**Assignment 1 – Module 3**

**Problem 3:**

**ANS:**

**The most detailed grain is the combination of individual product, individual customer, and date**

* **50000 members: member rows(given)**
* **350 franchises: franchises(given)**
* **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**
* **Days per year: 365**
* **Customer number (product) = 50000**
* **Customer number (service) = 50000**
* **Customer number (special event) = 200\*150=30000**
* **Fact table size (merchandize 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\*200=60000 purchases per year (including special events)**
* **Sparsity estimate:** 
  + **1 - ( fact table size / product of dimensions )**
  + **(1 – ( 450000 / (500\*50000\*365) ) = 0.9995**
  + **The data cube has mostly missing cells with slightly more than 0.0005% of cells with non-zero values.**
  + **1 - ( fact table size / service of dimensions )**
  + **(1 – ( 100000 / (20\*50000\*365) ) = 0.997**
  + **The data cube has mostly missing cells with slightly more than 0.003% of cells with non-zero values.**
  + **1 - ( fact table size / special events of dimensions )**
  + **(1 – ( 60000 / (30000\*365) ) = 0.995**
  + **The data cube has mostly missing cells with slightly more than 0.005% of cells with non-zero values.**