1. (True/False) Machine Learning is a subset of Artificial Intelligence
   1. True
2. (True/False) Deep Learning is a subset of Machine Learning
   1. True
3. (True/False) Machine Learning consists in programming computers to learn from real-time human interactions
   1. False
4. (True/False) AI Winters happened mostly due to the lack of understanding behind the theory of neural networks
   1. True
5. Most modern applications that use computer vision, use models that were trained using this discipline:
   1. Deep Learning
6. In the Machine Learning Workflow, the main goal of the Data Exploration and Preprocessing step is to:
   1. Determine how to clean your data such that you can use it to train a model
7. What is the goal of supervised learning?
   1. Predict the labels.
8. What is deep learning?
   1. Deep learning is machine learning that involves deep neural networks.
9. When is a standard machine learning algorithm usually a better choice than using deep learning to get the job done?
   1. When working with small data sets.
10. What is a Turing test?
    1. It tests a machine's ability to exhibit intelligent behavior.
11. What are some of the different milestones in deep learning history?
    1. Geoffrey Hinton’s work, AlexNet, and TensorFlow
12. What is artificial intelligence?
    1. Any program that can sense, reason, act, and adapt.
13. What are two spaces within AI that are going through drastic growth and innovation?
    1. Computer vision and natural language processing.
14. Why did AI flourish so much in the last years?
    1. Faster and inexpensive computers and data storage
15. How does Alexa use artificial intelligence?
    1. Recognizes our voice and answers questions.
16. What are the first two steps of a typical machine learning workflow?
    1. Problem statement and data collection.
17. Which statement about the Pandas read\_csv function is TRUE?
    1. It can read both tab-delimited and space-delimited data.
18. Which of the following is a reason to use JavaScript Object Notation (JSON) files for storing data?
    1. Because they are cross-platform compatible.
19. The data below appears in 'data.txt', and Pandas has been imported. Which Python command will read it correctly into a Pandas DataFrame?

63.03 22.55 39.61 40.48 98.67 -0.25 AB

39.06 10.06 25.02 29 114.41 4.56 AB

68.83 22.22 50.09 46.61 105.99 -3.53 AB

* 1. pandas.read\_csv('data.txt', header=None, sep=' '

1. (True/False) Outliers must be very extreme to noticeably impact the fit of a statistical model.
   1. False
2. (True/False) Outliers should always be replaced, since they never contain useful information about the data.
   1. False
3. Which residual-based approach to identifying outliers compares running a model with all data to running the same model, but dropping a single observation?
   1. Externally-studentized residuals
4. What is a CSV file?
   1. CSV files are rows of data or values separated by commas.
5. What are residuals?
   1. Residuals are the difference between the actual values and the values predicted by a given model.
6. If removal of rows or columns of data is not an option, why must we ensure that information is assigned for missing data?
   1. Most models will not accept blank values in our data.
7. What are the two main data problems companies face when getting started with artificial intelligence/machine learning?
   1. Lack of relevant data and bad data
8. What does SQL stand for and what does it represent?
   1. SQL stands for Structured Query Language, and it represents a set of relational databases with fixed schemas.
9. What does NoSQL stand for and what does it represent?
   1. NoSQL stands for Not-only SQL, and it represents a set of databases that are not relational, therefore, they vary in structure.
10. What is a JSON file?
    1. JSON stands for JavaScript Object Notation, and it is a standard way to store the data across platforms.
11. What is meant by the Messy Data?
    1. Duplicated or unnecessary data.
    2. Inconsistent text and typos.
    3. Missing data.
12. What is an outlier?
    1. Outlier is an observation in dataset that is distant from most other observations.
13. How do we identify outliers in our dataset?
    1. We can identify outliers both visually and with statistical calculations.
14. From the options listed below, select the option that is NOT a valid exploratory data approach to visually confirm whether your data is ready for modeling or if it needs further cleaning or data processing:
    1. Create a correlation heatmap to confirm the sign and magnitude of correlation across your features.
15. These are two of the most common variables for data visualization:
    1. matplotlib and seaborn
16. (True/False) You can use the pandas library to use plots.
    1. True
17. Which scaling approach converts features to standard normal variables?
    1. Standard scaling
18. Which variable transformation should you use for ordinal data?
    1. Ordinal encoding
19. What are polynomial features?
    1. They are higher order relationships in the data.
20. What does Boxcox transformation do?
    1. It transforms the data distribution into more symmetrical bell curve
21. Select three important reasons why EDA is useful.
    1. To determine if the data makes sense, to determine whether further data cleaning is needed, and to help identify patterns and trends in the data
22. What assumption does the linear regression model make about data?
    1. This model assumes a linear relationship between predictor variables and outcome variables.
23. What is skewed data?
    1. Data that is distorted away from normal distribution; may be positively or negatively skewed.
24. Select the two primary types of categorical feature encoding.
    1. Nominal encoding and ordinal encoding
25. Which scaling approach puts values between zero and one?
    1. Min-max scaling
26. Which variable transformation should you use for nominal data with multiple different values within the feature?
    1. One-hot encoding

# Practice Quiz: Estimation and Inference, and Hypothesis Testing

1. (True/False) In general, the population parameters are unknown.
   1. True
2. (True/False) Parametric models have **finite** number of parameters.
   1. True
3. The most common way of estimating parameters in a parametric model is:
   1. using the maximum likelihood estimation

# Practice Quiz: Hypothesis Testing

1. A p-value is:
   1. the smallest significance level at which the null hypothesis would be rejected
2. Type 1 Error 1 is defined as:
   1. Saying the null hypothesis is false, when it is actually true
3. You find through a graph that there is a strong correlation between Net Promoter Score and the visual time that customers spend on a website. Select the TRUE assertion:
   1. There is an underlying factor that explains this correlation, but manipulating the time that customers spend on a website may not affect the Net Promoter Score they will give to the company

# Graded Quiz: Module 4 - Inferential Statistics and Hypothesis Testing

1. Which one of the following is common to both machine learning and statistical inference?
   1. Using sample data to infer qualities of the underlying population distribution.
2. Which one of the following describes an approach to customer churn prediction stated in terms of probability?
   1. Predicting a score for individuals that estimates the probability the customer will leave.
3. What is customer lifetime value?
   1. The total purchases over the time which the person is a customer.
4. What is customer lifetime value?
   1. It provides an estimate of the variable’s probability distribution.
5. The outcome of rolling a fair die can be modelled as a \_\_\_\_\_\_\_ distribution.
   1. uniform
6. Which one of the following features best distinguishes the Bayesian approach to statistics from the Frequentist approach?
   1. Bayesian statistics incorporate the probability of the hypothesis being true.
7. Which of the following best describes what a hypothesis is?
   1. A hypothesis is a statement about a population.
8. A Type 2 error in hypothesis testing is \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_:
   1. incorrectly accepting the null hypothesis.
9. Which statement best describes a consequence of a type II error in the context of a churn prediction example? Assume that the null hypothesis is that customer churn is due to chance, and that the alternative hypothesis is that customers enrolled for greater than two years will not churn over the next year.
   1. You incorrectly conclude that customer churn is by chance
10. Which of the following is a statistic used for hypothesis testing?
    1. The likelihood ratio.