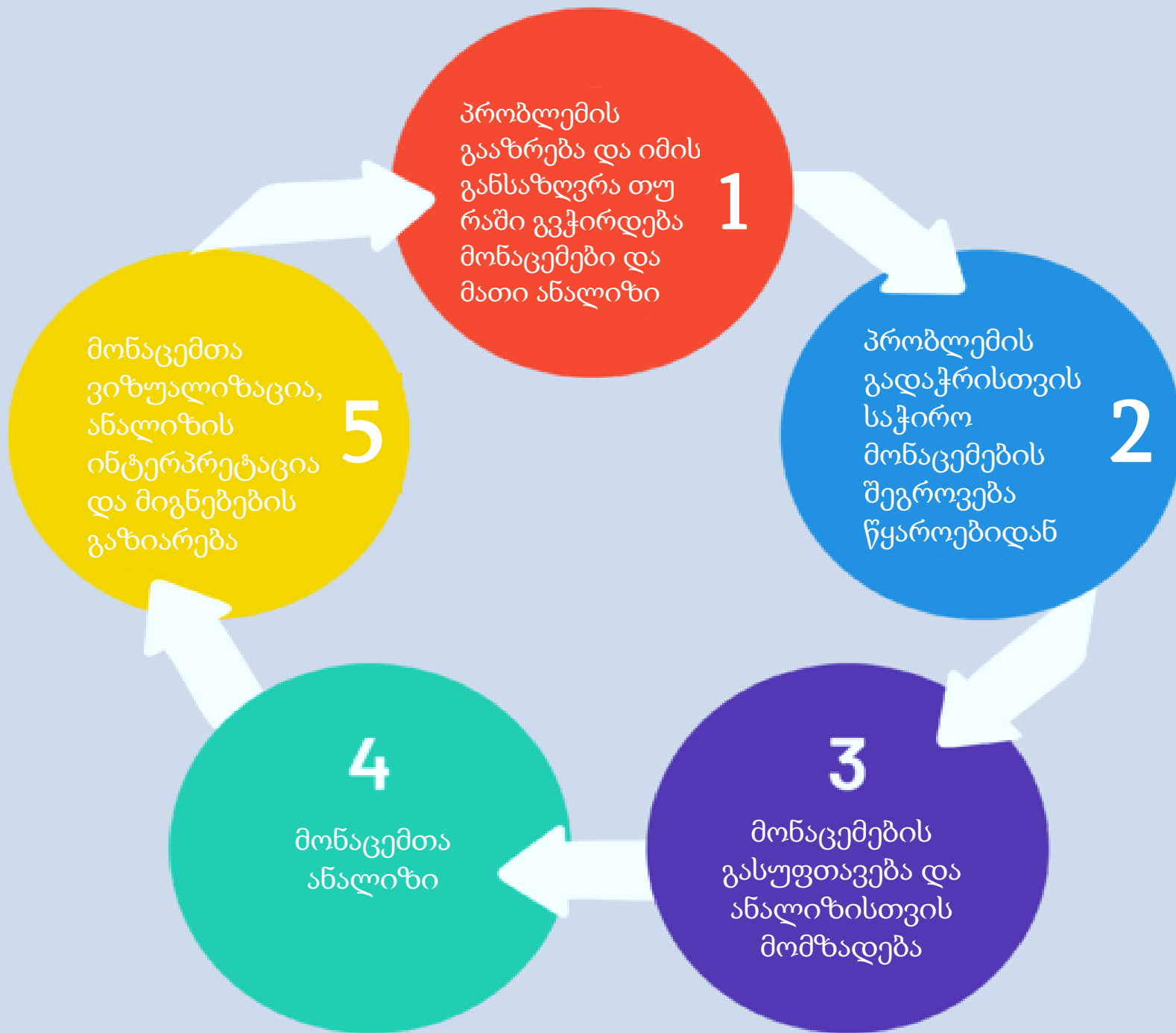




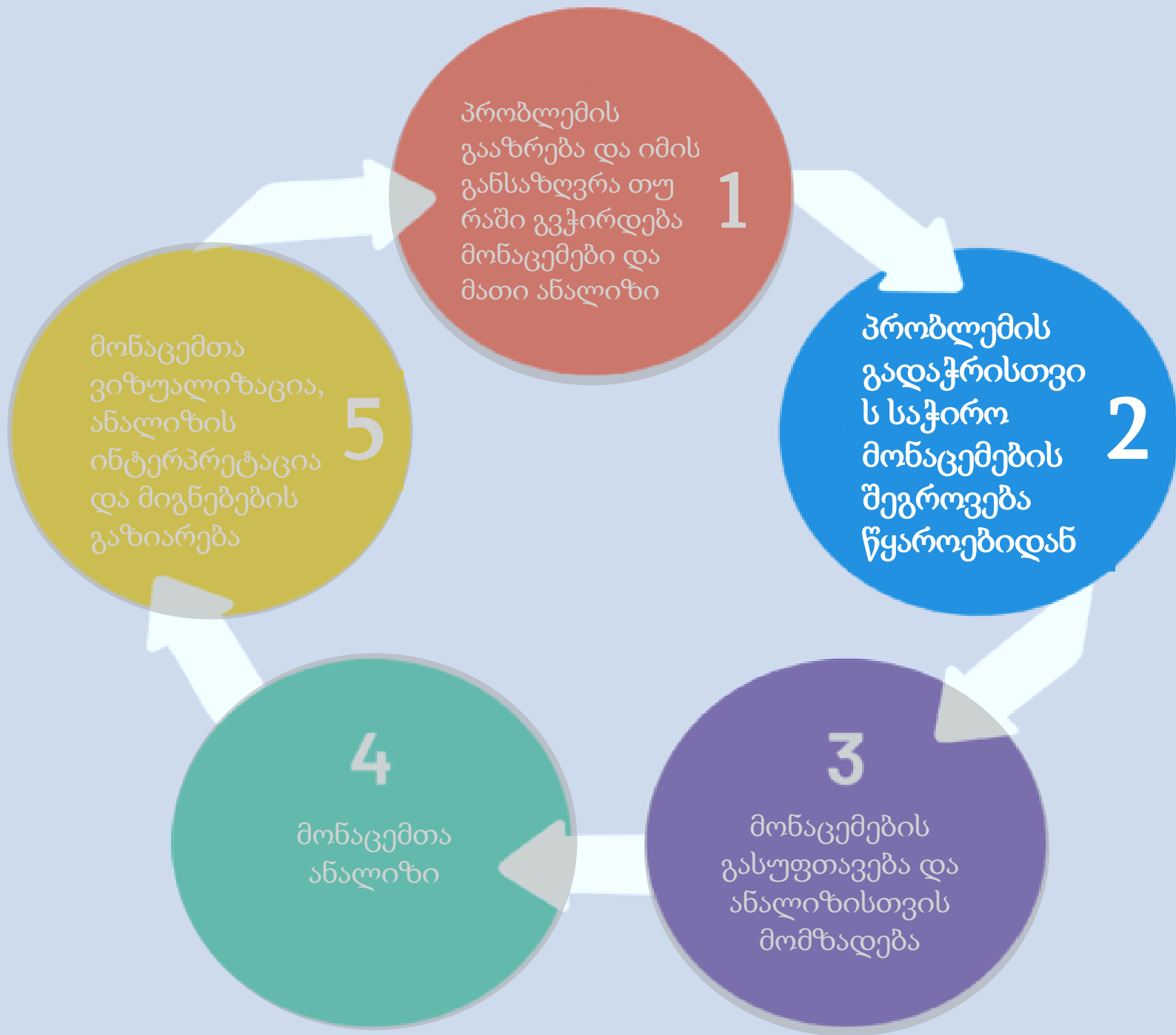
მონაცემთა ანალიტიკა

ნუცა აბაზაძე
2022

მონაცემთა ანალიზის ციკლი



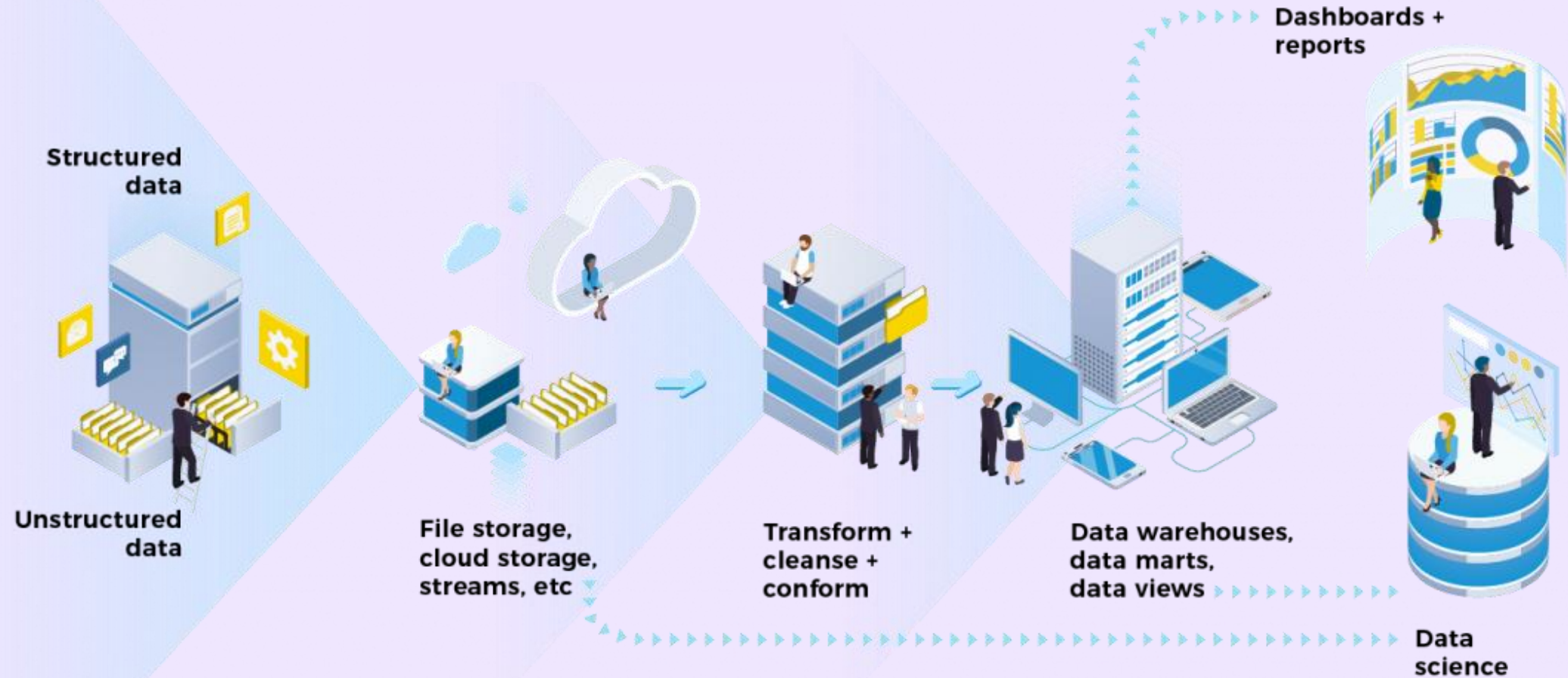
მონაცემთა ანალიზის ციკლი



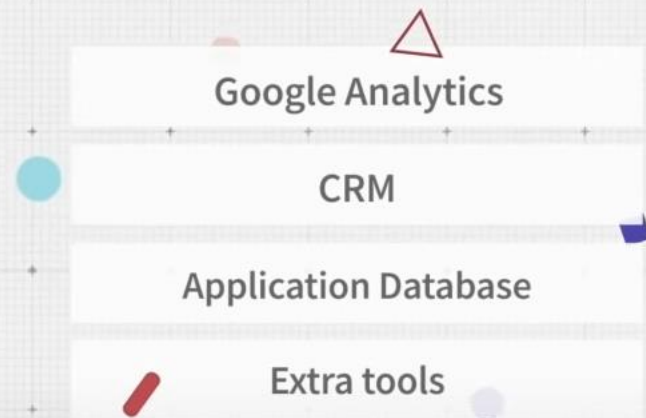
2

მონაცემთა
ეკონომიკა,
ფორმატები და
ტიპები

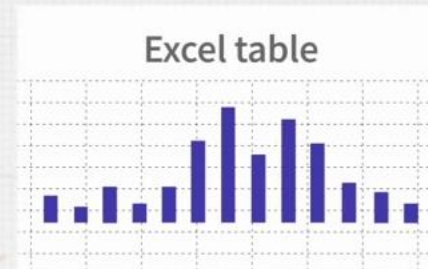
DATA PIPELINE/INFRASTRUCTURE/ECOSYSTEM



DATA PIPELINE V1 😊



**Spreadsheet
beast**



**Analytic
user**

DATA PIPELINE V2 😊



Data Engineering

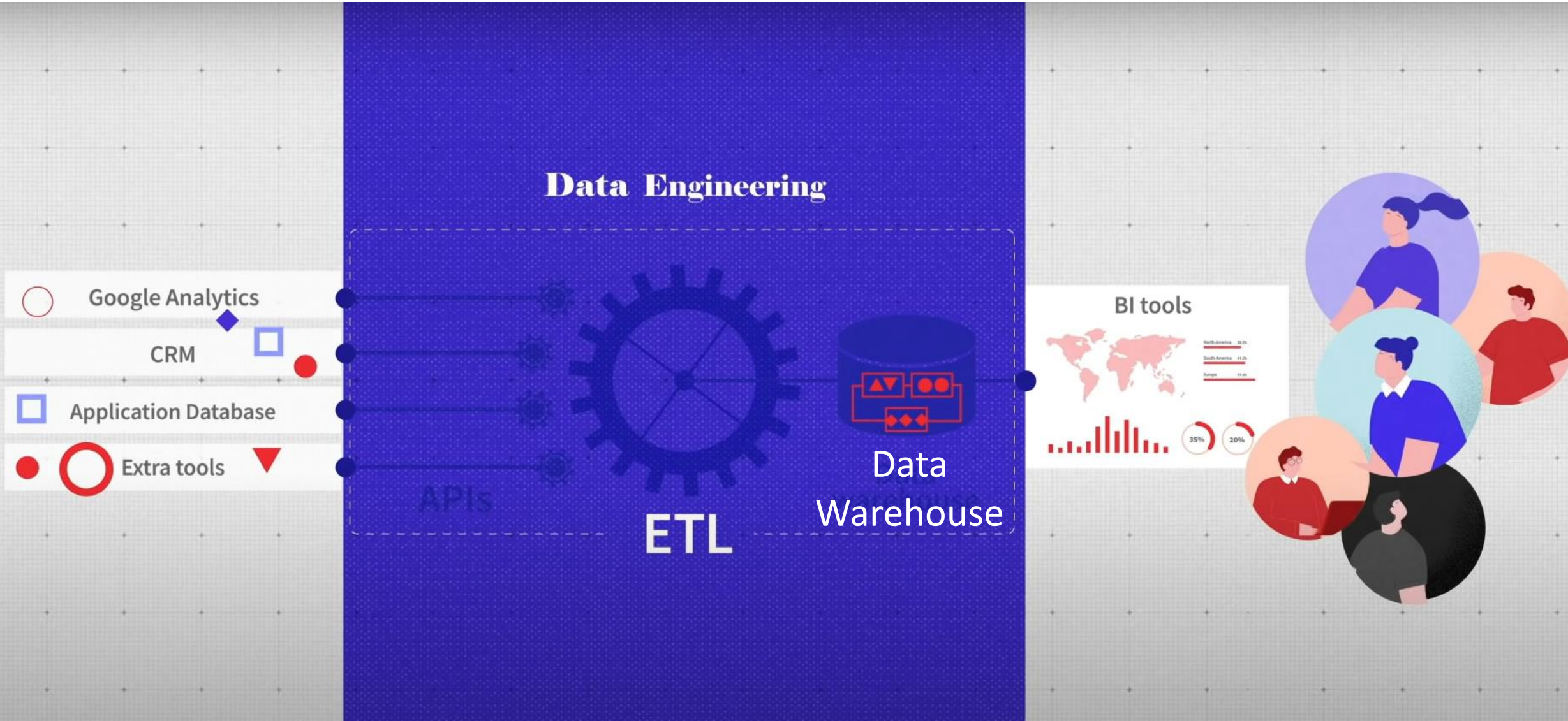
- Google Analytics
- CRM
- Application Database
- Extra tools



BI tools

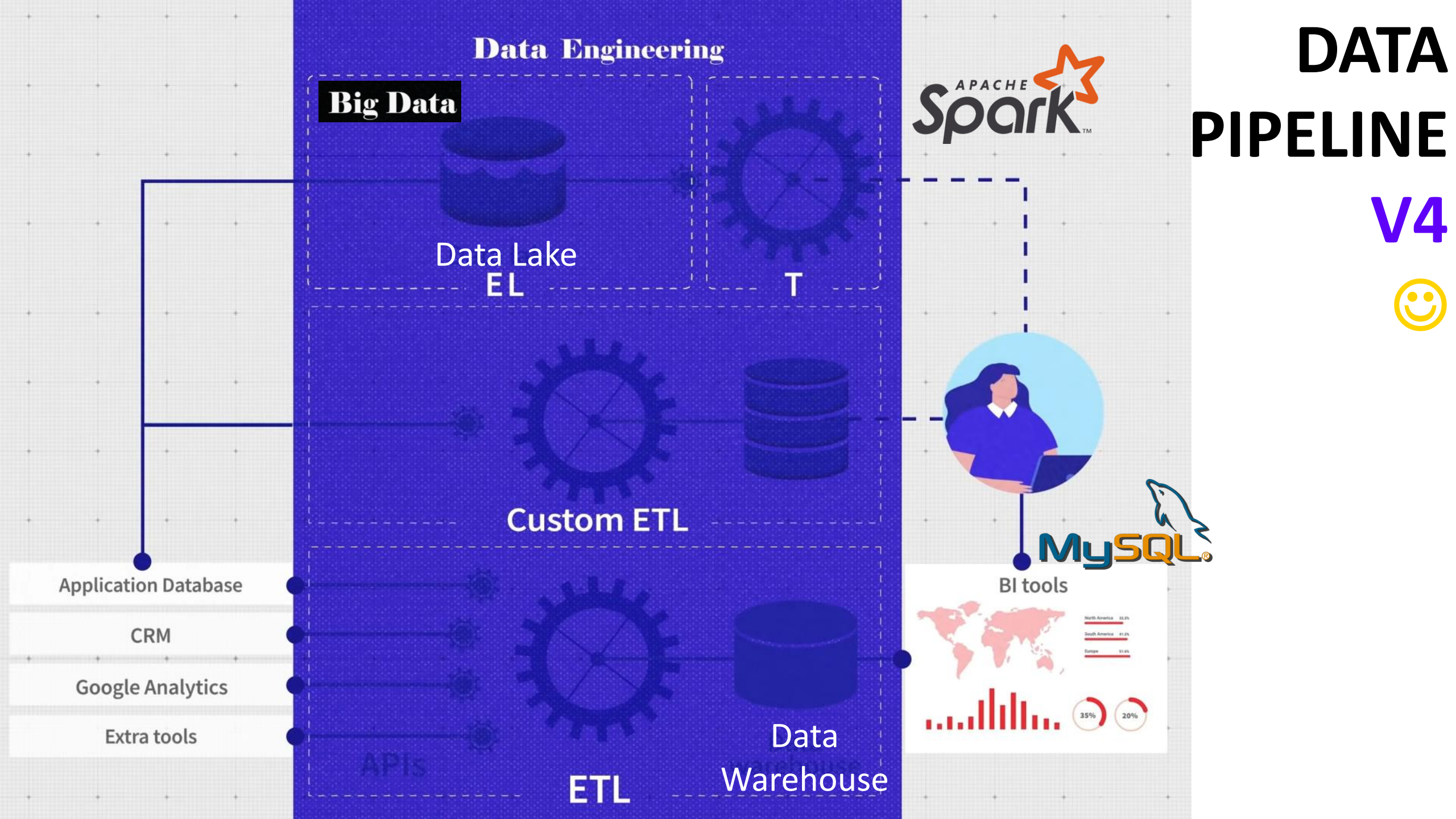


DATA PIPELINE V3 😊



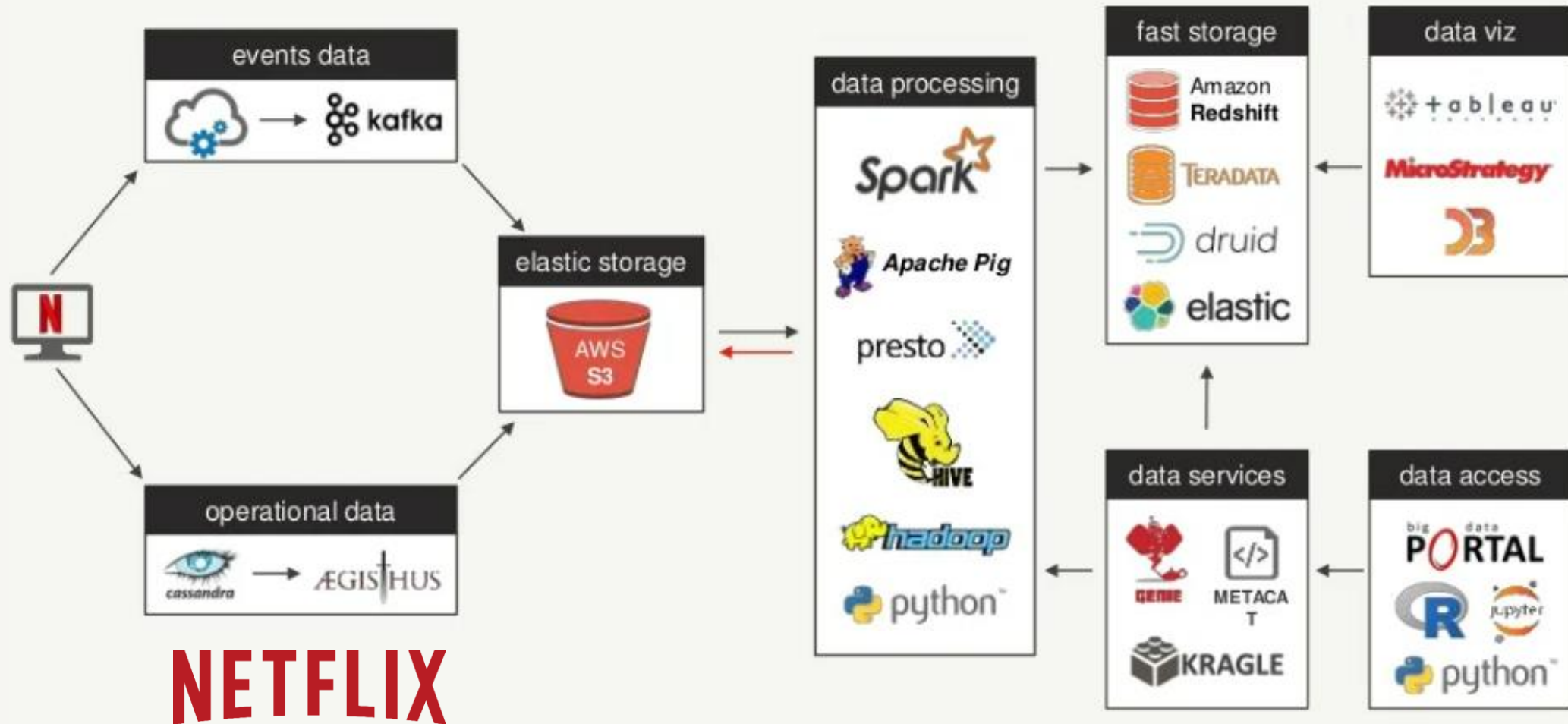
DATA PIPELINE

V4



DATA PIPELINE V5 😊

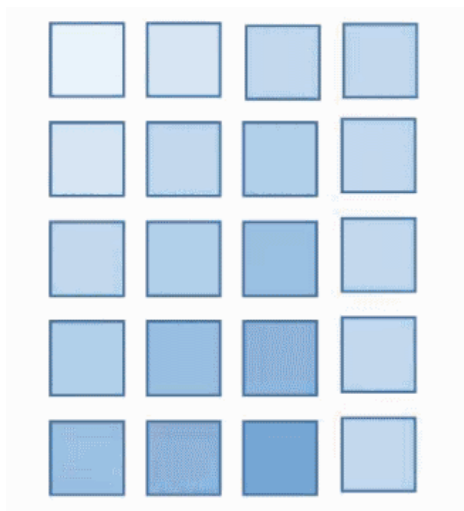
CLOUD + BIG DATA





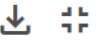
სტრუქტურირებული VS არასტრუქტურირებული

სტრუქტურირებული მონაცემები



ცხრილის ფორმატში
შენახული მონაცემები

hotel_bookings.csv (16.86 MiB)

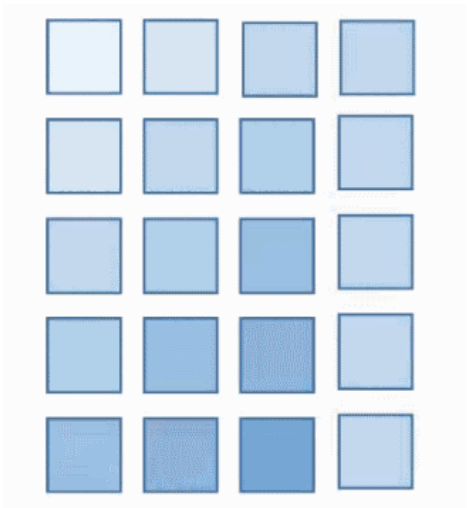


Detail Compact Column

6 of 32 columns

▲ hotel	# is_canceled	▲ arrival_dat...	# adults	▲ country	# is_repeate...
Resort Hotel	0	July	2	PRT	0
Resort Hotel	0	July	2	PRT	0
Resort Hotel	0	July	1	GBR	0
Resort Hotel	0	July	1	GBR	0
Resort Hotel	0	July	2	GBR	0
Resort Hotel	0	July	2	GBR	0
Resort Hotel	0	July	2	PRT	0
Resort Hotel	0	July	2	PRT	0
Resort Hotel	1	July	2	PRT	0
Resort Hotel	1	July	2	PRT	0
Resort Hotel	1	July	2	PRT	0
Resort Hotel	0	July	2	PRT	0

სტრუქტურირებული
მონაცემები

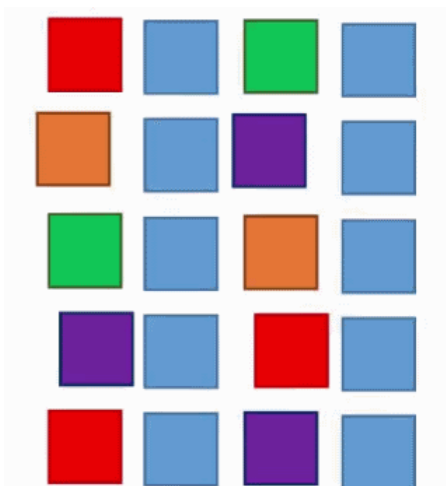


ცხრილის ფორმატში
შენახული მონაცემები

sales data-set.csv (13.26 MiB)

Detail	Compact	Column		
# Store	# Dept	A Date	# Weekly_Sa...	✓ IsHoliday
1	1	05/02/2010	24924.5	FALSE
1	1	12/02/2010	46039.49	TRUE
1	1	19/02/2010	41595.55	FALSE
1	1	26/02/2010	19403.54	FALSE
1	1	05/03/2010	21827.9	FALSE
1	1	12/03/2010	21043.39	FALSE
1	1	19/03/2010	22136.64	FALSE
1	1	26/03/2010	26229.21	FALSE
1	1	02/04/2010	57258.43	FALSE
1	1	09/04/2010	42960.91	FALSE
1	1	16/04/2010	17596.96	FALSE

ნახევრად სტრუქტურირებული მონაცემები



[Chicken Tikka Masala Recipe - Allrecipes.com](#)



[allrecipes.com/recipe/chicken-tikka-masala/](#) ▼

★★★★★ Rating: 4.4 - 1,705 reviews - 2 hrs 20 mins - 404 cal

This is an easy recipe for **Chicken Tikka Masala** - Chicken marinated in yogurt and spices and then served in a tomato cream sauce.

Curry Stand Chicken Tikka ... - 106 Photos - Authentic Chicken Tikka Masala

[Chicken Tikka Masala - Bon Appétit](#)

[www.bonappetit.com/recipe/chicken-tikka-masala](#) ▼

★★★★★ Rating: 5 - 3 votes - 5 mins - Calories (kcal) 610 cal

INGREDIENTS: 6 garlic cloves, finely grated; 4 teaspoons finely grated peeled ginger; 4 teaspoons ground turmeric; 2 teaspoons garam **masala**; 2 teaspoons ...

[Chicken tikka masala | BBC Good Food](#)

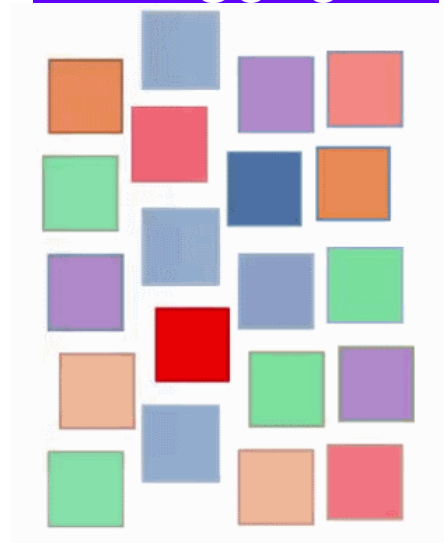


[www.bbcgoodfood.com/recipes/12798/](#) ▼

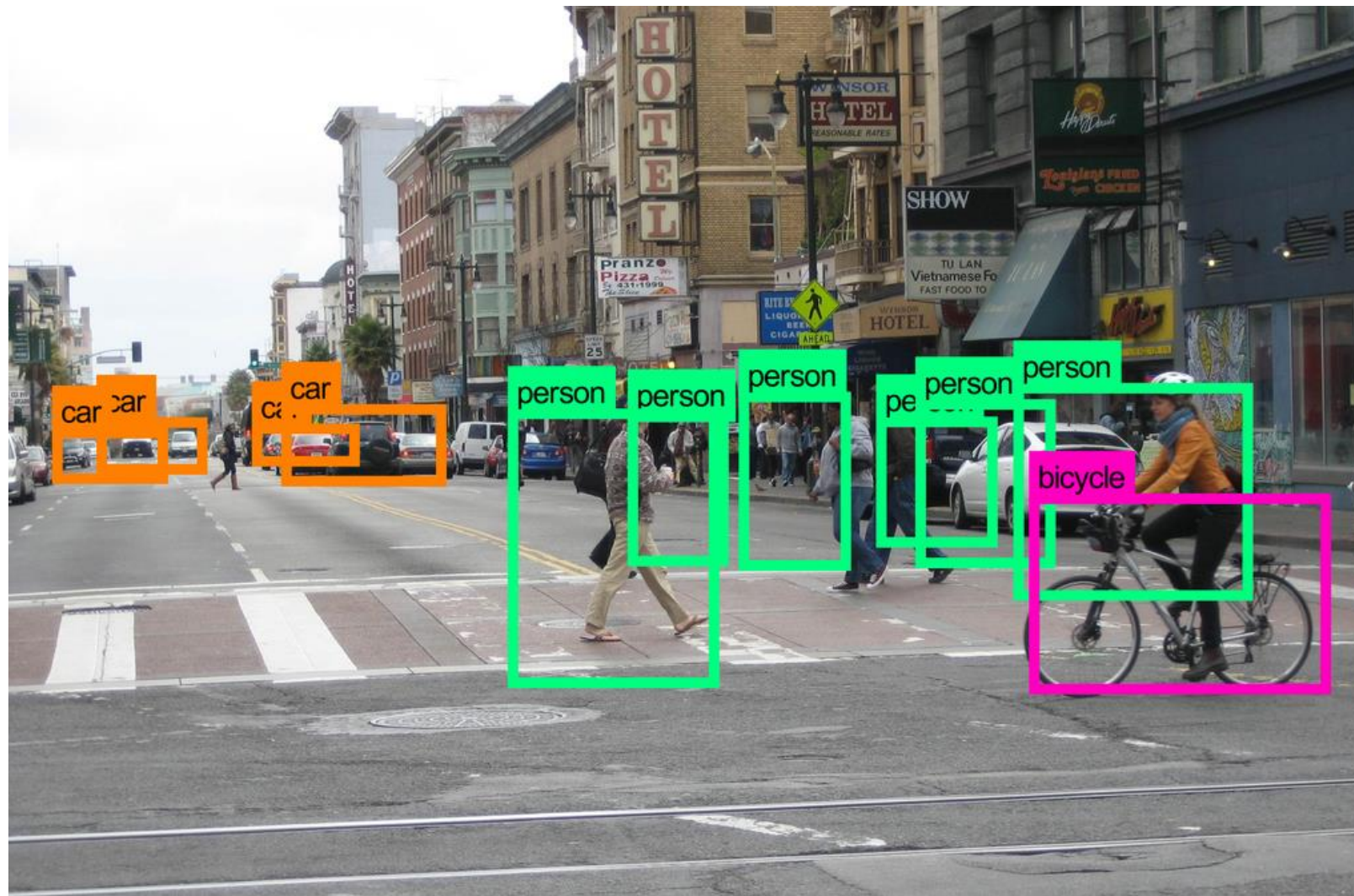
★★★★★ Rating: 4.5 - 104 votes - 1 hr 5 mins - 345 cal

This takeaway favourite is freezer-friendly and quick to reheat, giving you the chance to get ahead and save money, from BBC Good Food magazine.

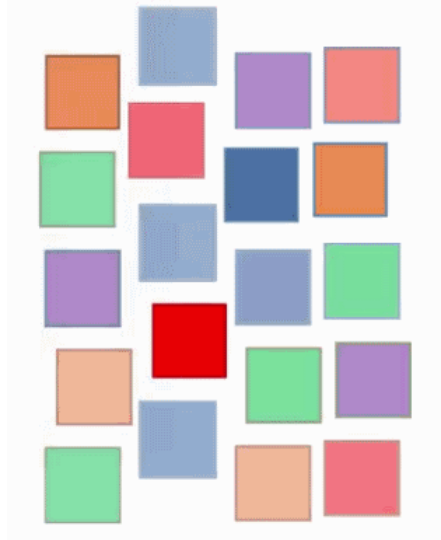
არა
სტრუქტურირებული
მონაცემები



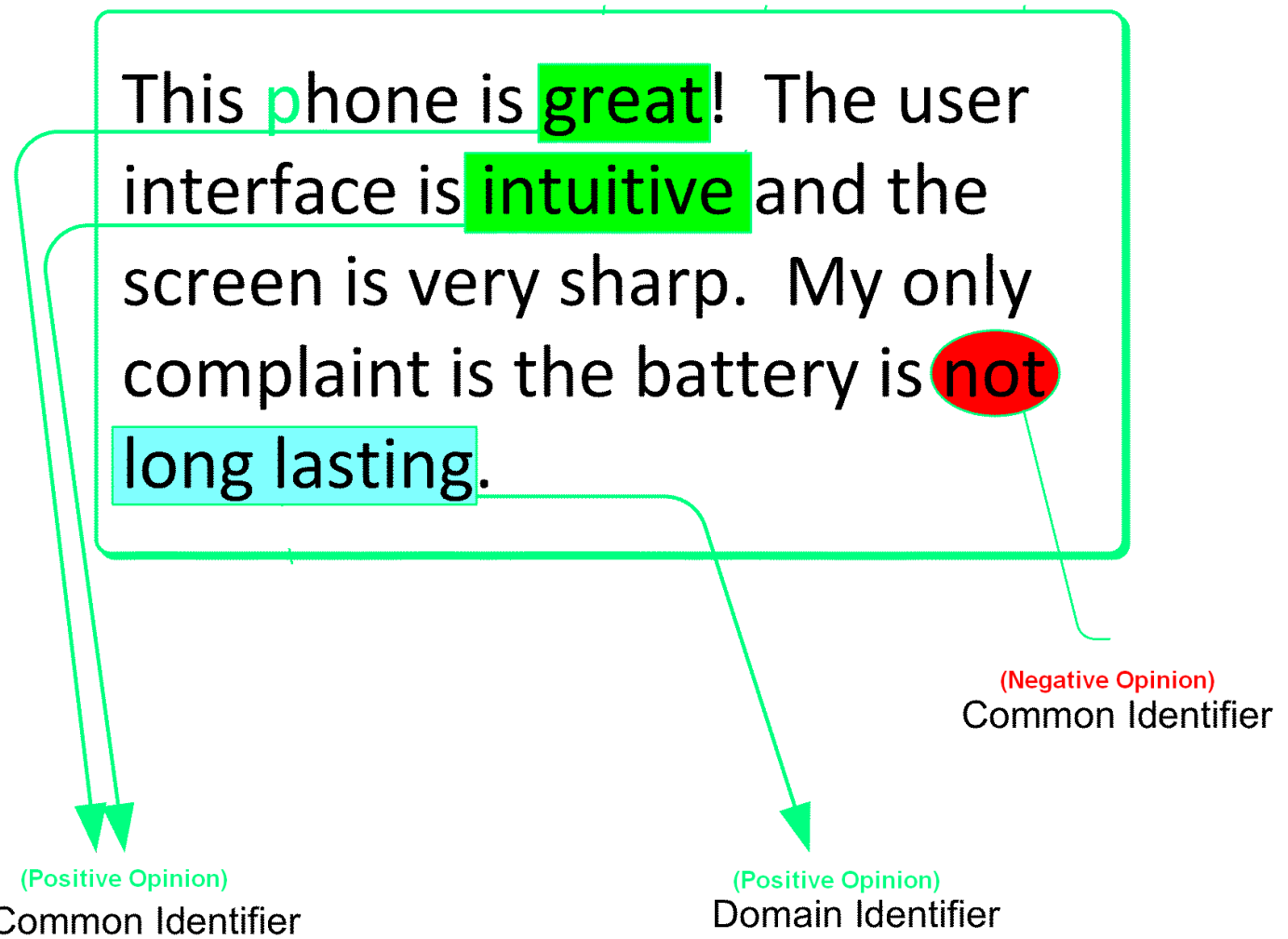
ფოტოები, ვიდეოები,
აუდიო ჩანაწერები,
ტექსტები



არა
სტრუქტურირებული
მონაცემები



ფოტოები, ვიდეოები,
აუდიო ჩანაწერები,
ტექსტები



მონაცემთა ბაზები - Databases

RELATIONAL

Posts (id, Title)

1	Title
---	-------

Comments

01	1	Comment 1
02	1	Comment 2

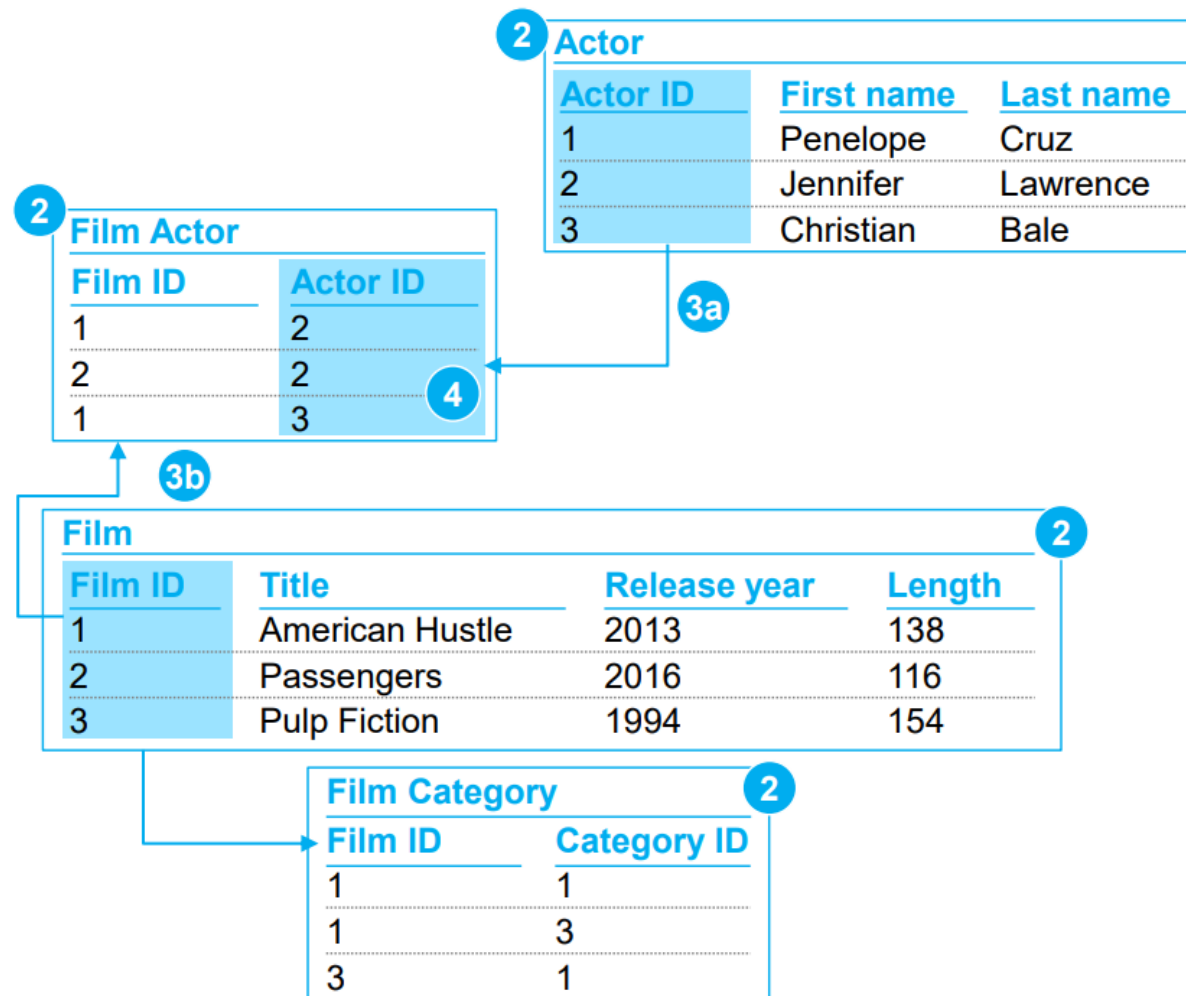
NON-RELATIONAL

Posts (id, Title, Comments / Image)

1	Title	Comment 1
		Comment 2
		Comment 3
<hr/>		
2	Title 2	Image

რელაციური მონაცემთა ბაზები (RDBMS)

1. რელაციური მონაცემთა ბაზა არის სხვადასხვა ცხრილების ერთობლიობა
2. ინფორმაცია გადანაწილებულია სხვადასხვა ცხრილში
3. ამ ცხრილების გადაბმა შესაძლებელია სპეციალური ველების მეშვეობით
4. ყველა ცხრილს აქვს ერთი ისეთი ველი რომლის მიხედვითაც ცხრილში არის უნიკალური ჩანაწერები



მონაცემთა ბაზები შედგება ცხრილებისგან

customer

item

sale

A 4x4 grid of orange squares, totaling 16 squares. The grid is composed of four rows and four columns of identical orange squares, separated by thin white lines.

ცხრილი შედგება სვეტებისგან (Columns) და მწკრივებისგან (Rows)

id	name	address	city	state	zip
1	Bill Smith	123 Main St	Hope	CA	98765
2	Mary Smith	123 Dorian St	Harmony	AZ	98765
3	Bob Smith	123 Laugh St	Humor	CA	98765

მწკრივს შეგვიძლია დავარქვათ ჩანაწერი

id	name	address	city	state	zip
1	Bill Smith	123 Main St	Hope	CA	98765
2	Mary Smith	123 Dorian St	Harmony	AZ	98765
3	Bob Smith	123 Laugh St	Humor	CA	98765

sale

id	item_id	cust_id	quan	price
1	1	2	3	2995
2	2	2	1	1995
3	1	1	1	2995

item

id	name	description
1	Pixels	64 RGB
2	Humor	Especially dry
3	Beauty	Inner beauty

customer

id	name	address	city	state	zip
1	Bill Smith	123 Main St	Hope	CA	98765
2	Mary Smith	123 Dorian St	Harmony	AZ	98765
3	Bob Smith	123 Laugh St	Humor	CA	98765

What is SQL

Structured Query Language



1

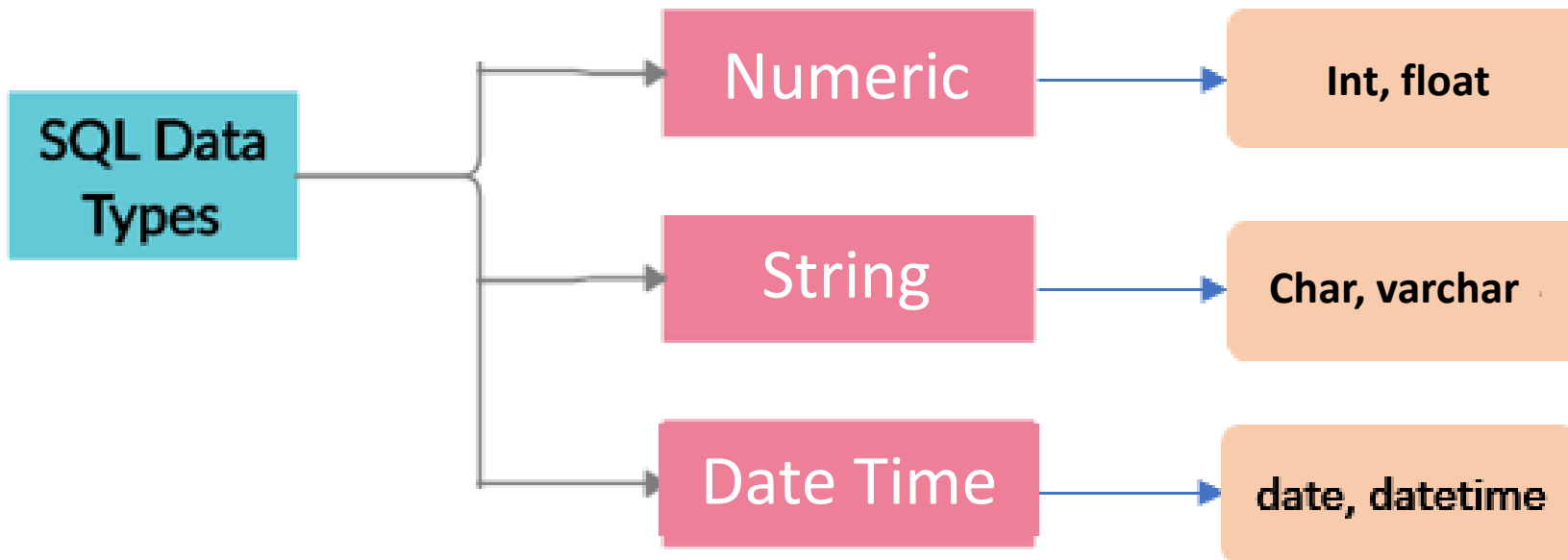
მონაცემთა ბაზებთან
კომუნიკაციისთვის
საჭირო ენა

2

თუ გვინდა რომ
წამოვიღოთ, შევქმათ
ან დავამუშავოთ
მონაცემები ბაზებში
უნდა გამოვიყენოთ
SQL

3

SQL-ის პრინციპების გაგების
შემთხვევაში, მონაცემთა
ანალიტიკოსის დანარჩენი
ამოცანების გაგება გაცილებით
მარტივია



- **integer:** 321 # whole number
- **float:** 321.0 # decimal

'a literal SQL string'

- **string:** "321" # character(s)

Id	Desc	DateTime
1	May 22th 2018, noon	2018-05-22 12:00:00.000
2	June 13th 2017, 9:00 morning	2017-06-13 09:00:00.000
3	April 22nd 2000, 3:00 afternoon	2000-04-22 15:00:00.000

<https://github.com/nutsaabazadze>

რესურსები SQL-ის სავარჯიშოდ

<https://datalemur.com/>

https://sqlzoo.net/wiki/SQL_Tutorial

<https://sqlbolt.com/>