Thi Hong Yen Truong

ID: 614745

**W2D1**

**Compute Relative Frequencies**

Number of Input-Splits: 2

Number of Reducers: 1

Input Split 0

15 91 80 12 19 80 18

17 15 80 18 19 18

Input Split 1

19 15 80 18 19 18

18 15 18 18 88 18

**(Please use the same window definition, we have been using)**

**OPTIONAL NO EXTRA CREDIT:** In-Mapper Combining Pair Approach, In-Mapper Combining Stripe Approach.

|  |  |
| --- | --- |
| Input-Split 0 | Input-Split 1 |
| 15 91 80 12 19 80 18 | 19 15 80 18 19 18 |
| 17 15 80 18 19 18 | 18 15 18 18 88 18 |

1. Pair Approach

|  |  |
| --- | --- |
| W(X) 0 | W(X) 1 |
| W(15) = {91, 80, 12, 19, 80, 18}  W(91) = {80, 12, 19, 80, 18}  W(80) = {12, 19}  W(12) = {19, 80, 18}  W(19) = {80, 18}  W(80) = {18}  W(17) = {15, 80, 18, 19, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = { 18 } | W(19) = {15, 80, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = {18}  W(18) = {15}  W(15) = { 18, 18, 88, 18}  W(18) = {88}  W(88) = {18} |
| Mapper 0 Output | Mapper 1 Output |
| ((15, 91), 1)  ((15, \*), 1)  ((15, 80), 1)  ((15, \*), 1)  ((15, 12), 1)  ((15, \*), 1)  ((15, 19), 1)  ((15, \*), 1)  ((15, 80), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((91, 80), 1)  ((91, \*), 1)  ((91, 12), 1)  ((91, \*), 1)  ((91, 19), 1)  ((91, \*), 1)  ((91, 80), 1)  ((91, \*), 1)  ((91, 18), 1)  ((91, \*), 1)  ((80, 12), 1)  ((80, \*), 1)  ((80, 19), 1)  ((80, \*), 1)  ((12, 19), 1)  ((12, \*), 1)  ((12, 80), 1)  ((12, \*), 1)  ((12, 18), 1)  ((12, \*), 1)  ((19, 80), 1)  ((19, \*), 1)  ((19, 18), 1)  ((19, \*), 1)  ((80, 18), 1)  ((80, \*), 1) | ((19, 15), 1)  ((19, \*), 1)  ((19, 80), 1)  ((19, \*), 1)  ((19, 18), 1)  ((19, \*), 1)  ((15, 80), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((15, 19), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((80, 18), 1)  ((80, \*), 1)  ((80, 19), 1)  ((80, \*), 1)  ((80, 18), 1)  ((80, \*), 1)  ((18, 19), 1)  ((18, \*), 1)  ((19, 18), 1)  ((19, \*), 1) |
| ((17, 15), 1)  ((17, \*), 1)  ((17, 80), 1)  ((17, \*), 1)  ((17, 18), 1)  ((17, \*), 1)  ((17, 19), 1)  ((17, \*), 1)  ((17, 18), 1)  ((17, \*), 1)  ((15, 80), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((15, 19), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((80, 18), 1)  ((80, \*), 1)  ((80, 19), 1)  ((80, \*), 1)  ((80, 18), 1)  ((80, \*), 1)  ((18, 19), 1)  ((18, \*), 1)  ((19, 18), 1)  ((19, \*), 1) | ((18, 15), 1)  ((18, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((15, 88), 1)  ((15, \*), 1)  ((15, 18), 1)  ((15, \*), 1)  ((18, 88), 1)  ((18, \*), 1)  ((88, 18), 1)  ((88, \*), 1) |
| Reducer Input | Reducer Output |
| ((12, \*), [1, 1, 1])  ((12, 18), [1])  ((12, 19), [1])  ((12, 80), [1])  ((15, \*), [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1])  ((15, 12), [1])  ((15, 18), [1, 1, 1, 1, 1, 1, 1, 1])  ((15, 19, [1, 1, 1])  ((15, 80), [1, 1, 1, 1])  ((15, 88), [1])  ((15, 91), [1])  ((17, \*), [1, 1, 1, 1, 1])  ((17, 15), [1])  ((17, 18), [1, 1])  ((17, 19), [1])  ((17, 80), [1])  ((18, \*), [1, 1,1 ,1])  ((18, 15), [1])  ((18, 19), [1, 1])  ((18, 88), [1])  ((19, \*), [1, 1, 1, 1, 1, 1, 1])  ((19, 15), [1])  ((19, 18), [1, 1, 1, 1])  ((19, 80), [1, 1])  ((80, \*), [1, 1, 1, 1, 1, 1, 1, 1, 1])  ((80, 12), [1])  ((80, 18), [1, 1, 1, 1, 1])  ((80, 19), [1, 1, 1])  ((88, \*), [1])  ((88, 18), [1])  ((91, \*), [1, 1, 1, 1, 1])  ((91, 12), [1])  ((91, 18), [1])  ((91, 19), [1])  ((91, 80), [1, 1]) | Sum 3  ((12, 18), 1/3)  ((12, 19), 1/3)  ((12, 80), 1/3)  Sum 18  ((15, 12), 1/18)  ((15, 18), 8/18)  ((15, 19), 3/18)  ((15, 80), 4/18)  ((15, 88), 1/18)  ((15, 91), 1/18)  Sum 5  ((17, 15), 1/5)  ((17, 18), 2/5)  ((17, 19), 1/5)  ((17, 80), 1/5)  Sum 4  ((18, 15), 1/4)  ((18, 19), 2/4)  ((18, 88), 1/4)  Sum 7  ((19, 15), 1/7)  ((19, 18), 4/7)  ((19, 80), 2/7)  Sum 9  ((80, 12), 1/9)  ((80, 18), 5/9)  ((80, 19), 3/9)  Sum 1  ((88, 18), 1)  Sum 5  ((91, 12), 1/5)  ((91, 18), 1/5)  ((91, 19), 1/5)  ((91, 80), 2/5) |

2. Stripe Approach

|  |  |
| --- | --- |
| W(X) 0 | W(X) 1 |
| W(15) = {91, 80, 12, 19, 80, 18}  W(91) = {80, 12, 19, 80, 18}  W(80) = {12, 19}  W(12) = {19, 80, 18}  W(19) = {80, 18}  W(80) = {18}  W(17) = {15, 80, 18, 19, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = { 18 } | W(19) = {15, 80, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = {18}  W(18) = {15}  W(15) = { 18, 18, 88, 18}  W(18) = {88}  W(88) = {18} |
| Mapper 0 Output | Mapper 1 Output |
| (15, [(91, 1), (80, 2), (12, 1), (19, 1), (18, 1)])  (91, [(80, 2), (12, 1), (19, 1), (18, 1)])  (80, [(12, 1), (19, 1)])  (12, [(19, 1), (80, 1), (18, 1)])  (19, [(80, 1), (18, 1)])  (80, [(18, 1)])  (17, [(15, 1), (80, 1), (18, 2), (19, 1)])  (15, [(80, 1), (18, 2), (19, 1)])  (80, [(18, 2), (19, 1)])  (18, [(19, 1)])  (19, [(18, 1)]) | (19, [(15, 1), (80, 1), (18, 1)])  (15, [(80, 1), (18, 2), (19, 1)])  (80, [(18,2), (19, 1)])  (18, [(19, 1)])  (19, [(18, 1)])  (18, [(15, 1)])  (15, [(18, 3), (88,1)])  (18, [(88,1)])  (88, [(18,1)]) |
| Reducer Input | Reducer Output |
| (12, [[(19, 1), (80, 1), (18, 1)]])  (15, [[(91, 1), (80, 2), (12, 1), (19, 1), (18, 1)], [(80, 1), (18, 2), (19, 1)], [(80, 1), (18, 2), (19, 1)], [(18, 3), (88,1)]])  (17, [[(15, 1), (80, 1), (18, 2), (19, 1)]])  (18, [[(19, 1)], [(19, 1)], [(88, 1)]])  (19, [[(80, 1), (18, 1)], [(18, 1)], [(15, 1), (80, 1), (18, 1)], [(18, 1)]])  (80, [[(12, 1), (19, 1)], [(18, 1)], [(18, 2), (19, 1)], [(18,2), (19, 1)]])  (88, [[(18,1)]])  (91, [[(80, 2), (12, 1), (19, 1), (18, 1)]]) | (12, [19/117, 80/117, 18/117])  (15, [91/712, 320/712, 12/712, 57/712, 144/712, 88/712])  (17, [15/150, 80/150, 36/150, 19/150])  (18, [38/126, 88/126])  (19, [160/247, 72/247, 15/247])  (80, [12/159, 57/159, 90/159])  (88, [1])  (91, [160/209, 12/209, 19/209, 18/209]) |

3. In-Mapper Combining Pair Approach

|  |  |
| --- | --- |
| W(X) 0 | W(X) 1 |
| W(15) = {91, 80, 12, 19, 80, 18}  W(91) = {80, 12, 19, 80, 18}  W(80) = {12, 19}  W(12) = {19, 80, 18}  W(19) = {80, 18}  W(80) = {18}  W(17) = {15, 80, 18, 19, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = { 18 } | W(19) = {15, 80, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = {18}  W(18) = {15}  W(15) = { 18, 18, 88, 18}  W(18) = {88}  W(88) = {18} |
| Mapper 0 Output | Mapper 1 Output |
| ((15, 91), 1)  ((15, \*), 1)  ((15, 80), 3)  ((15, \*), 3)  ((15, 12), 1)  ((15, \*), 1)  ((15, 19), 2)  ((15, \*), 2)  ((15, 18), 3)  ((15, \*), 3)  ((91, 80), 2)  ((91, \*), 2)  ((91, 12), 1)  ((91, \*), 1)  ((91, 19), 1)  ((91, \*), 1)  ((91, 18), 1)  ((91, \*), 1)  ((80, 12), 1)  ((80, \*), 1)  ((80, 19), 2)  ((80, \*), 2)  ((12, 19), 1)  ((12, \*), 1)  ((12, 80), 1)  ((12, \*), 1)  ((12, 18), 1)  ((12, \*), 1)  ((19, 80), 1)  ((19, \*), 1)  ((19, 18), 2)  ((19, \*), 2)  ((17, 15), 1)  ((17, \*), 1)  ((17, 80), 1)  ((17, \*), 1)  ((17, 18), 1)  ((17, \*), 1)  ((17, 19), 1)  ((17, \*), 1)  ((17, 18), 1)  ((17, \*), 1)  ((80, 18), 3)  ((80, \*), 3)  ((18, 19), 1)  ((18, \*), 1) | ((19, 15), 1)  ((19, \*), 1)  ((19, 80), 1)  ((19, \*), 1)  ((19, 18), 2)  ((19, \*), 2)  ((15, 80), 1)  ((15, \*), 1)  ((15, 18), 5)  ((15, \*), 5)  ((15, 19), 1)  ((15, \*), 1)  ((80, 18), 2)  ((80, \*), 2)  ((80, 19), 1)  ((80, \*), 1)  ((18, 19), 1)  ((18, \*), 1)  ((18, 15), 1)  ((18, \*), 1)  ((15, 88), 1)  ((15, \*), 1)  ((18, 88), 1)  ((18, \*), 1)  ((88, 18), 1)  ((88, \*), 1) |
| Reducer Input | Reducer Output |
| ((12, \*), [1, 1, 1])  ((12, 18), [1])  ((12, 19), [1])  ((12, 80), [1])  ((15, \*), [1,3,1,2,3, 1, 5,1,1])  ((15, 12), [1])  ((15, 18), [3, 5])  ((15, 19, [2, 1])  ((15, 80), [3, 1])  ((15, 88), [1])  ((15, 91), [1])  ((17, \*), [1, 1, 1, 1, 1])  ((17, 15), [1])  ((17, 18), [1, 1])  ((17, 19), [1])  ((17, 80), [1])  ((18, \*), [1, 1,1 ,1])  ((18, 15), [1])  ((18, 19), [1, 1])  ((18, 88), [1])  ((19, \*), [1, 2, 1, 1, 2])  ((19, 15), [1])  ((19, 18), [2, 2])  ((19, 80), [1, 1])  ((80, \*), [1, 2, 3, 2, 1])  ((80, 12), [1])  ((80, 18), [3, 2])  ((80, 19), [2, 1])  ((88, \*), [1])  ((88, 18), [1])  ((91, \*), [2, 1, 1, 1])  ((91, 12), [1])  ((91, 18), [1])  ((91, 19), [1])  ((91, 80), [2]) | Sum 3  ((12, 18), 1/3)  ((12, 19), 1/3)  ((12, 80), 1/3)  Sum 18  ((15, 12), 1/18)  ((15, 18), 8/18)  ((15, 19), 3/18)  ((15, 80), 4/18)  ((15, 88), 1/18)  ((15, 91), 1/18)  Sum 5  ((17, 15), 1/5)  ((17, 18), 2/5)  ((17, 19), 1/5)  ((17, 80), 1/5)  Sum 4  ((18, 15), 1/4)  ((18, 19), 2/4)  ((18, 88), 1/4)  Sum 7  ((19, 15), 1/7)  ((19, 18), 4/7)  ((19, 80), 2/7)  Sum 9  ((80, 12), 1/9)  ((80, 18), 5/9)  ((80, 19), 3/9)  Sum 1  ((88, 18), 1)  Sum 5  ((91, 12), 1/5)  ((91, 18), 1/5)  ((91, 19), 1/5)  ((91, 80), 2/5) |

4. In-Mapper Combining Stripe Approach

|  |  |
| --- | --- |
| W(X) 0 | W(X) 1 |
| W(15) = {91, 80, 12, 19, 80, 18}  W(91) = {80, 12, 19, 80, 18}  W(80) = {12, 19}  W(12) = {19, 80, 18}  W(19) = {80, 18}  W(80) = {18}  W(17) = {15, 80, 18, 19, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = { 18 } | W(19) = {15, 80, 18}  W(15) = {80, 18, 19, 18}  W(80) = {18, 19, 18}  W(18) = {19}  W(19) = {18}  W(18) = {15}  W(15) = { 18, 18, 88, 18}  W(18) = {88}  W(88) = {18} |
| Mapper 0 Output | Mapper 1 Output |
| (15, [(91, 1), (80, 3), (12, 1), (19, 2), (18, 3)])  (91, [(80, 2), (12, 1), (19, 1), (18, 1)])  (80, [(12, 1), (19, 2), (18, 3)])  (12, [(19, 1), (80, 1), (18, 1)])  (19, [(80, 1), (18, 2)])  (17, [(15, 1), (80, 1), (18, 2), (19, 1)])  (18, [(19, 1)]) | (19, [(15, 1), (80, 1), (18, 2)])  (15, [(80, 1), (18, 5), (19, 1), (88, 1)])  (80, [(18,2), (19, 1)])  (18, [(19, 1), (15, 1), (88, 1)])  (88, [(18,1)]) |
| Reducer Input | Reducer Output |
| (12, [[(19, 1), (80, 1), (18, 1)]])  (15, [[(91, 1), (80, 3), (12, 1), (19, 2), (18, 3)], [(80, 1), (18, 5), (19, 1), (88, 1)]])  (17, [[(15, 1), (80, 1), (18, 2), (19, 1)]])  (18, [[(19, 1)], [(19, 1), (15, 1), (88, 1)]])  (19, [[(80, 1), (18, 2)], [(15, 1), (80, 1), (18, 2)]])  (80, [[(12, 1), (19, 2), (18, 3)], [(18,2), (19, 1)]])  (88, [[(18,1)]])  (91, [[(80, 2), (12, 1), (19, 1), (18, 1)]]) | (12, [(19, 1), (80, 1), (18, 1)])  (15, [(91, 1), (80, 4), (12, 1), (19, 3), (18, 8), (88, 1)])  (17, [(15,1), (80, 1), (18, 2), (19, 1)])  (18, [(19,2), (15,1), (88, 1)])  (19, [(80,2), (18, 4), (15, 1)])  (80, [(12,1), (19,3), (18,5)])  (88, [(18,1)])  (91, [(80,2), (12,1),(19,1),(18,1)]) |