```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Create sample data
data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Emma'],
    'Age': [20, 22, None, 19, 20],  # Introduce a missing value
    'Grade': [85, 78, 90, 92, 88],
    'Favorite_Subject': ['Math', 'English', 'Science', 'History',
'Art']
}

# Create DataFrame
df = pd.DataFrame(data)

# Display the sample dataset
print("Sample Dataset:")
print(df)

# Data Cleaning

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

# Drop rows with missing values
df_cleaned = df.dropna()

# Remove duplicates if any
df_cleaned = df_cleaned.drop_duplicates()

# Display the cleaned dataset
print("\nCleaned Dataset:")
print(df_cleaned)

# Data Visualization

# Plotting age distribution
plt.hist(df_cleaned['Age'].values, bins=5, color='skyblue',
edgecolor='black')
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.grid(True)
plt.show()
```

```
Sample Dataset:
      Name   Age   Grade Favorite_Subject
0    Alice  20.0      85             Math
1      Bob  22.0      78          English
2  Charlie   NaN      90          Science
3    David  19.0      92          History
4     Emma  20.0      88              Art

Missing Values:
Name                0
Age                 1
Grade               0
Favorite_Subject    0
dtype: int64

Cleaned Dataset:
     Name   Age   Grade Favorite_Subject
0   Alice  20.0      85             Math
1     Bob  22.0      78          English
3   David  19.0      92          History
4    Emma  20.0      88              Art
```

Age Distribution