

## Modern Data Scientist:

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Equally difficult, is finding people who understand who a data scientist is. Here is a cheat sheet describing the modern data scientist. Contrary to popular belief, **data scientists are not Unicorns**, so all skills may not apply.

### Math & Statistics:

- Machine learning
- Statistical modeling
- Experimental design
- Bayesian inference
- Supervise learning:
  - Decision trees
  - Random forests
  - Logistic regression
- Unsupervised learning:
  - Clustering
  - Dimensionality reduction
- Optimization:
  - Gradient descent & variants

### Programming & Database:

- Computer science fundamentals
- Scripting language e.g. Python
- Statistical computing package e.g. Python, R
- Databases: SQL & NoSQL
- Relational algebra
- Parallel databases & query processing
- MapReduce
- Hadoop & Hive/Pig
- Custom reduces
- Experience with [XaaS](#) like AWS

### Domain Knowledge & Soft Skills:

- Passionate about the business
- Curious about data
- Influence without authority
- Hacker mindset
- Problem solver
- Strategic
- Proactive
- Creative
- Innovative
- Collaborative

### Communication & Visualization:

- Able to engage with management
- Story telling skills
- Translate data driven insights into decisions and actions
- Visual art design
- Python packages like Matplotlib, Bokeh & Seaborn
- R packages like GGPlot or lattice
- Knowledge of other visualization tools:
  - Flare
  - D3.js
  - Tableau