Přiložený datový soubor obsahuje volně dostupné informace z meteorologické stanice v pražském Klementinu. Proveď základní analýzu dat a přidej k tomu popis.

GPT

Data soubor obsahuje historické informace o meteorologické stanici v pražském Klementinu, včetně začátku měření, místa měření, výšky nad terénem, provozovatele a dalších podrobností o měřicích přístrojích a metodách. Záznamy se zdají obsahovat různé typy informací v různých sloupcích, včetně některých, které nejsou jasně označeny nebo jsou nevyplněny (např. mnoho sloupců s názvem "Unnamed").

Pro hlubší analýzu je třeba zjistit, která část dat obsahuje relevantní a strukturované meteorologické záznamy, pokud jsou v souboru přítomny. Vypadá to, že prvních pár řádků obsahuje historické informace o samotné stanici a jejích pozorovatelích, ne meteorologická data jako taková.

Pokud hledáte analýzu konkrétních meteorologických dat (např. teplot, srážek atd.), mohlo by být nutné prozkoumat další části souboru, kde by taková data mohla být uložena. Mohu prozkoumat soubor dále, abych našel relevantní část pro analýzu, pokud existuje. Chcete, abych pokračoval? ​

Soubor obsahuje celkem tři listy a nejvíce mě zajímá list s názvem "data", který by měl obsahovat údaje o měření teploty v jednotlivých dnech od roku 1775 až do roku 2022. Chtěl bych vytvořit program v jazyce Python, který dokáže tato data načíst do vhodné datové struktury a poté uživateli umožní získávat různé analytické informace s využitím funkcí, ještě lépe objektů a metod. Navrhni, prosím, nejlepší řešení.

GPT

Chcete-li zpracovávat a analyzovat data o teplotě z listu "data" ve vašem souboru Excel můžeme vytvořit Python program, který používá knihovnu pandas pro načtení a manipulaci s daty, a objektově orientované programování (OOP) pro strukturování analytických funkcí. Zde je návrh řešení:

Načtení dat: Nejprve načteme data z listu "data" souboru Excel do pandas DataFrame. DataFrame je vhodná datová struktura pro práci s tabulkovými daty.

Definice třídy pro analýzu: Vytvoříme třídu TemperatureAnalytics, která bude obsahovat metody pro různé typy analýz, jako je výpočet průměrné, maximální a minimální teploty, analýzu trendů teploty atd.

Metody pro analýzu: Pro každou analytickou operaci definujeme metodu uvnitř třídy TemperatureAnalytics. Tyto metody mohou zahrnovat:

get\_average\_temperature(year): Vrátí průměrnou teplotu pro zadaný rok.

get\_max\_temperature(year): Vrátí nejvyšší teplotu a datum, kdy k ní došlo, pro zadaný rok.

get\_min\_temperature(year): Vrátí nejnižší teplotu a datum, kdy k ní došlo, pro zadaný rok.

get\_monthly\_averages(year): Vrátí průměrnou teplotu pro každý měsíc v zadaném roce.

analyze\_temperature\_trends(start\_year, end\_year): Analyzuje trendy teplot mezi zadanými roky.

Interaktivní uživatelské rozhraní: Vytvoříme jednoduché textové rozhraní, které umožní uživatelům zadávat požadavky a získávat výsledky analýz.

Začneme s načtením dat a definicí základní struktury třídy:

Data byla úspěšně načtena a obsahují sloupce pro rok (rok), měsíc (měsíc), den (den), průměrnou teplotu (T-AVG), maximální teplotu (TMA), minimální teplotu (TMI), srážky (SRA) a případný flag (Flag).

Nyní vytvoříme třídu TemperatureAnalytics, která bude obsahovat metody pro provedení různých analytických operací s těmito daty: ​​

Analýza byla úspěšně provedena pro rok 2022 s následujícími výsledky:

Průměrná teplota za rok 2022 je 12.41°C.

Nejvyšší teplota zaznamenaná v roce 2022 byla 36.8°C, a to 19. června.

Nejnižší teplota zaznamenaná v roce 2022 byla -7.2°C, a to 14. prosince.

Třída TemperatureAnalytics je flexibilní a umožňuje provádět různé analýzy teplotních dat, jako je výpočet průměrů, maxím a minim pro zvolené roky, stejně jako analýzu měsíčních průměrů a trendů teploty. Tento přístup může být dále rozšířen o další metody podle potřeb analýzy.

Pro interaktivní práci s tímto analytickým nástrojem by mohlo být užitečné vytvořit jednoduché textové rozhraní nebo webové rozhraní, které by uživatelům umožnilo specifikovat typ analýzy a roky pro analýzu.

Pokud máte další specifické požadavky nebo dotazy ohledně analýzy, neváhejte se zeptat. ​​

Výborně, nyní bych požádal o výpis daného kódu v Pythonu.

GPT

Samozřejmě, zde je kompletní kód Pythonu, který jsme připravili pro analýzu meteorologických dat z pražského Klementina, včetně načtení dat a definice třídy TemperatureAnalytics s jejími metodami:

import pandas as pd

# Načtení dat z Excel souboru

data\_path = 'cesta/k/souboru/PKLM\_pro\_portal.xlsx' # Upravte cestu k vašemu souboru

data\_sheet\_name = 'data'

temperature\_data = pd.read\_excel(data\_path, sheet\_name=data\_sheet\_name)

# Definice třídy pro analýzu teplot

class TemperatureAnalytics:

def \_\_init\_\_(self, data):

self.data = data

def get\_average\_temperature(self, year):

yearly\_data = self.data[self.data['rok'] == year]

return yearly\_data['T-AVG'].mean()

def get\_max\_temperature(self, year):

yearly\_data = self.data[self.data['rok'] == year]

max\_temp = yearly\_data['TMA'].max()

date\_of\_max\_temp = yearly\_data[yearly\_data['TMA'] == max\_temp][['rok', 'měsíc', 'den']].iloc[0]

return max\_temp, date\_of\_max\_temp

def get\_min\_temperature(self, year):

yearly\_data = self.data[self.data['rok'] == year]

min\_temp = yearly\_data['TMI'].min()

date\_of\_min\_temp = yearly\_data[yearly\_data['TMI'] == min\_temp][['rok', 'měsíc', 'den']].iloc[0]

return min\_temp, date\_of\_min\_temp

def get\_monthly\_averages(self, year):

yearly\_data = self.data[self.data['rok'] == year]

return yearly\_data.groupby('měsíc')['T-AVG'].mean()

def analyze\_temperature\_trends(self, start\_year, end\_year):

trend\_data = self.data[(self.data['rok'] >= start\_year) & (self.data['rok'] <= end\_year)]

annual\_average\_temperatures = trend\_data.groupby('rok')['T-AVG'].mean()

return annual\_average\_temperatures

# Vytvoření instance třídy a provedení analýzy

temperature\_analytics = TemperatureAnalytics(temperature\_data)

# Příklad použití

average\_temp\_2022 = temperature\_analytics.get\_average\_temperature(2022)

max\_temp\_2022, date\_of\_max\_temp\_2022 = temperature\_analytics.get\_max\_temperature(2022)

min\_temp\_2022, date\_of\_min\_temp\_2022 = temperature\_analytics.get\_min\_temperature(2022)

print(f"Průměrná teplota v roce 2022: {average\_temp\_2022}°C")

print(f"Maximální teplota v roce 2022: {max\_temp\_2022}°C, datum: {date\_of\_max\_temp\_2022['den']}.{date\_of\_max\_temp\_2022['měsíc']}.{date\_of\_max\_temp\_2022['rok']}")

print(f"Minimální teplota v roce 2022: {min\_temp\_2022}°C, datum: {date\_of\_min\_temp\_2022['den']}.{date\_of\_min\_temp\_2022['měsíc']}.{date\_of\_min\_temp\_2022['rok']}")