

## Data Science Trivia: Questions, Answers, and Facts

### Printable Trivia Pack

Data Science trivia highlights the people, methods, and milestones that shaped a discipline at the crossroads of statistics, computing, and real-world problem solving. From foundational concepts to playful facts about algorithms, data, and discovery, it offers a family-friendly mix of easy, funny, and more challenging questions grounded in the field's evolution.

#### HARD DATA SCIENCE TRIVIA

1. Which method became widely known after the 1995 work of Cortes and Vapnik?

Answer: Support vector machine

2. What was the publication year of the backpropagation paper that reignited interest in neural networks?

Answer: 1986

3. In what year did the ImageNet Large Scale Visual Recognition Challenge begin?

Answer: 2010

4. Which model won the ImageNet competition in 2012?

Answer: AlexNet

5. What system defeated Lee Sedol in Seoul in 2016?

Answer: AlphaGo

6. The clustering algorithm introduced in 1996 was which one?

Answer: DBSCAN

7. Robert Tibshirani introduced which method in 1996?

Answer: LASSO

8. Which resampling method did Bradley Efron introduce in 1979?

Answer: Bootstrap

9. What method was proposed by Maurice Quenouille before the bootstrap arrived?

Answer: Jackknife

10. Name the algorithm that was published in 1953 and later became foundational in Monte Carlo methods.?

Answer: Metropolis algorithm

11. What optimization theorem was published by Wolpert and Macready in 1997?

Answer: No Free Lunch theorem

12. The ROC curve was originally developed in connection with what wartime application area?

Answer: Radar signal detection during World War II

13. In model evaluation shorthand, what does AUC stand for?

Answer: Area under the curve

14. Chronology check: which came first, the Metropolis algorithm or the Gibbs sampler entering mainstream statistics?

Answer: The Metropolis algorithm came first.

#### FUNNY DATA SCIENCE TRIVIA

1. What classic data-analysis lesson says your chart should get a turn before your hot take does?

Answer: Always plot your data.

2. Which computing-era warning from the 1950s politely says bad input leads to bad output, even if the computer looks very confident?

Answer: garbage in, garbage out

3. Which language was created by Ross Ihaka and Robert Gentleman in the 1990s, giving statisticians yet another reason to argue about plotting defaults?

Answer: R

4. Guido van Rossum first released which language in 1991, eventually enabling many notebooks, scripts, and suspiciously optimistic comments?

Answer: Python

5. Which Python library created by Wes McKinney became famous for making tables slightly less scary?

Answer: pandas

6. What plotting package took its name from the phrase 'Grammar of Graphics,' as if charts needed their own textbook title?

Answer: ggplot2

7. Which Python visualization library was created by Michael Waskom, helping many people make prettier plots than their first draft deserved?

Answer: seaborn

8. Which package name is literally short for 'Numerical Python,' in case the full title felt too committed?

Answer: NumPy

9. Which package name expands to 'Scientific Python,' sounding exactly as serious as it intends to?

Answer: SciPy

10. What term describes searching through data until something finally looks significant enough to wave around dramatically?

Answer: data dredging

11. Finish the standard statistics warning: correlation does not imply \_\_\_\_\_.?

Answer: causation

12. Nassim Nicholas Taleb popularized which lesson about misleading trends, starring a bird whose confidence is not matched by its timeline?

Answer: the turkey illusion

13. Who is credited with the line, 'All models are wrong, but some are useful'â a quote that manages to be rude and helpful at the same time?

Answer: George Box

#### FUN DATA SCIENCE TRIVIA

1. Which statistics classic starts with the wonderfully suspicious word "How"?

Answer: How to Lie with Statistics

2. Which book snagged the 2012 National Academies Communication Award: "Thinking, Fast and Slow" or "The Signal and the Noise"?

Answer: Thinking, Fast and Slow

3. Who created the TeX typesetting system, giving technical writing a very precise personality?

Answer: Donald Knuth

4. During World War II, who made major contributions to sequential analysis?

Answer: Abraham Wald

5. What term did Arthur Samuel coin in 1959?

Answer: Machine Learning

6. Which came first: the perceptron or the coining of the term "machine learning"?

Answer: The perceptron

7. Frank Rosenblatt introduced which early learning model in 1958?

Answer: perceptron

8. Which method was described in 1951 by Evelyn Fix and Joseph Hodges?

Answer: k-nearest neighbors

9. A 1977 paper by Dempster, Laird, and Rubin formalized what algorithm?

Answer: expectation-maximization algorithm

10. Which method did Leo Breiman introduce in 2001?

Answer: random forest

11. Which algorithm was developed by Larry Page and Sergey Brin?

Answer: PageRank

12. What competition put a cool \$1 million on the table for improving movie recommendation

accuracy?

Answer: Netflix Prize

13. Who is known for public outreach on both data science and astronomy?

Answer: Kirk Borne

#### DATA SCIENCE FAMILY TRIVIA

1. In what year was the first edition of Francis Anscombe's famous quartet published?

Answer: 1973

2. How many data points are in each dataset of Anscombe's quartet?

Answer: 11

3. Who wrote the book "The Visual Display of Quantitative Information"?

Answer: Edward Tufte

4. Which member of Hans Rosling's family helped co-author "Factfulness" with him?

Answer: Ola Rosling

5. Besides Hans Rosling and Ola Rosling, which co-author helped write "Factfulness"?

Answer: Anna Rosling R^nnlund

6. Who popularized the term "debugging" after a moth was found in a computer relay?

Answer: Grace Hopper

7. True or false: The 2016 film "Hidden Figures" follows mathematicians who worked for NASA.?

Answer: True

8. If you visited the Jet Propulsion Laboratory, which U.S. state would you be in?

Answer: California

9. Los Alamos National Laboratory is located in which state?

Answer: New Mexico

10. Which state is home to Oak Ridge National Laboratory?

Answer: Tennessee

11. Bletchley Park is in which country?

Answer: England

12. The University of Cambridge is located in which city and country?

Answer: Cambridge, England

#### EASY DATA SCIENCE TRIVIA

1. Who mapped London cholera cases around the Broad Street pump in 1854?

Answer: John Snow

2. Which pioneer used polar area diagrams to present mortality data from the Crimean War?

Answer: Florence Nightingale

3. Who wrote the 1925 book "Statistical Methods for Research Workers"?

Answer: Ronald Fisher

4. Who published "Computing Machinery and Intelligence" in 1950?

Answer: Alan Turing

5. Claude Shannon worked at which famous research organization?

Answer: Bell Labs

6. Which researcher is strongly associated with Statistical Learning Theory?

Answer: Vladimir Vapnik

7. Who co-founded Coursera in 2012?

Answer: Andrew Ng

8. Which 2018 Turing Award co-winner is Geoffrey Hinton?

Answer: Geoffrey Hinton

9. Who shared the 2018 Turing Award alongside other deep learning leaders?

Answer: Yoshua Bengio

10. Who introduced generative adversarial networks in 2014?

Answer: Ian Goodfellow

11. Which R package creator is Hadley Wickham known for?

Answer: ggplot2

12. Who co-founded the Gapminder Foundation?

Answer: Hans Rosling

13. Which statistics website was founded by Nate Silver?

Answer: FiveThirtyEight

Source: <https://triviagong.com/themes/data-science>