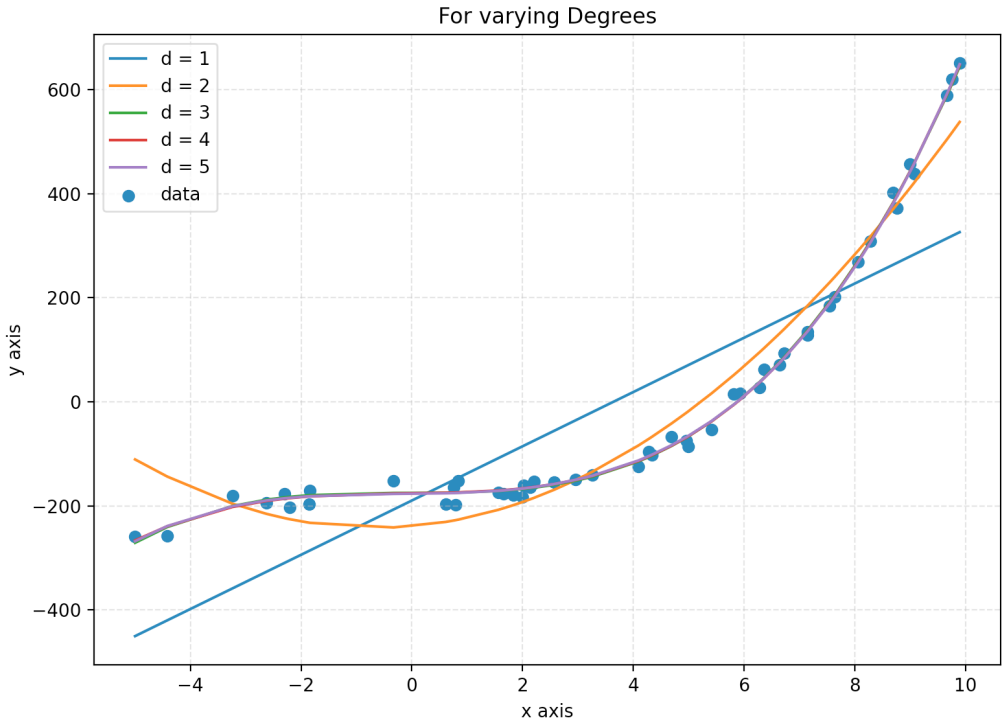
Name: Rahul Maheswaran
Github Username: rahulmaheswaran
PUID: 0029004787 (rmaheswa@purdue.edu)
Instructor: Chris Brinton
Problem 1 Writeup

1 Estimated Functions

$$\hat{y_1}(x) = 52.1581x - 189.866$$
$$\hat{y_2}(x) = 7.00158x^2 + 9.30386x - 239.334$$
$$\hat{y_3}(x) = 0.820138x^3 + 0.261767x^2 - 0.0103277x - 175.277$$
$$\hat{y_4}(x) = 0.00598796x^4 + 0.755218x^3 + 0.23456x^2 + 1.17636x - 175.880$$
$$\hat{y_5}(x) = 0.00085312x^5 - 0.00469804x^4 + 0.752811x^3 + 0.526085x^2 + 0.965916x - 176.837$$

2 Data Visualization



3 Discussion

The data seems best follow a 3rd order polynomial which can be seen from the low error between the regression function $\hat{y_3}(x)$ and the data in the plot above. Though $\hat{y_3}(x)$, $\hat{y_4}(x)$, $\hat{y_5}(x)$ overlap, $\hat{y_3}(x)$ is a better fit as it is more generalized and prevents over-fitting errors, when used to predict other values.

If we measured a new data point $x = 2$, the corresponding predicted value would be $\hat{y_3}(x) = -167.68922055890684$.