

3. The most detailed grain is the combination of individual product or service, individual customer, and date (for special events, only customer and date).

- ☐ 50000 members: sum of member rows
- ☐ 350 franchises: sum of franchises
- ☐ 450,000 items sold merchandises (Contains rows) per year
- ☐ 500 Unique merchandise items
- ☐ 100,000 ServicePurchase rows per year
- ☐ 20 Unique ServCategory rows
- ☐ 300 SpecialEvents Worksheet rows per year per franchise with 200 franchises using this spreadsheet
- ☐ 150 unique customers per special event worksheet
- ☐ Merchandise Product sales(item level): 450,000
- ☐ Days per year: 365
- ☐ Customer number (product) = 50000
- ☐ Customer number (service) = 50000
- ☐ Customer number (special event) = $200 \times 150 = 30000$
- ☐ Fact table size (merchandise product sales) is determined - 450000 purchases per year (including merchandise product)
- ☐ Fact table size (service sales) is determined - 100000 purchases per year (including service)
- ☐ Fact table size (special event sales) is determined - $300 \times 200 = 60000$ purchases per year (including special events)
- ☐ Sparsity estimate:
 - o $1 - (\text{fact table size} / \text{product of dimensions})$
 - o $(1 - (450000 / (500 \times 50000 \times 365))) = 0.9995$
 - o The data cube has mostly missing cells with slightly more than 0.0005% of cells with non-zero values.

 - o $1 - (\text{fact table size} / \text{service of dimensions})$
 - o $(1 - (100000 / (20 \times 50000 \times 365))) = 0.997$
 - o The data cube has mostly missing cells with slightly more than 0.003% of cells with non-zero values.

 - o $1 - (\text{fact table size} / \text{special events of dimensions})$
 - o $(1 - (60000 / (30000 \times 365))) = 0.995$
 - o The data cube has mostly missing cells with slightly more than 0.005% of cells with non-zero values.