import numpy as np

import matplotlib.pyplot as plt

from sklearn.datasets import fetch\_california\_housing

from sklearn.linear\_model import LinearRegression

# Load data

data = fetch\_california\_housing()

X = data.data[:, [0]] # Use only 'MedInc' feature(column 0)

y = data.target

# Train linear regression model

model = LinearRegression()

model.fit(X, y)

# Predict line

x\_range = np.linspace(X.min(), X.max(),100).reshape(-1, 1)

y\_pred\_line = model.predict(x\_range)

# Plot data and regression line

plt.figure(figsize=(8, 5))

plt.scatter(X, y, alpha=0.3, label="Data",color='skyblue')

plt.plot(x\_range, y\_pred\_line, color='red',label="Best Fit Line")

plt.title("Linear Regression: Median Income vs House Value")

plt.xlabel("Median Income")

plt.ylabel("House Value (in $100,000s)")

plt.legend()

plt.grid(True)

plt.show()