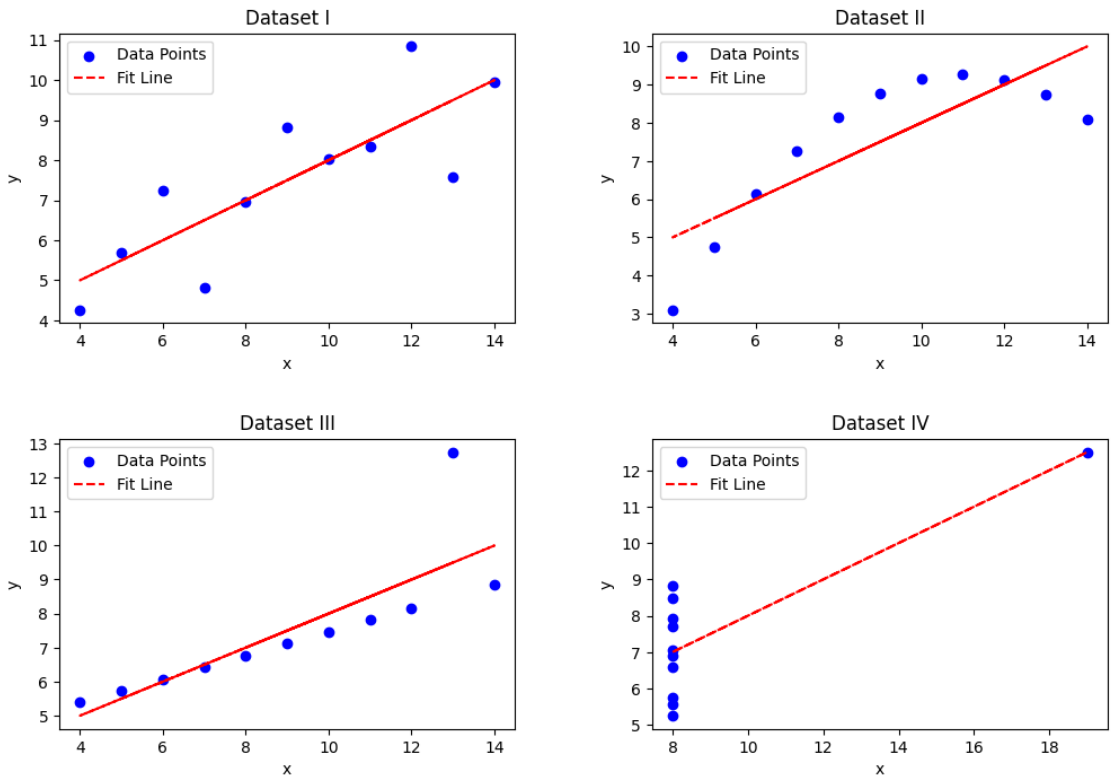
Data Visualization V2

October 1, 2025

[3]:
```
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np

# Load Anscombe's quartet directly from seaborn
df = sns.load_dataset("anscombe")

# Create 4 subplots
fig, axs = plt.subplots(2, 2, figsize=(12, 8))
titles = ["Dataset I", "Dataset II", "Dataset III", "Dataset IV"]

for i, dataset in enumerate(df["dataset"].unique()):
    data = df[df["dataset"] == dataset]
    ax = axs[i//2, i%2]

    # Scatter plot
    ax.scatter(data["x"], data["y"], color="blue", label="Data Points")

    # Best-fit regression line
    m, b = np.polyfit(data["x"], data["y"], 1)
    ax.plot(data["x"], m*data["x"] + b, color="red", linestyle="--", label="Fit␣
 ↪Line")

    ax.set_title(titles[i])
    ax.set_xlabel("x")
    ax.set_ylabel("y")
    ax.legend()

# Clean layout + main title
fig.suptitle("Anscombe's Quartet", fontsize=16, fontweight="bold")
plt.subplots_adjust(hspace=0.4, wspace=0.3)
plt.show()
```


Anscombe's Quartet
Dataset I
Dataset II
Dataset III
Dataset IV
Data Points
Fit Line
x
y

[ ]: