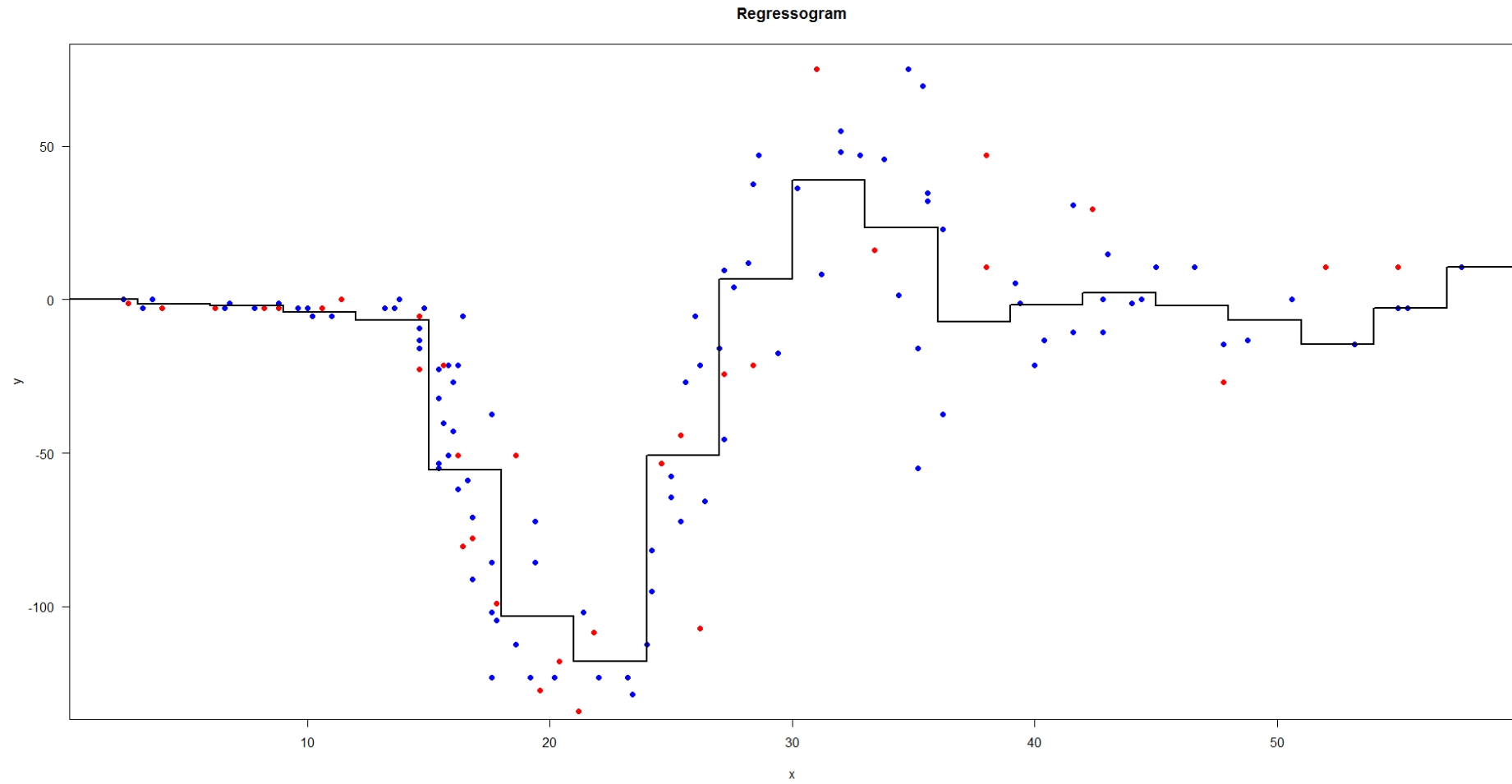


Introduction to Machine Learning HW4

- 1- I first read the csv file and separated the first 100 data points into x_{train} and y_{train} , and remaining 33 to x_{test} and y_{test} , as instructed.
- 2- I defined given bin width and origin variables. I also checked $\max(x_{train}, x_{test}) = 57.6$ and defined the maximum value as 60 for convenience. Then I calculated left and right borders for regressogram.
- 3- In regressogram, I checked for every data point whether it is in the same bin with any value introduced in previous section (left and right borders). I multiplied this vector with y_{train} values and then divided to sum of the vector, which returned the regressogram. Then I plotted the data and calculated the RMSE.
- 4- In Running Mean and Kernel Smoother, I used very similar approaches which differ only on the bin definitions. Running Mean uses absolute $\frac{1}{2}$ times of the bin width and is linear while Kernel Smoother uses Gaussian kernel. Again, I plotted and calculated RMSEs for each method.
- 5- Lastly, I printed the results obtained in RMSE calculations.

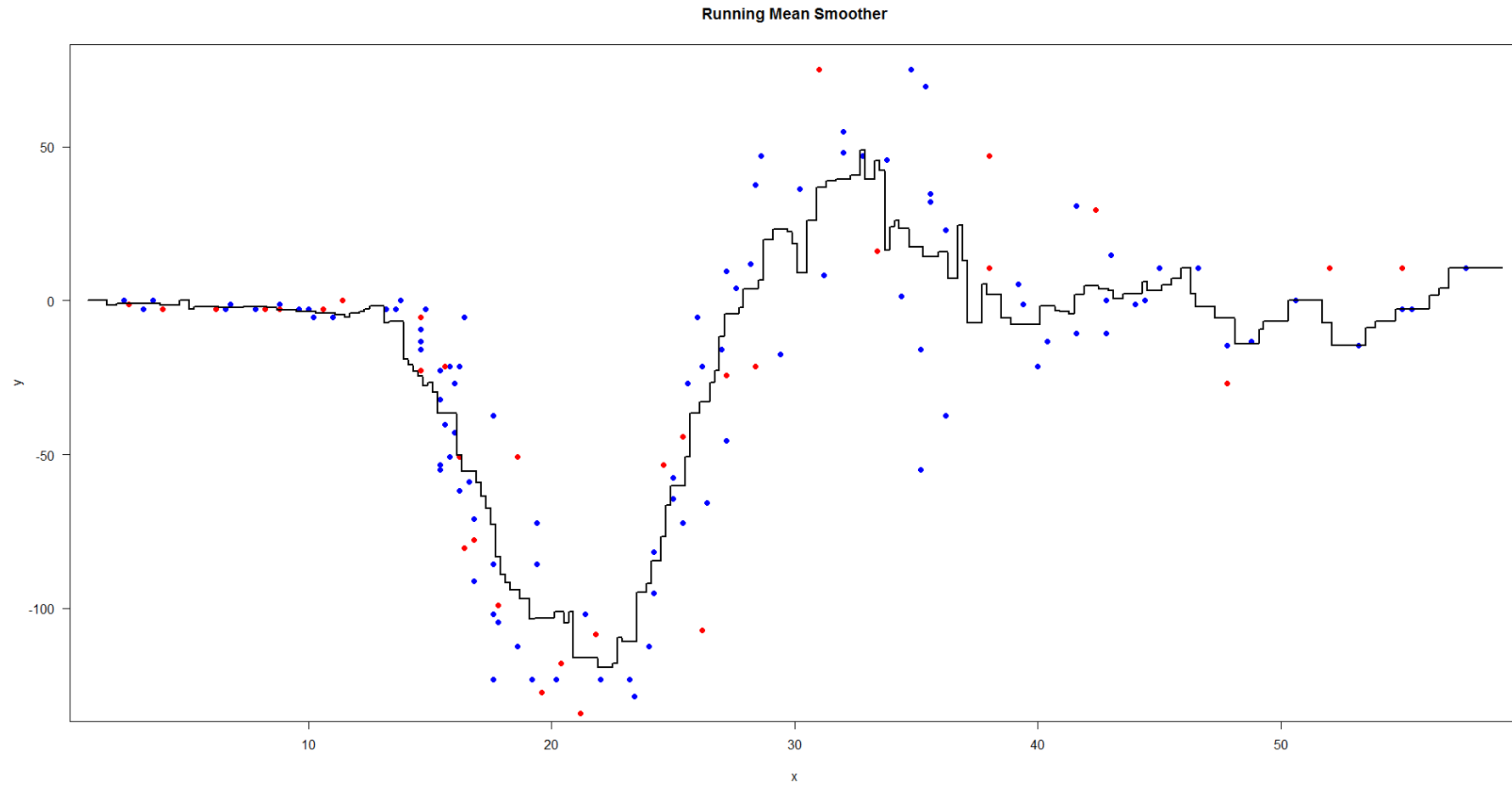
I shared the results next page in case the .r file does not work properly.

07.11.2018
Oğuzhan Akan
30223



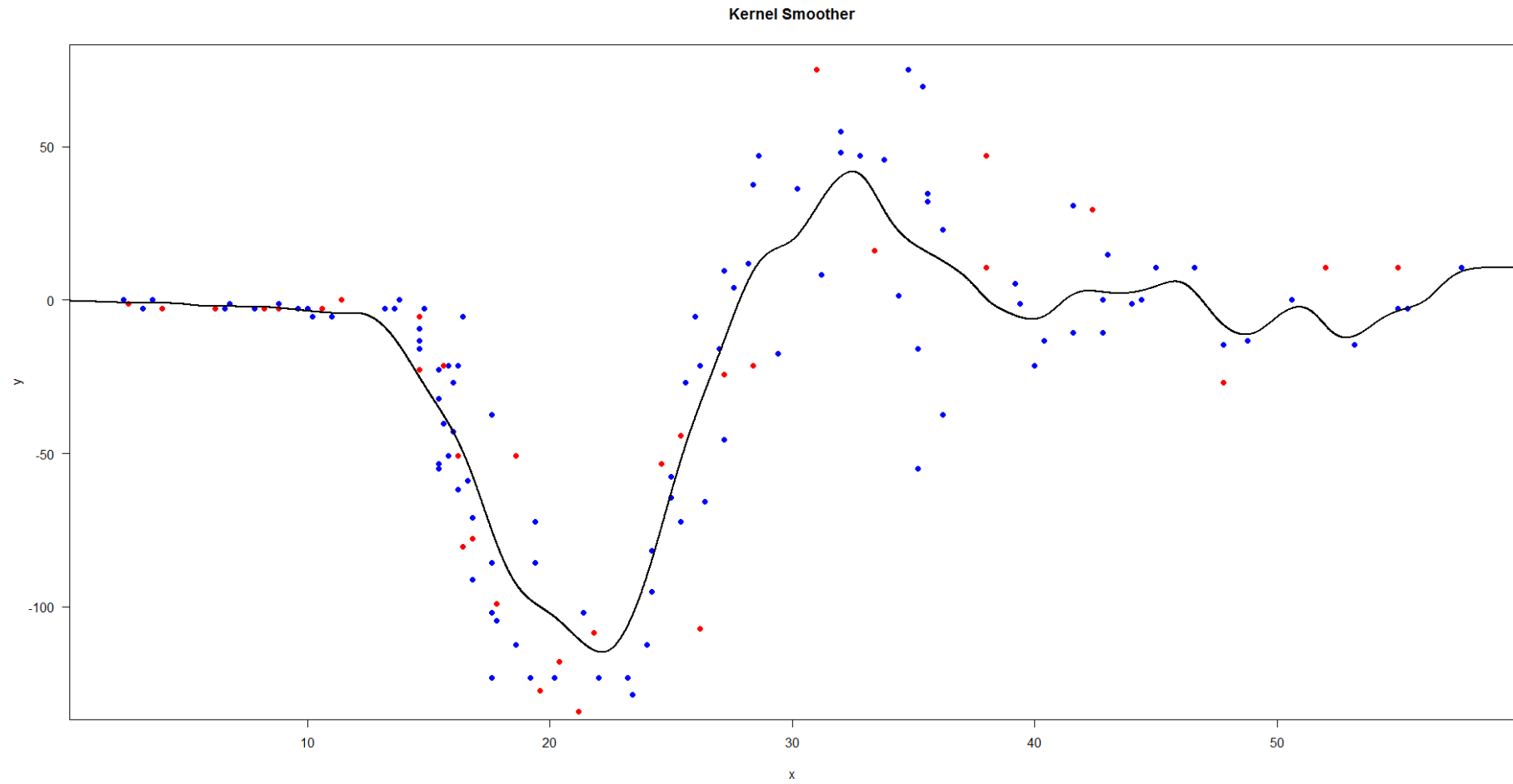
Regressogram => RMSE is 24.726 when h is 3

07.11.2018
Oğuzhan Akan
30223



Running Mean => RMSE is 23.84032 when h is 3

07.11.2018
Oğuzhan Akan
30223



Kernel Smoother => RMSE is 24.17056 when h is 1