Kode program yang akan diajarkan di sini bisa diunduh di tautan [berikut ini](https://colab.research.google.com/drive/1ABfaCT_uth1mGqGKkmHfapi3WDCvvLcn?usp=sharing). Silahkan buat salinan dari kode ini dan jalankan di Google Colab Anda.

**Tujuan**

Pada latihan kali ini kita akan memprediksi harga rumah berdasarkan jumlah kamar.

**Tahapan Latihan**

Berikut adalah tahapan latihan yang akan Anda lakukan:

1. Impor library yang dibutuhkan.
2. Buat dataset *dummy*dengan Numpy Array.
3. Buat plot dari model.

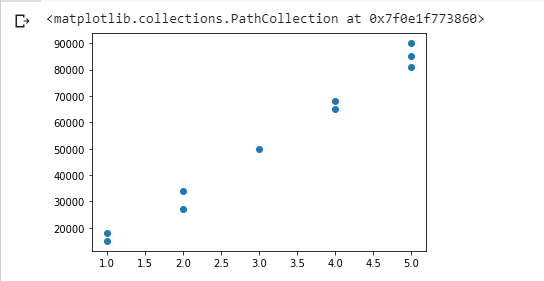
**Codelab**

Pertama kita mengimpor*library*yang diperlukan. Lalu buat data *dummy*menggunakan numpy array.

1. import numpy as np
3. #buat data jumlah kamar
4. bedrooms = np.array([1,1,2,2,3,4,4,5,5,5])
6. #data harga rumah. asumsi dalam dollar
7. house\_price = np.array([15000, 18000, 27000, 34000, 50000, 68000, 65000, 81000,85000, 90000])

Selanjutnya, kita bisa mencoba menampilkan data tersebut dalam bentuk *scatter plot*. Jumlah kamar pada sumbu X adalah variabel independen dan harga rumah pada sumbu Y  adalah variabel dependen.

1. # menampilkan scatter plot dari dataset
2. import matplotlib.pyplot as plt
3. %matplotlib inline
5. plt.scatter(bedrooms, house\_price)

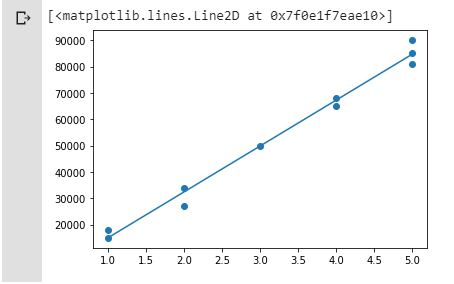
Tampilan dari kode tersebut sebagai berikut.  


Lalu pada *cell* berikutnya, kita bisa mulai melatih model kita dengan memanggil fungsi LinearRegression.fit() pada data kita. Fungsi ini untuk melatih model regresi linier dari *library*SKLearn.

1. from sklearn.linear\_model import LinearRegression
3. # latih model dengan Linear Regression.fit()
4. bedrooms = bedrooms.reshape(-1, 1)
5. linreg = LinearRegression()
6. linreg.fit(bedrooms, house\_price)

Terakhir kita bisa melihat bagaimana model kita menyesuaikan dengan data yang kita miliki dengan membuat plot dari model kita.

1. # menampilkan plot hubungan antara jumlah kamar dengan harga rumah
2. plt.scatter(bedrooms, house\_price)
3. plt.plot(bedrooms, linreg.predict(bedrooms))

Hasilnya sebagai seperti di bawah ini:  


Model regresi linier adalah salah satu model *machine learning* yang paling sederhana. Model ini memiliki kompleksitas rendah dan bekerja sangat baik pada dataset yang memiliki hubungan linier. Jadi, ketika Anda menemui masalah yang terlihat memiliki hubungan linier, regresi linier dapat menjadi pilihan pertama sebagai model untuk dikembangkan.