

OTVÁRANIE A ČÍTANIE SÚBOROV

💡 Otváranie & Čítanie TXT Súborov

🔗 with

```
with open([adresa_súboru], [režim('r'/'w')], encoding=[encoding]) as file:
    content = file.read()
    print(content)
```

⚙️ Nainštaluj si Rainbow CSV Extension

- 🔍 **Čo to je?** Extension na prácu s CSV súbormi.
- 💡 **Prečo to potrebujem?** CSV súbory sú riadky s hodnotami oddelenými čiarkami. Táto extension automaticky zafarbí každý 'stĺpec' inou farbou pre lepšiu viditeľnosť.
- 🔗 **Kde to zoženiem?** VSCode → ctrl + shift + X → vyhľadaj & nainštaluj 'Rainbow CSV'

💡 Otváranie & Čítanie CSV Súborov

🔗 with

```
import csv

with open([adresa_súboru], [režim], [encoding]) as file:
    reader = csv.reader(file)

    for row in reader:
        print(row)
```

ANALÝZA TABULIEK S pandas

💡 Premeň CSV súbory na pandas dataframes

🔗 read_csv

```
import pandas as pd

df = pd.read_csv([adresa_súboru])
```

💡 Zobraz základné info o tabuľke s pandas

SYNTAX	ČO TO ROBÍ
df.head(k)	zobrazí prvých k riadkov tabuľky, default je 5
df.tail(k)	zobrazí posledných k riadkov tabuľky, default je 5
df.shape	vráti tuple (počet_riadkov, počet_stĺpcov)

💡 Manipulácia so stĺpcami v pandas

SYNTAX	Čo to Robí
<code>df[col_name]</code>	vráti všetky hodnoty v danom stĺpci ako iterovateľnú sériu (series)
<code>df[df[col_name] == val]</code>	porovná všetky hodnoty v danom stĺpci s hodnotou, vráti len riadky, ktoré prejdú
<code>df[new_col_name] = series</code>	priradí elementy series jeden po druhom do riadkov dataframeu ako nový stĺpec
<code>df.sort_values(col_name)</code>	zoradí hodnoty stĺpca v vzostupnom/abecednom poradí

💡 Vyhľadávanie konkrétnych buniek v pandas

SYNTAX	Čo to Robí
<code>df.loc[k, col_name]</code>	vráti bunku v k-tom riadku daného stĺpca
<code>df.iloc[m, n]</code>	vráti bunku v m-tom riadku a n-tom stĺpci

💡 Zoskupovanie hodnôt v pandas

<code>df.groupby(col_name).func()</code>	zoskupí unikátne hodnoty z col_name a aplikuje agregáčnú funkciu
<code>df.groupby(col1)[col2].func()</code>	zoskupí unikátne hodnoty v col1, aplikuje agregáčnú funkciu na col2

💡 Ďalšie štatistiky s pandas

<code>df.describe()</code>	zobrazí základné štatistiky pre všetky numerické stĺpce
<code>df[col_name].median()</code>	vypíše medián daného stĺpca
<code>df[col_name].mean()</code>	vypíše priemer daného stĺpca
<code>df[col_name].std()</code>	vypíše štandardnú odchýlku daného stĺpca
<code>df[col_name].sum()</code>	vypíše súčet daného stĺpca
<code>df[col_name].min()</code>	vypíše minimálnu hodnotu v danom stĺpci
<code>df[col_name].max()</code>	vypíše maximálnu hodnotu v danom stĺpci

Zisti viac o veciach, ktoré sa dajú robiť s pandas dataframe [tu](#) 