



Hello World Bootcamp

SQL Databases





Course Outline



Content

- What is Database?
- What is SQL?
- ER Diagram
- CRUD
- Basic SQL Clauses
- JOIN Multiple Tables
- Intro to BigQuery

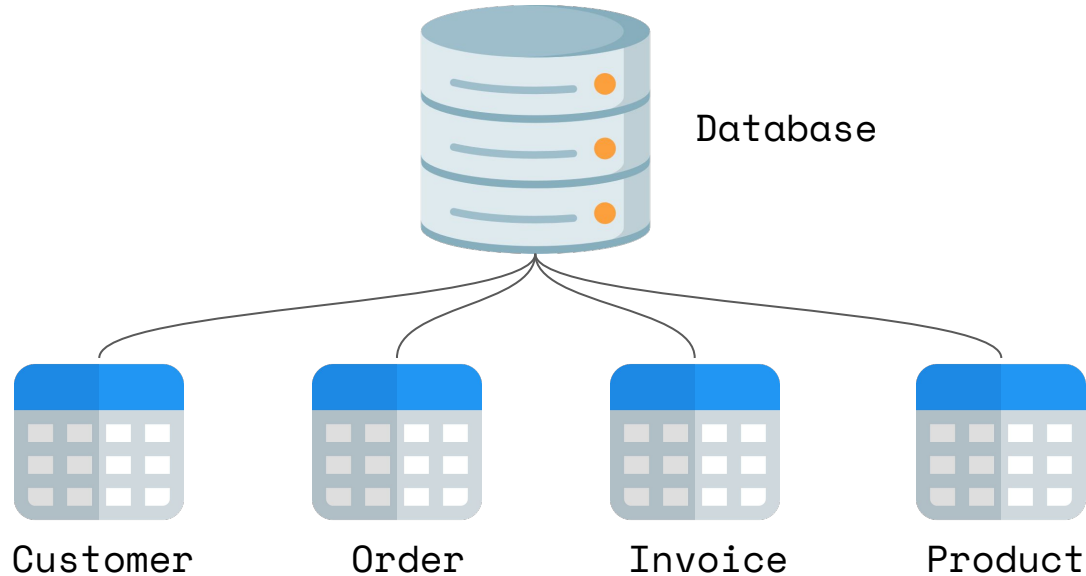




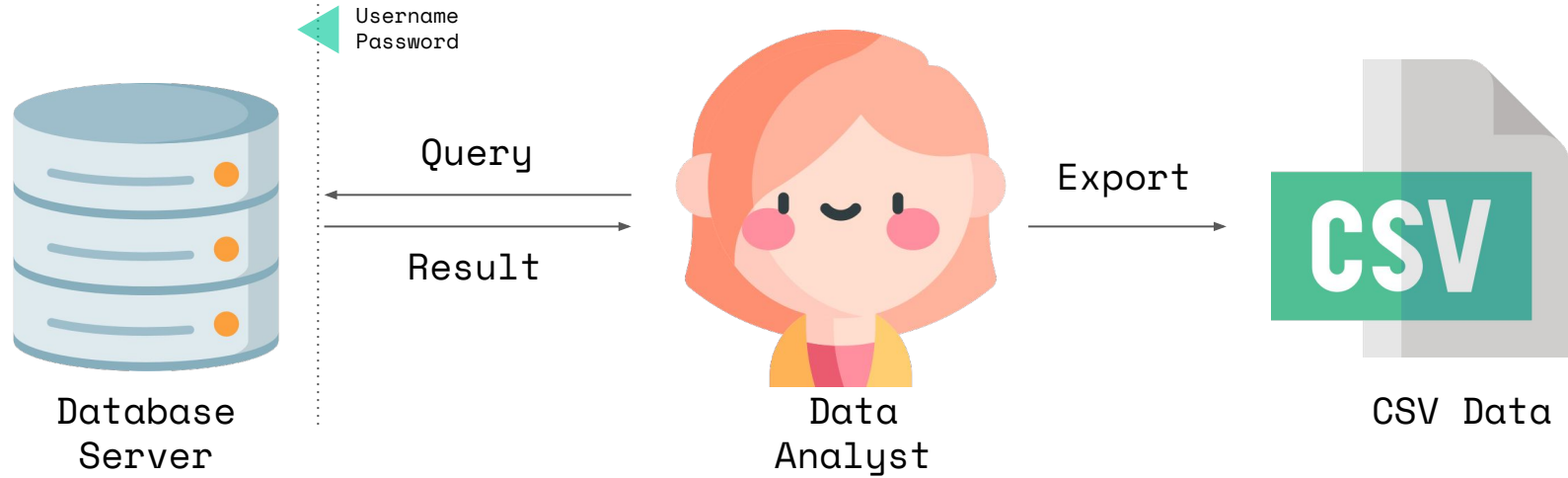
What is Database?



What is Database?



Data Analyst Workflow



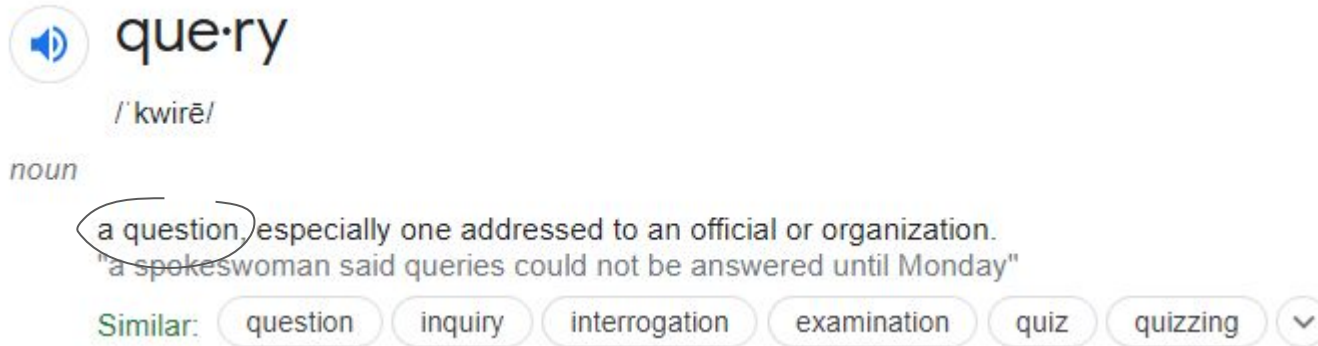


What is SQL?



What is SQL

Structured Query Language



Top programming languages for data scientists in 2022



Javier Candales Luna • January 21, 2022

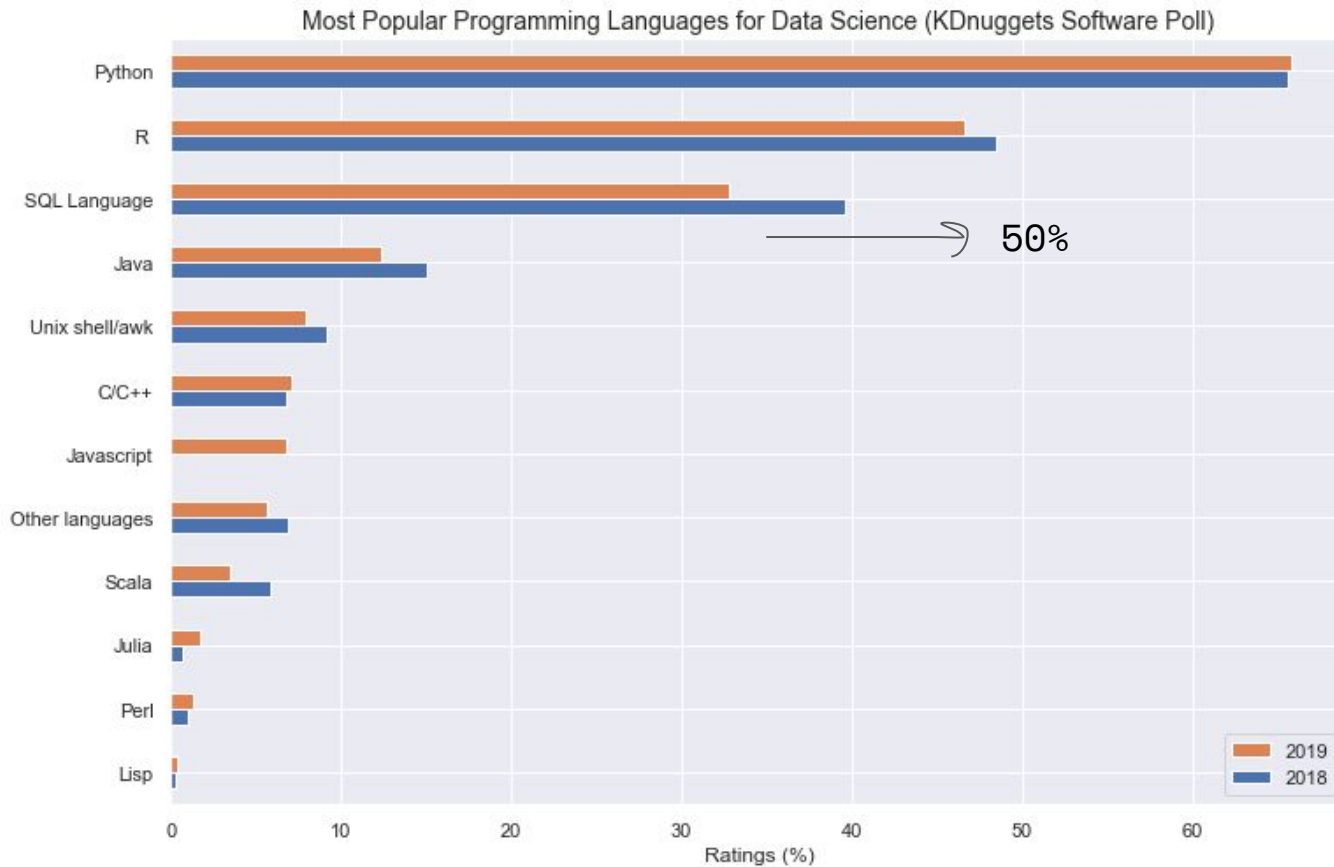


In this article, we will look at some of the top data science programming languages for 2022, and present the strengths and capabilities of each of them.

- Python
- R
- SQL
- Java
- Julia
- Scala
- C/C++
- JavaScript
- Swift
- Go
- MATLAB
- SAS

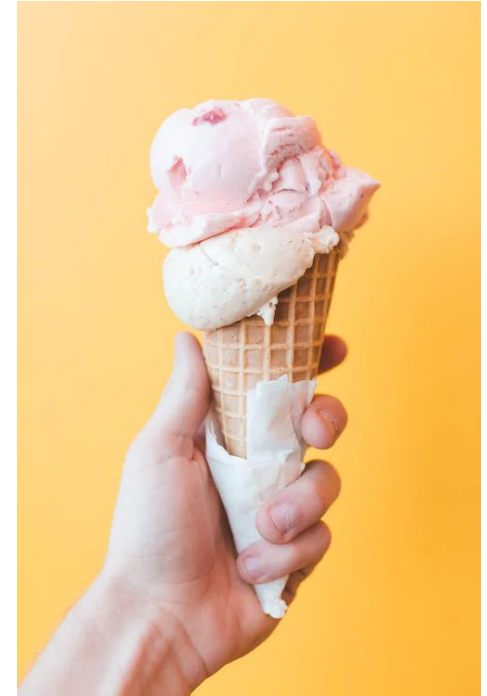
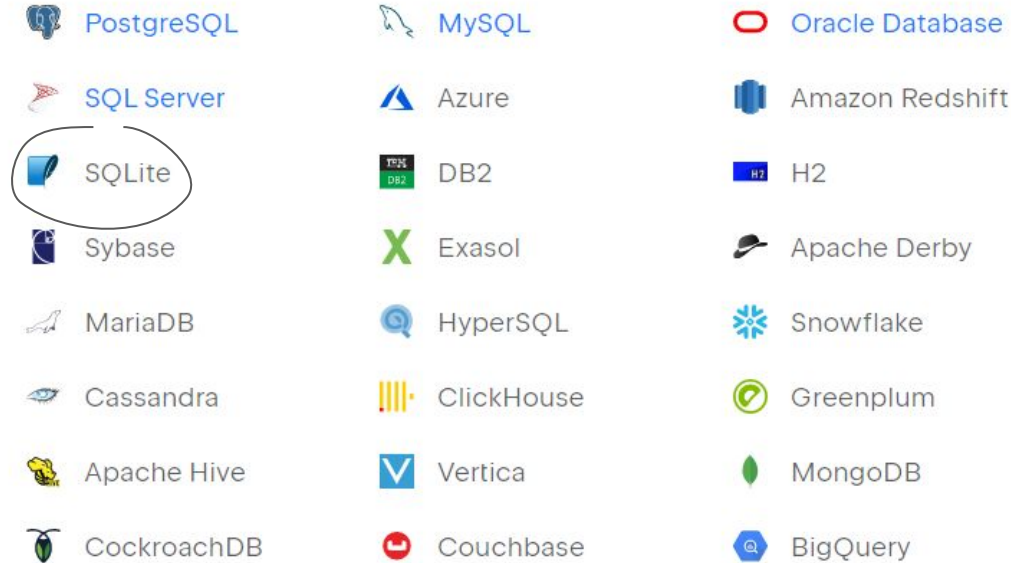
Top Three Tools for Data Science





SQL Flavours

A lot of flavours



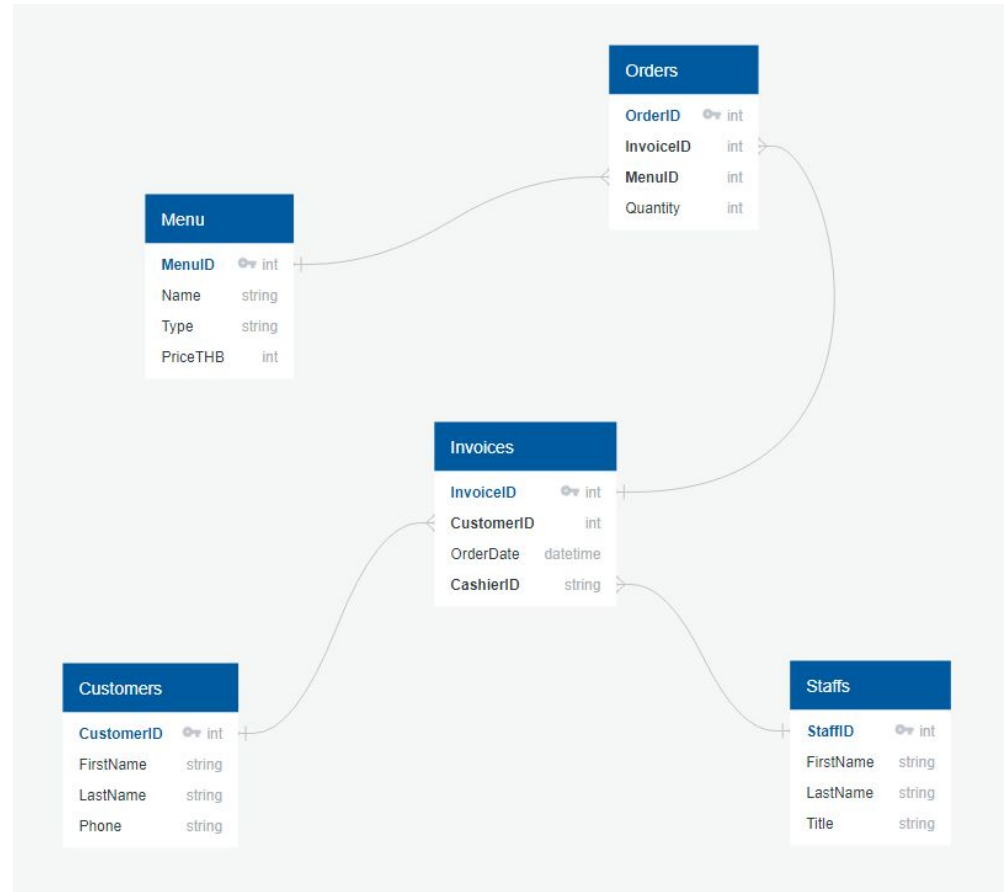


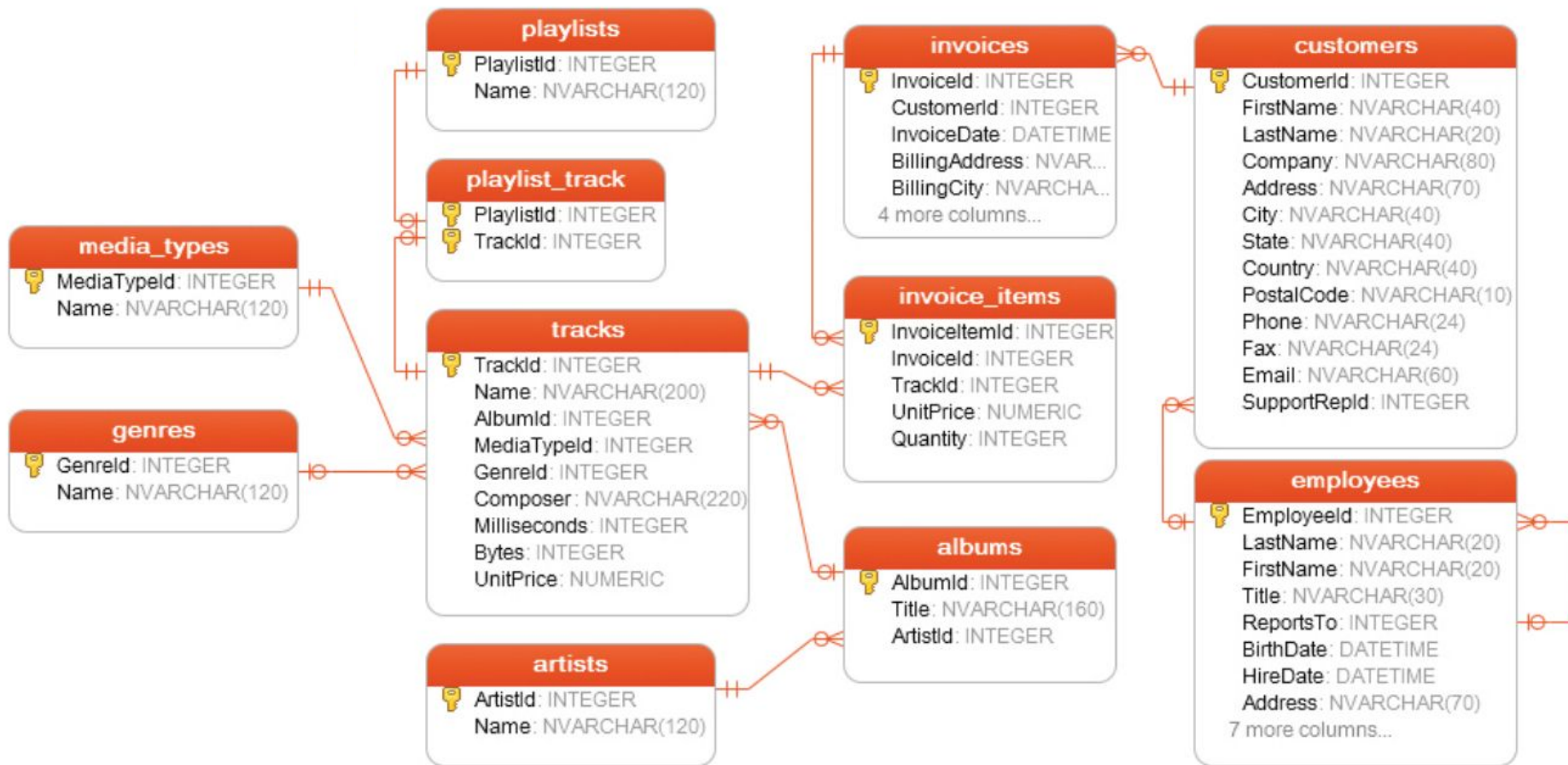
ER Diagram

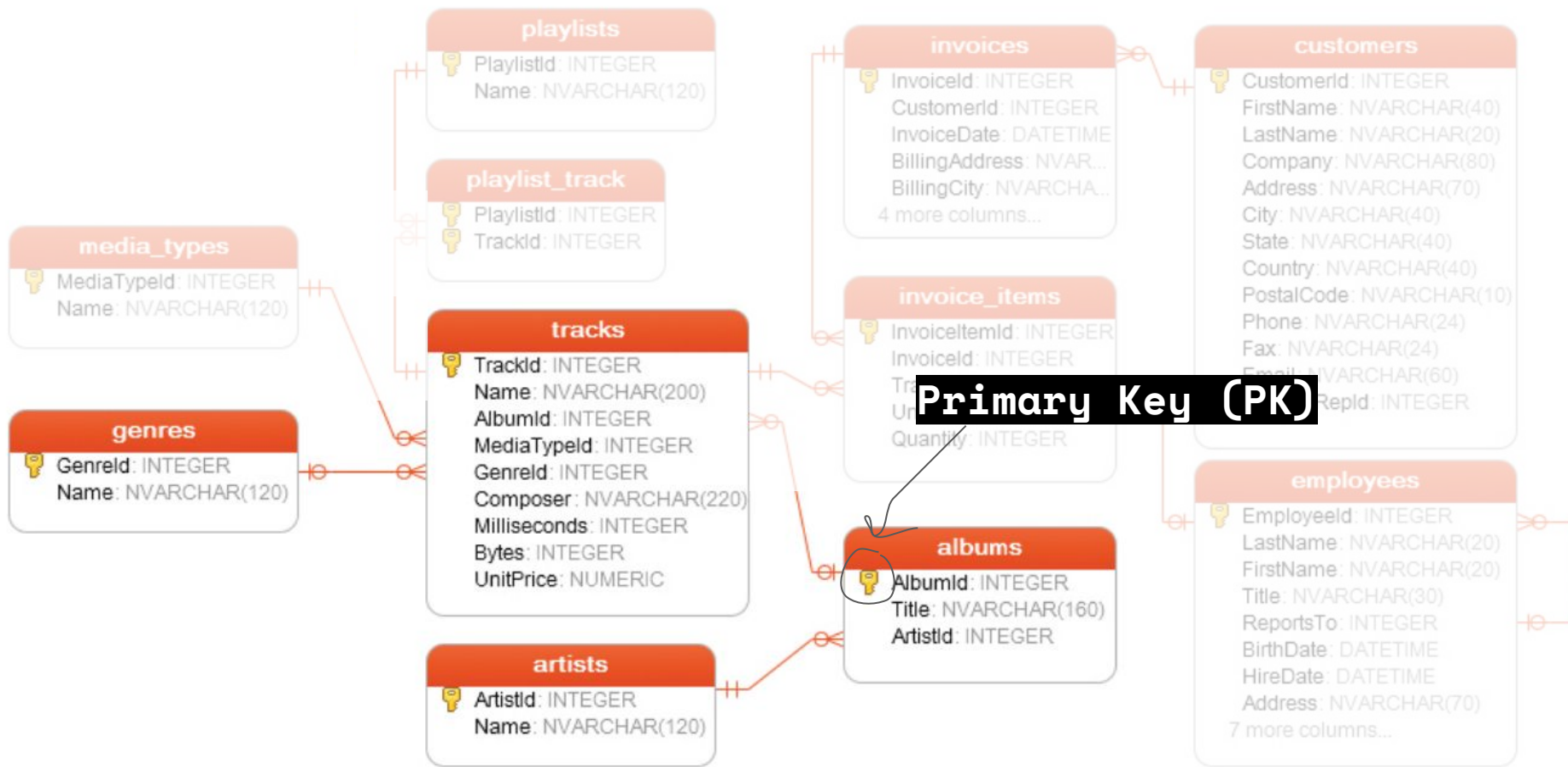


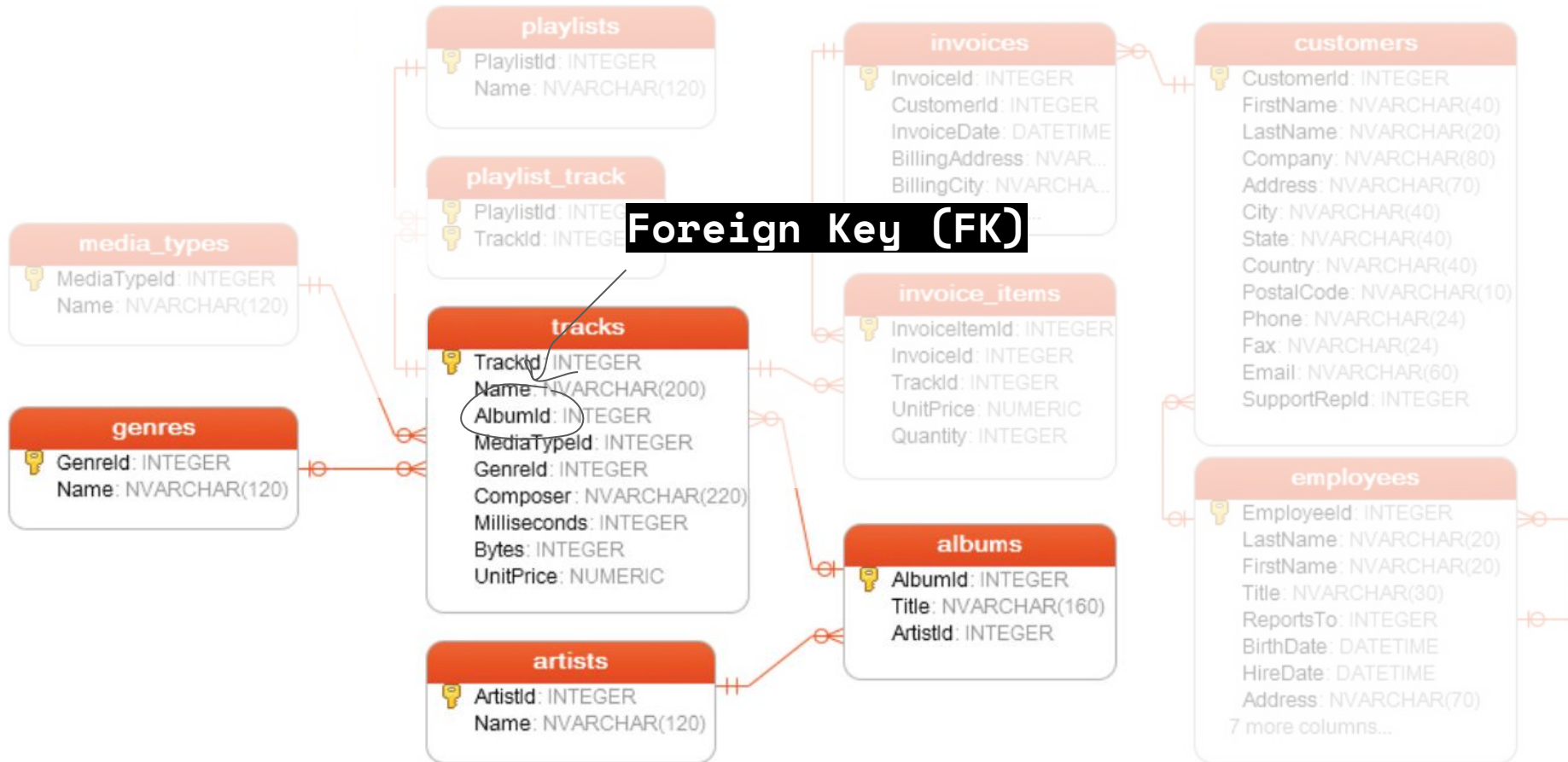
ER Diagram

Entity Relationship Diagram

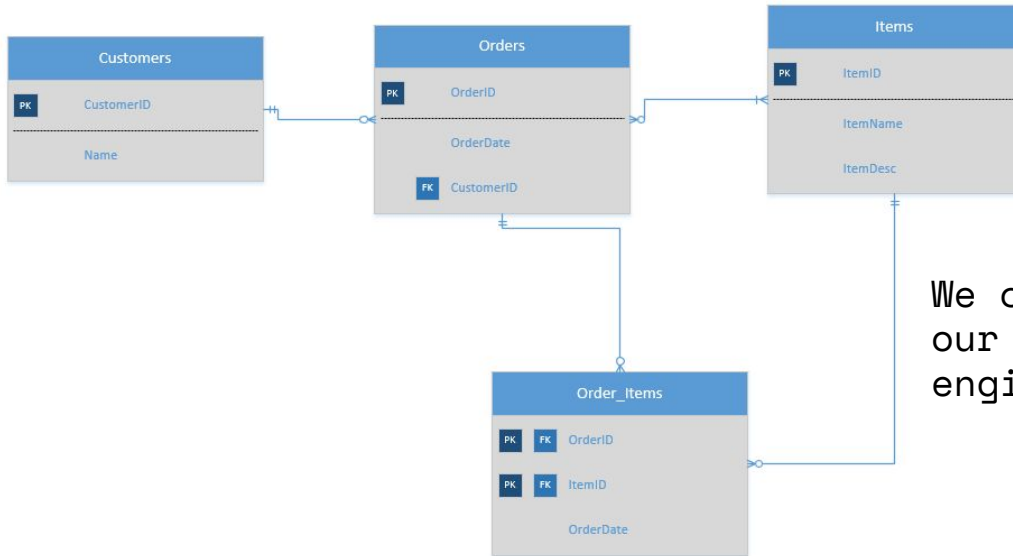








Data Model



We can **model relationships** in our data (Usually done by data/engineering team)





CRUD Operations



CRUD

- CREATE
- RETRIEVE (SELECT)
- UPDATE
- DELETE



CREATE

```
CREATE TABLE customer (  
    id int,  
    name text,  
    city text,  
    email text,  
    avg_spending real  
);
```



INSERT INTO

```
INSERT INTO customer VALUES  
  (1, "toy", "BKK", "toy@mail.com", 500.25),  
  (2, "joe", "BKK", "joe@mail.com", 125.50),  
  (3, "ann", "LON", "ann@mail.com", 999.55),  
  (4, "ken", "LON", "mae@mail.com", 658.20);
```



RETRIEVE aka. SELECT

```
SELECT * FROM customer;
```



Select every column from customer table



DROP TABLE

```
DROP TABLE customer;
```





Basic SQL for Data Analysts



Basic SQL Clauses

- ❑ SELECT
- ❑ FROM
- ❑ WHERE
- ❑ GROUP BY + AGGREGATE FUNCTIONS()
- ❑ HAVING
- ❑ ORDER BY



Simple Data Manipulation

Select Columns

ID	Name	City	Email	Spending
1001	Toy	BKK	toy@mail.com	1500
1002	Anna	LON	anna@mail.com	1200
1003	Marry	LON	marry@mail.com	900
1004	Ken	JPN	ken@mail.com	2500



Simple Data Manipulation

Filter rows

ID	Name	City	Email	Spending
1001	Toy	BKK	toy@mail.com	1500
1002	Anna	LON	anna@mail.com	1200
1003	Marry	LON	marry@mail.com	900
1004	Ken	JPN	ken@mail.com	2500



Simple Data Manipulation

Select columns + Filter rows

ID	Name	City	Email	Spending
1001	Toy	BKK	toy@mail.com	1500
1002	Anna	LON	anna@mail.com	1200
1003	Marry	LON	marry@mail.com	900
1004	Ken	JPN	ken@mail.com	2500

Select columns: id name and city

Filter: customers in London





SELECT



SELECT

```
SELECT * FROM customers;
```



SELECT

```
SELECT  
    firstname,  
    lastname,  
    country,  
    email  
FROM customers;
```



Rename Column using AS

```
SELECT
    firstname AS fname,
    lastname AS lname,
    country,
    email
FROM customers;
```



LIMIT

```
SELECT *  
FROM customers  
LIMIT 10;
```





WHERE



WHERE (filter rows)

```
SELECT
    firstname,
    lastname,
    country,
    email
FROM customers
WHERE country = 'USA';
```



WHERE IN

```
SELECT
    firstname,
    lastname,
    country,
    email
FROM customers
WHERE country IN ('USA', 'France', 'Belguim');
```



WHERE BETWEEN

```
SELECT
    firstname,
    lastname,
    country,
    email
FROM customers
WHERE customerid BETWEEN 10 AND 15;
```



WHERE Multiple Conditions

```
SELECT
    firstname,
    lastname,
    country,
    state
FROM customers
WHERE country = 'USA' AND state = 'CA';
```



WHERE LIKE

```
SELECT
```

```
    firstname,
```

```
    lastname,
```

```
    country,
```

```
    state
```

```
FROM customers
```

```
WHERE firstname LIKE 'J%';
```

% matches any character



WHERE LIKE

```
SELECT
    firstname,
    lastname,
    country,
    state
FROM customers
WHERE email LIKE '%@gmail.com';
```

find gmail users



WHERE Multiple Tables

```
SELECT *  
FROM artists, albums  
WHERE artists.artistid = albums.artistid;
```





Aggregate Functions



Aggregate Functions


```
SELECT  
    AVG(total),  
    SUM(total),  
    MIN(total),  
    MAX(total),  
    COUNT(total)  
FROM invoices;
```



AGG + GROUP BY

```
SELECT  
    country,  
    COUNT(*) AS n  
FROM customers  
GROUP BY country;
```

Tips: columns in group by should be presented in SELECT too





Aggregate Functions



ORDER BY

```
SELECT *  
FROM customers  
ORDER BY country DESC;
```



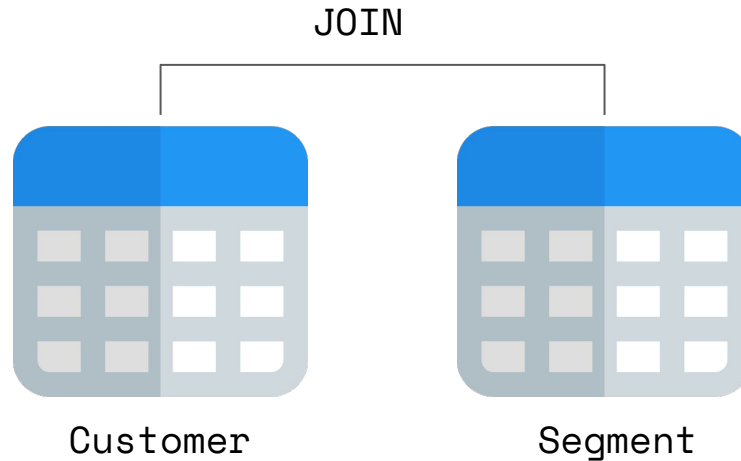


JOIN Multiple Tables



What is JOIN?

Getting data from multiple tables



SQL JOIN = VLOOKUP()

Customer

ID	Name	City
1001	Toy	BKK
1002	Anna	LON
1003	Marry	LON

Segment

ID	SegName	Cust_ID
1	Deal Hunter	1001
2	Price Sensitive	1002
3	Premium	1003

Join PK=FK



Result Set

Cool!

ID	Name	City	SegName
1001	Toy	BKK	Deal Hunter
1002	Anna	LON	Price Sensitive
1003	Marry	LON	Premium

ง่ายจนงง!



JOIN Example

```
SELECT * FROM table1  
JOIN table2  
ON table1.pk = table2.fk;
```



Primary Key



Foreign Key



JOIN syntax

```
SELECT * FROM customer  
JOIN segment  
ON customer.id = segment.cust_id;
```

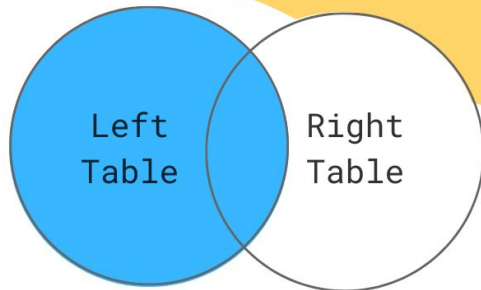


Four Join Types

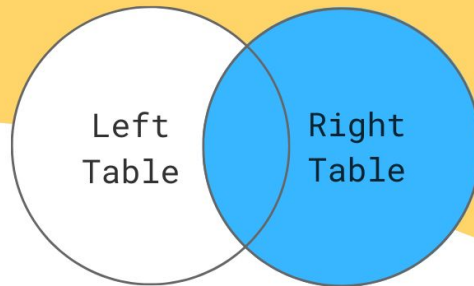
- ❏ INNER JOIN
- ❏ LEFT JOIN
- ❏ RIGHT JOIN
- ❏ FULL JOIN



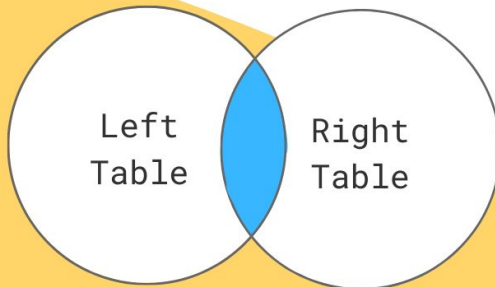
LEFT JOIN



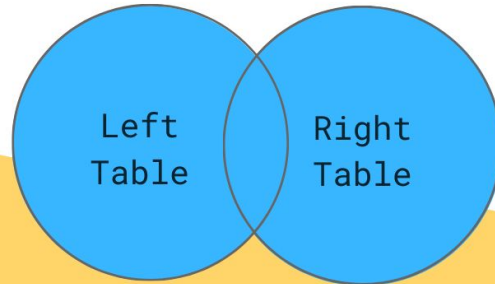
RIGHT JOIN



INNER JOIN



FULL JOIN



INNER JOIN (default)

Customer

ID	Name	City
1001	Toy	BKK
1002	Anna	LON
1003	Marry	LON
1004	Ken	JPN

Segment

ID	SegName	Cust_ID
1	Deal Hunter	1001
2	Price Sensitive	1002
3	Premium	1003

Join PK=FK



Result Set

inner join

Only Matched Rows Return

ID	Name	City	ID	SegName	Cust_ID
1001	Toy	BKK	1	Deal Hunter	1001
1002	Anna	LON	2	Price Sensitive	1002
1003	Marry	LON	3	Premium	1003



LEFT JOIN

Customer

ID	Name	City
1001	Toy	BKK
1002	Anna	LON
1003	Marry	LON
1004	Ken	JPN

Segment

ID	SegName	Cust_ID
1	Deal Hunter	1001
2	Price Sensitive	1002
3	Premium	1003

Join PK=FK



Result Set

left join

All rows in left table will be in the result set

ID	Name	City	ID	SegName	Cust_ID
1001	Toy	BKK	1	Deal Hunter	1001
1002	Anna	LON	2	Price Sensitive	1002
1003	Marry	LON	3	Premium	1003
1004	Ken	JPN	NULL	NULL	NULL



RIGHT JOIN

Customer

ID	Name	City
1001	Toy	BKK
1002	Anna	LON
1003	Marry	LON
1004	Ken	JPN

Segment

ID	SegName	Cust_ID
1	Deal Hunter	1001
2	Price Sensitive	1002
3	Premium	1003

Join PK=FK



Result Set

right join

All rows in right table will be in the result set

ID	SegName	Cust_ID	ID	Name	City
1	Deal Hunter	1001	1001	Toy	BKK
2	Price Sensitive	1002	1002	Anna	LON
3	Premium	1003	1003	Marry	LON



FULL JOIN

Customer

ID	Name	City
1001	Toy	BKK
1002	Anna	LON
1003	Marry	LON
1004	Ken	JPN

Segment

ID	SegName	Cust_ID
1	Deal Hunter	1001
2	Price Sensitive	1002
3	Premium	1003
99	New Segment	9999

Join PK=FK



Result Set full join

All rows in left table will be in the result set

ID	Name	City	ID	SegName	Cust_ID
1001	Toy	BKK	1	Deal Hunter	1001
1002	Anna	LON	2	Price Sensitive	1002
1003	Marry	LON	3	Premium	1003
1004	Ken	JPN	NULL	NULL	NULL
NULL	NULL	NULL	99	New Segment	9999

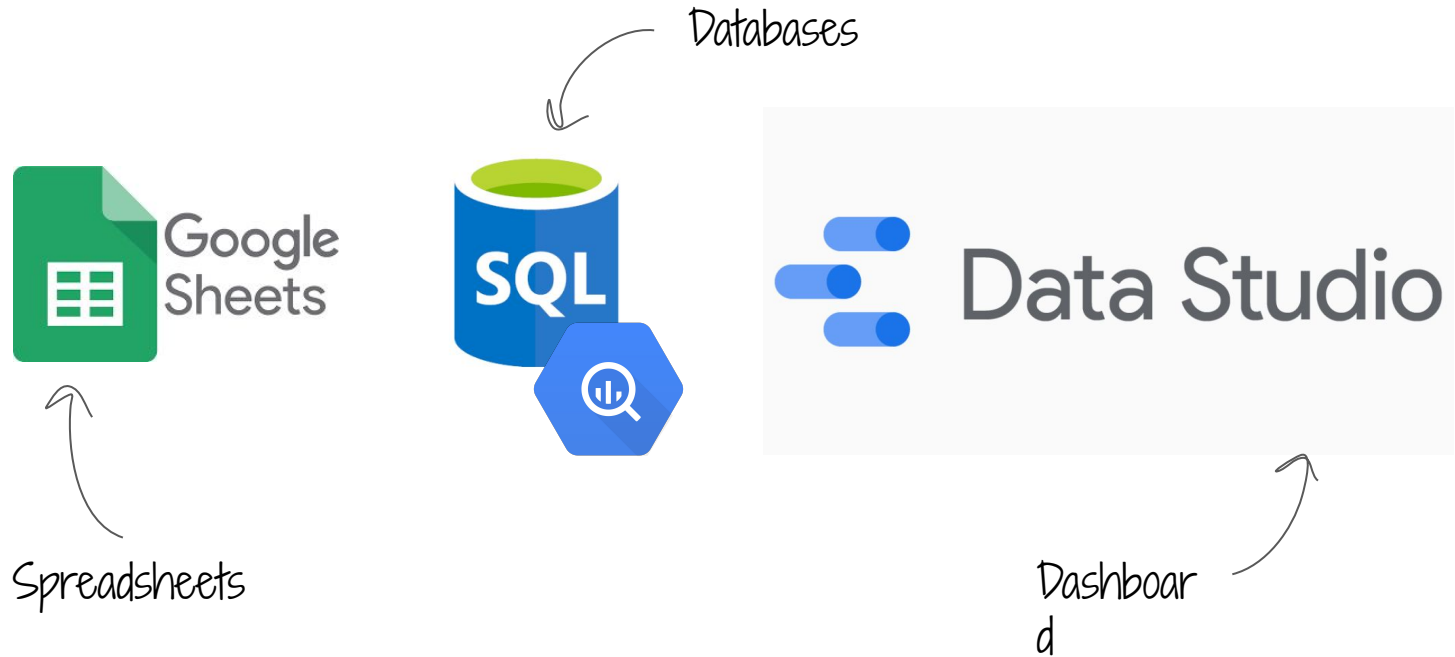




Modern Data Warehouse



Core Data Analytics Stack



Core Data Analytics Stack



Store



Analyze



Present



Separate Compute/ Storage

Get \$300 in free credits and free usage of 20+ products →

Dream, build, and transform with Google Cloud

Build apps faster, make smarter business decisions, and connect people anywhere.

[Go to console](#)

[Contact sales](#)

What's new For developers

EVENT

Today, meet tomorrow. Join us October 11-13 for Google Cloud Next



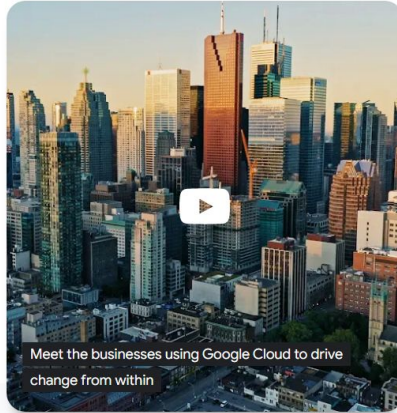
2-MINUTE QUIZ

How data-driven is your company culture?



REPORT

Rank your digital transformation against global leaders



Compute



Storage



Databases



Data analytics



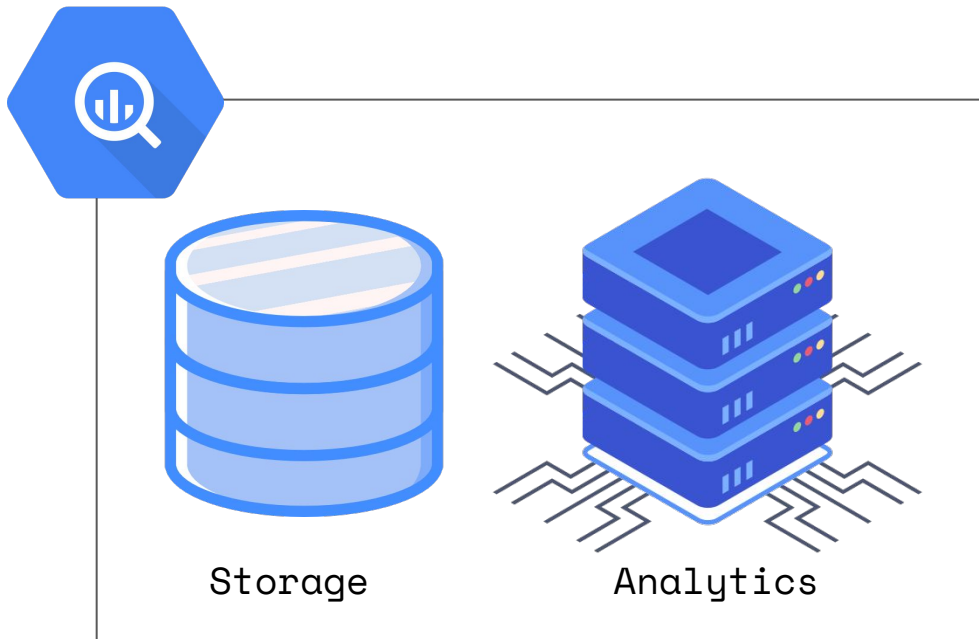
AI and machine learning



Networking



BigQuery



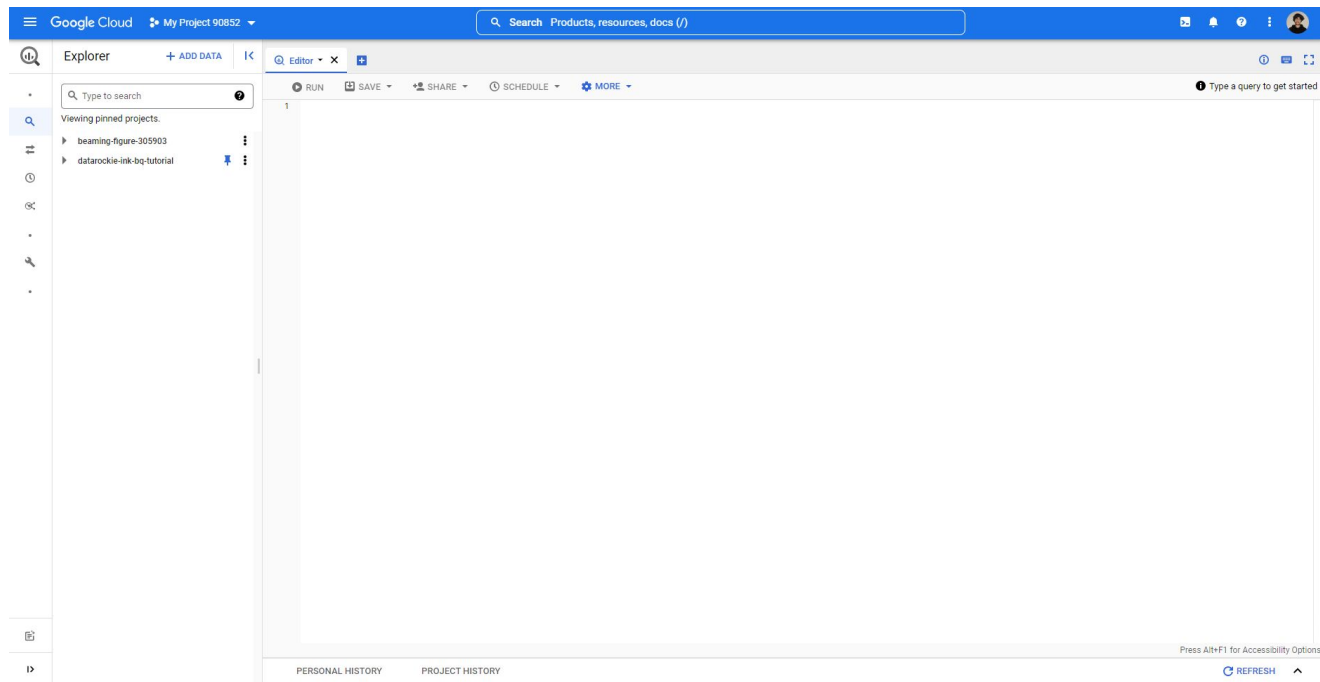
What is BigQuery?



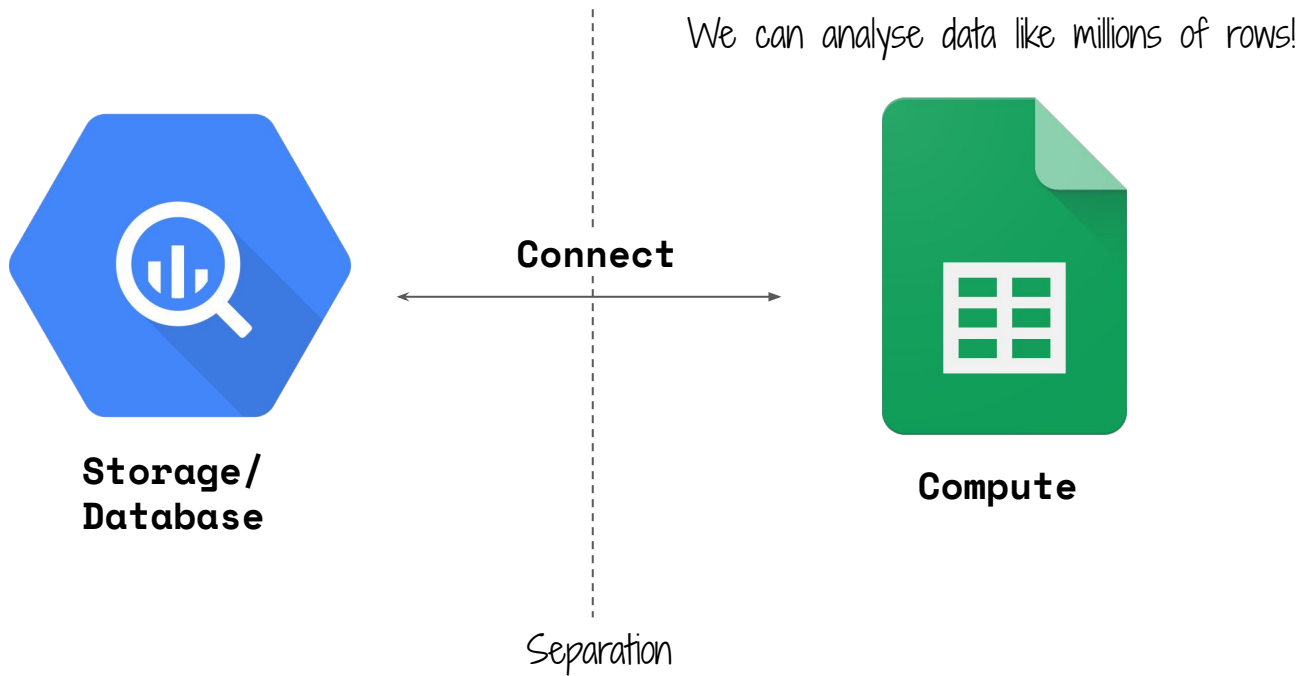
BigQuery is a fully managed enterprise data warehouse that helps you manage and analyze your data with built-in features like machine learning, geospatial analysis, and business intelligence. BigQuery's serverless architecture lets you use SQL queries to answer your organization's biggest questions with zero infrastructure management. BigQuery's scalable, distributed analysis engine lets you query terabytes in seconds and petabytes in minutes.



BigQuery UI



BigQuery <> Spreadsheets



BigQuery in Sheets

Untitled spreadsheet ☆ 📄 🌐

File Edit View Insert Format Data Tools Extensions Help Last edit was seconds ago

100%

store 10K Rows Refresh options Schedule refresh Connection settings Learn more

Chart Pivot table Function Extract + Calculated column Column stats

PREVIEW

Row_ID	Order_ID	Order_Dat	Ship_Date	Ship_Mod	Customer	Customer	Segment	Country	City	State	Postal_Co	Region	Product_ID	Category	Sub_Cate	Prodi
6231	CA-2017-127656	7/11/2017	7/17/2017	Standard Class	NW-18400	Natalie Webber	Consumer	United States	Waterloo	Iowa	50701	Central	OFF-AR-1000114	Office Supplies	Art	Staples
4190	CA-2016-157714	9/26/2016	10/1/2016	Second Class	CS-12175	Charles Sheldon	Corporate	United States	Iowa City	Iowa	52240	Central	OFF-PA-1000402	Office Supplies	Paper	Hammm
3426	CA-2015-153381	9/24/2015	9/28/2015	Standard Class	DE-13255	Deanna Eno	Home Office	United States	Dubuque	Iowa	52001	Central	OFF-BI-1000152	Office Supplies	Binders	Acco P
3427	CA-2015-153381	9/24/2015	9/28/2015	Standard Class	DE-13255	Deanna Eno	Home Office	United States	Dubuque	Iowa	52001	Central	FUR-CH-1000094	Furniture	Chairs	Hon Ol
8001	US-2015-151407	11/8/2015	11/12/2015	Standard Class	RD-19585	Rob Dowd	Consumer	United States	Dubuque	Iowa	52001	Central	TEC-PH-1000381	Technology	Phones	Cisco S
8416	CA-2017-107265	4/6/2017	4/12/2017	Standard Class	ML-17755	Max Ludwig	Home Office	United States	Marion	Iowa	52302	Central	OFF-PA-1000047	Office Supplies	Paper	Easy-st
9251	CA-2016-105354	12/2/2016	12/6/2016	Standard Class	PW-19030	Pauline Webber	Corporate	United States	Marion	Iowa	52302	Central	OFF-BI-1000110	Office Supplies	Binders	GBC W
9915	CA-2017-160927	1/29/2017	1/31/2017	Second Class	TM-21010	Tamara Manning	Consumer	United States	Marion	Iowa	52302	Central	OFF-PA-1000384	Office Supplies	Paper	Xerox 1
9916	CA-2017-160927	1/29/2017	1/31/2017	Second Class	TM-21010	Tamara Manning	Consumer	United States	Marion	Iowa	52302	Central	OFF-PA-1000017	Office Supplies	Paper	Xerox 1
9917	CA-2017-160927	1/29/2017	1/31/2017	Second Class	TM-21010	Tamara Manning	Consumer	United States	Marion	Iowa	52302	Central	OFF-ST-1000159	Office Supplies	Storage	Tenex f
9918	CA-2017-160927	1/29/2017	1/31/2017	Second Class	TM-21010	Tamara Manning	Consumer	United States	Marion	Iowa	52302	Central	FUR-FU-1000001	Furniture	Furnishings	DAX Ve
1009	US-2017-106705	12/26/2017	1/1/2018	Standard Class	PO-18850	Patrick O'Brill	Consumer	United States	Burlington	Iowa	52601	Central	OFF-PA-1000156	Office Supplies	Paper	Recycle
9210	CA-2017-142776	12/11/2017	12/14/2017	Second Class	RS-19870	Roy Skaria	Home Office	United States	Burlington	Iowa	52601	Central	OFF-EN-1000314	Office Supplies	Envelopes	Pastel i
9211	CA-2017-142776	12/11/2017	12/14/2017	Second Class	RS-19870	Roy Skaria	Home Office	United States	Burlington	Iowa	52601	Central	OFF-BI-1000201	Office Supplies	Binders	Wilson
771	CA-2017-104220	1/30/2017	2/5/2017	Standard Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	OFF-BI-1000103	Office Supplies	Binders	Cardini
772	CA-2017-104220	1/30/2017	2/5/2017	Standard Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	TEC-PH-100046	Technology	Phones	AT&T 8
773	CA-2017-104220	1/30/2017	2/5/2017	Standard Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	OFF-BI-1000030	Office Supplies	Binders	GBC In
774	CA-2017-104220	1/30/2017	2/5/2017	Standard Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	OFF-BI-1000391	Office Supplies	Binders	DXL Ar
775	CA-2017-104220	1/30/2017	2/5/2017	Standard Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	OFF-AR-1000464	Office Supplies	Art	Boston
				Class	BV-11245	Benjamin Venier	Corporate	United States	Des Moines	Iowa	50315	Central	FUR-FU-1000255	Furniture	Furnishings	C-Line i
				y	LC-16870	Lena Cacioppo	Consumer	United States	Des Moines	Iowa	50315	Central	OFF-PA-1000002	Office Supplies	Paper	Xerox 2
					DM-13345	Denise Monton	Corporate	United States	Des Moines	Iowa	50315	Central	OFF-PA-1000384	Office Supplies	Paper	Xerox 1

Preview of 10K rows 7:56 AM Refresh preview

+ Sheet 1 schedules Extract 2 Extract 1 store





Hello World Bootcamp

SQL Databases

