

File Edit View Insert Format Tools Data Window Help

Office for National Statistics

Interim Life Tables, United Kingdom

Interim Life Tables, United Kingdom [Back to contents](#)

Period: expectation of life
Based on data for the year 2014-2016

Age	Males					Females				
	q_x	m_x	e_x	l_x	d_x	q_x	m_x	e_x	l_x	d_x
0	0.000592	0.000376	100000.0	987.6	79.89	0.000494	0.000454	100000.0	405.4	81.28
1	0.000409	0.000267	99442.4	405.5	75.32	0.000384	0.000384	99544.6	28.2	80.64
2	0.000242	0.000162	99407.8	24.0	75.36	0.000182	0.000186	99506.4	18.5	79.87
3	0.000184	0.000119	99377.9	18.3	74.37	0.000147	0.000147	99407.8	14.8	79.88
4	0.000127	0.000077	99358.5	12.6	73.38	0.000111	0.000117	99472.1	11.7	77.88
5	0.000114	0.000114	99346.9	11.3	72.40	0.000097	0.000097	99461.6	9.7	76.70
6	0.000123	0.000123	99335.6	12.0	71.40	0.000118	0.000110	99451.9	11.6	75.71
7	0.000095	0.000095	99323.0	9.4	70.41	0.000094	0.000094	99440.9	8.3	74.72
8	0.000114	0.000114	99314.2	11.3	69.42	0.000087	0.000087	99432.6	8.6	73.72
9	0.000119	0.000118	99302.0	11.8	68.43	0.000089	0.000089	99424.0	8.9	72.72
10	0.000120	0.000120	99291.1	11.9	67.44	0.000084	0.000084	99417.1	8.3	71.73
11	0.000126	0.000126	99279.2	12.5	66.44	0.000091	0.000091	99408.6	9.0	70.74
12	0.000147	0.000147	99266.6	14.6	65.45	0.000119	0.000119	99399.7	11.8	69.75
13	0.000170	0.000170	99252.0	16.9	64.46	0.000127	0.000127	99387.9	12.8	68.76
14	0.000188	0.000188	99235.2	19.2	63.47	0.000152	0.000152	99375.3	15.1	67.76
15	0.000257	0.000256	99216.0	25.4	62.48	0.000164	0.000164	99362.1	18.3	66.77
16	0.000309	0.000309	99195.3	33.8	61.49	0.000219	0.000219	99346.9	21.7	65.78
17	0.000396	0.000395	99176.9	55.1	60.52	0.000281	0.000281	99324.1	24.0	64.80
18	0.000466	0.000466	99151.9	63.4	59.55	0.000271	0.000271	99303.2	26.8	63.81
19	0.000567	0.000567	99128.5	66.0	58.58	0.000291	0.000290	99273.2	28.9	62.82
20	0.000720	0.000720	99097.4	71.9	57.63	0.000384	0.000384	99244.3	29.2	61.83
21	0.000872	0.000872	99059.2	66.5	56.67	0.000506	0.000506	99216.3	27.8	60.84
22	0.000976	0.000949	99014.7	74.1	55.71	0.000529	0.000529	99188.5	27.9	59.85
23	0.001011	0.000911	98963.6	80.1	54.75	0.000598	0.000598	99160.7	29.6	58.86
24	0.000979	0.000738	98908.6	72.9	53.80	0.000529	0.000529	99131.0	27.8	57.87
25	0.000924	0.000621	98857.7	65.2	52.83	0.000463	0.000463	99102.2	26.4	56.88
26	0.000778	0.000578	98792.5	76.7	51.88	0.000342	0.000342	99072.8	33.9	55.89
27	0.000664	0.000464	98716.9	84.1	50.92	0.000264	0.000264	99043.6	36.1	54.90
28	0.000578	0.000378	98635.0	75.3	49.96	0.000206	0.000206	99013.0	38.2	53.91
29	0.000488	0.000288	98549.2	67.8	49.00	0.000162	0.000162	98981.6	37.6	52.92
30	0.000348	0.000147	98420.4	30.0	48.04	0.000124	0.000124	98949.7	42.9	51.93
31	0.000268	0.000095	98316.2	96.7	47.08	0.000075	0.000075	98918.4	42.0	51.94

Office for National Statistics

Home About Us Contact Us

