**Exercise 1:** data smoothing
Dataset:

| Day | Price |
|---|---|
| 1 | 10 |
| 2 | 12 |
| 3 | 13 |
| 4 | 15 |
| 5 | 14 |
| 6 | 16 |
| 7 | 18 |

**Instructions:**

- Calculate the 3-Day SMA for each day starting from Day 3:
- Calculate the 3-Day WMA for each day starting from Day 3, with weights w3=3; w2=2 and w1=1.

**Exercise** 2: Feature Selection Using Variance, Covariance, and Entropy
Given the next dataset, answer the following questions.

| Customer | Age Group (X1) | Income Level (X2) | Browsing Time (X3) | Purchase (Y) |
|---|---|---|---|---|
| 1 | 25 | 40000 | 30 | 0 |
| 2 | 35 | 60000 | 40 | 1 |
| 3 | 30 | 55000 | 35 | 1 |
| 4 | 45 | 30000 | 20 | 0 |
| 5 | 40 | 45000 | 50 | 1 |

**Part 1:** Feature Selection Using Variance and Covariance

1. Give an opinion about calculating the covariance between a numerical feature and a binary class feature.
2. Given the Point Biserial Correlation formula defined by: n1, n2, number of observations of both class 1 and 2. n,s number of observations and s standard deviation of the feature. X1, X2 the mean of values where the class is 1, and mean of values where the class is 0.

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_0}{s}\sqrt{\frac{n_1 n_0}{n^2}}$$

3. Select Features Based on correlation:
    a. Calculate Point Biserial Correlation of each feature with the target class.
    b. Which feature shows the strongest relationship with the target variable, and which feature has the most?

**Part 2:** Feature Selection Using Entropy and Information Gain

1. Convert Categorical Features to Ordinal Data.
2. Calculate the Entropy of the Target Variable (Y).
3. Calculate the Conditional Entropy for Each Feature
4. Calculate the Information Gain for Each Feature
5. Compare the Results:
    - Compare the results from Part 1 with Part 2 (Entropy and Information Gain).
    - Which method provided more useful insights for feature selection in predicting Purchase (Y)?

**Exercise 3:** Lecture questions

1. What is the purpose of feature engineering in a data science project? Why is it considered an essential step?
2. What is the difference between feature selection and feature extraction? Provide examples of each.
3. Why might removing irrelevant features (feature selection) improve model performance? What issues can irrelevant features cause?
4. How would you decide which features to select if you have a large dataset with many features? What criteria or methods could you use?
5. Explain how domain knowledge can assist in selecting or extracting relevant features. Why is it important in feature engineering?
6. What are some methods for selecting features in a dataset? Describe one technique and explain how it helps improve the model.
7. Why might dimensionality reduction techniques, like Principal Component Analysis (PCA), be useful in feature extraction? What are some benefits and drawbacks of using PCA?
8. In your opinion, what types of features are usually most important for models: those based on original data, selected features, or extracted features? Explain your reasoning.
9. What is an exhaustive search algorithm, and how might it be used in feature selection? What are the limitations of using an exhaustive approach?
10. Explain what a metaheuristic algorithm is and provide an example of one commonly used in feature selection. How do metaheuristics address some of the limitations of exhaustive search?
11. What are some potential disadvantages of using metaheuristic algorithms in feature selection? Are there cases where these techniques might perform poorly