Lack of Access to Healthcare in Underserved Communities

PROJECT CODE:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
from scipy.stats import linregress
np.random.seed(42)
num_regions = 20  # Increased the number of regions for
broader analysis

# Generate random data
regions = [f'Region {i+1}' for i in range(num_regions)]
population = np.random.randint(5000, 60000, num_regions)
healthcare_facilities = np.random.randint(1, 25,
num_regions)

# Additional metrics
population_density = np.random.randint(50, 1200,
num_regions)  # People per square km
avg_distance_to_clinic = np.random.randint(5, 150,
num_regions)  # in km
avg_waiting_time = np.random.randint(10, 240,
num_regions)  # Waiting time in minutes
healthcare_quality_score = np.random.randint(1, 10,
num_regions)  # Score out of 10
medical_staff_count = np.random.randint(10, 500,
num_regions)  # Number of staff
budget_allocation = np.random.randint(1_000_000,
50_000_000, num_regions)  # Budget in USD
```

```
# Calculating people per facility and staff per 1000
population
people_per_facility = population / healthcare_facilities
staff_per_1000 = (medical_staff_count / population) * 1000
budget_per_capita = budget_allocation / population
data = pd.DataFrame({
    'Region': regions,
    'Population': population,
    'Facilities': healthcare_facilities,
    'People per Facility': people_per_facility,
    'Population Density': population_density,
    'Avg Distance to Clinic (km)': avg_distance_to_clinic,
    'Avg Waiting Time (min)': avg_waiting_time,
    'Healthcare Quality Score': healthcare_quality_score,
    'Medical Staff Count': medical_staff_count,
    'Staff per 1000 Population': staff_per_1000,
    'Budget Allocation (USD)': budget_allocation,
    'Budget per Capita (USD)': budget_per_capita
})
corr_matrix = data.corr(numeric_only=True)
print("\n📊 Correlation Matrix:")
print(corr_matrix)
sns.set(style="whitegrid")
fig, ax = plt.subplots(4, 1, figsize=(16, 24))

# 📊 Plot 1: People per Facility by Region
barplot1 = sns.barplot(
    x="People per Facility",
    y="Region",
    data=data.sort_values(by='People per Facility',
ascending=False),
    palette="coolwarm",
    ax=ax[0]
)

for index, value in enumerate(data.sort_values(by='People per Facility', ascending=False)['People per Facility']):
    barplot1.text(value + 500, index, f'{value:.2f}', color='black', va="center", fontsize=10)

ax[0].set_title("Healthcare Access: People per Facility", fontsize=16, fontweight='bold')
ax[0].set_xlabel("People per Healthcare Facility", fontsize=12)
ax[0].set_ylabel("Region", fontsize=12)

# 📊 Plot 2: Staff per 1000 Population vs. Healthcare Quality
scatter1 = sns.scatterplot(
    x="Staff per 1000 Population",
    y="Healthcare Quality Score",
    hue="People per Facility",
    size="Population",
    data=data,
    palette="viridis",
    sizes=(100, 1000),
    ax=ax[1]
)

ax[1].set_title("Staff per 1000 Population vs. Healthcare Quality", fontsize=16, fontweight='bold')
ax[1].set_xlabel("Staff per 1000 People", fontsize=12)
ax[1].set_ylabel("Healthcare Quality Score", fontsize=12)

# 📊 Plot 3: Distance to Clinic vs. Budget per Capita
```

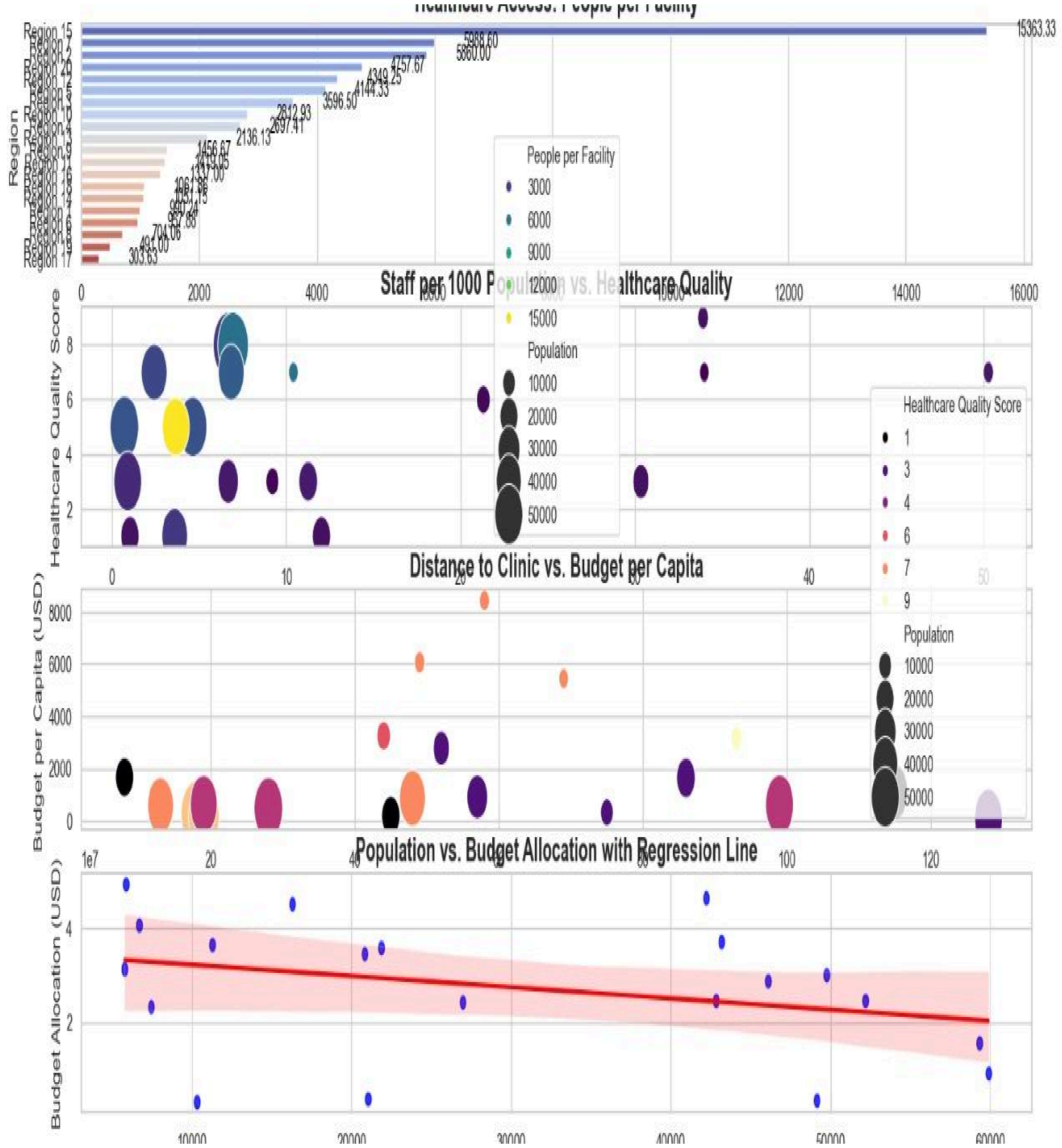
```
scatter2 = sns.scatterplot(
    x="Avg Distance to Clinic (km)",
    y="Budget per Capita (USD)",
    hue="Healthcare Quality Score",
    size="Population",
    data=data,
    palette="magma",
    sizes=(100, 1000),
    ax=ax[2]
)

ax[2].set_title("Distance to Clinic vs. Budget per Capita", fontsize=16, fontweight='bold')
ax[2].set_xlabel("Average Distance to Clinic (km)", fontsize=12)
ax[2].set_ylabel("Budget per Capita (USD)", fontsize=12)

# 📊 Plot 4: Budget Allocation vs. Population with Linear Regression
sns.regplot(
    x="Population",
    y="Budget Allocation (USD)",
    data=data,
    scatter_kws={"color": "blue"},
    line_kws={"color": "red"},
    ax=ax[3]
)
ax[3].set_title("Population vs. Budget Allocation with Regression Line", fontsize=16, fontweight='bold')
ax[3].set_xlabel("Population", fontsize=12)
ax[3].set_ylabel("Budget Allocation (USD)", fontsize=12)
plt.tight_layout()
plt.show()
```

```
print("\n📝 Full Dataset with Additional Metrics:")
print(data.sort_values(by='People per Facility',
ascending=False))
```

OUTPUT:


Healthcare Access: People per Facility
Region
Region 15
Region 7
Region 2
Region 20
Region 12
Region 5
Region 3
Region 10
Region 4
Region 13
Region 9
Region 17
People per Facility
Population
Staff per 1000 Population vs. Healthcare Quality
Healthcare Quality Score
Healthcare Quality Score
Population
Distance to Clinic vs. Budget per Capita
Budget per Capita (USD)
Population vs. Budget Allocation with Regression Line
Budget Allocation (USD)
1e7