



SUT/Faculty of  
Industrial  
Engineering

# Python Project

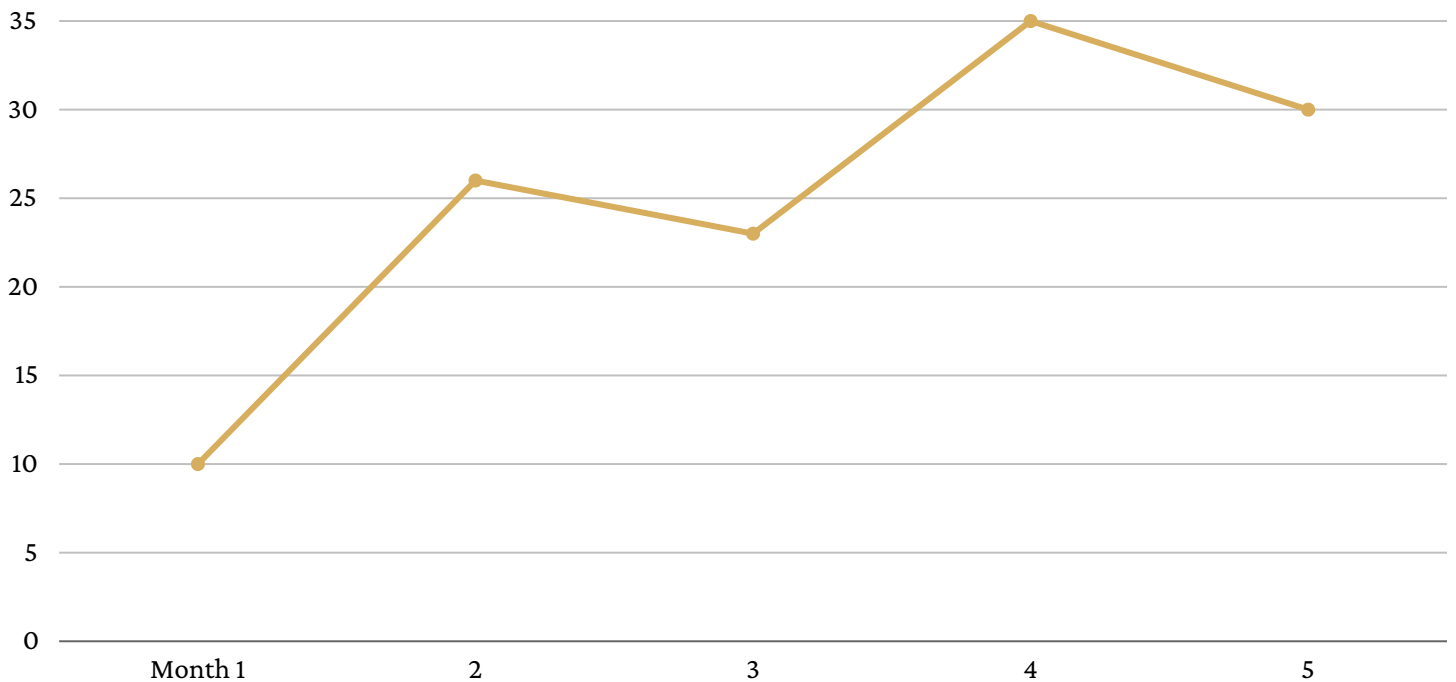
Date:07/06/2024

Spring 2024

## Project report

*Instructor : Dr.Rezapour*

Leili Motahari(99171214)



## 1) Cleaning Data :

In this section, we want to replace a null data with an appropriate value. In this section, the alternative method is to use the average of the available data.

In this section, we want to replace a null data with an appropriate value. In this section, the alternative method is to use the average of the available data

First, read the file and then use the apply method to fill the Nan values with the average value on the first column. The method used will not consider Nan values in the average calculation, and as a result, we do not have an error in this section.

## 2) Staff Salaries

In this section, we plan to predict employee salaries until 2050 by implementing a linear regression. For this, we open the new file and read its data.

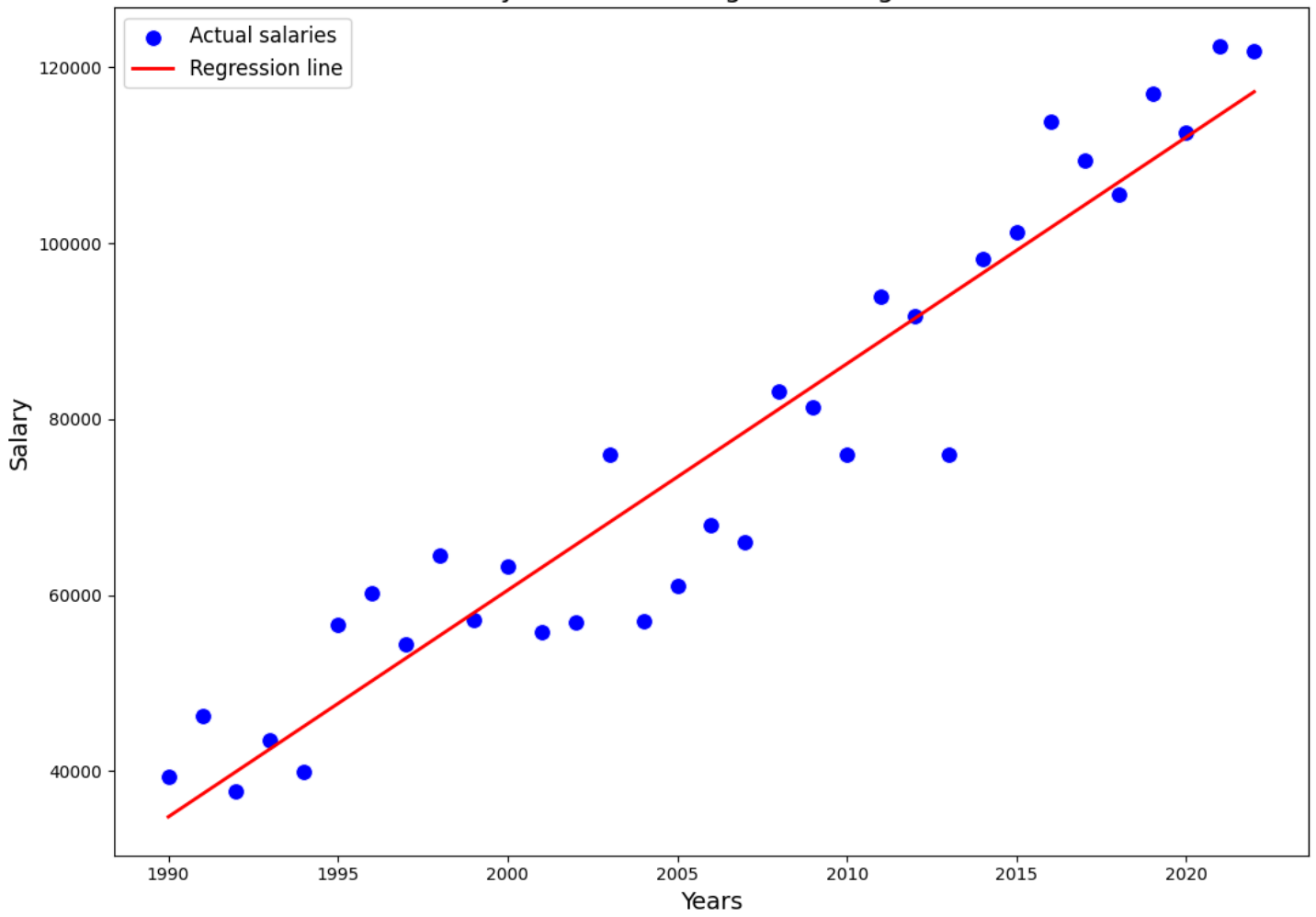
## 3) Regression Plot on Salary

In this section, we intend to draw a plot of the applied regression and the salary data given in the initial file and measure the accuracy of the model.

Three criteria are used to measure the accuracy of the model and the result is written based on the first criterion.

The  $R^2$  score indicates how well your model is performing; consider the placement dataset. If the  $R^2$  score is one, then the regression line is perfect with no error. In this case, it's close to 1, so based on the  $R^2$  score, we can claim that the model performs well. The MSE shows that we have a little bit of data that is widely dispersed around its central moment (mean).

Salary Prediction using Linear Regression



<b>bias</b>	2.8221896200469047e-11
<b>Mean Squared Error</b>	58816116.22780049
<b>R_squared</b>	0.910946

The  $R^2$  score indicates how well your model is performing; consider the placement dataset. If the  $R^2$  score is one, then the regression line is perfect with no error. In this case, it's close to 1, so based on the  $R^2$  score, we can claim that the model performs well. The MSE shows that we have a little bit of data that is widely dispersed around its central moment (mean).

#### 4) Material 1 Price Prediction

In this section, we want to predict the price of raw material 1, which is imported, with 1% inflation in the coming years. For ease of work, the price of this material in the coming years is calculated with the factor  $(F/P, 5\% + 1\%, n)$ .

#### 5) Material Worth Prediction

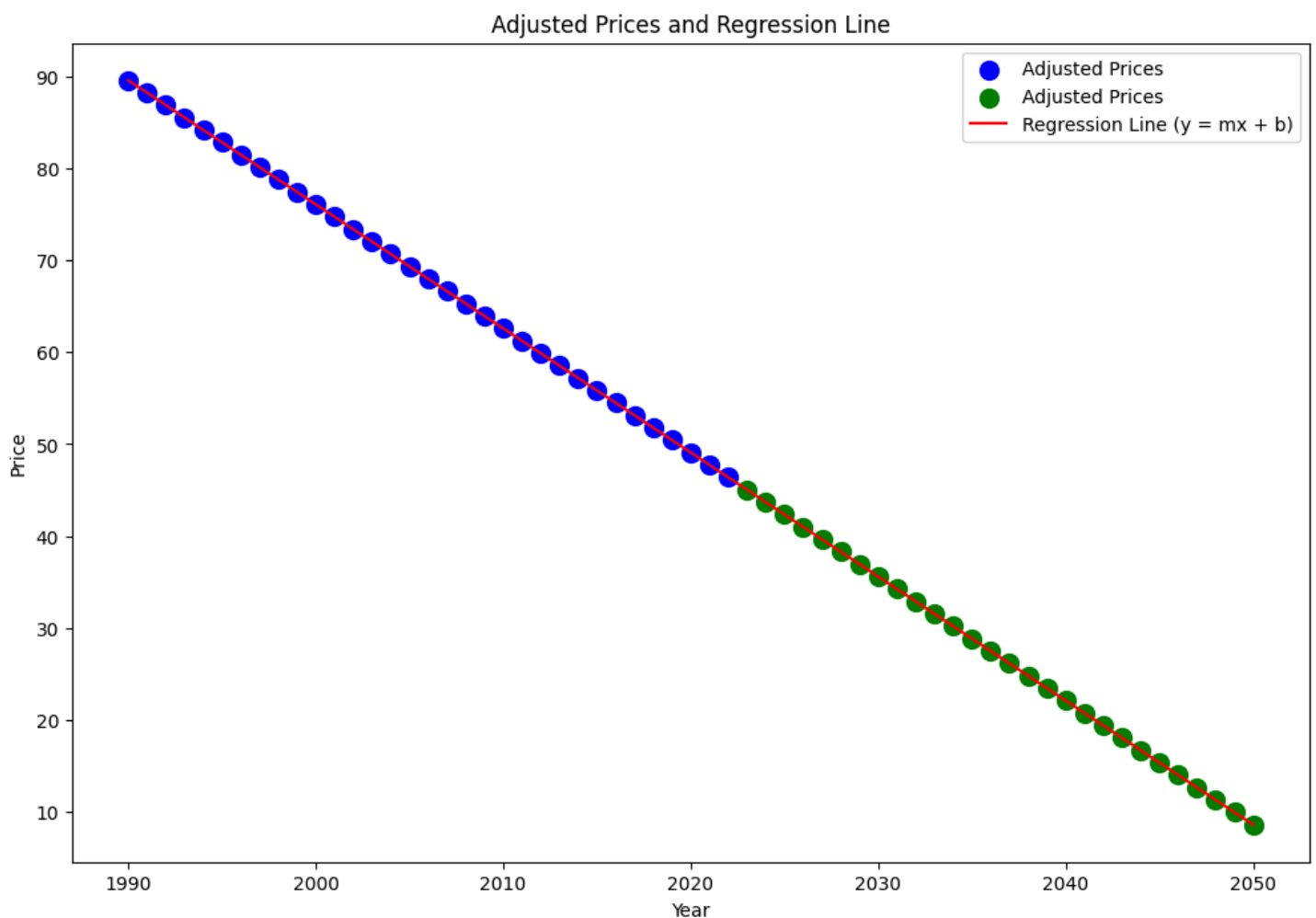
The demand for raw material 2 is decreasing, which causes its value to decrease. In this section, to evaluate the value of this material, we convert its price to the base year and apply a regression to the new data. The demand for raw material 2 is decreasing, which causes its value to decrease. In this section, to evaluate the value of this material, we convert its price to the base year and apply a regression to the new data. To convert the price to the price in the base year, the factor  $(P/F, 5\%, n)$  is used

## 6) Regression on Material2 new prices

According to the previous explanation,  $R^2\_Score$  is equal to one, which indicates that there is no error in the regression line.

The value of bias and MSE is also very small and shows the high accuracy of the model.

This result is not far-fetched because using the data pre-pinned by this model, data related to 2023 and later have been written.



<b>bias</b>	1.93361e-14
<b>Mean Squared Error</b>	2.02925e-26
<b>R_squared</b>	1

According to the previous explanation, R2\_Score is equal to one, which indicates that there is no error in the regression line.

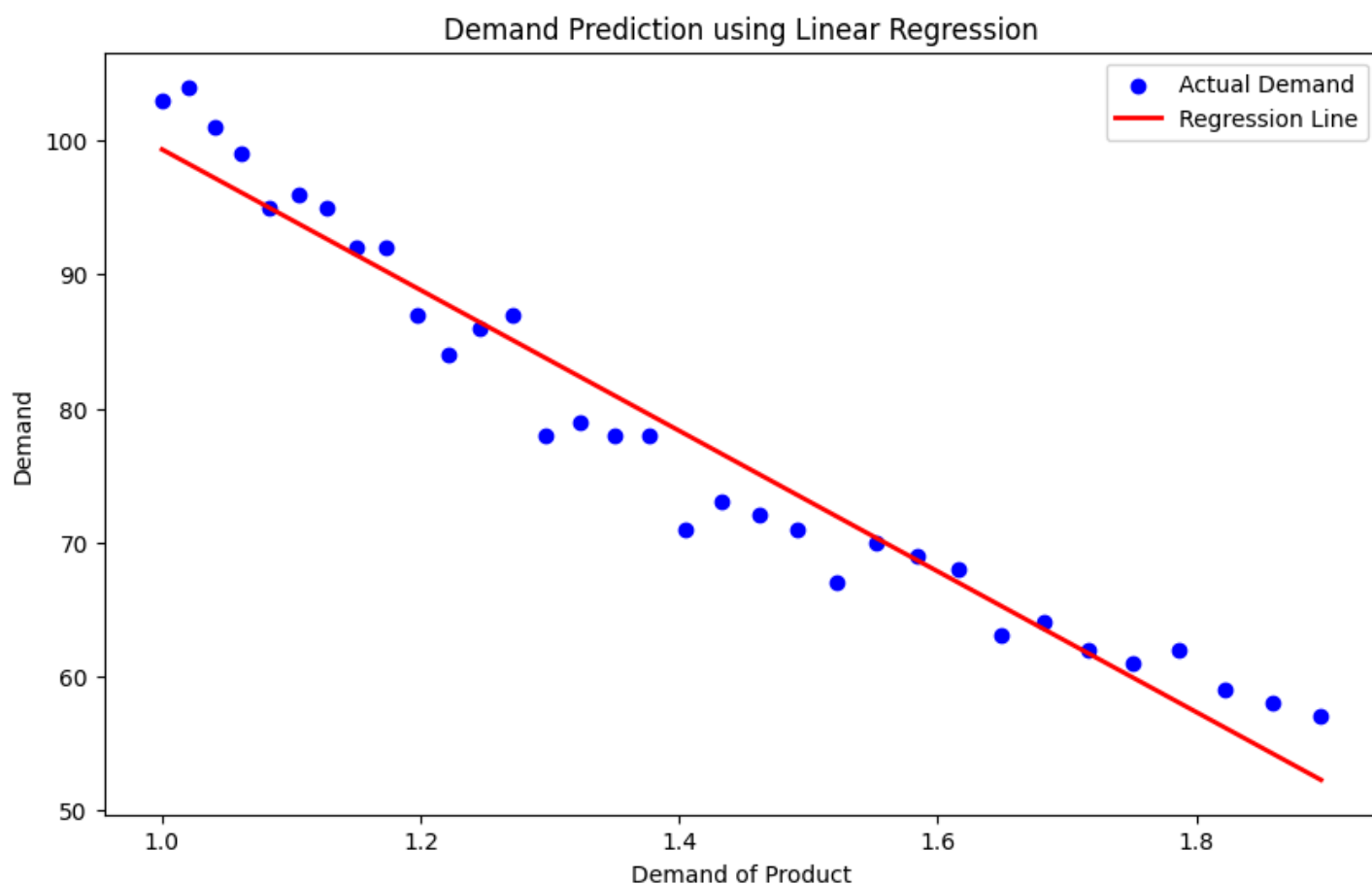
The value of bias and mSE is also very small and shows the high accuracy of the model.

This result is not far-fetched because using the data pre-pinned by this model, data related to 2023 and later have been written.

## 8) Regression on demand

In this section, we intend to use regression to find the relationship between demand and Then, using the obtained model, predict the amount of demand until 2050.

## 9) Regression plot for demand



<b>bias</b>	4.306319610667274e-15
<b>Mean Squared Error</b>	10.02409927424594
<b>R_squared</b>	0.951178

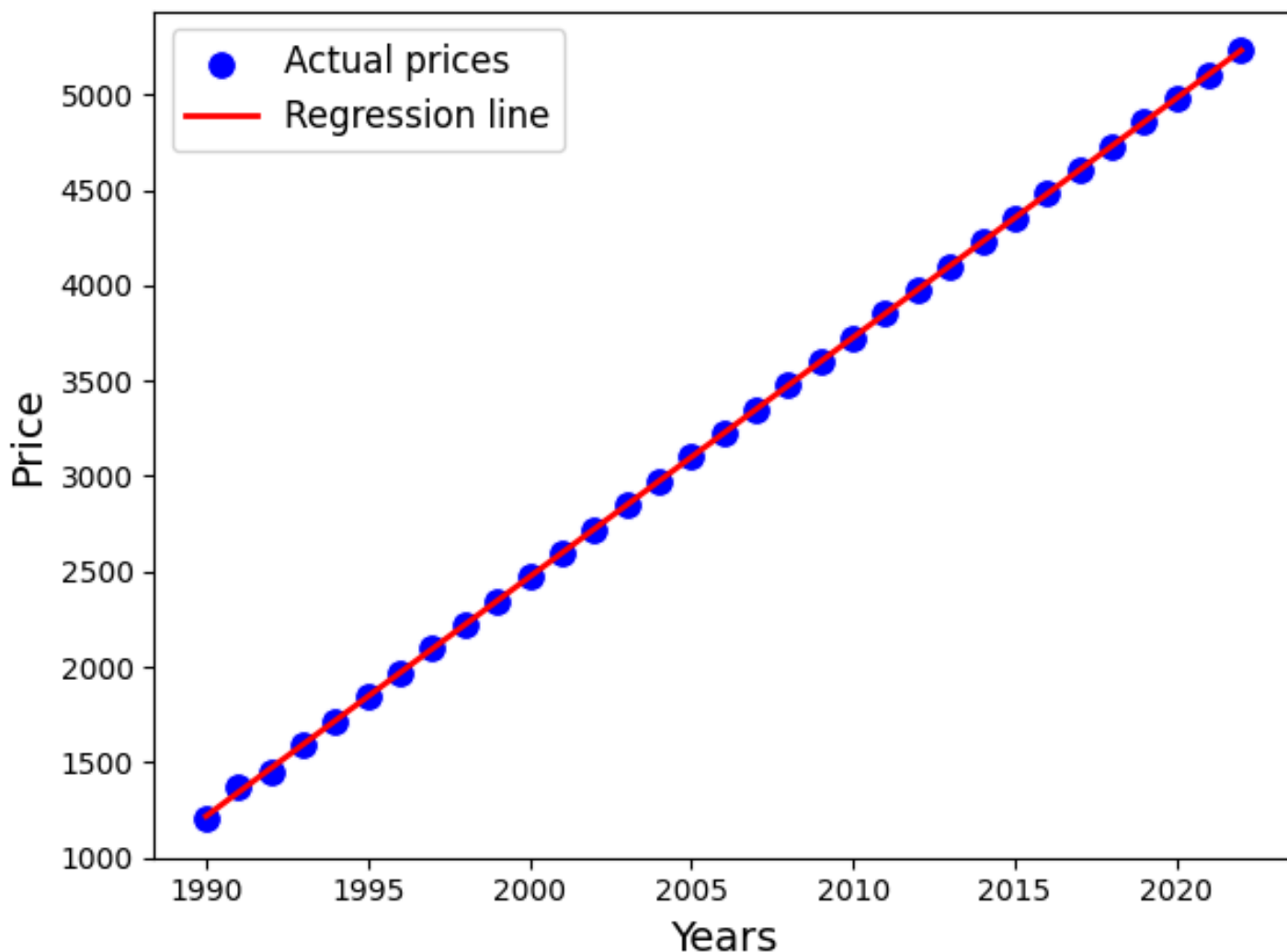
R2\_Score is close to one, which indicates that the model fits the data with good accuracy.

The amount of bias is also very small and shows the high accuracy of the model.

A larger value of MSE indicates the dispersion of demand

The demand for this product is decreasing and it has a range between 57 and 103. The decreasing trend occurs at a relatively good speed, but this criterion and dispersion occurred due to the decrease in demand.

## 11) Regression plot of final product price



<b>bias</b>	0
<b>Mean Squared Error</b>	49.5
<b>R_squared</b>	0.999965

Low error rate and R2\_score shows that the model is well fitted.