

1. Understand and explain the risks of overfit and underfit, and use first principles to help avoid overfit
 2. Use polynomial models to make predictions and interpret the results in the context of real world scenarios
 3. Think through ethical considerations for using polynomial models to make predictions
 4. Build polynomial models and other models using Excel trendline
-

Example Problem:

Remember the problem from yesterday: A marketing team is tracking the cumulative, or total, sales of a product since its launch. The data for the first 10 days are below.

1. Use Excel to try polynomial trendlines of degrees 2 through 6.
2. Record the equations and R^2 values for each.
3. Which model gives the highest R^2 ?
4. Which model behaves the way you would expect for *cumulative* sales?
5. Why might the more complex model be misleading? How could this be problematic to a decision maker?

Day	Cumulative Sales
0	0
1	25
2	60
3	100
4	140
5	170
6	190
7	205
8	220
9	232
10	240

Practice Problem:

You've been asked to analyze the relationship between the years of experience (in years) an employee has and their corresponding salary (in dollars). Using the data below create a higher order polynomial model than what you created yesterday. Use that model to make a prediction for the salary of an employee with 9 and 12 years. Do these predictions seem better or worse? Based on your analysis, what model would you recommend to a company VP? Justify your choice, considering the mathematical fit and real-world expectations.

Years	Salary (\$)
1	50,000
2	55,000
3	65,000
4	80,000
5	110,000
6	150,000
7	200,000

A marketing analyst wants to model how many new customers are gained based on monthly advertising spending. Using the data below develop a model and predict how many customers would be gained with \$7,500, \$10K, and \$13K in advertising.

Ad Spend (\$K)	New Customers
1	22
2	39
3	55
4	61
5	64
6	69
7	68
8	72
9	71