Bayesian networks come from the theory of Rev. Bayes and it relates conditional and marginal probabilities of events A and B. A is considered the “prior” or marginal probability and does not consider any information dealing with B, but event B does not need to occur after event A. One example of this would be what is the probability that it rained given that it is wet outside.

Bayesian inference is useful for modeling random variables like a demographic statistic or a business KPI. It is most useful when data is limited or precisely knowing how likely certain facts are then others. Since Bayesian inference is used when you are uncertain about some fact, the data is used help understand. 2 distributions are used prior and posterior. Prior reflects beliefs before seeing the data and posterior reflects beliefs after seeing the data.

There are typically 4 prior distribution:

1. Informative; empirical – data from previous experiments are reflected in the prior beliefs
2. Informative; non-empirical – there is some inherent reason to prefer certain values over others
3. Informative; domain-knowledge – no supporting data, but know some fact are truer than others because of expert level status
4. Non-informative – beliefs have little or no effect

A key benefit of Bayesian networks is that they can be built from human knowledge or machine learned from data. They are also bidirectional meaning that they can be made from the ways or a combination of the two both building off one another.