PR3. Identify the grain in your dimensional design using the business needs as a guideline. You should then indicate relative storage requirements for the grain using the statistics for the data sources. Using the cardinality estimates provided, you should determine either the fact table size or sparsity and then compute the unknown grain size variable. For example, you should compute sparsity if the fact table size is given.

1. Dimensions

Franchise: 350 (ERD).
Customer: 50,000 (ERD) + (150/spreadsheet * 200 spreadsheets) = 80,000.
Sales Item: 500 (ERD – Merchandise) + 20 (ERD – ServCategory) + 1 (spreadsheet – Special Events) = 521.

2. Fact Table Size (per year)
   450,000 (ERD – Contains (merchandise)) + 100,000 (ERD – ServicePurchase) + 300 * 200 (spreadsheets) = 610,000

3. Sparsity estimate
   1 - (fact table size / product of dimensions)
   1-(610,000/(350 * 80,000 * 521 * 365) = 1-
   (610,000/10,220,000,000) = 1-0.0000596885 = 0.9999403