

---

## PYTHONDA MA'LUMOTLAR TAHLILI

---

**Hamroyev Bobirjon Baxritdinovich**

Osiyo xalqaro universiteti,  
“Umumtexnik fanlar” kafedrası o'qituvchisi

**Annotatsiya:** Ushbu maqolada Python dasturlash tilining ma'lumotlar tahlili sohasidagi qo'llanilishi haqida batafsil ma'lumot berilgan. NumPy, Pandas, Matplotlib, Seaborn va boshqa kutubxonalar yordamida ma'lumotlarni yig'ish, tozalash, tahlil qilish, vizualizatsiya qilish va ilmiy hisoblash imkoniyatlari yoritilgan. Shuningdek, ushbu sohada Pythonning afzalliklari, cheklovlari va kelajak istiqbollari muhokama qilingan.

### **Kirish**

Ma'lumotlar tahlili hozirgi zamonaviy texnologiyalar rivojlanishida muhim o'rin tutadi. Hozirgi kunda sun'iy intellekt, katta ma'lumotlar (Big Data), va avtomatlashtirish sohalari ma'lumotlar tahlilisiz rivojlana olmaydi. Python dasturlash tili ma'lumotlarni boshqarish va tahlil qilish jarayonlarini osonlashtiradigan kuchli vosita sifatida tanilgan.

Pythonning o'ziga xosligi uning soddaligida va kuchli kutubxonalarida namoyon bo'ladi. NumPy, Pandas, Matplotlib, va Seaborn kabi kutubxonalar Pythonni nafaqat dasturiy jarayonlarni avtomatlashtirishda, balki katta hajmdagi ma'lumotlar bilan ishlashda ham muhim vositaga aylantiradi.

### **Pythonning ma'lumotlar tahlilidagi o'ri**

Python ochiq manbaliligi va keng funktsionalligi bilan boshqa dasturlash tillariga nisbatan ma'lumotlar tahlilida ko'proq qo'llaniladi. Quyida Pythonning ma'lumotlar tahlilidagi asosiy imkoniyatlari keltirilgan:

**Ma'lumotlarni yig'ish va tozalash:** Python yordamida API orqali ma'lumotlarni yig'ish va birlamchi tozalash oson amalga oshiriladi. **Ma'lumotlar tahlili:** Kutubxonalar yordamida katta hajmdagi ma'lumotlarni qayta ishlash va tahlil qilish mumkin. **Vizualizatsiya:** Matplotlib va Seaborn kabi kutubxonalar grafik va diagrammalar orqali tahlil natijalarini tasvirlashga imkon beradi. **Statistik modellashtirish:** StatsModels va SciPy kutubxonalari statistik tahlil va modellashtirish uchun keng imkoniyat yaratadi. **Avtomatlashtirish:** Python skriptlari jarayonlarni avtomatlashtirish va real vaqt tahlilini amalga oshirishda foydalaniladi.

### **Asosiy kutubxonalar**

Python ma'lumotlar tahlilida turli kutubxonalarga boy. Quyida ularning asosiylari keltirilgan: **NumPy** NumPy – bu sonli hisoblashlar uchun asosiy kutubxona bo'lib, katta o'lchamli massivlar va matritsalar bilan ishlashni osonlashtiradi.

### **Imkoniyatlari:**

Ko'p o'lchovli massivlar yaratish.

Algebraik va trigonometrik hisoblash.

Statistik tahlillar uchun tayyor funksiyalar.

### **Matematik operatsiyalar:**

NumPy orqali quyidagi amallarni tez va samarali bajarish mumkin:

Qo'shish, ayirish, ko'paytirish va bo'lish kabi asosiy arifmetik operatsiyalar.

Algebraik operatsiyalar (matritsa determinantini hisoblash, o'zaro teskari matritsalarini topish).  
Statistik hisob-kitoblar (o'rtacha qiymat, dispersiya, mediana va h.k.).

**Moslashuvchanlik:** NumPy massivlari o'lchamini osongina o'zgartirish va ularga ma'lumot qo'shish mumkin. Masalan, massivni qayta shakllantirish (reshape funksiyasi) yoki uning o'lchamini kengaytirish (hstack, vstack) oson amalga oshiriladi.

**Pandas** Pandas ma'lumotlarni boshqarish va qayta ishlash uchun qulay vositadir. Ma'lumotlar jadvallarini boshqarish uchun asosiy kutubxona sifatida foydalaniladi.

### **Imkoniyatlari:**

Ma'lumotlarni CSV, JSON, va SQL formatlaridan o'qish va saqlash.

Filtrlash, guruhlash va tozalash funksiyalari.

Vaqt qatorlarini tahlil qilish imkoniyati.

**Matplotlib** Matplotlib statistik va grafik chizish vositasi bo'lib, chiziqli grafiklar, histogramlar va boshqa diagrammalar yaratish uchun ishlatiladi.

### **Imkoniyatlari:**

Grafik va diagrammalarni formatlash.

2D va 3D grafiklar yaratish.

Maxsus grafik sozlamalari.

**Seaborn** Seaborn Matplotlib asosida yaratilgan va statistik tahlil natijalarini vizualizatsiya qilishni soddalashtiradi.

### **Imkoniyatlari:**

Yuqori sifatli statistik grafiklar yaratish.

Korrelyatsiyalarni ko'rsatish.

Estetik jihatdan yuqori grafikalar.

**StatsModels** StatsModels statistik tahlil va regressiyalarni amalga oshiruvchi kuchli kutubxonadir.

### **Imkoniyatlari:**

Vaqt qatorlarini tahlil qilish.

ANOVA va regressiya tahlillari.

Model natijalarini baholash va vizualizatsiya qilish.

**Afzalliklar va cheklovlar** Pythonning ma'lumotlar tahlilidagi kuchli tomonlari va ba'zi kamchiliklari quyidagilardan iborat:

**Afzalliklar:** Soddalik va o'rganish uchun qulaylik. Boy kutubxonalar va vositalar mavjudligi.

Ochiq manba bo'lgani sababli bepul foydalanish imkoniyati. **Cheklovlar:** Og'ir hisoblashlar uchun Pythonning sekin ishlashi. Katta hajmdagi ma'lumotlar bilan ishlashda ko'proq xotira talab qilishi. **Kelajak istiqbollari**

Python dasturlash tili ma'lumotlar tahlilida o'z dolzarbligini yo'qotmaydi.

**Big Data:** Python katta ma'lumotlarni qayta ishlashda Spark va Hadoop kabi tizimlar bilan birga ishlatiladi. **Sun'iy intellekt:** TensorFlow va PyTorch kabi kutubxonalar orqali Python AI tizimlarida keng qo'llaniladi. **Ma'lumotlarni avtomatlashtirish:** Python avtomatlashtirish vositalarini takomillashtirishda foydalaniladi.

### **Xulosa**

Python ma'lumotlar tahlili uchun mos va kuchli dasturlash tili hisoblanadi. Unga asoslangan kutubxonalar yordamida nafaqat ma'lumotlarni boshqarish, balki murakkab tahlillarni ham amalga oshirish mumkin.

### **Foydalanilgan adabiyotlar:**

1. McKinney, W. Python for Data Analysis. O'Reilly Media, 2017.
2. VanderPlas, J. Python Data Science Handbook. O'Reilly Media, 2016.
3. Oliphant, T. Guide to NumPy. 2nd Edition, 2015.
4. Grus, J. Data Science from Scratch. O'Reilly Media, 2019.