Policy based methods differ from value based methods in the sense that with policy based methods, you do not have to worry about training your neural network to make estimates on the current value of a given state and action your agent. Instead, the goal is to select an action based on a specific policy such that instead of selecting an action based on a random policy distribution, your agent will select actions deterministically. An advantage of this compared to other value-based algorithms such as epsilon-greedy Q-learning is that it eliminates the chance of being greedy by selecting a poor action. Because the current policy is approaching an optimal policy, then the agent will eventually select the best action without leaving it up to chance. Based on my current understanding, the differences between the two is as follows:

1. Value-based methods do not perform value estimates iteratively whereas policy based methods do in order to reach convergence. In the case of TRPO, because the algorithm is based around A2C, the policy is trained under a certain number of time steps. This isn't the case with value-based methods.
2. Value-based methods reach convergence slower compared to policy based methods.
3. Value based methods can solve harder control problems whereas policy based methods are more apt for simple problems.
4. Policy based methods are more apt for continuous action spaces.
5. Value based methods perform explicit exploration whereas policy-based have innate exploration.
6. It is easier to train off-policy for value-based whereas policy-based works well with supervised learning.

Baselines are used in order to reduce variance and increase stability in order to reduce the size of the policy update. TRPO is an example of such that introduces different ways in order to reduce the size of the policy update such as normalizing the advantage function and computing the clipped loss objective or by computing the KL-penalized objective and subtracting the loss function based on that. In that sense, the KL-penalized objective and clipped surrogate objectives are actually baselines.