Amanda Cox

Amanda Cox won the American Statistical Association's award for excellence in statistical reporting this year. But she denies that she is a statistician, or that what she does is statistics. Julian Champkin talked to her.

Most people who take a statistics degree end up doing – well, statistics. Deep sea diving, or hotel management, or driving a train are not usual career paths for them. Nor are graphic design, or journalism. Meet Amanda Cox. She did not become a statistician; after her masters degree she became a graphic designer and a journalist.

Not a statistician, then? "I don't deal with uncertainty in a formal enough way to call what I do statistics, or myself a statistician." That is what she recently told the website Simply Statistics (`http://bit.ly/SH0biI`). It cannot be said that she has failed statistics in any way, though; earlier this year she won the American Statistical Association's award for

excellence in statistical reporting. She has also won the top award at Malofiej, which you may not have heard of but which newspaper graphic designers have: it is the largest and most prestigious international infographics contest. Graphic artists as well as statisticians love what she does.

Amanda works at the *New York Times*. "My technical title is 'graphics editor', but no one knows what this means. On good days, what we do is 'journalism'", she says. She is part of the graphics team there; and the newspaper's ways of presenting data to its readers in charts, diagrams and pictorial forms generally have won near-universal acclaim. Its graphics are clear, they are imaginative, they are innovative; and if it matters, they frequently tend to be beautiful as well.

Amanda's graphics have been praised by no less a statistical graphic authority than Edward Tufte: "A superb piece of work that provides enormous historical context, describes rich variation not just a recent average, combines words and graphics, uses annotation to call out important points, and contextualizes recent changes in market volatility." That is what he said of one of her pieces of work, called 'The Pulse of Uncertainty', which can be found at `http://nyti.ms/dzKBop`. And Tufte continued: "It integrates traditional news reporting with high-resolution (sparkline-like) graphics, and makes no distinction among words, numbers, graphics – the idea is *whatever it takes to explain something*." All that was about a timeline of daily changes in the Dow Jones average – which is not a set of data that sets the world on fire or that is usually considered to call for anything but bog-standard graphing.

Better perhaps to call them visualisations rather than strictly graphics. A fair number are interactive; her work appears on the *New York Times* website as well as in print. "It's about a 50–50 split", she says. From her home in Brooklyn Heights she travels daily by subway to the newspaper's office where I caught her by phone before the paper's morning conference ended and everyone started rushing about to research, write and draw

Graphics are coming of age as a way of communicating data. Amanda Cox is an expert

the next day's pages. Some graphics are created at speed – daily newspapers are frenetic places. Sometimes she gets a bit more notice. But "every project expands to fill all the time that you can get for it", she says.

What led her from the calm and well-ordered discipline (I assume) of the stats department to the rather more bustling and urgent atmosphere of the newsroom? What makes her graphics so good at conveying information to the reader over his or her breakfast, given also that the said reader may know nothing, and think that he or she cares less, about the topic concerned, let alone about the data that lies behind the topic?

Graphics convey quantitative information, but they are works of art. It helps if they are beautiful. Was her background artistic, then, or in numbers? "My parents were accountants. So it was a strictly numbers background rather than artistic. In middle school I was more math inclined. There was a bit of interest in art, but nothing you could call a passion." The art came through happy chance. "In college I graduated a little early, so I had to take an extra half-credit course most semesters. The only half-credits the school offered were in art or in gym. Racquetball and canoeing were fine, but wood sculpture was amazing…." Correlation may not be causation, but so many careers are shaped by semi-accidents like that.

Born in 1980 in Michigan, she took her college degree in maths and economics and her MS in statistics from the University of Washington. "I dropped out of my PhD programme in the second year. After the first year, I wasn't sure what I was doing. My brain wasn't big enough for measure theory. So I applied for random things" – among them, an internship at the *New York Times*. "I returned to grad school, had a better time, and in 2005 joined the *New York Times* more permanently. Journalism is fun. The role is to be curious about things, ask interesting questions. It's a good mandate." It is the same mandate that statisticians have.

"It is nice here – a pretty small community of people." There are around 25 in the graphics department – but, unusually for a newspaper, the graphics desk counts as a full department on a par with sports, metro, or the foreign desks. "So although lots of what we do is making visualisations of data that reporters bring us on stories that other departments have originated, my boss also attends the daily morning editorial meeting and puts up our own ideas for graphics that we should do. So

some topics we decide, some are decided for us."

The job is about communicating to non-statisticians, non-economists, non-specialists – to people who are eating their breakfast or who are on the train on their way to work. "We take our mission as 'clear and compelling' communication." How clear? Her charts, and her department's charts, are original, non-formulaic. How much work does she expect her readers to put into reading one of them? "It depends on the graphic. There is a school that says 'every graphic should be immediate: if you cannot understand it at a glance then you haven't got it right'. I don't necessarily subscribe to that. We can ask for some effort if the story demands it. I'd say we should make it 'as immediately obvious *as possible*' with the last two words emphasised. There is room for subtlety."

Above all, there is room for creativity. How does one help the creative process along? "It is sketching with data. Trying out different forms of presenting it and seeing what happens. A lot of the sketches you throw away because they don't work."

"What I have to do is add an explanatory layer on top of the data that is there. It adds interest, it adds relevance, it gives perspective. It is a trick we use a lot." It is more easily done with printed visualisations than with interactive ones; and the two media demand different treatments.

"On-line, you can keep much more of the data, layer upon layer of it, and ask the user to dig down to find it. Print enforces space constraints. You have to concentrate on the data that is actually important. You throw away a greater proportion of the information; on screen, you can keep it."

"The first thing anyone does with a graphic is to look to see where they are on it; so putting a 'You are here' spot on a graphic helps you orient yourself, helps you see the bigger picture, and where you yourself are within in it. Because, for all of us, our own personal experience is quite limited. Our social circles are limited too. Background helps to widen our perspective, to see ourselves in a wider context, of the next neighbourhood to ours, of the kid who has just left school, of the mainly-black neighbourhood a few blocks north from where we live."

As with an interactive graphic at `nyti.ms/1ndlhL`. If you click on your own grey line on it you can see the unemployment among white middle-class professional males.

(That, I suppose, is the average demographic of statisticians, though I would love to be proved wrong.) Click on a different grey line and you get the jobless rate among black men with a high-school degree, or among Hispanic women aged 15–24, or among almost any ethnic, educational and age mix that you can think of – and it is very obvious that, however tough things may be for the first group of people, they are an awful lot tougher for the second and third. And that also makes it personal, and relevant to every reader.

Which last two are key. "We deal in moderately abstract concepts that are not really abstract at all. Like inflation," she says. "*Shopping* makes a lot more sense to me, and to most people. Connect it to something I am familiar with. So illustrate a story about inflation with data about the prices of things that people buy: bread, dairy, seasonal vegetables." As she did with "All of Inflation's Little Parts", reproduced in part in Figure 1. (The whole graphic is rather bigger; the *New York Times* has bigger pages than we do.) One concrete example is worth a dozen concepts.

"Graphics provide a context", she says. The unemployment graphic helped relate the situation in your neighbourhood to one maybe the wrong side of the tracks. That is personal context, or social. There are others. "'Million' and 'billion' mean much the same to me, and I suspect to very many readers. Very often they are just shorthand for 'a lot'. Part of our job is bringing the very small or the very large onto a scale that means something. A bar chart of month-by-month casualties of the Iraq war tells you little of the scale of it; put it beside one of the Vietnam war and you get a comparision; add one of the Second World War, and the context is clearer still."

Adding the background she has compared to music – a main voice, with background singers adding the harmonies and richness. You could call it making the data sing. She has done that literally as well. "Really small intervals of time are difficult to understand in the same way as really big numbers are difficult to understand. Two-tenths of a second means not very much, but those sorts of times separate the first three places in many sports." Set out in a table they mean nothing. An interactive graphic on speed-skating in the Winter Olympics instead bleeps out the winning intervals – four bleeps in under a second – and you can tell instantly how comfortably or otherwise the winner was in the lead. The amazing thing is

that the skiers – who race separately, possibly hours apart, under very different conditions – are just as close as the speed-skaters.

"And with a little bit more data you can do something smarter. Don't plot just their Olympic times, but their times in each of five previous competitions." For one competitor, that gave four equally-pitched tones then a lower one for the Olympic race – meaning you suddenly see – sorry hear – how that competitor turned in an unexpectedly poor performance. "It is revealing the unexpected, the piece of data that is odd, unusual, that stands out – that's another thing that a good visualisation should do."

"And I love revealing patterns." An electoral graphic, a map, showed vote split in counties across the country; and it revealed a thin blue arc of Democrat voters that stretched from Carolina down through the southern states and then back again up to the Mississippi. "Someone asked why the pattern was there. I didn't know. But a reader wrote online that it reminded him of his own expertise, which was cotton-farming in the 1880s – the blue arc followed the cotton farms. Another reader followed that up again: the arc was areas that were under water millions of years ago in the Cretaceous. So the sea creatures died, affecting the soil, and the soil affected what you could grow, and what you could grow affected the people who lived there, and the people who lived there affected the elections. We were not doing it, readers on the internet were, but revealing relationships, revealing patterns like that – that's fun."

Analysis is all. One of the slides on a presentation she gives reads: "Nothing really important is headlined 'Here is some data. Hope you find something interesting in it.'" Annotation is critical. Editing is critical. Sometimes a graphic just has to be a list of facts – but it is so much better if you can make it tell a story, give it a beginning, a middle and an end. "I do think the traditional news graphic has succeeded if the pieces add up to more than the sum of the individual parts – if it reveals a structure."

An unexpected one was an apparently simple interactive histogram of how states voted in the 2008 primaries (http://nyti.ms/i6clTp). Boxes for each state pile up either side of the middle line to show which way they voted. Again, the interactivity makes it more fun: you can click on different demographics and the boxes move across the screen and rearrange themselves accordingly. Click on "Women" and you find more women in all states voting for Hillary Clinton than for Barack Obama; on "Under age 30" or "Blacks" and the boxes trundle across to the Obama side. All much as expected, until you click on "Post-graduate" – and you find that most of the boxes move to the Obama side but the box for Arkansas goes for a little walk and ends up lonely and all on its own, favouring Clinton by 44 percentage points. "When you think about it, it is obvious", says Amanda. "Lots of people with a post-graduate education in Arkansas are probably connected to the Clintons." It is a no-brainer, of course; but the graphic helped make it a no-brainer. And printing ten versions of the same chart, side by side, one for each demographic, would have been absurd.
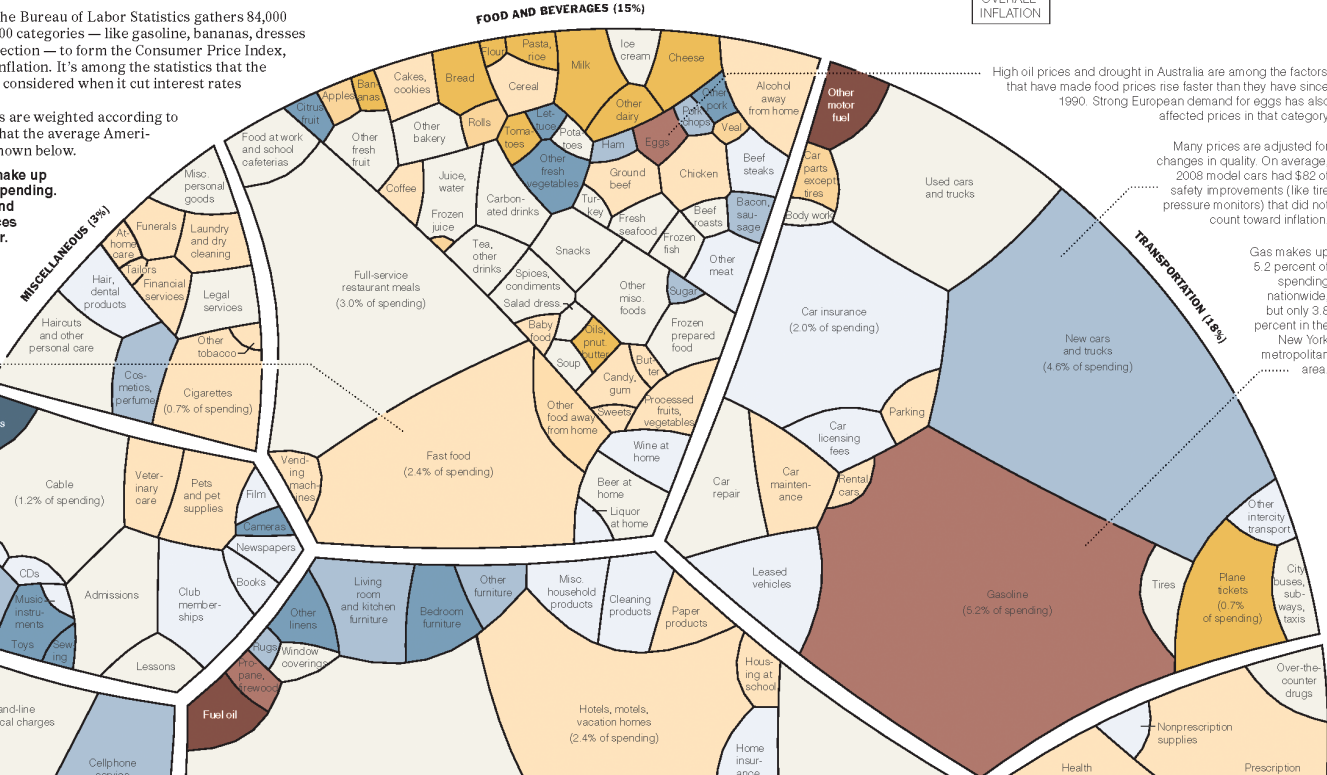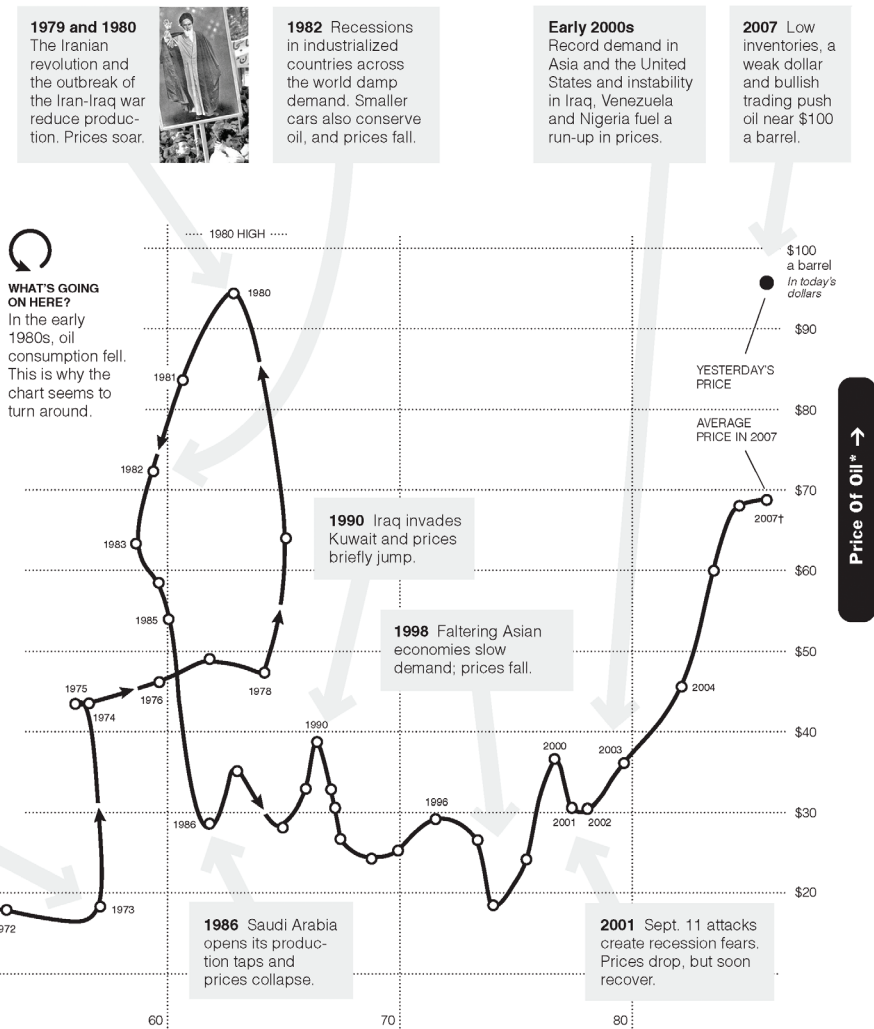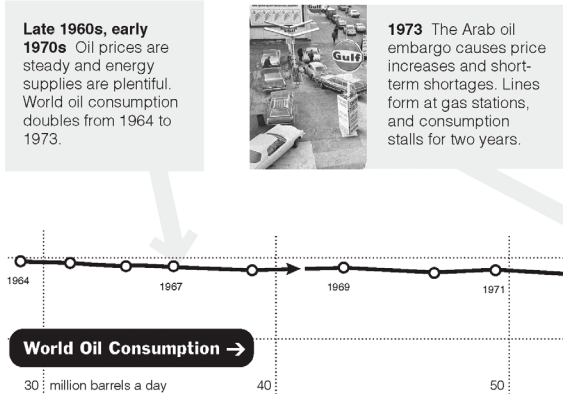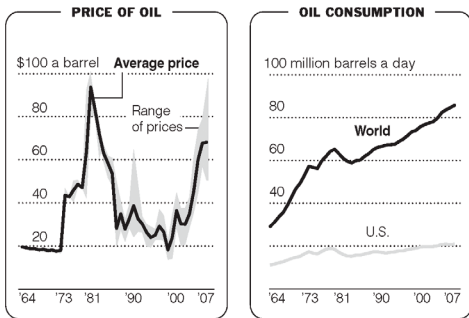

Figure 1. Inflation and shopping: which items you spend most on, and which have gone up the most, captured in a single graphic. Section of an Amanda Cox graphic from the *New York Times*, May 4th 2008. See http://nyti.ms/XUzA for complete view. Reproduced courtesy of *New York Times*

Oil's Roller Coaster Ride

This chart tracks the relationship between oil prices and oil consumption since 1964. Global oil consumption is shown on the horizontal axis and oil prices are shown on the vertical axis. So, when consumption is increasing and prices are flat, the line moves straight right. And when prices are rising and demand stops growing, the line moves straight up.

1979 and 1980 The Iranian revolution and the outbreak of the Iran-Iraq war reduce production. Prices soar.

1982 Recessions in industrialized countries across the world damp demand. Smaller cars also conserve oil, and prices fall.

Early 2000s Record demand in Asia and the United States and instability in Iraq, Venezuela and Nigeria fuel a run-up in prices.

2007 Low inventories, a weak dollar and bullish trading push oil near $100 a barrel.

PRICE OF OIL

$100 a barrel   **Average price**

Range of prices

OIL CONSUMPTION

100 million barrels a day

World

U.S.

WHAT'S GOING ON HERE? In the early 1980s, oil consumption fell. This is why the chart seems to turn around.

1980 HIGH

$100 a barrel
In today's dollars

$90

YESTERDAY'S PRICE

AVERAGE PRICE IN 2007

$80

Price Of Oil* →

1990 Iraq invades Kuwait and prices briefly jump.

1998 Faltering Asian economies slow demand; prices fall.

Late 1960s, early 1970s Oil prices are steady and energy supplies are plentiful. World oil consumption doubles from 1964 to 1973.

1973 The Arab oil embargo causes price increases and short-term shortages. Lines form at gas stations, and consumption stalls for two years.

1986 Saudi Arabia opens its production taps and prices collapse.

2001 Sept. 11 attacks create recession fears. Prices drop, but soon recover.

World Oil Consumption →

30 million barrels a day

* Average annual price of West Texas Intermediate crude oil, adjusted for inflation using the Consumer Price Index. Posted prices (not spot prices) are shown before 1983.

† Consumption forecast as of Nov. 6

Sources: Energy Information Administration; Federal Reserve; Bureau of Labor Statistics; Rocky Mountain Institute

AMANDA COX/THE NEW YORK TIMES; PHOTOGRAPHS BY THE ASSOCIATED PRESS

Figure 2. A graphic that repays study. At first sight it is a time-line – but time cannot flow backwards. The x-axis is actually world oil consumption; time flows along the black wiggly line. *New York Times*, November 9th 2007. Reproduced courtesy of *New York Times*

The traditional news graphic, used in every newspaper and by every statistician, is the bar chart. "Everything is a bar chart. I hate bar charts because I can make a bar chart out of anything. It is a bit of a lazy way out. Sometimes bar charts don't respect what is unique about the data. Every set of data is special, and there is a way of bringing that specialness out. Bar charts do not find or respect that. I feel much the same about timelines. Certain things I can create a really good timeline for, others will remain mediocre whatever you do. Take the same data and do something different with it.

"Generic solutions are rarely awesome. There is something non-generic to the work that we do. Our assumption is that the audience does not necessarily care about the data that goes into the graphic. That's different from

an academic paper, where you assume that the academic who is reading it does care, and cares enough to tease out what an uninspiring diagram means. For newspapers that assumption has to be a little more relaxed – push it towards the awesome side. It is still about respecting the data, but respecting the audience as well."

Has she three simple hints for beginners – easy ways to make diagrams better? "A goal of graphing, I think, is to see things that you wouldn't expect. Try out different solutions. Throw away a lot of ideas. You have to get to a point where you are fluent enough in something that trying out lots of different ideas isn't a barrier."

Do not despair: the other two hints are simpler. "What distinguishes a lot of our graphics from, say, business graphs is where the text goes. Incorporate it within the graphic,

even if it is brief and explained somewhere else in the text. At the point where something interesting happens, write it in."

More practically still, simpler still, and my favourite: "Put a verb in your headline. Even 'is' will do." If a picture is worth a thousand words, one word can help out the picture. Just help the poor reader with a sentence that makes sense. As an editor who has wrestled with and tried to make sense of many a verbless submission, can I add my plea to hers?

It is 10.30, and the newsmen are about to come out of morning conference with the day's lead stories. She will have to find ways to make them clear to 1½ million readers, and has not too many hours to do it. So I ring off. Before I do, one last question:

"My favourite graphic? I don't know. That is a work in progress. It is still to come."