Assignment-4

1. **Optimal logic query plan**
- The optimized logical query works in this manner, first it starts with the Product relation and moves according to the condition on attributes Production-year which should be after 2000 and the Rating which should be 5. It then projects the Product on name and company-name. It then selects the Company relation while projecting the attributes name and state and finally it joins the projected product and company relations and project the resulting table on attributes name of the product and state of the company.

2. **Query Optimization**
- To calculate the cost I used the hash join formula:

For example:
RS = 3 * B(R) + 3 * B(S)
Where B(R) = T(R)/5   , since we can have at most 5 tuples.

For size I used the Selinger formula :

For example:
(T(R) * T(S)) / (max (V(R,B), V(S,B)) * max(V(R,C), V(S,C)))

Query	Size	Cost	Plan
RS	200	4200	SR
RW	40k	3600	WR
RU	40k	3000	UR
SW	60k	3000	WS
WU	20k	1800	UW
RSW	4000	38400	R(SW)
RSU	2000	25800	S(RU)
RWU	4000	14400	R(UW)
SWU	6000	13800	S(UW)
RSWU	80k	36k	(UW)(UR)

Assignment-4

3. **Consider the following classes of schedules: serializable and 2PL. For each of the following schedules, state which of the preceding classes it belongs to. If you cannot decide whether a schedule belongs to a certain class based on the listed actions, explain briefly your reasons**

1. **T1:R(X), T2:R(Y), T3:W(X), T2:R(X), T1:R(Y)**
   o It can be categorized as serializable since it does not creates a cycle of conflicts. It can be be also said 2PL since it doesn't have conflicting commands with locks.

2. **T1:R(X), T1:R(Y), T1:W(X), T2:R(Y), T3:W(Y), T1:W(X), T2:R(Y)**

   o It does not seems to be serializable because it Reads Y again in T2 at last from T3 which forms a cycle and therefore it cannot be 2PL as well because it has to be serializable to be a 2PL schedule.