

PANDA KUTUBXONASI VA UNING QO‘LLANILISHI

Tojimamatov Israil Nurmamatovich

Farg‘ona davlat universiteti Amaliy matematika

va informatika kafedrası katta o‘qituvchisi

E-mail: israiltojimatov@gmail.com

Mamatxonova Gulasalxon Saidjon qizi

Farg‘ona davlat universiteti Amaliy matematika

yo‘nalishi 3-bosqich 23.07-guruh talabasi

E-mail: tulanovagulixon@gmail.com

Annotatsiya: Ushbu maqolada Panda kutubxonasi va uning dasturiy taqdimotlarda qo‘llanilishi haqida ilmiy tahlil berilgan Kutubxonaning maqsadi va imkoniyatlari, ma'lumotlar bilan ishlashdagi samaradorligi, real loyihalarda qo‘llanilish usullari yoritilgan Shuningdek kutubxona yordamida ma'lumotlarni tozalash, transformatsiya qilish va vizualizatsiya qilish jarayonlari batafsil ko‘rib chiqilgan Maqola ilmiy tadqiqotlar va amaliy misollar asosida tuzilgan

Kalit so‘zlar: Panda kutubxonasi ma'lumotlar bilan ishlash Python ma'lumotlar tahlili DataFrame tozalash vizualizatsiya

Аннотация: В данной статье рассматривается библиотека Panda и ее применение в программных проектах Статья включает анализ возможностей библиотеки эффективность работы с данными и методы использования в реальных проектах Кроме того подробно обсуждаются процессы очистки данных трансформации и визуализации на основе научных исследований и практических примеров

Ключевые слова: библиотека Panda работа с данными Python анализ данных DataFrame очистка визуализация

Abstract: This article explores the Panda library and its applications in software projects. The study presents an analysis of the library features, efficiency in data handling and practical usage in real projects. Moreover, processes of data cleaning, transformation and visualization are discussed in detail. The article is based on scientific research and practical examples.

Keywords: Panda library, data handling, Python data analysis, DataFrame, cleaning, visualization.

Kirish

Ma'lumotlar zamonaviy ilm-fan va texnologiya sohalarida muhim resurs hisoblanadi. Ma'lumotlarni yig'ish, tahlil qilish, tozalash va vizualizatsiya qilish jarayonlari samarali ishlash uchun kuchli vositalarni talab qiladi. Shu jihatdan Python dasturlash tili va uning kutubxonalari ma'lumotlar bilan ishlashni sezilarli darajada soddalashtiradi.

Panda kutubxonasi aynan shu ehtiyojni qondirishga mo'ljallangan va keng ko'lamli imkoniyatlar taqdim etadi. Kutubxona asosan ma'lumotlarni saqlash, qayta ishlash va tahlil qilish uchun mo'ljallangan DataFrame va Series kabi ma'lumot tuzilmalari bilan ishlash imkonini beradi. Shu bilan birga, u ma'lumotlarni tozalash, qayta shakllantirish, filtr va guruhlash funksiyalarini osonlik bilan bajarishga yordam beradi.

Kirish qismida maqolaning maqsadi ham belgilab olinadi: Panda kutubxonasining nazariy asoslarini, uning dasturiy imkoniyatlarini va real loyihalarda qo'llanilishini ilmiy tarzda tahlil qilish. Shuningdek, kutubxonadan foydalanish orqali ma'lumotlar bilan ishlash jarayonlarining samaradorligini oshirish usullari ko'rib chiqiladi. Maqola o'quvchilarga Python dasturlash tilida ma'lumotlar tahlili bo'yicha amaliy ko'nikmalarni shakllantirishda ham yordam beradi.

Panda kutubxonasi va uning asosiy imkoniyatlari

Panda Python dasturlash tilida ma'lumotlar bilan ishlash uchun yaratilgan kuchli kutubxona bo'lib, uning asosiy maqsadi ma'lumotlarni tezkor va samarali qayta ishlashni ta'minlashdir. Kutubxona DataFrame va Series kabi asosiy ma'lumot tuzilmalari bilan ishlash imkonini beradi. DataFrame ikki o'lchovli jadval ko'rinishidagi ma'lumotlar bilan ishlash uchun mo'ljallangan bo'lsa, Series bir o'lchovli ma'lumotlar ketma-ketligini

ifodalaydi Bu tuzilmalar foydalanuvchiga ma'lumotlarni saralash, guruhlash, filtrlash va statistik tahlil qilish imkonini beradi

Ma'lumotlarni import qilish va eksport qilish Panda kutubxonasining eng qulay xususiyatlaridan biridir Kutubxona CSV, Excel, SQL va boshqa formatlardagi fayllardan ma'lumotlarni osonlik bilan o'qish va saqlash imkonini beradi Bu esa katta hajmdagi ma'lumotlar bilan ishlashda va real loyihalarda samaradorlikni oshiradi

Ma'lumotlarni tozalash va transformatsiya qilish jarayoni ham Panda kutubxonasining asosiy imkoniyatlaridan hisoblanadi Ma'lumotlar bazasida mavjud bo'lishi mumkin bo'lgan bo'sh qiymatlar, takroriy yozuvlar va noto'g'ri formatlar kutubxona yordamida tezkorlik bilan aniqlanadi va tuzatiladi Bundan tashqari, ustunlar qo'shish, nomlarini o'zgartirish, ma'lumotlarni normalizatsiya qilish kabi operatsiyalarni ham samarali bajarish mumkin

Panda kutubxonasi shuningdek ma'lumotlarni guruhlash va agregatsiya qilish imkoniyatlarini ham taqdim etadi Bu funksiyalar yordamida foydalanuvchi ma'lumotlarni ma'lum mezonlarga ko'ra guruhlab, ularning statistik ko'rsatkichlarini hisoblay oladi Masalan, o'rtacha qiymat, jami summasi, maksimal va minimal qiymatlarni aniqlash kabi amallarni osonlik bilan bajarish mumkin

Vizualizatsiya va tahlil bo'yicha Panda kutubxonasi boshqa kutubxonalar bilan birgalikda ishlash imkoniyatiga ega Matplotlib va Seaborn kabi kutubxonalar bilan integratsiyasi foydalanuvchiga ma'lumotlarni grafik ko'rinishda taqdim etish, trendlarni aniqlash va ma'lumotlarni tushunarli tarzda vizualizatsiya qilish imkonini beradi Bu esa ilmiy tadqiqotlarda va biznes loyihalarida qaror qabul qilish jarayonini sezilarli darajada yaxshilaydi

Real loyihalarda Panda kutubxonasi qo'llanilishi keng ko'lamga ega Masalan, moliya sohasida narxlar o'zgarishini tahlil qilish, savdo va marketing bo'yicha ma'lumotlarni segmentlash, sog'liqni saqlash sohasida bemorlar ma'lumotlarini tahlil qilish, ilmiy tadqiqotlarda eksperiment natijalarini qayta ishlash kabi amaliyotlarda Panda kutubxonasi muhim vosita sifatida ishlatiladi

Shu bilan birga, kutubxona ma'lumotlar bilan ishlash jarayonlarini tezlashtirish va dasturchi ishini yengillashtirishga xizmat qiladi Katta hajmdagi ma'lumotlar bilan ishlashda Python dasturlash tilining imkoniyatlarini kengaytiradi va foydalanuvchi uchun intuitiv va qulay interfeys yaratadi

Panda kutubxonasi o'zining qulayligi, samaradorligi va keng imkoniyatlari tufayli ma'lumotlar tahlili, ilmiy tadqiqotlar va biznes loyihalarida asosiy vositalardan biriga aylangan Kutubxonaning har bir funksiyasi foydalanuvchi ehtiyojlariga moslashgan va real dunyo muammolarini hal qilishda muhim rol o'ynaydi

Panda kutubxonasi Python dasturlash tilining eng ommabop va kuchli vositalaridan biri bo'lib, uning asosiy maqsadi ma'lumotlar bilan ishlashni soddalashtirish va samaradorligini oshirishdir Kutubxona ma'lumotlar tahlili, tozalash, transformatsiya qilish va vizualizatsiya qilish jarayonlarida keng qo'llaniladi DataFrame va Series tuzilmalari foydalanuvchiga ma'lumotlarni osonlik bilan boshqarish, saralash, filtrlash va statistik tahlil qilish imkonini beradi DataFrame ikki o'lchovli ma'lumotlar jadvallarini ifodalaydi va har bir ustun turli turdagi ma'lumotlarni o'z ichiga olishi mumkin Series esa bir o'lchovli ketma-ketlik ko'rinishida ma'lumotlarni saqlaydi

Ma'lumotlar bilan ishlash jarayonida Panda kutubxonasi turli formatdagi ma'lumotlarni import qilish va eksport qilishni osonlashtiradi CSV, Excel, SQL va boshqa formatlardagi fayllardan ma'lumotlarni tezkorlik bilan o'qish va qayta yozish imkonini beradi Bu xususiyat loyihalarda katta hajmdagi ma'lumotlar bilan samarali ishlashni ta'minlaydi Shu bilan birga kutubxona ma'lumotlarni tozalash imkoniyatlarini ham taqdim etadi Bo'sh qiymatlar, takroriy yozuvlar, noto'g'ri formatlangan ma'lumotlarni aniqlash va tuzatish operatsiyalari foydalanuvchiga qulay tarzda bajariladi Bundan tashqari, ustunlar nomlarini o'zgartirish, yangi ustunlar qo'shish va ma'lumotlarni normalizatsiya qilish kabi operatsiyalar ham kutubxona yordamida tezkorlik bilan amalga oshiriladi

Panda kutubxonasi ma'lumotlarni guruhlash va agregatsiya qilish imkoniyatlarini ham taqdim etadi Bu funksiyalar ma'lumotlarni ma'lum mezonlarga ko'ra guruhlab, statistik ko'rsatkichlarni hisoblash imkonini beradi Foydalanuvchi o'rtacha qiymat, jami

summasi, maksimal va minimal qiymatlarni aniqlash kabi amallarni samarali tarzda bajarishi mumkin Ma'lumotlarni guruhlash va agregatsiya qilish tahlil jarayonini sezilarli darajada tezlashtiradi

Vizualizatsiya va tahlil jarayonida Panda kutubxonasi boshqa kutubxonalar bilan integratsiyada ishlaydi Matplotlib va Seaborn bilan birgalikda foydalanuvchiga ma'lumotlarni grafik ko'rinishda taqdim etish, trendlarni aniqlash va ma'lumotlarni tushunarli tarzda vizualizatsiya qilish imkonini beradi Bu esa ilmiy tadqiqotlar, biznes loyihalari va ma'lumotlarga asoslangan qaror qabul qilish jarayonlarini sezilarli darajada yaxshilaydi

Real loyihalarda Panda kutubxonasi turli sohalarda keng qo'llaniladi Moliyaviy tahlilda narxlar o'zgarishini kuzatish va prognozlash, savdo va marketing bo'yicha ma'lumotlarni segmentlash, sog'liqni saqlash sohasida bemorlar ma'lumotlarini tahlil qilish, ilmiy tadqiqotlarda eksperiment natijalarini qayta ishlash kabi amaliyotlarda kutubxona asosiy vosita sifatida ishlatiladi Shu bilan birga kutubxona foydalanuvchiga intuitiv interfeys yaratadi va katta hajmdagi ma'lumotlar bilan ishlash jarayonini sezilarli darajada tezlashtiradi

Panda kutubxonasi o'zining qulayligi, keng imkoniyatlari va samaradorligi tufayli ma'lumotlar tahlili sohasida standart vositalardan biriga aylangan Kutubxonaning har bir funksiyasi foydalanuvchining ehtiyojlariga moslashgan va real dunyo muammolarini hal qilishda muhim rol o'ynaydi Shu bilan birga, kutubxona dasturchilar va tadqiqotchilar uchun ilmiy va amaliy loyihalarda samarali vosita bo'lib xizmat qiladi

Panda kutubxonasi foydalanuvchiga ma'lumotlarni tezkorlik bilan o'qish va qayta ishlash imkonini beradi Masalan, CSV fayldan ma'lumotlarni import qilish jarayoni quyidagicha amalga oshiriladi

```
import pandas as pd
```

```
data = pd.read_csv('malumotlar.csv')
```

Ushbu kod orqali foydalanuvchi CSV faylni DataFrame ko'rinishida o'qiy oladi va keyingi tahlil jarayonlari uchun tayyorlaydi DataFrame tuzilmasi ma'lumotlarni ustunlar va qatorlar ko'rinishida saqlaydi va har bir ustun turli ma'lumot tiplarini qabul qilishi mumkin

Ma'lumotlarni tozalash jarayoni ham Panda kutubxonasi yordamida osonlik bilan bajariladi Masalan, bo'sh qiymatlarni aniqlash va ularni o'zgartirish quyidagi tarzda amalga oshiriladi

```
data.fillna(0, inplace=True)
```

Bu kod barcha bo'sh qiymatlarni nol bilan to'ldiradi Shu bilan birga takroriy yozuvlarni aniqlash va o'chirish ham muhim qadam hisoblanadi

```
data.drop_duplicates(inplace=True)
```

Guruhlash va agregatsiya funksiyalari ham tahlil jarayonini sezilarli darajada tezlashtiradi Masalan, foydalanuvchi ma'lumotlarni kategoriya bo'yicha guruhlab, ularning o'rtacha qiymatini hisoblashni xohlasa

```
grouped = data.groupby('kategoriya')['qiymat'].mean()
```

Bu misol orqali ma'lumotlarni ma'lum mezon bo'yicha guruhlab, statistik tahlil qilish mumkin Shuningdek Panda kutubxonasi vizualizatsiya imkoniyatlarini ham taklif qiladi Matplotlib bilan integratsiya qilish orqali foydalanuvchi DataFrame ma'lumotlarini grafik ko'rinishda taqdim etishi mumkin

```
import matplotlib.pyplot as plt
```

```
data['qiymat'].plot(kind='bar')
```

```
plt.show()
```

Ushbu jarayonlar ilmiy tadqiqotlar va amaliy loyihalarda muhim ahamiyatga ega Chunki ma'lumotlarni vizual ko'rinishda taqdim etish orqali trendlar va asosiy xususiyatlarni aniqlash osonlashadi

Real loyihalarda Panda kutubxonasi turli sohalarda keng qo'llaniladi Moliyaviy tahlil, marketing ma'lumotlarini segmentlash, sog'liqni saqlash sohasida bemorlar

ma'lumotlarini tahlil qilish va ilmiy tadqiqotlarda eksperiment natijalarini qayta ishlash kabi jarayonlarda kutubxona asosiy vosita sifatida ishlatiladi. Shu bilan birga, kutubxona foydalanuvchiga qulay va intuitiv interfeys taqdim etadi, ma'lumotlar bilan ishlash jarayonini sezilarli darajada tezlashtiradi.

Shu tarzda Panda kutubxonasi ilmiy va amaliy loyihalarda asosiy vosita sifatida foydalanuvchilarni qo'llab-quvvatlaydi. Har bir funksiyasi foydalanuvchi ehtiyojlariga moslashgan va real dunyo muammolarini hal qilishda muhim rol o'ynaydi.

Panda kutubxonasi foydalanuvchiga ma'lumotlar ustida turli filtrlar va indekslash operatsiyalarini bajarish imkonini beradi. Bu jarayon ma'lumotlar tahlilining muhim qismi bo'lib, foydalanuvchiga ma'lum shartga javob beruvchi yozuvlarni tezkorlik bilan tanlab olish imkonini beradi. Misol uchun, DataFrame tuzilmasida ma'lum ustun qiymatiga mos yozuvlarni filtrlash quyidagicha amalga oshiriladi:

```
filtered_data = data[data['yosh'] > 30]
```

Ushbu kod yosh ustunidagi qiymati 30 dan katta bo'lgan yozuvlarni ajratadi. Shu bilan birga, foydalanuvchi bir nechta shartlarni birlashtirib murakkab filtrlarni yaratishi mumkin:

```
filtered_data = data[(data['yosh'] > 30) & (data['kategoriya'] == 'A')]
```

Bu filtr yordamida foydalanuvchi yosh ustuni 30 dan katta bo'lgan va kategoriya ustuni "A" bo'lgan yozuvlarni tanlab oladi. Pandada indekslash ham muhim operatsiya bo'lib, foydalanuvchi DataFrame ustunlari yoki qatorlarini indekslash orqali ma'lumotlarga tezkor kirish imkoniyatiga ega.

```
data.set_index('id', inplace=True)
row = data.loc[1001]
```

Yuqoridagi kod orqali foydalanuvchi 1001 raqamli indeksga mos yozuvga tezkorlik bilan murojaat qiladi. Indekslish va filtrlar birgalikda ishlatilganda katta hajmdagi ma'lumotlarni tez va samarali tahlil qilish imkonini beradi. Shu bilan birga, bu jarayon

ilmiy tadqiqotlarda va real loyihalarda qaror qabul qilish uchun asosiy vosita sifatida ishlatiladi

Katta hajmdagi loyihalarda turli manbalardan olingan ma'lumotlarni birlashtirish muhim jarayon hisoblanadi. Panda kutubxonasi foydalanuvchiga bir nechta DataFrame obyektlarini birlashtirish va ularni yagona tuzilma shakliga keltirish imkonini beradi. Bu esa ma'lumotlar tahlilini soddalashtiradi va loyihalarning murakkabligini kamaytiradi.

```
merged_data = pd.merge(data1, data2, on='id', how='inner')
```

Yuqoridagi kod ikki DataFrame obyektini "id" ustuni bo'yicha birlashtiradi va faqat ikkala DataFrame'da mavjud yozuvlarni oladi. Shu bilan birga, foydalanuvchi turli usullarda birlashtirishni amalga oshirishi mumkin. Masalan, chap birlashtirish, o'ng birlashtirish yoki to'liq birlashtirish.

```
merged_data = pd.merge(data1, data2, on='id', how='left')
```

Ma'lumotlarni birlashtirish orqali foydalanuvchi bir nechta manbalardan olingan yozuvlarni yagona tahlil qilish imkoniga ega bo'ladi. Bu esa ilmiy tadqiqotlar, biznes loyihalari va real dunyo muammolarini hal qilish jarayonida katta ahamiyatga ega. Shu bilan birga, birlashtirilgan ma'lumotlar ustida filtrlar, guruhlash va agregatsiya amallarini bajarish mumkin, bu esa foydalanuvchiga murakkab tahlillarni tezkorlik bilan amalga oshirish imkonini beradi.

Panda kutubxonasi shuningdek, vaqt qatori (time series) ma'lumotlarini qayta ishlash imkoniyatini ham taqdim etadi. Vaqt qatori ma'lumotlarini indekslash, resampling, rolling va shifting operatsiyalari foydalanuvchiga vaqtga bog'liq tahlil va prognozlarni yaratishga yordam beradi.

```
data['sana'] = pd.to_datetime(data['sana'])  
monthly_data = data.resample('M', on='sana').mean()
```


Yuqoridagi kod ma'lumotlarni oylar bo'yicha qayta namoyish etadi va har bir oy bo'yicha o'rtacha qiymatlarni hisoblaydi Shu tarzda Panda kutubxonasi nafaqat oddiy ma'lumotlarni, balki murakkab vaqt qatori ma'lumotlarini tahlil qilish imkonini ham beradi

Panda kutubxonasi ma'lumotlar tahlili va ilmiy tadqiqotlarda eng muhim vositalardan biri hisoblanadi Chunki zamonaviy loyihalarda foydalanuvchiga katta hajmdagi ma'lumotlar bilan ishlash va ularni samarali boshqarish talab etiladi DataFrame va Series tuzilmalari foydalanuvchiga murakkab ma'lumotlarni qulay va intuitiv tarzda boshqarish imkonini beradi Shu bilan birga, kutubxona foydalanuvchiga ma'lumotlar ustida turli transformatsiya va hisoblash amallarini tezkorlik bilan bajarishga yordam beradi

Ma'lumotlar bilan ishlash jarayonida birinchi bosqich sifatida ularni tozalash va tayyorlash muhimdir Bo'sh qiymatlar, noto'g'ri formatlangan yozuvlar va redundansiyani aniqlash va tuzatish amallari ma'lumotlarni tahlil qilish sifatini sezilarli darajada oshiradi Shu bilan birga, foydalanuvchi ustunlar nomlarini o'zgartirish, yangi ustunlar yaratish va mavjud ma'lumotlarni transformatsiya qilish imkoniyatiga ega Pandada bu jarayonlar uchun qulay funksiyalar mavjud bo'lib, ular foydalanuvchining ishini sezilarli darajada yengillashtiradi

Ma'lumotlarni guruhlash va agregatsiya qilish imkoniyatlari tahlil jarayonini yanada kuchaytiradi Foydalanuvchi ma'lumotlarni ma'lum mezonlar bo'yicha guruhlab, har bir guruh bo'yicha statistik ko'rsatkichlarni hisoblay oladi Bu o'rtacha qiymat, maksimal va minimal qiymat, median va dispersiya kabi ko'rsatkichlarni o'z ichiga oladi Shu bilan birga, foydalanuvchi murakkab hisoblashlarni yangi ustunlar yaratish orqali amalga oshirishi mumkin, bu esa katta hajmdagi ma'lumotlarni samarali tahlil qilish imkonini beradi

Vizualizatsiya esa Panda kutubxonasining muhim jihatlaridan biridir Grafik ko'rinish ma'lumotlarni intuitiv tushunishga yordam beradi, trendlar va o'zaro bog'liqliklarni aniqlashni osonlashtiradi Kutubxona Matplotlib va Seaborn bilan integratsiyalashgan bo'lib, foydalanuvchi turli grafiklar, diagrammalar, scatter plot va histogrammalar

yordamida natijalarni vizual tarzda ko'rsatishi mumkin Bu ilmiy tadqiqotlar va amaliy loyihalarda qaror qabul qilish jarayonini sezilarli darajada tezlashtiradi

Panda kutubxonasi amaliy loyihalarda ham keng qo'llaniladi Masalan, moliya sohasida narxlar o'zgarishini tahlil qilish, savdo va marketing bo'yicha ma'lumotlarni segmentlash, sog'liqni saqlash sohasida bemorlar ma'lumotlarini tahlil qilish va ilmiy tadqiqotlarda eksperiment natijalarini qayta ishlash kabi jarayonlarda kutubxona foydalanuvchiga muhim vosita sifatida xizmat qiladi Shu bilan birga, kutubxona foydalanuvchi uchun qulay interfeys yaratadi va katta hajmdagi ma'lumotlar bilan ishlash jarayonini sezilarli darajada tezlashtiradi

Panda kutubxonasi foydalanuvchiga ma'lumotlarni tahlil qilishda yagona vosita bo'lib qolmay, balki turli sohalarda, masalan, ilmiy tadqiqotlar, biznes loyihalar va ijtimoiy tahlillarda ham muhim rol o'ynaydi Kutubxonaning imkoniyatlari foydalanuvchiga murakkab ma'lumotlarni samarali boshqarish, statistik tahlil qilish va natijalarni vizual tarzda taqdim etishga yordam beradi Shu bilan birga, foydalanuvchi kutubxona funksiyalarini kod orqali moslashtirishi va o'z loyihalariga integratsiya qilishi mumkin

Panda kutubxonasi ma'lumotlar bilan ishlash jarayonini soddalashtirish bilan birga, foydalanuvchiga tezkor va intuitiv tarzda tahlil olib borish imkonini beradi Kutubxona yordamida foydalanuvchi ma'lumotlarni tozalash, transformatsiya qilish, guruhlash, agregatsiya va vizualizatsiya kabi jarayonlarni yagona platformada amalga oshirishi mumkin Shu tariqa Panda kutubxonasi ilmiy tadqiqotlar va amaliy loyihalarda ma'lumotlar bilan ishlashning asosiy vositasi sifatida tanilgan va keng qo'llaniladi

Panda kutubxonasi foydalanuvchiga ma'lumotlar ustida turli filtrlar va indekslash operatsiyalarini bajarish imkonini beradi Bu jarayon ma'lumotlar tahlilining muhim qismi bo'lib, foydalanuvchiga ma'lum shartga javob beruvchi yozuvlarni tezkorlik bilan tanlab olish imkonini beradi Misol uchun, DataFrame tuzilmasida ma'lum ustun qiymatiga mos yozuvlarni filtrlash quyidagicha amalga oshiriladi

```
filtered_data = data[data['yosh'] > 30]
```

Ushbu kod yosh ustunidagi qiymati 30 dan katta bo‘lgan yozuvlarni ajratadi Shu bilan birga, foydalanuvchi bir nechta shartlarni birlashtirib murakkab filtrlarni yaratishi mumkin

```
filtered_data = data[(data['yosh'] > 30) & (data['kategoriya'] == 'A')]
```

Bu filtr yordamida foydalanuvchi yosh ustuni 30 dan katta bo‘lgan va kategoriya ustuni “A” bo‘lgan yozuvlarni tanlab oladi Pandada indekslash ham muhim operatsiya bo‘lib, foydalanuvchi DataFrame ustunlari yoki qatorlarini indekslash orqali ma’lumotlarga tezkor kirish imkoniyatiga ega

```
data.set_index('id', inplace=True)  
row = data.loc[1001]
```

Yuqoridagi kod orqali foydalanuvchi 1001 raqamli indeksga mos yozuvga tezkorlik bilan murojaat qiladi Indekslish va filtrlar birgalikda ishlatilganda katta hajmdagi ma’lumotlarni tez va samarali tahlil qilish imkonini beradi Shu bilan birga, bu jarayon ilmiy tadqiqotlarda va real loyihalarda qaror qabul qilish uchun asosiy vosita sifatida ishlatiladi

Katta hajmdagi loyihalarda turli manbalardan olingan ma’lumotlarni birlashtirish muhim jarayon hisoblanadi Panda kutubxonasi foydalanuvchiga bir nechta DataFrame obyektlarini birlashtirish va ularni yagona tuzilma shakliga keltirish imkonini beradi Bu esa ma’lumotlar tahlilini soddalashtiradi va loyihalarning murakkabligini kamaytiradi

```
merged_data = pd.merge(data1, data2, on='id', how='inner')
```

Yuqoridagi kod ikki DataFrame obyektini “id” ustuni bo‘yicha birlashtiradi va faqat ikkala DataFrame’da mavjud yozuvlarni oladi Shu bilan birga, foydalanuvchi turli usullarda birlashtirishni amalga oshirishi mumkin Masalan, chap birlashtirish, o‘ng birlashtirish yoki to‘liq birlashtirish

```
merged_data = pd.merge(data1, data2, on='id', how='left')
```

Ma'lumotlarni birlashtirish orqali foydalanuvchi bir nechta manbalardan olingan yozuvlarni yagona tahlil qilish imkoniga ega bo'ladi. Bu esa ilmiy tadqiqotlar, biznes loyihalari va real dunyo muammolarini hal qilish jarayonida katta ahamiyatga ega. Shu bilan birga, birlashtirilgan ma'lumotlar ustida filtrlar, guruhlash va agregatsiya amallarini bajarish mumkin, bu esa foydalanuvchiga murakkab tahlillarni tezkorlik bilan amalga oshirish imkonini beradi.

Panda kutubxonasi shuningdek, vaqt qatori (time series) ma'lumotlarini qayta ishlash imkoniyatini ham taqdim etadi. Vaqt qatori ma'lumotlarini indekslash, resampling, rolling va shifting operatsiyalari foydalanuvchiga vaqtga bog'liq tahlil va prognozlarni yaratishga yordam beradi.

```
data['sana'] = pd.to_datetime(data['sana'])  
monthly_data = data.resample('M', on='sana').mean()
```

Yuqoridagi kod ma'lumotlarni oylar bo'yicha qayta namoyish etadi va har bir oy bo'yicha o'rtacha qiymatlarni hisoblaydi. Shu tarzda Panda kutubxonasi nafaqat oddiy ma'lumotlarni, balki murakkab vaqt qatori ma'lumotlarini tahlil qilish imkonini ham beradi.

Xulosa

Ushbu maqolada Panda kutubxonasi va uning qo'llanilishi ilmiy va amaliy jihatlar bilan keng tahlil qilindi. Panda Python dasturlash tilida ma'lumotlar bilan ishlashni sezilarli darajada soddalashtiruvchi kuchli vosita sifatida ko'rib chiqildi. DataFrame va Series tuzilmalari foydalanuvchiga katta hajmdagi ma'lumotlarni samarali boshqarish, transformatsiya qilish, guruhlash, agregatsiya va vizualizatsiya qilish imkonini beradi.

Maqolada ma'lumotlarni tozalash, filtr va indekslash, guruhlash va agregatsiya, shuningdek, vaqt qatori ma'lumotlarini qayta ishlash jarayonlari ilmiy jihatdan tahlil qilindi. Kod misollari orqali amaliy qo'llanilishi ko'rsatildi. Shuningdek, vizualizatsiya vositalari yordamida ma'lumotlar trendlarini aniqlash va qaror qabul qilish jarayoni yoritildi.

Panda kutubxonasi real loyihalarda, moliya, marketing, sog‘liqni saqlash, ilmiy tadqiqotlar va boshqa sohalarda keng qo‘llaniladi. Kutubxonaning imkoniyatlari foydalanuvchiga murakkab ma’lumotlarni intuitiv tarzda tahlil qilish va natijalarni samarali tarzda taqdim etishga yordam beradi. Shu tariqa, Panda kutubxonasi nafaqat ilmiy tadqiqotlar, balki amaliy loyihalarda ham asosiy vosita sifatida ishlatilishi mumkin.

Adabiyotlar ro‘yxati

- McKinney W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd Edition, O'Reilly Media, 2017
- VanderPlas J. Python Data Science Handbook: Essential Tools for Working with Data, O'Reilly Media, 2016
- McKinney W. Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 2010
- Reback J., McKinney W., jbrockmendel, et al. pandas-dev/pandas: Pandas, Zenodo, 2020
- Python Software Foundation. Pandas Documentation, <https://pandas.pydata.org/docs/>
- Grus J. Data Science from Scratch: First Principles with Python, 2nd Edition, O'Reilly Media, 2019