

List of my best posts – 2023

- 1) Why Data Scientists love Notebooks while Software engineers hate it -
<https://bit.ly/3RY36FV>
- 2) Everything you wanted to know about LASSO and RIDGE Regression -
<https://bit.ly/3voUNKk>
- 3) The biggest myth - Normal Distribution is the most prevalent distribution in nature -
<https://bit.ly/41A18yR>
- 4) AUROC is all about Ranking? - <https://bit.ly/41BreBk>
- 5) Myth - Statistical Inference belongs to 20th Century - <https://bit.ly/3vgfusd>
- 6) Data Science projects that are best avoided in your resume – <https://bit.ly/48hksmX>
- 7) Causal Inference Resources - <https://bit.ly/4aOrIsn>
- 8) Which Error is worse - Type 1 Error or Type 2 Error? - <https://bit.ly/48iKZR2>
- 9) People to follow on Stack exchange and Cross Validated - <https://bit.ly/48ywRCy>
- 10) Rules of Thumb in Statistics - <https://bit.ly/3TEzYVF>
- 11) Why Statisticians make for good Data Scientists - <https://bit.ly/48xYyvy>
- 12) The levels of Linear Regression understanding - <https://bit.ly/48cF8MK>
- 13) Multivariate \neq Multivariable - <https://bit.ly/3vdg1uX>
- 14) What is the meaning of p value and t value in statistical tests? - <https://bit.ly/4aDIEkY>
- 15) Why Data Science aspirants are not finding Data Science Jobs - <https://bit.ly/3RxqHMi>
- 16) The relationship between Curse of Dimensionality and Degrees of Freedom -
<https://bit.ly/3TFmDfv>
- 17) Linear Regression is a classic example of 'So-much-has-been-written-yet-so-little-has-been-understood' - <https://bit.ly/3RYtfvI>

- 18) Sticking to Defaults - The reason your Data Science project could fail -
<https://bit.ly/3TB2aZv>
- 19) The secret to cracking Data Science Problems (hint: Probability Distributions) -
<https://bit.ly/48xkbMI>
- 20) Data Dredging and Algorithm Dredging -<https://bit.ly/3RWP1bK>
- 21) Why does π show up in Normal Distribution? - <https://bit.ly/3RBj7QP>
- 22) Why was ANOVA invented? - <https://bit.ly/3tyoR5S>
- 23) How useful is F test in Linear Regression? - <https://bit.ly/3TyfttE>
- 24) The Mean and Conditional Mean - <https://bit.ly/41BlUum>
- 25) The key to mastering Statistics/Data Science - <https://bit.ly/48rKUKV>
- 26) Red pill moments in Data Science - <https://bit.ly/41GB7hs>
- 27) The need for Statistical Assumptions - <https://bit.ly/3GZrwsA>
- 28) How do you know you are applying the right ML algorithm or statistical test? -
<https://bit.ly/3TEaJ5O>
- 29) Bootstrapping: a way to learn about sampling distribution - <https://bit.ly/47f2Oit>
- 30) A high t statistic does not indicate strong relationship of Independent variables with
the Dependent variable - <https://bit.ly/48wHiGZ>
- 31) Non parametric does not mean NO parameters rather it means many parameters -
<https://bit.ly/3RXvcBc>
- 32) R bloggers – a rich resource on statistics <https://bit.ly/48cApuK>
- 33) Why most Data Science projects start with a high probability of failure -
<https://bit.ly/3twg2cQ>
- 34) Data Science Consulting - Success Mantra <https://bit.ly/3vdiRjB>
- 35) Data Discipline - <https://bit.ly/47dTb3x>

- 36) Polynomial regression is a linear model - <https://bit.ly/3v7Njve>
- 37) Why Linear Regression is not all about predictions - <https://bit.ly/48Bylwd>
- 38) Seeing Statistical Tests through the lens of Signal vs Noise - <https://bit.ly/48wLeHB>
- 39) Prediction \neq Estimation - <https://bit.ly/48v9R7J>
- 40) The magic numbers in Statistics - <https://bit.ly/48uM47O>
- 41) $P(A|B) \neq P(B|A)$ - <https://bit.ly/3NINTGb>
- 42) Statistics' Physics connection – Moments <https://bit.ly/3S0ZenS>
- 43) PCA is not exactly a dimension reduction technique <https://bit.ly/48uPbww>
- 44) The Big XY problem in Data Science - <https://bit.ly/48BC2SB>
- 45) Statistics - First understand and then predict; Machine Learning - First predict and then understand <https://bit.ly/3RD91PI>
- 46) To explain or to predict - <https://bit.ly/3GZUIiQ>
- 47) How Marketing Mix Modeling (MMM) helped me learn Linear Regression from first Principles - <https://bit.ly/3tHtMRU>
- 48) Fit model to data not data to model - <https://bit.ly/41DYjNd>
- 49) How Bayesian methodology exacerbates Aleatoric and Epistemic uncertainties - <https://bit.ly/47bMOh3>
- 50) Why Everyone isn't a Bayesian - <https://bit.ly/4aDPwig>
- 51) Bayesian MMM is not a silver bullet for MMM's Multicollinearity issue - <https://bit.ly/4auN2Tz>
- 52) The Richard McElreath's Quartet - <https://bit.ly/3NHAM8j>
- 53) Why we report both confidence Interval and Prediction Interval in our MMM models- <https://bit.ly/41AVJHC>
- 54) Why do we need assumptions in Statistics? <https://bit.ly/3NEpnWM>

55) What the word 'confidence' in 'Confidence Interval' Signifies <https://bit.ly/3NMhP4j>

56) In MMM, you don't need to train/test split your data <https://bit.ly/3TFqjhr>

57) Explaining the 'Hourglass' shape of Confidence Interval <https://bit.ly/3RBoNu7>

58) Rows > Columns <https://bit.ly/3NJiZ4R>

59) Mean is not a prediction model <https://bit.ly/3H3chhW>

60) Dynamic Time Warping and its Innovative Use Cases in Marketing Science

<https://bit.ly/48cHCee>

61) Multicollinearity and Statistical Power <https://bit.ly/4ayvEgx>

62) Selecting MMM models via AIC? Some key pointers <https://bit.ly/41JRnyc>

63) How we build robust MMM models with help of Bootstrapping <https://bit.ly/3vfCojk>

64) Decile Analysis - An interesting NON-CLASSIFICATION use of Logistic Regression

<https://bit.ly/3GYN7kE>

65) Why Heteroscedasticity matters in Marketing Mix Modeling (MMM)

<https://bit.ly/48dKW8L>

66) The utility of F test in Marketing Mix Modeling (MMM) <https://bit.ly/48eAggu>

67) Validating MMM Models the right way <https://bit.ly/48wNxKP>

68) Calibration vs Validation <https://bit.ly/47eF7GT>

69) In Bayesian MMM, the 'lost in translation' problem is more pronounced

<https://bit.ly/41BdV3Q>

70) How a statistical technique that helped solve German Tank problem during WW2 is

helping us get accurate attribution in Marketing Mix Modeling (MMM)

<https://bit.ly/3twH6Zm>

71) Why customers should not expect their Marketing Mix Models (MMM) to have very

high R squared value <https://bit.ly/41Cr5h3>

- 72) Don't make million-dollar decision based on $\text{CORREL}(x, y)$ <https://bit.ly/3NGY5yV>
- 73) It is good to choose coefficients to magnify the signal, but it is far better to choose them to cancel out the noise <https://bit.ly/3tz7AcJ>
- 74) The meaning of 'statistic' in statistics explained through Matryoshka Dolls - <https://bit.ly/3tz7Fx3>
- 75) Poisson Distribution is not about modeling *rare* events <https://bit.ly/3vb6OTD>
- 76) Why confidence intervals in a way are barometer of your Marketing Mix Models (MMM)? <https://bit.ly/48cOrfE>
- 77) Looking to get started on MMM? Know your type of Data <https://bit.ly/48vf6Er>
- 78) Don't Stepwise Regression your MMM model <https://bit.ly/3vgKNDb>