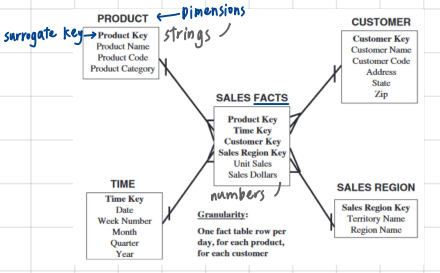
*Star schema

1. Process : business case
2. Grain : lowest level of detail
3. Dimensions
4. Facts
5. Duration



Business Subject: Hotel Occupancy
Grain: One row per specific flight (by a single aircraft from a specific departure to arrival airport)

Dimension Attributes	Dimensions			
	Date	Aircraft	Departure Airport	Arrival Airport
	Date	Aircraft ID	Airport name	Airport name
	Year	Aircraft Model	Airport city	Airport city
	Quarter	Aircraft Capacity	Airport country	Airport country
	Month			
	Day of Month			

Facts
Total Passengers
Total Revenue

*Types of Fact Table

	Transaction	Periodic Snapshot	Accumulating Snapshot
Business process modeled	Events ("income and expanse statement")	Periodic updates on status of events ("balance statement")	Elapsed time between events
Grain	One row per "event".	One row per "period" – time interval when the snapshots are taken	One row per "item" – thing being tracked through the business process
Primary key	Combination of foreign keys and/or a natural primary key (transaction #) and/or a surrogate primary key	Date (or appropriate time interval) and other foreign keys	ID of item going through business process (loan application #, help desk #, etc)
Fact types	Generally, fully additive except for "per-unit" facts (like unit price)	Balance/level facts are generally semi-additive, being non-additive across a date dimension	Generally, fully additive
Density	Sparse	Dense	Sparse

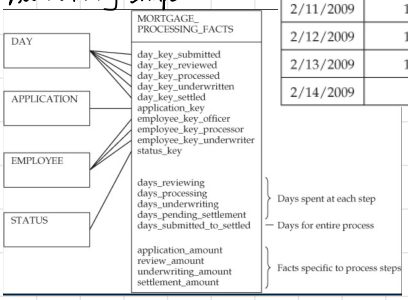
Transaction

Day	Transaction Type	Transaction Amount
2/1/2009	Initial Deposit	2000.00
2/2/2009	Withdrawal	(20.00)
2/3/2009	Check	(35.50)
2/3/2009	Check	(17.02)
2/6/2009	Check	(75.00)
2/6/2009	Deposit	75.00
2/7/2009	Check	(800.00)
2/10/2009	Check	(68.29)
2/14/2009	Withdrawal	(100.00)

Periodic Snapshot

Day	Balance
2/1/2009	2000.00
2/2/2009	1980.00
2/3/2009	1927.48
2/4/2009	1927.48
2/5/2009	1927.48
2/6/2009	1927.48
2/7/2009	1127.48
2/8/2009	1127.48
2/9/2009	1127.48
2/10/2009	1059.19
2/11/2009	1059.19
2/12/2009	1059.19
2/13/2009	1059.19
2/14/2009	959.19

Accumulating Snapshot



*Fact Additivity

- Additive e.g. sales revenue, number of unit sold
- Semi-additive e.g. account balance, inventory levels
- Non-additive e.g. interest rates, percentage margins

*Semi-structured data

Convert the following to JSON format:

Student ID	First Name	Last Name	Age	GPA
1	John	Doe	20	3.5
2	Alice	Smith	19	3.9
3	Bob	Johnson	21	3.2

dataframe ← [
row ← {
column names ← "student_id": 1, → value
"first_name": "John",
"last_name": "Doe",
"age": 20,
"gpa": 3.5
},
{
"student_id": 2,
"first_name": "Alice",
"last_name": "Smith",
"age": 19,
"gpa": 3.9
}
]

JSON Schema

```
{
   "$schema": "http://json-schema.org/draft-04/schema#",
   "title": "Product",
   "description": "A product from Acme's catalog",
   "type": "object",

   "properties": {

      "id": {
         "description": "The unique identifier for a product",
         "type": "integer"
      },

      "name": {
         "description": "Name of the product",
         "type": "string"
      },

      "price": {
         "type": "number",
         "minimum": 0,
         "exclusiveMinimum": true
      }
   },

   "required": ["id", "name", "price"]
}
```

JSON

```
[
   {
      "id": 2,
      "name": "An ice sculpture",
      "price": 12.50
   },

   {
      "id": 3,
      "name": "A blue mouse",
      "price": 25.50
   }
]
```

Name	Email	Phone	Street	City	State	Zip Code
John Doe	john@example.com	+12345678 90				
Alice Smith	alice@example.com					
Bob Johnson	bob@example.com		123 Main St	Anytown	CA	12345
Eve Brown	eve@example.com	+98765432 10	456 Elm St	Sometown	NY	54321

```
{
   "contacts": [
      {
         "name": "John Doe",
         "email": "john@example.com",
         "phone": "+12345678 90"
      },
      {
         "name": "Alice Smith",
         "email": "alice@example.com"
      },
      {
         "name": "Bob Johnson",
         "email": "bob@example.com",
         "address": {
            "street": "123 Main St",
            "city": "Anytown",
            "state": "CA",
            "zip_code": "12345"
         }
      },
      {
         "name": "Eve Brown",
         "email": "eve@example.com",
         "phone": "+98765432 10",
         "address": {
            "street": "456 Elm St",
            "city": "Sometown",
            "state": "NY",
            "zip_code": "54321"
         }
      }
   ]
}
```