Candidate Generation: An SQL Implementation

- Suppose the items in F_{k-1} are listed in an order
- Step 1: self-joining F_{k-1}

  insert into C_k

  select $p.item_1, p.item_2, \ldots, p.item_{k-1}, q.item_{k-1}$

  from F_{k-1} as p, F_{k-1} as q

  where $p.item_1 = q.item_1, \ldots, p.item_{k-2} = q.item_{k-2}, p.item_{k-1} < q.item_{k-1}$
- Step 2: pruning

  for all *itemsets c in* C_k do

  for all *(k-1)-subsets s of c* do

  if *(s is not in* F_{k-1}*)* **then delete** *c* **from** C_k