


Source:

<https://www.planning.vic.gov.au/guides-and-resources/Data-spatial-and-insights/victorian-population-data>


```
1 # Import libraries
2 import numpy as np
3 import matplotlib.pyplot as plt
4 from sklearn.linear_model import LinearRegression

1 # Historical data: Year (X) and Population (Y)
2 # You can replace these arrays with your actual data
3 years = np.array([2017, 2018, 2019, 2020, 2021, 2022, 2023]).reshape(-1, 1) # Independent variable (X)
4 population = np.array([6358210, 6479695, 6590050, 6567195, 6566092, 6718947, 6865358]) # Dependent variable (Y)
```

```
1 # Training
2 # Create a Linear Regression model
3 model = LinearRegression()
4 # Fit the model to the data
5 model.fit(years, population)
```

 `LinearRegression()`

```
1 # Predicting
2 # Predict future populations (e.g., 2025, 2030)
3 future_years = np.array([2025, 2026, 2027, 2028, 2029, 2030]).reshape(-1, 1)
4 predicted_population = model.predict(future_years)
5
6 # Output predictions
7 for year, pop in zip(future_years, predicted_population):
8     print(f"Predicted population in {year[0]}: {pop:.2f}")
```

 Predicted population in 2025: 6945076.36
 Predicted population in 2026: 7015647.43
 Predicted population in 2027: 7086218.50
 Predicted population in 2028: 7156789.57
 Predicted population in 2029: 7227360.64
 Predicted population in 2030: 7297931.71

```
1 # Visualize the data and regression line
2 plt.scatter(years, population, color='blue', label='Historical Data')
3 plt.plot(years, model.predict(years), color='red', label='Linear Regression Line')
4 plt.scatter(future_years, predicted_population, color='green', label='Predicted Data')
5
6 plt.xlabel('Year')
7 plt.ylabel('Population')
8 plt.title('Victoria Population Prediction using Linear Regression')
9 plt.legend()
10 plt.grid(True)
11 plt.show()
```

