Writing for Significance

Mario Cortina Borja
University College London

Brian Tarran
Royal Statistical Society

February, 2020

Outline

A brief history of Significance

- Significance is a bimonthly magazine for anyone interested in statistics and the analysis and interpretation of data
- Its aim is to communicate and demonstrate in an entertaining, thought–provoking and non–technical way the practical use of statistics in all walks of life and to show informatively and authoritatively how statistics benefit society

- An official magazine of the Royal Statistical Society (RSS) and the American Statistical Association (ASA)
- Published by Wiley–Blackwell
- It's **not** a peer–reviewed journal
- Established in 2004; originally, published quarterly; from 2011 onwards published bimonthly
- It has a website, `https://www.significancemagazine.com/` since 2010 – papers can be published there and/or in the printed version
- Started as a magazine of the RSS in 2004; became an official magazine of the ASA in 2011
- Articles can be accessed on line free of charge one year after publication

- Editors: Helen Joyce, Julian Champkin, Brian Tarran
- Editorial board: currently 10 members from RSS (7 UK–based, one each in Canada, Mexico and South Africa), and 10 from ASA (9 US–based, 1 in China)
- It's distributed to all ASA and RSS members, 26,000 copies
- Website `https://www.significancemagazine.com/` is widely read

- Articles are largely non–technical and hence accessible and appealing, not only to members of the profession, but also to all users of statistics
- As well as promoting the discipline and covering topics of professional relevance, Significance contains a mixture of statistics in the news, case–studies, reviews of existing and newly developing areas of statistics, the application of techniques in practice and problem solving, all with an international flavour

Why?

- Significance is a showcase for statistics: it's written by statisticians for anyone interested in the analysis and interpretation of data
- It's about busting myths, answering questions, supporting decisions
- It's not peer-reviewed but publishing in it should help you achieve wider impact/engagement with your work

- Interesting, engaging and easy to read articles
- A strong story with sound statistics and fault–free analysis

- Ask yourself: "What's the best way to showcase the power of statistics?"
- You might want to: bust apart a popular myth,answer a burning question, provide evidence to support decision–making
- When thinking about a topic for your article, remember: you can write about anything – as long as it has some connection with statistics
- Make sure it's a topic you're interested in, personally and professionally
- If you're not interested in what you're writing about, your audience definitely won't be

- Articles can be on any subject – we have published articles about statistical graphics, plastic pollution, birthday problems, Rugby world cup, Ryanair's seating allocation policy, fake news, marathon records, James Joyce's novels, the director of analytics for the Miami Dolphins football team, Florence Nightingale, UK Prime Ministers' ages...
- Important thing is to get the style right – **it's not a journal**
- Articles should be written in a journalistic, magazine style

- Don't use jargon – explain ideas and concepts, and use analogies to help illustrate things that might be a little too abstract in isolation
- A strong conclusion is needed to give people something to take away and think about
- Never forget – you are telling a story; you're not just presenting your work
- Never lose focus of what you're writing about

- Keep the number of figures, tables, and references to a minimum
- Avoid equations, unless it's essential, e.g. the article is about an equation!
- You may use data sources and references which you should not normally use in theses or peer–reviewed articles, e.g. Wikipedia or websites
- Think of photos that may accompany the article

- Ask family or friends to review your work – identify what needs explaining or simplifying
- Try to answer any questions they have in a second draft, so that when submitted to our editorial board those same questions aren't coming up
- After editorial board review, make sure you address all comments made by reviewers – even if it is to disagree with a suggestion that's been made
- They prefer to know you've read the comments and considered them, even if they aren't all taken on board

- Your paper will be reviewed (but **not** peer–reviewed) by one or two members of the editorial board, and the editor
- This process is similar to what you go through in a journal though the reviewers understand the difference between peer–review and and not–peer–review – and so should you
- The editor may reject your paper before sending it to members of the editorial board, exactly as it happens in journals

- Too much technical detail/mathematics
- Lack of story and story structure
- Lack of wider relevance – you need to relate the importance of your work to non–experts, which means that some of the more theoretical areas of statistics are not really appropriate for us

- Send it somewhere else!
- There are magazines with a more mathematical slant, e.g. Chalkdust (UK), Laberintos e Infinitos (Mexico), Mathematics Magazine (USA), Mathematical Gazette (UK)...
- The ASA publishes another magazine: Chance

Could Harold Shipman, one of the most prolific serial killers of all time, had been stopped by using industrial quality control techniques?

Figure 2. Cumulative excess death certificates signed by Shipman, for people older than 64 and who died at home or in his practice
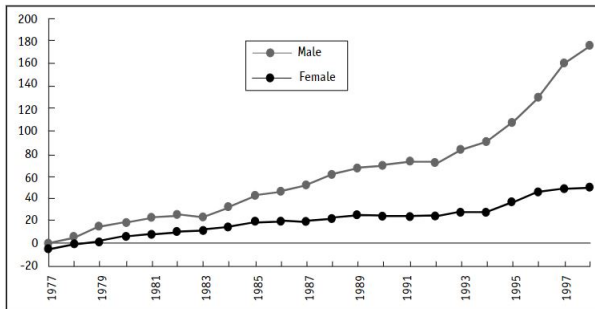
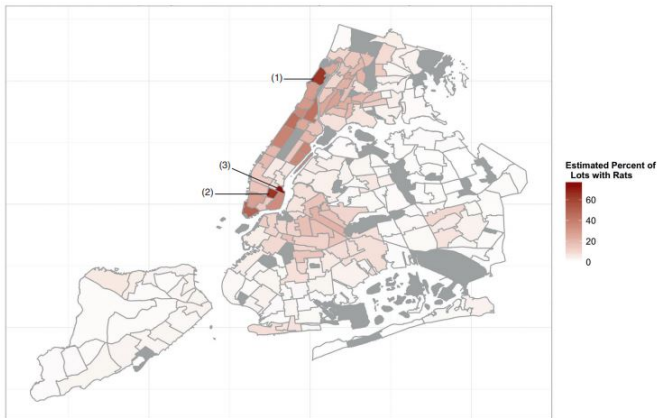(This is one of the most viewed papers in the history of the magazine)

Figure 4. The estimated percentage of New York City rat-inhabited lots by neighbourhood (NTA) as estimated by lot comparison for the 2010–2011 study period. Washington Heights (1), the East Village (2) and Stuyvesant Town-Peter Cooper Village (3) are identified as having large concentrations of rat-inhabited lots

On July 8, at the Estádio Mineirão in Belo Horizonte, Brazil lost 7-1 to Germany in the first semi–final of the 2014 World Cup. Shocking, tragic, stunning are among the adjectives applied to the match's result but can we call it surprising?

Table 1: Weaver's interpretation of SI values

<5	Not surprising
10	Begins to be surprising
1,000	Definitely surprising
1,000,000	Very surprising
1,000,000,000,000	Miracle!

Table 2: Frequencies of goals scored against Brazil in 103 World Cup finals' matches

Goals	0	1	2	3	4	5	6	7
Frequency	44	35	15	6	2	0	0	1

Table 3: Surprise index values from the negative binomial model

Goals	0	1	2	3	4	5	6	7	8
Surprise Index	0.7	1.0	2.1	5.2	14.3	24.0	129.3	412.0	1348.5

Figure 1: Empirical probabilities and probabilities from the negative binomial model
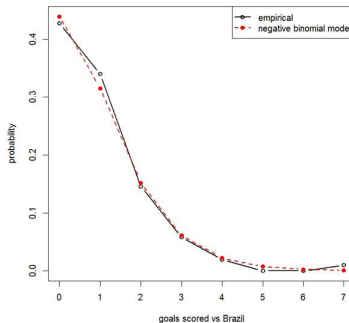
The long modern human growth period has neutralised the advantage of having a fast period of early root growth, so the root spurt has presumably remained where it was at the last point in our evolutionary history at which it conferred some benefit
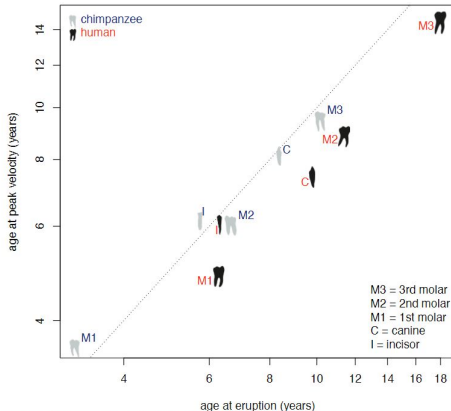

Figure 4. The mean age at peak velocity plotted against the mean age at eruption for five tooth types in chimpanzees (grey teeth, blue text) and humans (black teeth, red text). The age axes (years) are on logarithmic scales, and the line of equality, joining points where age at peak velocity equals age at eruption, is shown
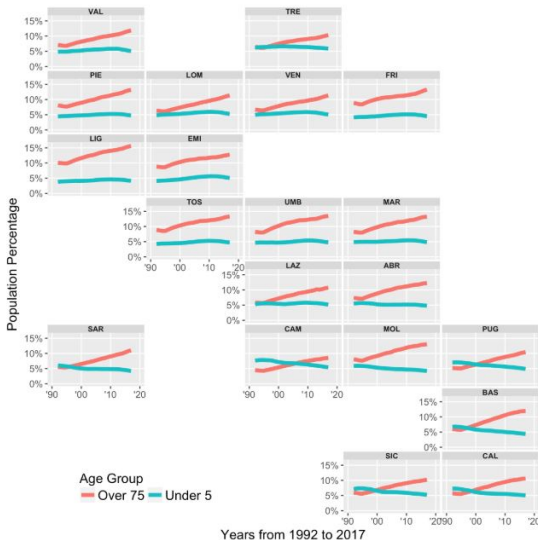
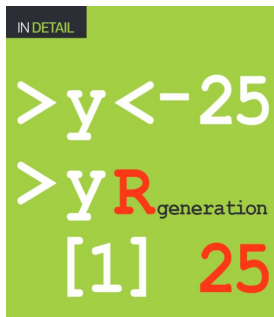How often does the President of the USA say the word "America" in his State of the Union address?
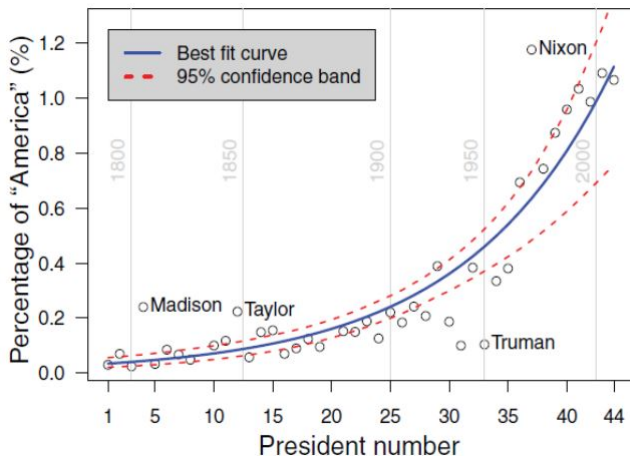
Stella Cangelosi, Luciana Dalla Valle & Julian Stander (2017) *Visualising regional data using the* `geofacet` *R package*

Examples

IN DETAIL

```
>y<-25
>y R generation
[1] 25
```

ROBERT GENTLEMAN

ROSS IHAKA

PETER DALGAARD

HADLEY WICKHAM

JULIA SILGE

GABRIELA DE QUEIROZ

SIGNIFICANCE

Letisha Smith (2018) *Cooking up statistics: the science and the art*

Winner of the 2018 Award for Statistical Excellence in Early–Career Writing.
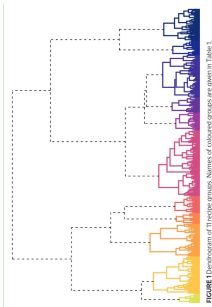Using machine learning to help streamline meal plans

TABLE 1 Details of each recipe group, including total number of recipes within each group, ingredients and example recipes. The final column reports the expected percentage of shared ingredients when selecting recipes at random within each group.

Name	Total	Ingredients	Recipes	Expected % of shared ingredients if 3 recipes are selected at random
Shrimp	30	shrimp, garlic, salt, olive oil, pasta, cheese, tomato	Basil shrimp, Spicy grilled shrimp, Lemon ginger shrimp, Shrimp Florentine with zoodles, Linguine with seafood and sun-dried tomatoes	36
Global vegetarian and dishes with prominent veggie sides	115	tomatoes, olive oil, garlic, onion	Baked eggplant parmesan, Quinoa black bean burgers, Chicken with quinoa and veggies, Avocado shrimp ceviche–estillo Sarita, Marrakesh vegetable curry	26
Breadcrumbs for coating, stuffing, and patties	101	garlic, breadcrumbs, cheese, egg, onion	Yummy lemon salmon burgers, Mozzarella stuffed pesto turkey meatballs, Crab stuffed flounder, Chef John's ricotta meatballs, Asparagus and mozzarella stuffed chicken breasts, Garlic lemon double stuffed chicken	27
Tex-Mex	33	cheese (monterrey jack, cheddar), chicken, soup, tortilla	Chicken enchiladas II, Fabulous wet burritos, Chicken chimichangas with sour cream sauce, Jimmy's Mexican pizza, Taco lasagne, Black bean and corn quesadillas	38
Italian	53	cheese (mozzarella, parmesan) tomato sauce, ground beef, ricotta, basil, olive oil	Baked spaghetti, Chicken and spinach Alfredo lasagne, Spinach cheese manicotti, Italian grilled cheese sandwiches, Mushroom risotto	34
Meat marinade	203	salt, garlic, pepper, olive oil, onion, pork, chicken, beef	Salmon with lemon and dill, Spinach stuffed chicken breasts, Grilled tri-tip with Oregon herb rub, Prune and olive chicken	17
Poultry especially in pies	54	chicken, cheese, cream, salt, butter	Chicken pot pie III, BBQ chicken pizza, Honey mustard grilled chicken	22
Chicken in broth	47	chicken, salt, butter, flour, broth, vegetables (carrot, mushroom, celery)	Polish chicken and dumplings, Chicken Marsala, Chef John's salt roasted chicken	28
Slow cooked	131	onion, salt, butter, flour, potatoes, carrot, chicken, corn	Slow cooker turkey breast, Honey baked chicken II, Campbell's slow cooker savoury pot roast, Mississippi roast slow cooker pepperoncini pot roast	13
BBQ, meatloaf, sloppy joes	44	brown sugar, ketchup, mustard, barbecue sauce, ground beef, pork, vinegar	Texas pork ribs, Meatball nirvana, Chef John's turkey sloppy joes, Slow cooker Texas pulled pork, Scott Hibbs' amazing whisky grilled baby back ribs	37
Pan-Asian	98	soy sauce, garlic, ginger, chicken, vinegar, lemon	Thai pineapple chicken curry, Asian pork tenderloin, Japanese beef stir fry, House fried rice, Teriyaki marinated chicken	26

*Values were estimated by drawing 10 000 samples from each group.

FIGURE 1 Dendrogram of 11 recipe groups. Names of coloured groups are given in Table 1

Gian Carlo Di–Luvi (2019) *How does it make you feel?*

Quantifying emotional content in popular songs is a challenge; this paper attempts a solution, using text mining to determine the net sentiment of Grammy–winning Albums of the Year
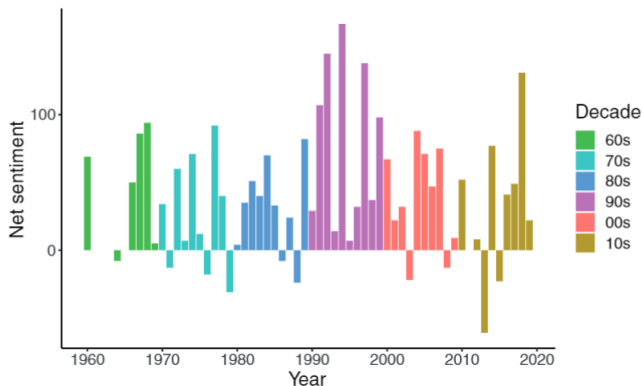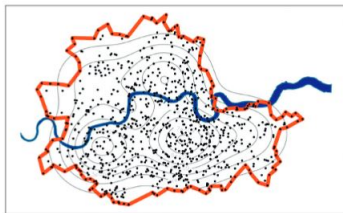

FIGURE 4 Net sentiment scores of Grammy Albums of the Year, 1960–2019. Missing values in the

A new analysis of the classic dataset collected by RD Clarke, a British actuary who sought to determine whether the apparent clustering of V–1 strikes on London during the Second World War was the result of targeting or randomness

Number of V–1s in circle (k)	Expected number of circles (Poisson)	Observed number of circles
0	469.07	508
1	355.09	306
2	134.40	121
3	33.91	51
4	6.42	14
5 and over	0.97	0
	1,000	1,000

FIGURE 3 Using the whole bomb map to apply Clarke's randomness test. We fit a boundary (red line) around the V–1 hits (black points) and sample 1,000 circles of area 0.25 km² each from within it. The results do not fit a Poisson distribution well with λ = 0.757 (lower than Clarke's value) because the density of the V–1s is not constant. Fitting the density of the V–1s using a technique called a Gaussian mixture model – which estimates the number of centred distributions which must be added together to explain the observed distribution – suggests three main centres of density (overlaid contours in grey). However, the data is limited by the boundaries of the London County map.

- Read the magazine
- Read its authors' guidelines:
  `https://rss.onlinelibrary.wiley.com/hub/journal/17409713/author-guidelines`
- Choose a topical issue
- Approach your subject from a statistical angle
- Minimise the number of equations, and citations, and technical jargon
- Start with a strong title, then a strong introductory sentence, and end with a strong, rounding sentence
- **Write, not Right!**