**BİM453 Introduction to Machine Learning Term Project**

**Final Report**

Betülnaz Hayran-28354853660

İshak Çoban-12422490116

**Introduction**

Our Project is about weather forecasting. Our system will use the previous air temperatures as an input. Then we will select the day we want to get the predicted weather of and by using linear regression we will get the predicted weather of the chosen day. After getting the predicted weather, our project will give a recommendation about clothes based on the predicted weather of the chosen day.

**Method**

Our methods are linear regression and recommendation system.

Simple Linear regression:

We used the weather dataset, in this dataset we have used the day and mean temperature columns for input. After getting the day we chose as input, the output would be the chosen day’s estimated temperature. We set the x value to ‘day’ array, and the y value to ‘temperature’ array. After doing linear regression math, we found the b0 and b1 in the formula of the regression line which is Y=b0+b1\*X. After founding b0 and b1 and choose the x value (the day we want it’s estimated temperature) we find the y value (chosen day’s estimated temperature). The program shows a graphic of the regression line.

Recommendation system:

Before all about things, we divide and classify all temperatures.

**6 °C**(smaller than six celcius) **cold**

**6 °C** - **15 °C** **warm**

**15 °C** - **25 °C hot**

**25 °C**(bigger than twenty-five celcius) **too hot**

We determine 5 clothes that will be preferred by users and convert these clothes to numbers that are between 1 to 5.

**1 jeans**

**2 socks**

**3 dress**

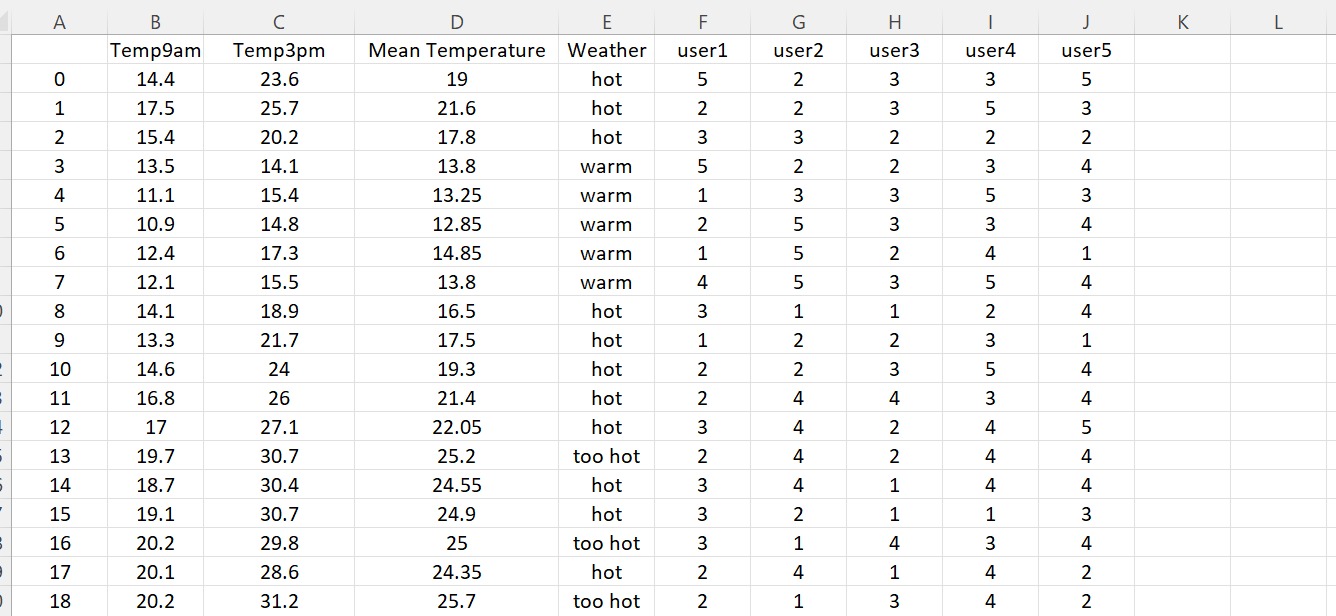
**4 skirt**

**5 t-shirt**

We need first method’s result. Then this result is classified according to the above information(cold, warm etc.). After weather condition was determined, this method will control users’ preferences about clothes in the latest 10 same conditions. Cloth/clothes which the most preferred will presented to user.

**Experimental Work**

Here’s a photo of the part of our dataset:



We use D column(Mean Temperature) for regression. We find these values by getting mean of B(Temp9am) and C(Temp3pm) columns. After that, we classify mean temperatures and write on E column(Weather).

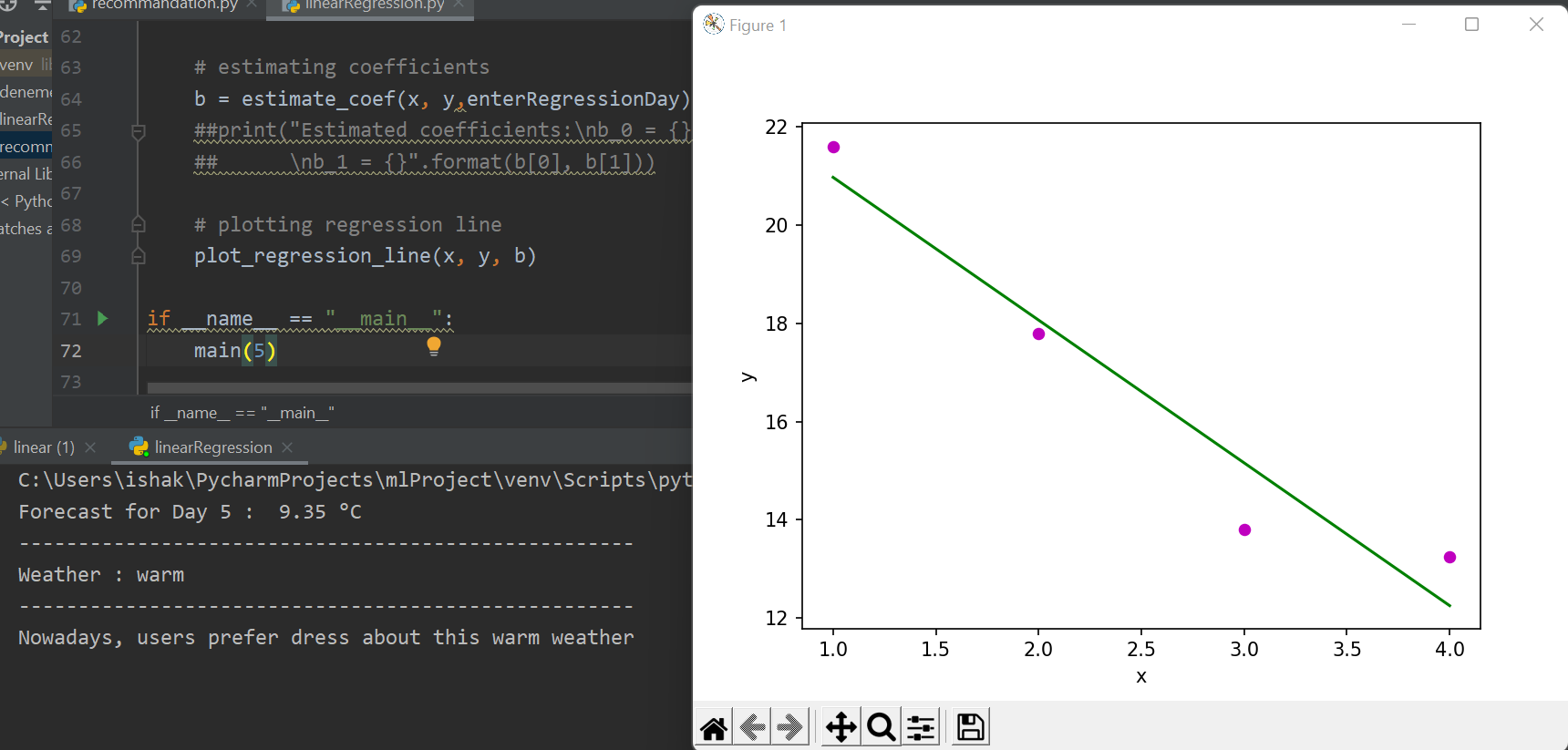
Our dataset has 5 users and these users have cloth choices about that day.

If we use linear regression;

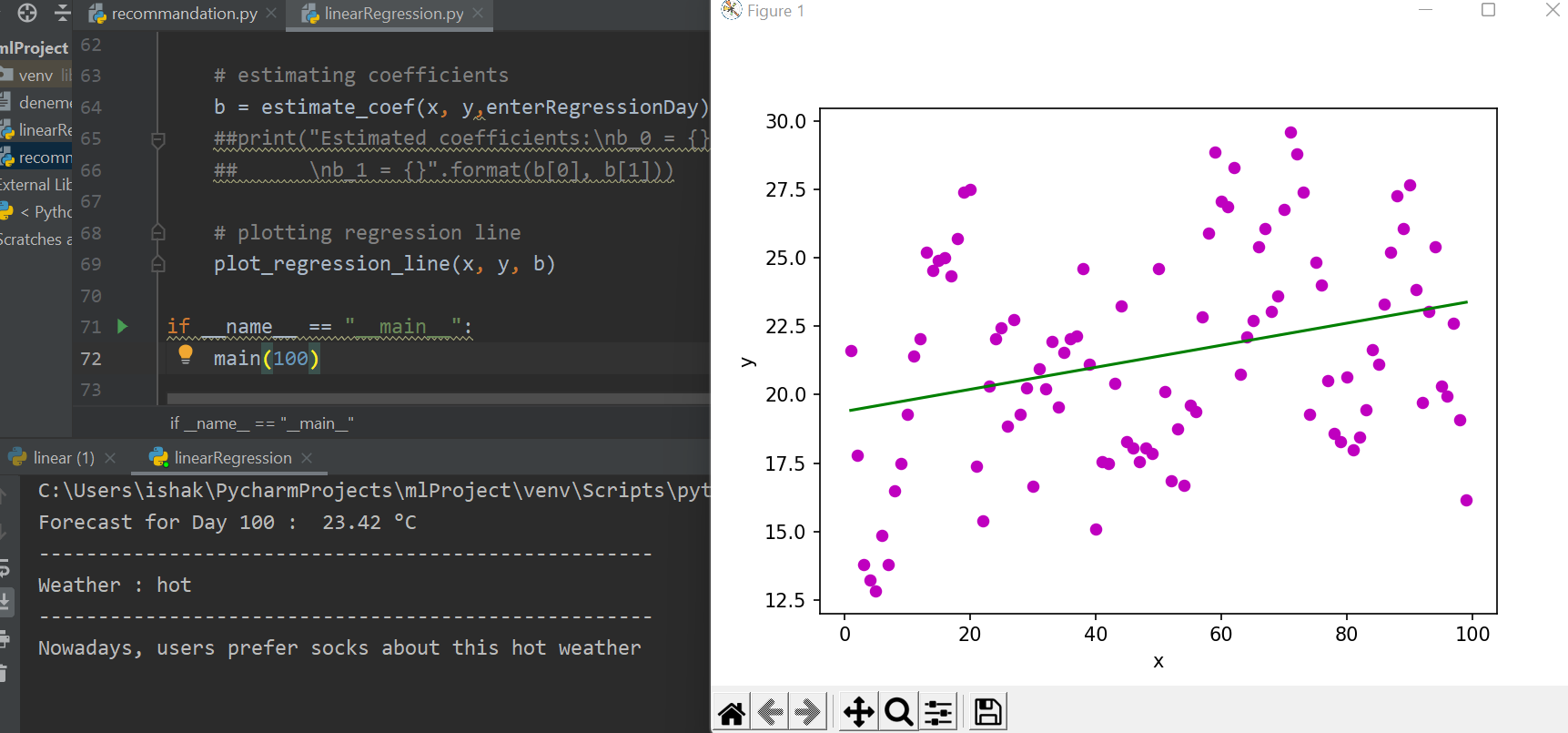
Here’s an example of the output of our code when the x is 5.

Here x is the day that will be to make regression.

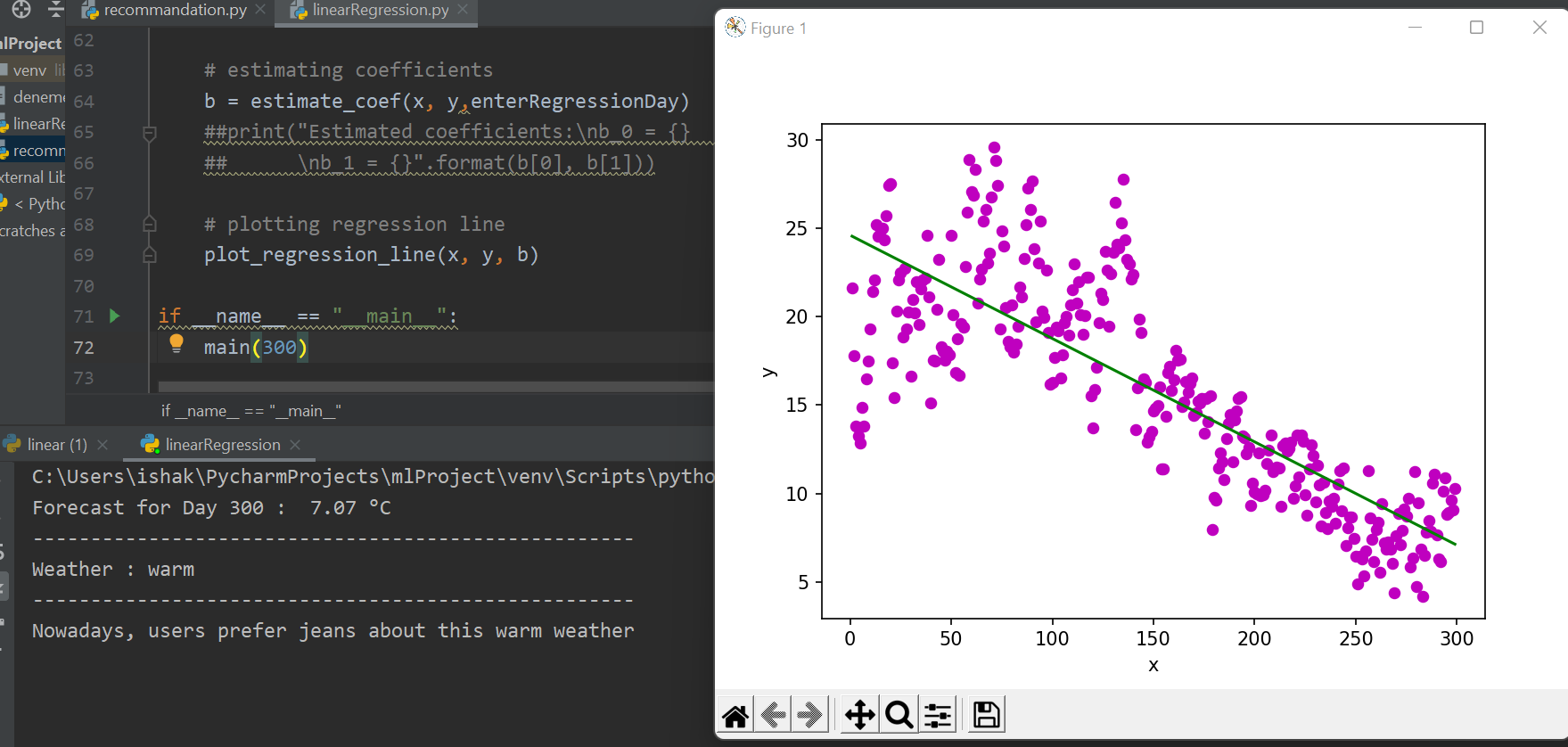
To find the 5th day we wrote 5 in the main method.



Here’s an example of the output of our code when the input (x) is 100.



Here’s an example of the output of our code when the input (x) is 300.



If we use recommendation method;

For day 3;

Text

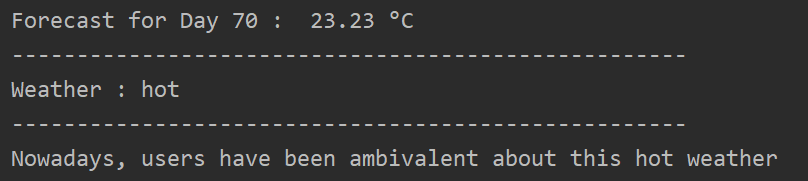
Description automatically generated

For day 20

Text

Description automatically generated with medium confidence

For day 70



**Conclusion:**

The estimated weather in 5th day is 9.35 but the mean temperature in the 5th day is 12.85.The estimated weather in 100th day is 23.42 but the mean temperature in the 100th day is 16.25.The estimated weather in 300th day is 7.07 but the mean temperature in the 300th day is 11.15.

There are some errors between the real and estimated value as we predicted. The reason for this error is the weather inputs does not increase or decrease in a linear way and that’s why the weather prediction can may not have accurate outputs as we wish.

There are some situations in the second method. As we know, the result which got with regression cannot sometimes find same weather in previous values after classification to weather. Therefore, system prints **dataset has not values.** Apart from that if the most preferred cloth is bigger than 1, system recommend each clothes. However, the most preferred cloth is bigger than 3,dataset has an undecided frame and system recommends **users have been ambivalent.**

**Contribution:**

Betülnaz’s part (%50): Linear regression system and writing report.

İshak’s part (%50): Recommendation system and writing report.

**References:**

<https://www.geeksforgeeks.org/linear-regression-python-implementation/>

<https://www.kaggle.com/zaraavagyan/weathercsv>

**Ethical Statement:**

While preparing the project, we did not take any action that could be considered as plagiarism.