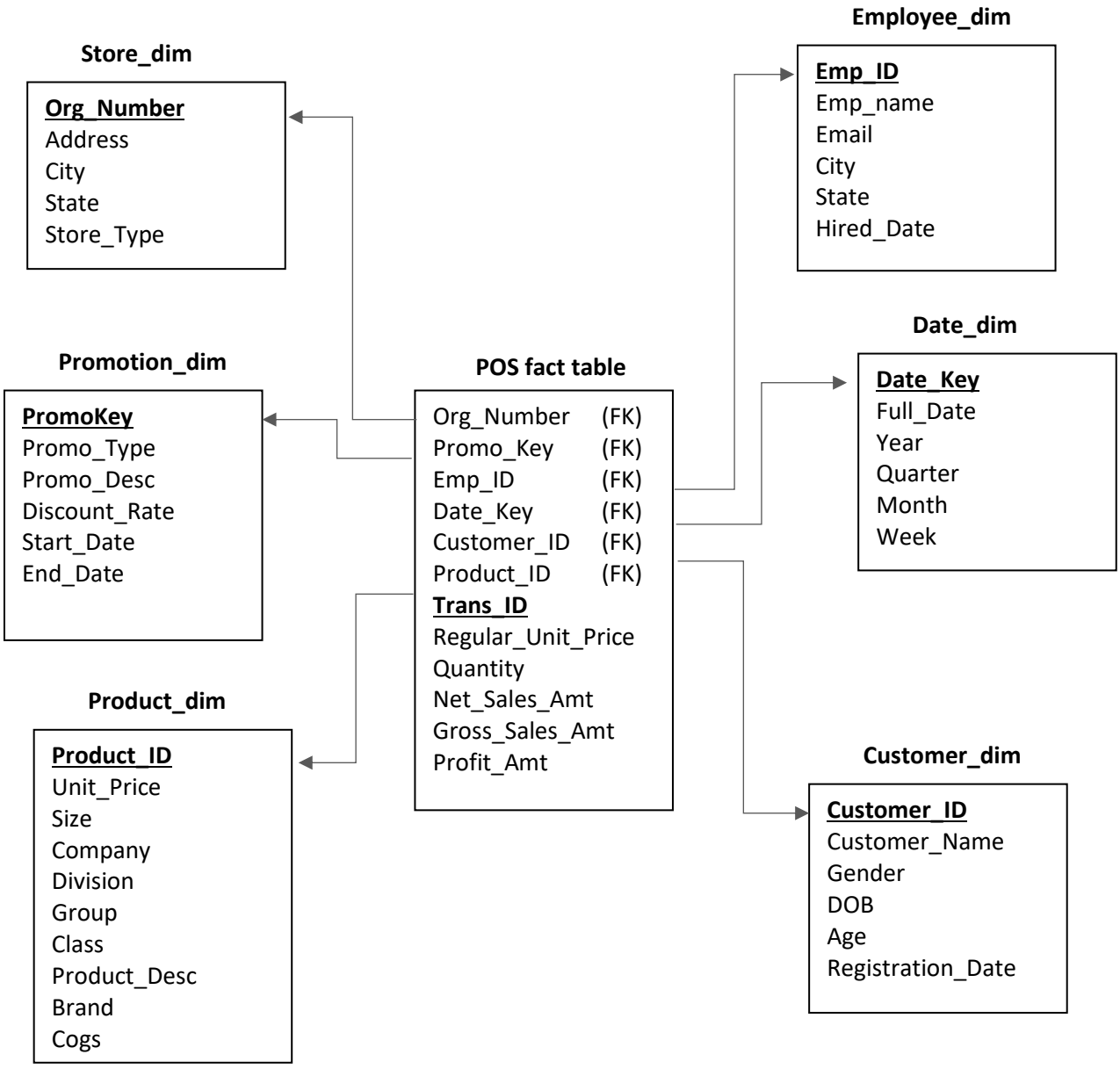
Retail Analytics: Data Modeling

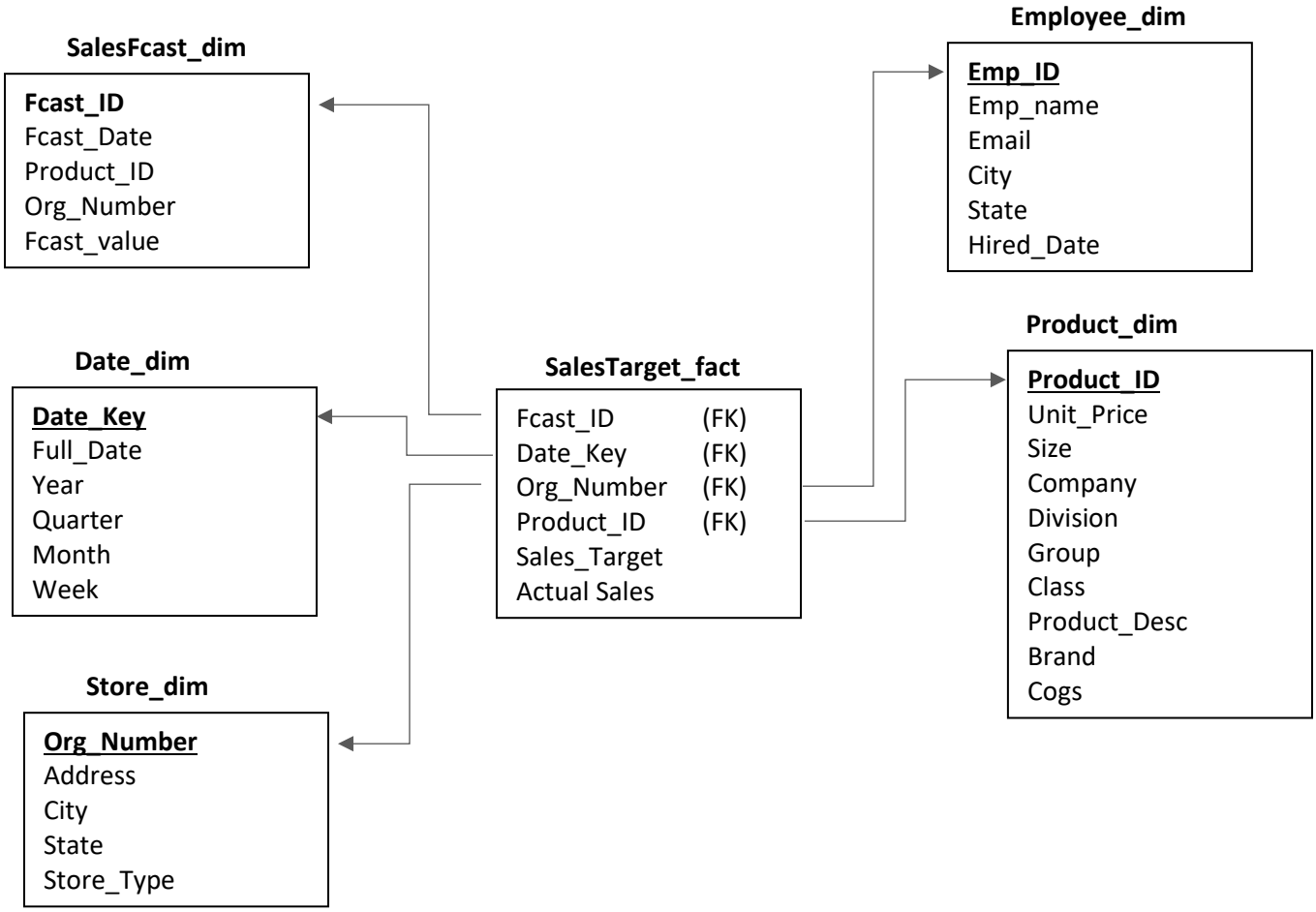
By Minal Singh, Priyanka Phalake, Soumya Tejas
University at Buffalo, The State University of New York

In order to meet the business requirements, we are designing two fact tables, and seven dimensions as per Kimball's approach, in the star-schema format.



Retail Analytics: Data Modeling

By Minal Singh, Priyanka Phalake, Soumya Tejas
University at Buffalo, The State University of New York



IT arranges the data in a way that it can satisfy specific business query. The steps involved are:

Select the business process: A major operational process that is supported by some kind of legacy system from which data can be collected. In this case, it is Pont of Sales(POS) and sales forecast.

Define grain: The fundamental level of data represented in a fact table for the selected process. In this case, it is each transaction, and each forecast.

Define dimension: Choose the dimensions that apply to each fact table record. Here, the dimensions selected are date, employee, customer, product, store, promotion, forecast.

DATE DIMENSION: This dimension table is a very familiar dimension across all the industry. Qualitative attributes bear more significance for this dimension (eg. Sales on weekends compared to sales on working days). We have included columns such as Holiday indicator, Weekend indicator, etc.

Retail Analytics: Data Modeling

By Minal Singh, Priyanka Phalake, Soumya Tejas
University at Buffalo, The State University of New York

DateKey
FullDate
Calender Year
Calender YearMonth
Calendar YearMonth Num
Calendar Semester
Calendar Quarter
Calendar Month Number
Month Long Name
Month Short Name
Calendar Week Number
Calendar Week Name
Calendar 2-Week
Day Number of Year
Day Number of Month
Day Number of Week
Day Name Long
Weekday Indicator
….
Fiscal Day Name of Year

PRODUCT DIMENSION: This dimension table includes all the details of the products available at each store, online or offline. It is used to describe the SKUs in a store and consists many descriptive attributes of each SKU. The analyst can browse through dimension attributes and we can roll up and drill down using attributes as constraints even if they don't belong to the merchandize hierarchy.

Category
Prod_ID
Product_Desc
Brand
Unit_Price
Size

CUSTOMER DIMENSION: This dimension would include all the details of the customers making purchases or returns at each store, online or offline. This would help us in more analysis of which customers purchase what types of products and which customers help in generating more profit for the organization.

Retail Analytics: Data Modeling

By Minal Singh, Priyanka Phalake, Soumya Tejas
University at Buffalo, The State University of New York

customer_id
first_name
last_name
date_of_birth
email
gender
phone
city
street_address
zip_code
state
state_abbr
country

STORE DIMENSION: This is the primary geographic dimension which describes every store in a retail chain. Stores can be rolled up using attributes like zip code, county, state, country, etc. Store dimension is mainly used in the corporate reporting and the necessary components of the store dimension from multiple operational sources assemble general assembled at headquarters.

ROW_WID
ORG_NUM
W_CURR_CODE
ST_ADDRESS
C_STATE_PROV_CODE
CITY_CODE
POSTAL_CODE
COUNTRY_REGION_NAME
COUNTRY_REGION_CODE
PHONE_NUM
WEB_ADDRESS
ORG_TYPE_CODE
MGR_NAME
W_COUNTRY_CODE
COUNTY_CODE
W_INSERT_DT
W_UPDATE_DT

Retail Analytics: Data Modeling

By Minal Singh, Priyanka Phalake, Soumya Tejas
University at Buffalo, The State University of New York

EMPLOYEE DIMENSION: This dimension table includes all the data corresponding to an employee of the organization. This might help us to get insight the employee strength at each store which might be the cause of an issue such as lesser sales which might have a solution in increasing the number of employees.

EMPLOYEE_ID
LNAME
FNAME
CITY
STATE
PHONE
SEX
SALARY
DATE_OF_BIRTH
DATE_OF_HIRED

PROMOTIONS DIMENSION: This dimension table includes all the details of the promotions available to the customers, their type, description and their validity. Common promotion activities are end of the season sales, provisional price reductions, festival offers, coupons, etc. Promotion effectiveness is a management concern for achieving volume and profitability.

PROMOKEY
PROMO_TYPE
PROMO_DESC
DISCOUNT_RATIO
START_DATE
END_DATE