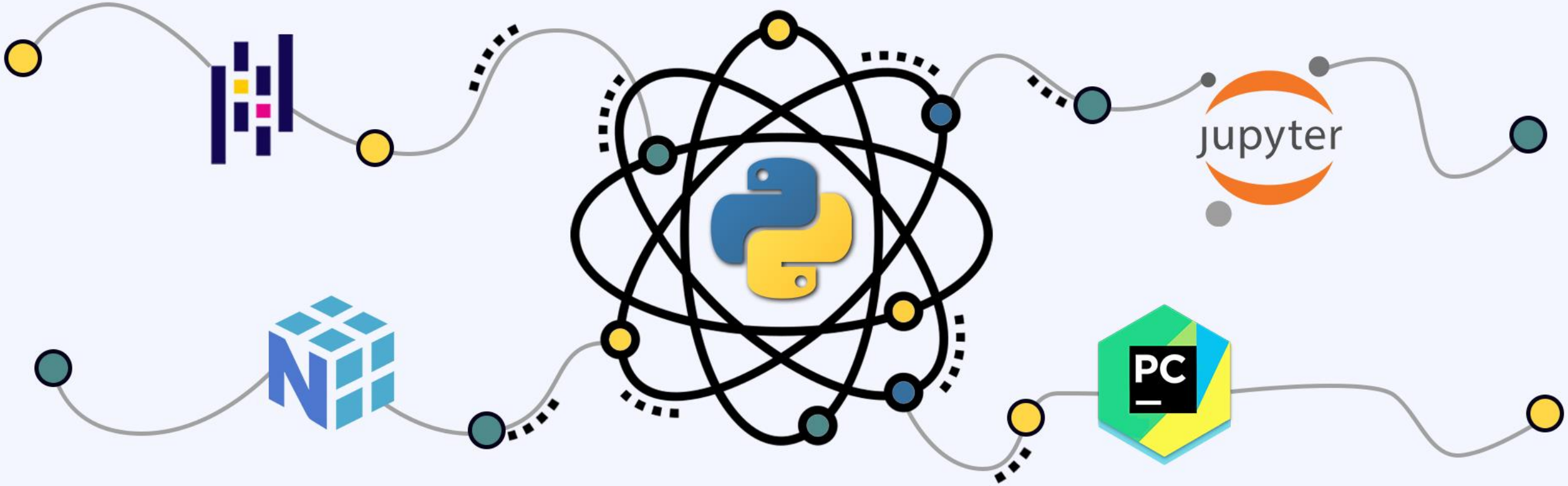
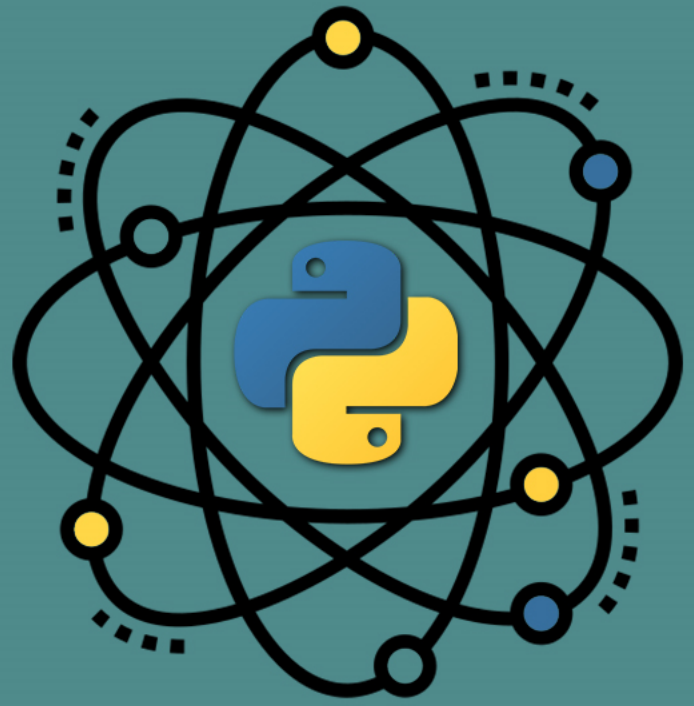


MOHAMMAD DEHGHANI PTYHON FOR DATA SCIENCE

زبان پایتون در علم داده

مدرس: محمد دهقانی



زبان پایتون در علم داده

مدرس:
محمد دهقانی

جلسه: دوم
موضوع: Python vs R

مواردی که در این جلسه بررسی می شوند:

1. معرفی پایتون
2. معرفی R
3. مقایسه هر دو

Python vs. R:

In many ways, the two open source languages are very similar. Free to download for everyone, both languages are well suited for data science tasks. The main difference is that Python is a general-purpose programming language, while R has its roots in statistical analysis.

Tip: Once you've learned one programming language, it's typically easier to learn another one.

About Python:

Python is a general-purpose, object-oriented programming language that emphasizes code readability through its generous use of white space. Released in 1989, Python is easy to learn and a favorite of programmers and developers. In fact, Python is one of the most popular programming languages in the world, just behind Java and C.

Plus, Python is particularly well suited for deploying machine learning at a large scale. Its suite of specialized deep learning and machine learning libraries includes tools like scikit-learn, Keras, Pytorch and TensorFlow, which enable data scientists to develop sophisticated data models that plug directly into a production system.

Guido van Rossum (Inventor of Python) •

About R :

R is an open source programming language that's optimized for statistical analysis and data visualization. Developed in 1992, R has a rich ecosystem with complex data models and elegant tools for data reporting. At last count, more than 13,000 R packages were available via the Comprehensive R Archive Network (CRAN) for deep analytics.

Training and evaluating machine learning and deep learning algorithms(With the help of keras)

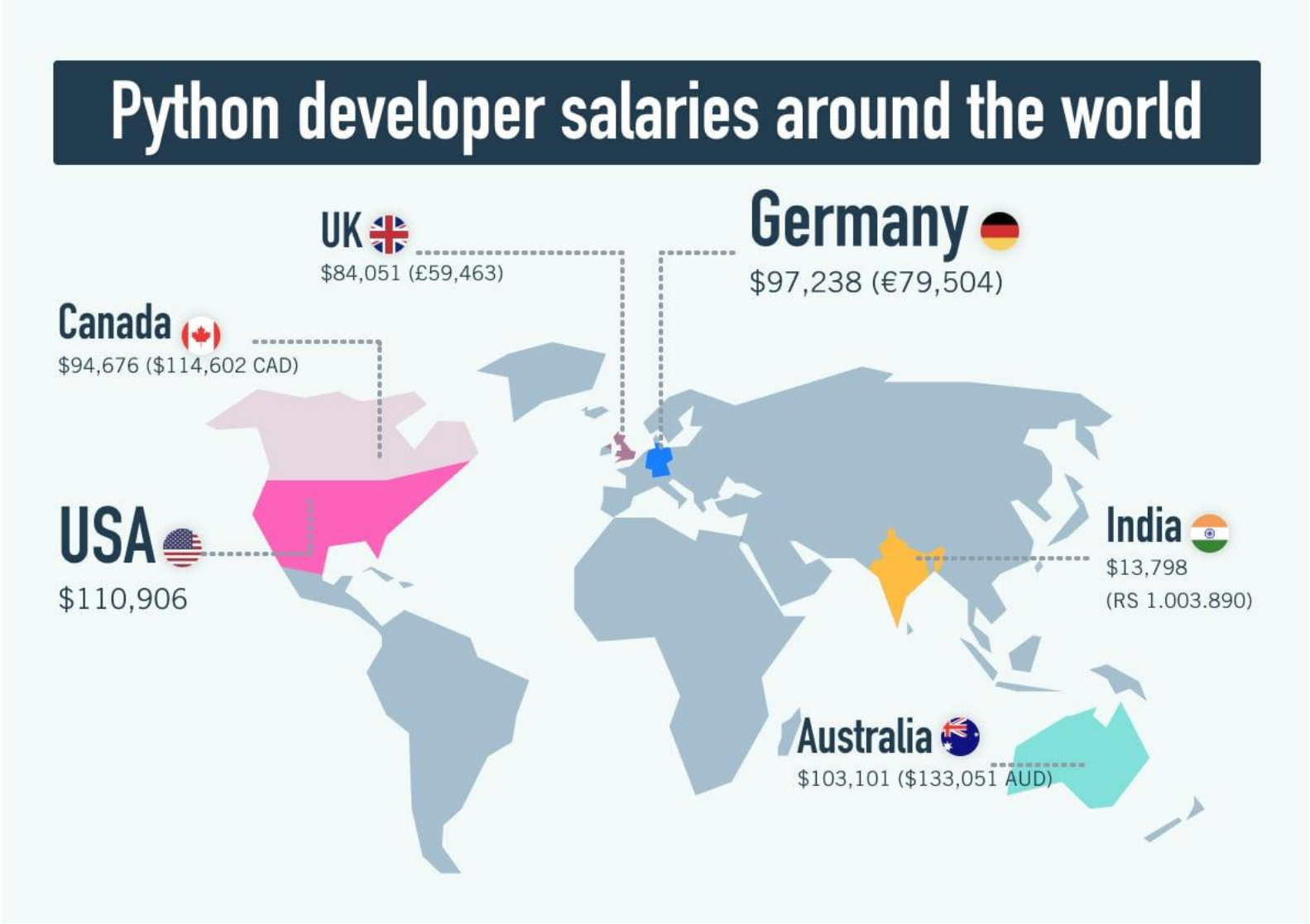

ROBERT GENTLEMAN
ROSS IHAKA

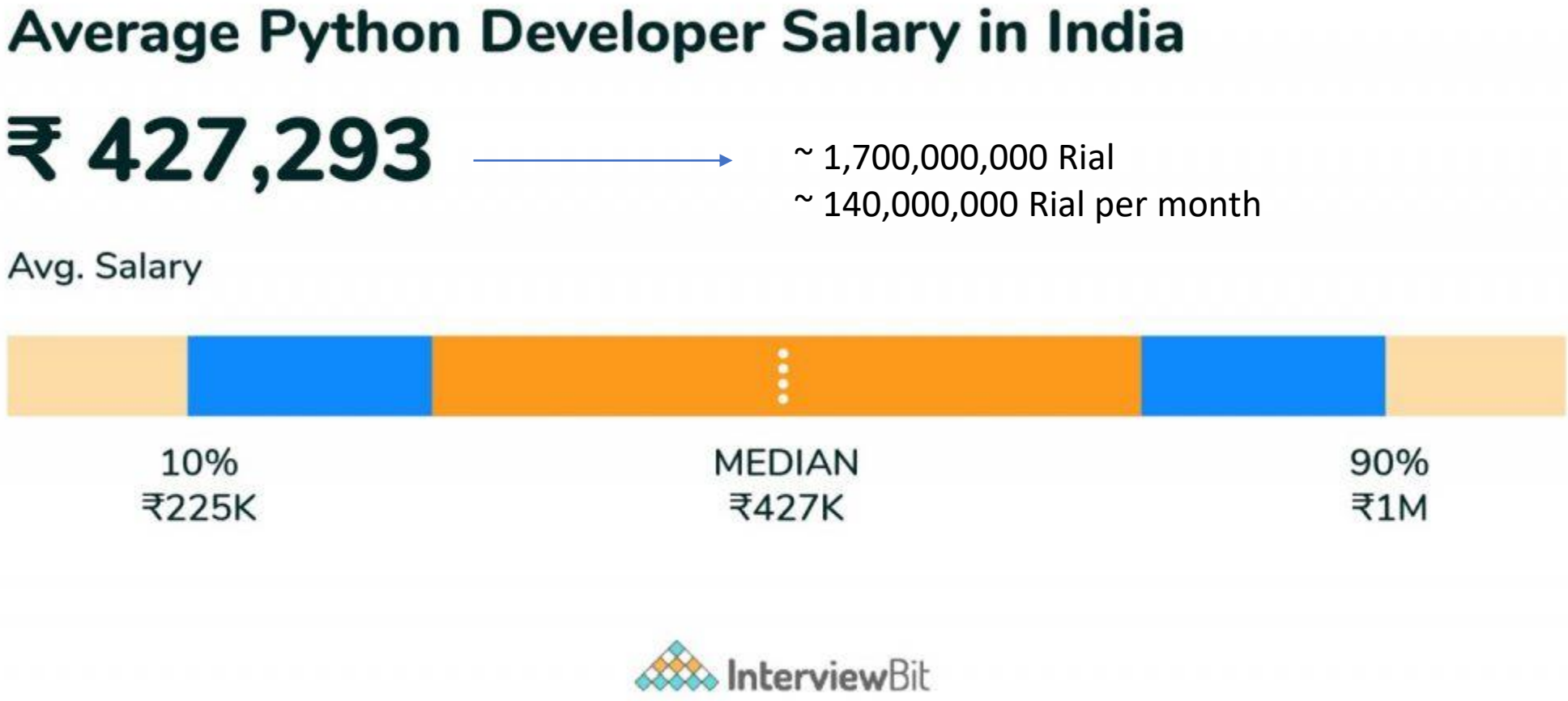
Top Paying and Most Popular Programming Languages in 2020

Rank by Average Salary	
1. Python	$119,000
2. JavaScript	$117,000
3. Java	$104,000
4. C	$103,000
5. C++	$102,000
6. C#	$97,000
7. PHP	$94,000
8. SQL	$92,000

Rank by Volume of Job Openings	
1. Python	50,000
2. SQL	50,000
3. Java	45,000
4. JavaScript	38,000
5. C++	29,000
6. C#	21,000
7. PHP	13,000
8. C	9,000

Source:codeplatoon.org


Python developer salaries around the world
UK
$84,051 (£59,463)
Germany
$97,238 (€79,504)
Canada
$94,676 ($114,602 CAD)
USA
$110,906
India
$13,798
(RS 1.003.890)
Australia
$103,101 ($133,051 AUD)



~ 1,700,000,000 Rial
~ 140,000,000 Rial per month

| AI |

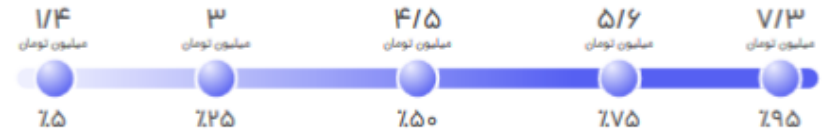
| Data Mining |

| Image Processing |

تعداد شرکت کنندگان: ۶۴ نفر

نحوه تفسیر نمودارهای گزارش

روی نمودارها ۵ عدد به صورت درصدی وجود دارد که بالای هر کدام از این درصدها یک عدد به میلیون تومان نوشته شده است. بطور مثال در نمودار زیر بالای ٪۷۵، عدد ۵/۶ میلیون تومان نوشته شده است. این بدین معناست که ٪۷۵ از افراد این جامعه آماری، کمتر از ۵/۶ میلیون تومان و ٪۲۵، بیش از ۵/۶ میلیون تومان دریافت می کنند.



برای فهم بهتر گزارش لطفا بخش **راهنمای گزارش** در صفحه ۸ را مطالعه کنید.

| Django, etc. |

تعداد شرکت کنندگان: ۴۶۱ نفر

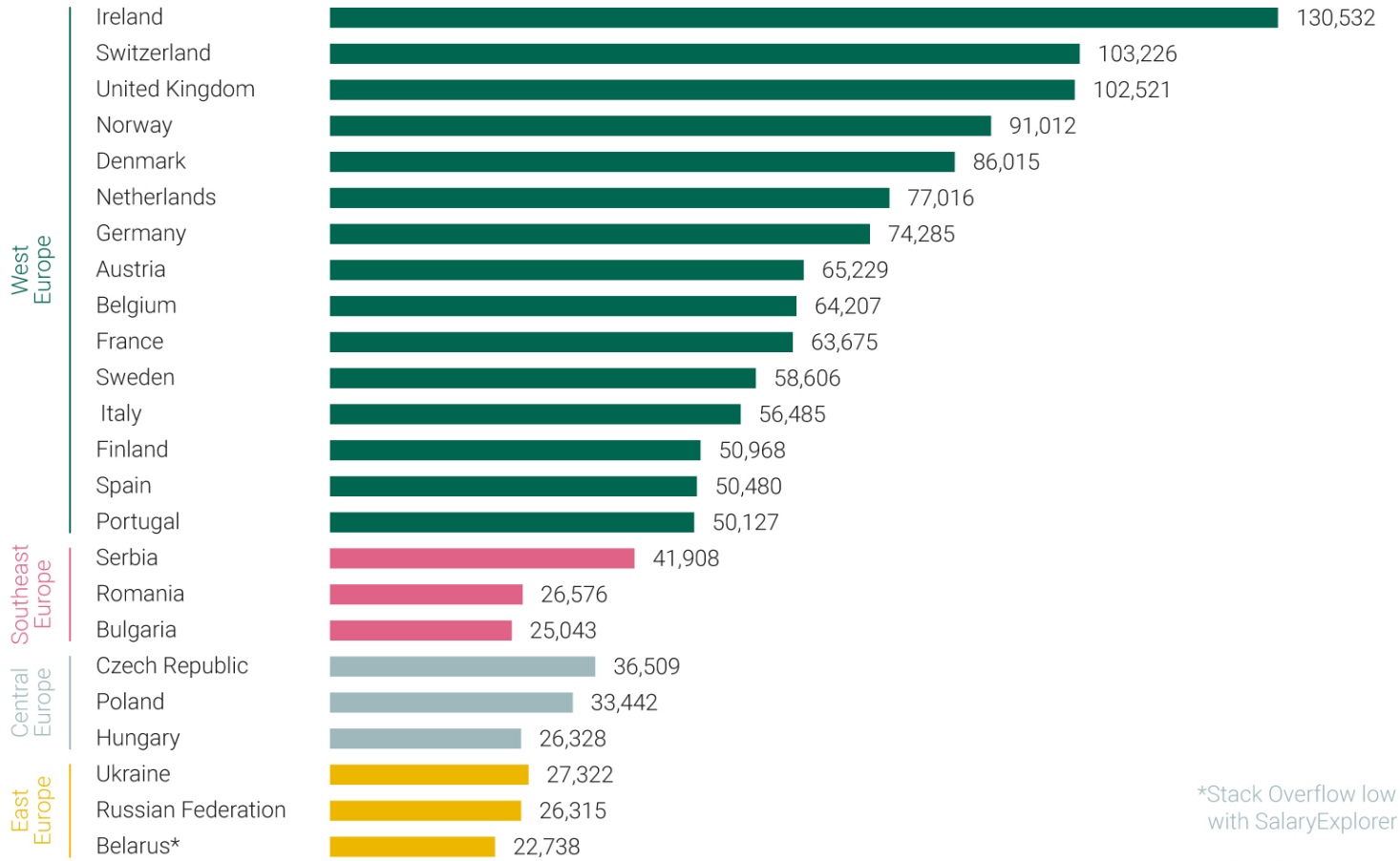
نحوه تفسیر نمودارهای گزارش

روی نمودارها ۵ عدد به صورت درصدی وجود دارد که بالای هر کدام از این درصدها یک عدد به میلیون تومان نوشته شده است. بطور مثال در نمودار زیر بالای ٪۷۵، عدد ۵/۶ میلیون تومان نوشته شده است. این بدین معناست که ٪۷۵ از افراد این جامعه آماری، کمتر از ۵/۶ میلیون تومان و ٪۲۵، بیش از ۵/۶ میلیون تومان دریافت می کنند.

برای فهم بهتر گزارش لطفا بخش **راهنمای گزارش** در صفحه ۸ را مطالعه کنید.

Python Developer Average Annual Salaries 2020: Europe, USD

BEETROOT

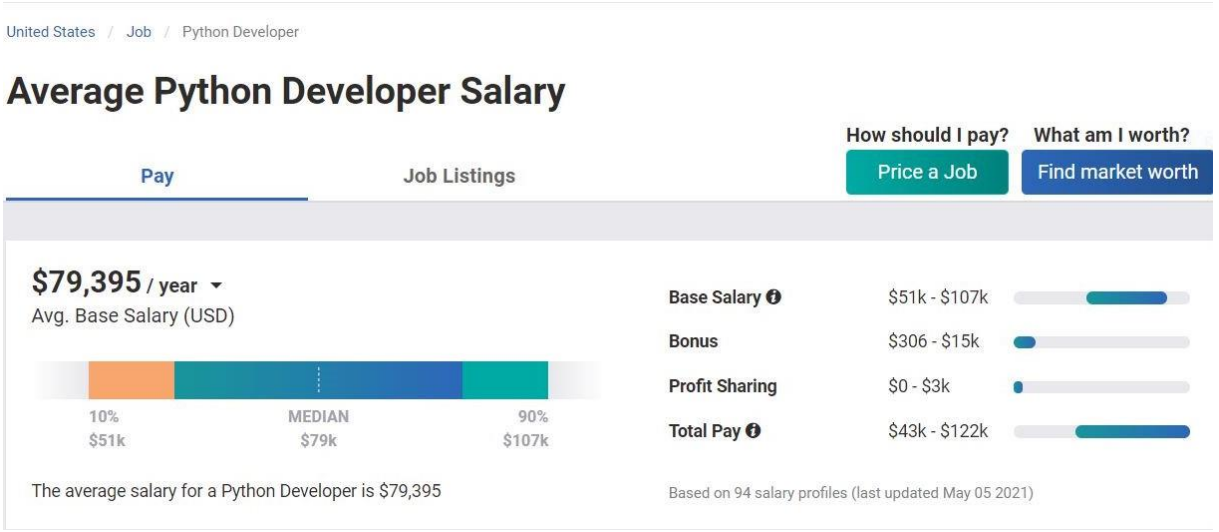


*Stack Overflow low confidence data averaged with SalaryExplorer data (Glassdoor data n/a)

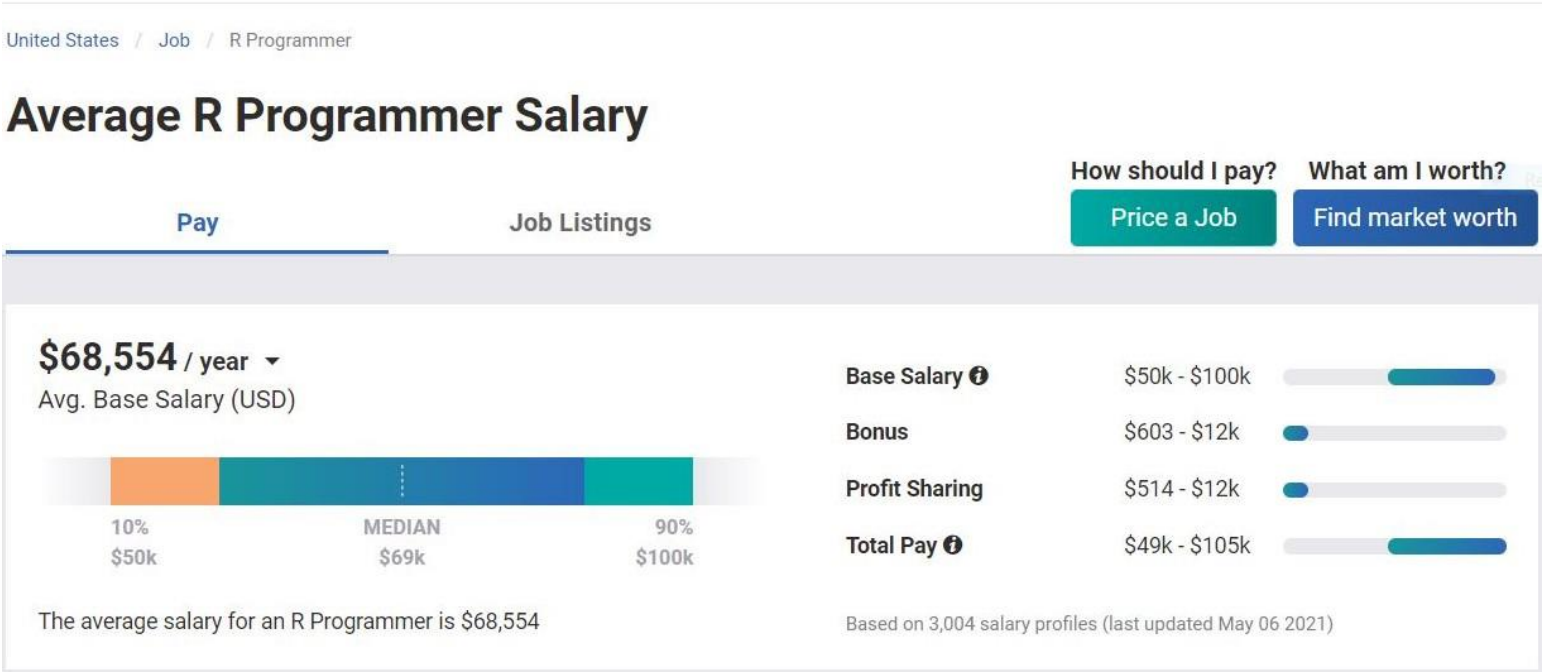
Source: insights.stackoverflow.com, glassdoor.com

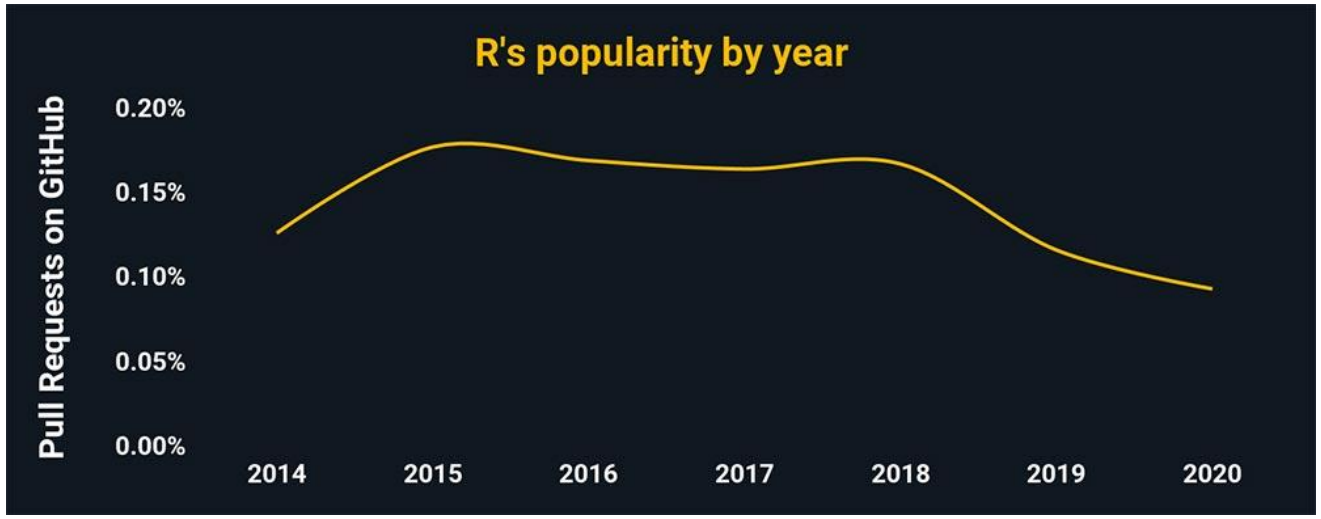
Python developer salary by industry

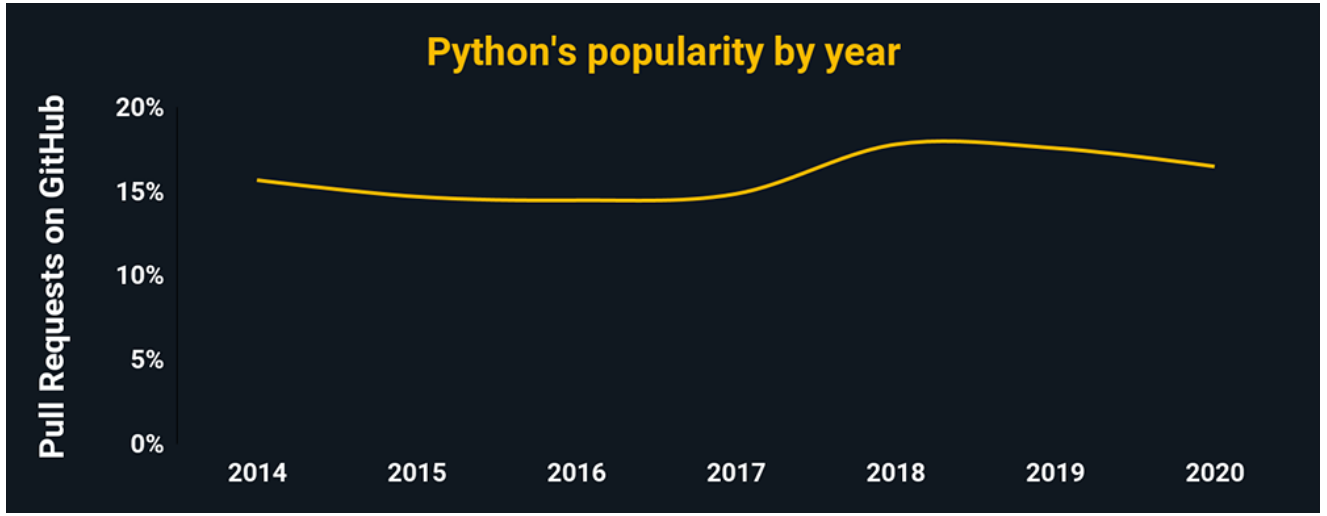
Banking and financial services:	$92,338
Energy	$85,338
Biotech and pharma	$78,503
Insurance, accounting and legal	$76,973
Travel and tourism	$74,602
Healthcare	$73,513
Marketing and advertising	$72,721
Automotive	$71,322
Government	$70,630
Education	$65,111



Source: PayScale.com







Source: flatironschool.com

Google Trends | Compare

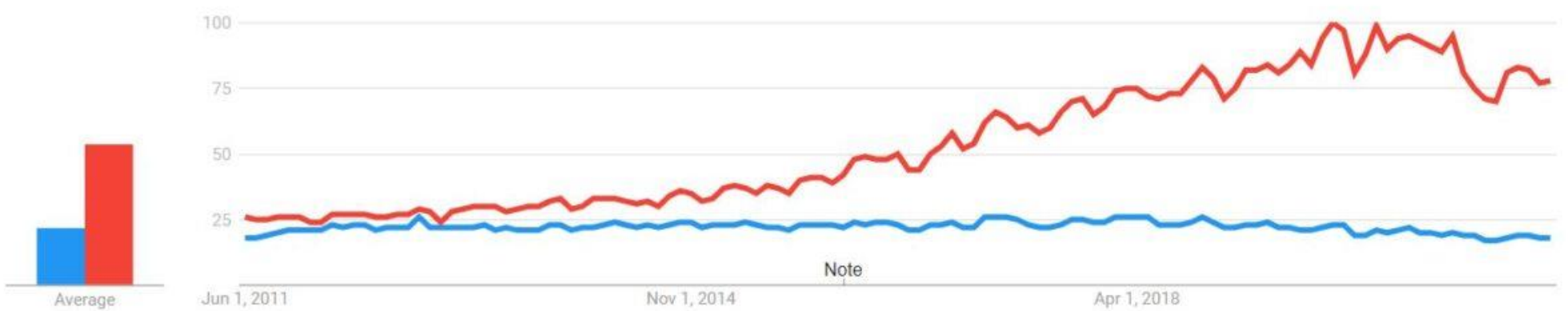
R
Programming language

Python
Programming language

+ Add comparison

Worldwide | 5/26/11 - 6/26/21 | All categories | Web Search

Interest over time



User loyalty, Python vs. R

91% OF PYTHON USERS REMAIN LOYAL TO PYTHON

10% switch from R to Python

5% switch from Python to R

74% OF R USERS REMAIN LOYAL TO R

Source: KDnuggets polls 2016

Usability

Coding and debugging is easy because of the simple syntax.

Statistical models can be written with only a few lines.

Ecosystem

The Python Package Index (PyPi) and Anaconda are repositories of Python software with all libraries. Users can contribute to these repositories.

Search through these sources easily with Rdocumentation. Packages are collections of R functions, data, and compiled code.

Ease of Learning

Python’s focus on readability and simplicity.

R is easier to learn when you start out, but the intricacies of advanced functionalities makes it more difficult to develop expertise.

Advantages

- Handling massive amounts of data
- Building deep learning models
- Performing non-statistical tasks, like web scraping, saving to databases, and running workflows

- Widely considered the best tool for making beautiful graphs and visualizations.
- Great for statistical analysis.

Disadvantages

- Visualizations are more convoluted in Python than in R, and results are not as eye-pleasing or informative.
- Python doesn’t have as many libraries for data science as R.

- Finding the right packages to use in R may be time consuming.
- There are many dependencies between R libraries.
- R can be considered slow if code is written poorly.
- Not as popular as Python for deep learning and NLP.

IDE

- There are many Python IDEs to choose from which drastically reduce the overhead of organizing code, output, and notes files. Jupyter Notebooks and Spyder are popular, and Jupyter Lab is gaining traction.
- RStudio is the most popular R IDE. It's available in two formats: RStudio Desktop for running locally as a regular desktop application and RStudio Server for access via web browser while running on a remote Linux server.

I also agree

- R users mainly consists of Scholars and R&D professionals while Python users are mostly Programmers and Developers
- R is difficult to learn at the beginning while Python is Linear and smooth to learn
- R is mainly used when the data analysis tasks require standalone computing or analysis on individual servers.
- R programming is better suited for statistical learning, with unmatched libraries for data exploration and experimentation. Python is a better choice for machine learning and large-scale applications, especially for data analysis within web applications.