1)

a) No – this might be part of the preprocessing of data, but not the datamining step

b) No – this would be considered input data, or gathering input data

c) No – harder, but would not say so. Data mining refers to extraction of patterns or useful information, that was not previously known. Although the sum of sales might be considered a ”unknown” quantity, I would not include simple arithmetic in data mining. No patterns etc is discovered.

d) No – even though you can treat this as a data mining problem (input sequence of head/tail) and try to find a pattern to predict, this does not satisfy the “useful information, not previously known” criterion.

e) Yes – looking for patterns in historical data can be considered a data mining task, as it takes some preprocessed input data, analyzes it for patterns etc. and gives some output based on the findings.

f) Yes – analyzing the data and try to find abnormalities or outliers may be considered a data mining task, the monitoring part would be input data, and the data mining task would be looking for outliers.

2)

a) Continuous, ratio

b) Discrete, ordinal

c) Continuous, ratio

d) Discrete…, nominal

e) Discrete, nominal.

f) Discrete, ordinal

3)

**Visualization** is applied both in preprocessing and postprocessing. For preprocessing, visualization of data is an easy way to get an understanding of distribution, outliers etc for each of the input variables. In the postprocessing, it is an internpretable way to explore and showcase results.

**Dimensonality reduction** is done in the preprocessing step, to isolate important variables, reduce noise in the input data and improve effiecency.

**Machine learning** is applied in the data mining step. Machine learning techniques are used to find patterns and transform data into knowledge. It can do so extremely fast and effecetivly compared to what a human could do without these tools.

4)

a)