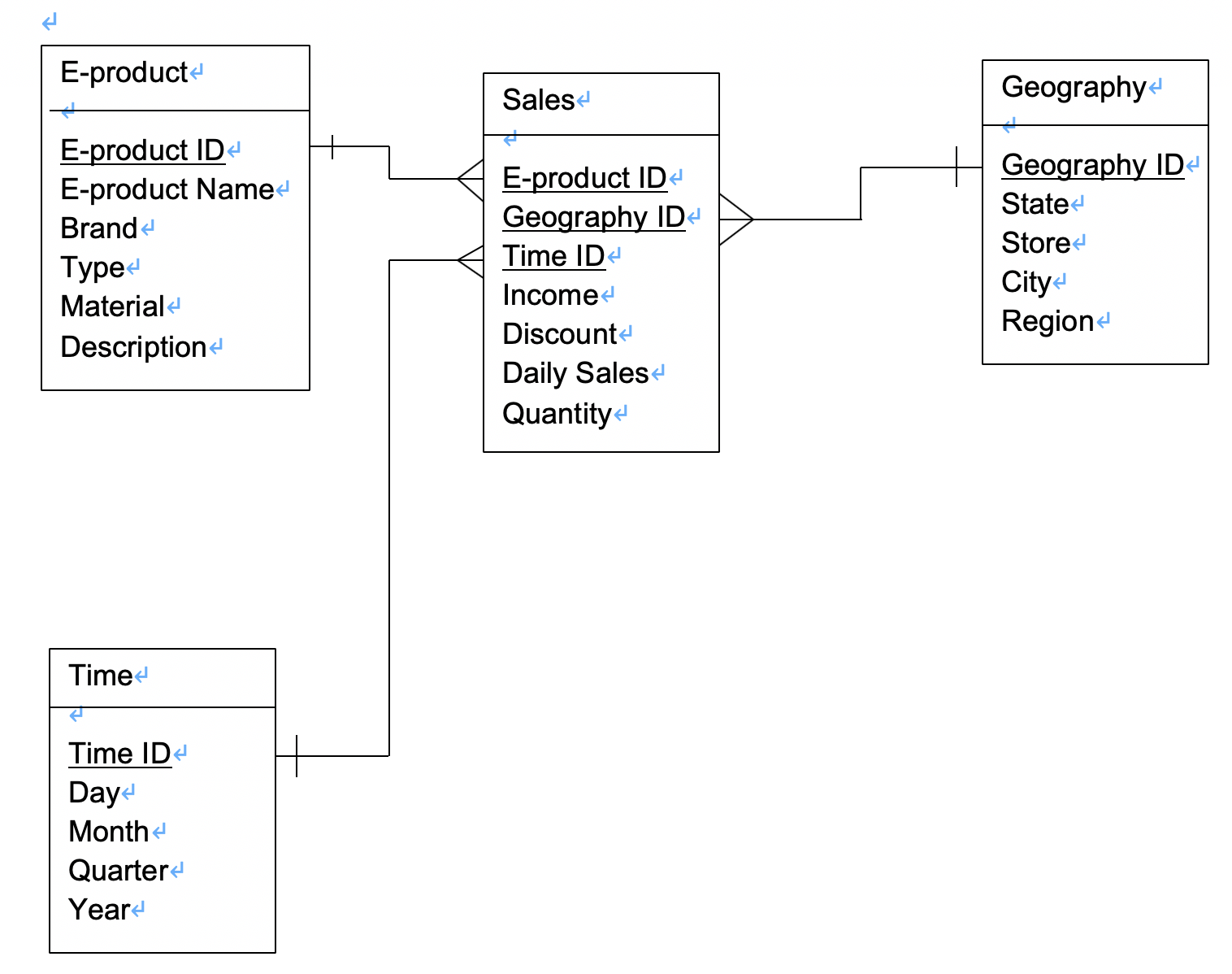
**Assignment 1**

Name：Wentao Ma

Student Number：13308679

**•Star Schema**



**•** Explain what your facts, measures, and dimensions are, and how they are linked with each other.

Fact: Sales. That is because the Sales is the focus which is closely related to decision-making. In addition, I identify the Sales table as the fact table.

Measures: I identify the income, discount and the daily sales of electronics products are the measures. Because these things are attributes that describe the sales which is identified as the fact.

Dimensions: There are three dimensions in this star schema, ‘e-product, geography and time’, respectively. They are all discrete attributes which determine the granularity adopted to represent the fact (sales).

These three dimension tables (E-product table, Geography table and Time table) are all linked to the fact table (Sales table) by using the different foreign keys. For example, the ‘E-product table’ is linked to the Sales table through the attribute ‘E-product ID’. There is a one-to-many relationship between them, which means that for different values in the ‘Sales table’, only one corresponding value appears in the ‘E-product table’. Similarly, ‘Geography ID’ links the ‘Geography table’ with the ‘Sales table’ as one-to many relationship. ‘Time ID’ links the ‘Time table’ with the ‘Sales table’ as one-to many relationship, too.

**• Produce a query which can be executed on your star schema and addresses the CEO’s requirement.**

SELECT E.E-product Name, SUM(S.Quantity)

FROM (

SELECT E.E-product Name, SUM(S.Quantity),

RANK() OVER (ORDER BY SUM(S.Quantity) DESC) AS Rank

FROM Sales S, E-product P, Geography G, Time T

WHERE S.E-product ID = P.E-product ID AND

S.Geography ID = G.Geography ID AND

S.Time ID = T.Time ID AND

T.Day = 26 AND

T.Month = ‘December’ AND

T.Year >= 2014 AND T.Year < 2019 AND

G.State = ‘NSW’)

WHERE Rank <= 9;

**• Provide explanations to your query regarding its execution.**

Above all, I use RANK () operation in the sub query to create a column named Rank by descending order, then I use an operation call SUM() to calculate the total sum quantity in last five boxing days. Next, I use 3 equations of IDs to link Sales table and the other 3 dimension tables. At the same time, I used G.State and T.Day/Month/Year to limited the location and time.

Secondly, the main query is for showing top 9 products’ names and quantities limited from the subquery. This whole query can finally meet the CEO’s requirement and tackle the given issue.