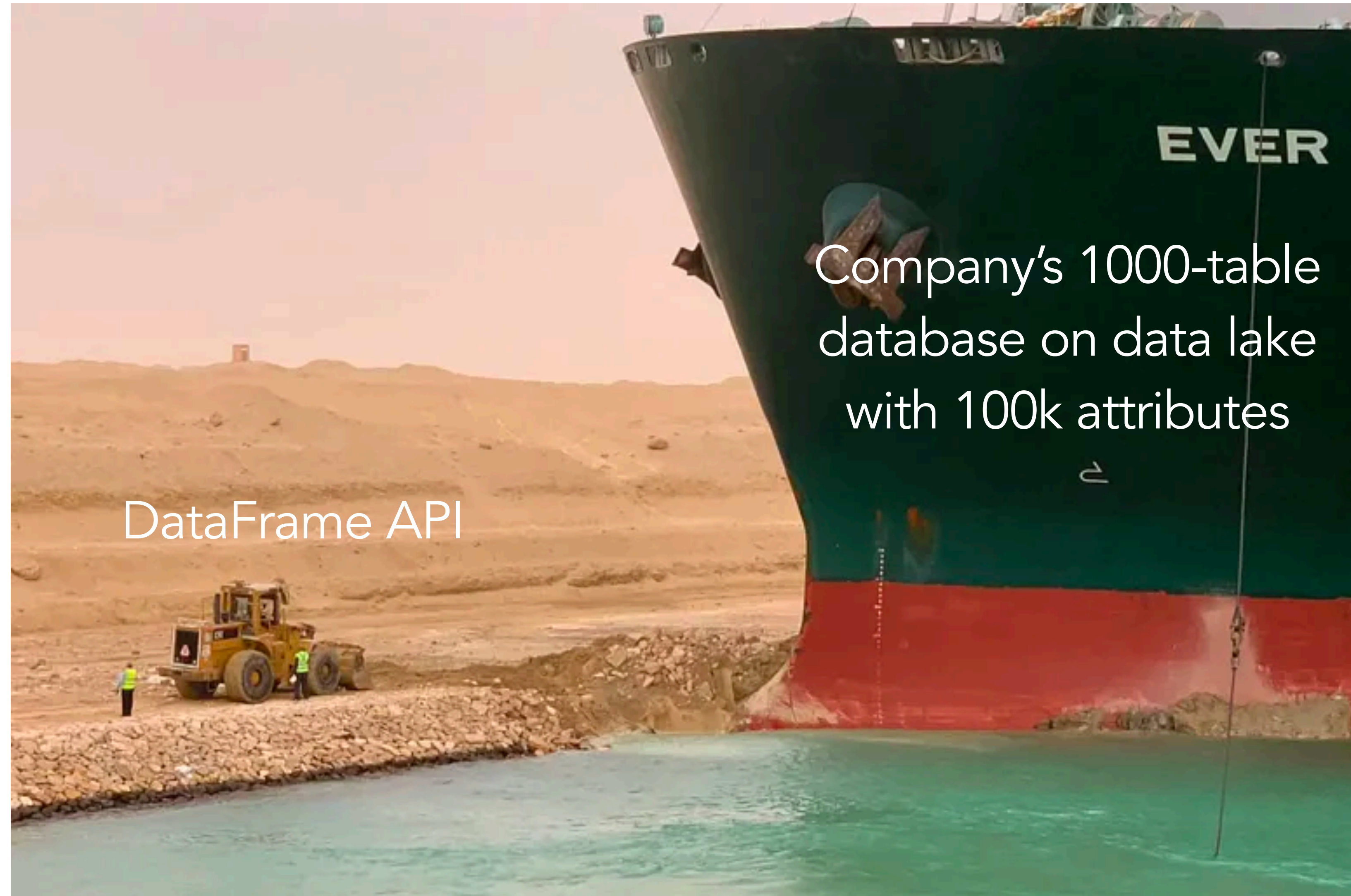


# DSC 204a Scalable Data Systems

- Haojian Jin



Company's 1000-table  
database on data lake  
with 100k attributes

DataFrame API

# Where are we in the class?

## Foundations of Data Systems (2 weeks)

- Digital representation of Data → Computer Organization → Memory hierarchy → Process → Storage

## Scaling Distributed Systems (3 weeks)

- Cloud → Network → **Distributed storage** → Partition and replication (HDFS) → Distributed computation

## Data Processing and Programming model (5 weeks)

- Data Models evolution → Data encoding evolution → → IO & Unix Pipes → Batch processing (MapReduce) → Stream processing (Spark)

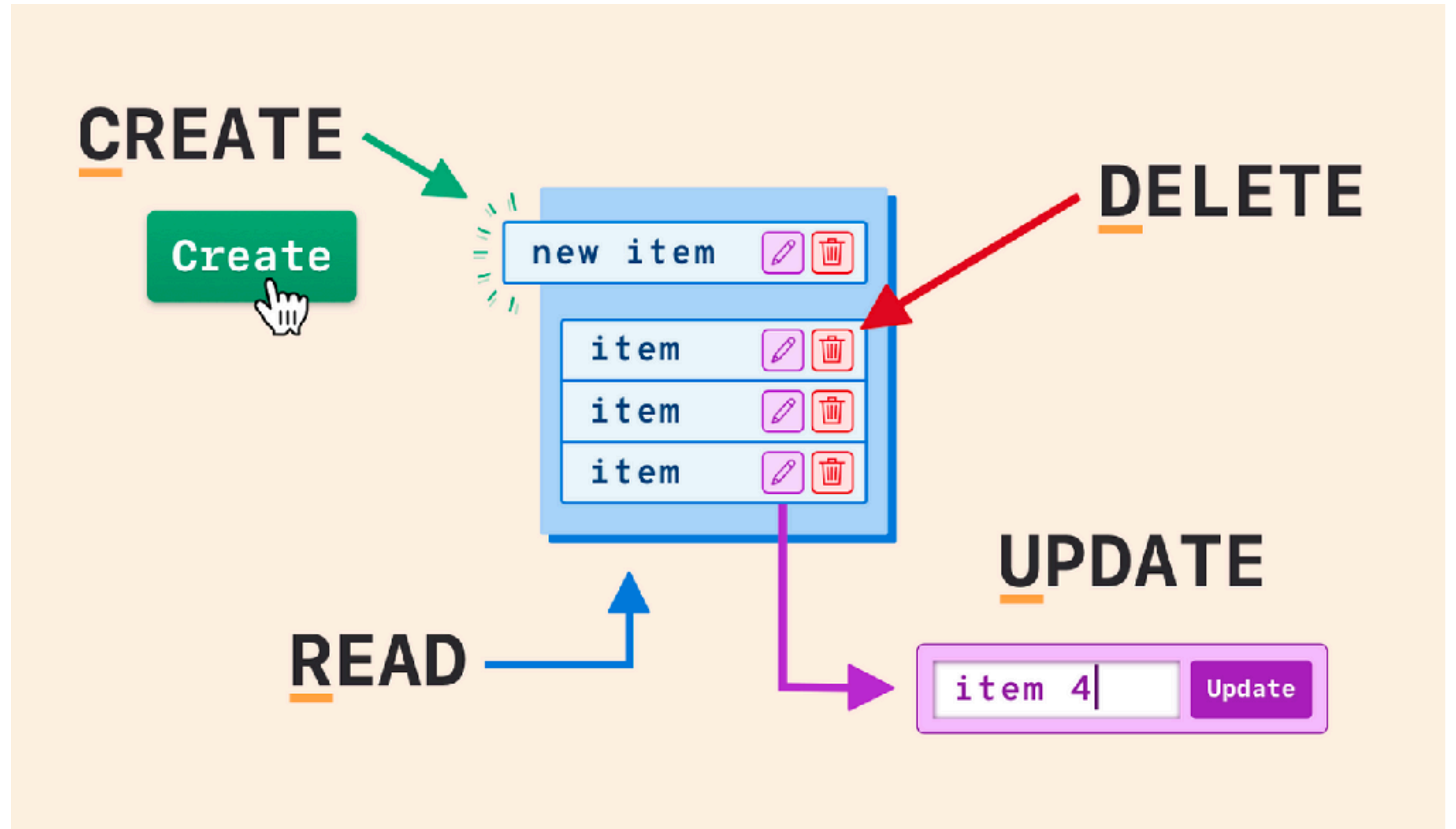
# Today's topic: Column-oriented storage

- OLTP v.s. OLAP
- Data warehousing
- Schemas for Analytics
- Column-oriented storage
- Data cubes and materialized views

# CRUD

**I'm a Database Developer,  
all what I do is**

	<b><i>C</i>reate</b>
	<b><i>R</i>ead</b>
	<b><i>U</i>pdate</b>
	<b><i>D</i>elete</b>





# Database transactions

- Make sale
- Place an order
- Pay an employee's salary
- Comment a blog post
- Act in games
- Add/remove contact to an address book

Online transaction processing (OLTP)

# Walmart Beer and Diaper (1988)



Forbes 1988

- Unexpected correlation:
  - Sales of diapers and beer

# Data analytics

- What was the total revenue of each of our stores in Jan?
- How many more bananas than usual did we sell during our latest data?
- Which brand of baby food is most often purchased together with brand X diapers?

Online analytic processing (OLAP)

# OLTP v.s. OLAP

Property	Transaction processing systems (OLTP)	Analytic systems (OLAP)
Main read pattern	Small number of records per query, fetched by key	Aggregate over large number of records



# OLTP v.s. OLAP

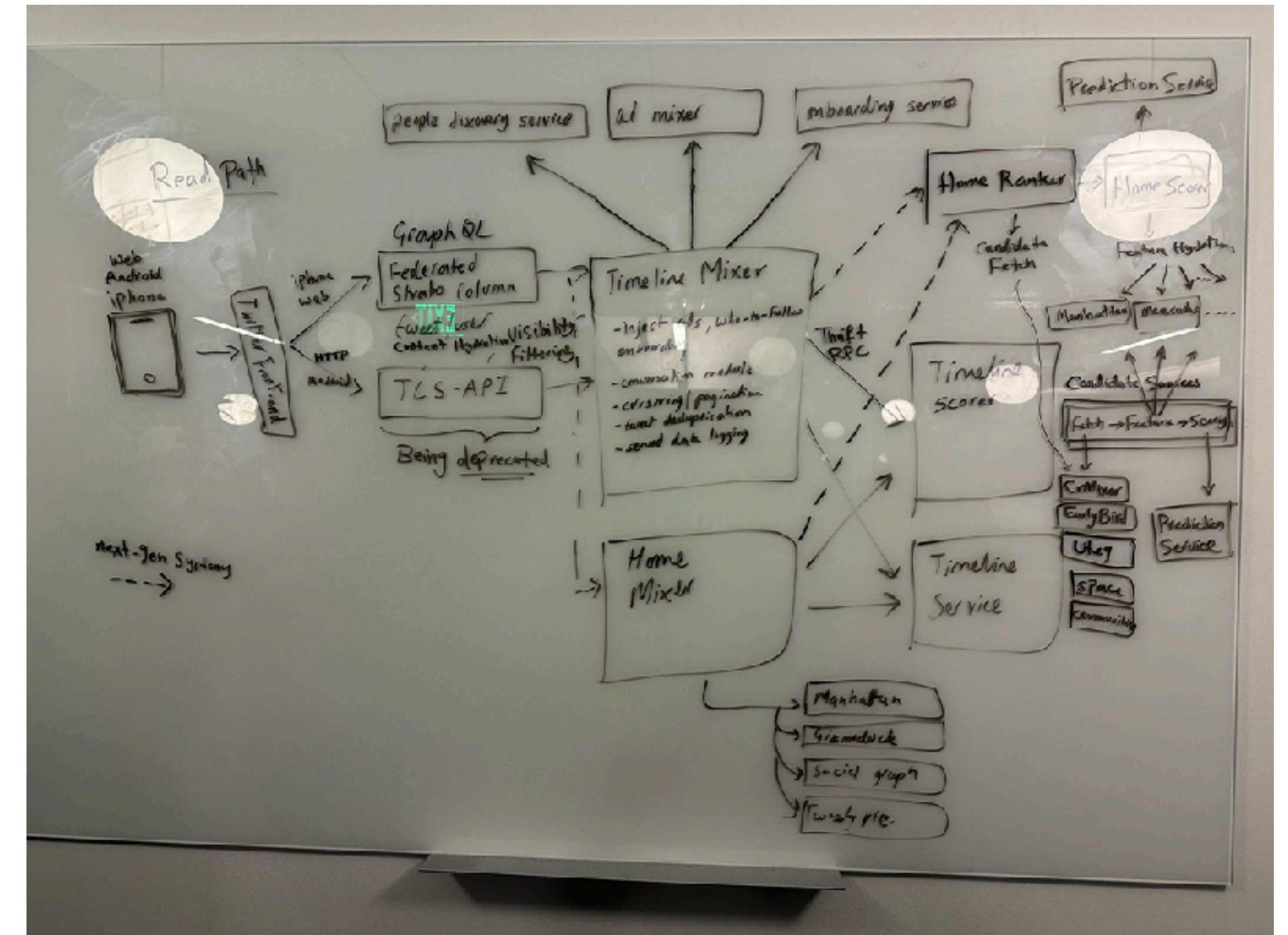
Property	Transaction processing systems (OLTP)	Analytic systems (OLAP)
Main read pattern	Small number of records per query, fetched by key	Aggregate over large number of records
Main write pattern	Random-access, low-latency writes from user input	Bulk import (ETL) or event stream
Primarily used by	End user/customer, via web application	Internal analyst, for decision support
What data represents	Latest state of data (current point in time)	History of events that happened over time
Dataset size	Gigabytes to terabytes	Terabytes to petabytes

# Today's topic: Column-oriented storage

- OLTP v.s. OLAP
- **Data warehousing**
- Schemas for Analytics
- Column-oriented storage
- Data cubes and materialized views



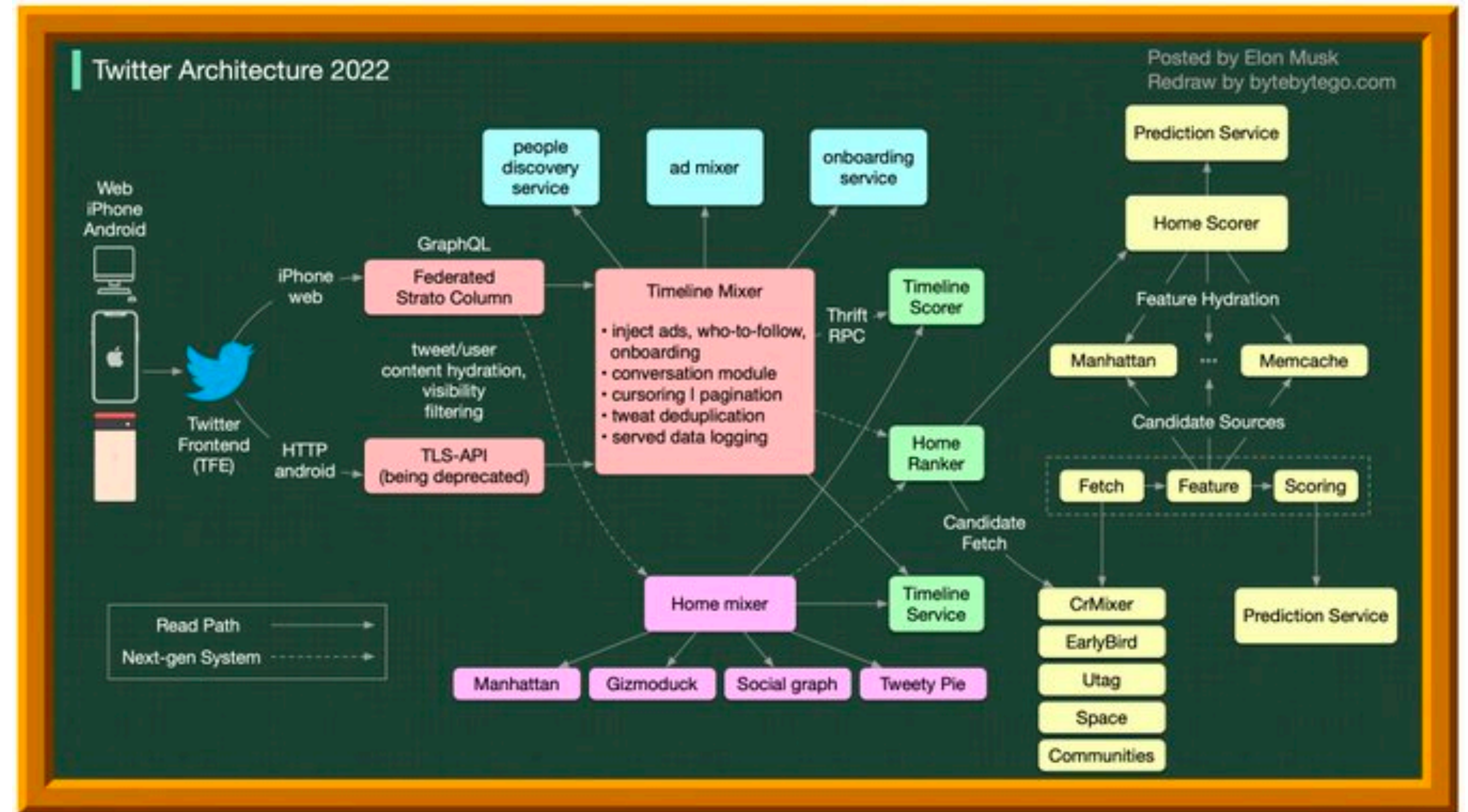
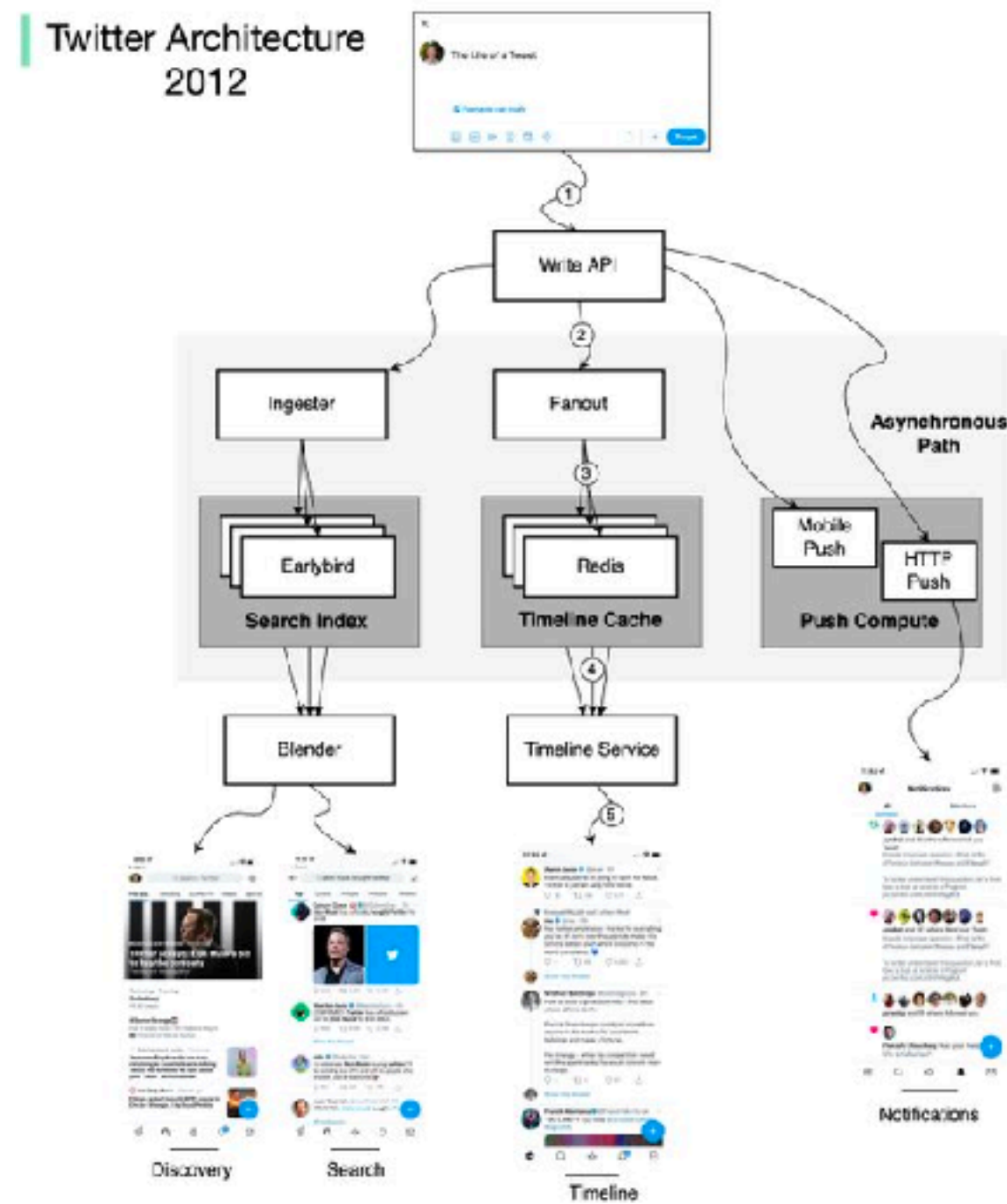
# Transaction systems are complex.



Elon Musk's Twitter System Design Diagram Explained  
[https://www.youtube.com/watch?v=\\_Y5aGCOkymQ](https://www.youtube.com/watch?v=_Y5aGCOkymQ)



# Transaction systems need to be highly available.



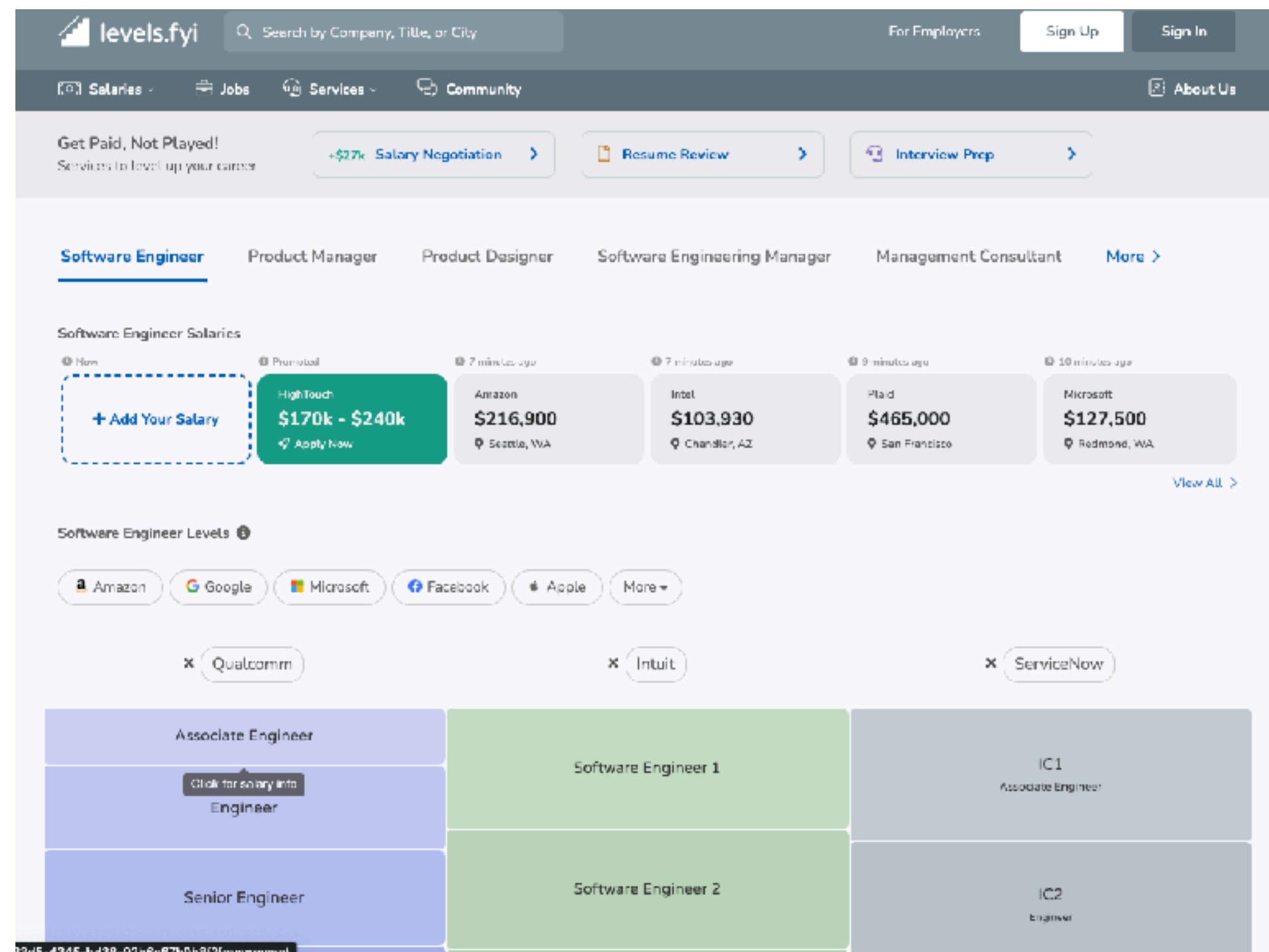
- Low latency.
- Highly available.
- **Ad hoc analytic queries are expensive.** <https://twitter.com/alexxubyte/status/1594008281340530688>

# Data warehouse

- A separate database that analysts can query to their hearts' content, without affecting OLTP operations.
- Maintain a read-only copy for analytic purposes.
- Only exist in almost all large enterprises.



# Small companies?

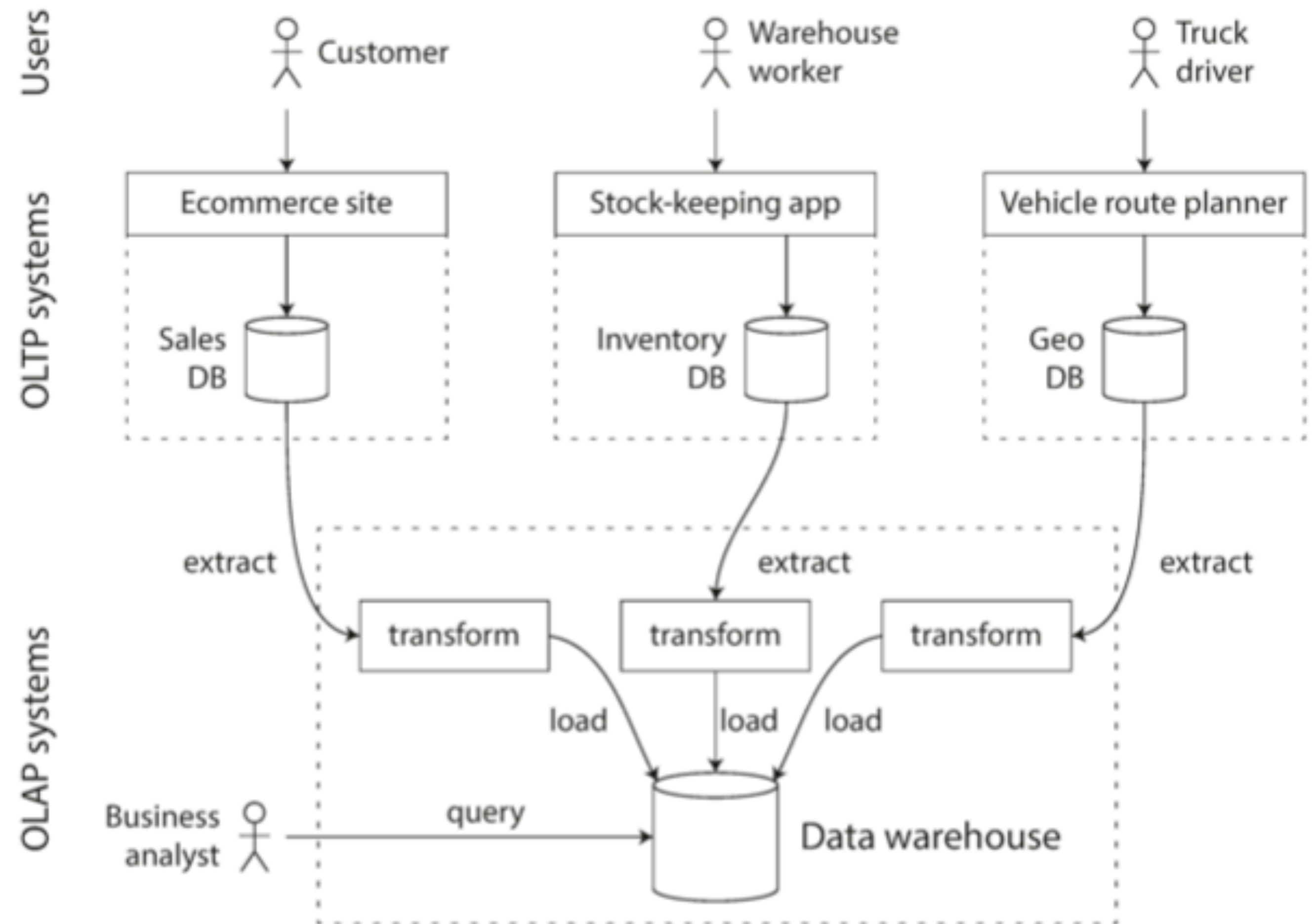


**How Levels.fyi scaled to millions of users with Google Sheets as a backend**

Our philosophy to scaling is simple, avoid premature optimization

# Extract-Transform-Load (ETL)

- Extract
  - Periodica data dump
  - Continuous streaming
- Transform
  - Analysis-friendly schema
  - Data cleaning
- Load into a data warehouse



# Why data warehouse?

- Separation of concerns
  - Performance (reliability, latency)
  - Expertise requirement, management
- The indexes in last lecture (e.g., SSTable, B-tree) are not good for reading and writing a single record.
  - But **are not good at answering analytic queries.**

# How do you interact with OLAP & OLTP

- SQL query interface
  - *Select \* from*
  - “A database system can be considered mature when it has an SQL query interface”.
  - Both OLAP and OLTP
- OLAP:
  - More and more codeless user interfaces.

# Data wrangler

**Data Wrangler**

Transform Script Import Export

- Split data repeatedly on newline into rows
- Split split repeatedly on ','
- Promote row 0 to header
- Delete empty rows

Text Columns Rows Table Clear

Extract from Year after 'in '

Extract from Year after ' in '

Cut from Year after 'in '

Cut from Year after ' in '

Split Year after 'in '

Split Year after ' in '

	Year	extract	#	Property_crime_rate
0	Reported crime in Alabama	Alabama		
1	2004			4029.3
2	2005			3900
3	2006			3937
4	2007			3974.9
5	2008			4081.9
6	Reported crime in Alaska	Alaska		
7	2004			3370.9
8	2005			3615
9	2006			3582
10	2007			3373.9
11	2008			2928.3
12	Reported crime in Arizona	Arizona		
13	2004			5073.3
14	2005			4827
15	2006			4741.6
16	2007			4502.6
17	2008			4087.3
18	Reported crime in Arkansas	Arkansas		
19	2004			4033.1
20	2005			4068
21	2006			4021.6
22	2007			3945.5
23	2008			3843.7
24	Reported crime in California	California		
25	2004			3423.9
26	2005			3321
27	2006			3175.2
28	2007			3032.6
29	2008			2940.3
30	Reported crime in Colorado	Colorado		

SIGCHI 2011

<https://www.trifacta.com/>  
\$400 million, Feb 2022



# Today's topic: Column-oriented storage

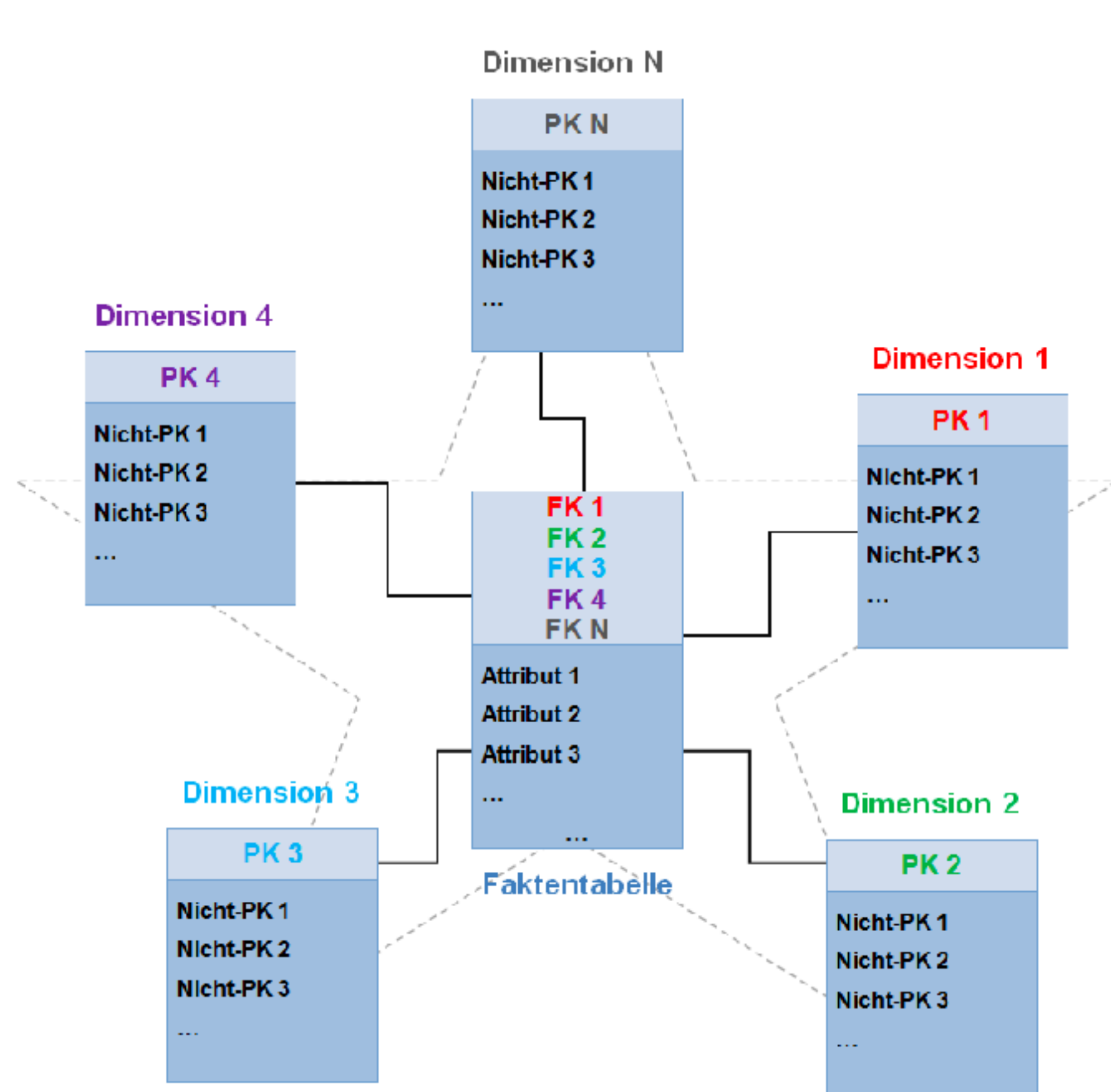
- OLTP v.s. OLAP
- Data warehousing
- **Schemas for Analytics**
- Column-oriented storage
- Data cubes and materialized views

# Data analytic queries

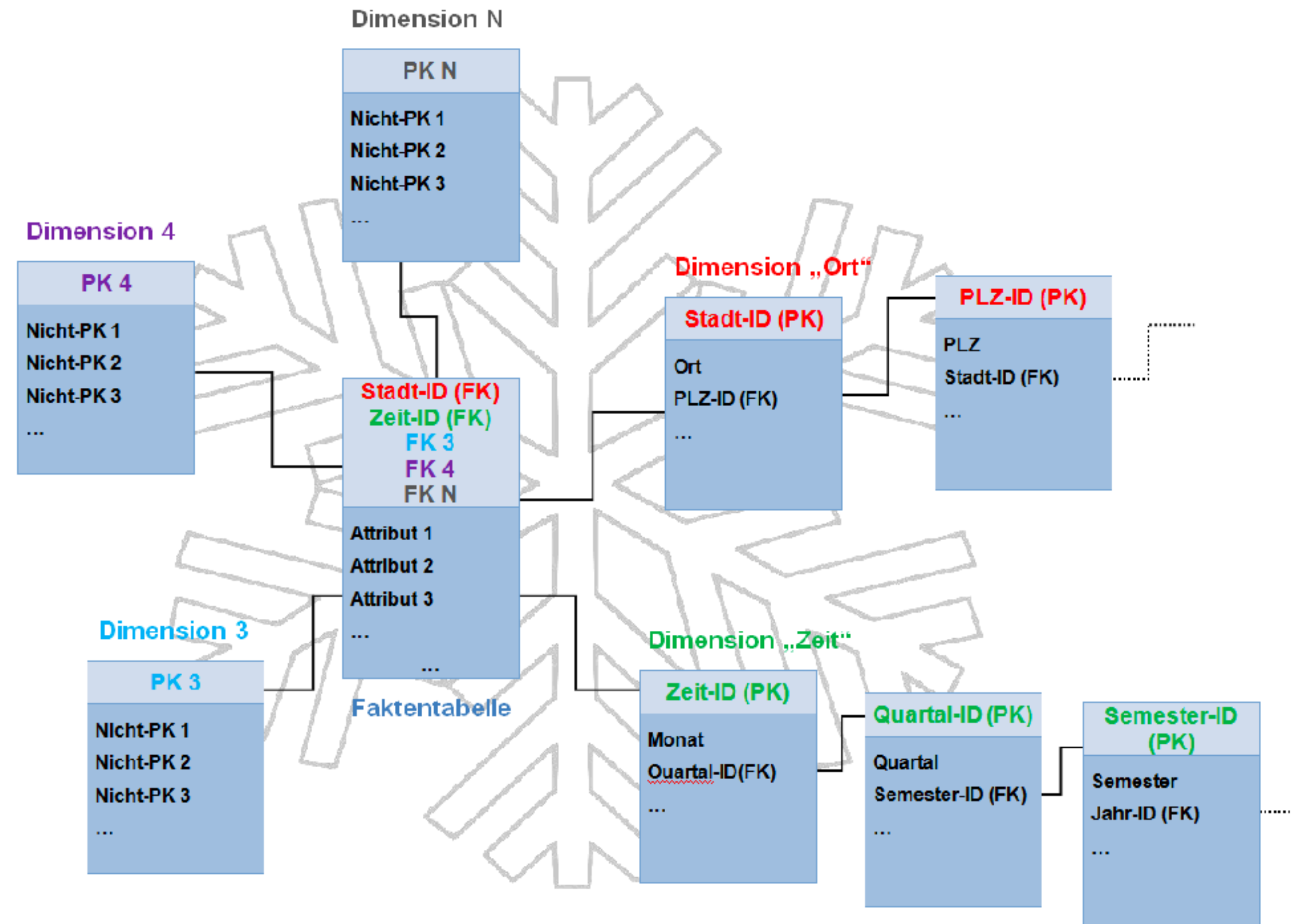
- What was the total revenue of each of our stores in Jan?
- How many more bananas than usual did we sell during our latest data?
- Which brand of baby food is most often purchased together with brand X diapers?

# Data model

- Relation (SQL)
- Document (NoSQL)
- Graph (GraphQL)
- Network
- Hierarchy
- Stars
- Snowflake



Star



Snowflake

# Star schema

- Fact table in the middle
  - A collection of events
  - e.g., click events, page views, retail sales
- Two types of columns
  - Attributes
  - References to dimension tables.
- Fact table: event meta data
- Dimensions: who, what, where, when, how, and why of the event.

dim\_product table

product_sk	sku	description	brand	category
30	OK4012	Bananas	Freshmax	Fresh fruit
31	KA9511	Fish food	Aquatech	Pet supplies
32	AB1234	Croissant	Dealicious	Bakery

dim\_store table

store_sk	state	city
1	WA	Seattle
2	CA	San Francisco
3	CA	Palo Alto

fact\_sales table

date_key	product_sk	store_sk	promotion_sk	customer_sk	quantity	net_price	discount_price
140102	31	3	NULL	NULL	1	2.49	2.49
140102	69	5	19	NULL	3	14.99	9.99
140102	74	3	23	191	1	4.49	3.89
140102	33	8	NULL	235	4	0.99	0.99

dim\_date table

date_key	year	month	day	weekday	is_holiday
140101	2014	jan	1	wed	yes
140102	2014	jan	2	thu	no
140103	2014	jan	3	fri	no

dim\_customer table

customer_sk	name	date_of_birth
190	Alice	1979-03-29
191	Bob	1961-09-02
192	Cecil	1991-12-13

dim\_promotion table

promotion_sk	name	ad_type	coupon_type
18	New Year sale	Poster	NULL
19	Aquarium deal	Direct mail	Leaflet
20	Coffee & cake bundle	In-store sign	NULL



# Example: dim\_date table

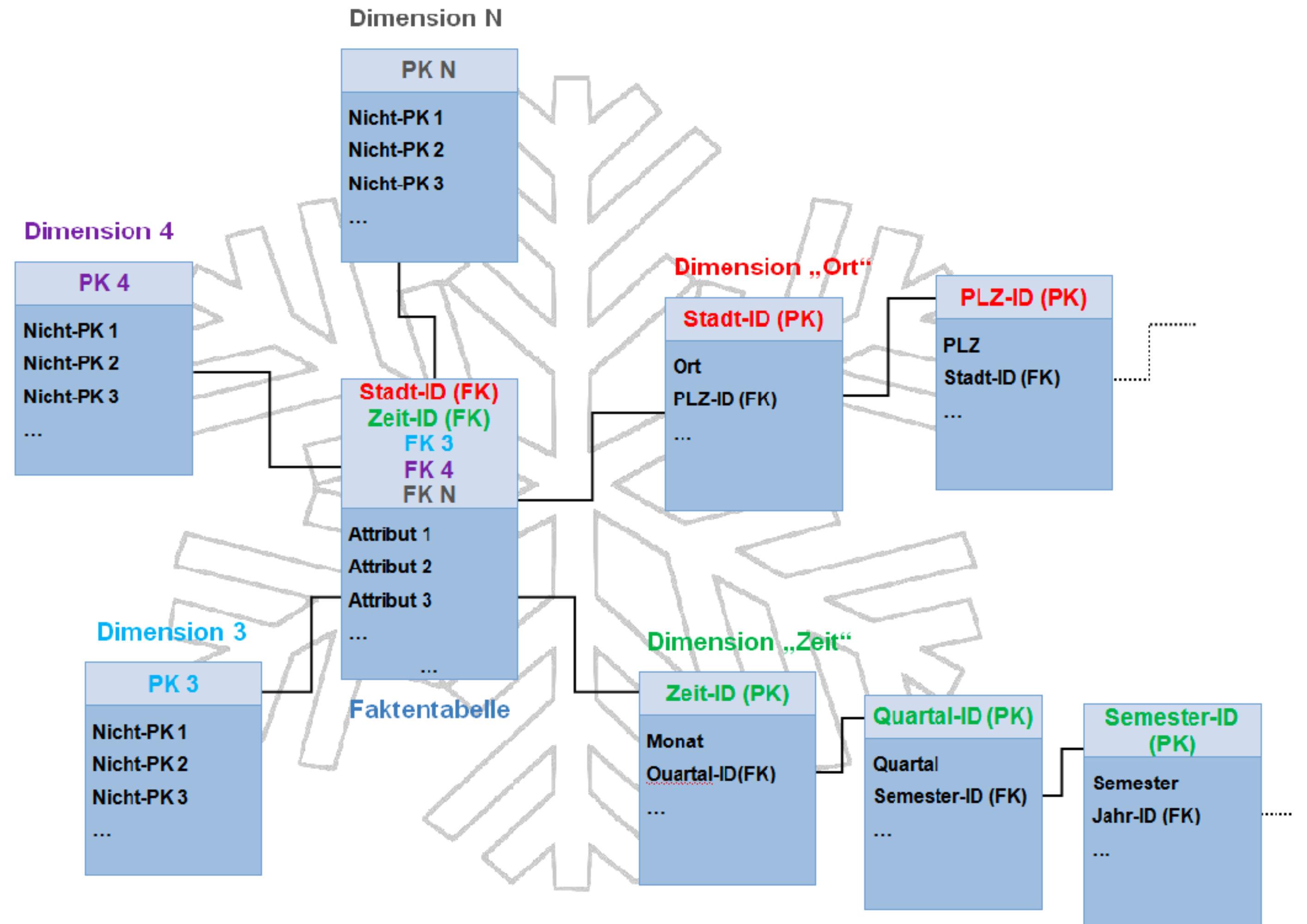
- Speed up the analysis.
- Easier development.

dim\_date table

date_key	year	month	day	weekday	is_holiday
140101	2014	jan	1	wed	yes
140102	2014	jan	2	thu	no
140103	2014	jan	3	fri	no

# Snowflake schema

- More normalized data model



# Database normalization

- Storage redundancy
- Easy interactions.

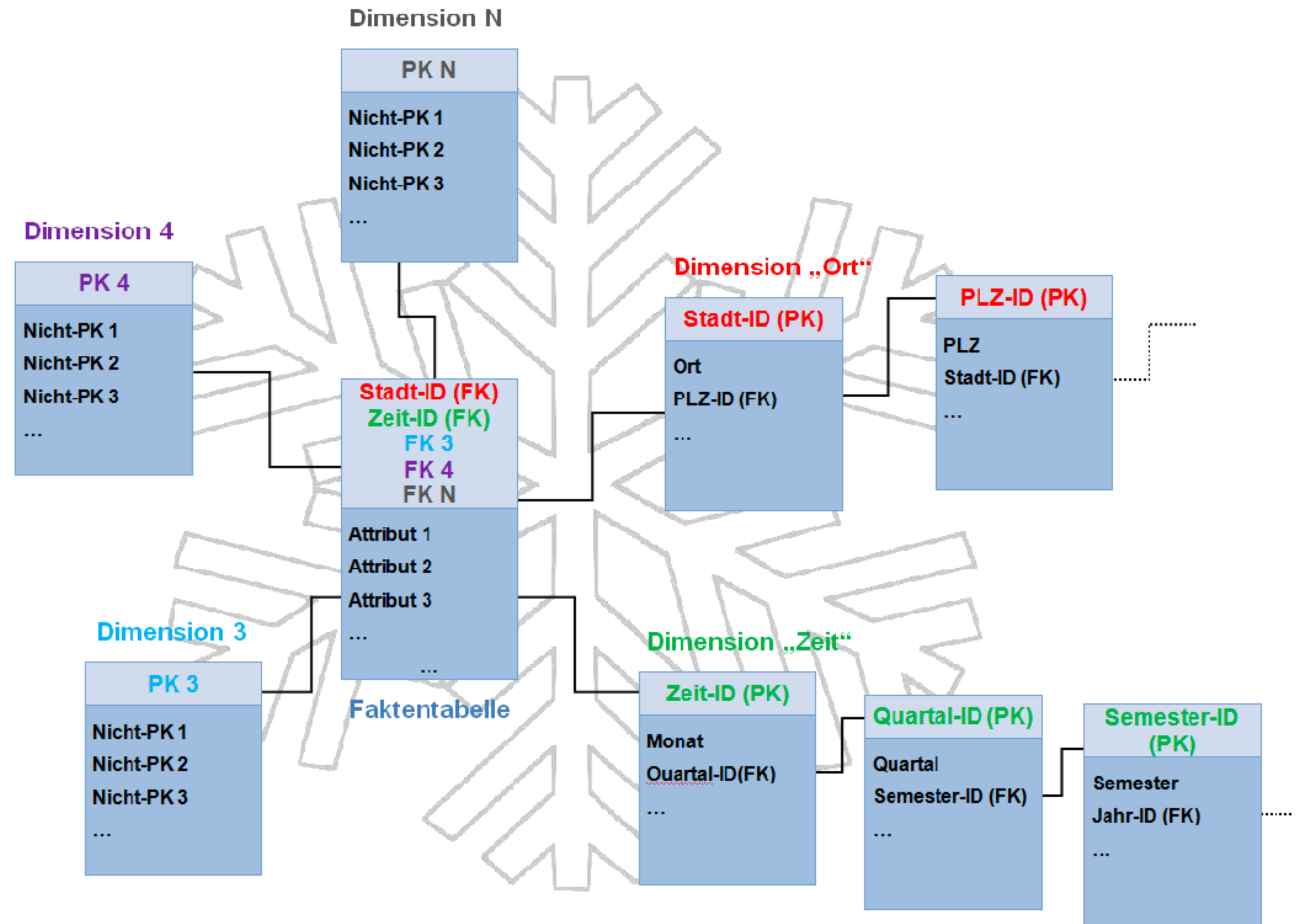
EmpID	Employee	Age	Dept
1001	ABC	30	Sales,Finance
1002	CDE	30	Sales,Finance,DevOps

DeptID	DeptName
1	Sales
2	Finance
3	DevOps

EmpID	Employee	Age	DeptID
1001	ABC	30	1
1001	ABC	30	2
1002	CDE	40	1
1002	CDE	40	2
1002	CDE	40	3

# Snowflake schema tradeoffs

- More storage efficiency
- Enforces data quality
- Slower
- Lots of overhead upon initial setup
- High maintenance costs



# Today's topic: Column-oriented storage

- OLTP v.s. OLAP
- Data warehousing
- Schemas for Analytics
- **Column-oriented storage**
- Data cubes and materialized views



# Data scale

- Fact tables
  - Hundreds of columns
  - Trillions of rows
  - Petabytes of data
- Dimension tables
  - Million of rows.
  - Can be wide. But less common.

# How many columns do we need?

- What was the total revenue of each of our stores in Jan?
- How many more bananas than usual did we sell during our latest data?
- Which brand of baby food is most often purchased together with brand X diapers?

dim\_product table

product_sk	sku	description	brand	category
30	OK4012	Bananas	Freshmax	Fresh fruit
31	KA9511	Fish food	Aquatech	Pet supplies
32	AB1234	Croissant	Dealicious	Bakery

dim\_store table

store_sk	state	city
1	WA	Seattle
2	CA	San Francisco
3	CA	Palo Alto

fact\_sales table

date_key	product_sk	store_sk	promotion_sk	customer_sk	quantity	net_price	discount_price
140102	31	3	NULL	NULL	1	2.49	2.49
140102	69	5	19	NULL	3	14.99	9.99
140102	74	3	23	191	1	4.49	3.89
140102	33	8	NULL	235	4	0.99	0.99

dim\_date table

date_key	year	month	day	weekday	is_holiday
140101	2014	jan	1	wed	yes
140102	2014	jan	2	thu	no
140103	2014	jan	3	fri	no

dim\_customer table

customer_sk	name	date_of_birth
190	Alice	1979-03-29
191	Bob	1961-09-02
192	Cecil	1991-12-13

dim\_promotion table

promotion_sk	name	ad_type	coupon_type
18	New Year sale	Poster	NULL
19	Aquarium deal	Direct mail	Leaflet
20	Coffee & cake bundle	In-store sign	NULL



