

PROVA FINALE PYTHON (MODULO 4)

PUNTO 1 - Si richiede di verificare le dimensioni del dataset e le diciture presenti nell'intestazione.

```
In [1]: import csv

with open("./owid_covid_data.csv", newline="", encoding="ISO-8859-1") as filecsv:
    lettore = csv.reader(filecsv, delimiter=",")
    header = next(lettore)
    print(header)
```

```
['iso_code', 'continent', 'location', 'date', 'total_cases', 'new_cases', 'new_cases_smoothed', 'total_deaths', 'new_deaths', 'new_deaths_smoothed', 'total_cases_per_million', 'new_cases_per_million', 'new_cases_smoothed_per_million', 'total_deaths_per_million', 'new_deaths_per_million', 'new_deaths_smoothed_per_million', 'reproduction_rate', 'icu_patients', 'icu_patients_per_million', 'hosp_patients', 'hosp_patients_per_million', 'weekly_icu_admissions', 'weekly_icu_admissions_per_million', 'weekly_hosp_admissions', 'weekly_hosp_admissions_per_million', 'total_tests', 'new_tests', 'total_tests_per_thousand', 'new_tests_per_thousand', 'new_tests_smoothed', 'new_tests_smoothed_per_thousand', 'positive_rate', 'tests_per_case', 'tests_units', 'total_vaccinations', 'people_vaccinated', 'people_fully_vaccinated', 'total_boosters', 'new_vaccinations', 'new_vaccinations_smoothed', 'total_vaccinations_per_hundred', 'people_vaccinated_per_hundred', 'people_fully_vaccinated_per_hundred', 'total_boosters_per_hundred', 'new_vaccinations_smoothed_per_million', 'new_people_vaccinated_smoothed', 'new_people_vaccinated_smoothed_per_hundred', 'stringency_index', 'population_density', 'median_age', 'aged_65_older', 'aged_70_older', 'gdp_per_capita', 'extreme_poverty', 'cardiovasc_death_rate', 'diabetes_prevalence', 'female_smokers', 'male_smokers', 'handwashing_facilities', 'hospital_beds_per_thousand', 'life_expectancy', 'human_development_index', 'population', 'excess_mortality_cumulative_absolute', 'excess_mortality_cumulative', 'excess_mortality', 'excess_mortality_cumulative_per_million']
```

```
In [3]: # Per calcolare la dimensione ho utilizzato un modulo di os.path utilizzando la

import os
file_size = os.path.getsize(r'./owid_covid_data.csv')
print('File Size:', file_size, 'bytes')
```

File Size: 83189533 bytes

PUNTO 2 - Si chiede poi per ogni continente di trovare il numero di casi totali avvenuti in quello stesso continente; si chiede di non considerare eventuali locazioni che nel dataset non appartengono ad alcun continente.

```
In [4]: import pandas as pd

# Carica il file CSV
dataframe = pd.read_csv('owid_covid_data.csv', low_memory=False)

# Raggruppa i dati per continente e calcola la somma dei casi
casi_per_continente = dataframe.groupby('continent')['total_cases'].sum()
```

```
# Stampa i risultati
print(casi_per_continente)
```

```
continent
Africa      8.812017e+09
Asia        1.196259e+11
Europe      1.260114e+11
North America 7.199848e+10
Oceania     5.110476e+09
South America 4.311459e+10
Name: total_cases, dtype: float64
```

PUNTO 3 - Sempre riguardo i casi di COVID totali, si chiede di sviluppare una funzione che prenda in input il dataset e due nomi di continenti, e che ne confronti i seguenti relativi descrittori statistici: valori minimo e massimo, media, e percentuale rispetto al numero dei casi totali nel mondo (in questo caso calcolati anche sulle locazioni senza indicazione di continente).

```
In [5]: import pandas as pd

def calcola_statistiche(dataset, continenti):
    # Carica il dataset utilizzando pandas
    df = pd.read_csv(dataset, low_memory=False)

    # Filtra i dati per i continenti specificati
    continenti_df = df[df['continent'].isin(continenti)]

    # Calcola le statistiche per i casi totali nel mondo
    casi_totali_mondo = df['total_cases'].sum()
    casi_totali_continente = continenti_df['total_cases'].sum()
    percentuale_casi = (casi_totali_continente / casi_totali_mondo) * 100

    # Calcola il valore minimo, massimo e la media per i casi totali nel continente
    minimo_casi = continenti_df['total_cases'].min()
    massimo_casi = continenti_df['total_cases'].max()
    media_casi = continenti_df['total_cases'].mean()

    # Restituisce i risultati
    return {
        'minimo_casi': minimo_casi,
        'massimo_casi': massimo_casi,
        'media_casi': media_casi,
        'percentuale_casi': percentuale_casi
    }

# Utilizzo
dataset = 'owid_covid_data.csv'
continenti = ['Europe', 'Asia']
risultati = calcola_statistiche(dataset, continenti)

print("Valore minimo dei casi nel continente:", risultati['minimo_casi'])
print("Valore massimo dei casi nel continente:", risultati['massimo_casi'])
print("Media dei casi nel continente:", risultati['media_casi'])
print("Percentuale dei casi nel continente rispetto al mondo:", risultati['percentuale_casi'])
```

Valore minimo dei casi nel continente: 1.0

Valore massimo dei casi nel continente: 99256991.0

Media dei casi nel continente: 2170400.8957641195

Percentuale dei casi nel continente rispetto al mondo: 15.469317917910724

PUNTO 4 - Si chiede poi di effettuare lo stesso tipo di analisi – anche in questo caso sviluppando una funzione ad hoc – per il numero di vaccinazioni totali per ogni continente.

```
In [8]: import pandas as pd

def calcola_statistiche(dataset):
    # Carica il dataset utilizzando pandas
    df = pd.read_csv(dataset, low_memory=False)

    # Calcola il numero totale di casi nel mondo
    casi_totali_mondo = df['total_cases'].sum()

    # Calcola le statistiche per i numeri totali di vaccinazioni per ogni continente
    statistiche_per_continente = {}
    continenti = df['continent'].unique()

    for continente in continenti:
        dati_continente = df[df['continent'] == continente]
        vaccinazioni_totali = dati_continente['total_vaccinations'].sum()

        # Calcola la percentuale di vaccinazioni rispetto ai casi totali nel mondo
        percentuale_vaccinazioni = (vaccinazioni_totali / casi_totali_mondo) * 100

        statistiche_per_continente[continente] = {
            'minimo_vaccinazioni': dati_continente['total_vaccinations'].min(),
            'massimo_vaccinazioni': dati_continente['total_vaccinations'].max(),
            'media_vaccinazioni': dati_continente['total_vaccinations'].mean(),
            'percentuale_vaccinazioni': percentuale_vaccinazioni
        }

    # Restituisce i risultati
    return statistiche_per_continente

# Utilizzo
dataset = 'owid_covid_data.csv'
risultati = calcola_statistiche(dataset)

# Stampa i risultati per ogni continente
for continente, statistiche in risultati.items():
    print("Continente:", continente)
    print("Valore minimo delle vaccinazioni:", statistiche['minimo_vaccinazioni'])
    print("Valore massimo delle vaccinazioni:", statistiche['massimo_vaccinazioni'])
    print("Media delle vaccinazioni:", statistiche['media_vaccinazioni'])
    print("Percentuale delle vaccinazioni rispetto ai casi totali nel mondo:", statistiche['percentuale_vaccinazioni'])
    print()
```

Continent: Asia
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 3491077000.0
Media delle vaccinazioni: 208028270.6204165
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 231.51808739502982

Continent: nan
Valore minimo delle vaccinazioni: nan
Valore massimo delle vaccinazioni: nan
Media delle vaccinazioni: nan
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 0.0

Continent: Europe
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 192221468.0
Media delle vaccinazioni: 26869434.68464302
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 40.97846964326252

Continent: Africa
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 116606863.0
Media delle vaccinazioni: 7871299.604517059
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 3.094686373153005

Continent: Oceania
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 63681652.0
Media delle vaccinazioni: 12457436.752842313
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 1.5870895720299212

Continent: North America
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 676728782.0
Media delle vaccinazioni: 64286967.071168765
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 33.39246303054999

Continent: South America
Valore minimo delle vaccinazioni: 0.0
Valore massimo delle vaccinazioni: 486436436.0
Media delle vaccinazioni: 62241777.59088674
Percentuale delle vaccinazioni rispetto ai casi totali nel mondo: 23.914421347876033

PUNTO 5 - Alla fine, basandosi sui calcoli fatti, il committente chiede di stilare un breve (tre o quattro righe) paragrafo testuale riassuntivo sulle statistiche di casi e vaccinazioni, che si concentri solo sulle differenze esistenti tra Europa, Sud America e Oceania.

ANALISI FINALE - Le statistiche relative ai casi e alle vaccinazioni evidenziano alcune differenze significative tra Europa, Sud America e Oceania. Per quanto riguarda i casi

totali, l'Europa presenta il valore massimo, seguita dal Sud America e infine dall'Oceania, che registra il valore minimo. In termini di vaccinazioni totali, invece, l'Europa mostra un valore massimo, seguita dall'Oceania e dal Sud America. Tuttavia, se consideriamo la percentuale delle vaccinazioni rispetto ai casi totali nel mondo, emerge che l'Oceania ha la percentuale più alta, seguita dall'Europa e infine dal Sud America. Questi dati evidenziano le diverse situazioni e strategie di gestione della pandemia in questi continenti.