# بسم الله الرحمن الرحیم

# 

**استاد محمد احمد زاده**

**درس مباحث ویژه**

**هم گروه عالیه خرمی سرریگانی و فرشته رنجبری**

: Python Programming

۱. چرا Python زبان برنامه‌نویسی محبوب علم داده است؟

پایتون به دلیل سادگی در خوانایی، یادگیری آسان و مجموعه گسترده‌ای از کتابخانه‌ها، به یکی از محبوب‌ترین زبان‌های برنامه‌نویسی برای علم داده تبدیل شده است. ویژگی‌های کلیدی آن عبارتند از:

کتابخانه‌های غنی: شامل NumPy، Pandas، Matplotlib، Scikit-learn و TensorFlow که به تحلیل داده و یادگیری ماشین کمک می‌کنند.

انعطاف‌پذیری بالا: امکان استفاده در برنامه‌های تحت وب، یادگیری ماشین، پردازش تصویر، و اتوماسیون را فراهم می‌کند.

محیط‌های تعاملی: مانند Jupyter Notebook که به دانشمندان داده کمک می‌کند تا به‌صورت تعاملی کد را اجرا کرده و تحلیل کنند.

جامعه کاربری گسترده: به دلیل محبوبیت زیاد، منابع یادگیری و پشتیبانی قوی برای این زبان وجود دارد.

۲. Pandas و NumPy چه تفاوتی دارند؟

هر دو کتابخانه برای پردازش داده در پایتون استفاده می‌شوند، اما تفاوت‌های مهمی دارند:

NumPy برای محاسبات عددی و کار با آرایه‌های چندبعدی بهینه شده است. پردازش برداری سریع و عملکرد بالایی دارد.

Pandas برای تحلیل داده‌های جدولی (DataFrame) به کار می‌رود و شامل ابزارهای پیشرفته برای فیلتر کردن، دسته‌بندی و تجزیه‌وتحلیل داده‌ها است.

به‌طور کلی، NumPy بر عملیات عددی در آرایه‌ها تمرکز دارد، در حالی که Pandas برای پردازش داده‌های ساختاریافته و سری‌های زمانی مناسب‌تر است.

۳. چرا Matplotlib برای تجسم داده‌ها استفاده می‌شود؟

Matplotlib یک کتابخانه قدرتمند برای رسم نمودارهای متنوع است که در علم داده و یادگیری ماشین کاربرد دارد. ویژگی‌های مهم آن عبارتند از:

امکان ایجاد نمودارهای خطی، ستونی، دایره‌ای و پراکندگی.

قابلیت سفارشی‌سازی بالا برای نمایش بهتر داده‌ها.

ادغام با Jupyter Notebook برای نمایش تعاملی نمودارها.

توانایی ذخیره نمودارها در فرمت‌های مختلف مانند PNG، PDF و SVG.

۴. چرا Seaborn برای تجسم داده‌های پیشرفته کاربرد دارد؟

Seaborn بر پایه Matplotlib ساخته شده و قابلیت‌های بهتری برای نمایش داده‌ها دارد، از جمله:

ایجاد نمودارهای آماری پیچیده با کد کمتر.

پشتیبانی از گراف‌های توزیع داده مانند KDE Plot و Box Plot.

قابلیت تعامل بهتر با DataFrameهای Pandas.

سبک‌های بصری جذاب‌تر نسبت به Matplotlib.

به دلیل این ویژگی‌ها، Seaborn در تحلیل داده‌های آماری و کشف الگوهای پنهان در داده‌ها بسیار مفید است.

۵. چگونه یک Function در Python تعریف کنید؟

در پایتون، توابع با استفاده از کلمه کلیدی def تعریف می‌شوند. مثلاً:

def greet(name):

return f"Hello, {name}!"

print(greet("Ali"))

ویژگی‌های کلیدی توابع در پایتون:

می‌توانند مقدار بازگشتی داشته باشند (return).

می‌توانند چندین ورودی بپذیرند.

امکان تعریف توابع بازگشتی (Recursive Functions) وجود دارد.

می‌توانند توابع ناشناس (lambda) داشته باشند.

۶. چرا از List Comprehension در Python استفاده می‌شود؟

List Comprehension یک روش کارآمد برای ایجاد لیست‌ها با استفاده از سینتکس کوتاه‌تر است. مثال:

numbers = [x\*\*2 for x in range(10)]

print(numbers) # [0, 1, 4, 9, 16, 25, 36, 49, 64, 81]

مزایای آن:

افزایش خوانایی کد: کد مختصرتر و تمیزتر است.

سرعت بیشتر: در مقایسه با حلقه‌های for عملکرد بهتری دارد.

امکان فیلتر کردن داده‌ها: می‌توان از شرط‌ها در آن استفاده کرد.

۷. چگونه یک فایل CSV را در Python بخوانید؟

پایتون به کمک کتابخانه Pandas می‌تواند فایل‌های CSV را به راحتی پردازش کند:

import pandas as pd

df = pd.read\_csv("data.csv")

print(df.head()) # نمایش پنج سطر اول فایل CSV

دلایل استفاده از Pandas برای خواندن CSV:

مدیریت بهتر داده‌ها در قالب DataFrame.

پشتیبانی از انواع داده‌ها و تبدیل خودکار آن‌ها.

امکان فیلتر کردن و تغییر داده‌ها به‌صورت ساده.

۸. XML و JSON چه تفاوتی دارند؟

هر دو برای ذخیره و انتقال داده استفاده می‌شوند، اما تفاوت‌های کلیدی دارند:

| ویژگی | JSON | XML | |--------|------|-----| | ساختار | سبک‌تر و خواناتر | پیچیده‌تر | | سرعت پردازش | سریع‌تر | کندتر | | استفاده در وب | پرکاربرد در APIها | کمتر در APIها استفاده می‌شود | | پشتیبانی از داده‌های تو در تو | بله | بله | | خوانایی برای انسان | آسان‌تر | سخت‌تر |

JSON به دلیل سادگی و ادغام بهتر با زبان‌های برنامه‌نویسی مدرن، معمولاً برای تبادل داده‌ها در APIها و برنامه‌های تحت وب استفاده می‌شود.

.