

1, Describe the relationship between the Represent and the Critique Stage

In represent stage, we figure out how we going to represent data and the basic form of it. The relationship between these two is that we do an analytical evaluation of what works and what doesn't to the visualization we decided in the represent stage.

2, List the advantages of filtering data before mining.

Reduce workload for mining stage.

Reduce the probabilities of making errors in mining stage.

3, I can describe what happens in the Filter stage

In this stage, we filter the data and remove portions that are not relevant to our use.

4, List 4 basic statistics methods for describing data

Measure of frequency.

Measure of central tendency.

Measure of dispersion or variation.

Measure of position.

5, Define & give an example of a secondary data source

Secondary data source means data that is collected by someone other than the primary user.

An example of that is Wikipedia.