شرح مسئله:

شما یک فایل CSV به نامstudents.csv دارید که اطلاعات دانش‌آموزان یک کلاس در آن ذخیره شده است. این فایل دارای ستون‌های زیر است (ممکن است برخی مقادیر خالی وجود داشته باشد):

student\_id: شماره دانش‌آموزی

name: نام دانش‌آموز

age: سن

grade: مقطع تحصیلی

math\_score: نمره ریاضی

science\_score: نمره علوم

هدف: شما باید این داده‌ها را بخوانید، تمیز کنید، تحلیل‌های اولیه انجام دهید و به سوالات پاسخ دهید.

مراحل تمرین

گام ۱: راه‌اندازی و خواندن داده‌ها

1. کتابخانه pandas را ایمپورت کن.

2. فایل students.csv را بخوان و آن را در یک DataFrame به نام df ذخیره کن.

3. از ستون student\_id به عنوان ایندکس DataFrame استفاده کن.

گام ۲: بررسی اولیه داده‌ها

1. پنج سطر اول داده‌ها را نمایش بده.

2. خلاصه اطلاعات DataFrame را با استفاده از info() چاپ کن.

3. آمار توصیفی داده‌های عددی را با describe() نمایش بده.

گام ۳: مدیریت مقادیر خالی

1. تعداد مقادیر خالی در هر ستون را پیدا کن.

2. مقادیر خالی ستون‌های عددی (age, math\_score, science\_score) را با میانگین همان ستون پر کن.

3. مقادیر خالی ستون‌های غیرعددی (name, grade) را با مقدار "Unknown" پر کن.

گام ۴: تغییر ساختار داده‌ها

1. نام ستون‌ها را به فارسی تغییر بده (یا برعکس):

name → نام

age → سن

grade → مقطع

math\_score → نمره\_ریاضی

science\_score → نمره\_علوم

2. ستون student\_id را حذف کن (چون قبلاً آن را به عنوان ایندکس انتخاب کردی).

گام ۵: فیلتر و انتخاب داده‌ها

1. فقط دانش‌آموزانی را نمایش بده که سن آن‌ها بیشتر از ۲۰ سال است.

2. فقط ستون‌های نام و نمره\_ریاضی را برای تمام دانش‌آموزان انتخاب کن.

3. با استفاده از iloc، ردیف‌های ۵ تا ۱۰ و ستون‌های ۱ تا ۳ را انتخاب کن.

4. با استفاده از loc، ردیف‌های با ایندکس ۱۰۱ تا ۱۰۵ و ستون‌های نام و سن را انتخاب کن.

گام ۶: مرتب‌سازی

1. داده‌ها را بر اساس نمره\_ریاضی به صورت نزولی مرتب کن.

2\*. داده‌ها را ابتدا بر اساس سن (صعودی) و سپس بر اساس نمره\_علوم (نزولی) مرتب کن.

گام ۷: سوالات تحلیلی

1. چند دانش‌آموز در مقطع "دهم" هستند؟

2. میانگین نمره ریاضی تمام دانش‌آموزان چقدر است؟

3. دانش‌آموزی که بالاترین نمره علوم را دارد، چه نام دارد؟

نکته:

فرض کن فایل students.csv در همان پوشه‌ای قرار دارد که کد تو اجرا می‌شود. اگر با مشکل خواندن فایل مواجه شدی، می‌توانی یک DataFrame نمونه با داده‌های فرضی درست کنی و مراحل را روی آن اجرا کنی.