به نام خدا

Python 🡪 Pandas 🡪 Professional

import pandas as pd

کتابخانه پانداز را در پروژه ایمپورت می کنیم و با نام مستعار خطاب می کنیم

Mydict = {

‘Name’ : [‘Ali’, ‘Negar’, ‘Sara’, ‘Mahsa’],

‘Age’ : [37, 45, 29, 32],

‘Score’ : [19, 20, 17.5, 18.25]

}

Df = pd.DataFrame(mydict)

ابتدا یک دیکشنری درست کرده و باید با همین نظام ترتیب و قائده مقدار ها را قرار داد.

سپس تابع دیتا فریم را از پانداز گرفته و دیکشنری بالا را داخل آن قرار می دهیم. اگر آبجکتی که در آن دیتا فریم را ساخته ایم در محیط مناسب به صورت کونسول کال کنیم – دیتا فریم شده اطلاعات را مشاهده می کنیم.

Df[[‘Name’, ‘Age’]]

در اینجا می توان عناوین داخل ستون نام و سن را مشاهده کرد

Df.iloc[:, 0:2]

در اینجا نیز می توان دو ستون اول را عناوین داخل شان را مشاهده کرد

Sdf = df.sort\_values(‘Age’)

اینجا ستون سن را مرتب کردیم اما چون می دانیم رو خود دیتا فریم اصلی تاثیر نمی گذارد – آن را در یک متغیر دیگر ذخیره می کنیم

df.sort\_values(‘Age’, inplace= True)

در اینجا نیز مانند بالا عمل کدیم اما چون میی خواهیم مرتب شده ستون سن در خود فریم اصلی ذخیره شود پارامتری که نوشته شده را برابر با درست یا صحیح قرار می دهیم.

Df[‘Age’].max()

Df[‘Score’].min()

در اینجا اول ستون سن ها بزرگترینشان را بر می گرداند و سپس ستون امتیازات کوچکترینشان را بر می گرداند.

Df.mean()

در اینجا از ستون های عددی میانگین می گیرد و میانگین آن ها را بر می گرداند

Df[[‘Score’]].mean()

در اینجا ما فقط از ستون خاصی میانگین میگیریم. مثلا در اینجا ستون امتیازات.

Df[[‘Score’]].describel()

تمام جزئیات مربوط به ستون امتیازات رو نمایش می دهد.

Df[df[‘Age’] < 35]

ردیف هایی که سن انها کمتر از سی و پنج است را نمایش می دهد.

Dfn = df.rename(columns={‘Score’ : ‘Grade’})

نام ستون امتیازات رو تغییر می دهیم به نمره

Df[‘Name’] = df[‘Name’].replace({‘Ali’ : ‘Ali reza’})

در ستون اسم هر جا علی دید – تبدیل به علی رضا می کند

a = np.arange(12).reshape(4, 3)

df = pd.DataFrame(a, columns= [‘A’, ‘B’, ‘C’])

df.drop(columns ‘B’)

می توان یک آرایه را در دیتافریم قرار داد و سپس با پارامتر نوشته شده نام ستون ها را وارد کرد

سپس در اینجا ستون ب را حذف کردیم

Df.drop([‘A’, ‘C’], axis= 1)

می توان به روش دیگر نیز ستون را حذف کرد اما با پارامتر نوشته شده و دادن عدد یک به معنای سطر. زیرا نام ستون ها را در سطر اول مشاهده می کنیم.

برای حذف سطر هم از همین روش ها می رویم اما فقط در قالب لیست نام سطر ها را می نویسیم.

Df.to\_excel(‘d:student.xlsx’, index= False)

S = pd.read\_excel(‘d:student.xlsx’)

کد اول دیتا فرم های پانداز را به فایل اکسل تبدیل می کند. ایندکس که نوشته شده به معنای این است که کد سطر ها که از صفر شروع میشود باشد یا خیر که ما برابر نادرست یا خیر قرار داده ایم.

و کد دوم فایل اکسل را می خواند و به دیتافرم های پانداز تبدیل می کند.

نکته: هر گاه از کد زیر در دیکشنری که می خواهیم آن را به دیتا فرم های پانداز تبدیل کنیم – بیاوریم... به معنای جای خالیاست و می توانیم با نامپای در او تغییراتی ایجاد کنیم.

Np.NaN

Frame[frame[‘col2’].isna()]

در اینجا میاد و در ستون دوم هرجا به کلمه کلیدی خالی رسید – ردیف اون رو میاورد.

Frame[frame[‘col1’].notna()]

در اینجا میاد و بررسی می کند در ستون اول کجا کلمه کلیدی خالی بکار رفته است – و ردیف اون رو حذف کرده – و اصلی رو نشون میده

Dt = df.groupby(‘Team’)

Gt.first()

در اینجا دیتافریم دی اف را بر اساس تیم ها گروه و دسته می کنیم. دستور بعدی – اولین آیتم هر گروه را بر می گرداند.

Gt.get\_group(‘A’)

Gt.get\_group(‘B’)

آیتم های داخل هر گروه که گفتیم را بر می گرداند.

Gt.sum()

این دستور مجموع ستون های عددی هر گروه را بر می گرداند.

Gt.agg([‘max’])

از هر گروه بزرگترین آیتم را بر می گرداند.

Gt.agg([‘max’, ‘min’])

در این جا بزرگترین و کوچکترین را برمیگرداند.