3. **Distribution skew** occurs when the distribution of the training dataset is significantly different from the distribution of the serving dataset, and is typically caused by: (check all that apply).

☐ A data source that provides some feature values is modified between training and serving time.

☐ There is different logic for generating features between training and serving. For example, if you apply some transformation only in one of the two code paths.

☑ Faulty sampling method that selects a sample for training which is not representative of serving data distribution.

✓ **맞습니다**

Spot on! A faulty sampling mechanism that chooses a non-representative subsample is an example of distribution skew.

☑ Different data sources for training and serving data.

✓ **맞습니다**

Way to go! Data sources between training and serving often change and so this is another case of distribution skew.

☑ Trend, seasonality, changes in data over time.

✓ **맞습니다**

Keep it up! Data distributions between training and serving often change and so this is another case of distribution skew.