

دوره
علم داده،
یادگیری ماشین و
هوش مصنوعی

جلسه صفر:
آشنایی با دوره علم داده،
یادگیری ماشین و
هوش مصنوعی

ارائه‌کننده:
دکتر فرزاد مینویی

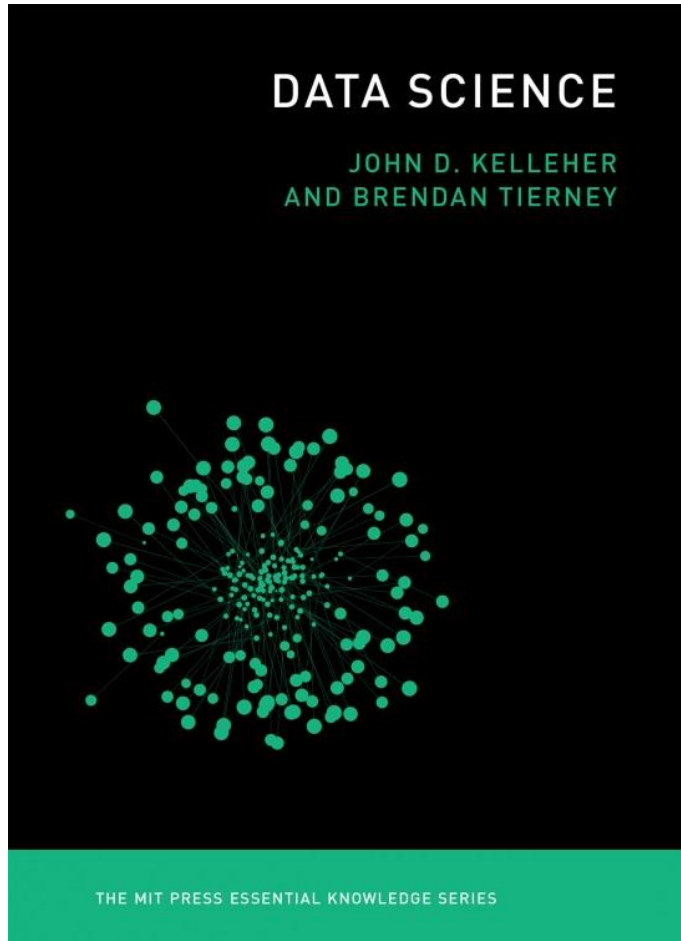


اهداف جلسه صفر

- درباره علم داده
- آشنایی با دوره
- آشنایی با پلتفرم کافه تدریس
- پرسش و پاسخ

علم داده (Data Science) چیست؟

علم داده شامل مجموعه‌ای از اصول، تعریف مسئله، الگوریتم‌ها و فرآیندهاست که هدف آن استخراج الگوهای غیرواضح و مفید از مجموعه داده‌های بزرگ است.



سه تحول عمده در حوزه فناوری



- تولید حجم انبوهی از داده



5 MB Hard Disk Drive - 1956

© Go Physics



1 TB Micro SD Card - 2020

- افزایش توان سخت‌افزاری برای ذخیره‌سازی این حجم از داده‌ها با قیمت پایین

- پیشرفت روش‌های محاسباتی و الگوریتم‌ها برای تجزیه و تحلیل داده‌های با حجم بالا

موزه‌های دانشی مورد نیاز برای ورود به علم داده

- برنامه‌نویسی و فناوری اطلاعات (IT)
- ریاضیات
- آمار
- یادگیری ماشین
- دانش تخصصی حوزه (برای مثال کسب و کار)
- نمایش داده
- مهارت‌های ارتباطی و ارائه
- استدلال کردن

Source: Write Like an Amazonian

March 5

Published: 2018

Amazon Writing Style Tip #2

Be objective: avoid adjectives and adverbs

- Adjectives are imprecise and don't contribute to making a decision
- Most Amazonians react negatively to buzz words and qualifications without data

Subjective Better → Objective

• Sales increased significantly in Q4, due to use of holiday promotions.

• We made the application much faster.

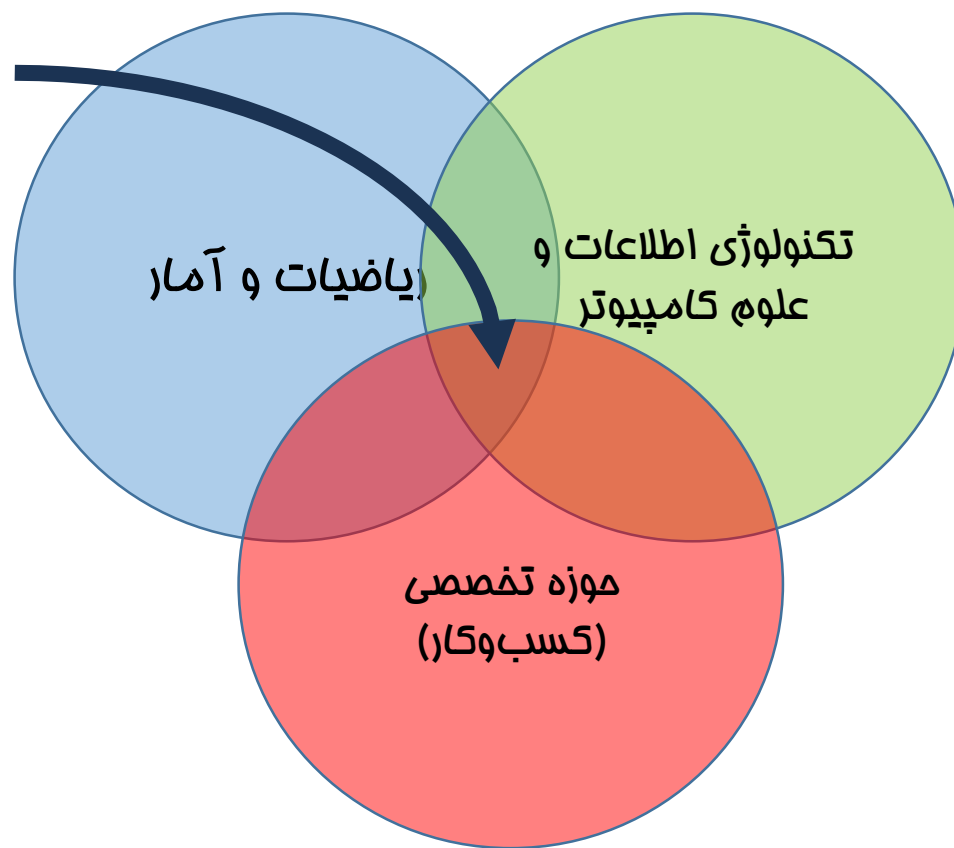
• This will make the endeavor extremely successful.

• Unit sales increased by 40% in Q4 2011, compared to Q4 2010, because of holiday promotions.

• We reduced server side trailing 90-day latency from 10ms to 1ms.

• This will increase output by 2.5%.

علم داده



علم داده از منظر دانشگاه و صنعت

• از منظر دانشگاه

- در حال حاضر در دانشگاه‌های دنیا، عنوان اصلی کسی "دانشمند داده" نیست!

• از منظر صنعت

- دانشمند داده (Data Scientist)
- مهندس داده (Data Engineer)
- تحلیل‌گر داده (Data Analyst)
- کارشناس هوش تجاری (Business Intelligence Developer)
- مهندس یادگیری ماشین (Machine Learning Engineer)
- توسعه‌دهنده هوش مصنوعی (Artificial Intelligence Developer)
- مدیر تیم داده (Data Team Lead)

پیشنهاد مطالعه

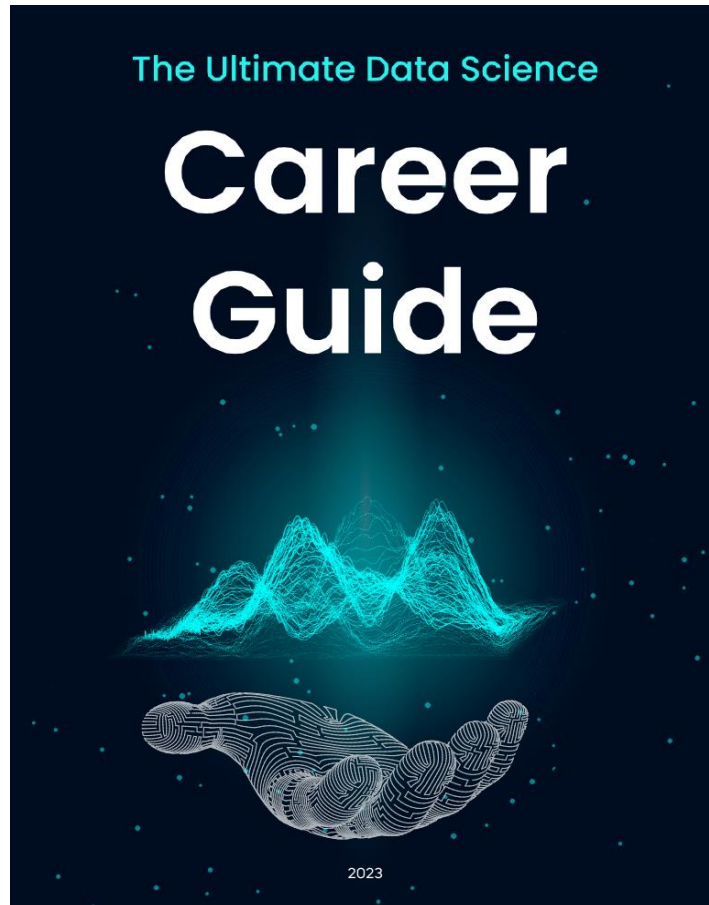


TABLE OF CONTENTS	
I	What Makes Data Science a Good Career Choice? 1. Why Is Data Science Important? 07 2. Who Can Benefit from Data Science? 08
III	What Are the Top Data Science Roles to Explore? 1. What is a Data Scientist? 13 2. What is a Data Analyst? 17 3. What is a Business Intelligence Analyst? 20 4. What is a Data Architect? 24 5. What is a Data Engineer? 27
V	Which Industry Should You Consider for Your Data Science Career? 1. Finance and Insurance 38 2. Data Science in Big Tech Companies 40 3. Healthcare 41 4. Manufacturing 42
VII	Ready to Take the Next Step? 1. Building a Data Science Career 48
IX	Conclusion 1. Not Sure Where to Start 57
II	What Is the Data Science Job Outlook? 1. What is the Future of Data Science Jobs? 10 2. What Are the Most In-Demand Data Science Skills? 10
IV	How to Get a Job in Data Science? 1. How to Write a Compelling Data Science Resume? 32 2. How to Build a Data Science Project Portfolio That Stands Out? 33 3. How to Prepare for a Data Science Interview? 35 4. How to Land Your First Data Science Job? 36
VI	Types of Organizations to Consider 1. Data Science Consulting Companies 44 2. Data Science Startups 45 3. Government Data Science Jobs 46
VIII	Resources 1. Recommended Articles 50 2. Recommended Books 51 3. Portfolio & Interview Preparation 52 4. Useful Tools & Websites 53 5. Recommended Youtube Channels 54 6. Great LinkedIn Content Creators 55

اهداف آموزشی دوره چیست؟

- یادگیری مفاهیم اصلی حوزه علم داده، یادگیری ماشین و هوش مصنوعی
- آشنایی با منطق الگوریتم‌های یادگیری ماشین و هوش مصنوعی
- نحوه اجرا و پیاده‌سازی الگوریتم‌های یادگیری ماشین در پایتون
- نحوه فکر کردن به مسائل از منظر رویکرد داده‌محور

محتوای آموزشی چگونه در اختیار شما قرار می‌گیرد؟

- کلاس‌های آموزشی آنلاین
- فایل‌های ارائه و کد
- مقالات و کتاب‌های مرجع
- کانال و گروه تلگرام برای پرسش و پاسخ

محورهای دوره آموزشی علم داده، یادگیری ماشین و هوش مصنوعی

- برنامه‌نویسی مقدماتی با پایتون
- مفاهیم پایه جبرخطی و آمار
- کتابخانه‌های مهم پایتون برای تحلیل داده
- Numpy، Pandas، Matplotlib، Scipy و Statsmodels
- مدل‌های یادگیری ماشین
 - یادگیری نظارت‌شده
 - یادگیری نظارت‌نشده
 - یادگیری تقویتی
- مدل‌های یادگیری عمیق
- تحلیل سری‌های زمانی
- مروری بر مفاهیم پایگاه داده و زبان کوئری‌نویسی SQL
- مقدمه‌ای بر توسعه محصولات داده‌محور

ویژگی‌های دوره آموزشی علم داده، یادگیری ماشین و هوش مصنوعی

- تدریس مفاهیم و تئوری‌های علم داده و یادگیری ماشین به زبان ساده
- تمرکز روی ارتقاء توانمندی حل مسئله با رویکرد داده‌محور
- بهره‌گیری از موردکاوی‌های دنیای واقعی
- تمرین‌ها و کوئیزهای متنوع و پروژه پایانی دوره
- امکان پرسش و پاسخ با استاد درس و سایر دانشجویان

مسیر پیش رو

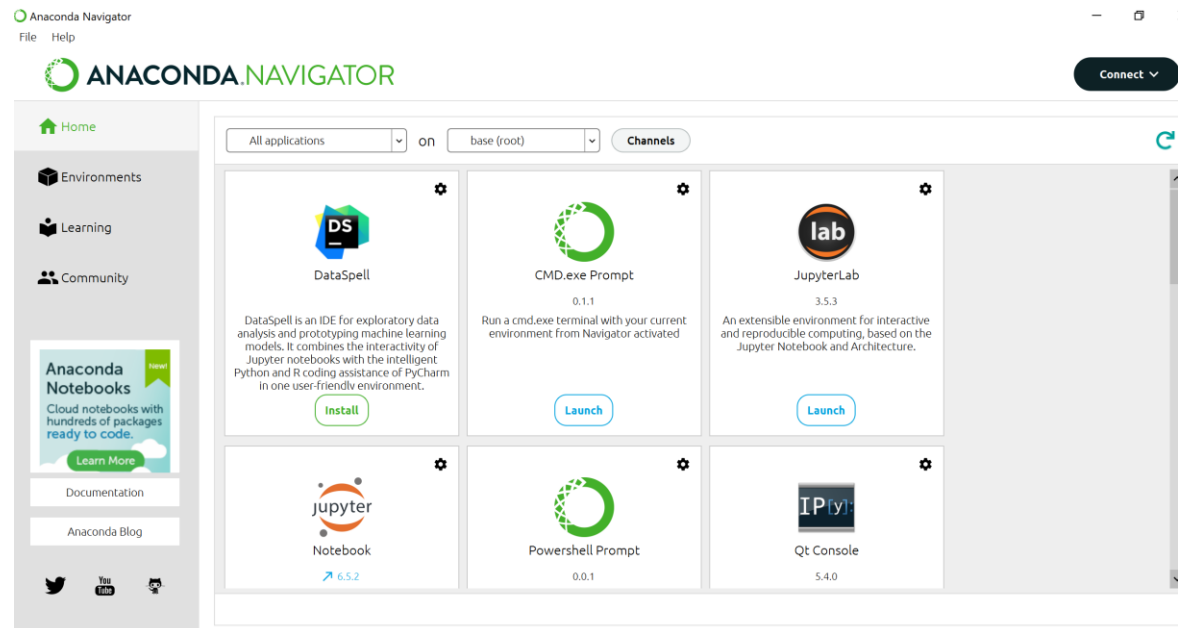
- جلسه اول: نصب آناکوندا، شروع برنامه نویسی با پایتون
- جلسه دوم: ساختارهای داده، عبارات شرطی و حلقه ها در پایتون
- جلسه سوم: ماژول ها و کتابخانه، توابع و کلاس در پایتون
- جلسه چهارم: مروری بر مفاهیم جبرخطی و ماتریس ها و محاسبات با کتابخانه Numpy در پایتون
- جلسه پنجم: آماده سازی داده با کتابخانه Pandas در پایتون
- جلسه ششم: تصویرسازی داده با کتابخانه Matplotlib در پایتون
- جلسه هفتم: تحلیل اکتشافی داده ها (Data Exploratory Analysis)
- جلسه هشتم: داده های گمشده و داده پرت
- جلسه نهم: مقدمه ای بر مفاهیم یادگیری ماشین
- جلسه دهم: الگوریتم kNN Regression و پیاده سازی در پایتون
- جلسه یازدهم: الگوریتم رگرسیون خطی
- جلسه دوازدهم: پیاده سازی الگوریتم رگرسیون خطی در پایتون
- جلسه سیزدهم: الگوریتم درخت تصمیم، جنگل تصادفی و تقویت گرادیان برای حل مسائل رگرسیون و پیاده سازی در پایتون
- جلسه چهاردهم: پیاده سازی الگوریتم های مبتنی بر درخت تصمیم در پایتون
- جلسه پانزدهم: الگوریتم رگرسیون لجستیک

مسیر پیش رو

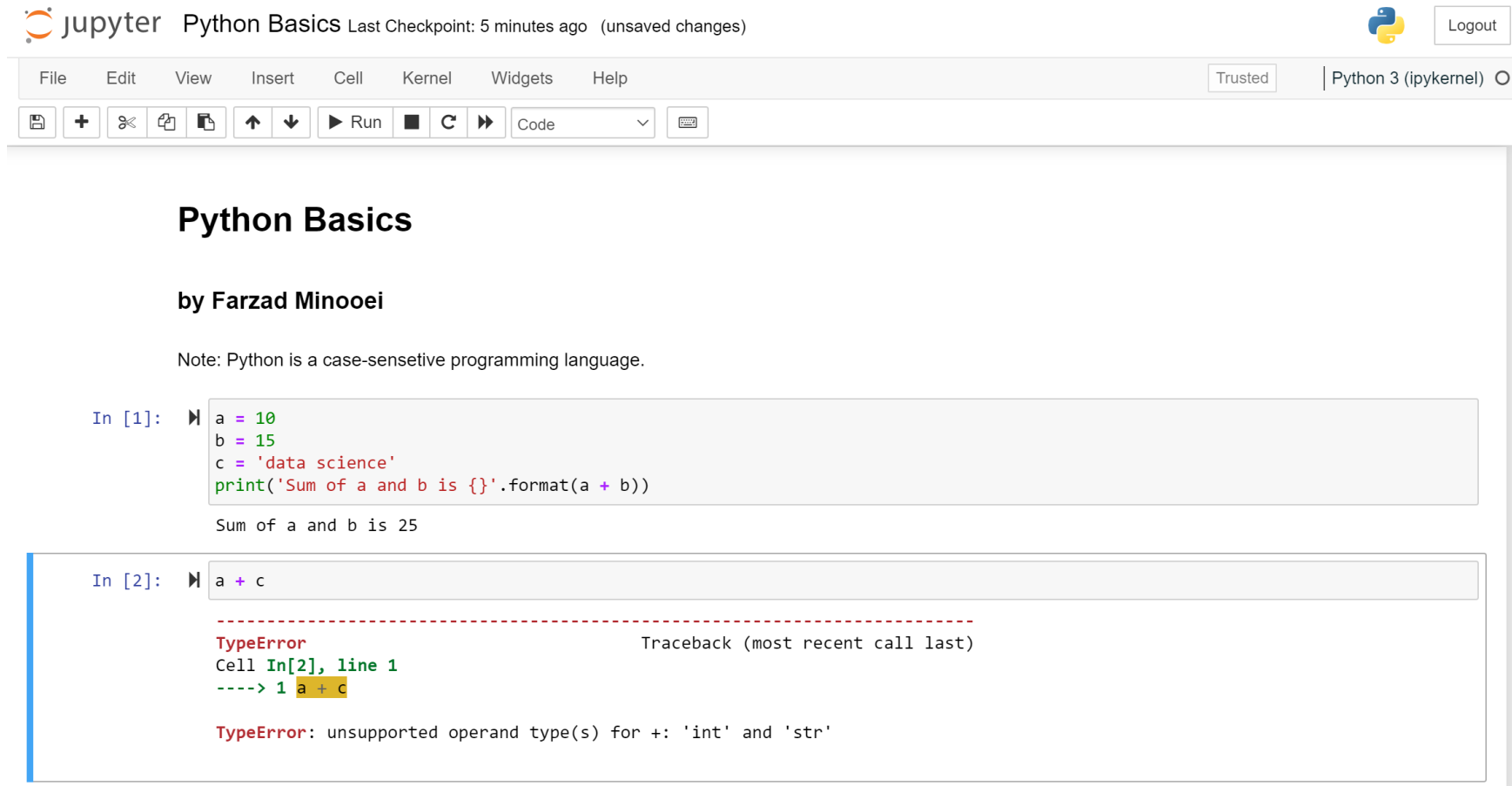
- جلسه شانزدهم: الگوریتم‌های kNN، درخت تصمیم، جنگل تصادفی و ماشین بردار پشتیبان برای حل مسائل دسته‌بندی
- جلسه هفدهم: پیاده‌سازی الگوریتم‌های دسته‌بندی در پایتون
- جلسه هجدهم: الگوریتم‌های نظارت‌نشده؛ فوشه‌بندی و کاهش بعد
- جلسه نوزدهم: پیاده‌سازی الگوریتم‌های نظارت‌نشده در پایتون
- جلسه بیستم: شبکه عصبی مصنوعی و یادگیری عمیق با کتابخانه TensorFlow در پایتون
- جلسه بیست و یکم: شبکه عصبی کانولوشنی
- جلسه بیست و دوم: موردکاوی پردازش تصویر در پایتون
- جلسه بیست و سوم: تحلیل سری‌های زمانی
- جلسه بیست و چهارم: پیاده‌سازی تحلیل سری‌های زمانی در پایتون
- جلسه بیست و پنجم: مقدمه‌ای بر تحلیل متن
- جلسه بیست و ششم: شبکه‌های عصبی بازگشتی
- جلسه بیست و هفتم: مقدمه‌ای بر یادگیری تقویتی
- جلسه بیست و هشتم: موردکاوی یادگیری تقویتی در پایتون
- جلسه بیست و نهم: آشنایی با مفاهیم پایگاه داده و انبار داده، نصب MySQL
- جلسه سی‌ام: زبان کوئری‌نویسی SQL
- جلسه سی و یکم: آشنایی با مفاهیم مهندسی یادگیری ماشین (Machine Learning Engineering)
- جلسه سی و دوم: موردکاوی توسعه محصول داده‌محور (Data Product) در پایتون

آماده‌سازی محیط توسعه برای پایتون


- Anaconda
 - <https://www.anaconda.com/>
- Jupyter Notebook
 - <https://jupyter.org/>



آماده‌سازی محیط توسعه برای پایتون



The screenshot shows a Jupyter Notebook titled "Python Basics" with a last checkpoint of 5 minutes ago. The interface includes a menu bar (File, Edit, View, Insert, Cell, Kernel, Widgets, Help) and a toolbar with icons for saving, adding cells, and running code. The notebook content includes a title "Python Basics" by Farzad Minooei, a note about Python being case-sensitive, and two code cells. The first cell runs successfully, printing the sum of a and b. The second cell attempts to add a string and an integer, resulting in a `TypeError`.

jupyter Python Basics Last Checkpoint: 5 minutes ago (unsaved changes)  Logout

File Edit View Insert Cell Kernel Widgets Help Trusted | Python 3 (ipykernel) O

Python Basics

by Farzad Minooei

Note: Python is a case-sensitive programming language.

In [1]: `a = 10
b = 15
c = 'data science'
print('Sum of a and b is {}'.format(a + b))`

Sum of a and b is 25

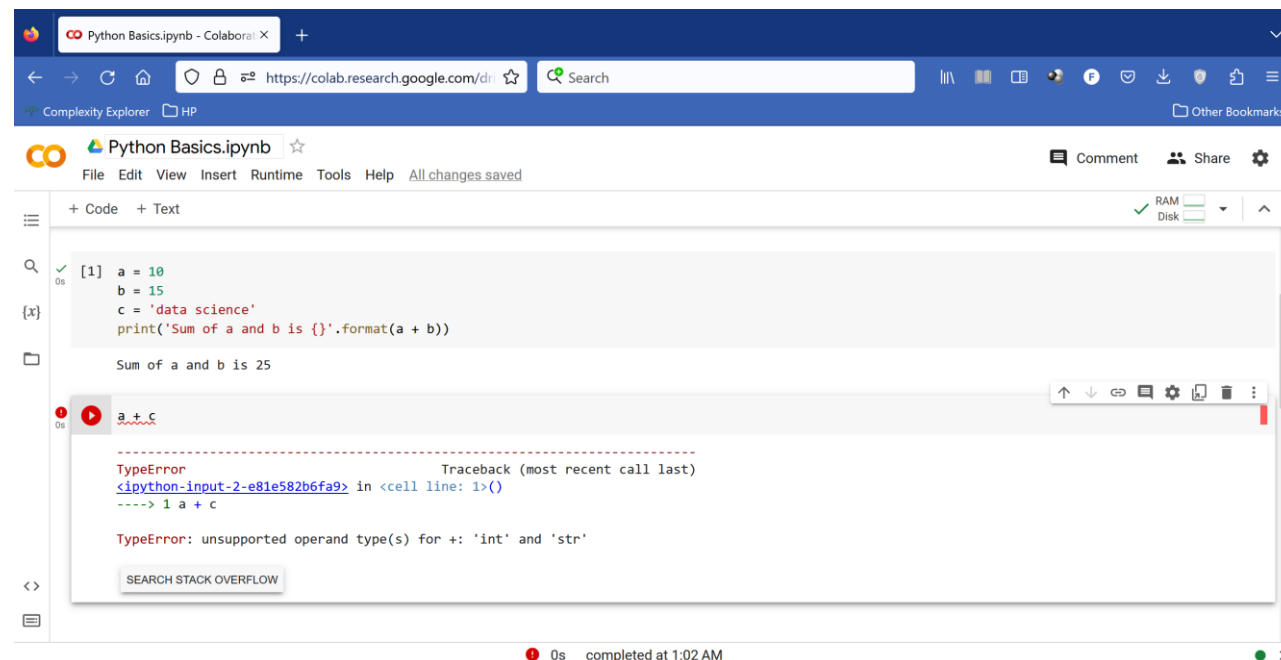
In [2]: `a + c`

Traceback (most recent call last)
TypeError
Cell In[2], line 1
----> 1 a + c

TypeError: unsupported operand type(s) for +: 'int' and 'str'

آماده‌سازی محیط توسعه برای پایتون

- Google Colab
 - <https://colab.research.google.com/>



The screenshot shows a Google Colab notebook titled "Python Basics.ipynb". The code cell contains the following Python code:

```
[1] a = 10
    b = 15
    c = 'data science'
    print('Sum of a and b is {}'.format(a + b))
```

The output of the code cell is "Sum of a and b is 25". Below the code cell, there is a runtime error message:

```
Traceback (most recent call last)
<ipython-input-2-e81e582b6fa9> in <cell line: 1>()
----> 1 a + c

TypeError: unsupported operand type(s) for +: 'int' and 'str'
```

The error message indicates a `TypeError` because the code attempts to add an integer (`a`) and a string (`c`).

پیشنهاد مطالعه



W3schools

Tutorials ▾ Exercises ▾ Certificates ▾ Services ▾ Search... Q ⓘ </> Spaces 🛒 Get Certified Sign Up Log in

HTML CSS JAVASCRIPT SQL PYTHON JAVA PHP HOW TO W3.CSS C C++ C# BOOTSTRAP REACT MYSQL JQUERY EX >

Python Tutorial

- Python HOME
- Python Intro
- Python Get Started
- Python Syntax
- Python Comments
- Python Variables
- Python Data Types
- Python Numbers
- Python Casting
- Python Strings
- Python Booleans
- Python Operators
- Python Lists
- Python Tuples
- Python Sets
- Python Dictionaries
- Python If...Else

Python Introduction

< Previous

Next >

What is Python?

Python is a popular programming language. It was created by Guido van Rossum, and released in 1991.

It is used for:

- web development (server-side),
- software development

ADVERTISEMENT

Build your career. Get Full Access.

Lifelong access to all W3Schools courses and certifications!

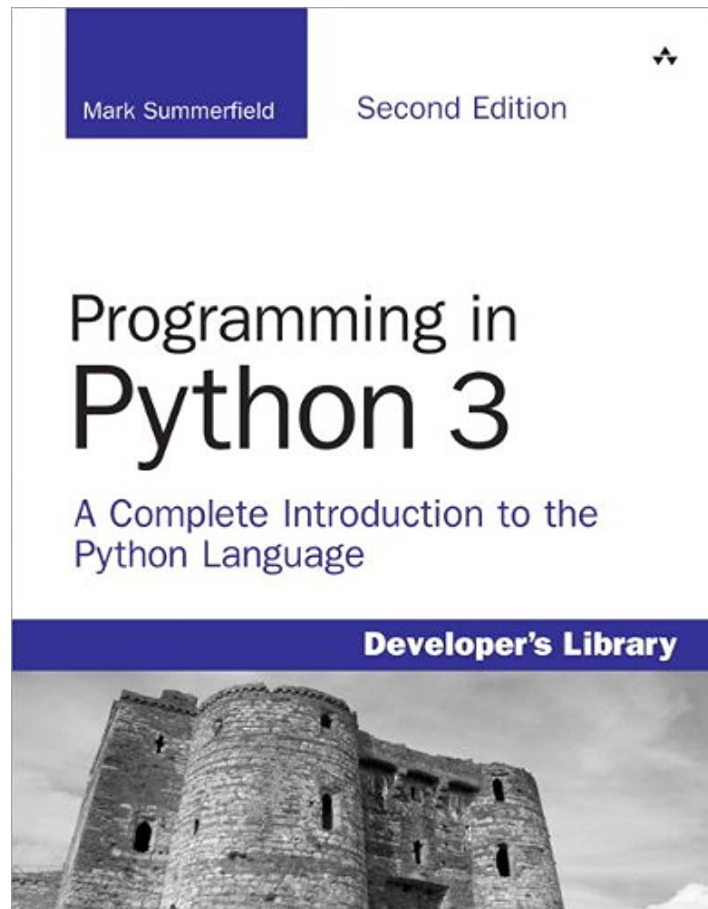
SAVE 770\$

Start today

W3schools

https://www.w3schools.com/python/python_intro.asp

پیشنهاد مطالعه



<https://cs.smu.ca/~porter/csc/227/ProgrammingInPython3.pdf>

چند توصیه

- تمرین، تمرین و تمرین
- مطالعه
- سعی و خطا کردن
- حفظ مداومت

پرسش و پاسخ

