- Can have many or no correct answers
- How to know if we have good predictors?
  - Aim for no relation between the features, but features should be related to the thing we want to predict.
  - Correlation
    - Look at point positions in the data set (rank)
- How to know if we have done a good job?
- Spearman Coefficient:
  - Look at natural order of numbers
    - 10, 20, 30, 40 ,50 → 1,2,3,4,5
    - 1,0,2,3,4 → 2,1,3,4,5
    - From formula: -1,1,0,0,0
    - Why is it multiplied by 6?
      - Things should be in the range of -1 to 1. Consider maximum value of sum of squares. Only occurs if distance is maximal at a time.
- Correlation vs Causation:
  - Correlation:
    - Temperatures and ice cream sales are positively correlated

        - Temp increase does not cause ice cream sales to rise
    - Sleeping with shoes on is correlated with waking up with a headache
        - Third factor that causes correlation: going to bed drunk
- Causation:
    - Testing is necessary to know causation for sure
        - Tricky to test causality
- Testing:
    - Split data into train/test sets
    - Use training sets to find patterns and create model
    - Use testing set to evaluate the model on data it has never seen before
- Overfitting vs Underfitting
    - Overfitting: Model is too tailored to the data
    - Underfitting: Model may be more complex but not well equipped
- The type of mistake matters (mistakes on outliers, false positives/negatives)