

51 Awesome Data Science Blogs You Need To Check Out

Articles Career Advice 15 min read

Blog (<https://365datascience.com/blog>) / Articles (<https://365datascience.com/blog/#articles>) / 51 Awesome Data Science Blogs You Need To Check Out

Blogs, they're everywhere. Blogs about travel, blogs about pets, blogs about blogs. And data science is no exception. Data science blogs are a dime a dozen and with so many, where do you start when you need to find the most valuable information for your needs?

That's what we're here to help with.

Data science is a rapidly growing field and it should come as no surprise that many people have something to say on the subject. So, this article is going to briefly review 51 data science blogs that we thought were pretty awesome and give you a snippet of what they are all about.

Note: There is no ranking order for these, it is less about 'Which is the best blog?' and more about 'All of these blogs are awesome but which one offers what I need?'.

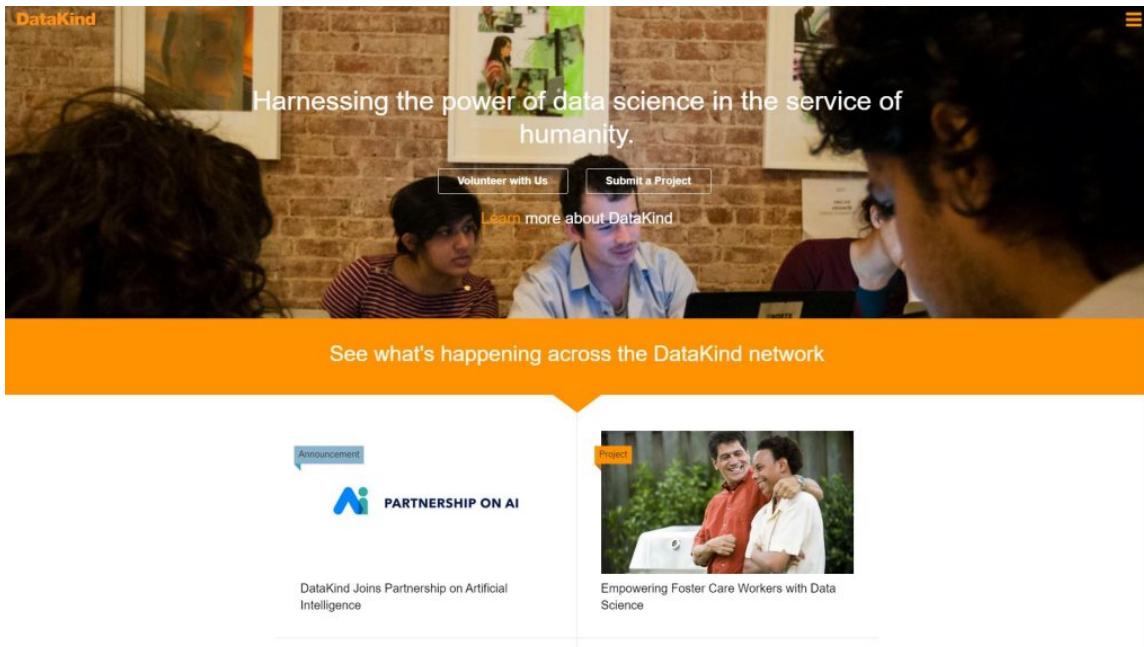
We will give:

- A brief description of each blog (from my perspective).
- Their Alexa rank (which uses stats on visitors and page views over the last three months to give a global traffic ranking; this is updated daily so be aware that this can change from the time of writing).
- The frequency with which the blog is updated.
- An interesting article from the site being reviewed.



Hope this helps!

Blog name: DataKind (<http://www.datakind.org/>)



(<https://365datascience.com/wp-content/uploads/2018/11/datakind-e1542114505843.jpg>)

Description: Using data to save the world. That's what DataKind is all about. Using the same algorithms and techniques which companies use to make a profit and using them to create a sustainable world where all human needs are met – That's founder Jake Porway's vision and it seems they are doing a great job. Check out their blog to see what they get up to and how they help everything from Foster Care to Disaster prevention. A wonderful and ethical company and blog

Alexa Rank: 638,290

Post frequency: 1-2 per week

Article to check out: Follow the amazing efforts to protect the endangered banded leaf monkey from extinction. Have a read of this article to see how the RBLWG are using machine vision to compare photos of the monkeys so they can more accurately judge population numbers and geolocation to assess habitat.

Protecting Singapore's Endangered Banded Leaf Monkey

(<http://www.datakind.org/blog/protecting-singapores-endangered-banded-leaf-monkey>)



Blog name: Probably Overthinking It (<https://www.allendowney.com/blog/>)

(<https://365datasience.com/wp-content/uploads/2018/11/Probably-overthingking-it-min-e1542113487898.jpg>)

Description: Allen Downey is an author and professor at Olin College. He writes generally about data science and Bayesian statistics. He has built a nice clear blog (although this is his new blog, he links to his old one which has a large archive). Allen is active on twitter and often goes through problems with his followers. If you fancy trying to solve some problems or are just looking for some general data science and statistics reading, this is a good place to go.

Alexa Rank: 1,005,753

Post frequency: 2-4 per month

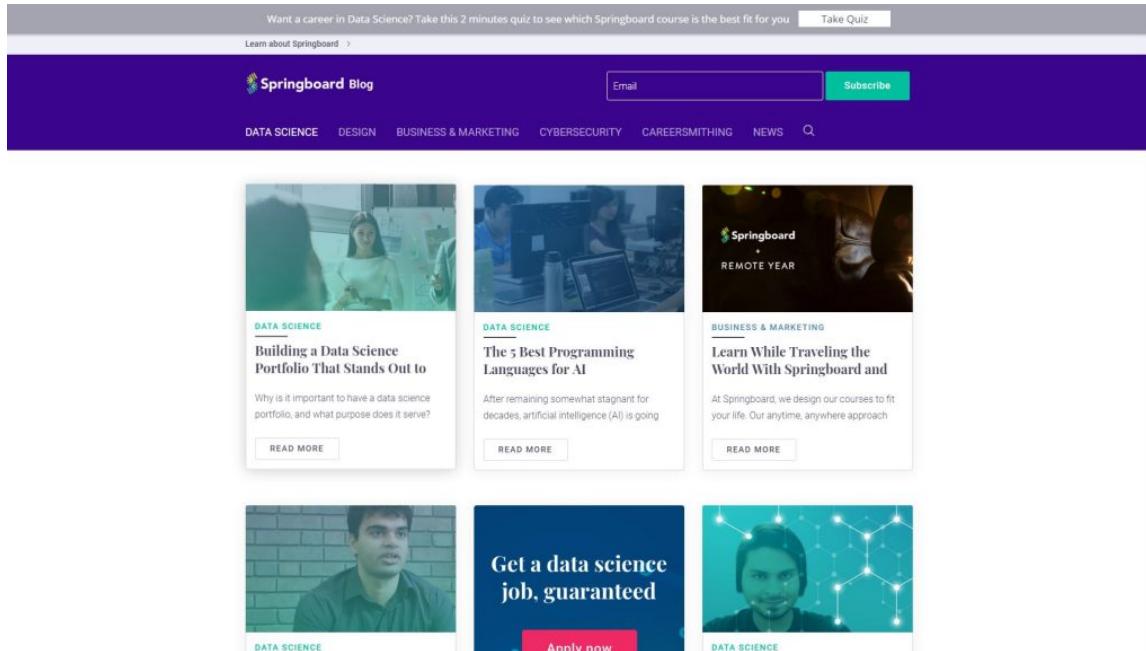
Article to check out: If you like dungeons and dragons (the game, obviously) and/or Bayesian statistics problems, then have a look at this problem. The solution is posted below in Jupyter so you can compare.

The Dungeons and Dragons Problem

(<https://www.allendowney.com/blog/2018/10/02/the-dungeons-and-dragons-problem/>)



Blog name: Springboard (<https://www.springboard.com/blog/category/data-science/>)



(<https://365datascience.com/wp-content/uploads/2018/11/Springboard-min-e1542113532340.jpg>)

Description: Springboard covers a few topics but it doesn't detract from its data science articles. Some really nice posts here with an attractive presentation. It contains career tips, lists and data science as a general topic. It is an inviting blog for almost anybody, whether you're just getting into AI or are already a deep learning veteran.

Alexa Rank: 43,323

Post frequency: Daily

Article to check out: Everybody enjoys a good TED talk and if you're reading this then you also enjoy data science. Sometimes those things combine and sometimes someone is kind enough to compile the 11 best data science TED talk out there. How nice of T.J DeGroat to do that – he even gives a brief synopsis of each one. Plenty to keep you engaged.

11 Must-Watch TED Talks on Data Science (<https://www.springboard.com/blog/11-must-watch-ted-talks-on-data-science/>)

Blog name: BAIR (<https://bair.berkeley.edu/blog/>)





Visual Reinforcement Learning with Imagined Goals

Vitchyr Pong* and Ashvin Nair* Sep 6, 2018

We want to build agents that can accomplish arbitrary goals in unstructured complex environments, such as a personal robot that can perform household chores. A promising approach is to use deep reinforcement learning, which is a powerful framework for teaching agents to maximize a reward function. However, the typical reinforcement learning paradigm involves training an agent to solve an individual task with a manually designed reward. For example, you might train a robot to set a dinner table by designing a reward function based on the distance between each plate or utensil and its goal location. This setup requires a person to design the reward function for each task, as well as extra systems like object detectors, which can be expensive and brittle. Moreover, if we want machines that can perform a large repertoire of chores, we would have to repeat this RL training procedure on each new task.



(<https://365datasience.com/wp-content/uploads/2018/11/BAIR-min-e1542113318169.jpg>)

Description: Berkeley Artificial Intelligence Research (BAIR)'s blog is for their researchers to communicate research findings and perspectives on the AI field. Written by students and faculty, this blog aims to reach both expert and general audiences and they achieve this aim with an abundance of articles by some of the greatest minds in AI research. Something for everyone interested in AI.

Alexa Rank: 1298

Post frequency: 1-2 per week

Article to check out: Imitation is a key part of human intelligence. This article shows how researchers have developed AI to imitate actions in a variety of ways. Absolutely fascinating and it raises many questions about the future of AI and the amazing work done here. A must read.

One-Shot Imitation from Watching Videos

(<https://bair.berkeley.edu/blog/2018/06/28/daml/>)

Blog name: KDnuggets
(<https://www.kdnuggets.com/news/index.html>)



The screenshot shows the KDnuggets News homepage. At the top, there's a yellow header bar with the KDnuggets logo, a search bar, and social media links for Twitter, Facebook, and LinkedIn. Below the header, the main content area has a yellow header "News / Blog". It features a list of "Latest" news articles with titles like "ODSC Europe Schedule Launched + Premium Data Science Training + ODSC West - Jul 24, 2018.", "SocialCops: Sr Data Analyst - Jul 24, 2018.", and "New Online MS in Business Analytics from Drexel - Jul 24, 2018.". There are also sections for "More Recent Stories" and "Top Stories". A sidebar on the right shows a vertical stack of social media icons (Facebook, Twitter, LinkedIn, etc.) with a "3.1K SHARES" counter.

(<https://365datascience.com/wp-content/uploads/2018/11/1-KD-nuggets-e1542113095741.jpg>)

Description: On the off chance you haven't heard of *KDnuggets*. It's one of the best blogs out there. They feature plenty of quality guest posts (including some from us) and although it can look a bit overwhelming, you are sure to find something good to read. Go check it out.

Alexa Rank: 20,825

Post frequency: Daily

Article to check out: Well this article by us is the obvious choice really. *KDnuggets* awarded us the Platinum badge for it, which was a great honour.

Data Science vs Machine Learning vs Data Analytics vs Business Analytics

(<https://www.kdnuggets.com/2018/05/data-science-machine-learning-business-analytics.html>)

Blog name: O'Reilly
(<https://www.oreilly.com/topics/data>)

(<https://365datasience.com/wp-content/uploads/2018/11/6-OReilly-e1542113594764.jpg>)

Description: A professional and in-depth look into data and data science that is in no way intimidating. The contributors here clearly know what they are doing. Posts include videos, podcasts, and superbly written articles. Very easy to binge read on this site.

Alexa Rank: 10,527

Post frequency: Daily

Article to check out: With data growing faster than we can comprehend, no-one can be totally sure of what's to come. This article gives a fresh take on this idea.

Data's day of reckoning (<https://www.oreilly.com/ideas/datas-day-of-reckoning>)



(<https://365datascience.com/complete-data-science-training/>)

Blog name: *Adit Deshpande* **(<https://adeshpande3.github.io/adeshpande3.github.io/>)**

The Last 5 Years In Deep Learning

We've come quite a long way

Interest over time

deep learning
Search term

READ MORE

Why Machine Learning Is A Metaphor For Life

There are more similarities than you think

(<https://365datascience.com/wp-content/uploads/2018/11/3-adeshpande-e1542113586255.jpg>)

Description: This blog has been quiet for a few months and it is a little light in terms of the number of posts, but oh boy it is quite the opposite when it comes to the posts themselves. *Adit Deshpande* has really put in the effort and it shows by the number of comments saying how much he has helped the comment author. Head there now to get some fantastic machine and deep learning tutorials.

Alexa Rank: 150,484 (Github's Alexa rank)

Post frequency: It varies

Article to check out: *Adit* wrote a very entertaining post about using deep learning to train a chatbot to talk like you. A lot of fun and you can try it for yourself.

How I Used Deep Learning To Train A Chatbot To Talk Like Me (Sorta)

(<https://adeshpande3.github.io/adeshpande3.github.io/How-I-Used-Deep-Learning-to-Train-a-Chatbot-to-Talk-Like-Me>)

Blog name: Andrew Gelman (<http://andrewgelman.com/>)

The screenshot shows a blog post titled "Revisiting "Is the scientific paper a fraud?"" by Andrew. The post discusses the nature of scientific papers and their formalized structure. It includes several paragraphs of text, a sidebar with recent comments, and a navigation bar at the top.

(<https://365datasience.com/wp-content/uploads/2018/11/20-Andrew-Gelman-e1542113612342.jpg>)

Description: Eleven authors on this blog mean it is packed full of statistical content. Each contributor adds a personal touch to their writing with some genuinely funny posts (and comments). You'll find general opinions, scientific reviews, tutorials and a healthy amount of sarcasm. Also, they have a 'zombies' category... braiiins!

Alexa Rank: 149,833

Post frequency: Daily

Article to check out: If you've ever read or written a scientific paper then you will relate to this post.

Revisiting "Is the scientific paper a fraud?"

(<http://andrewgelman.com/2018/07/29/revisiting-scientific-paper-fraud/>)

Blog name: Hyndsoft (<https://robjhyndman.com/hyndsoft/>)

The screenshot shows the homepage of the Hyndsoft blog. At the top, there's a navigation bar with links to 'Rob J Hyndman', 'Hyndsoft blog', 'Publications', 'Software', 'Seminars', 'Teaching', 'Research team', and 'About'. Below the navigation, there's a section titled 'All Hyndsoft posts by date' with a link to 'UserR!2018 talks' from July 2018. This post discusses talks from the userR! conference held in Brisbane. Another post listed is 'Seasonal decomposition of short time series' from July 2018, which addresses a common error in R's decomposition functions. A third post, 'A forecast ensemble benchmark', from June 2018, compares various forecasting methods. A fourth post, 'Forecasting in NYC: 25-27 June 2018', from April 2018, announces a workshop in New York. On the right side of the page, there's a sidebar titled 'Topics covered' which is a tag cloud. The most prominent tags include 'forecasting', 'r', 'statistics', 'research', 'computing', 'data-science', 'conference', 'publications', 'teaching', and 'research-team'. Other smaller tags visible include 'graphics', 'hts', 'humour', 'jif', 'sf2017', 'jobs', 'journals', 'kaggle', 'latex', 'mathematics', 'maxima', 'monash-university', 'obituary', 'organization', 'otexts', 'phd', 'poetry', 'prizes', 'productivity', 'progress', 'publishing', 'refereeing', 'references', 'reproducible-research', 'research-team', 'ropensci', 'seminars', 'stackexchange', and 'supervision'.

(<https://365datasience.com/wp-content/uploads/2018/11/61-Hyndsoft->

e1542113244800.jpg)

Description: 'Forecasting' is Rob Hyndman's most used tag. So, if that interests you then you're in luck. He also covers statistics, computing, conferences and more. The blog is aimed at his research team but hopes researchers in other statistical mathematical science fields will find some use in it. If this is you, then go take a look – Rob's door is wide open.

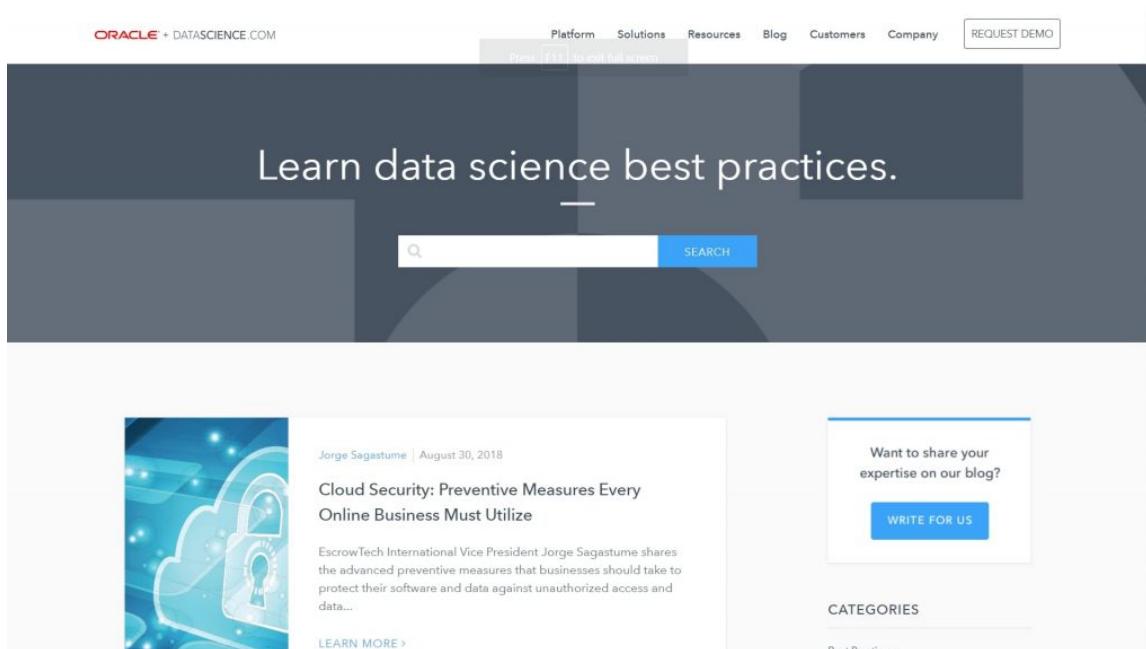
Alexa Rank: 207,870

Post frequency: 2-3 per month

Article to check out: Forecasting is the main thing here, so one of the latest posts on it will give you an idea of Rob's style.

A forecast ensemble benchmark (<https://robjhyndman.com/hyndsoft/benchmark-combination/>)

Blog name: Oracle and Data Science (<https://www.datascience.com/blog>)



(<https://365datasience.com/wp-content/uploads/2018/11/Oracle-and-data-science-e1542113481182.jpg>)

Description: Recently bought by Oracle this is perhaps one of *the* top data science blogs. Stylish, quick and easy to use; This a hub of data science information a must for anyone interested in the world of data.

Alexa Rank: 71,618

Post frequency: 2-3 per week

Article to check out: Some great reasons for using search engine tools for data analysis. Nicely written and genuinely useful for anyone who wants to analyse data.

7 Benefits of Using Search Engine Tools for Data Analysis

(<https://www.datascience.com/blog/search-engine-tools-for-data-analysis>)

Blog name: Revolutions
(<http://blog.revolutionanalytics.com/>)

(<https://365datascience.com/wp-content/uploads/2018/11/Revolutions-e1542113507740.jpg>)

Description: *Revolutions* is a blog dedicated to delivering news and valuable information to members of the R community. If you like to keep up to date with the latest R goings-on then this blog is worth keeping in your bookmarks

Alexa Rank: 128,146

Post frequency: Every working day

Article to check out: More of a video than an article. Take a look at Microsoft Azure in this informative and entertaining video on machine learning.

Video: Azure Machine Learning in plain English

(<http://blog.revolutionanalytics.com/2018/08/aml-video.html>)

Blog name: Airbnb Data blog
(<https://medium.com/airbnb-engineering/tagged/data-science>)

The screenshot shows the Airbnb Engineering & Data Science blog on Medium. The sidebar includes a 'TAGGED IN' section for 'Data Science', the blog's logo, a brief description of the team, follower count (55K), and links to other platforms like Twitter and Facebook. The main content area displays an article by Shijjing Yao titled 'Categorizing Listing Photos at Airbnb', which discusses how large-scale deep learning models are changing the way we think about listing photos. The article has 2.1K views and 10 responses.

(<https://365datasience.com/wp-content/uploads/2018/11/8-Airbnb-e1542113600594.jpg>)

Description: Airbnb's Data blog is the place where their data scientists go to write about how the company uses data to provide their service. It's an open look into one of the most successful businesses in recent history. It has great insights, not only if you use the service, but for anyone interested in real-world uses of data. With many contributors, you're sure to find an interesting post.

Alexa Rank: 260 (this is Medium's Alexa rank)

Post frequency: Bi-monthly

Article to check out: Although one of the older articles, it gives a great overview of how Airbnb does things.

How Airbnb uses Machine Learning to Detect Host Preferences

(<https://medium.com/airbnb-engineering/how-airbnb-uses-machine-learning-to-detect-host-preferences-18ce07150fa3>)

Blog name: The Data Incubator
(<https://blog.thedataincubator.com/>)

The Data Incubator

“ DATA ARE BECOMING THE NEW RAW MATERIAL OF BUSINESS ”
— THE ECONOMIST

 The Data Incubator Blog

Alumni Spotlights Data Science in 30 Minutes Data Sources Rankings Technical Blogs Uncategorized


Telling Truth from Hype When Hunting for Data Science Work
Scams and fake offers are an unfortunate reality of online job searches across industries.


The Future of Management: Human Resource Analytics
Technology is making its way into the HR department so that organizations can use their data to improve.


The Benefits of Active Learning for Data Science Skills
The success of active learning has led institutions of higher learning to implement it in their programs


How Foundations Student


2018 Data Sources for Cool Data


Data Science in 30 Minutes: Using Data Science to Predict

<https://365datasience.com/wp-content/uploads/2018/11/The-Data-Incubator-min-e1542113552158.jpg>

Description: The Data Incubator takes PhD graduates from academia and trains them to be work ready. This is an invaluable service to the data science world and, as it happens, so is their blog. With interviews from some amazing minds and articles on data science as a whole alongside more specifically written tutorials and comparisons to help anyone with an interest or passion for data – especially if you're a recent graduate, this place could be just what you need.

Alexa Rank: 163,747

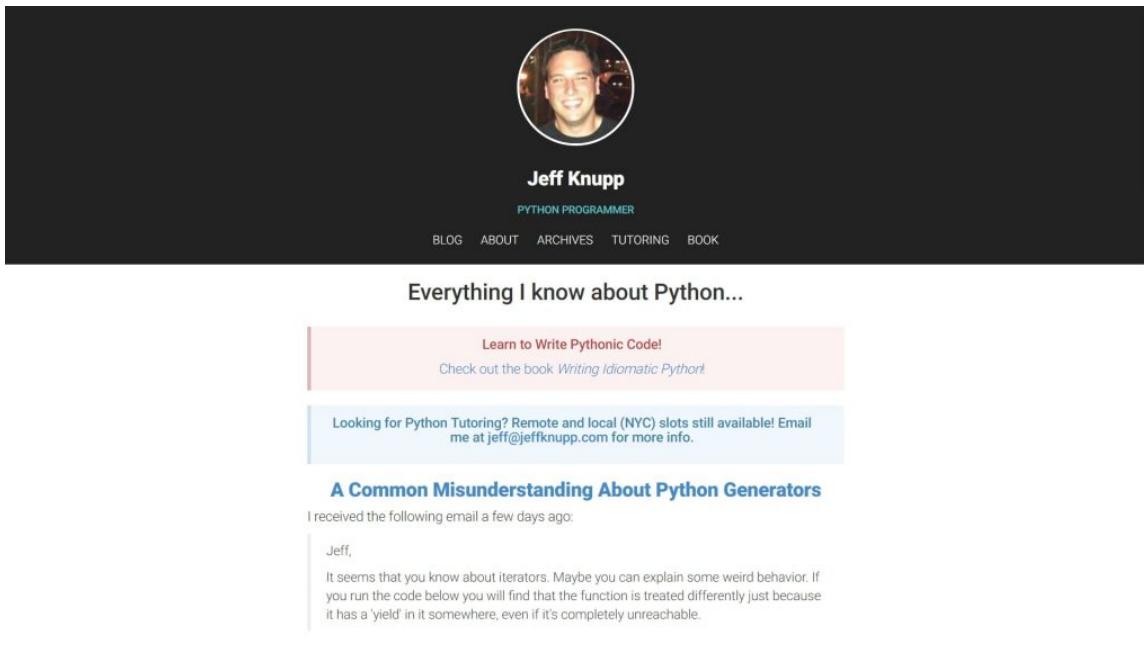
Post frequency: 1-2 per week

Article to check out: Want to gain more data science skills – this article is a great place to start. It gives you a rundown of the benefits of active learning for data science skills. Data incubator use the technique themselves, so they know what they're talking about.

The Benefits of Active Learning for Data Science Skills

(<https://blog.thedataincubator.com/2018/09/active-learning-for-data-science/>)

Blog name: Jeff Knupp (<https://jeffknupp.com/>)

A screenshot of Jeff Knupp's blog homepage. At the top is a circular profile picture of Jeff Knupp, a smiling man with short hair. Below the picture is his name, "Jeff Knupp", and the title "PYTHON PROGRAMMER". A navigation bar below the title includes links for "BLOG", "ABOUT", "ARCHIVES", "TUTORING", and "BOOK". The main content area features a heading "Everything I know about Python...". Below this is a pink box containing the text "Learn to Write Pythonic Code! Check out the book *Writing Idiomatic Python!*". Another box below it says "Looking for Python Tutoring? Remote and local (NYC) slots still available! Email me at jeff@jeffknupp.com for more info." A section titled "A Common Misunderstanding About Python Generators" follows, with a note that "I received the following email a few days ago:" and a snippet of an email from "Jeff" discussing iterators and weird behavior.

(<https://365datasience.com/wp-content/uploads/2018/11/11-Jeff-Knupp-e1542113606379.jpg>)

Description: Jeff Knupp is a Python programmer and that's just what his blog focuses on. He has a nice, personal writing style and discusses a wide range of python functions on which he goes into more detail in his book. I would recommend if you are looking to pick up some useful python tips that you can try out.

Alexa Rank: 146,974

Post frequency: 1-2 per 3 months

Article to check out: This post on using python to count cards is a good example of Jeff's engaging writing style. Take a look to see an interesting (and perhaps profitable) use of the programming language.

Counting Cards With Python (<https://jeffknupp.com/blog/2018/01/02/counting-cards-with-python/>)

Blog name: The Clever Machine
(<https://theclevermachine.wordpress.com/>)



Derivation: Maximum Likelihood for Boltzmann Machines

SEP 23 posted by [dustinstansbury](#)

In this post I will review the gradient descent algorithm that is commonly used to train the general class of models known as Boltzmann machines. Though the primary goal of the post is to supplement another post on restricted Boltzmann machines, I hope that those readers who are curious about how Boltzmann machines are trained, but have found it difficult to track down a complete or straight-forward derivation of the maximum likelihood learning algorithm for these models (as I have), will also find the post informative.

First, a little background: Boltzmann machines are stochastic neural networks that can be thought of as the probabilistic extension of the Hopfield network. The goal of the Boltzmann machine is to model a set of observed data in terms of a set of visible random variables v and a set of latent/unobserved random variables h . Due to the relationship between Boltzmann machines and neural networks, the random variables are often referred to as "units." The role of the visible units is to approximate the true distribution of the data, while the role of the latent variables is to extend the expressiveness of the model by capturing underlying features in the observed data. The latent variables are often referred to as hidden units, as they do not result directly from the observed data and are generally marginalized over to obtain the likelihood of the observed data, i.e.

$$p(v; \theta) = \sum_h p(v, h; \theta).$$

where $p(v, h; \theta)$ is the joint probability distribution over the visible and hidden units based on the current model parameters θ . The general Boltzmann machine defines $p(v, h; \theta)$ through a set of weighted, symmetric connections between all visible and hidden units (but no connections from any unit to itself). The graphical model for the general Boltzmann machine is shown in Figure 1.

(<https://365datasience.com/wp-content/uploads/2018/11/The-clever-Machine-min-e1542113545670.jpg>)

Description: As Dustin Stansbury, the creator of this technical blog, puts it “the blog will be posts about analysis techniques, algorithms, theory, or things I think are cool, exciting, or helpful.” And it does exactly what it says on the tin. This blog delves deep into long complex articles. If you’re looking for a light read, this is not the place, if you’re looking for in-depth tutorials, then you’ve found what you are looking for.

Alexa Rank: 555.599

Post frequency: 2 per Month

Article to check out: This article is called “A Gentle Introduction to Artificial Neural Networks” – so it will give you a base idea of what Dustin considers ‘gentle’ and if you too consider it ‘gentle’ then don’t worry because he even links to what I can only refer to as ‘The mind-blowing’.

A Gentle Introduction to Artificial Neural Networks

(<https://theclevermachine.wordpress.com/2014/09/11/a-gentle-introduction-to-artificial-neural-networks/>)

Blog name: Dataiku (<https://blog.dataiku.com/>)



(<https://365datasience.com/wp-content/uploads/2018/11/Dataiku-mini1542113409603.jpg>)

Description: Dataiku is a central hub for analytics and machine learning and their blog shows this with regular updates on the world of analytics. News, webinars and loads of blog posts – This is a good place for anyone interested in analytics

Alexa Rank: 172,517

Post frequency: 1-3 per week

Article to check out: Deep learning is everywhere these days, and this post takes you through exactly what it is while avoiding the jargon. A great guide

A Deep (But Jargon and Math Free) Dive Into Deep Learning

(<https://blog.dataiku.com/deep-learning-essentials>)

Blog name: Cloudera
(<https://www.cloudera.com/more/news-and-blogs.html>)

The screenshot shows the Cloudera website's press releases page. At the top, there's a navigation bar with links for PRODUCTS, SOLUTIONS, DOWNLOADS, MORE, and social media icons. Below the navigation is a banner image of a subway tunnel with many framed pictures on the walls. Overlaid on the banner is the text "Press, news, and blogs". Underneath the banner, the heading "Recent press releases" is displayed. Three recent press releases are listed:

- Sep 10, 2018 - Cloudera and Tech Data forge strategic partnership in Australia and New Zealand
- Sep 10, 2018 - Cloudera Delivers Its Most Powerful Machine Learning and Analytics Platform - Cloudera Enterprise 6
- Sep 05, 2018 - Cloudera Reports Second Quarter Fiscal Year 2019 Financial Results

A "PRESS RELEASES" button is located below the list. On the right side of the page, there's a "Cloudera in the news" section with a small orange button labeled "CHAT".

(<https://365datasience.com/wp-content/uploads/2018/11/Cloudera-min-e1542113338351.jpg>)

Description: Cloudera has a couple of blogs. The first is an engineering blog which covers best practices, how-tos, use cases, and internals from Cloudera Engineering and its community. The second, a VISION blog, is a wider picture on a range of topics but at the heart is what Cloudera stands for – Security. A company comprised of some big names from Google, Facebook, Hadoop, Yahoo! and Oracle their blog has a lot to offer especially if you have an interest in the company.

Alexa Rank: 17,604

Post frequency: Daily

Article to check out: This article goes in-depth into how Cloudera and Accenture joined forces to fight against insider threats. Illegal insider-actions can cost companies millions, so it's a problem that is worth knowing something about, especially if you've got a business with some shifty employees.

Taking out the threat from the inside (<http://vision.cloudera.com/taking-out-the-threat-from-the-inside/>)

Blog name: Aylien (<http://blog.aylien.com/>)

The screenshot shows the Aylien blog homepage. At the top, there's a navigation bar with links for Products, Research, Blog, Company, and Contact. Below that is a secondary navigation bar with links for General, Product, Data Science, Research, a search icon, and a 'Get Started' button. The main content area features a section titled 'Latest Posts' with three cards. Each card includes a small profile picture, the author's name, the post title, and the publication date. Below this is a section titled 'What are you looking for?' with a 'Browse Our Categories' link. A horizontal navigation bar at the bottom allows users to switch between Product, Data Science, Research, and General categories.

(<https://365datasience.com/wp-content/uploads/2018/11/Aylien-min-e1542113310518.jpg>)

Description: There is something here for everyone and anyone interested in Text Analysis, NLP and Deep Learning. Organized in a real user-friendly way, it's clear this blog is made with the reader in mind. Articles are written with quite specific topics, but they are written well and if you're looking for something, in particular, it'll be easy to find it. If not, then you're sure to find something worth reading.

Alexa Rank: 178,014

Post frequency: 1-2 per week

Article to check out: Chatbots are a fascinating part of AI and this article takes a look at the history of them. A genuinely great read for anyone interested in chatbots and their evolution.

An Interactive History of Chatbots (<http://blog.aylien.com/interactive-history-chatbots/>)

Blog name: Subconscious musings
(<https://blogs.sas.com/content/subconsciousmusings/>)

The screenshot shows the Subconscious Musings blog page on the SAS website. The header includes the SAS logo and navigation links for Industry Solutions, Products, Learn, Support, Customer Stories, Partners, and About SAS. The main content area features three blog posts:

- Advanced Analytics** (November 5, 2018): Discusses maps, models, and analytic problem framing.
- Artificial Intelligence | Machine Learning** (October 31, 2018): Improving model interpretability with LIME.
- Advanced Analytics | Machine Learning** (October 17, 2018): Four machine learning strategies for solving real-world problems.

On the right side, there's an "About" section with a bio for Radhika Kulkarni, Advanced Analytics RAD VP, and her colleagues at SAS. It also includes social media sharing icons (LinkedIn, Twitter, Facebook, YouTube, RSS) and a "Tags" section with various labels like advanced analytics, AI interpretability, analytics, etc.

(<https://365datascience.com/wp-content/uploads/2018/11/Subconscious-musings-min-e1542113539704.jpg>)

Description: Radhika Kulkarni and her team at SAS (one of the bigger players in analytics, they are worth knowing about) set out to discuss all things related to advanced analytics that solve many of the challenges that face businesses and organizations. Plenty of interesting topics on this blog, many based on SAS, others that just discuss analytics in general but all are relevant and well written. Great blog if analytics, AI and deep learning are important to you.

Alexa Rank: 6040

Post frequency: 1-3 per month

Article to check out: Machine learning is great! We know that, but it's how it applies to real-world problems that really makes it special. This short but sweet article on just that takes you through 4 strategies in an easy-to-digest way.

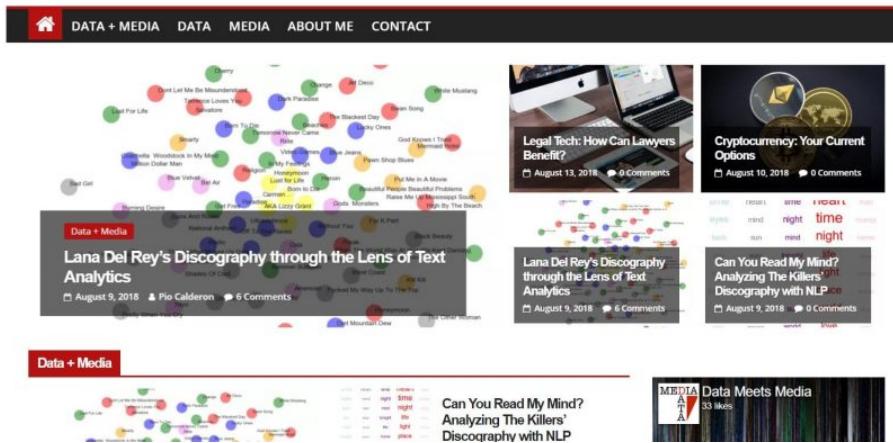
Four machine learning strategies for solving real-world problems

(<https://blogs.sas.com/content/subconsciousmusings/2018/10/17/four-machine-learning-strategies-for-solving-real-world-problems/>)

Blog name: Data Meets Media
(<http://datameetsmedia.com/>)

DATA meets MEDIA

television, movies, data science.



(<https://365datascience.com/wp-content/uploads/2018/11/Data-Meets-Media-min-e1542113389581.jpg>)

Description: Data Meets Media is a blog by Pio Calderon, who takes his two main interests and combines them in interesting ways, using data to analyse anything from the lyrics of 'The Killers' discography to plot summaries of Oscar-winning films. He also provides posts on the topics separately where you can find pieces on cryptocurrencies, AI, data science, movie reviews and more. The blog is just over a year old and already has plenty to interest most people.

Alexa Rank: 1,297,224

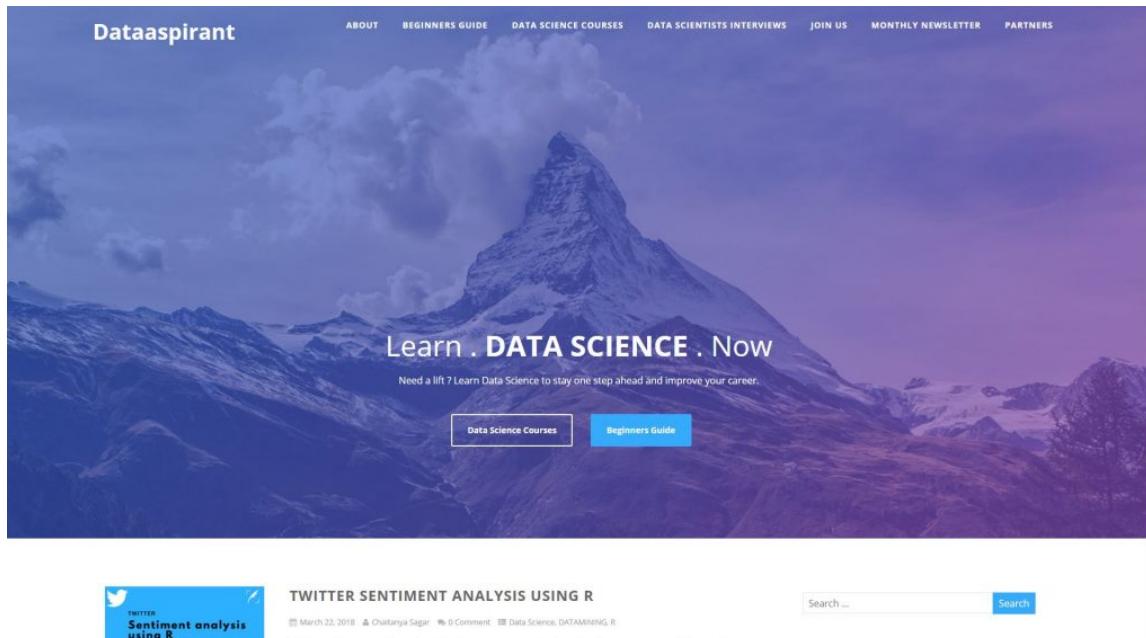
Post frequency: 4-5 per month

Article to check out: Have you ever wondered about the colour dynamics of the Oscar-winning film La La Land? No, me neither. However, Pio here has put together a post on that very topic, one of many things that you didn't know you'd be interested in.

The Color Dynamics of La La Land (<http://datameetsmedia.com/color-dynamics-la-la-land/>)

Blog name: Data Aspirant (<http://dataaspirant.com/>)





(<https://365datascience.com/wp-content/uploads/2018/11/data-aspirant-e1542114229452.jpg>)

Description: Data Aspirant is run by Saimadhu Polamuri who is a self-taught data scientist and educator. Data Aspirant is a portal for data science beginners, so here is where you will find tutorials on all areas of data science. A great place to get started or brush up on your skills.

Alexa Rank: 160,140

Post frequency: 1-2 per month

Article to check out: A long in-depth list of R tips. If you want to improve your R skills it's worth a read. Well written and easy to follow, a good taster for what else is on the blog.

10 SMART R PROGRAMMING TIPS TO BECOME BETTER R PROGRAMMER

(<http://dataaspirant.com/2017/11/04/smart-r-programming-tips/>)

Blog name: Facebook Research
(<https://research.fb.com/blog/>)



(<https://365datasience.com/wp-content/uploads/2018/11/Facebook-research-min-e1542113436956.jpg>)

Description: Facebook is making leaps and bounds in the AI, machine learning and data science fields. This blog keeps you up to date with everything the Facebook research team are doing. Plus plenty of insights into economics, natural language processing, connectivity, security and plenty more.

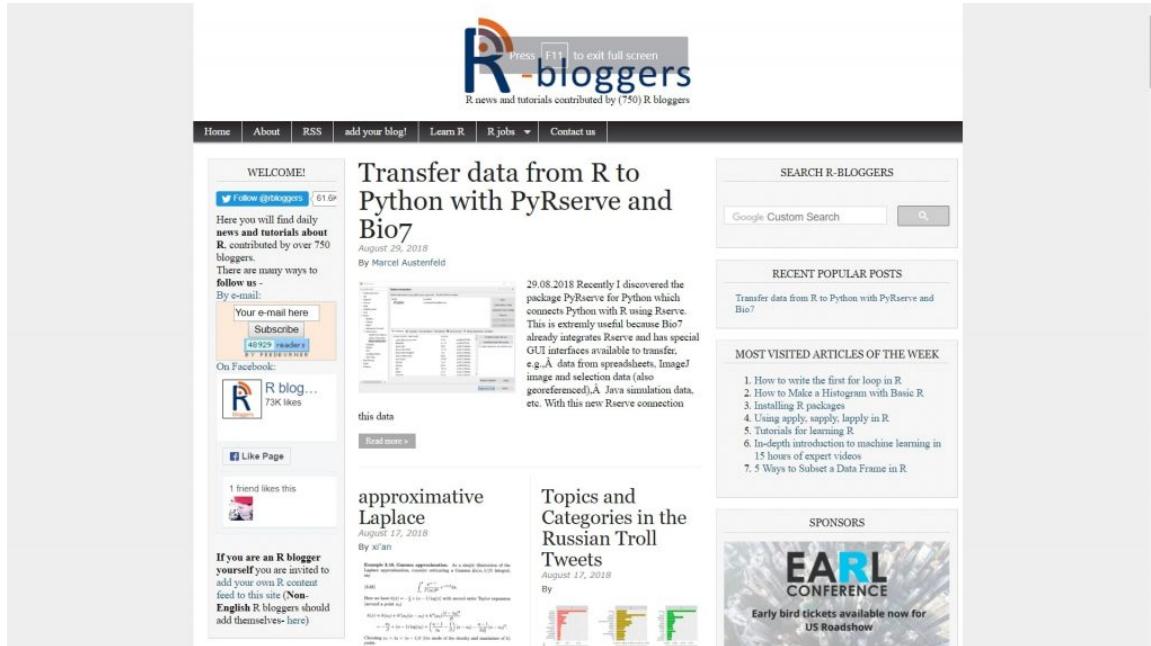
Alexa Rank: 8912

Post frequency: 1-2 per week

Article to check out: Facebook researchers use AI to turn whistles into orchestral music, and power other musical "translations" – Another article you wouldn't expect to be written, but that's the world we live in. Read this to see what steps researchers are taking with AI technology that could transform the music industry.

Facebook researchers use AI to turn whistles into orchestral music, and power other musical "translations" (<https://research.fb.com/facebook-researchers-use-ai-to-turn-whistles-into-orchestral-music-and-power-other-musical-translations/>)

Blog name: R-bloggers (<https://www.r-bloggers.com/>)



(<https://365datascience.com/wp-content/uploads/2018/11/R-bloggers-e1542113500865.jpg>)

Description: 750 contributors writing about, if you hadn't guessed it, R. R-Bloggers is about encouraging bloggers to empower R users. There is a strong community here and if you want to get involved in R or you already 'R' then there will be something for you here.

Alexa Rank: 14,541

Post frequency: Daily

Article to check out: A nice easy-to-digest tutorial for anyone looking to write loops. There are more advanced topics available but this is a good article to get a feel for the site.

How to write the first for loop in R (<https://www.r-bloggers.com/how-to-write-the-first-for-loop-in-r/>)

Blog name: Edwin Chen (<http://blog.echen.me/>)

Edwin Chen

Pilot. Ex: math and linguistics at MIT, speech recognition at MSR, quant trading at Clarium, stats at Amazon, ads at Twitter, analytics at Dropbox, data science at Google.

I work on AI, human computation, and data.

hello@echen.me

Quora
Twitter
Github

Atom / RSS

Recent Posts

- Moving Beyond CTR: Better Recommendations Through Human Evaluation
- Propensity Modeling, Causal Inference, and Discovering Drivers of Growth
- Product Insights for Airbnb

Moving Beyond CTR: Better Recommendations Through Human Evaluation

Imagine you're building a recommendation algorithm for your new online site. How do you measure its quality, to make sure that it's sending users relevant and personalized content? Click-through rate may be your initial hope... but after a bit of thought, it's not clear that it's the best metric after all.

Take Google's search engine. In many cases, improving the quality of search results will *decrease* CTR! For example, the ideal scenario for queries like *When was Barack Obama born?* is that users never have to click, since the question should be answered on the page itself.

Or take Twitter, who one day might want to recommend you interesting tweets. Metrics like CTR, or even number of favorites and retweets, will probably optimize for showing quick one-liners and pictures of funny cats. But is a Reddit-like site what Twitter really wants to be? Twitter, for many people, started out as a news site, so users may prefer seeing links to deeper and more interesting content, even if they're less likely to click on each individual suggestion overall.

Or take eBay, who wants to help you find the products you want to buy. Is CTR a good measure? Perhaps not: more clicks may be an indication that you're having trouble finding what you're looking for. What about revenue? That might not be ideal either: from the user perspective, you want to make your purchases at the lowest possible price, and by optimizing for revenue, eBay may be turning you towards more expensive products that make you a less frequent customer in the long run.

And so on.

So on many online sites, it's unclear how to measure the quality of personalization and recommendations using metrics like CTR, or revenue, or dwell time, or whatever. What's an engineer to do?

Well, consider the fact that many of these are *relevance* algorithms. Google wants to show you relevant search results. Twitter wants to show you relevant tweets and ads. Netflix wants to recommend you relevant movies. LinkedIn wants to find you relevant people to follow. So **why, so often, do we never try to measure the relevance of our models?**

I'm a big fan of *man-in-the-machine* techniques, so to get around this problem, I'm going to talk about a **human evaluation** approach to measuring the performance of personalization and discovery products. In particular, I'll use the example of related book suggestions on Amazon as I walk through the rest of this post.

Amazon, and Moving Beyond Log-Based Metrics

(Let's continue motivating a bit why log-based metrics are often imperfect measures of relevance and quality, as this is an important but difficult-to-understand point.)

(<https://365datasience.com/wp-content/uploads/2018/11/Edwin-Chen-e1542113429663.jpg>)

Description: Edwin Chen takes his own experience in AI and data and writes some really interesting articles with wonderful insights. From Twitter to Amazon, Edwin has fascinating things to say on a multitude of topics, complete with easy to follow pictures and a very readable writing style.

Alexa Rank: 353.089

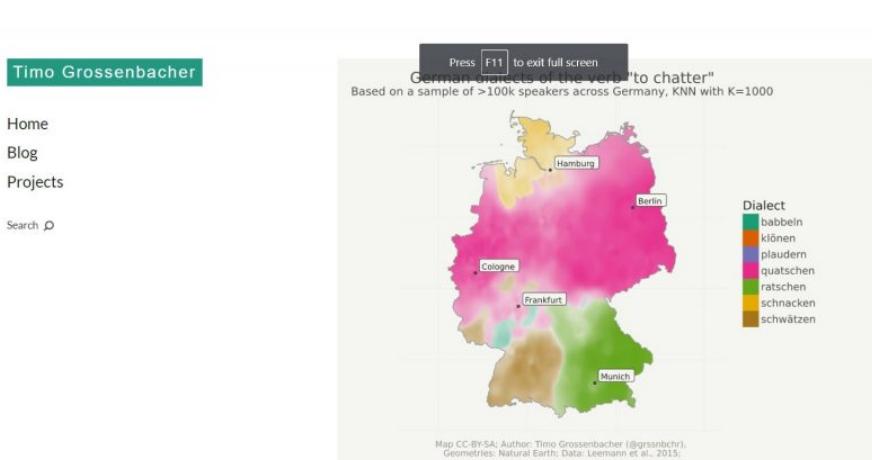
Post frequency: n/a

Article to check out: A scientific look at recommendation algorithms of popular online bookstores. Should these sites incorporate human evaluation? Read to find out.

Moving Beyond CTR: Better Recommendations Through Human Evaluation

(<http://blog.echen.me/2014/10/07/moving-beyond-ctr-better-recommendations-through-human-evaluation/>)

Blog name: Timo Grossenbacher
(<https://timogrossenbacher.ch/blog/>)



MARCH 17, 2018 / COMMENTS 2

Categorical spatial interpolation with R

How I used the kknn and ggplot2 packages together with some parallel computation to spatially interpolate several hundred thousand points.

<https://timogrossenbacher.ch/2018/03/categorical-spatial-interpolation-with-r/>

(<https://365datascience.com/wp-content/uploads/2018/11/Timo-Grossenbacher-e1542113559332.jpg>)

Description: Timo combines Geographic information, science and data journalism to create awesome visualisations. He's even written a book with over 80 maps called "Grüezi, Moin, Servus!". It's in German but luckily he shows you how he performed his studies on his blog. He prefers R, so if you lean toward that language and have a thing for geographics then check Timo's blog out.

Alexa Rank: 2,545,850

Post frequency: 1-2 per month

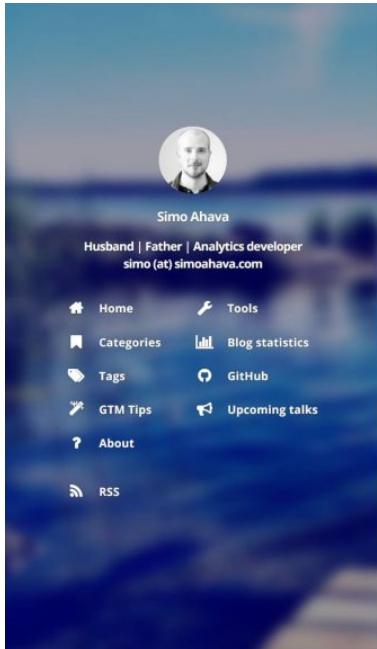
Article to check out: Timo's latest post on spatial interpolation is a great tutorial. And his studies involve dialects around Germany, so if you're into linguistics you'll get some extra enjoyment out of this.

Categorical spatial interpolation with R

(<https://timogrossenbacher.ch/2018/03/categorical-spatial-interpolation-with-r/>)

Blog name: Simo Ahava
(<https://www.simoahava.com/about-simo-ahava/>)





#GTMTips: Two Simple Data Model Tricks
August 3, 2018 in Analytics | 2 Comments

One of the more difficult concepts in Google Tag Manager is the data model. In essence, the data model is what Google Tag Manager uses to populate the Data Layer variable. You might be tempted to think that it's the same thing as the dataLayer array, but it's not. The data model is a representation of the keys and values you push into dataLayer. Whenever you push any key into dataLayer, GTM grabs this key and updates the corresponding key in its data model with the new value, or in the case of objects and arrays merges the old and the new value together.

[Continue reading](#)

#GTMTips: Get Position Index Of Visible Element
August 2, 2018 in Analytics | 3 Comments

It's time for a very simple #GTMTips article (I know, I always write that these are simple tips, but then they escalate into complex behemoths). Today, we'll cover a nifty trick you can use with the Element Visibility trigger in Google Tag Manager. This tip was inspired by a question from Eugen Potog in the Google Tag Manager Facebook group. The use case is that you have an Element Visibility trigger firing for a number of elements all sharing the same CSS selector.

[Continue reading](#)

(<https://365datascience.com/wp-content/uploads/2018/11/Simo-Ahava-e1542113513204.jpg>)

Description: Analytics is the name of the game here. Simo Ahava has been a development expert for Google analytics and is willing to help anyone who has questions about analytics or Google Tag Manager. He also dabbles in digital marketing and his blog is welcoming and well maintained. He does his best to simplify complex topics while writing thorough tutorials.

Alexa Rank: 120,426

Post frequency: 4-5 per month

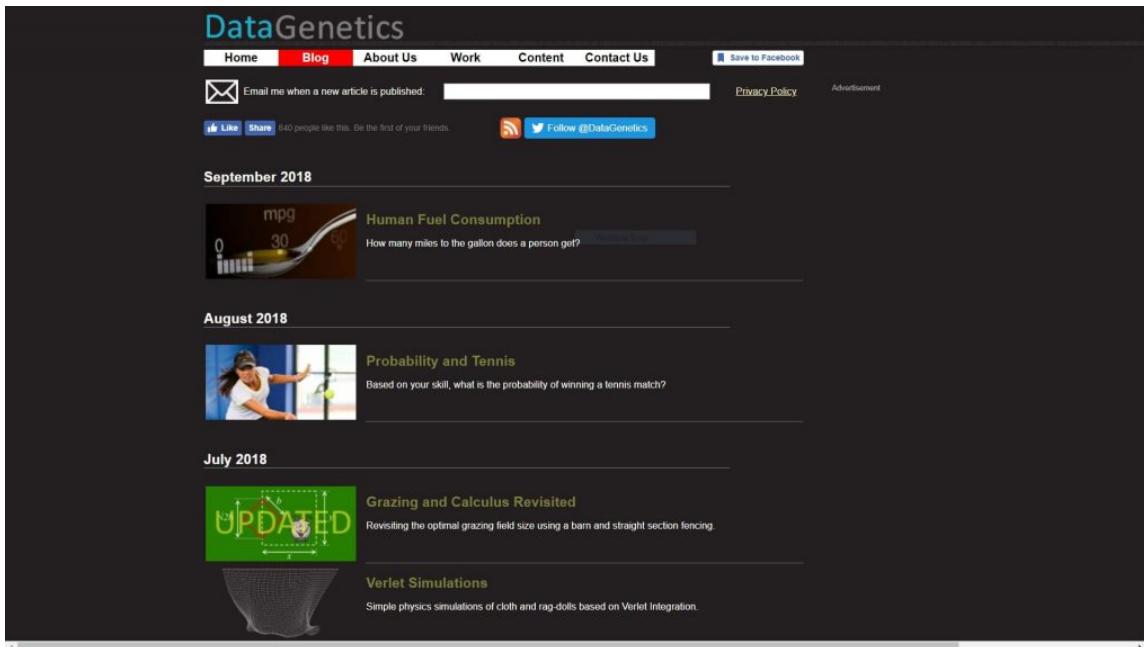
Article to check out: This in-depth tutorial on how to track browsing behaviour was inspired by a question from one of Simo's followers on twitter. This goes to show the sense of community on the blog. Written in collaboration with Jethro Nederhof, this is a must read for any google analytics enthusiast.

Track Browsing Behavior In Google Analytics

(<https://www.simoahava.com/analytics/track-browsing-behavior-in-google-analytics/>)

Blog name: DataGenetics (<http://www.datagenetics.com>)





(<https://365datascience.com/wp-content/uploads/2018/11/data-genetics-e1542113378544.jpg>)

Description: Nick Berry works at Facebook as a data scientist and is passionate about data privacy. His blog covers a range of topics within data and statistics. It is easy to read and you are more than likely to learn something new. If you want to know your chances of winning a tennis match or how to turn water into wine, then you had better check this out.

Alexa Rank: 224,998

Post frequency: 1-6 per month

Article to check out: A Statistical look at the probability of winning a tennis match. An interesting look at the game that goes into great detail.

Probability and Tennis (<http://datagenetics.com/blog/august12018/index.html>)

Blog name: Algobeans | Data Analytics Tutorials & Experiments for the Layman (<https://algobeans.com>)





(<https://365datasience.com/wp-content/uploads/2018/11/26-algobeans-e1542113618377.jpg>)

Description: As the name suggests, the focus of this blog is to give simplified data science tutorials for the newbie. Not for data science pros but if you're still trying to get to grips on some concepts of data science and statistics then this is for you. Well explained and presented rather nicely.

Alexa Rank: 823,831

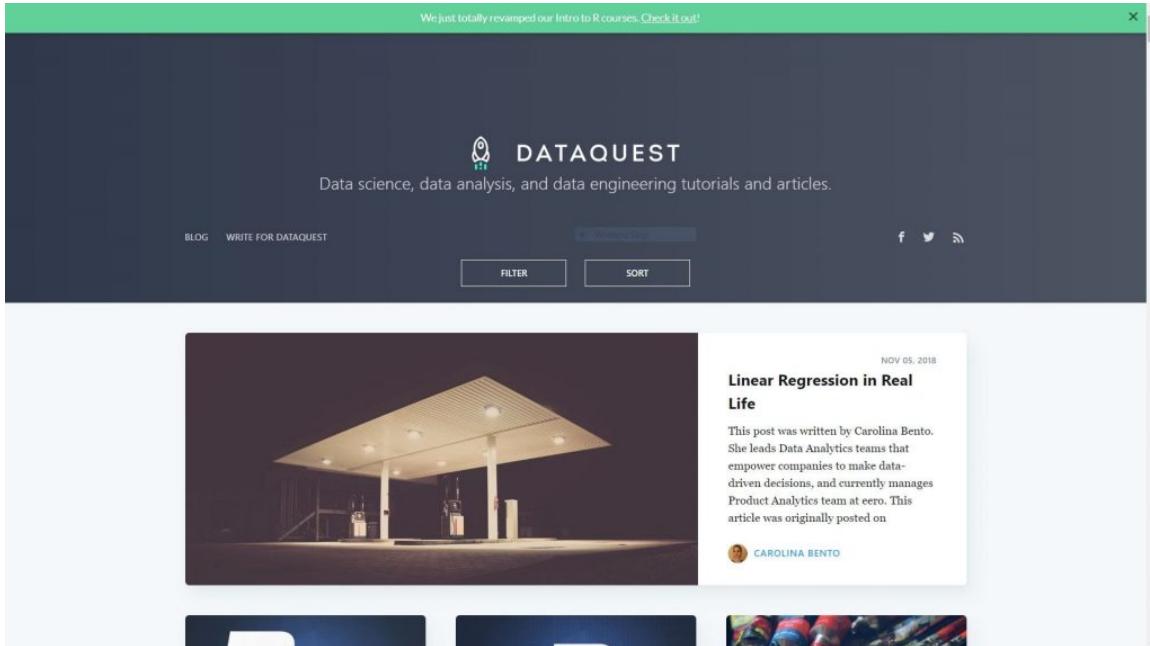
Post frequency: 1-3 per 6 months

Article to check out: A/B testing is useful in many situations. And these guys explain the concept very well.

LAYMAN'S GUIDE TO A/B TESTING (<https://algobeans.com/2017/07/19/laymans-guide-to-ab-testing/>)

Blog name: DataQuest (<https://www.dataquest.io/blog/>)





(<https://365datascience.com/wp-content/uploads/2018/11/Dataquest-min-e1542113415371.jpg>)

Description: Dataquest has a lovely, clean looking blog with plenty of interesting blog articles. They are mostly geared toward aspiring data scientists as that is their demographic. Their goal is to train people to become data scientists so you'll find plenty of tutorials and job hunting advice – practical stuff you can put to good use. They have plenty of guest writers but not just anyone it seems as all the blogs are well written and engaging. A great blog for anyone just getting into data science or looking to brush up their skills.

Alexa Rank: 24,864

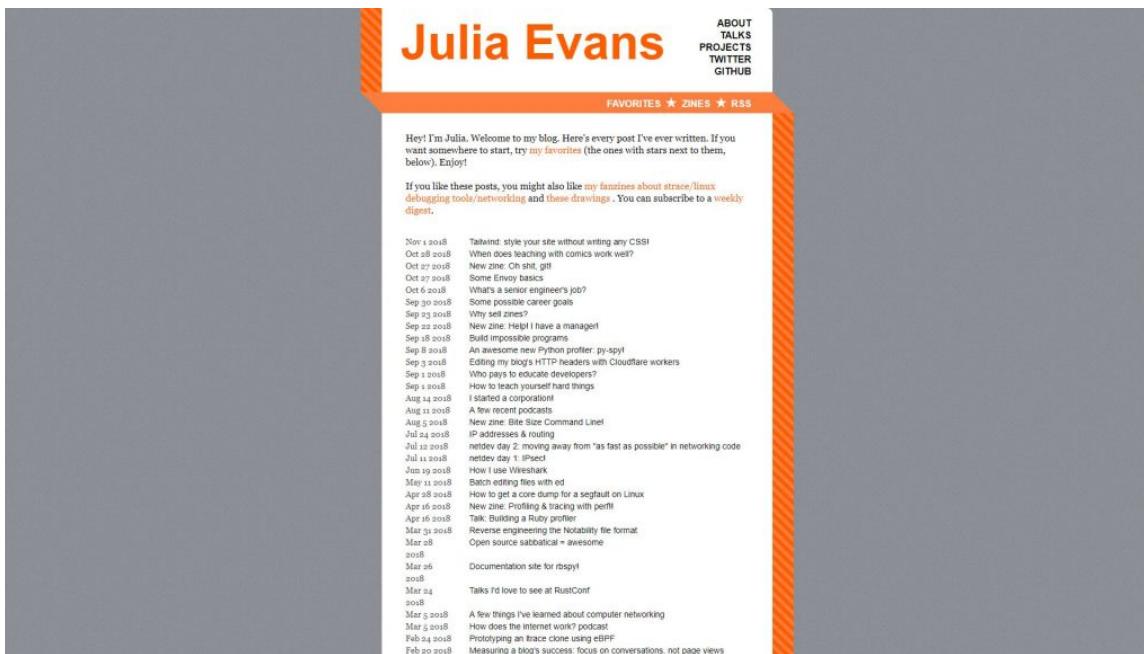
Post frequency: 2-3 per week

Article to check out: Python Vs R – of course, everyone has their opinion on which is better but who is right? Well, the best way to find out is to tackle the question like a data scientist – study the data. This is what this article does, it takes an unbiased look at a selection of functions and draws insightful conclusions that are sure to help anyone decide which program they would like use. A well-written and useful piece.

Python vs R: Head to Head Data Analysis (<https://www.dataquest.io/blog/python-vs-r/>)

Blog name: Julia Evans (<https://jvns.ca/>)





(<https://365datasience.com/wp-content/uploads/2018/11/Julia-Evans-min-e1542113448856.jpg>)

Description: Julia Evans is a programmer from Montreal. This blog is where she shares what she's learned recently. She has a huge collection of interesting posts that really show her passion for programming. Her writing is humble and invites a sense of discovering things together. She has open communication with her readers, which is great, and has also published paid and free informative programming zines which are very well made. A great blog for beginners and anyone who has an interest in programming.

Alexa Rank: 107,742

Post frequency: 3-4 per month

Article to check out: While her topics are wide-ranging it's her tutorials that show off her skills. This one on the basics of Envoy show how she puts her personality into her writing, in a way that invites you to continue reading and I am sure will entice you to read more of her blog and follow her journey.

Some Envoy basics (<https://jvns.ca/blog/2018/10/27/envoy-basics/>)

Blog name: A Blog From a Human-engineer-being
(<http://www.erogol.com/>)



The screenshot shows a blog post from 'A Blog From Human-engineer-being'. The post is titled 'Using WSL Linux on Windows 10 for Deep Learning Development.' and was published on July 4, 2018. The content discusses how WSL (Windows Subsystem for Linux) allows users to run Linux on Windows 10, enabling them to use favorite Linux tools like bash, zsh, and vim for development. It highlights the convenience of using GPU drivers and installing software like CUDA. The author also notes that while WSL is a great tool, it has limitations, such as not being able to run certain software directly. The sidebar includes sections for 'RECENT POSTS' and 'TAG CLOUD', along with a search bar and a 'SUBSCRIBE' button.

(<https://365datascience.com/wp-content/uploads/2018/11/29-Erogol-e1542113627521.jpg>)

Description: Although it won't win any awards for presentation, it certainly is a great place for anyone looking to expand their machine and deep learning knowledge. It is a little hard to find things but when you do, they're well-written guides on a plethora of topics. If you have a solid grasp of machine learning I think you can expand on that here.

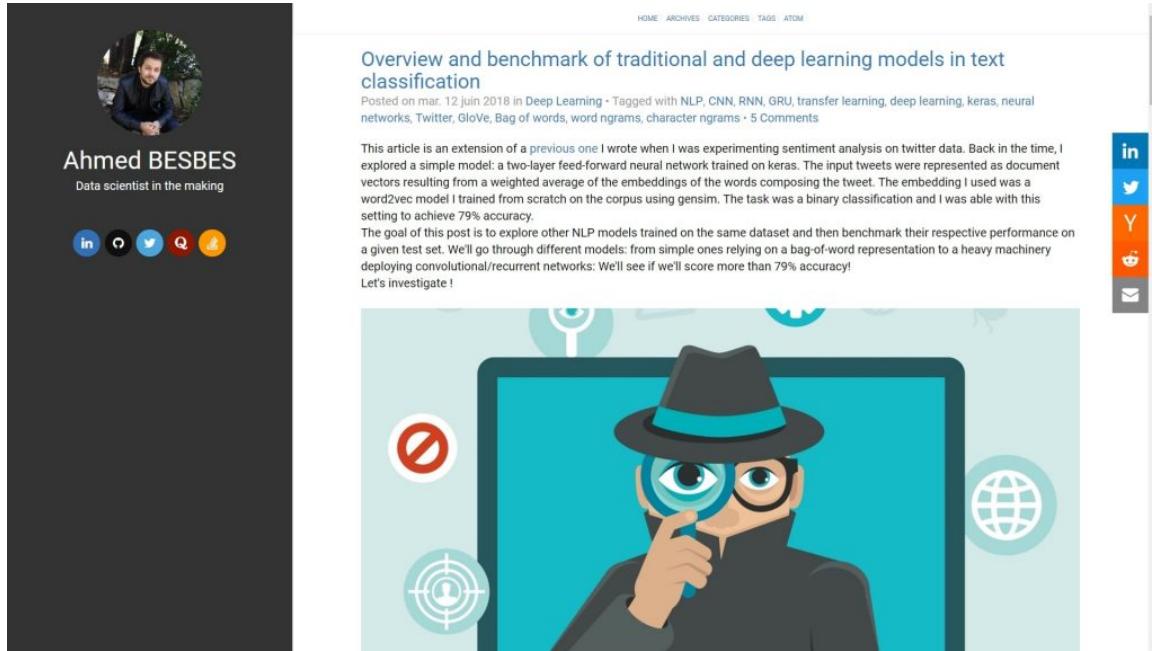
Alexa Rank: 609,158

Post frequency: 1- per 3 months

Article to check out: This history of machine learning is clearly the cornerstone post. It's a well written interesting piece that sets the stage for the rest of the blog.

Brief History of Machine Learning (<http://www.erogol.com/brief-history-machine-learning/>)

Blog name: Ahmed BESBES (<https://ahmedbesbes.com>)



A screenshot of a blog post by Ahmed BESBES. The post is titled "Overview and benchmark of traditional and deep learning models in text classification". It includes a bio section, social media links, and a detailed article about text classification models. A sidebar on the right features social sharing icons for LinkedIn, Twitter, YouTube, and others.

(<https://365datasience.com/wp-content/uploads/2018/11/47-Ahmed-BESBES-e1542113234676.jpg>)

Description: Ahmed is an up-and-coming data scientist. And I bet he'll be very successful. There are perhaps ten posts on his site that are based around what he has been learning. There is a KDnuggets silver award on there which is always a good sign of a quality post. A modest blog that is worth following.

Alexa Rank: 861,886

Post frequency: 1-2 per 6 months

Article to check out: A wonderful article that shows a real sense of community. Ahmed reached the top 4% in the Kaggle Titanic Challenge and rather than posting a simple bragging article, he explains to his readers how he did it so they can succeed too. Good Show Ahmed!

How to score 0.8134 in Titanic Kaggle Challenge (<https://ahmedbesbes.com/how-to-score-08134-in-titanic-kaggle-challenge.html>)

Blog name: Machine Learning Mastery
(<https://machinelearningmastery.com/blog/>)

The screenshot shows the homepage of [Machine Learning Mastery](https://machinelearningmastery.com). At the top, there's a navigation bar with links for Start Here, Blog, Topics, Ebooks, FAQ, About, and Contact. A search bar is also present. On the right side, there's a sidebar with a profile picture of Jason Brownlee and a welcome message: "Welcome to Machine Learning Mastery! Hi, I'm Jason Brownlee, PhD. I write tutorials to help developers (like you) get results with machine learning." Below this is a "POPULAR" section listing several articles with small thumbnail images and titles.

- How to Develop Multilayer Perceptron Models for Time Series Forecasting**
- How to Use the TimeseriesGenerator for Time Series Forecasting in Keras**
- A Gentle Introduction to LSTM Autoencoders**

Each article has a brief description, a "Continue Reading" link, and a timestamp. The sidebar also includes links for "Want more help?", "Get years of experience in a PDF.", and "Click to Discover the Ebook Catalog".

(<https://365datasience.com/wp-content/uploads/2018/11/Machine-Learning-MAstery-min-e1542113460390.jpg>)

Description: Jason Brownlee has a fascination with machine learning, and he wants to share his love for it with others. He has a very fleshed out blog with tons of tutorials for the machine learning enthusiast. His explanations are clear and extensive. Not for beginners but there are so many tutorials most will find something of value. Plus, he tells you right from the beginning what you will know after reading the tutorial – which is handy. Check this out if you want to learn something new.

Alexa Rank: 12,523

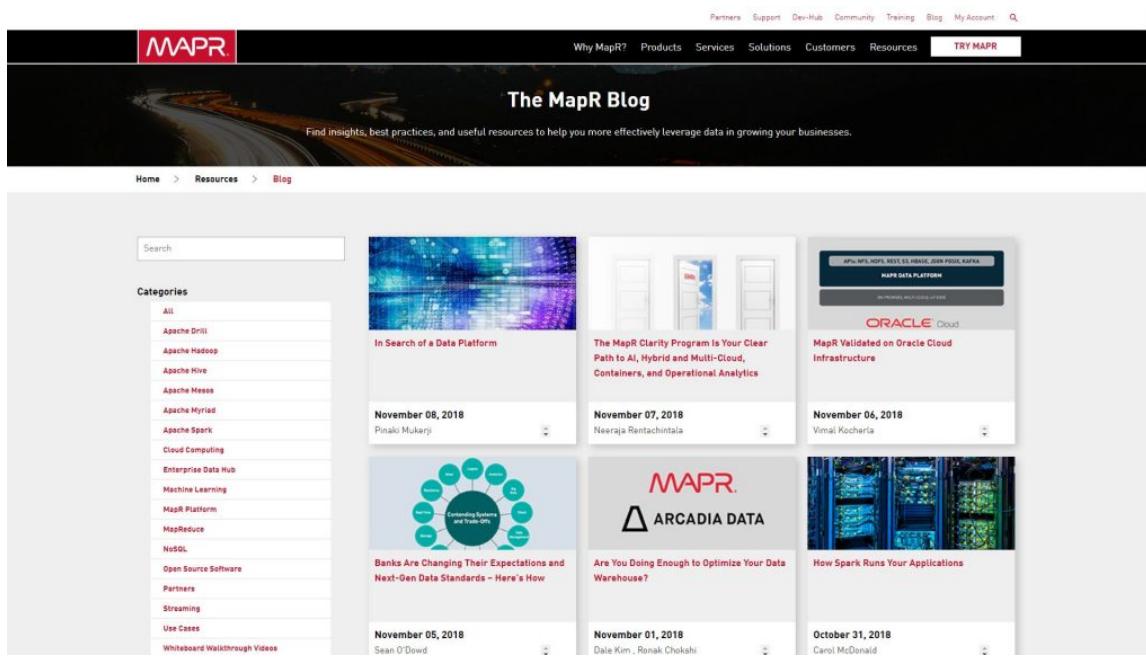
Post frequency: 1-2 per week

Article to check out: His tutorials mostly follow the same style, so after checking this out you'll know what to expect with his others. This is about how to code the student's t-test. Considering how widely used it is in statistics it's a good place to start and see if you like Jason's style.

How to Code the Student's t-Test from Scratch in Python

(<https://machinelearningmastery.com/how-to-code-the-students-t-test-from-scratch-in-python/>)

Blog name: MAPR (<https://mapr.com/blog/>)



(<https://365datascience.com/wp-content/uploads/2018/11/MAPR-min-e1542113467680.jpg>)

Description: MAPR has an extensive blog, covering loads of topics – cloud computing, open source software, and apache spark to name but a few. They have a ton of expert contributors and their posts are very inviting. As MAPR say, they are proud to enable Big Data experts to share their industry knowledge. Come learn from the best, and share the knowledge. And that's exactly what they do.

Alexa Rank: 70,269

Post frequency: 2-3 per week

Article to check out: The posts on this blog assume varying levels of experience. This one on AI and ML is a more introductory article which shows off this pleasant blog and is a great place to get started.

Artificial Intelligence and Machine Learning: What Are They and Why Are They Important?
(<https://mapr.com/blog/artificial-intelligence-and-machine-learning-what-are-they-and-why-are-they-important/>)

Blog name: PYImageSearch
(<https://www.pyimagesearch.com/>)

Creating GIFs with OpenCV
by Adrian Rosebrock on November 5, 2018 in [Faces](#), [Tutorials](#)

In this tutorial, you will learn how to create animated GIFs using OpenCV, Python, and ImageMagick. You'll then combine all of these techniques to build a meme generator with OpenCV! We all need a good laugh every now and then. And perhaps the best way to find LOLs is from memes. Some of my favorite [...] Continue Reading 18

Multi-object tracking with dlib
by Adrian Rosebrock on October 29, 2018 in [dlib](#), [Object Tracking](#), [Tutorials](#)

In this tutorial, you will learn how to use the dlib library to efficiently track multiple objects in real-time video. So far in this series on object tracking we have learned how to: Track single objects with OpenCV Track multiple objects utilizing OpenCV Perform single object tracking with dlib Track and count people entering a [...] Continue Reading 38

Object tracking with dlib
by Adrian Rosebrock on October 22, 2018 in [dlib](#), [Object Tracking](#), [Tutorials](#)

This tutorial will teach you how to perform object tracking using dlib and Python. After reading today's blog post you will be able to track objects in real-time video with dlib. A couple months ago we discussed centroid tracking, a simple, yet effective method to (1) assign unique IDs to each object in an image [...] Continue Reading 19

Deep learning, hydroponics, and medical marijuana
by Adrian Rosebrock on October 15, 2018 in [Deep Learning](#), [Keras](#), [Tutorials](#)

You're interested in deep learning and computer vision, but you don't know how to get started? Let me help. My new book will teach you all you need to know about deep learning. CLICK HERE TO MASTER DEEP LEARNING

Get your FREE 17 page Computer Vision, OpenCV, and Deep Learning Resource Guide PDF - Inside you'll find my hand-picked tutorials, books, courses, and libraries to help you master CV and DL. Download for Free!

Deep Learning for Computer Vision with Python Book -- OUT NOW!

DEEP LEARNING FOR COMPUTER VISION WITH PYTHON

Dr. Matthew Nielsen

Your interested in deep learning and computer vision, but you don't know how to get started? Let me help. My new book will teach you all you need to know about deep learning.

CLICK HERE TO MASTER DEEP LEARNING

Free 17-day crash course on Computer Vision, OpenCV, and Deep Learning

(<https://365datascience.com/wp-content/uploads/2018/11/pyimagesearch-min-e1542113494420.jpg>)

Description: Adrian Rosebrock specializes in image search engines and covers a bunch of skills – object tracking, face recognition, motion detection etc. His tutorials are informative, clear and offer a nice personal touch. An especially nice blog if images are your interest.

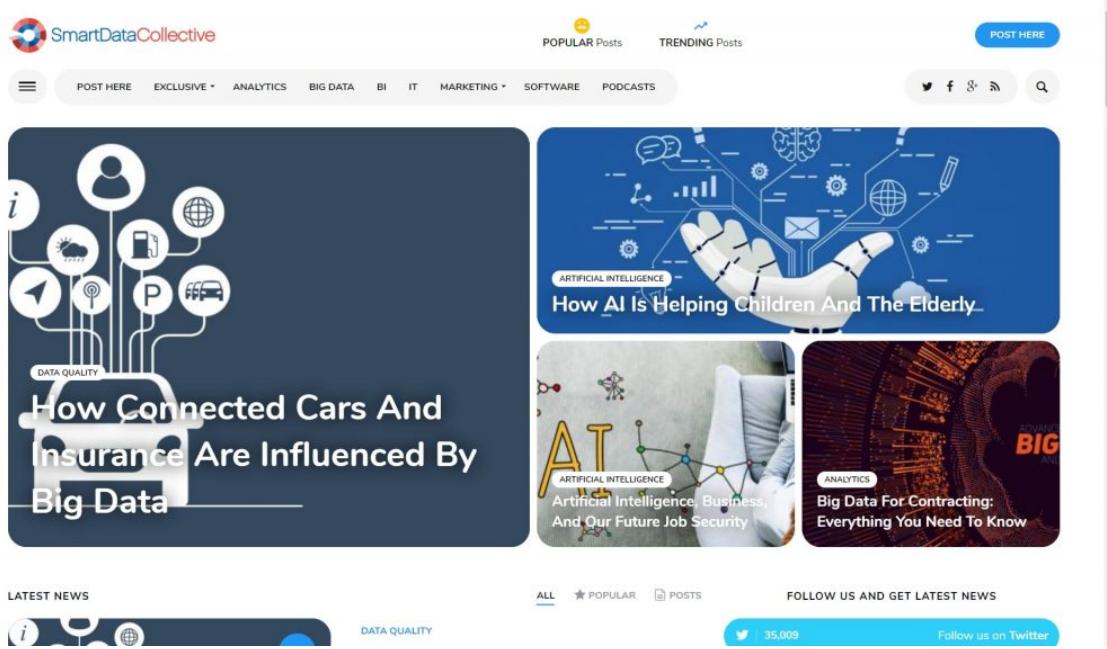
Alexa Rank: 20,005

Post frequency: 1-2 per week

Article to check out: In this tutorial, Adrian teaches you how to create animated GIFs with OpenCV. It's an entertaining and educational article. And if you don't like GIFs...Deal with it! (You'll get that joke if you read the post)

Creating GIFs with OpenCV (<https://www.pyimagesearch.com/2018/11/05/creating-gifs-with-opencv/>)

Blog name: Smart Data Collective
(<https://www.smartdatacollective.com/>)



(<https://365datasience.com/wp-content/uploads/2018/11/Smart-Data-Collective-e1542113526447.jpg>)

Description: A well-designed blog with plenty of contributors covering a range of topics. AI, analytics, blockchain, big data and more. Articles for the experienced enthusiast as well as the casual beginner. There is something here for everyone.

Alexa Rank: 307,887

Post frequency: Daily

Article to check out: The brighter side of AI. A light read on how AI is making positive changes for humanity.

How AI Is Helping Children And The Elderly (<https://www.smartdatacollective.com/how-ai-helping-children-elderly/>)

Blog name: Amazon AWS AI Blog
(<https://aws.amazon.com/blogs/ai/>)

#256: Using IoT For Smart Cities | July 29, 2018
#255: Building Deployment Pipelines - Interview with LEAP Dev | July 22, 2018
#254: Understanding Amazon Comprehend | July 15, 2018

Latest Posts

Maintaining Transport Layer Security all the way to your container part 2: Using AWS Certificate Manager Private Certificate Authority

<https://aws.amazon.com/podcasts/aws-podcast/TfI=1>

(<https://365datascience.com/wp-content/uploads/2018/11/35-AWS-e1542113633316.jpg>)

Description: The link here takes you to Amazon's AI blog, but there are over 25 more categories to choose from (although not all data science related). You will no doubt find something of interest. Most posts revolve around some Amazon project but with a plethora of expert guest bloggers, you are going to learn something for sure.

Alexa Rank: 10 (this is Amazon's Alexa ranking)

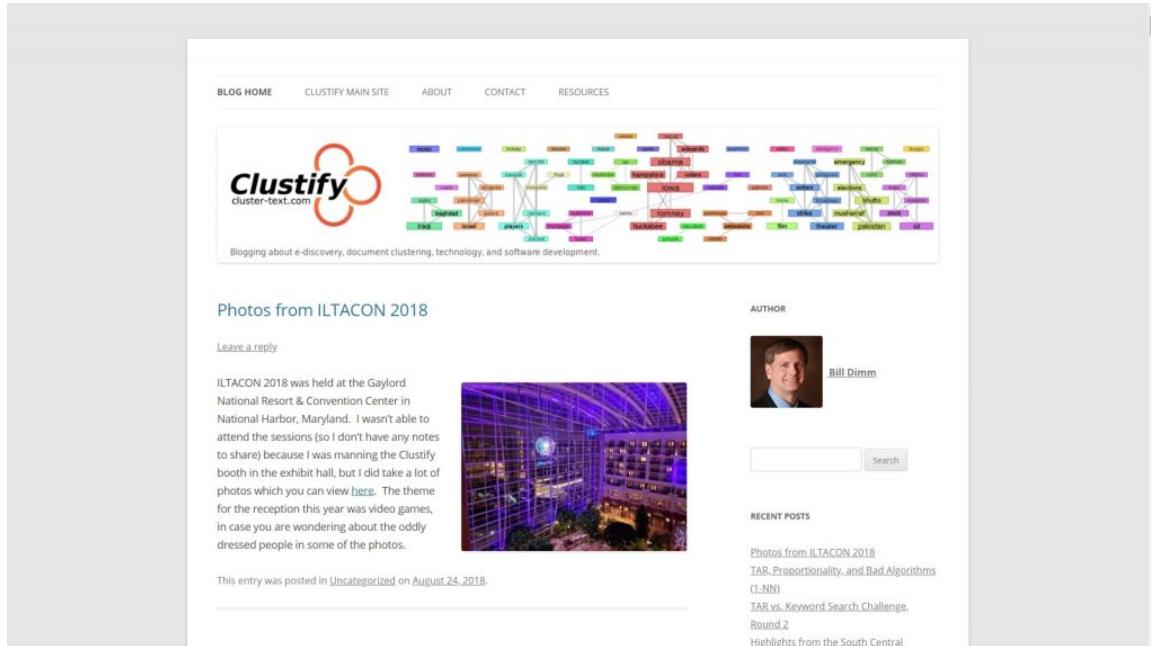
Post frequency: Daily

Article to check out: This tutorial on how to make a QnA bot using Alexa looks like a lot of fun.

Create a Question and Answer Bot with Amazon Lex and Amazon Alexa

(<https://aws.amazon.com/blogs/machine-learning/creating-a-question-and-answer-bot-with-amazon-lex-and-amazon-alexa/>)

Blog name: Clustify (<https://blog.cluster-text.com/>)



(<https://365datascience.com/wp-content/uploads/2018/11/Clustify-e1542113347377.jpg>)

Description: Clustify shares tips and insights about e-discovery, document clustering, predictive coding, information retrieval, computers, and software development. They keep up to date with conferences and retreats surrounding these areas. So if that is of interest to you then these articles will be right up your street.

Alexa Rank: 4,841,767

Post frequency: It varies

Article to check out: A post about a keyword search challenge during the NorCal eDiscovery & IG Retreat. An interesting experiment that sounds like it would have been great fun to participate in.

TAR vs. Keyword Search Challenge (<https://blog.cluster-text.com/2018/05/08/tar-vs-keyword-search-challenge/>)

Blog name: KritiKal Solutions Blog
(<http://www.kritikalsolutions.com/blog/>)





(<https://365datascience.com/wp-content/uploads/2018/11/KritiKal-e1542113454667.jpg>)

Description: KritiKal Solutions is a technology design house specializing in Product Development, R&D and Innovation. Their blog is heavy on machine learning and AI as well as their implications in the modern world. Written in a comprehensible manner and a well-presented design this is good for some casual reading.

Alexa Rank: 886,859

Post frequency: 1-2 per month

Article to check out: A few examples of where AI is being used in the real world. A nice reminder that whether we like it or not AI is already playing a huge part in our lives.

Top 5 Real-World Artificial Intelligence Examples

(<http://www.kritikalsolutions.com/blog/real-world-ai-examples/>)

Blog name: Freakonometrics
(<https://freakonometrics.hypotheses.org/>)

The screenshot shows the homepage of the Freakonomics blog. The header includes links for OpenEdition Books, Journals, Calenda, Hypotheses, Newsletters and alerts, and OpenEdition Premium. A search bar is at the top right. The main navigation menu on the left includes HOME, ABOUT, CONTACT, COURSES, and PUBLICATIONS. A sidebar on the left features a portrait of Arthur Charpentier and a bio in French. The main content area has a green header 'SUMMER BREAK' with a date (31/07/2018), author (ARTHUR CHARPENTIER), and a 'LEAVE A COMMENT' link. Below the header is a text message about being off for a few days due to moving. A large image of a truck heavily loaded with trash is displayed. A caption below the image says 'I will be back in September....'. On the right side, there's a sidebar titled 'freakonometrics' with various social media and academic links like email, Twitter, LinkedIn, RSS, Slideshare, Google Scholar, ResearchGate, DBLP, ORCID, CrossRef, RePEc, Scopus, Researcher ID, Academia, DZone, RBloggers, StatsBlog, and Web Page. It also includes a GitHub link: 'freakonometrics.github.io'. At the bottom right, there's a note in a red-bordered box: 'Because of non-updated links following the blog migration, you might have trouble with missing references.'

(<https://365datascience.com/wp-content/uploads/2018/11/38-freakonometrics-e1542113225675.jpg>)

Description: An academic blog by Professor Arthur Charpentier is an economist and mathematician. A huge back-catalogue of posts for just one person with some fascinating mathematical takes on a range of topics from the average lifespan to the number of friends you have. He also covers the lectures he gives and attends. Also, if you're a French speaker there are a few posts in French you can have a look at.

Alexa Rank: 28,764

Post frequency: 3-4 per week

Article to check out: The post on why on average people have fewer friends than their friends. Proved with maths.

GAME OF FRIENDSHIP PARADOX (<https://freakonometrics.hypotheses.org/53310>)

Blog name: DBMS2 (<http://www.dbms2.com/>)

The screenshot shows the DBMS2 blog homepage. At the top, there's a blue header with the DBMS2 logo and navigation links for Home, About, Contact, and Feeds. Below the header, a sidebar on the right contains links for Monash Research blogs, categories like 'Recent posts' and 'Categories', and sections for User consulting, Vendor advisory, and Monash Research highlights. The main content area features an article titled 'New legal limits on surveillance in the US'. The article discusses the Carpenter v. United States Supreme Court case and its implications for privacy. It includes several bullet points and a note about the Katz test and third-party doctrine. At the bottom of the article, there are links for 'Read more', 'Categories: GIS and geospatial, Surveillance and privacy', and 'Leave a Comment'.

(<https://365datascience.com/wp-content/uploads/2018/11/70-DBMS2-e1542113279431.jpg>)

Description: A general analytics and database management blog with a huge number of categories. Easy to navigate for what you want to find and some well-written articles. Some more casual than others. You'll find a number of articles focused on politics if that is of interest to you. There are even a few research blogs thrown in for good measure.

Alexa Rank: 2,480,873

Post frequency: 1-2 per month

Article to check out: One of the politically oriented posts will give you a sense of the writing style. Interesting read.

Technology implications of political trends

(<http://www.dbms2.com/2018/05/20/technology-implications-of-political-trends/>)

Blog name: Datafloq (<https://datafloq.com/read/>)

The screenshot shows the Datafloq homepage with a navigation bar at the top. Below the navigation is a search bar and a grid of four blog post cards. Each card includes a thumbnail image, the post title, a brief description, a 'READ MORE' button, and publication details (date and views). A cookie consent banner is visible at the bottom.

Big Data and IT Infrastructure: Analyzing Connections to Boost Enterprise...
The modern IT enterprise is comprised of myriad connection points. Here is how leaders can leverage big data to safeguard their systems.
[READ MORE](#)
03-09-2018 | 112 views

How can using big data improve your marketing?
Big data in marketing is leading to big changes in targeting, personalization, and other tactics.
[READ MORE](#)
03-09-2018 | 59 views

The IoT Has Become a Lifestyle Choice For Better Living Tomorrow
IoT has the potential to transform our society, economy and how we live our lives. Let's find how IoT has become a lifestyle choice for better living.
[READ MORE](#)
31-08-2018 | 174 views

How Data Visualization can Charge up Your Business Strategy
Data visuals identify your business needs and deliver relevant insights. Visuals should be appropriate enough to inspire and initiate apt action.
[READ MORE](#)
31-08-2018 | 320 views

In order to optimize the website and to continuously improve Datafloq, we use cookies. For more information click [here](#).
 Necessary
 Marketing
[Okay, I got it](#)

(<https://365datascience.com/wp-content/uploads/2018/11/Datafloq-e1542113402139.jpg>)

Description: Blockchain, VR, AI, IoT, Big data plus more covered here in some interesting and insightful articles. Opinion pieces on current trends seem to be the main bulk of articles here so if that's your thing, go right ahead.

Alexa Rank: 149,666

Post frequency: Daily

Article to check out: Blockchain is no doubt growing exponentially and it's affecting all walks of life. This article looks at the potential impact it could have on the film industry.

Will Blockchain become a game-changer for the film industry?

(<https://datafloq.com/read/will-blockchain-become-game-changer-film-industry/5433>)

Blog name: Cortana Intelligence and Machine Learning Blog
(<https://blogs.technet.microsoft.com/machinelearning/>)

(<https://365datascience.com/wp-content/uploads/2018/11/Microsoft-mini-1542113475382.jpg>)

Description: A machine learning blog by the developers of Cortana must hold some great information right? Well, it does, some great in-depth articles on machine and deep learning. And if you're a fan of Microsoft you'll definitely find some useful posts here.

Alexa Rank: 42 (Microsoft's Alexa rank)

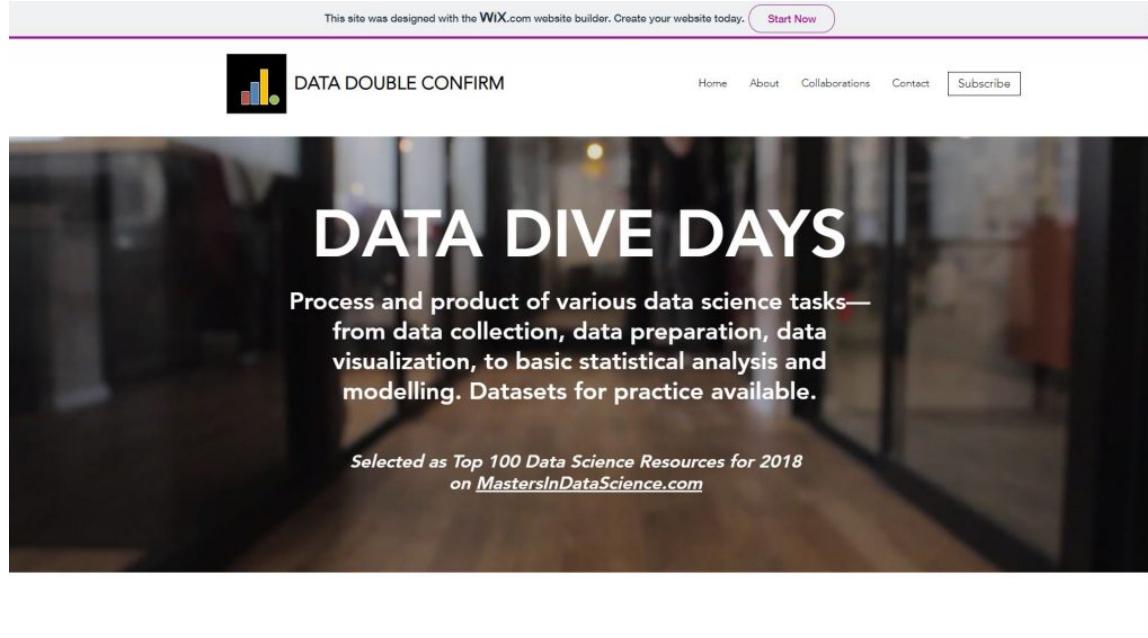
Post frequency: 1-2 per week

Article to check out: Using machine learning to match a persons clothing with shop images holds a lot of potential for advertising in the fashion industry. This article shows how it can be done.

Using Siamese Networks and Pre-Trained Convolutional Neural Networks (CNNs) for Fashion Similarity Matching

(<https://blogs.technet.microsoft.com/machinelearning/2018/07/10/how-to-use-siamese-network-and-pre-trained-cnns-for-fashion-similarity-matching/>)

Blog name: Data Double Confirm
(<https://projectosyo.wixsite.com/datadoubleconfirm>)



(<https://365datascience.com/wp-content/uploads/2018/11/Data-Double-Confirm-min-e1542113370384.jpg>) **Description:** A personal blog detailing the learning process of Hui Xiang Chua. With a B.Sc.(Hons) in statistics, an M.Sc in Business Analytics and a career as a research analyst, Hui works with a multitude of tools, such as R, Python, Tableau, SAS and SPSS. This means that if you are interested in developing your skills in any of these areas then following her blog is a good idea.

Alexa Rank: 507 (Wix's Alexa rank)

Post frequency: 4-5 per month

Article to check out: Everybody loves a good pie, but perhaps don't have the same love for pie charts. In this article, Hui opens up the discussion on pie charts and leads on to some more detailed articles on the topic if you're interested.

Pie Chart or Not (<https://projectosyo.wixsite.com/datadoubleconfirm/single-post/2018/07/31/Pie-Chart-or-Not>)

Blog name: Alex Perrier (<http://alexisperrier.com/>)





Hi, I'm Alexis Perrier, a Data Science consultant.
I help companies large and small, profit from machine learning.
I also teach and write about all things Data Science, from linear regression to deep learning with a penchant for NLP.
I work between Washington DC where I live
and Paris where I'm from.

Follow me on twitter @alexip.
Let's connect on LinkedIn or AngelList.

Recent posts

Feb 21, 2018

[Reduce GPU costs with startup scripts on the Google Cloud Engine](#)

Reduce GPU costs with on demand instances and startup scripts This post is about leveraging on demand capabilities of costly virtual instances on the Google Cloud Engine using startup scripts. Deep Learning is expensive Here's the situation: You're working on some large dataset, and you feel the irresistible urge to...

Feb 6, 2018

[iPhone addiction? Get a grip!](#)

Is it a DNA thing? My wife has a super power! She is totally immune to the constant nagging of her iPhone. She has an amazing ability to resist checking her emails every 5 minutes, texting back on the spot and playing the whack-a-notification game all day long. Maybe it's...

(<https://365datasience.com/wp-content/uploads/2018/11/67-Alex-Perrier-e1542113259251.jpg>)

Description: Alex Perrier is a data science consultant and writes some nice articles on Gsutil and Amazon's machine learning. Definitely something to be learned here and you get a sense of his personality. Seems like a nice guy.

Alexa Rank: 4,935,199

Post frequency: 1-2 per 3 months

Article to check out: A departure from data science but a fun blog on phone addiction that gives you a nice idea of Alex's writing.

iPhone addiction? Get a grip!

(https://alexisperrier.com/life/2018/02/06/control_iphone_addiction.html)

Blog name: Data Blogger (<https://www.data-blogger.com/>)



The screenshot shows a blog post titled "Scale out your Pandas DataFrame operations using Dask". The post was published on August 5, 2018, and has 0 comments. The content discusses the limitations of the apply method in Pandas and how Dask can be used for parallelization. Below the post is a social sharing section with icons for Facebook, Google+, LinkedIn, Twitter, and others. A "READ MORE" button is visible. The sidebar on the right includes a search bar, a newsletter subscription form, and sections for "Editors' Picks" and "Categories" which list various topics like Do-it-Yourself, Services, Technology, Web Technology, Book reviews, and Cryptocurrencies.

(<https://365datascience.com/wp-content/uploads/2018/11/Data-Blogger-min-e1542113356772.jpg>)

Description: Data Blogger has some great DIY articles which set this blog apart. Of course, there are articles on other topics too, mathematics, data science and software science to name a few. But if you want to pick up some skills, in python especially, then this is the place to go.

Alexa Rank: 824,153

Post frequency: 1-2 per month

Article to check out: Learn Python in one day, who could say no to that? In this 'how to' Kevin Jacob takes you through Python basics to get you started (of course it then links to further tutorials).

Python Tutorial: Learn Python in one Day (<https://www.data-blogger.com/2017/11/28/python-tutorial-learn-python-one-day/>)

Blog name: INSIGHT
(<https://blog.insightdatascience.com/>)

The screenshot shows the homepage of the Insight Data Science blog. At the top, there's a banner with the text "INSIGHT Your bridge to careers in data". Below the banner, there's a navigation bar with links to "ABOUT INSIGHT", "DATA SCIENCE", "DATA ENGINEERING", "HEALTH DATA", "AI", "DATA PM", "DEVOPS", and social media icons for Twitter, Facebook, and a "Follow" button. Two blog posts are visible:

- VerbiAge: Using NLP to help writers craft age-specific writing** by Sameeran Kunche (Jul 31). This post features a small image of colorful sticks.
- Seattle Insight Data Science: One Year Anniversary!** by Carina Martin (Jul 19). This post features a group photo of people.

On the right side of the page, there's a sidebar with a "Insight Data" logo, a brief description of the "Insight Fellows Program - Your bridge to a career in data", a "More information" link, a "FOLLOWERS" section showing 9.93K, and a "ELSEWHERE" section with links to email, Twitter, and Facebook.

(<https://365datasience.com/wp-content/uploads/2018/11/INSIGHT-min-e1542113442775.jpg>)

Description: Insight is a blog where loads of contributors write to give you, you guessed it, insights on all things data science. The company's goal is to help others progress in their data science careers, so expect plenty of helpful things in the blog for that purpose.

Alexa Rank: 184,919

Post frequency: 1-3 per week

Article to check out: Is there a formula for love? Issac Perron and the company worked with 'Happy Couple' and tried to predict relationship health from over 8,000 participants. Check out this article to see how that went.

Data science for couples: Relationship health predicted by how well you know each other, not how alike you are (<https://blog.insightdatasience.com/data-science-for-couples-relationship-health-predicted-by-how-well-you-know-each-other-not-how-bcfab5f9ca16>)

Blog name: Simply Stats (<https://simplystatistics.org/>)

The screenshot shows the homepage of the Simply Statistics blog. On the left is a sidebar with links to Home, About, Archive, Conferences, Courses, Interviews, Contributing, RSS, Twitter, GitHub, and a link to the 2011-2012 All rights reserved page. The main content area features three articles:

- The role of academia in data science education** by Rafael Irizarry (2016/11/01). A statistics blog by Rafa Irizarry, Roger Peng, and Jeff Leek. The post discusses the role of universities in training the data science workforce.
- Guest Post: Galin Jones on criteria for promotion and tenure in (bio)statistics departments** by Rafael Irizarry (2018/10/11). Editor's Note: I attended an ASA Chair's meeting and spoke about ways we could support junior faculty in data science. After giving my talk Galin Jones, Professor and Director of Statistics at University of Minnesota, and I had an interesting conversation about how they had changed their promotion criteria in response to a faculty candidate being unique. I asked him to write about his experience and he kindly contributed the following post.
- The economic consequences of MOOCs** by Jeff Leek (2018/10/06). Check out our new paper on the relationship between MOOC completion and economic outcomes! Last Monday we launched our Chromebook Data Science Program so that anyone with an internet connection, a web browser, and the ability to read and follow instructions could become a data scientist. Why did we launch another MOOC program? Aren't MOOCs dead? Well we didn't think so :). We have been pretty excited about MOOCs for a while now and now run five different MOOC programs through the Johns Hopkins Data Science Lab.

(<https://365datascience.com/wp-content/uploads/2018/11/Simply-stats-min-e1542113519781.jpg>)

Description: Simply Stats is a blog from three enthusiastic biostatistics professors who post about ideas they find interesting. They are pumped about the recent abundance of data and how statisticians are now scientists. If stats is your thing and you want to read 3 people write about there passion while sharing some fresh ideas surrounding statistics then this is the place.

Alexa Rank: 270,251

Post frequency: 3-2 per month

Article to check out: The role of academia in data science is a hot topic right now. This article is Rafael's take on how the two are connecting right now. A good read for anyone interested, Rafael opens up the conversation so feel free to leave a comment for some respectful debate.

The role of academia in data science education

(<https://simplystatistics.org/2018/11/01/the-role-of-academia-in-data-science-education/>)

Blog name: Data Camp
(<https://www.datacamp.com/community/blog>)

The screenshot shows the DataCamp community blog page. On the left, there's a sidebar with links to 'Community' (News, Tutorials, Cheat Sheets, Open Courses, Podcast - DataFramed, Chat), 'DataCamp' (Official Blog, Tech Thoughts), and social media links (Subscribe to RSS, Facebook, Twitter, LinkedIn, YouTube). The main area has a search bar and a 'Search' button. Below that is a 'ALL TAGS-' dropdown. A 'Official Blog' button is highlighted in green. The main content area displays five blog posts:

- Peter Bull discusses the importance of human-centered design in data science.** (Learning Data Science) by Hugo Bowne-Anderson (November 5th, 2018) with 4 likes.
- New Course: Analyzing Election and Polling Data in R** (R Programming) by Ryan Sheehy (November 6th, 2018) with 6 likes.
- Arnaub Chatterjee discusses artificial intelligence (AI) and machine learning (ML) in healthcare.** (Learning Data Science) by Hugo Bowne-Anderson (October 29th, 2018) with 8 likes.
- How DataCamp Handles Course Quality** (DataCamp) by Richie Cotton (October 25th, 2018) with 9 likes.
- Cassie Kozyrkov discusses decision making and decision intelligence!** (DataFramed) with 1 like.

(<https://365datasience.com/wp-content/uploads/2018/11/Data-Camp-mini-e1542113363898.jpg>)

Description: Another heavy hitter in the data science world, DataCamp has plenty of articles on everything analytical as well as updates on the company themselves. If you're looking to keep your skills up to date this is a great place to do that. It also includes interviews with players on the data science field and current events. A nice all-around blog that doesn't get too heavy (unless you want it too).

Alexa Rank: 4163

Post frequency: 1-2 per week

Article to check out: Andrew Gelman is the author of a blog we've featured on this list. Why not get to know him a little better with an entertainingly constructed interview by Hugo Bowne-Anderson. Statistics, data science, polling, and election forecasting are the topics up for discussion. Great stuff.

Andrew Gelman discusses election forecasting and polling. (Transcript)

(<https://www.datacamp.com/community/blog/election-forecasting-polling>)

Blog name: Dimensionless
(<https://dimensionless.in/blog/>)

The screenshot shows the homepage of 365 Data Science. At the top, there's a header with contact information (9923170071 / 8108094992, info@dimensionless.in) and social media links. Below the header, there's a navigation bar with links to HOME, COURSES, BLOG (which is highlighted in blue), CONTACTUS, and CORPORATETRAINING. A download icon is also present.

The main content area features several article cards:

- Data Science in Esports** by Kartik Singh | Nov 12, 2018 | Data Science. This article discusses the evolution of esports from a hobby to a professional sport.
- Top 5 Trends in Data Science** by Kartik Singh | Nov 9, 2018 | Data Science. This article lists five trends in data science, including its impact on business.
- Top 5 Trends in Data Science** by Kartik Singh | November 9, 2018 | Data Science. This article provides an introduction to data science and its trends.
- Top Data Science Hacks** by Kartik Singh | Nov 5, 2018 | Data Science. This article offers practical tips and hacks for data science students.
- Spam Detection with Natural Language Processing – Part 3** by Kartik Singh | Nov 5, 2018 | Data Science. This article focuses on spam detection using NLP techniques.

On the right side of the page, there's a sidebar with links to USEFUL LINKS (About Us, Blog, Contact Us, Terms & Conditions, Privacy Policy) and a Contact Us form with fields for Name and Email Address.

(<https://365datascience.com/wp-content/uploads/2018/11/Dimensionless-min-e1542113422115.jpg>)

Description: Dimensionless was born out of the fact that educational institutions are lagging behind when it comes to data science. They aim to be the solution to that and their blog is part of that vision. With many useful articles on making it in the data science field, it's one of the best places an aspiring data scientist should go.

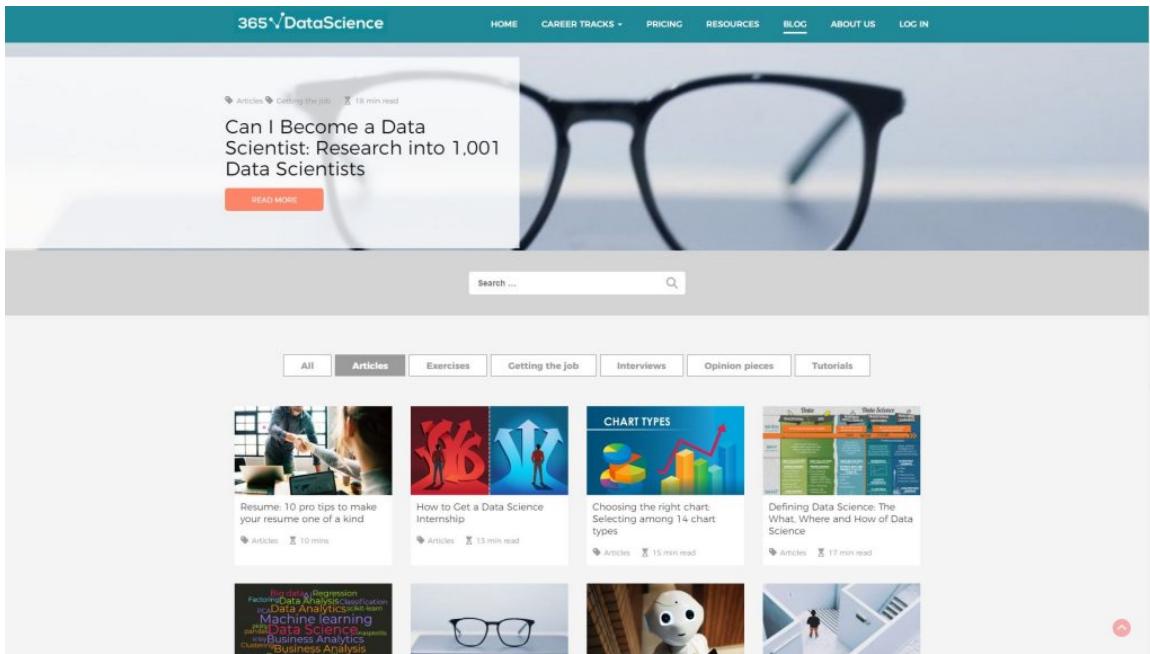
Alexa Rank: 784,263

Post frequency: 1-2 per week

Article to check out: Who wouldn't love some data science hacks, especially written by someone dedicated to helping students find their way through data science. Awesome stuff.

Top Data Science Hacks (<https://dimensionless.in/top-data-science-hacks/>)

Blog name: 365 Data Science
(<https://365datascience.com/blog/>)



(<https://365datasience.com/wp-content/uploads/2018/11/365-data-science-min-e1542113640631.jpg>)

Description: If you're going to check out one blog today, it really should be this one. Covering complex tutorials, career advice, interviews and general data science articles this blog has something for everyone. We know because we wrote it. Shameless plug I know, but I promised you a list of 51 of the top blogs and I really think ours deserves to be in it.

Alexa Rank: 283,726

Post frequency: 1-2 per week

Article to check out: One of our most popular articles. We researched dozens of books, hundreds of articles, and thousands of report pages on careers in data to give you *the most comprehensive data science career guide* out there. Check it out to see the result, I think you'll be surprised by the insights we got in it. Especially useful if you want to know if you could become a data scientist.

Starting a Career in Data Science: The Ultimate Guide

(<https://365datasience.com/career-data-science-ultimate-guide/>)

So, there's our list. 51 of our favourite data science blogs. As you can see they come in all shapes and sizes but we believe each of these has something to offer. We very much enjoyed looking through these projects of love from various professionals and enthusiasts.

in the data science field and we hope you have too.

Related Posts

Debunking 10 Misconceptions About AI (<https://365datascience.com/debunking-misconceptions-ai/>)

Shazmaan Malek (Scholarship Runner-Up) – AI and Ethics (<https://365datascience.com/shazmaan-malek-scholarship-application/>)

Choosing the right chart: Selecting among 14 chart types (<https://365datascience.com/chart-types-and-how-to-select-the-right-one/>)

The 5 Skills You Need To Match Any Data Science Job Description (<https://365datascience.com/5-skills-data-science-job/>)

Bayesian vs Frequentist Approach: Same Data, Opposite Results (<https://365datascience.com/bayesian-vs-frequentist-approach/>)

Related Posts



(<https://365datascience.com/5-business-basics-data-scientists/>)

5 Business Basics for Data Scientists
(<https://365datascience.com/5-business-basics-data-scientists/>)

7 min read

Articles, Career Advice



(<https://365datascience.com/andrea-scholarship-runner-up/>)

Andrea Uličná (Scholarship Runner-Up) – Free Will vs Recommender Systems
(<https://365datascience.com/andrea-scholarship-runner-up/>)

12 min read

Articles





(<https://365datasience.com/data-science-vs-ml-vs-data-analytics/>)

Data Science vs Machine Learning vs Data Analytics vs Business Analytics
[\(https://365datasience.com/data-science-vs-ml-vs-data-analytics/\)](https://365datasience.com/data-science-vs-ml-vs-data-analytics/)

11 min read

Articles



(<https://365datasience.com/john-oneil-scholarship-runner-up/>)

John O'Neill (Scholarship Runner-Up) – The Illusion of Choice
[\(https://365datasience.com/john-oneil-scholarship-runner-up/\)](https://365datasience.com/john-oneil-scholarship-runner-up/)

11 min read

Articles

6 comments



Nynitha Reddy

Reply

June 10, 2019 at 10:30 am (<https://365datasience.com/51-data-science-blogs/#comment-894>)

Thanks for sharing the valuable information about data science....It was very helpful to me



technorton

Reply

November 5, 2019 at 4:29 am (<https://365datasience.com/51-data-science-blogs/#comment-2566>)

It's very helpful list and thanks for the list..



**malikaa**[Reply](#)

November 7, 2019 at 5:45 am (<https://365datasience.com/51-data-science-blogs/#comment-2596>)

Thanks for the information. It's great

**shannu**[Reply](#)

November 13, 2019 at 5:06 am (<https://365datasience.com/51-data-science-blogs/#comment-2638>)

Thanks for sharing this innovate article.

**Emilia Jazz**[Reply](#)

November 18, 2019 at 1:38 pm (<https://365datasience.com/51-data-science-blogs/#comment-2683>)

Thanks for sharing. It's amazing.

**Samrid**[Reply](#)

November 28, 2019 at 12:45 pm (<https://365datasience.com/51-data-science-blogs/#comment-2724>)

These tips are really helpful. keep it up. Thanks for sharing the information.

Leave a Reply

Your email address will not be published.

COMMENT



//

NAME**EMAIL****POST COMMENT**

(<https://www.twinkl.com/the-dilmaur-channel/do-it-easy-Zhl1X8>)
(<https://365datascience.com/365datascience/>)

© 2020 365 Data Science. All Rights Reserved

^