



# DATA SCIENCE SKILLS

Kunaal Naik

# Credits

- Analytics Vidhya : [Article](#)
- Analytics Mag : [Article](#)
- SlideShare : [Complete Study](#)
- Educba : [Article](#)

# Major Comparison

## Data Scientist



Data Science is all about finding out new things, revelation of new data which will solve complex problems. Finding conclusions through statistics through mere observation and gradually reaching the perfect optimized solution is the job of a data scientist.

## Business Analyst



Business Analysts are a platform between IT and business stakeholders. They need to have a deep business knowledge and need to be involved in demanding questions to get value for money and bring value to developments done in IT industry.

# Requirement

## Data Scientist



A data scientist needs to have knowledge about all latest tools, SQL and if required they may need to code. They should have in depth knowledge of mathematics and statistics..

## Business Analyst



Business analysts may not require any technical knowledge. They must be comfortable in assessing changes, developing business cases and defining new requirements or changes in a project from functional perspective.

# History

## Data Scientist



Data analysis though seems to be a new craze these days, it dates back to 1962 when John Tukey wrote about 'The Future of Data Analysis'. Post that there were mentions about this and it started trending from 2006, through 2011 till now where data scientists are the most sought job profiles.

## Business Analyst



automate repetitive tasks, identify problems and deliver good quality technology at the expense of business needs. Through 1980s Business Analysts evolved to support business goals and be a mediator more effectively between IT Business Analysts came to the rising in 1970s when they started documenting all manual processes. They found the need to resources and business resources.



# Responsibilities

## Data Scientist



A data scientist has to handle and extract large amounts of data. This requires in depth knowledge of SQL to segregate datasets. They must have advanced knowledge of machine learning so that they can make changes in data by themselves and get a deeper insight..

## Business Analyst



Business Analysts need to gather and prepare requirements. They must prepare documents and also analyze and model all requirements. Post analysis they must take over the changes that are required and convey the same to IT team. Once changes are done they must perform acceptance testing to check if the requirements are met.

# Tools

## Data Scientist



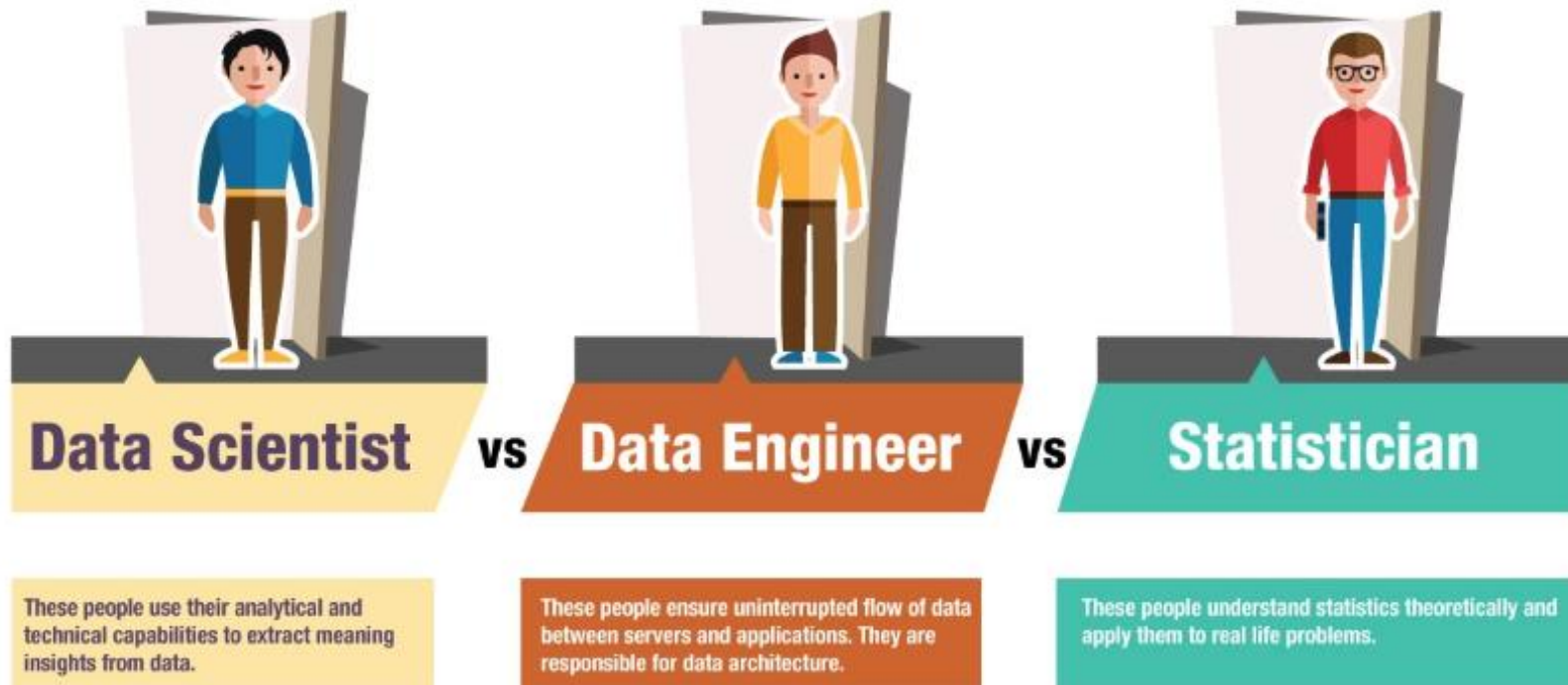
Tools of data scientists are none other than Data warehousing, Data visualization and machine learning.

## Business Analyst



There are various tools for business analysis like Blueprint, Axure, Bit impulse, etc. which make improve productivity.

# Comparison





# Responsibilities

Develop and plan required analytic projects in response to business needs.

Contribute to data mining architectures, modeling standards, reporting, and data analysis methodologies.

Collaborate with stakeholders to integrate data mining results with existing systems.

Monitor data mining system performance and implement efficiency improvements.

Design, construct, install, test and maintain highly scalable data management systems

Improve data foundational procedures, guidelines and standards

Integrate new data management technologies and software engineering tools into existing structures

Create custom software components (e.g. specialized UDFs) and analytics applications

Apply statistical theories and methods to solve practical problems of various industries

Determine methods for finding or collecting data

Design surveys or experiments or opinion polls to collect data

Analyze, interpret & undertake data analysis

Report conclusions from their analyses

# Skills

Programming, Mathematics, Business Understanding, Statistics, Data Visualization, Machine Learning, Attention to detail

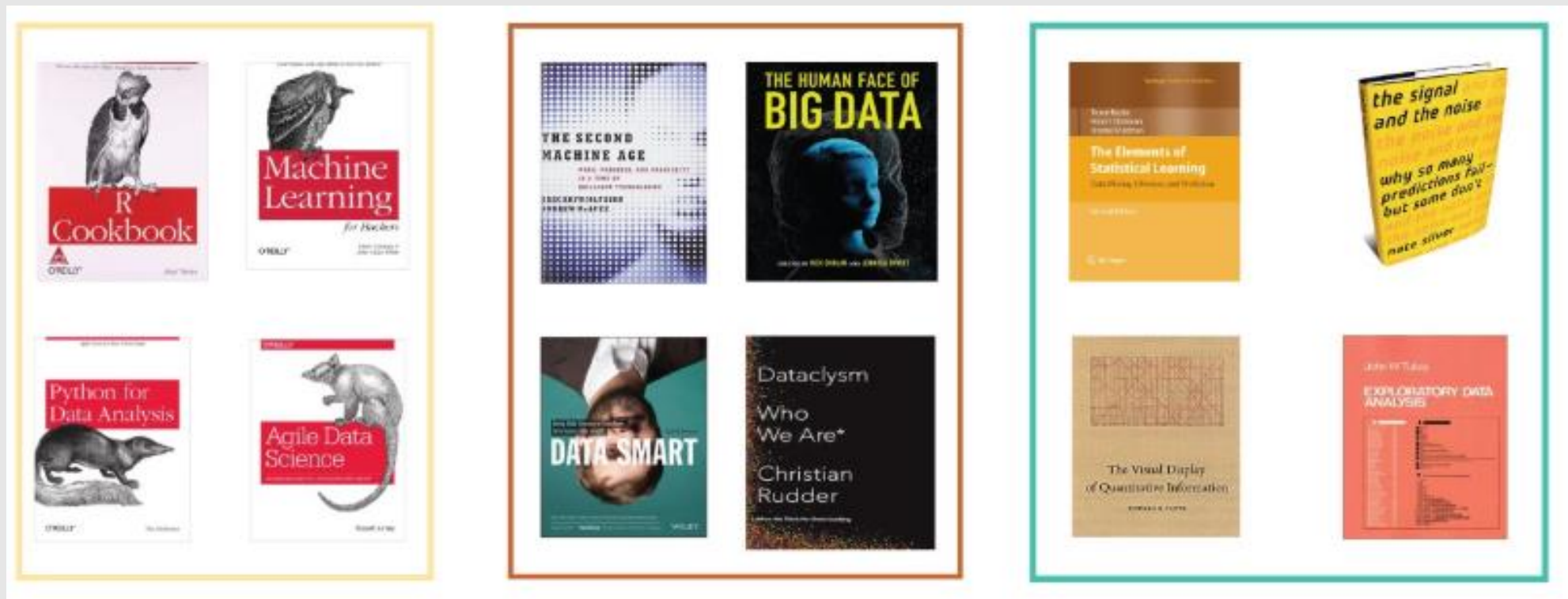
Database design, Production coding, Data collection, data warehousing, Data transformation, Work diligently with data

Technical and Analytics Skills, Mathematics, Operational Research, Writing skills, Ability to Analyze, Model and interpret data, Flair of explaining difficult concepts in simple manner

# Tools



# Books 😊



# Challenging Tasks

**Design  
Recommender  
Systems**

**Datawarehouse  
Design,  
Writing queries  
efficiently**

**Risk  
Pricing  
Model**



# Career Path



↑

- Chief Data Scientist
- Senior Data Scientist
- Data Scientist
- Junior Data Scientist



↑

- Data Architect
- BI Architect
- Senior Data Engineer
- Data Engineer



↑

- Private Statistical Consultant
- Statistical Manager
- Senior Statistician
- Applied Statistician
- Statistical Analyst

# Projects

- Beginner Projects All – [Analytics Vidhya Article](#)
- Tableau – [IPL Data](#), [Movies Data](#)
- Reporting Project : [Google Apps Data](#)

# Salaries Link

- [Cutshort](#)
- [Payscale](#)
- [Paysa](#)

# Attend Conferences

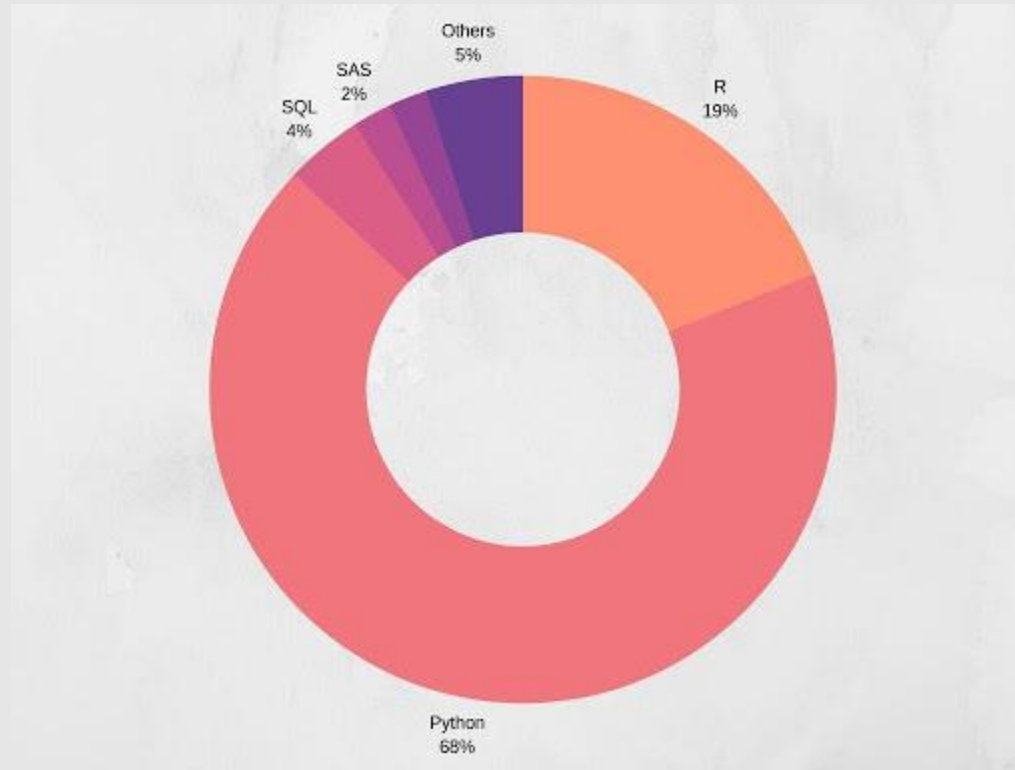
- Open Data Science Conference
- Analytics Summit
- Linked In Groups
- MeetUp Groups



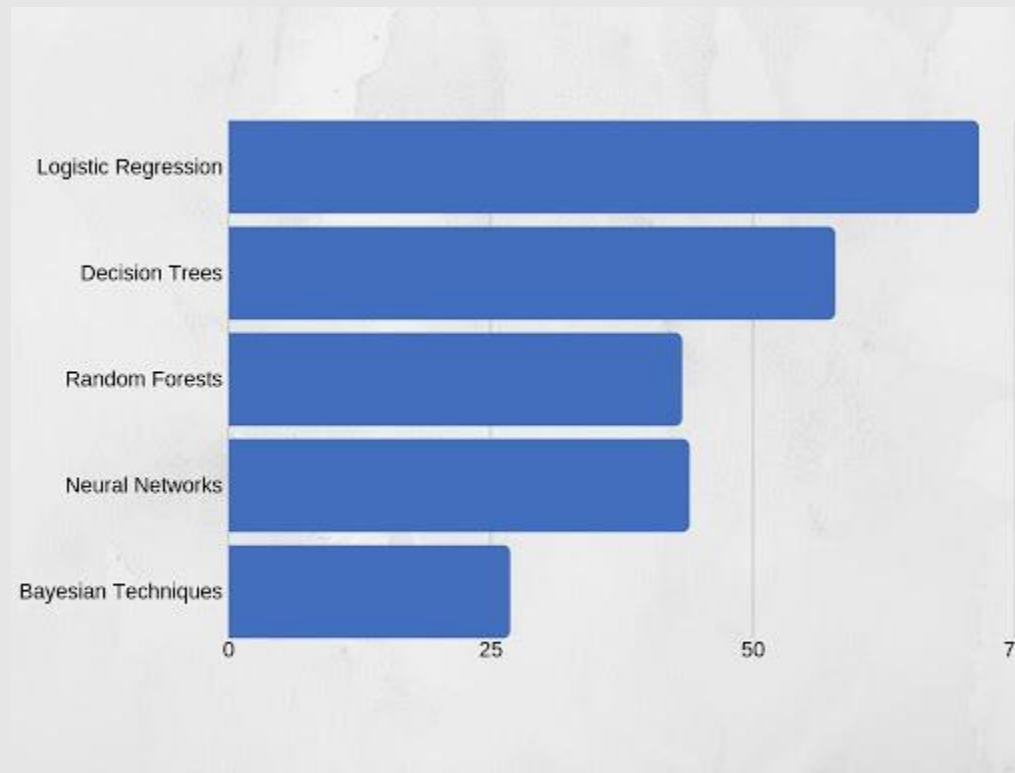
MORE DETAILS



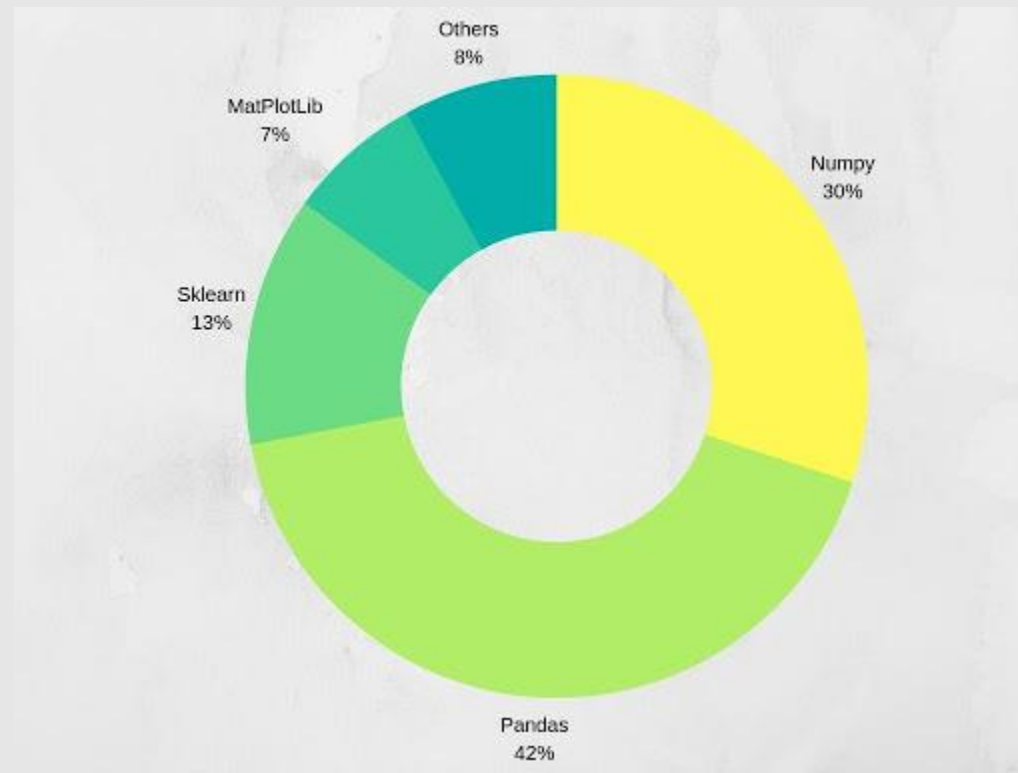
# Which Language Do Data Scientists Prefer For Statistical Modelling?



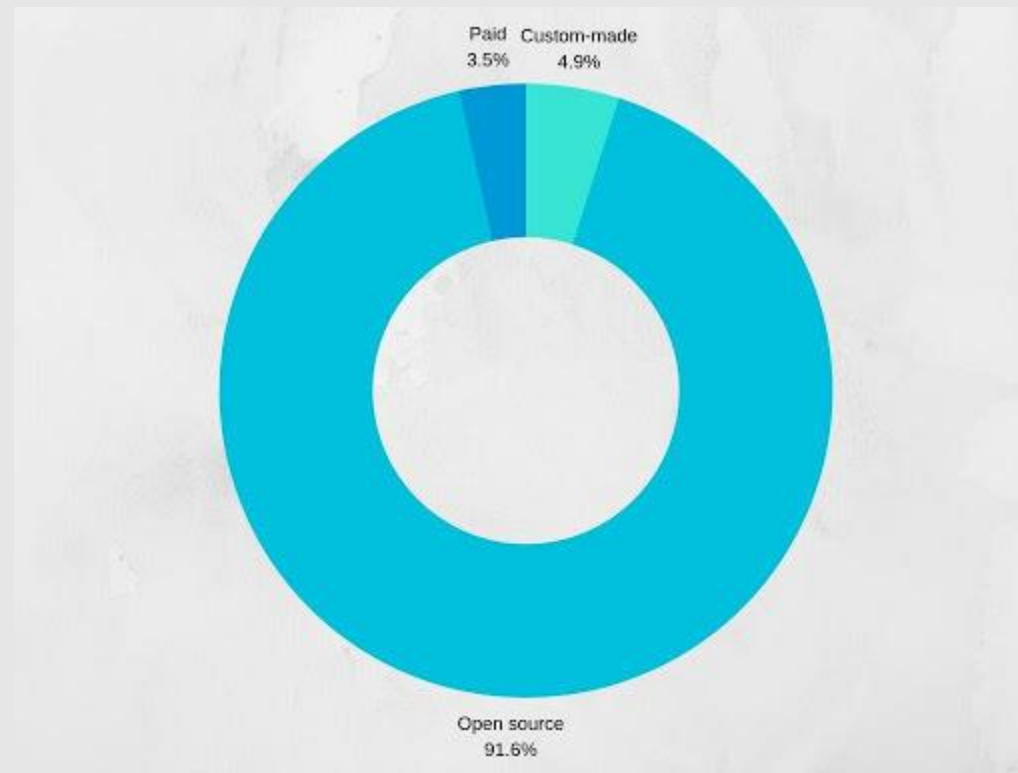
# Which Data Science Methods Are The Most Popular At Work?



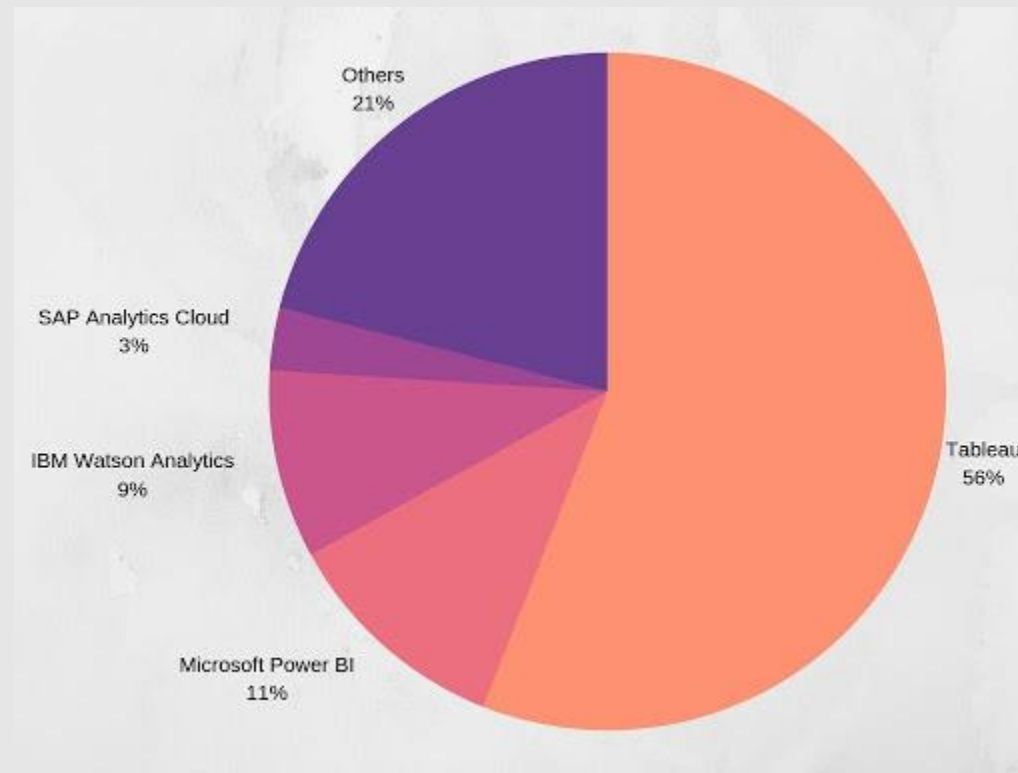
# Which Is The Most Popular Python General Purpose Library?



# Which Tools Do Data Scientists Prefer?

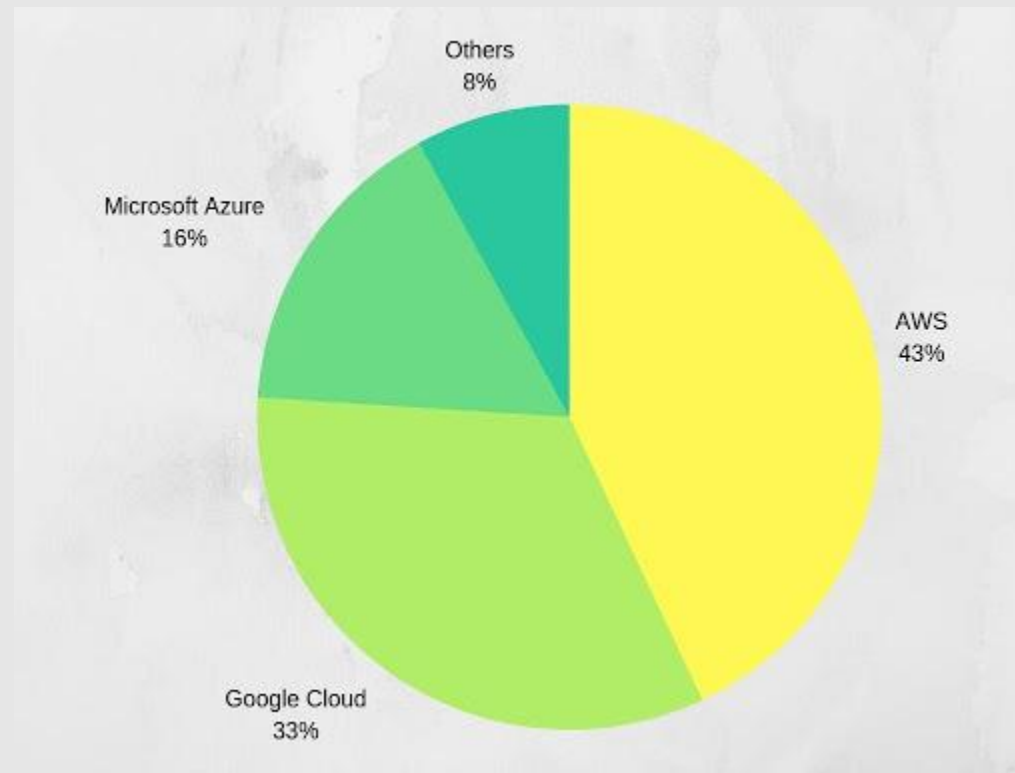


# Which Dashboard/Visualisation Tools Do Data Scientists Prefer?

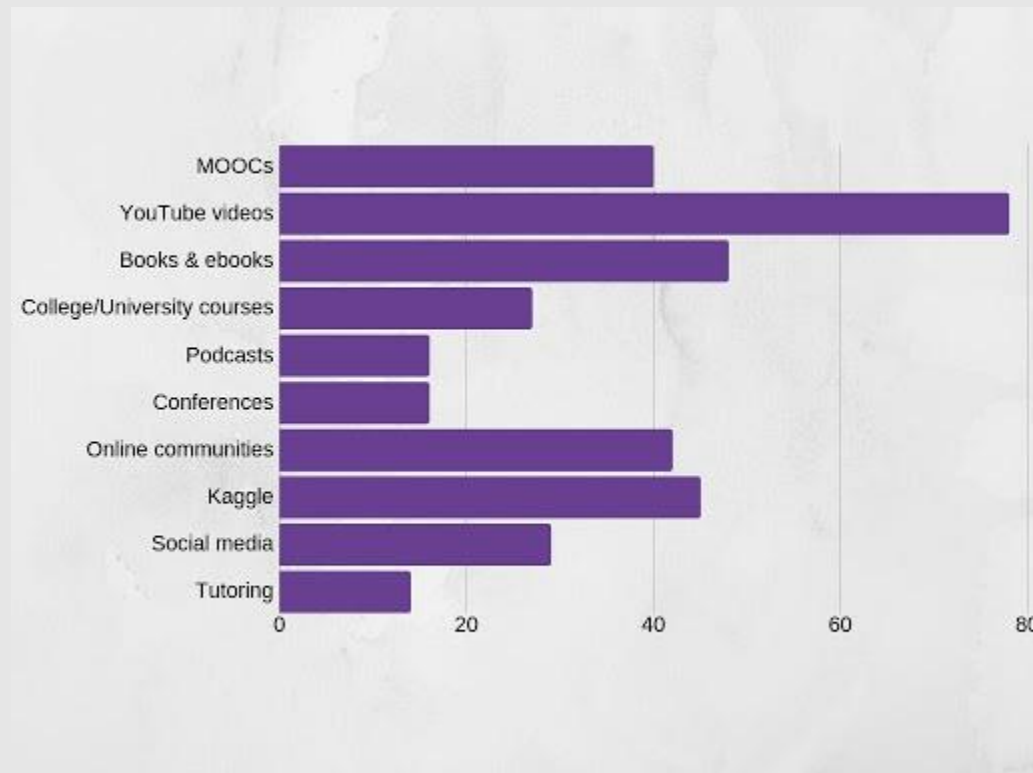




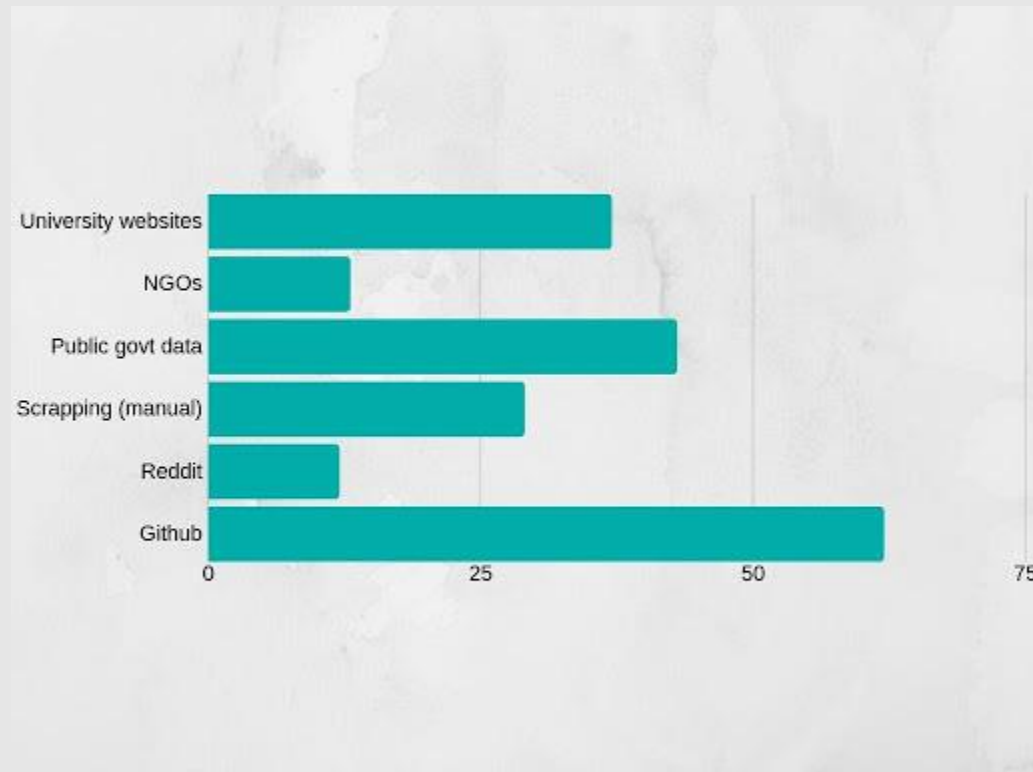
# Which Cloud Provider Do Data Scientists Prefer?



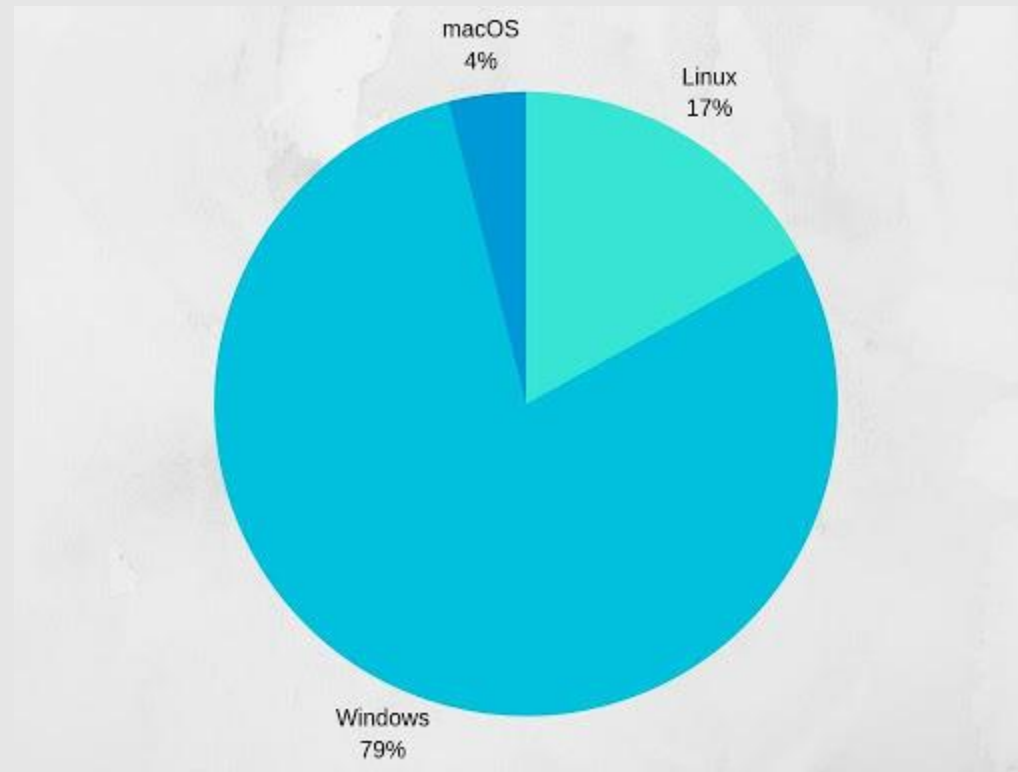
# What Kind Of Learning Resources Do Data Scientists Use To Keep Themselves Updated?



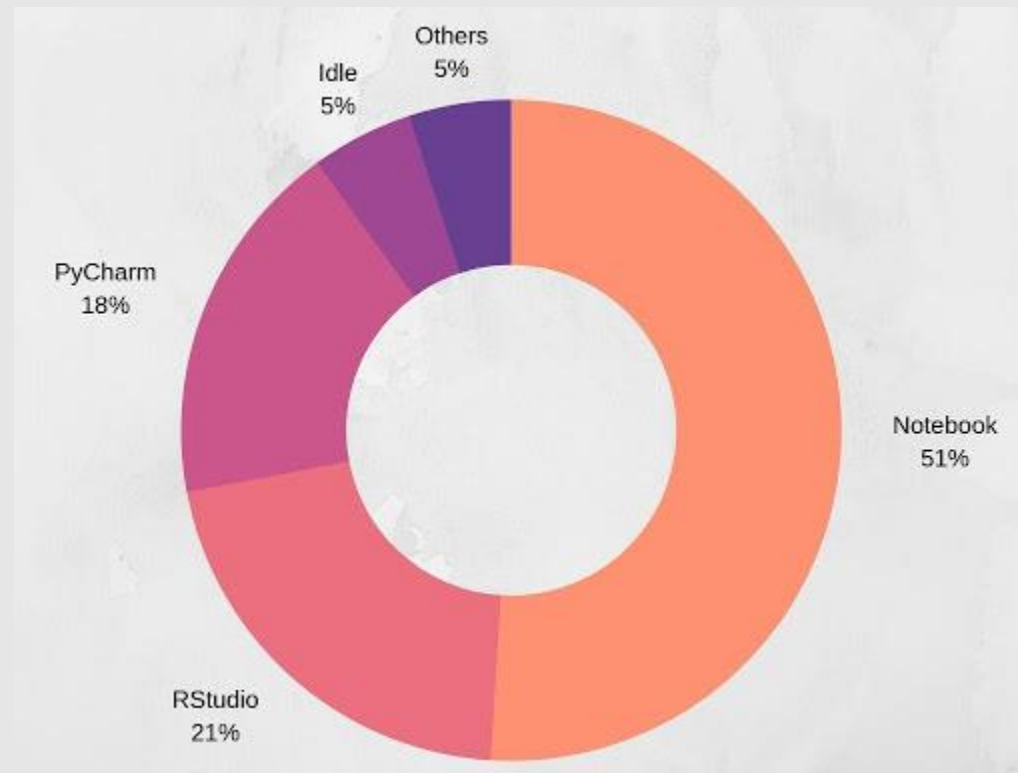
# Where Do Data Scientists Find Open Data?



# Which OS Do Most Data Scientists Use At Work?

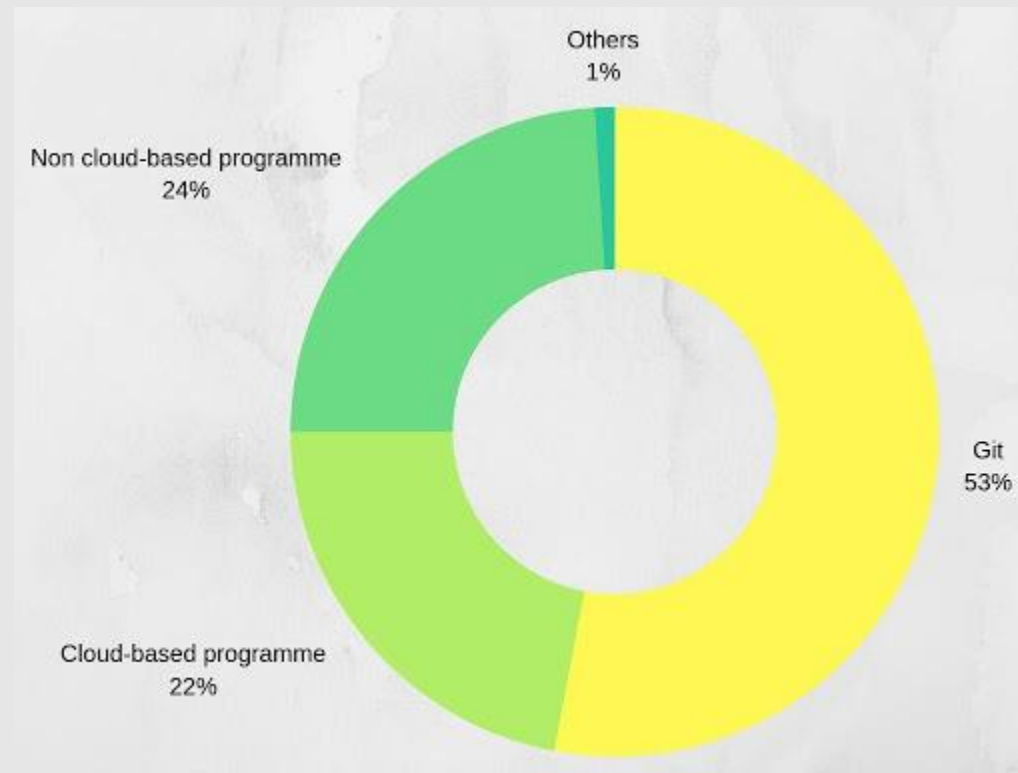


# Preferred Development Environment

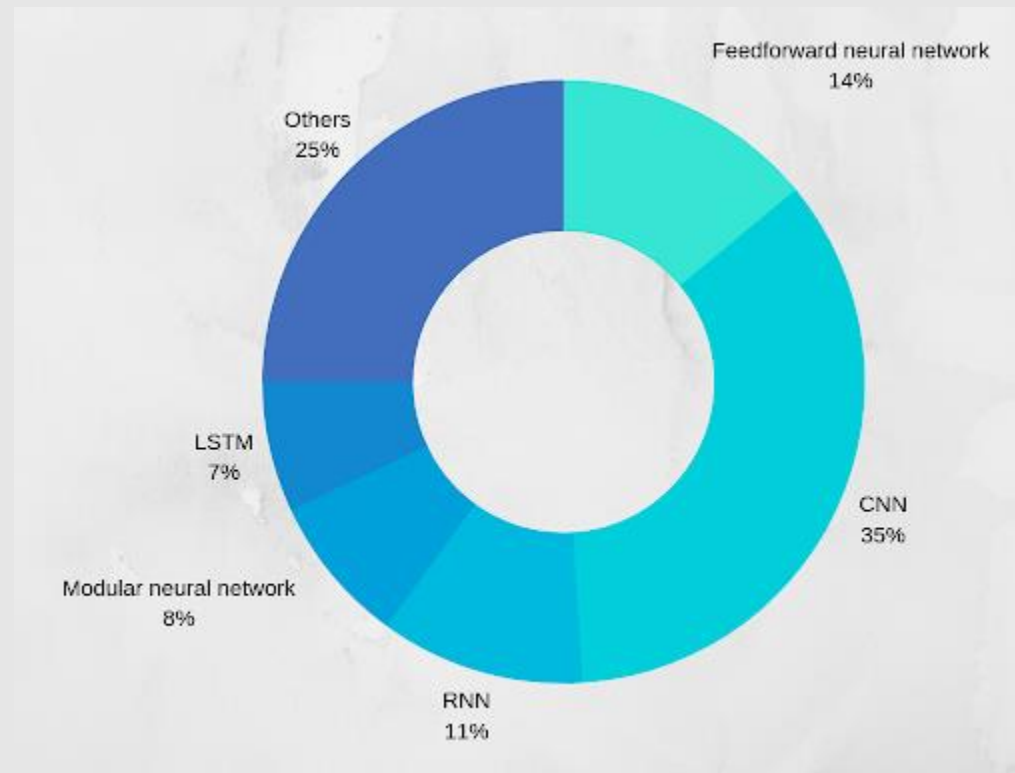




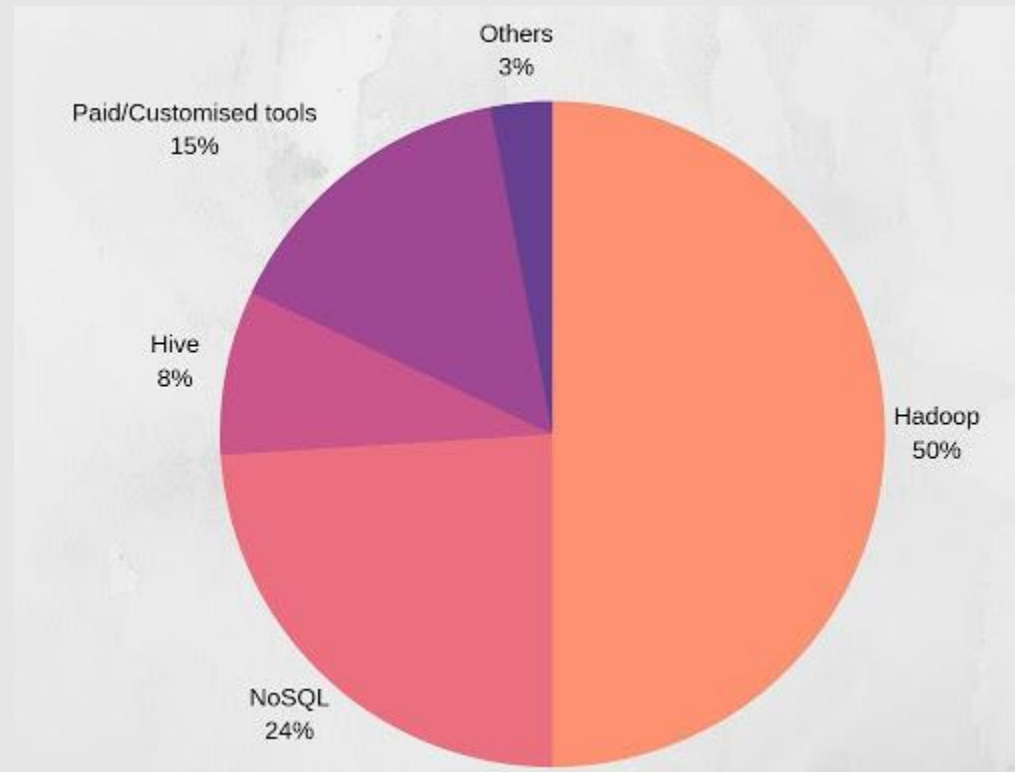
# How Is Code Shared At Your Workplace?



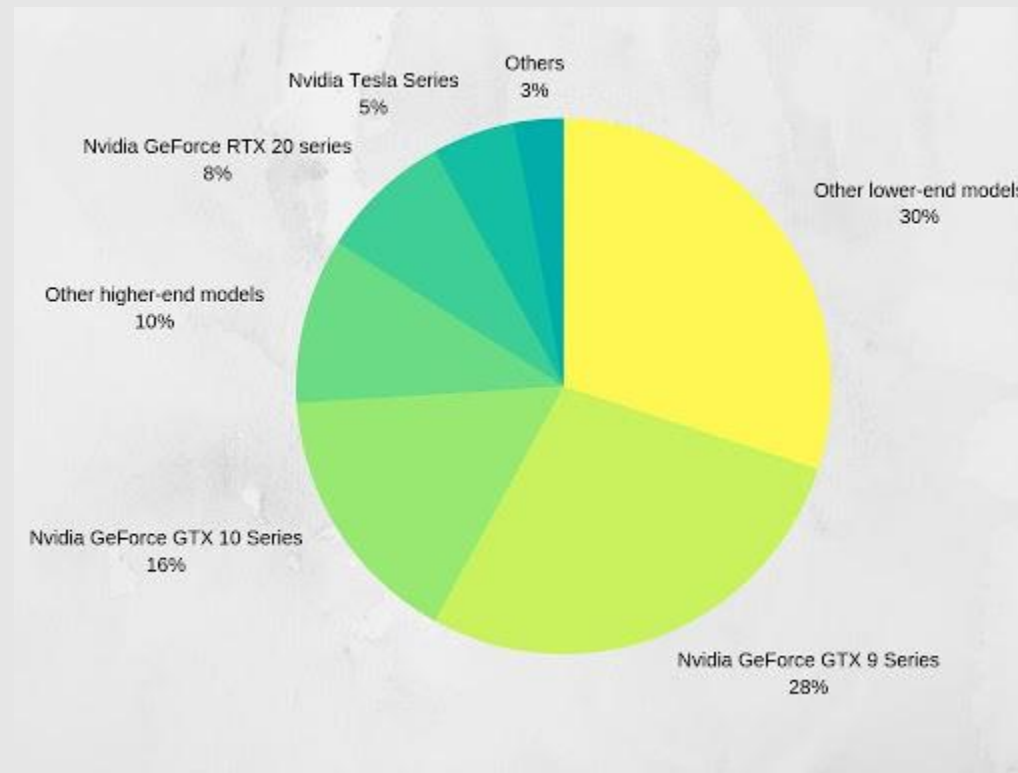
# What Is The Neural Network Architecture Data Scientists Use Most Frequently?



# Which Big Data Tool Have You Used The Most?



# Which GPUs Do Data Scientists Use At Work?



# Which Skills Are Most Important?

	BFSI	IT products and services	Manufacturing	E-commerce	Retail
Language	Python	Python	Python	Python	Python
DS Method	Logistic Regression	Logistic Regression	Logistic Regression	Logistic Regression	Logistic Regression
Python Library	Pandas	Pandas	Pandas	Numpy	Pandas
Data Visualisation Tool	Tableau	Tableau	Tableau	Tableau	Tableau
Cloud	AWS	AWS	AWS	Google Cloud	AWS
IDE	Notebook	Notebook	Notebook	RStudio	Notebook
NN Architecture	CNN	CNN	RNN	CNN	CNN
Big Data Tool	Hadoop	Hadoop	Hadoop	Paid/Customised tools	Hadoop