Introduction to Python for Information Systems

What will we cover today

- About the course
  - About me, About you, your expectations
  - How this course has been put together
  - Rubric
  - A Framework for praogramming

About me

- HELLO
  - My name is Pramod Gupta
- PhD in Electrical and Computer Engineering, from McMaster University, Canada
- Experience: Academics, NASA, EMC, GE, VISA etc.
- Been teaching Data Science related courses at various universities
- Independent data science consultant

Philosophy

- Sharing what I have Learned
- Let us make this an interactive experience. I would like this class to be as open and interactive
- Feel free to stop and ask for clarifications
- If you have a doubt, always better to ask
- Don’t be shy or hesitant
- Knowledge flow is bi-direction. Age is never a factor in knowledge. It is the experience.
- If there is any issue, please bring it to my notice as soon as possible so that we can address it.
- Goal: “ Learning to self-Learn

Your turn

Quick Introductions

- **Name, Profession**
- **Experience in analysis do you have?**
- Which analytics tools have you used/familiar
- What are your expectation and motivation from this course?
- What are the areas where you would like to apply and analyze (you can upload to the class forum)

Course Logistics

Goals and objectives

- Software is a huge field, and there is no way you can master it by reading a single book or taking a single course. The goal of this course is to give a solid foundation in the most important tools.

- This course is concerned with the nuts and bolts of Python. It is also a practical, modern introduction to scientific computing in Python, tailored for data-intensive applications.
- This course is also about the parts of Python you will need to effectively solve a broad set of real-world problems.

Goals and objectives

- This class is designed to provide you with the tools you need for solving real world problems using statistics and a better understanding of software techniques.

- We plan to achieve these goals by introducing you to the relevant statistical knowledge, how to use Python to solve problems through homework, discussion, class participation and project.

- At the end of the course:
  - You will have the tools to tackle a wide variety of real-world challenges, using the best parts of Python.
  - Be Trained how to present your project and results to a diverse audience

Course Description

- 15 total sessions
- 13-14 lecture days
- 1-2 sessions project presentation
- Each lecture will be divided into different topics and will consists of
  - Presentation
  - Demonstration
  - Hands-on activities
  - Discussion

Course Description

- Work Load
  - You are expected to put in 6-7 hrs. per week of work outside of class. A few of you will do well with less time than this, and a few of you will need more but do not worry.
  - Lecture notes, homework assignments and scripts will be available online on Canvas
  - Assignments should be submitted through Canvas.

Homework

- Homework
  - There will be 4-5 assignments and will comprise of problems related to the topic covered in the lecture. The objective of the assignments is to help you develop more in-depth understanding of the material covered in the lectures.
  - Do not copy someone else work, both parties will be penalized. Though you are welcome and encourage to work with each other.
  - Quizzes

About final project

- Aim: turning software techniques you learn in class to become your strength in dealing real world problems

- The project involves preparation of a report and presentation of the results during the last week of the class. The project will be done in groups of 2~3 students. If you already working on a research project in the area of interest you are encouraged to use dataset/topic from your research provided you make some extra effort for the class.

- Detailed instructions for the project will be posted later

A data analysis write up

Should tell a full story

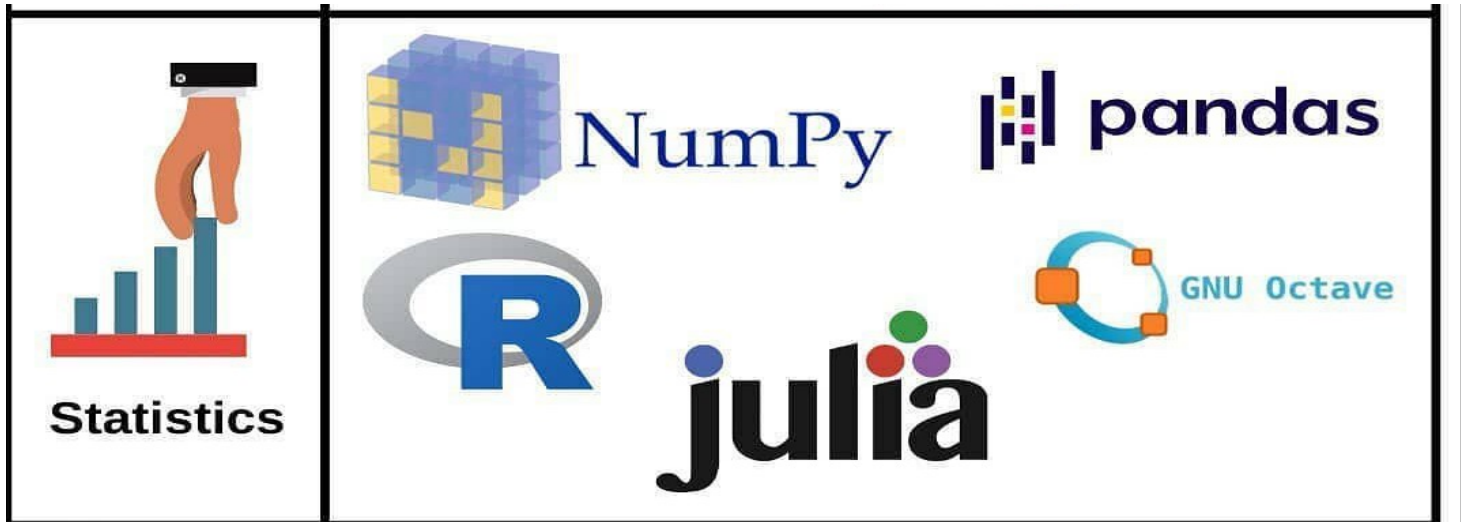
- Title, Introduction (Motivation)
- Method used (justification)
- Caveats
- Results
- References

Do not wait!
Get started early.
You will learn more that way

Who is this course for?

- Anyone who is interested in:
  - Helping companies make decisions aided by data
  - Refreshing some theory learned in school, but with a practical focus
  - Getting up to speed with new Open Source tools and libraries
  - Curious about the new technology

Data Science ToolBox



Data Science Toolbox

Data Visualization

Data-Driven Documents

Bokeh

plotly

tableau

RStudio Shiny

matplotlib

ML Tools

theano

scikit learn

Keras

PyTorch

dmlc XGBoost

DL4J

H_2O.ai

Data Science Toolbox

crystal reports

tableau SOFTWARE

BI Tools

IBM COGNOS

BRIO SOFTWARE

MicroStrategy®
Best In Business Intelligence™

Focus on the results, not the tools

- It is okay to use multiple tools
- Whatever you know to get the job done
- Specialized Tools: The best tool for the right job

Depending on the task, data scientists can avail of tools that are scalable, performant, require less code, and contain a lot of features. On the other hand this approach requires a lot more context-switching, and extra effort is needed to annotate long workflows.

Ben Lorica

What is Python

- **Python** is a general purpose, dynamic, high-level, and interpreted programming language. It supports Object Oriented programming approach to develop applications, created by Guido Rossum in 1989.
- It is simple and easy to learn and provides lots of high-level data structures.
- Python is *easy to learn* yet powerful and versatile scripting language, which makes it attractive for Application Development.
- Python's syntax and *dynamic typing* with its interpreted nature make it an ideal language for scripting and rapid application development.
- Python supports *multiple programming pattern*, including object-oriented, imperative, and functional or procedural programming styles.

What is Python

- Python is not intended to work in a particular area, such as web programming. That is why it is known as *multipurpose* programming language because it can be used with web, enterprise, 3D CAD, etc.
- We don't need to use data types to declare variable because it is *dynamically typed* so we can write a=10 to assign an integer value in an integer variable.
- Python makes the development and debugging *fast* because there is no compilation step included in Python development, and edit-test-debug cycle is very fast.
- It is ideally designed for rapid prototyping of complex applications.
- It has interfaces to many OS system calls and libraries and is extensible to C and C++.

What is Python

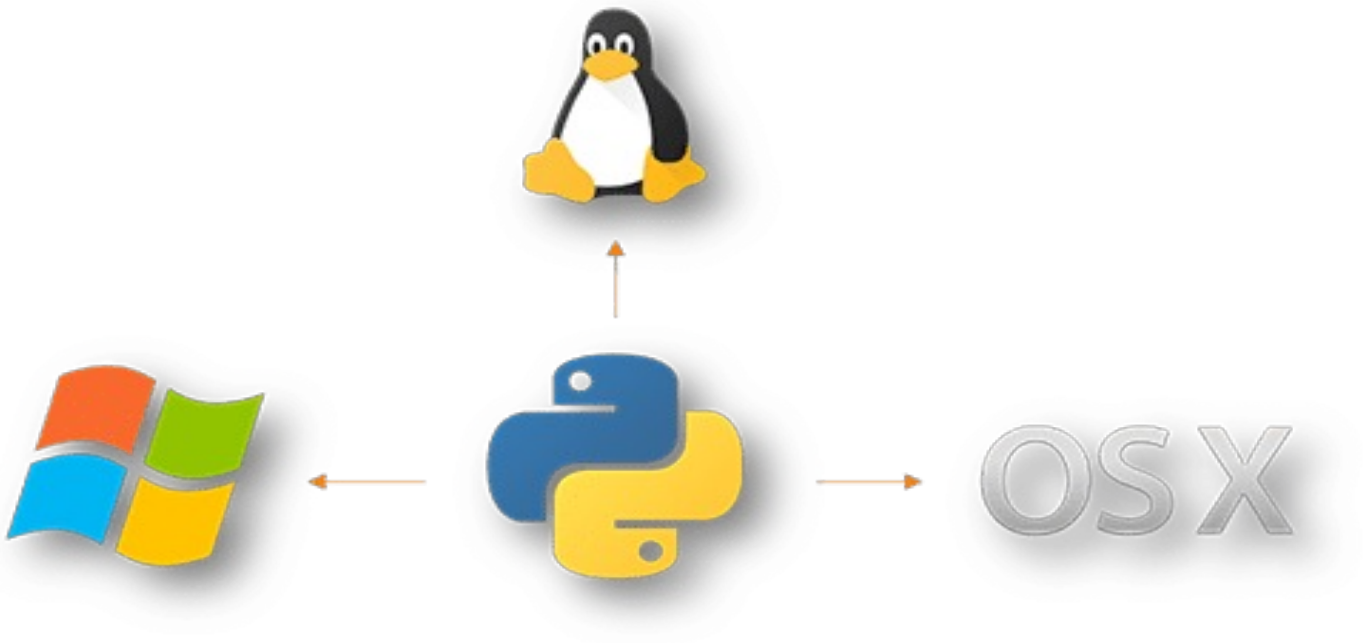
- It is a very popular language as many companies use the Python (NASA, Google, Facebook, Linkedin etc.).
- Python is widely used in Artificial Intelligence, Natural Language Generation, Neural Networks and other advanced fields.

What is Python

- It is a very popular language as many companies use the Python (NASA, Google, Facebook, Linkedin etc.).
- Python is widely used in Artificial Intelligence, Natural Language Generation, Neural Networks and other advanced fields.

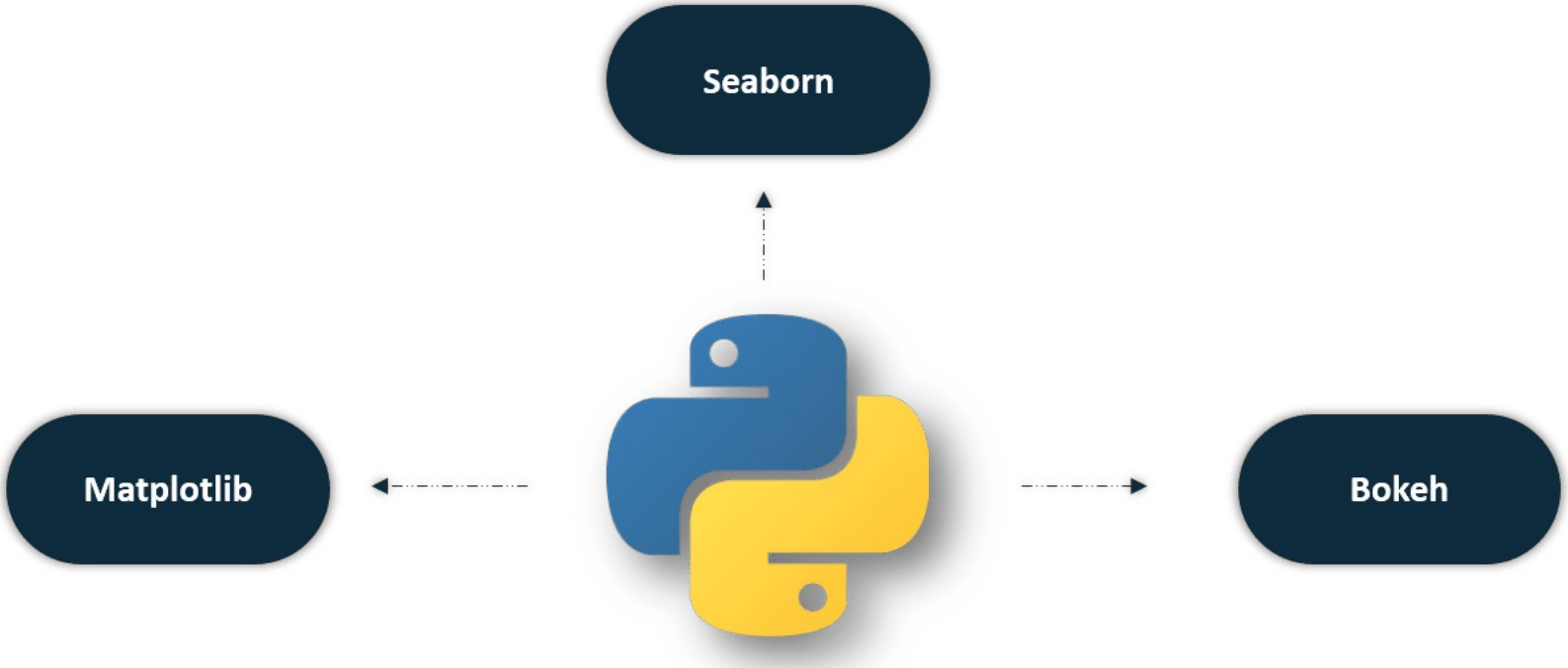
What is Python

- Python is also a cross-platform compatible language. So, what does this mean? Well, you can install and run Python on several operating systems. Whether you have a Windows, Mac, or Linux, you can rest assured that Python will work on all these operating systems.



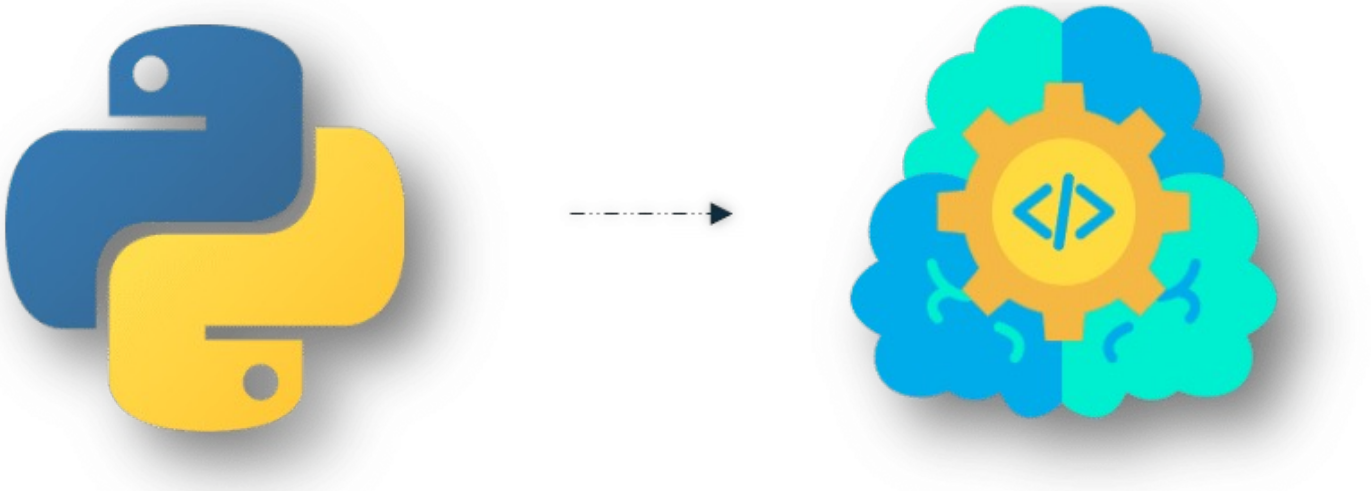
What is Python

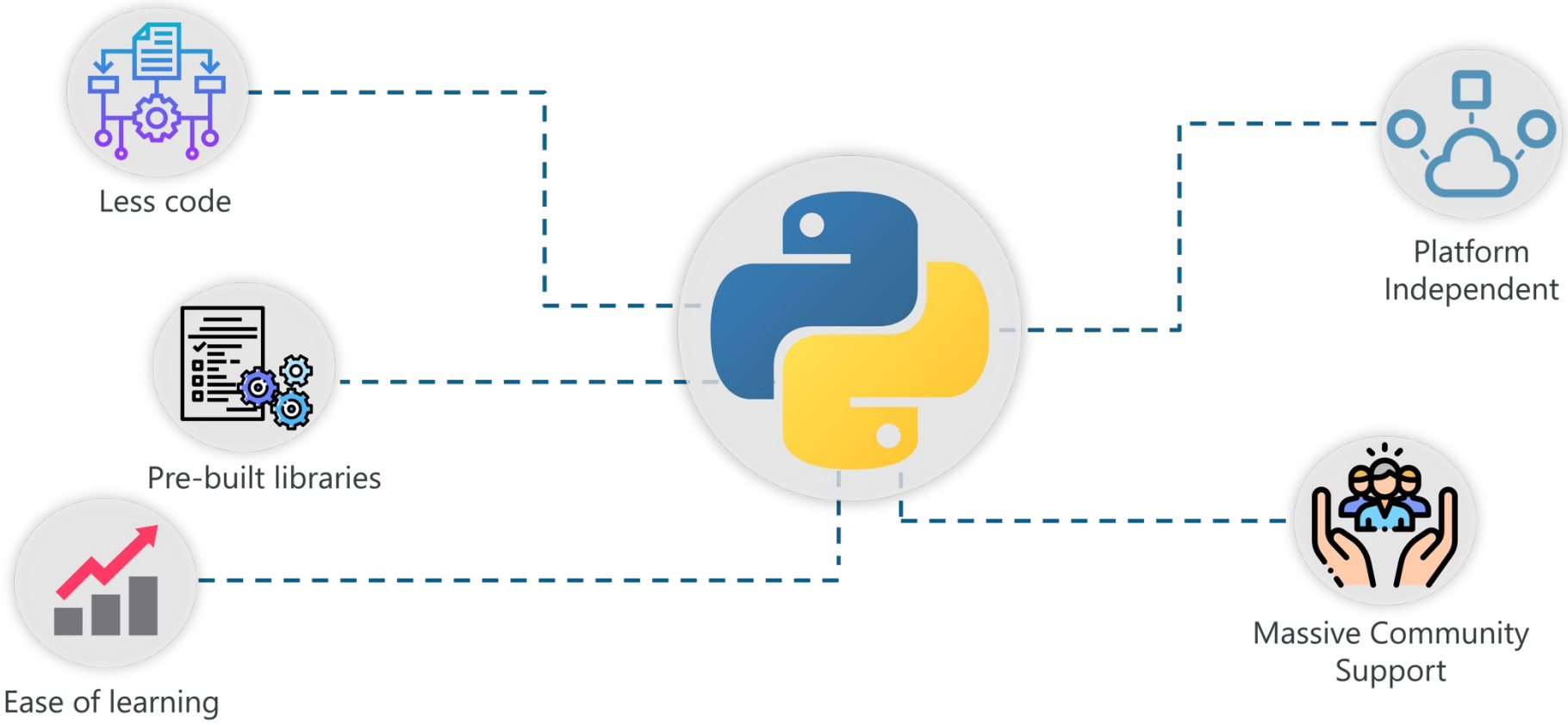
- In addition, Python is the most popular language for Machine learning and Deep Learning. Today, all top organizations are investing in Python to implement Machine Learning in the back-end.



What is Python

- In addition, Python is the most popular language for Machine learning and Deep Learning. Today, all top organizations are investing in Python to implement Machine Learning in the back-end.




Less code
Pre-built libraries
Ease of learning
Platform Independent
Massive Community Support

Why Python?

- Python provides great functionality to deal with mathematics, statistics and scientific functions. When it comes to data science application, it provides extensive libraries to deal with.
- It is not only open-source but also interpreted and high level tool
- Most importantly, Python is widely used in the scientific and research communities because of the ease of use, it's simple syntax makes it easy to adapt for people who even do not have an engineering background.
- Thanks to huge user base, just about every function that you might need for data analysis is available, often through open source extensions (known as packages) made available by the community.

Why Python?

- In summary, some reasons which go in favor of learning Python
  - Open Source – Free to install and use
  - Awesome online community
  - Very easy to learn
  - Libraries and Frameworks - Python has numerous libraries for different needs. Django and Flask are two of the most popular for web development and NumPy and SciPy and Pandas are for Data Science.

Though there are many advantages, it has few drawbacks too:

  - It is an interpreted language rather than compiled language, hence might take up more CPU time. However, given the savings in programmer time (due to ease of learning), it might still be good choice.
  - Python can be a challenging language for building highly concurrent, multithreaded applications, particularly applications with many CPU-bound threads. The reason for this is that it has what is known as the global interpreter lock (GIL), that prevents the interpreter from executing more than one Python instruction at a time.

Why to use Python

- The style of coding is easy
- The Python is free and open source. No need to pay any subscription charges.
- Python is effectively platform independent
- It is a full programing language
- Python is on the cutting edge, and expanding rapidly, 100,000+ modules are available

Why to use Python

- Python has unrivaled help resources. The community support is overwhelming. There are numerous forums to help you out.
- Python makes the best graphics
- One of the highly sought skill by analytics and data science companies
- comes standard with some of the most flexible and powerful graphics routines available anywhere

There are may more benefits. But , these are the ones which seems to be important and have kept me going.

Cont’d


10 reason to learn
Python
Ease of use
Extensive Support Libraries
Machine learning
Extensible
Embeddable
Raspberry Pi
Speedy
Web development
Server-side scripting
A less limited programming approach

Why to use Python

- Python has unrivaled help resources. The community support is overwhelming. There are numerous forums to help you out.
- Python makes the best graphics
- One of the highly sought skill by analytics and data science companies
- comes standard with some of the most flexible and powerful graphics routines available anywhere

There are may more benefits. But , these are the ones which seems to be important and have kept me going.

Python Popular Frameworks and Libraries

- Python has wide range of libraries and frameworks widely used in various fields such as machine learning, artificial intelligence, web applications, etc. We define some popular frameworks and libraries of Python as follows.
- **Web development (Server-side) -** Django Flask, Pyramid, CherryPy
- **GUIs based applications -** Tk, PyGTK, PyQt, PyJs, etc.
- **Machine Learning -** TensorFlow, PyTorch, **Scikit-learn**, Matplotlib, Scipy, etc.
- **Mathematics -** Numpy, Pandas, Scipy, etc.

R vs. Python

- Both are good options for Analysis
- R or Python + NumPy + SciPy + SciKit +Pandas
- R is stronger for stats. Python for General Programming
- ggplot2 > Matplotlib
- Easier to get started with R
- Python is easier to get “ into production”

Tips to Keep in Mind While Learning Python

The most common question asked by the beginners - **"What is the best way to learn Python"?** It is the initial and relevant question because first step in learning any programming language is to know how to learn.

The proper way of learning will help us to learn fast and become a good Python developer.

Tips to Keep in Mind While Learning Python

The most common question asked by the beginners - **"What is the best way to learn Python"?** It is the initial and relevant question because first step in learning any programming language is to know how to learn.

The proper way of learning will help us to learn fast and become a good Python developer. Make it clear why we want to learn Python?

The goal should be clear before learning the Python. Python is easy, a vast language as well. It includes numbers of libraries, modules, in-built functions and data structures. If the goal is unclear, then it will be a boring and monotonous journey of learning Python. Without any clear goal, you perhaps won't make it done.

Tips to Keep in Mind While Learning Python

- So, first figure out the motivation behind learning, which can anything be such as knowing something new, develop projects using Python, switch to Python, etc. Below are the general areas where Python is widely used. Pick any of them.
- Data Analysis and Processing
- Artificial Intelligence
- Games
- Hardware/Sensor/Robots
- Desktop Applications

Choose any one or two areas according to your interest and start the journey towards learning Python.

Tips to Keep in Mind While Learning Python

Keep Practicing

The next important step is to do the practice. It needs to implementing the Python concepts through the code. We should be consistence to our daily coding practice.

Consistency is the key of success in any aspect of life not only in programming. Writing code daily will help to develop muscle memory.

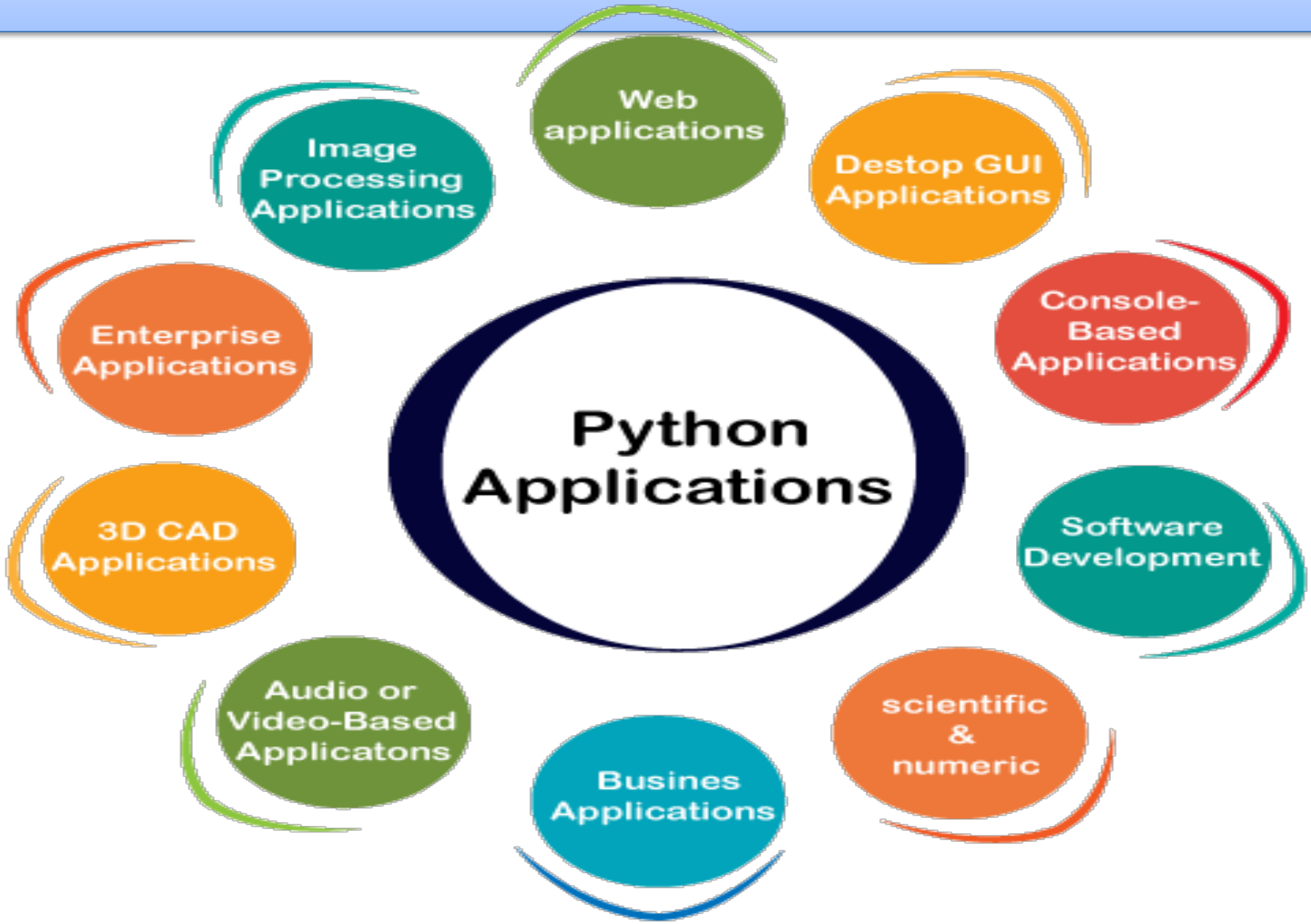
.

Where is Python Used?

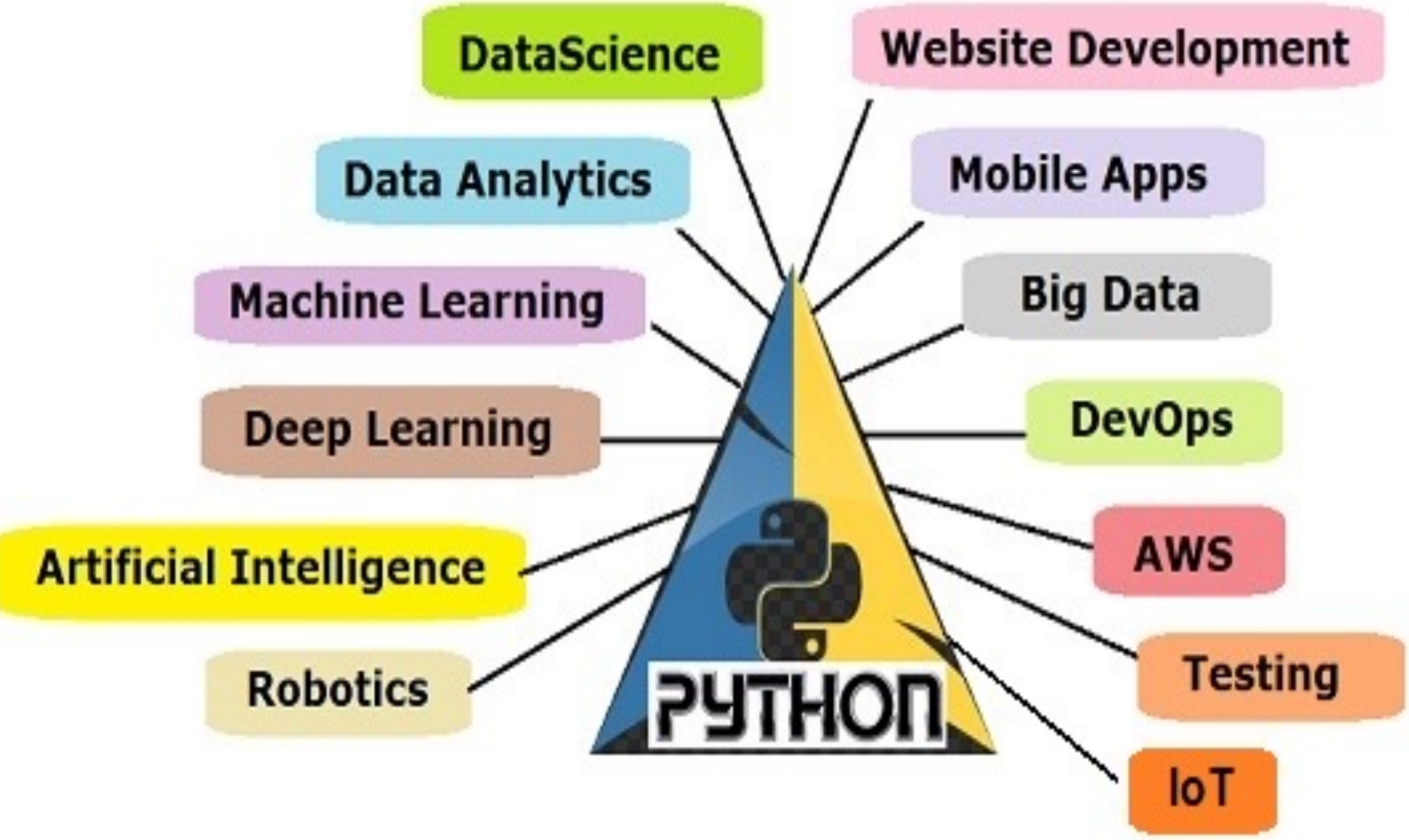
Python is a general purpose, open source, high-level programming language and also provides number of libraries and frameworks. Python has gained popularity because of its simplicity, easy syntax and user-friendly environment. The usage of Python as follows

Desktop Applications
Web Applications
Data Science
Artificial Intelligence
Machien Learning
Scientific Computing
Robotics
Internet fi Things
Gaming
Mobile Apps
Data Analysis and Preprocessing

Where is Python Used?



Where is Python Used?



How to install Python

There are two approaches to install Python?

- You can download Python directly from its project site "python.org/downloads/" and install individual components and libraries you want.

- Alternately, you can download and install a package, which comes with pre-installed libraries (Anaconda, anaconda.com/distribution). This method provides hassle free installation.

How to install Python

There are two approaches to install Python?

- You can download Python directly from its project site "python.org/downloads/" and install individual components and libraries you want.

- Alternately, you can download and install a package, which comes with pre-installed libraries (Anaconda, anaconda.com/distribution). This method provides hassle free installation.

Choosing a development environment

Once you have installed Python, there are various options for choosing an environment. Following are the most common options.

- Terminal / Shell based
- IDLE
- iPython/Jupyter notebook

Analytics philosophy

- You have to get your hands dirty
- Keep trying out things
- Download data, or some code, and try to run
- Make small tweaks
- Analysis is both a science and art.
- Understand how the analysis has been put together
- There is no way to know everything. Learning is the answer
  - You learn by observing and practicing

Starting with the end in mind

- Ask yourself these before you start the analysis
- What do I want to present?
- Which graphs will I create? and how many?
- What data will I need
- Where I can the data, i.e., source of the data
- Try a “mock plot” with dummy data
- Does it look like what I want

Resources to Learn Python

docs.python.org/3/tutorial

https://learnpython.org

The Quick Python Book
Naomi Ceder

Practical Programming: An Introduction to Computer Science Using Python
Paul Gries, Jennifer Campbell, Jason Montojo

Python for Data Analysis: Wes McKinney

Python Data Science Handbook: Jake VanderPlas