

1)

a)

$$T = \prod_{i=1}^6 (L_i + 1) = (1+1)^6 = 2^6 = 64$$

b) 157 Aggregate Cells Where Count >= 1

A1	A2	A3	A4	A5	A6								
B1	B2												
C1	C2												
<hr/>													
A1	A2	2^2											
*	*												
<hr/>													
B1	B2	2^2	A3	A4	A5	A6	2^4						
*	*		*	*	*	*	2^4						
<hr/>													
C1	C2	2^2											
*	*												
<hr/>													
<ul style="list-style-type: none"> - Subtract 2 from the $3 * 2^2$ calculation because *, * is counted 3 times - Subtract 3 from the end result to remove the base cells 													
$(3 * 2^2 - 2) * 2^4 = (3 * 4 - 2) * 16 = (12 - 2) * 16 = 10 * 16 = 160 - 3 = 157$													

c) Aggregated cells where count ≥ 3

*	*	A3	A4	A5	A6	1 * 1 * 2^4 = 32
		*	*	*	*	

Only *, * in the first 2 dimensions result the counts from all 3 base cells being aggregated

d) 4 Non star dimensions

*	*	A3	A4	A5	A6	
		*	*	*	*	

2)

a) There are 36 cuboids in the cube

Levels	Dimensions			
	State	Category	Price	Year
City				Quarter

The equation used for this cube is

$$T = \prod_{i=1}^4 (L + 1)$$

$$T = (2 + 1) * (1 + 1) * (1 + 1) * (2 + 1)$$

$$T = 3 * 2 * 2 * 3 = 6 * 6 = 36$$

2b) 56 distinct cells

2c) 34 distinct cells

2d) 33 distinct cells

2e) 1 distinct cell

2f) 1 distinct cell

3)

a) Each side is broken up into 3 partitions. The sizes are A = 900, B = 300, and C = 600. Each partition is reduced to 1/3 of its full value, meaning A_s = 300, B_s = 100, C_s = 200.

$$AB \text{ Plane} = 300 \times 100 = 3 * 10^4$$

$$AC \text{ Plane} = 300 \times 200 = 6 * 10^4$$

$$BC \text{ Plane} = 100 \times 200 = 2 * 10^4$$

If we count 1, 2, 3, 4, 5, 6..., we will have to hold the follow amounts of each plane in memory.

$$AB = 9$$

$$AC = 3$$

$$BC = 1$$

$$AB = 9 * 3 * 10^4 = 27 * 10^4$$

$$AC = 3 * 6 * 10^4 = 18 * 10^4$$

$$BC = 1 * 2 * 10^4 = 2 * 10^4$$

$$\text{Total} = (27 + 18 + 2) * 10^4 = 48 * 10^4$$

b) If we count 1, 4, 7, 10, 13, 16, 19, 22, 25, the following amounts of each plane are held in memory.

$$AB = 3$$

$$AC = 1$$

$$BC = 9$$

$$AB = 3 * 3 * 10^4 = 9 * 10^4$$

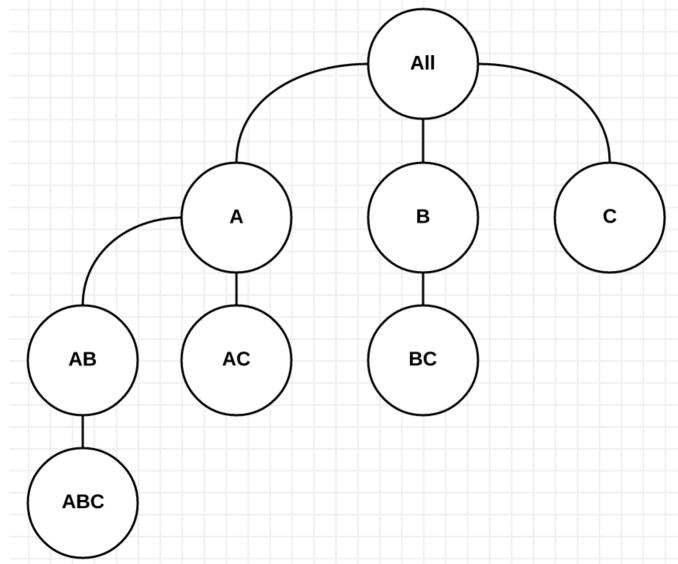
$$AC = 1 * 6 * 10^4 = 6 * 10^4$$

$$BC = 9 * 2 * 10^4 = 18 * 10^4$$

$$\text{Total} = (9 + 6 + 18) * 10^4 = 33 * 10^4$$

4)

a)



b)

$\text{min_support} = 4$. Expand if cell satisfies this criterion and stop if it is below this criterion.

$(*, *, *): 12$ - Expansion

$(A0, *, *): 12$ - Expansion

$(*, B0, *): 3$

$(*, B1, *): 3$

$(*, B2, *): 3$

$(*, B3, *): 3$

$(*, *, C0): 4$

$(*, *, C1): 4$

$(*, *, C2): 4$

$(A0, B0, *): 3$

$(A0, B1, *): 3$

$(A0, B2, *): 3$

$(A0, B3, *): 3$

$(A0, *, C0): 4$

$(A0, *, C1): 4$

$(A0, *, C2): 4$

16 Cells will be computed/considered

c)

$(*, *, *): 12$ - Expansion

$(B0, *, *): 3$

$(B1, *, *): 3$
 $(B2, *, *): 3$
 $(B3, *, *): 3$
 $(*, A0, *): 12$
 $(*, A0, C0): 4$
 $(*, A0, C1): 4$
 $(*, A0, C2): 4$
 $(*, *, C0): 4$
 $(*, *, C1): 4$
 $(*, *, C2): 4$

12 cells will be computed/considered

5)

- a) False: Operational update is a transactional concern and data warehouses are only concerned with loading data and accessing it
- b) False: Child cells are more specific than parent cells and B is more specific than A
- c) False: Drilling down shows more detailed information. Rolling up generalizes more information by aggregating 1 or more dimensions
- d) True: BUC is used to compute iceberg cubes that meet a minimum support requirement, which are a type of partial cube
- e) False: It is most effective when the product is moderate. If it is too high, the amount of data to hold in memory is too high.

6)

- a) Dice out a cube that contains only aggregated data for Europe, in the first 3 quarters of 2012, in the sports category. Then drill down the sports category to show each individual sports category

Dice: Region – Europe, Quater – 2012:1;2012:2,2012:3, Product - Sports

Drilldown: Product: *All Sports Products

Highest Revenue = Mountain Bike: \$975.74

Lowest Revenue = Termic Jacket: \$126.00



b) Slice a subset that only has aggregated data from North America. Drill down on browser and source.

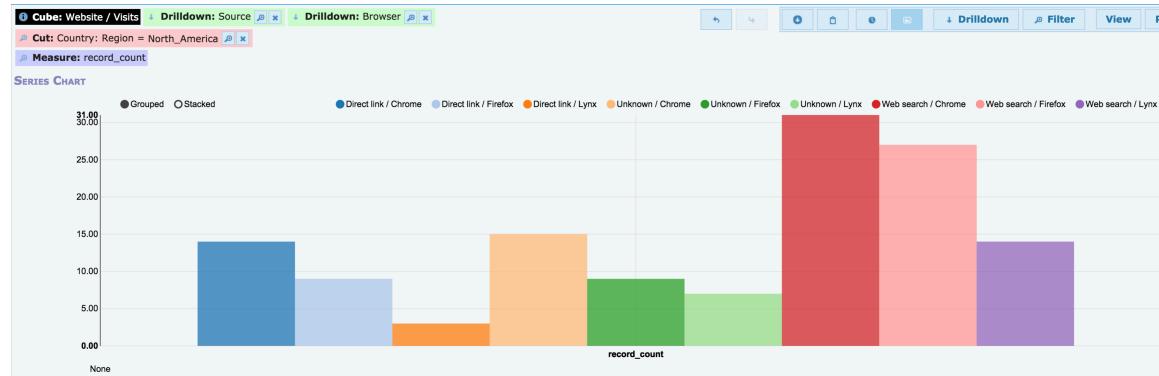
Slice: Region - North America

Drilldown: Source

Drilldown: Browser

Measure: visit_count

Most Visits: Web Search via Chrome – 31 Visits

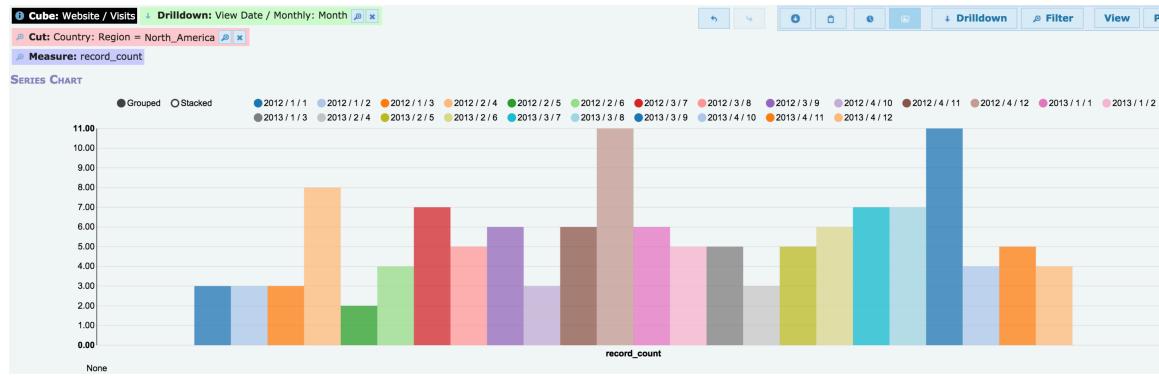


c)

Slice: Region - North America

Drilldown: View Date – Month

Measure: visit_count



d)

Webshop/Sales – This chart shows a comparison of total products sold in South American countries in Q4 between 2012 and 2013. This can help define what regions are increasing or decreasing performance and where products should be going.

Dice: 2012 – Q4, 2013 – Q4, Region – South America

Drilldown: Region – South America

Drilldown: Date – By year and quarter

Measure: quantity_sum

Horizontal Axis: year and quarter (Q4 only)



Website/Visits – This graph shows the comparison of total page views in the 4th quarter in the South American countries between 2012 and 2013. This can help the own decide where to focus marketing if one region is stronger than the other if they are trying to grow or expand in a region.

Dice: 2012 – Q4, 2013 – Q4, Region – South America

Drilldown: Region – South America

Drilldown: Date – By year and quarter

Measure: page_views

Horizontal Axis: year and quarter (Q4 only)

