

## Some Data Analysis Notes

A dataset is a collection of measurements that are used to learn about a population. Our data are the measurements we take (the sample) to try to characterize the overall population. The choice of statistical methods to be used to characterize a population based on the data we have in hand should be guided by what we know about the characteristics of the population. This process usually involves the scientific process of model building using both inductive and deductive reasoning. To select the right statistical method we need to know something about the characteristics of the data. One of the most frequent tasks when analyzing data is to describe and summarize those data in forms that convey their important characteristics.

### Characteristics of Water Resources Data

- A lower bound of zero
- The presence of outliers
- Positive skewness
- Non-normal distribution of data
- Censored (reported below or above a certain threshold)
- Seasonality
- Autocorrelation
- Dependence on other variables