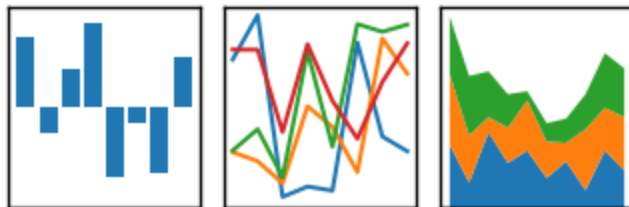


pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



১. প্যান্ডাস পরিচিতি (Introduction to Pandas)

প্যান্ডাস (Pandas) হল পাইথনের একটি শক্তিশালী লাইব্রেরি যা **ডাটা ম্যানিপুলেশন** এবং **বিশ্লেষণের** জন্য ব্যবহৃত হয়।

Pandas is the backbone of most data projects.

Pandas is derived from the term **"panel data"**, an econometrics term for data sets.

If you're thinking about data science as a career, then it is imperative that one of the first things you do is **learn pandas**.

ডাটা এনালাইসিসের জন্য একটা জনপ্রিয় লাইব্রেরি হচ্ছে প্যান্ডা / Pandas।

- এই চমৎকার লাইব্রেরি দিয়ে রিমোট ফাইল থেকে ডেটা লোড, লোকাল ফাইল থেকে বা ডাটাবেজ থেকে ডেটা লোড করা যায়।
- খুব সহজে ডেটার সামারি, স্ট্যাটিসটিক্স বা ডেটা সম্পর্কে আইডিয়া নেওয়া যায়।
- ডেটা ক্লিন করা, ডেটা ম্যানিপুলেট, ডেটা কন্সট্রাইন সহ দরকারি সব কাজই প্যান্ডার সাহায্যে করা যায়।
- বড় ডেটা সেট নিয়েও খুব সহজে কাজ করা যায়।
- এছাড়া ডেটা নিয়ে সাধারণ যে সব প্রশ্ন আসে আমাদের মাথায়, সে সব প্রশ্নের উত্তরও আমরা প্যান্ডার সাহায্যে বের করতে পারি।

২. প্যান্ডাস ডাটা স্ট্রাকচার (Pandas Data Structures)

প্যান্ডাসের দুটি প্রধান ডাটা স্ট্রাকচার আছে বা পাণ্ডাতে দুইটা প্রধান অবজেক্ট রয়েছে, DataFrame এবং Series।

- ◆ **Series:** এক-ডাইমেনশনাল ডাটা স্ট্রাকচার, যা একটি এক্সেল বা ডাটাবেজের কলামের মতো কাজ করে।
- ◆ **DataFrame:** দুই-ডাইমেনশনাল টেবিল, যা একাধিক সিরিজ সম্বলিত থাকে।

Series

	apples
0	3
1	2
2	0
3	1

+

Series

	oranges
0	0
1	3
2	7
3	2

=

DataFrame

	apples	oranges
0	3	0
1	2	3
2	0	7
3	1	2

A Series হল মূলত একটি কলাম, এবং a DataFrame হল একটি বহুমাত্রিক সারণী যা সিরিজের একটি সংগ্রহ নিয়ে গঠিত।

সংক্ষেপে:

- ◆ Series = এক কলামের ডাটা
- ◆ DataFrame = একাধিক কলামের টেবিল
- ◆ সহজে ডাটা ফিল্টার, বিশ্লেষণ এবং মডিফাই করা যায়