Report

4 steps:

1. Select the business process

“BP of fixing income from customers”

1. Declare the Grain

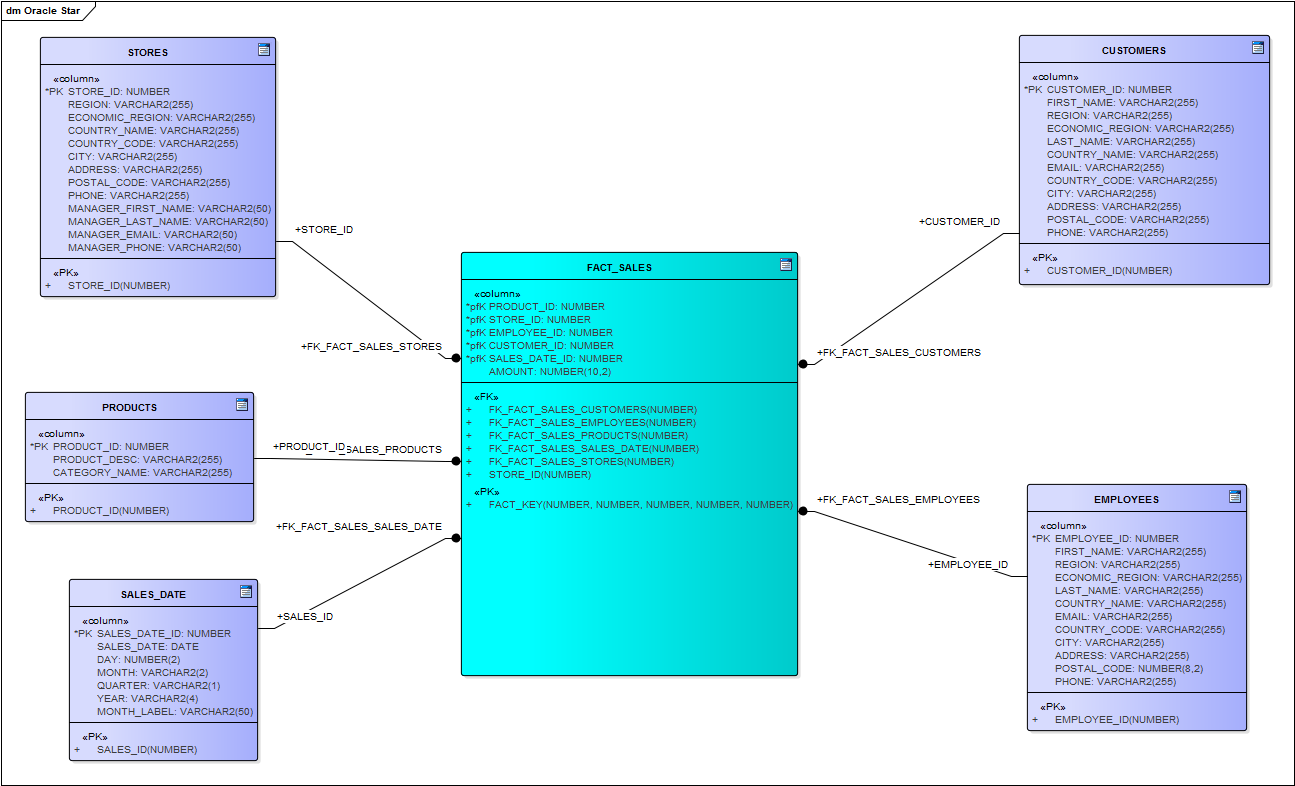
“Total amount of sale paid by customer in the appropriate office.”

1. Identify the Dimensions
2. Customers
3. Employees
4. Sales Date
5. Locations
6. Stores
7. Identify the Facts

“Total sales amount” as FACT\_SALES

It should be add that schemas were built with the help of SPARX Enterprise Architect 12 as it is one of the best tools to build database models. Not much attention was put on the length of the field types like it was presented in the task diagram.

**Star Schema:**



So as to put little description to my schemas, I’d like to add some comments:

* **dimensions “EMPLOYEES”, “CUSTOMERS”, “STORES”**

These 3 dimension have detailed information about entity’s addresses which was gained from “Geo Data” data.

* **dimension STORES**

Dimension “STORES” incudes information about not only about actual Store, but also about Manager of this store presented by fields “MANAGER\_NAME”, “MANAGER\_LAST\_NAME”, “MANAGER\_EMAIL”, “MANAGER\_PHONE”. These fields help to identify manager and there is no need to put such information as postal code, address, city, country, postal code to MANAGER because it would be exactly the same as in the STORE.

* **dimension SALES**

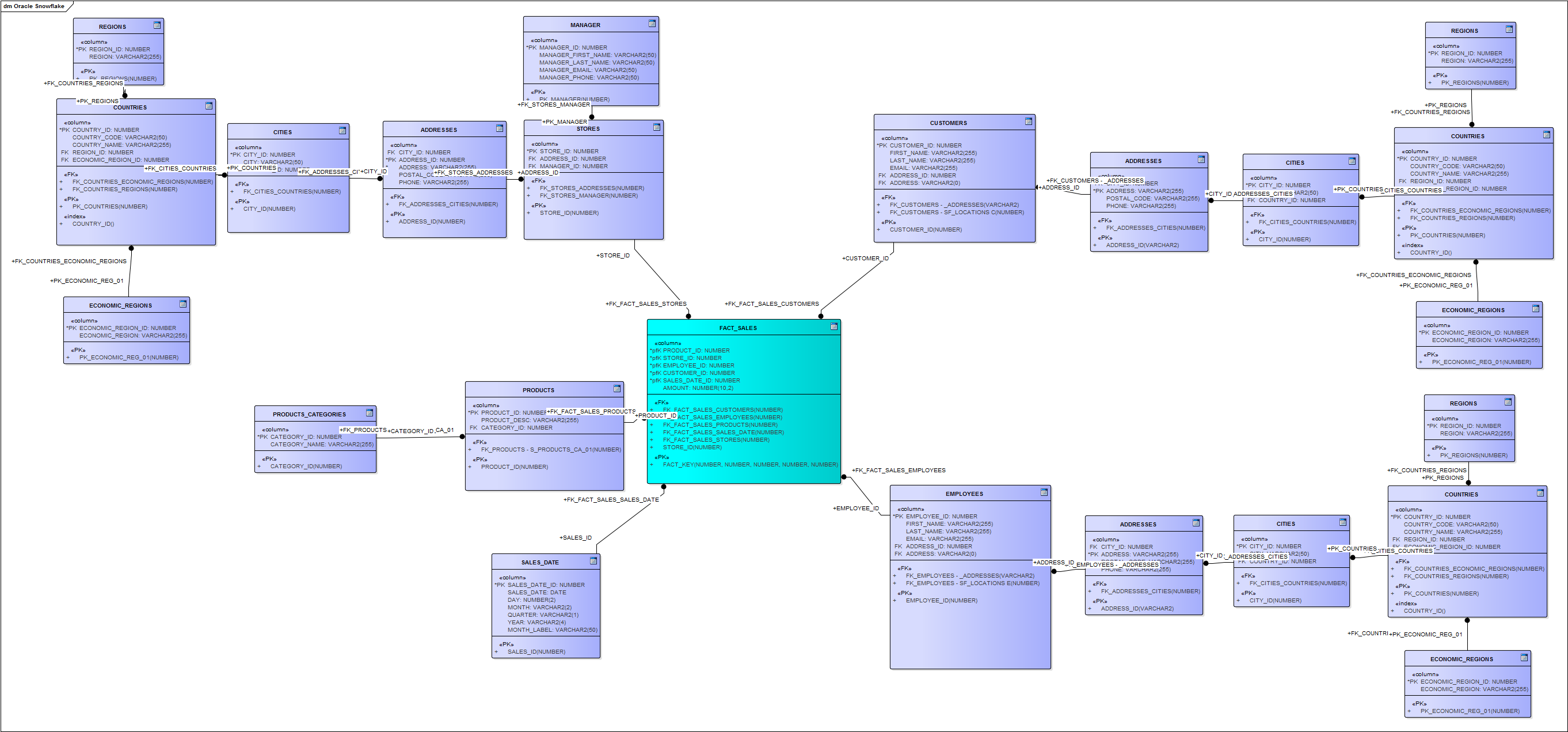
Dimension “SALES” presents data from SALES table divided by day, month, quarter, and year.

* **fact FACT\_SALES**

Fact table has a foreign key from every dimension which serve as a composite primary key for fact table.

One more thing about fact table is that it is composed from 2 tables: Payments and Sales, where there is no need to put Staff\_ID and Customer\_ID to the fact, because they are already in the SALES table (employee\_id, customer\_id). But there is one more field – AMOUNT that was put in the fact table and it should be generated as SUM of all PAYMENTS, because not only one payment can provide SALES AMOUNT, but few.

**Snowflake schema:**



*Snowflake diagram in .bpm can be found in “Reports” folder*

Some description about snowflake diagram:

* **fact table FACT\_SALES**

Fact table has no difference to Star Schema

* **dimensions “EMPLOYEES”, “CUSTOMERS”, “STORES”**

These dimensions were normalized through dividing them into 2: entity (presented by 1 table, *for example “CUSTOMERS”)* and it’s address (tables repeating from the Geo Data: *ADDRESSES, CITIES, …*). All of them have the same granularity and helps to provide schema with a 3rd normal form.

It should be said that it’s not the best way to store location for every entity (3 in our example) separately, but if we want to it more useful we should change our schema (for example to “galactic” schema).

* **dimension STORES**

dimension STORES was divided in one more table (compared to CUSTOMERS and EMPLOYEES) – MANAGER. MANAGER dimension stores information about manager in this store.

* **dimension PRODUCTS**

Dimension PRODUCTS was divided into 2 tables: Products, Product Category, so as to be normalized too.