



Co-funded by the
Erasmus+ Programme
of the European Union



Applied Data Science

Course Introduction

Prof. Gheith Abandah
أ.د. غيث علي عبندة

Outline

- Instructor Information
- Video: What is Data Science?
- Data Science Hierarchy of Needs
- Data Science Skills
- Textbook and References
- Course Objectives and Outcomes
- Course Outline
- Policies and Grading
- Important Dates

Instructor Information

- **Instructor:** Prof. Gheith Abandah
- **Email:** abandah@ju.edu.jo
- **Office:** CPE 406
- **Home page:** <http://www.abandah.com/gheith>
- **MS Team:** [Link](#)
- **Office hours:** Sun – Thu: 13:00 – 14:00

What is Data Science?

- YouTube Video from **Joma Tech**

What REALLY is Data Science? Told by a Data Scientist

<https://youtu.be/xC-c7E5PK0Y>

Data Science Hierarchy of Needs

THE DATA SCIENCE **HIERARCHY OF NEEDS**

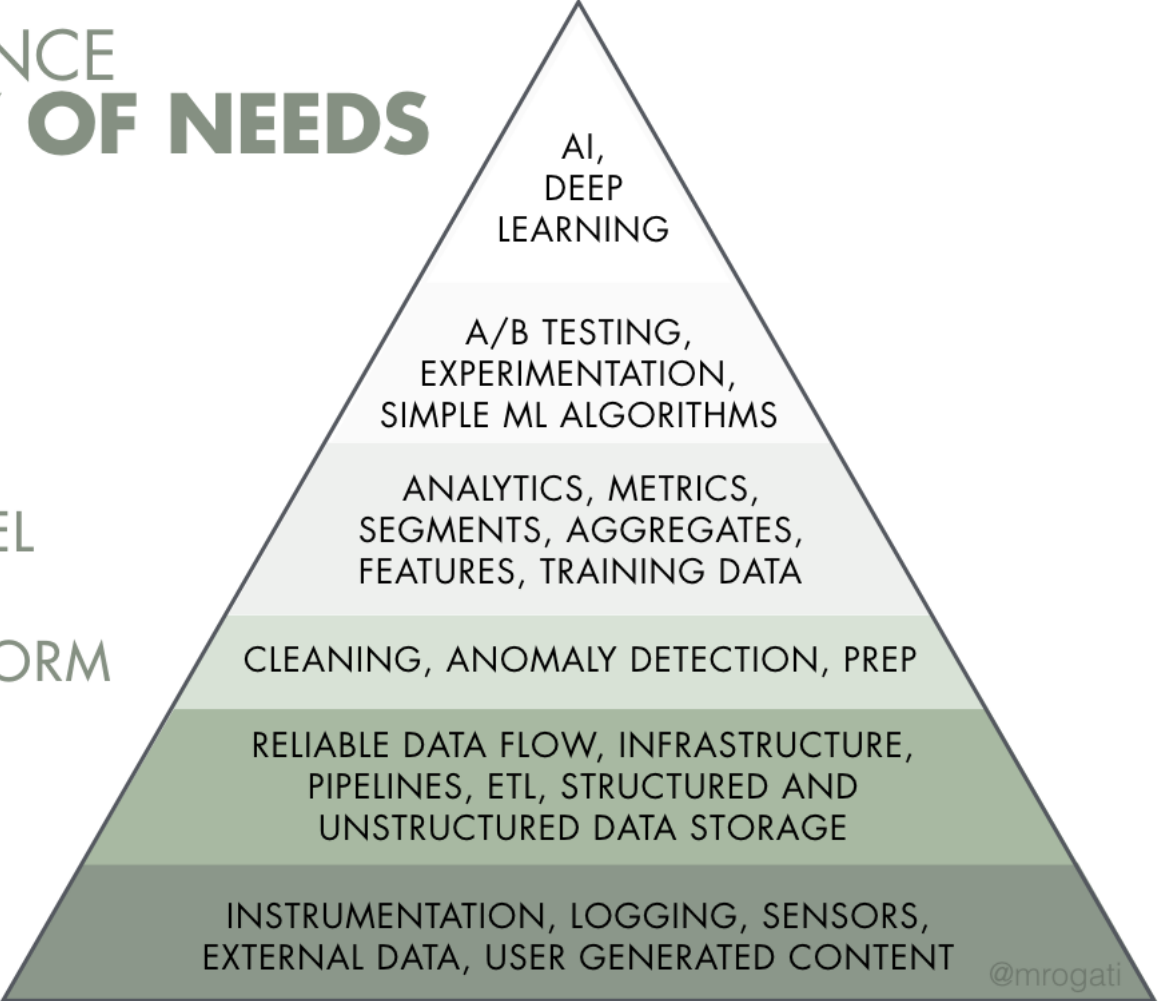
LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

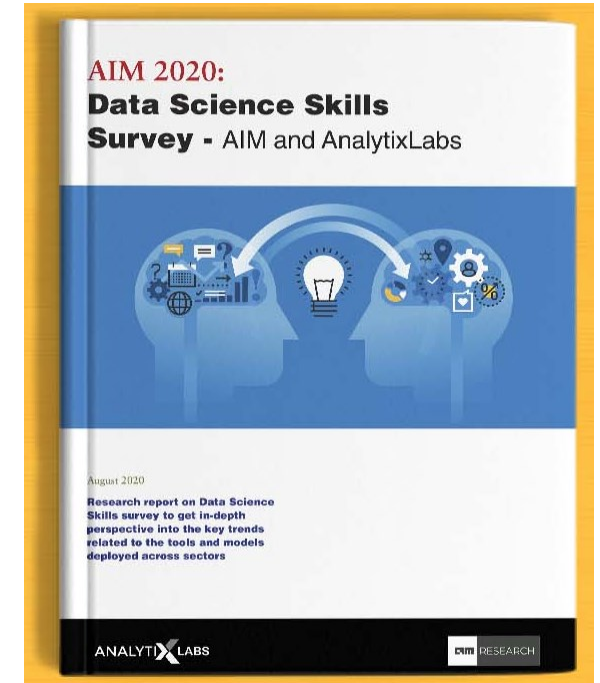
MOVE/STORE

COLLECT

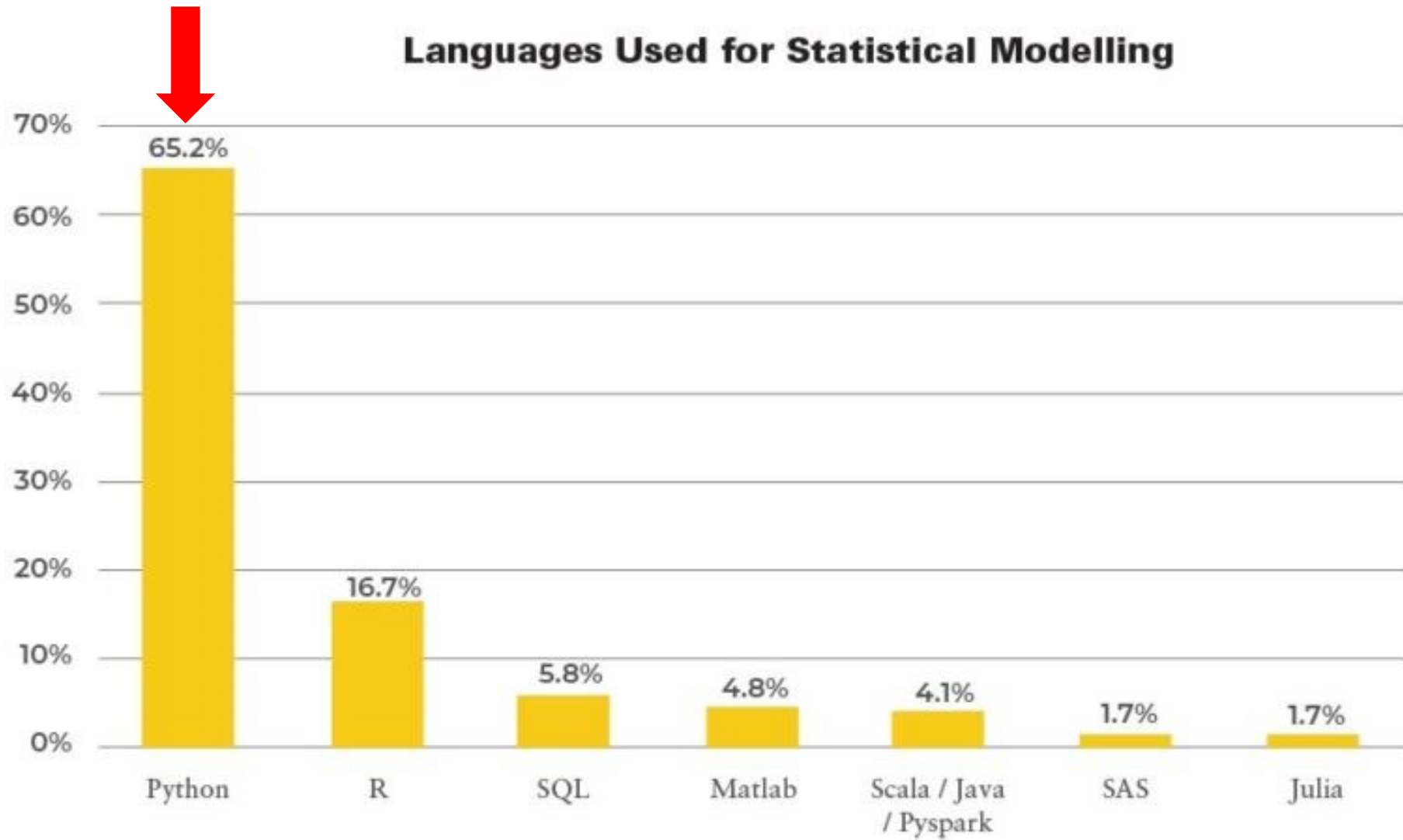


Data Science Skills

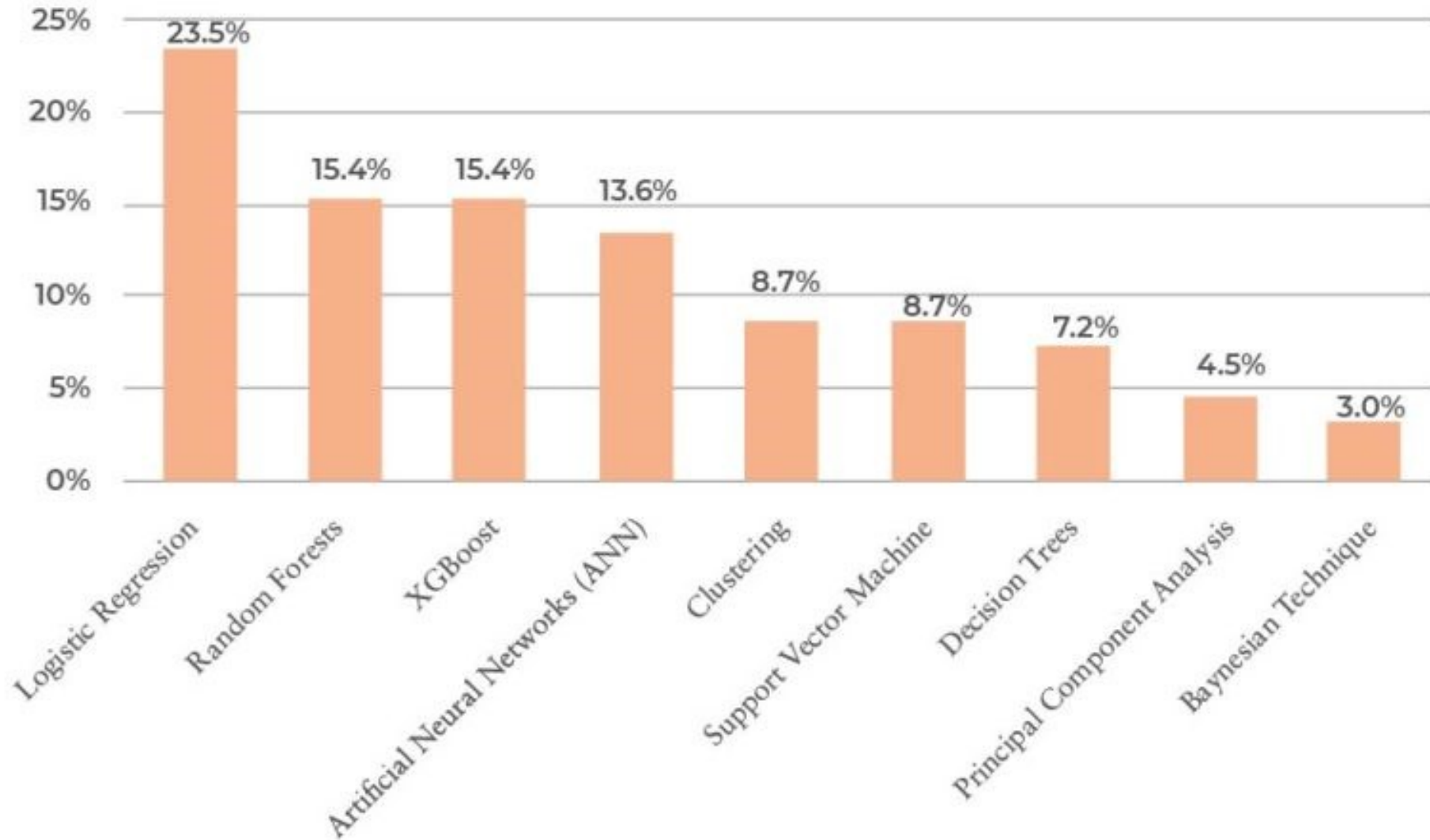
- **Data Science Skills Study 2020**
- <https://analyticsindiamag.com/aim-2020-data-science-skills-survey-aim-and-analytix-labs/>
- By AIM and AnalytixLabs
- Released on 17/8/2020



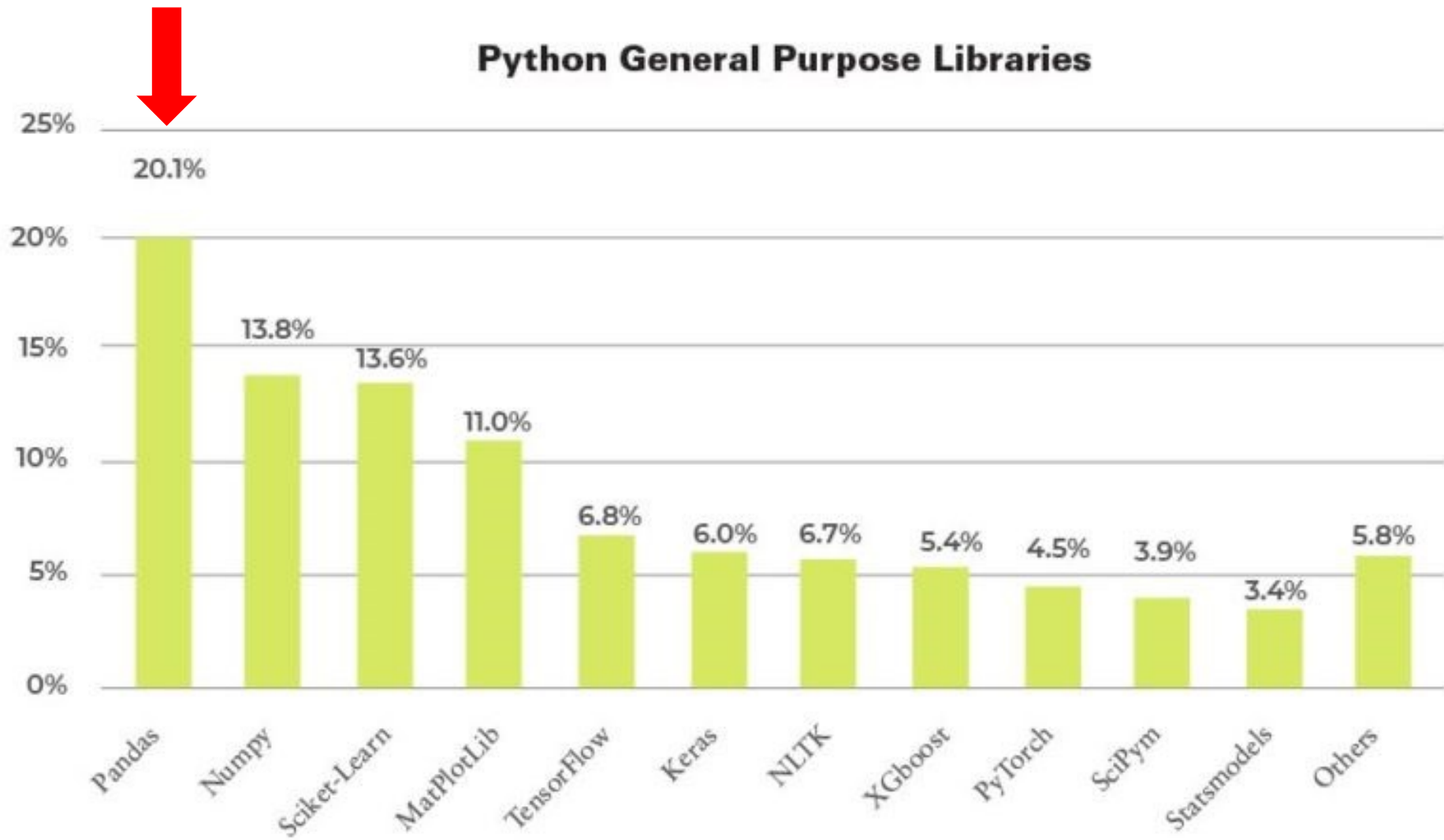
Languages Used for Statistical Modelling



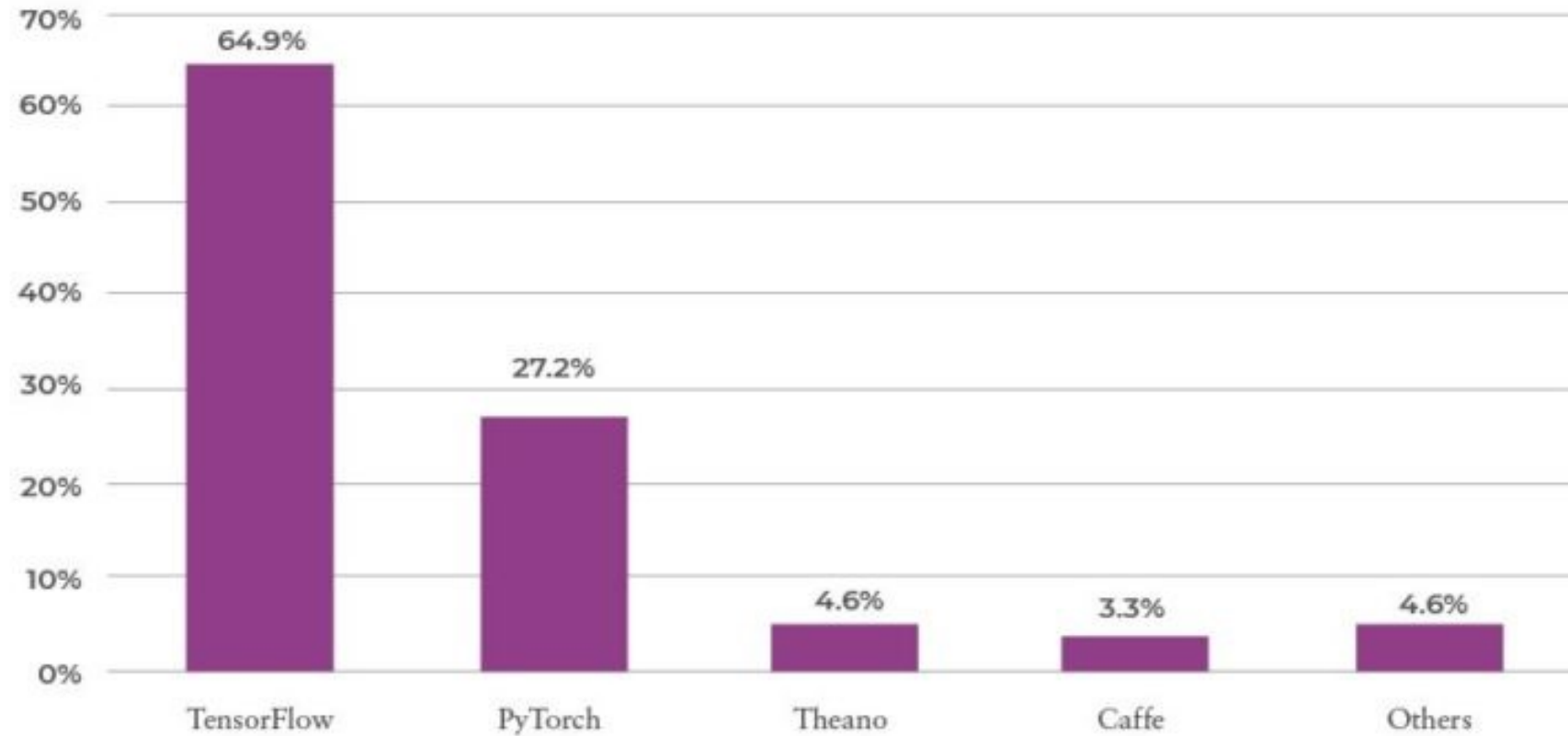
Data Science Models



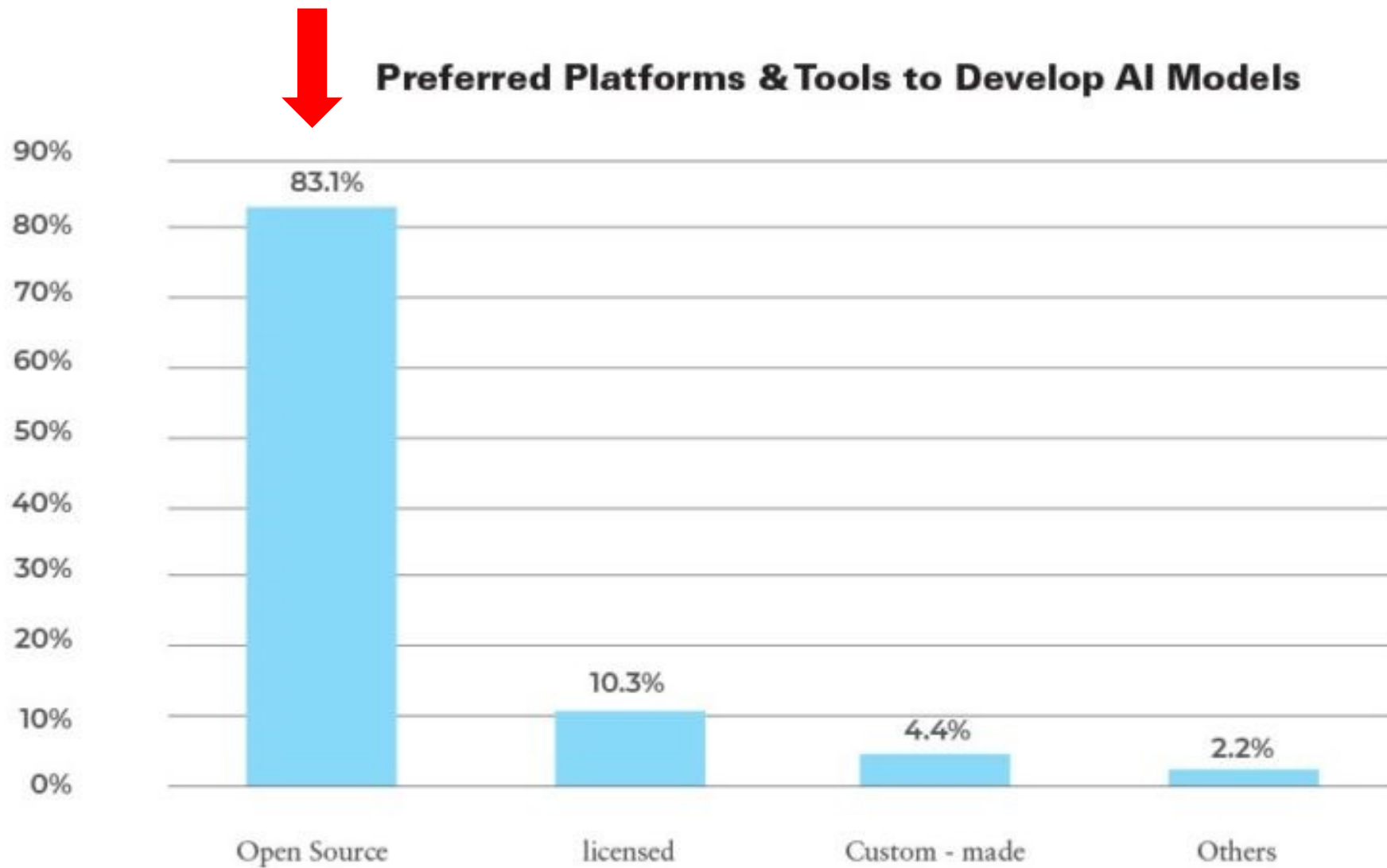
Python General Purpose Libraries



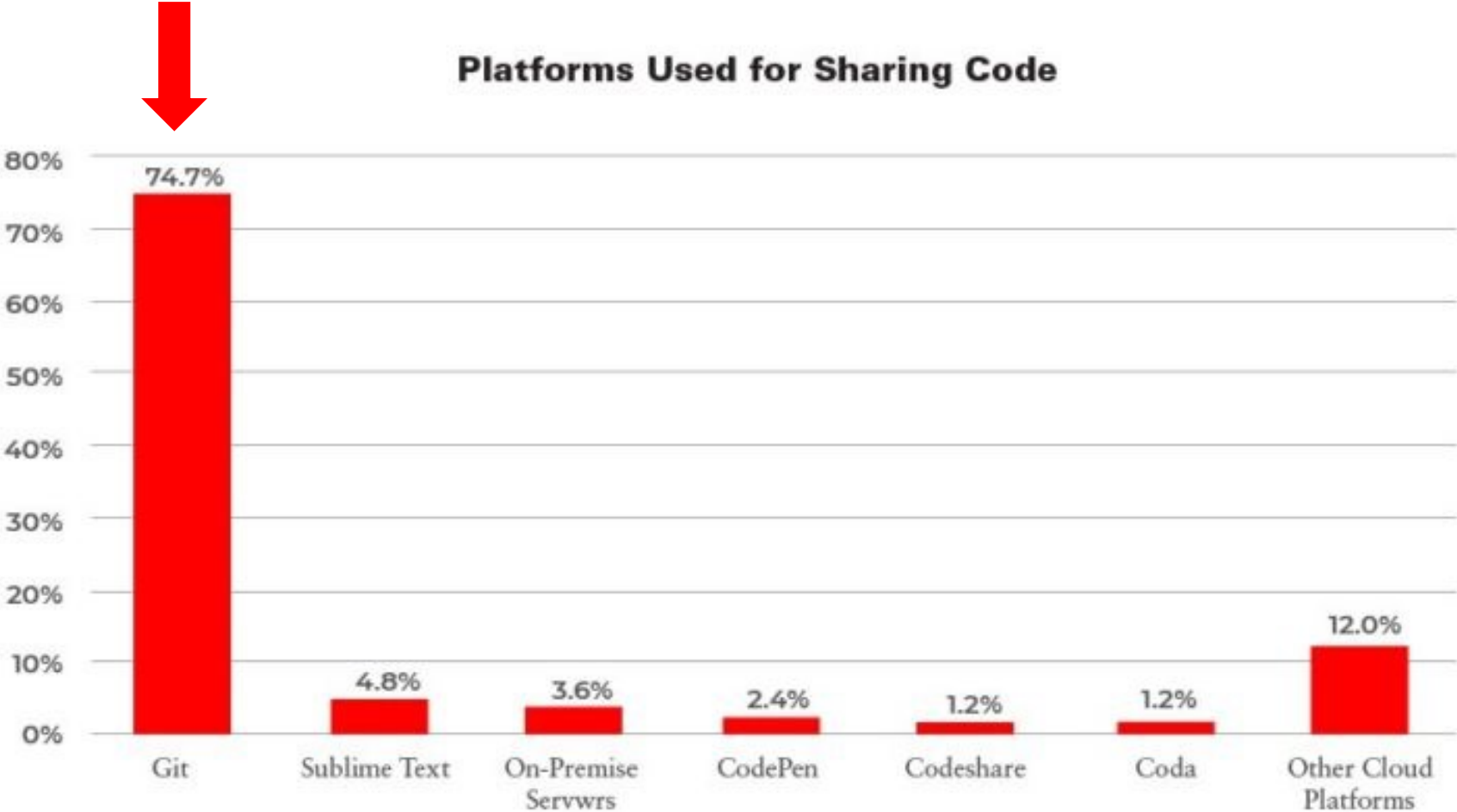
Python Frameworks Utilized for AI / Deep Learning Projects



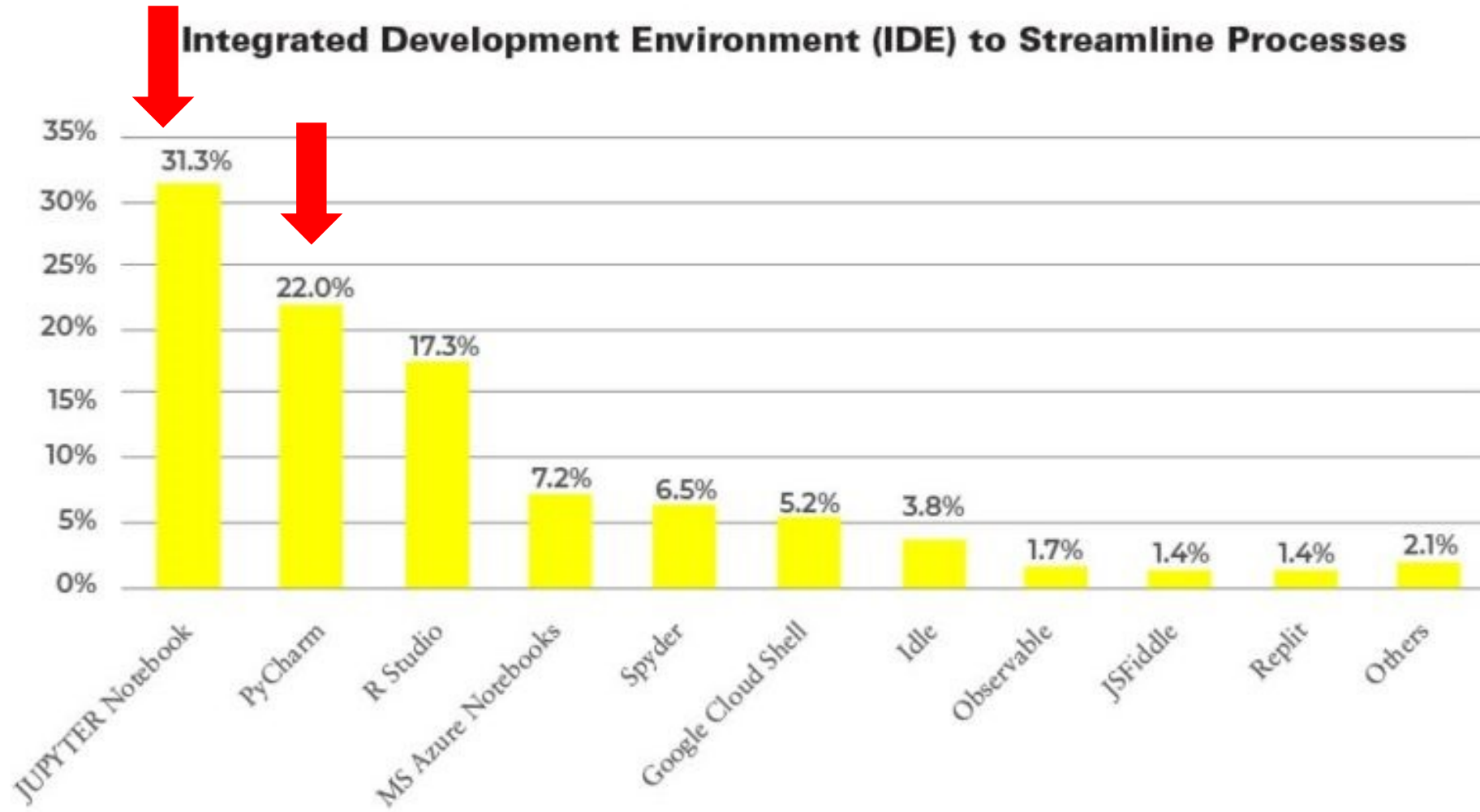
Preferred Platforms & Tools to Develop AI Models



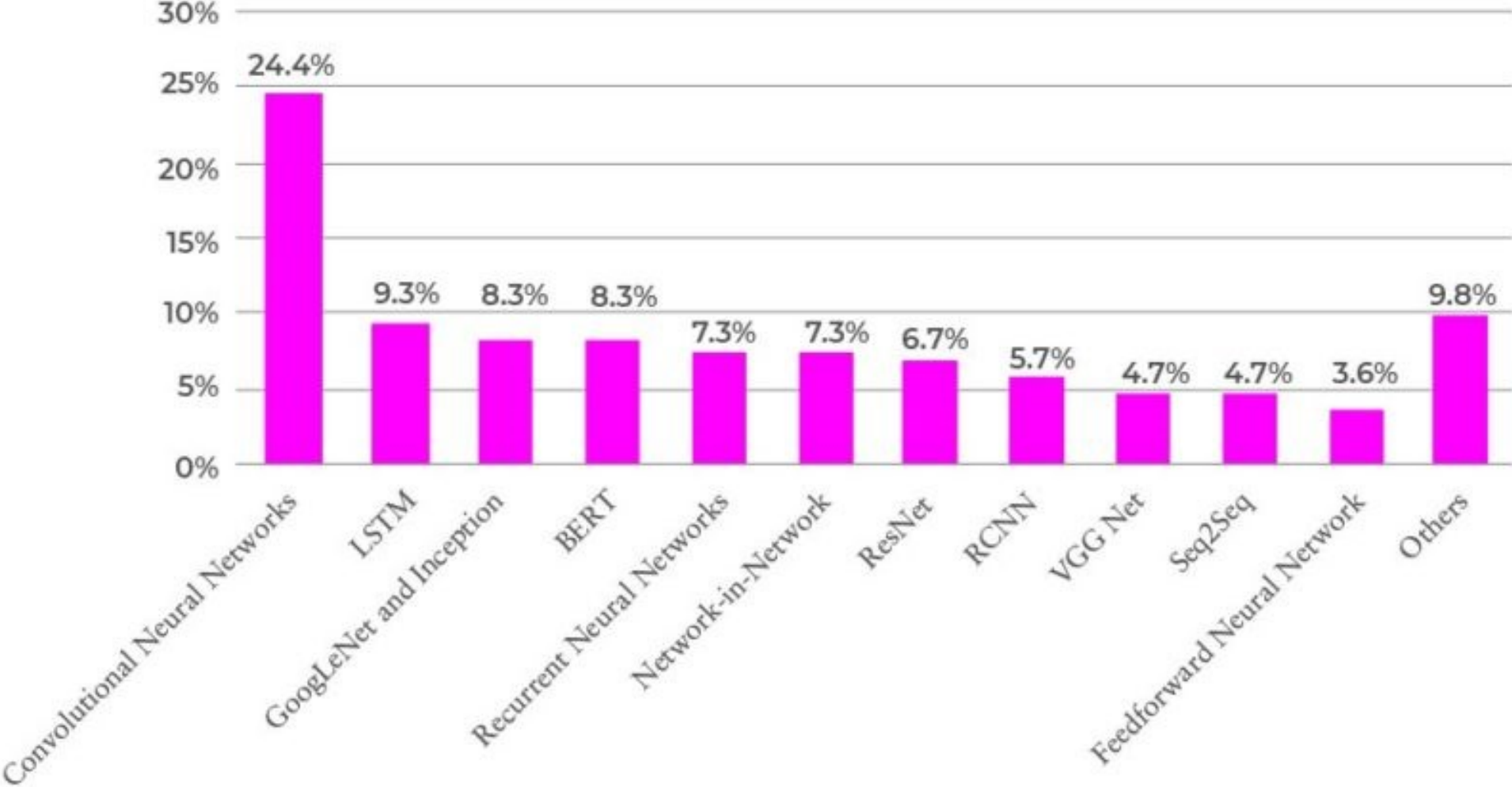
Platforms Used for Sharing Code



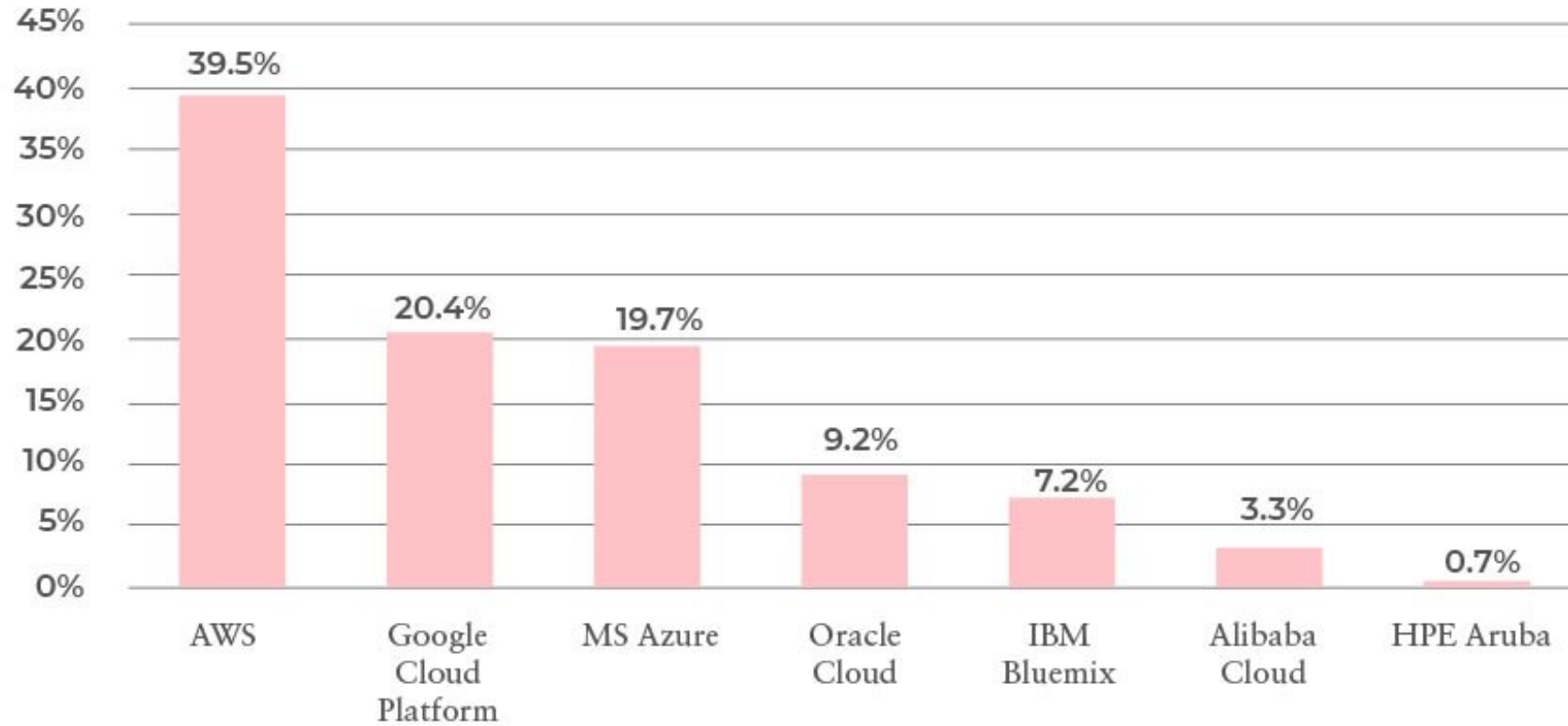
Integrated Development Environment (IDE) to Streamline Processes



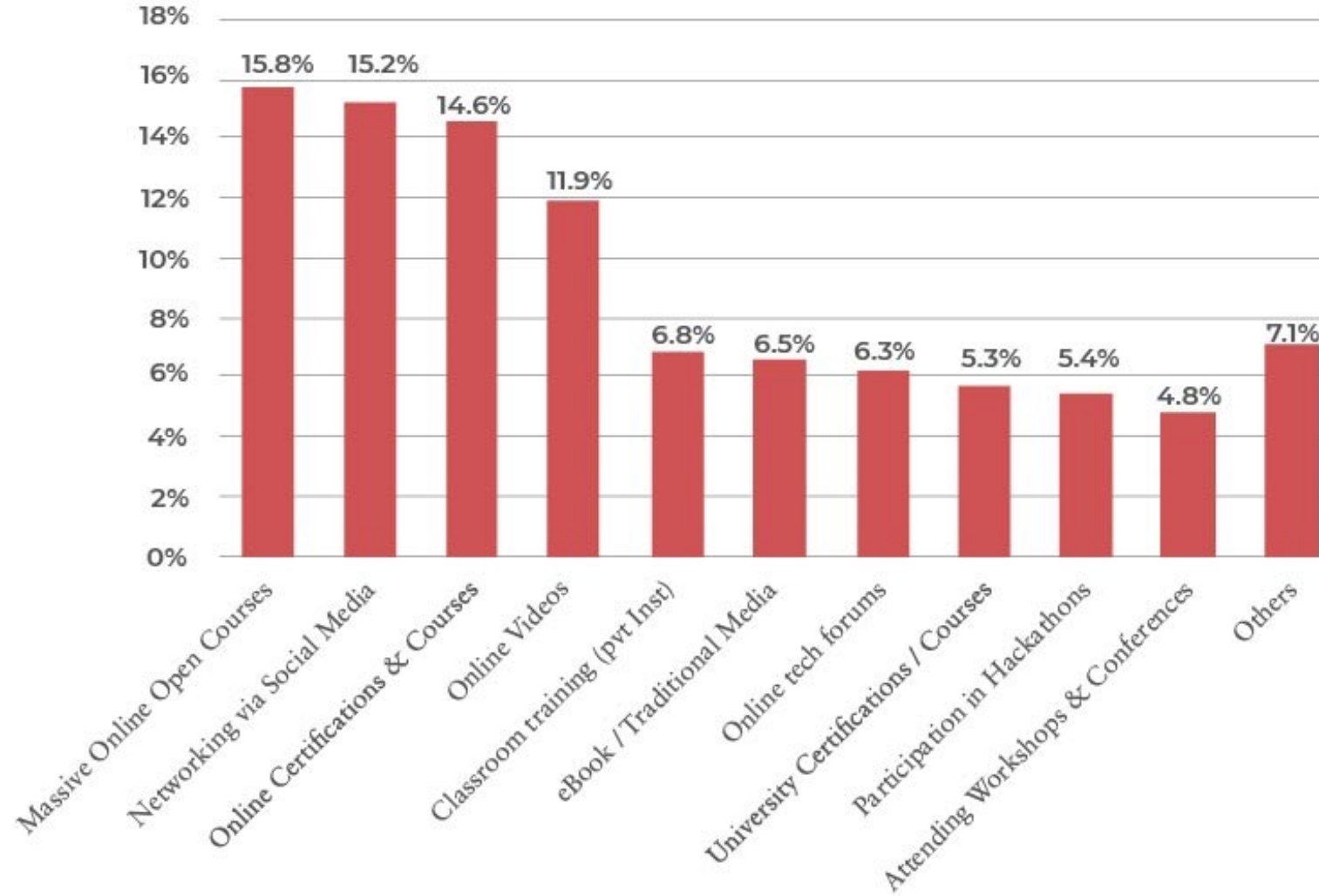
Neural Network Architectures



Cloud Service Platforms to Develop AI / ML Models

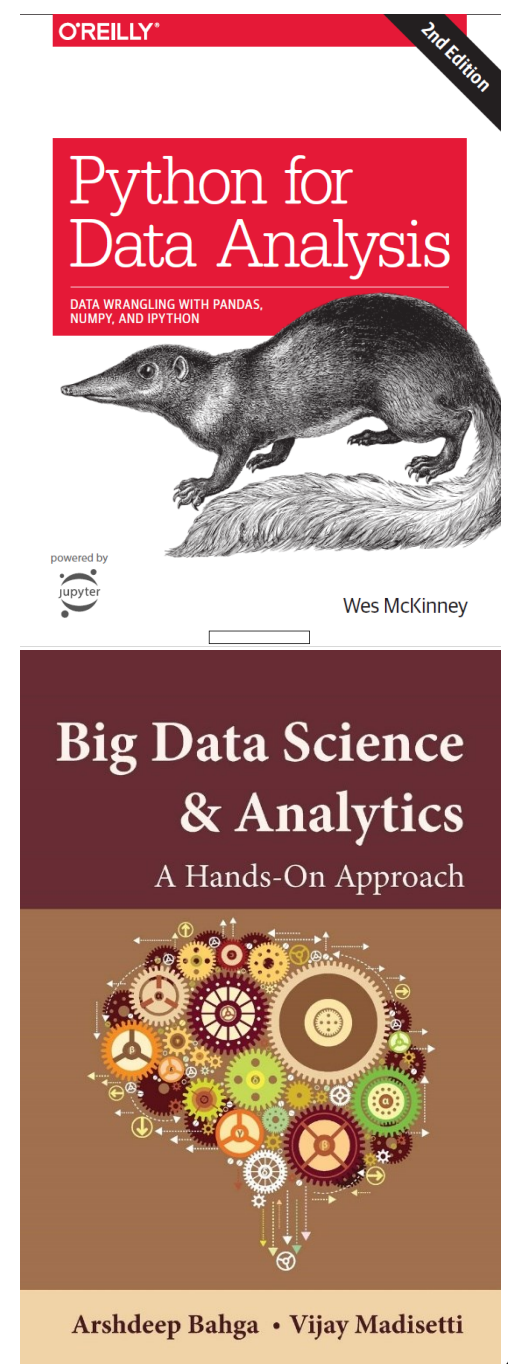


Learning Resources Utilized to Upskill



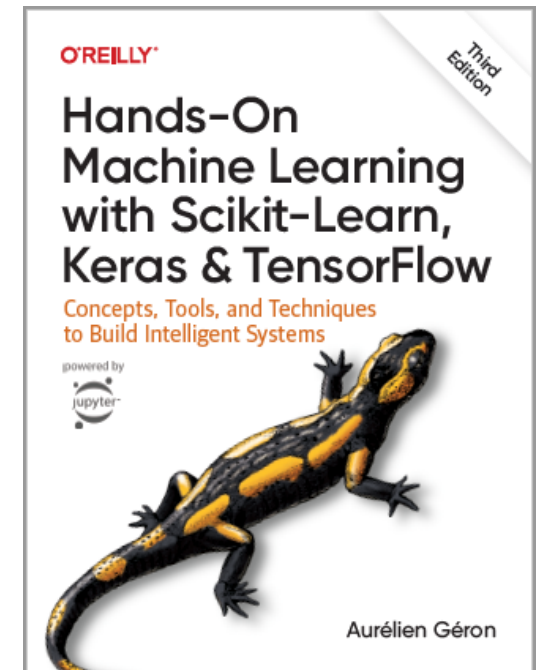
Textbooks

1. Wes McKinney, **Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython**, O'Reilly Media, 2nd Edition, 2018.
2. Arshdeep Bahga and Vijay Madisetti, **Big Data Analytics: A Hands-On Approach**, 2019.
3. Course web page:
http://www.abandah.com/gheith/?page_id=3031



References

1. Jake VanderPlas, **A Whirlwind Tour of Python**, O'Reilly Media, 2016.
2. Joel Gurs, **Data Science from Scratch**, O'Reilly Media, 2015.
3. Aurélien Géron, **Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts: Tools, and Techniques to Build Intelligent Systems**, 3rd Edition, O'Reilly Media, Oct 2022.



Course Objectives

- Introduce students to the practical techniques used in data analytics including loading, cleaning, preparation, wrangling, visualization, and analysis.
- Introduce students to the basic concepts and techniques in big data.

Course Outcomes

- Use Python and its specialized libraries to gain insight from data and solve problems.
- Know the main concepts and techniques used in handling big data and performing data analytics.

Course Outline

Week	Topic	ILO	Resources
1	Course Introduction	1	3
2+3	Pandas Data Structures, Essential Functionality & Descriptive Statistics	1	1
4+6	Data Loading, Storage and File Formats	1	1
6+7	Data Cleaning and Preparation	1	1
8	Data Wrangling: Join, Combine and Reshape	1	1
9+10	Plotting and Visualization with Matplotlib and Seaborn	1	1
11	Data Aggregation and Group Operations	1	1
12	Time Series	1	1
13	Introduction to Big Data, Architectures, and Patterns	2	2
14	MapReduce Patterns	2	2
15	Project Presentations	1+2	3

Policies

- Attendance is required
- Makeup exams need acceptable absence cause
- Late penalty is 25%
- All submitted work must be yours
- Cheating will not be tolerated
- Open-book exams
- Join the Microsoft Team at: [Link](#)
- Check department announcements at:
 - <https://www.facebook.com/profile.php?id=100087040924274>

Grading

Assessment tool	Mark	Topic(s)	Time
Homework Assignments	10%	Programming aspects	W2-W14
Midterm exam	30%	First 8 weeks	W8
Term project and presentation	20%	Practical and presentation aspects	W15
Final exam	40%	All material	W16
Total	100%		

Important Dates

Thu 2 Mar 2023	First Lecture
16-30 Apr 2023	Midterm Exam Period
Sun 4 Jun 2023	Last Date to Withdraw
Thu 1 Jun 2023	Last Lecture (project presentations)
8-20 Jun 2023	Final Exam Period