**Correlated items**

Best practice used to identify these

* What is best practice?

Evaluated against ?

**Cannibalization items:**

Create a Probabilistic classifier

Data labeled by comparing forecast with real sales during promotional period

Model trained and evaluated against the labels

Create a model that uses the features and the sale history to identify the cannibalization items?

Using both models together - evaluate against a well-known statistical model that only looks at sale history

* Need to decide on the statistical model

**Adaptive sampling?**