1. The key tasks involved in machine learning are data collection, data preprocessing, feature engineering, model selection, training, and evaluation. Data pre-processing involves cleaning, transforming, and normalizing the data to make it suitable for analysis.

2. Quantitative data refers to numerical data that can be measured and quantified. Qualitative data, on the other hand, refers to non-numerical data that cannot be measured in numbers. For example, age is quantitative data, while gender is qualitative data.

3. Sample data collection:

| Attribute | Data Type |

|-----------|----------|

| Name | String |

| Age | Numeric |

| Gender | Categorical |

| Height | Numeric |

| Education Level | Ordinal |

4. Machine learning data issues can arise due to missing data, outliers, irrelevant data, and noisy data. These issues can result in biased or inaccurate models.

5. Categorical data can be explored through frequency tables, bar charts, and pie charts. For example, a bar chart can be used to display the frequency of each category in a categorical variable.

6. Missing values can affect the learning process by reducing the accuracy and effectiveness of the model. The missing values can be filled in using imputation techniques, such as mean imputation, median imputation, or mode imputation.

7. The methods for dealing with missing data include deleting the rows or columns with missing data, imputing missing data using mean or median, and using machine learning algorithms that can handle missing data, such as K-nearest neighbors.

8. Data pre-processing techniques include normalization, transformation, feature scaling, feature selection, and dimensionality reduction. Dimensionality reduction involves reducing the number of features in the dataset, while feature selection involves selecting the most relevant features.

9.

i. IQR stands for interquartile range and is a measure of variability in the data. It is calculated as the difference between the third quartile and the first quartile. The IQR is used to identify outliers in the data.

ii. A box plot displays the quartiles of the dataset as well as outliers. The lower whisker will be longer than the upper whisker if the data is skewed to the left. Box plots can be used to identify outliers by plotting any data points that fall outside the whiskers.

10.

1. Data collected at regular intervals can be analyzed using time series analysis.

2. The gap between the quartiles is the interquartile range, which is used to identify outliers.

3. Cross-tabs can be used to examine the relationship between two categorical variables.

1.

1. Data with nominal values are categorical data that cannot be ordered, while ordinal data can be ordered.

2. Histograms are used to show the distribution of numeric data, while box plots display the quartiles and outliers.

3. The average is affected by outliers, while the median is not.