Networking Academy
CISCO

This certificate is awarded to

ARJUN ZORE

for successfully completing

Python Essentials 1

offered by MIT Academy of Engineering, alandi Pune
through the Cisco Networking Academy program.

Anuradha Pawar
Instructor
MIT Academy of Engineering, alandi Pune

08 Apr 2025
Completion Date

CISCO Networking Academy

This certificate is awarded to

ARJUN ZORE

for successfully completing

Python Essentials 2

offered by MIT Academy of Engineering, alandi Pune
through the Cisco Networking Academy program.

Anuradha Pawar

Anuradha Pawar
Instructor
MIT Academy of Engineering, alandi Pune

28 Apr 2025
Completion Date

Name: Arjun Kondiba Zore

Div: CS5

Roll No: 46

PRN: 202401100085

```python
# COVID-19 Data Analysis
# This script analyzes COVID-19 data with 20 different analysis questions

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Load the COVID-19 dataset
df = pd.read_csv('/content/country_wise_latest.csv')  # Replace with your
actual file path

# Display basic information about the dataset
print("Dataset Information:")
print(f"Shape: {df.shape}")
print("\nData Overview:")
print(df.head())
print("\nData Types:")
print(df.dtypes)
print("\nSummary Statistics:")
print(df.describe())

# Check for missing values
print("\nMissing Values:")
print(df.isnull().sum())

# Clean data if needed
# Fill missing values with 0 where appropriate
numeric_cols = ['Confirmed', 'Deaths', 'Recovered', 'Active', 'New cases',
                'New deaths', 'New recovered', 'Deaths / 100 Cases',
                'Recovered / 100 Cases', 'Deaths / 100 Recovered',
                'Confirmed last week', '1 week change', '1 week %
increase']
```

```python
df[numeric_cols] = df[numeric_cols].fillna(0)

# Now let's answer 20 different grain questions

# 1. What are the top 10 countries with the highest number of confirmed
cases?
print("\n1. Top 10 countries with highest confirmed cases:")
top_confirmed = df.nlargest(10, 'Confirmed')[['Country/Region',
'Confirmed']]
print(top_confirmed)

# 2. What are the top 10 countries with the highest death rates?
print("\n2. Top 10 countries with highest death rates (Deaths / 100
Cases):")
top_death_rates = df.nlargest(10, 'Deaths / 100 Cases')[['Country/Region',
'Deaths / 100 Cases']]
print(top_death_rates)

# 3. What are the top 10 countries with the highest recovery rates?
print("\n3. Top 10 countries with highest recovery rates (Recovered / 100
Cases):")
top_recovery_rates = df.nlargest(10, 'Recovered / 100
Cases')[['Country/Region', 'Recovered / 100 Cases']]
print(top_recovery_rates)

# 4. Distribution of cases by WHO Region
print("\n4. Distribution of confirmed cases by WHO Region:")
region_cases = df.groupby('WHO
Region')['Confirmed'].sum().sort_values(ascending=False)
print(region_cases)

# 5. Distribution of deaths by WHO Region
print("\n5. Distribution of deaths by WHO Region:")
region_deaths = df.groupby('WHO
Region')['Deaths'].sum().sort_values(ascending=False)
print(region_deaths)

# 6. Calculate and show countries with the highest active case percentage
print("\n6. Countries with highest percentage of active cases:")
df['Active_Percentage'] = (df['Active'] / df['Confirmed']) * 100
top_active_percentage = df.nlargest(10,
'Active_Percentage')[['Country/Region', 'Active_Percentage']]
print(top_active_percentage)

# 7. Weekly growth rate by country (top 10)
```

```python
print("\n7. Top 10 countries by weekly growth rate:")
top_growth = df.nlargest(10, '1 week % increase')[['Country/Region', '1
week % increase']]
print(top_growth)

# 8. Calculate mortality to recovery ratio
print("\n8. Top 10 countries by mortality to recovery ratio (Deaths / 100
Recovered):")
top_mortality_recovery = df.nlargest(10, 'Deaths / 100
Recovered')[['Country/Region', 'Deaths / 100 Recovered']]
print(top_mortality_recovery)

# 9. Identify countries with accelerating case growth
print("\n9. Countries with accelerating case growth (new cases > last week
average):")
df['Last_Week_Daily_Avg'] = df['Confirmed last week'] / 7
df['Acceleration_Factor'] = df['New cases'] / df['Last_Week_Daily_Avg']
accelerating = df[df['Acceleration_Factor'] > 1].nlargest(10,
'Acceleration_Factor')[['Country/Region', 'Acceleration_Factor']]
print(accelerating)

# 10. Calculate case fatality rate vs. recovery rate correlation
print("\n10. Correlation between case fatality rate and recovery rate:")
correlation = df['Deaths / 100 Cases'].corr(df['Recovered / 100 Cases'])
print(f"Correlation coefficient: {correlation}")

# 11. Find regions with highest average death rates
print("\n11. Average death rates by WHO Region:")
region_death_rates = df.groupby('WHO Region')['Deaths / 100
Cases'].mean().sort_values(ascending=False)
print(region_death_rates)

# 12. Calculate countries with highest new cases per existing cases
print("\n12. Countries with highest new cases relative to existing
cases:")
df['New_Case_Ratio'] = df['New cases'] / (df['Confirmed'] - df['New
cases'])
top_new_case_ratio = df.nlargest(10, 'New_Case_Ratio')[['Country/Region',
'New_Case_Ratio']]
print(top_new_case_ratio)

# 13. Distribution of active cases across WHO regions
print("\n13. Distribution of active cases across WHO regions:")
region_active = df.groupby('WHO
Region')['Active'].sum().sort_values(ascending=False)
```

```python
print(region_active)

# 14. Calculate recovery efficiency (recovered cases vs. time)
print("\n14. Top 10 countries by recovery efficiency (new recovered /
active):")
df['Recovery_Efficiency'] = df['New recovered'] / df['Active']
top_recovery_efficiency = df.nlargest(10,
'Recovery_Efficiency')[['Country/Region', 'Recovery_Efficiency']]
print(top_recovery_efficiency)

# 15. Weekly mortality growth rate
print("\n15. Countries with highest weekly death growth:")
df['Weekly_Death_Growth'] = df['New deaths'] * 7 / df['Deaths']
top_death_growth = df.nlargest(10,
'Weekly_Death_Growth')[['Country/Region', 'Weekly_Death_Growth']]
print(top_death_growth)

# 16. Identify countries with declining cases
print("\n16. Countries with declining cases (negative weekly change):")
declining = df[df['1 week change'] < 0].nsmallest(10, '1 week
change')[['Country/Region', '1 week change']]
print(declining)

# 17. Calculate case density (confirmed cases per region)
print("\n17. Case density by WHO Region:")
region_counts = df.groupby('WHO Region').size()
region_case_density = df.groupby('WHO Region')['Confirmed'].sum() /
region_counts
print(region_case_density)

# 18. Calculate deaths per active case
print("\n18. Top 10 countries by deaths per active case:")
df['Deaths_Per_Active'] = df['Deaths'] / df['Active']
top_deaths_per_active = df.nlargest(10,
'Deaths_Per_Active')[['Country/Region', 'Deaths_Per_Active']]
print(top_deaths_per_active)

# 19. Find countries with the lowest recovery rates but high confirmed
cases
print("\n19. Countries with high cases but low recovery rates:")
high_cases_threshold = df['Confirmed'].quantile(0.75)
df_high_cases = df[df['Confirmed'] > high_cases_threshold]
low_recovery = df_high_cases.nsmallest(10, 'Recovered / 100
Cases')[['Country/Region', 'Confirmed', 'Recovered / 100 Cases']]
print(low_recovery)
```

```
# 20. Calculate and show the ratio of new cases to new recovered cases
print("\n20. Countries with highest ratio of new cases to new recovered
cases:")
df['New_Cases_To_Recovered_Ratio'] = df['New cases'] / df['New
recovered'].replace(0, 0.1)
top_new_case_recovered_ratio = df.nlargest(10,
'New_Cases_To_Recovered_Ratio')[['Country/Region',
'New_Cases_To_Recovered_Ratio']]
print(top_new_case_recovered_ratio)


print("\nAnalysis complete.")
```

```
Dataset Information:
Shape: (187, 15)

Data Overview:
   Country/Region  Confirmed  Deaths  Recovered  Active  New cases  New deaths  \
0     Afghanistan      36263    1269      25198    9796        106          10
1         Albania       4880     144       2745    1991        117           6
2         Algeria      27973    1163      18837    7973        616           8
3          Andorra        907      52        803      52         10           0
4           Angola        950      41        242     667         18           1

   New recovered  Deaths / 100 Cases  Recovered / 100 Cases  \
0             18                3.50                  69.49
1             63                2.95                  56.25
2            749                4.16                  67.34
3              0                5.73                  88.53
4              0                4.32                  25.47

   Deaths / 100 Recovered  Confirmed last week  1 week change  \
0                    5.04                35526            737
1                    5.25                 4171            709
2                    6.17                23691           4282
3                    6.48                  884             23
4                   16.94                  749            201
```

```
dtype: int64

1. Top 10 countries with highest confirmed cases:
     Country/Region  Confirmed
173              US    4290259
23           Brazil    2442375
79            India    1480073
138          Russia     816680
154    South Africa     452529
111          Mexico     395489
132            Peru     389717
35            Chile     347923
177  United Kingdom     301708
81             Iran     293606

2. Top 10 countries with highest death rates (Deaths / 100 Cases):
     Country/Region  Deaths / 100 Cases
184           Yemen               28.56
177  United Kingdom               15.19
16          Belgium               14.79
85            Italy               14.26
61           France               13.71
77          Hungary               13.40
120     Netherlands               11.53
111          Mexico               11.13
157           Spain               10.44
183  Western Sahara               10.00
```

```
3. Top 10 countries with highest recovery rates (Recovered / 100 Cases):
     Country/Region  Recovered / 100 Cases
49         Dominica                 100.00
69          Grenada                 100.00
75         Holy See                 100.00
48         Djibouti                  98.38
78          Iceland                  98.33
24           Brunei                  97.87
121     New Zealand                  97.24
136           Qatar                  97.02
105        Malaysia                  96.60
110       Mauritius                  96.51

4. Distribution of confirmed cases by WHO Region:
WHO Region
Americas                8839286
Europe                  3299523
South-East Asia         1835297
Eastern Mediterranean   1490744
Africa                   723207
Western Pacific          292428
Name: Confirmed, dtype: int64
```

```
print("\nAnalysis complete.")
```

5. Distribution of deaths by WHO Region:
WHO Region
Americas   342732
Europe   211144
South-East Asia   41349
Eastern Mediterranean   38339
Africa   12223
Western Pacific   8249
Name: Deaths, dtype: int64

6. Countries with highest percentage of active cases:
         Country/Region   Active_Percentage
168   Timor-Leste   100.000000
117   Mozambique   99.353322
147   Serbia   97.750715
118   Namibia   94.085730
163   Syria   94.065282
161   Sweden   92.820707
32   Canada   92.319978
22   Botswana   91.204330
120   Netherlands   88.113381
76   Honduras   84.386402

7. Top 10 countries by weekly growth rate:
          Country/Region   1 week % increase
130   Papua New Guinea   226.32
63   Gambia   191.07
11   Bahamas   119.54
186   Zimbabwe   57.85
99   Libya   42.78
58   Ethiopia   42.52
22   Botswana   41.57
97   Lesotho   40.67
160   Suriname   37.44
41   Costa Rica   37.34

8. Top 10 countries by mortality to recovery ratio (Deaths / 100 Recovered):
       Country/Region   Deaths / 100 Recovered
32   Canada   inf
117   Mozambique   inf
147   Serbia   inf
161   Sweden   inf
163   Syria   inf
120   Netherlands   3259.26
177   United Kingdom   3190.26
184   Yemen   57.98
16   Belgium   56.28
61   France   37.20

9. Countries with accelerating case growth (new cases > last week average):

```
9. Countries with accelerating case growth (new cases > last week average):
    Country/Region   Acceleration_Factor
63          Gambia              3.062500
11         Bahamas              1.609195

10. Correlation between case fatality rate and recovery rate:
Correlation coefficient: -0.16891981705250228

11. Average death rates by WHO Region:
WHO Region
Europe                  4.198393
Eastern Mediterranean   3.563182
Americas                3.052571
Africa                  2.306458
South-East Asia         1.296000
Western Pacific         1.290000
Name: Deaths / 100 Cases, dtype: float64
```

```
12. Countries with highest new cases relative to existing cases:
             Country/Region   New_Case_Ratio
63                   Gambia         0.176895
11                  Bahamas         0.116959
22                 Botswana         0.077259
68                Greenland         0.076923
186                Zimbabwe         0.076433
21    Bosnia and Herzegovina        0.074844
91                   Kosovo         0.071707
37                 Colombia         0.067717
99                    Libya         0.059198
39       Congo (Brazzaville)        0.053325

13. Distribution of active cases across WHO regions:
WHO Region
Americas                4027938
Europe                  1094656
South-East Asia          637015
Africa                   270339
Eastern Mediterranean    251005
Western Pacific           77409
Name: Active, dtype: int64
```

```
14. Top 10 countries by recovery efficiency (new recovered / active):
              Country/Region  Recovery_Efficiency
48                  Djibouti             0.458333
144   Sao Tome and Principe             0.324786
185                   Zambia             0.291171
28                   Burundi             0.289474
5        Antigua and Barbuda             0.277778
152                 Slovenia             0.231092
103               Madagascar             0.203953
111                   Mexico             0.180204
109               Mauritania             0.159400
47                   Denmark             0.141805


15. Countries with highest weekly death growth:
              Country/Region  Weekly_Death_Growth
22                  Botswana             3.500000
63                    Gambia             1.750000
41                Costa Rica             0.669565
91                    Kosovo             0.605405
103               Madagascar             0.461538
99                     Libya             0.437500
119                    Nepal             0.437500
57                  Eswatini             0.411765
37                  Colombia             0.405150
39     Congo (Brazzaville)             0.388889
```

```
16. Countries with declining cases (negative weekly change):
    Country/Region  1 week change
88           Jordan              -47

17. Case density by WHO Region:
WHO Region
Africa                    15066.812500
Americas                 252551.028571
Eastern Mediterranean     67761.090909
Europe                    58920.053571
South-East Asia          183529.700000
Western Pacific           18276.750000
dtype: float64

18. Top 10 countries by deaths per active case:
    Country/Region  Deaths_Per_Active
24          Brunei                inf
143      San Marino                inf
110       Mauritius           5.000000
85           Italy           2.790875
48         Djibouti           2.416667
83          Ireland           2.308901
60          Finland           2.208054
126          Norway           2.040000
34             Chad           2.027027
123           Niger           1.916667
```

```
19. Countries with high cases but low recovery rates:
      Country/Region  Confirmed  Recovered / 100 Cases
32           Canada     116458                   0.00
161          Sweden      79395                   0.00
120     Netherlands      53413                   0.35
177  United Kingdom     301708                   0.48
16          Belgium      66428                  26.27
20          Bolivia      71181                  30.17
173              US    4290259                  30.90
133     Philippines      82040                  32.24
61           France     220352                  36.86
52            Egypt      92482                  37.67

20. Countries with highest ratio of new cases to new recovered cases:
         Country/Region  New_Cases_To_Recovered_Ratio
32              Canada                        6820.0
51             Ecuador                        4670.0
120        Netherlands                        4190.0
147             Serbia                        4110.0
31            Cameroon                        4020.0
161             Sweden                        3980.0
46             Czechia                        1920.0
182  West Bank and Gaza                       1520.0
156        South Sudan                         430.0
11             Bahamas                         400.0

Analysis complete.
```

https://youtu.be/swUnjzTHRlE?si=egiMp3EvfbFAAfcv