

Exam simulation

1 Problem specification: Travel agency

A travel agency organizes guided trips for tourists of different nationalities. The agency wants to know the main trends about trips with respect to the characteristics of the trips and the type of participants. The following is the relational schema of the operational database.

TRIP (CodeT, CodeDestination, CodeCategory, CodeGuide, Date, Price)

DESTINATIONS (CodeD, Name, Description, Type, CodeNation)

TRAVELING (CodeT, CodeParticipant)

PARTICIPANT (CodeP, Name, Surname, Age, Address, City, CodeNation)

NATION (CodeN, Name, Continent)

GUIDE (CodeG, Name, Surname, Age, Address, Sex, City, CodeNation)

TYPE (CodeTy, TypeName, Description)

CATEGORY (CodeC, CategoryName, Description)

Goal: Design the data warehouse for the travel agency.

2 A possible solution

ER model

The ER model is reported in Figure 1.

Facts, measures, dimensions

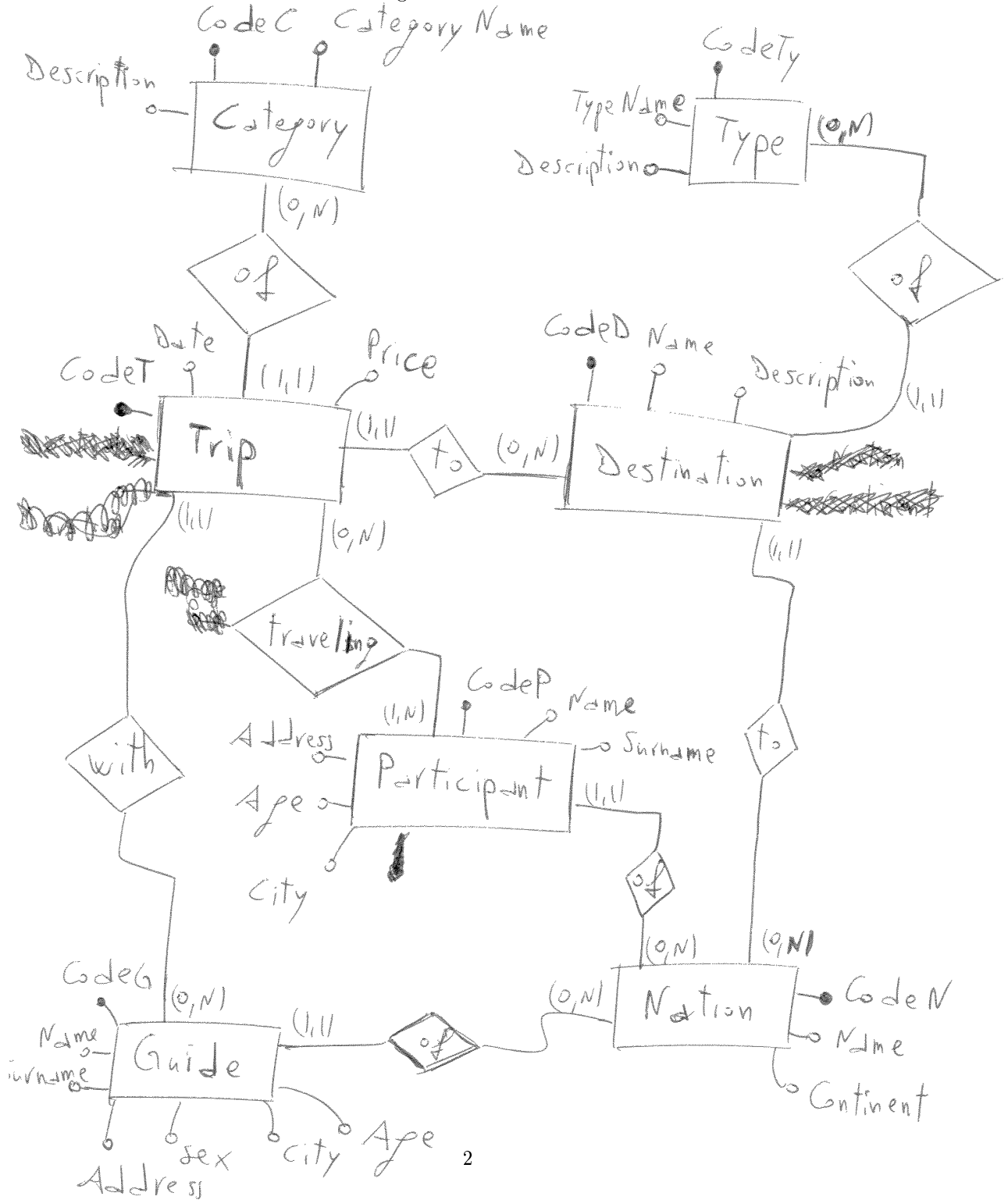
There is one fact: trips.

FACT trips

MEASURES NumberOfParticipant, NumberOfTrips, TotalIncome

DIMENSIONS Participant's Characteristics , Destination, Category, Guide,
Time

Figure 1: ER model



Attribute tree

The attribute tree associated to the "trips" fact is reported in Figure 2.

Fact model/fact schema

The fact model is reported in Figure 3.

Logical Snowflake schema

FACT_TRIPS (CodeC, CodeD, CodePT, CodeG, CodeT, NumTrips, NumParticipants, TotalIncome)

CATEGORY (CodeC, Name)

DESTINATION (CodeD, Name, TypeName, CodeN)

NATION (CodeN, Name, Continent)

PARTICIPANT_TYPE (CodePT, Age, City, CodeN)

GUIDE (CodeG, Age, Sex, City, CodeN)

TIME (CodeT, Date, DayOfTheWeek, Month, FourMonthPeriod, ThreeMonthPeriod, Semester, Year)

SQL queries

- Total income and total number of participants w.r.t. type of destination and year

```
SELECT D.TypeName, Ti.Year, SUM(FT.Income), SUM(FT.NumParticipants)
FROM FACT_TRIPS FT, Destination D, Time T
WHERE FT.CodeT = T.CodeT AND TF.CodeD = D.CodeD
GROUP BY D.TypeName, T.Year
```

- Total number of participants w.r.t. destination, semester and guide's nationality

```
SELECT D.Name, T.Semester, T.Year, G.Nation, SUM(FT.NumOfParticipants)
FROM FACT_TRIPS FT, Destination D, Time T, Guide G
WHERE FT.CodeT = T.CodeT
AND TF.CodeD = D.CodeD AND TF.CodeG = G.CodeG
GROUP BY D.Name, T.Semester, T.Year, G.Nation
```

- Average price per participant w.r.t. type of destination and year

```
SELECT D.TypeName, Ti.Year, SUM(FT.Income)/SUM(FT.NumOfParticipants)
FROM FACT_TRIPS FT, Destination D, Time T
WHERE FT.CodeT = T.CodeT AND TF.CodeD = D.CodeD
GROUP BY D.TypeName, T.Year
```

Figure 2: Attribute tree

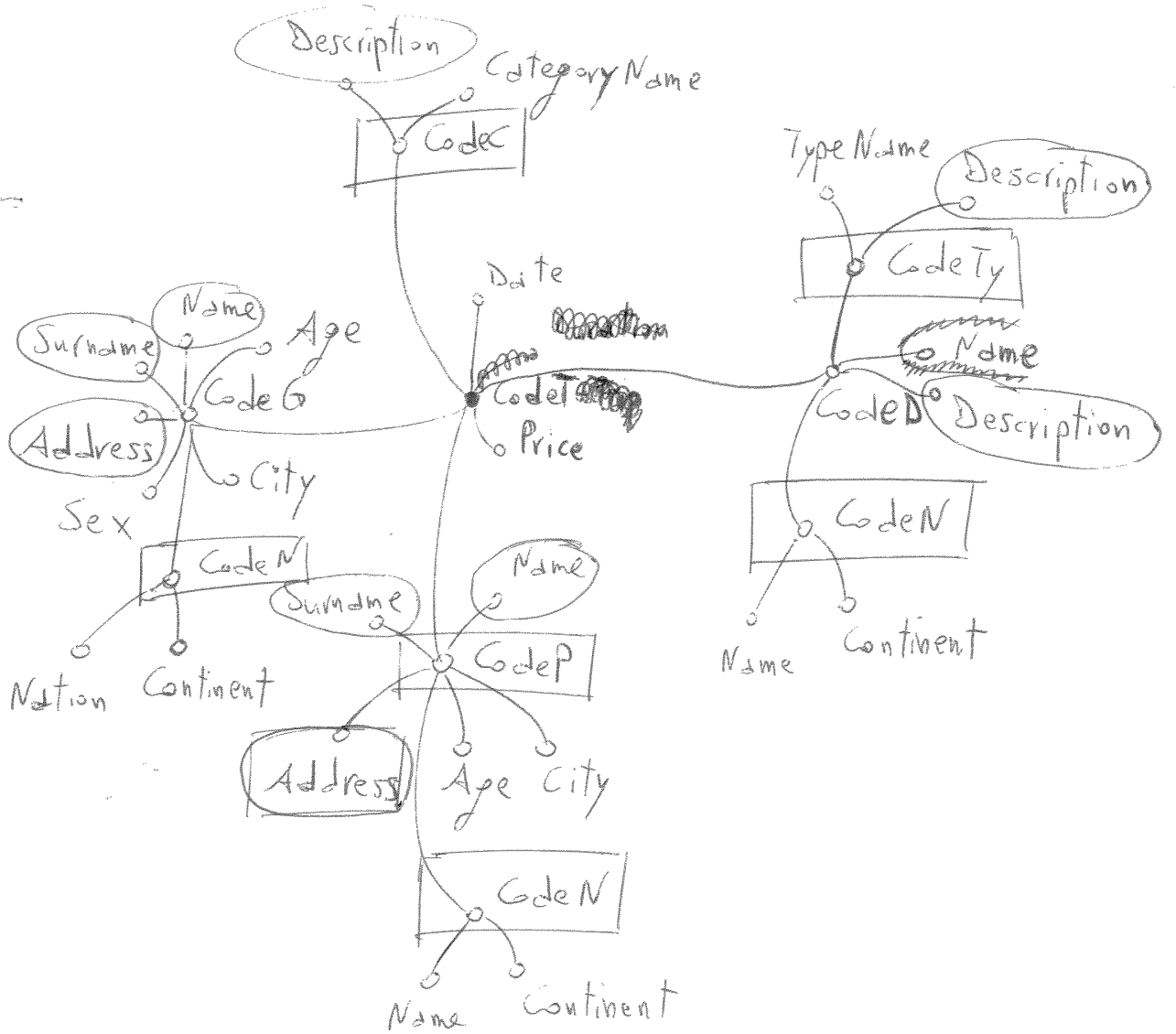
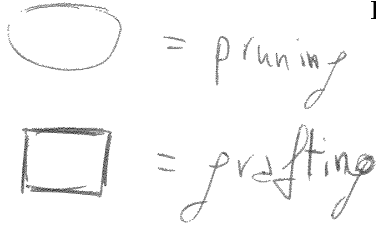
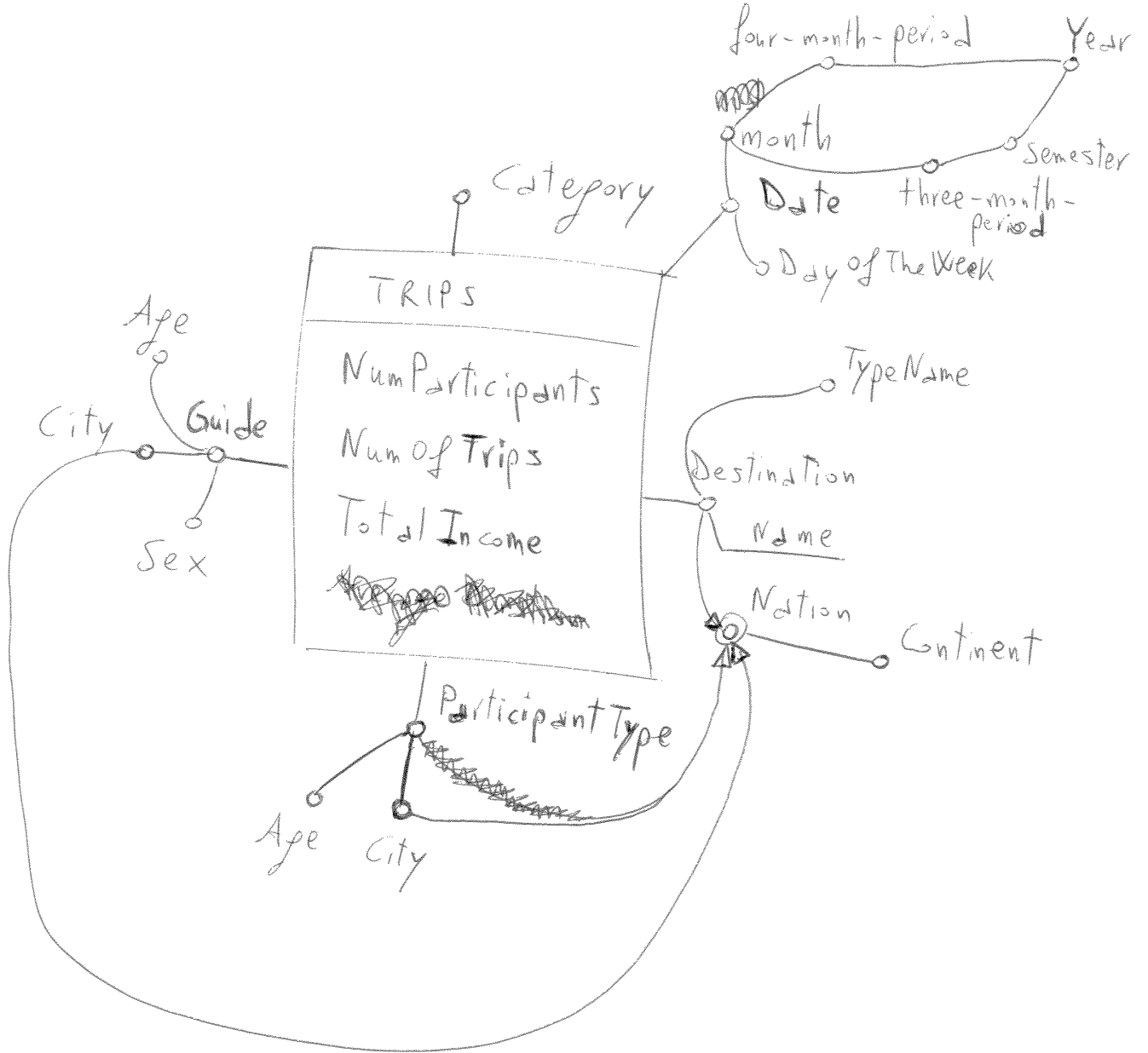


Figure 3: Fact model



Glossary.

NumParticipants → COUNT(*) from traveling, trip
 Num of Trips → COUNT(DISTINCT CODE) from trip, traveling
 Total Income → SUM(Price)
~~Price~~