nodeID(40) Filter isnotnull(s_store_sk#131) nodeID(38) FileScan parquet tpcds_100_parquet.store_sales[ss_sold_date_sk#108,ss_store_sk#115,ss_net_profit#130] nodeID(45) FileScan parquet tpcds_100_parquet.date_dim[d_date_sk#160,d_month_seq#163] nodeID(39) BroadcastExchange HashedRelationBroadcastMode(List(cast(input[0, int, false] as bigint)),false), [plan_id=151] nodeID(37) Filter (isnotnull(ss_store_sk#115) AND isnotnull(ss_sold_date_sk#108)) $nodeID(44) \ Filter (((isnotnull(d_month_seq\#163) \ AND \ (d_month_seq\#163 >= 1200)) \ AND \ (d_month_seq\#163 <= 1211)) \ AND \ isnotnull(d_date_sk\#160))$ nodeID(36) BroadcastHashJoin [ss_store_sk#115], [s_store_sk#131], Inner, BuildRight, false nodeID(42) BroadcastExchange HashedRelationBroadcastMode(List(cast(input[0, int, true] as bigint)), false), [plan_id=155] nodeID(35) Project [ss_sold_date_sk#108, ss_net_profit#130, s_state#155] nodeID(34) BroadcastHashJoin [ss_sold_date_sk#108], [d_date_sk#160], Inner, BuildRight, false nodeID(33) Project [ss_net_profit#130, s_state#155] nodeID(32) HashAggregate(keys=[s_state#155], functions=[partial_sum(ss_net_profit#130)]) nodeID(31) Exchange hashpartitioning(s_state#155, 200), ENSURE_REQUIREMENTS, [plan_id=160] nodeID(30) HashAggregate(keys=[s_state#155], functions=[sum(ss_net_profit#130)]) nodeID(29) Sort [s_state#155 ASC NULLS FIRST, _w2#98 DESC NULLS LAST], false, 0 nodeID(19) FileScan parquet tpcds_100_parquet.date_dim[d_date_sk#31,d_month_seq#34] nodeID(25) FileScan parquet tpcds_100_parquet.store[s_store_sk#59,s_county#82,s_state#83] nodeID(28) Window [rank(_w2#98) windowspecdefinition(s_state#155, _w2#98 DESC NULLS LAST, specifiedwindowframe(RowFrame, unboundedpreceding\$(), currentrow\$())) AS ranking#4], [s_state#155], [_w2#98 DESC NULLS LAST] $nodeID(18) \ Filter (((isnotnull(d_month_seq\#34) \ AND \ (d_month_seq\#34 >= 1200)) \ AND \ (d_month_seq\#34 <= 1211)) \ AND \ isnotnull(d_date_sk\#31))$ nodeID(24) Filter isnotnull(s_store_sk#59) nodeID(27) Filter (ranking#4 <= 5) nodeID(17) Project [d_date_sk#31] nodeID(23) Exchange hashpartitioning(s_state#83, 200), ENSURE_REQUIREMENTS, [plan_id=169] nodeID(15) FileScan parquet tpcds_100_parquet.store_sales[ss_sold_date_sk#8,ss_store_sk#15,ss_net_profit#30] nodeID(16) BroadcastExchange HashedRelationBroadcastMode(List(cast(input[0, int, true] as bigint)),false), [plan_id=148] nodeID(22) Sort [s_state#83 ASC NULLS FIRST], false, 0 nodeID(26) Project [s_state#155] nodeID(14) Filter (isnotnull(ss_sold_date_sk#8) AND isnotnull(ss_store_sk#15)) nodeID(13) BroadcastHashJoin [ss_sold_date_sk#8], [d_date_sk#31], Inner, BuildRight, false nodeID(21) SortMergeJoin [s_state#83], [s_state#155], LeftSemi nodeID(12) Project [ss_store_sk#15, ss_net_profit#30] nodeID(20) BroadcastExchange HashedRelationBroadcastMode(List(cast(input[0, int, false] as bigint)),false), [plan_id=173] nodeID(11) BroadcastHashJoin [ss_store_sk#15], [s_store_sk#59], Inner, BuildRight, false nodeID(10) Project [ss_net_profit#30, s_state#83, s_county#82] nodeID(9) Expand [[ss_net_profit#30, s_state#83, s_county#82, 0], [ss_net_profit#30, s_state#83, null, 1], [ss_net_profit#30, null, null, 3]], [ss_net_profit#30, s_state#193, s_county#194, spark_grouping_id#192L] nodeID(8) HashAggregate(keys=[s_state#193, s_county#194, spark_grouping_id#192L], functions=[partial_sum(ss_net_profit#30)]) nodeID(7) Exchange hashpartitioning(s_state#193, s_county#194, spark_grouping_id#192L, 200), ENSURE_REQUIREMENTS, [plan_id=179] nodeID(6) HashAggregate(keys=[s_state#193, s_county#194, spark_grouping_id#192L], functions=[sum(ss_net_profit#30)]) nodeID(5) Exchange hashpartitioning(_w1#207, _w2#208, 200), ENSURE_REQUIREMENTS, [plan_id=182] nodeID(4) Sort [_w1#207 ASC NULLS FIRST, _w2#208 ASC NULLS FIRST, _w3#209 DESC NULLS LAST], false, 0 nodeID(3) Window [rank(_w3#209) windowspecdefinition(_w1#207, _w2#208, _w3#209 DESC NULLS LAST, specifiedwindowframe(RowFrame, unboundedpreceding\$(), currentrow\$())) AS rank_within_parent#2], [_w1#207, _w2#208], [_w3#209 DESC NULLS LAST] nodeID(2) Project [total_sum#0, s_state#193, s_county#194, lochierarchy#1, rank_within_parent#2] nodeID(1) TakeOrderedAndProject(limit=100, orderBy=[lochierarchy#1 DESC NULLS LAST,CASE WHEN (lochierarchy#1 = 0) THEN s_state#193 END ASC NULLS FIRST, output=[total_sum#0,s_state#193,s_county#194,lochierarchy#1,rank_within_parent#2])

nodeID(0)

nodeID(41) FileScan parquet tpcds_100_parquet.store[s_store_sk#131,s_state#155]

nodeID(43) Project [d_date_sk#160]