



# Parte 2. Base de Datos

## Descripción de los datos

Base de datos de los viajes en taxi de Chicago:

La tabla `neighborhoods`, con información sobre los barrios de la ciudad:

- `neighborhood_id`: código del vecindario.
- `name`: nombre del vecindario.

La tabla `cabs`, con información sobre los taxis:

- `cab_id`: código único del vehículo.
- `vehicle_id`: identificador técnico del vehículo.
- `company_name`: la compañía dueña del automóvil.

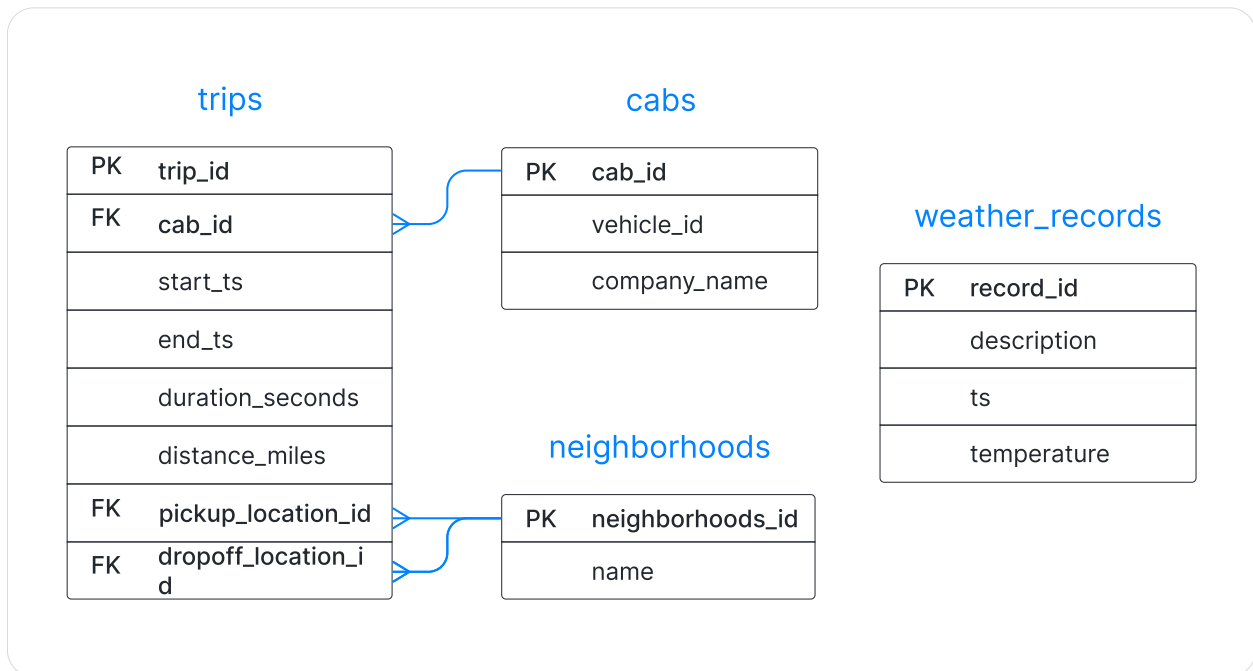
La tabla `trips`, con información sobre viajes:

- `trip_id`: código de viaje.
- `cab_id`: el código del automóvil usado para el viaje.
- `start_ts`: fecha y hora del inicio del viaje (tiempo redondeado a la hora más próxima).
- `end_ts`: fecha y hora del fin del viaje (tiempo redondeado a la hora más próxima).
- `duration_seconds`: la duración del viaje en segundos.
- `distance_miles`: la distancia del viaje en millas.
- `pickup_location_id`: el código del vecindario donde inició el viaje.
- `dropoff_location_id`: el código del vecindario donde terminó el viaje.

La tabla `weather_records` (registros meteorológicos), con información sobre el clima:

- `record_id`: código del registro de observación meteorológica.
- `ts`: fecha y hora de la observación (tiempo redondeado a la hora más próxima).
- `temperature`: temperatura a la hora de la observación.
- `description`: una breve descripción de las condiciones meteorológicas (p. ej., lluvia ligera o nubes dispersas).

## Disposición de la tabla



No existe un enlace directo entre las tablas `trips` y `weather_records` en la base de datos. Puedes vincular estas tablas por la hora de inicio ( `trips.start_ts` ) y el momento de las observaciones meteorológicas ( `weather_records.ts` ).

## Ejercicio 1

Tienes una base de datos con los viajes en taxi. El plan era tener 10 550 vehículos disponibles, lo que cubre la demanda del usuario; sin embargo, el equipo recibió muchas quejas de que no había vehículos suficientes. ¿Cuántos taxis hay actualmente en las calles? La información sobre todos los automóviles suficientes está en la tabla `cabs` .

1. Ve al servidor remoto.
2. Conéctate a la base de datos `chicago_taxi` con el nombre de usuario `morty` y la contraseña `smith` .
3. Cuenta el número total de automóviles en la tabla `cabs` . Recuerda que un automóvil podría pertenecer a distintas compañías.

Esto es lo que debes adjuntar en la respuesta:

1. Número de automóviles.
2. La solicitud que usaste para resolver el problema.

### Resultados del ejercicio 1:

Número de automóviles:

- 5529

La solicitud que utilizaste para resolver la tarea.

```
*SELECT COUNT (cab_id) AS cab_id FROM cabs;
```

## Ejercicio 2

En la tabla `cabs`, cuenta el número de automóviles de cada compañía. Ordena los valores en orden descendente. El equipo piensa que algunas compañías no tuvieron suficientes automóviles disponibles.

Devuelve las compañías con menos de 100 automóviles. Llama `cnt` (contados) al campo con el número de automóviles, y `company_name` al campo con el nombre de la compañía.

Para resolver el problema, utiliza el operador `HAVING`, una analogía de `WHERE` para las funciones agregadas. Estudia la documentación para aprender cómo funciona el operador:

(<https://postgrespro.com/docs/postgrespro/11/queries-table-expressions> (materiales en inglés))

Esto es lo que debes adjuntar en la respuesta:

1. Una lista de compañías con menos de 100 automóviles.
2. La solicitud que usaste para resolver el problema.

Nota: la consola muestra una lista incompleta. Para verla en su totalidad, presiona Enter o usa las flechas de tu teclado.

### Resultados del ejercicio 2:

Lista de las compañías con menos de 100 automóviles.

*	company_name	cnt
-----+-----		
	Nova Taxi Affiliation Llc	97
	Patriot Taxi DbA Peace Taxi Associat	89
	Blue Diamond	85

Checker Taxi Affiliation		81
Chicago Medallion Management		80
Chicago Independents		69
24 Seven Taxi		67
Checker Taxi		60
American United		55
Chicago Medallion Leasing INC		53
Top Cab Affiliation		49
KOAM Taxi Association		48
Chicago Taxicab		38
Norshore Cab		34
Gold Coast Taxi		20
Metro Group		20
Service Taxi Association		18
5 Star Taxi		14
American United Taxi Affiliation		8
Metro Jet Taxi A		8
Setare Inc		7
Leonard Cab Co		5
4615 - 83503 Tyrone Henderson		1
5062 - 34841 Sam Mestas		1
4623 - 27290 Jay Kim		1
5997 - 65283 AW Services Inc.		1
2092 - 61288 Sbeih company		1
1469 - 64126 Omar Jada		1
2733 - 74600 Benny Jona		1
2192 - 73487 Zeymane Corp		1
5006 - 39261 Salifu Bawa		1
3556 - 36214 RC Andrews Cab		1
3721 - Santamaria Express, Alvaro Santamaria		1
2809 - 95474 C & D Cab Co Inc.		1
2241 - 44667 - Felman Corp, Manuel Alonso		1
3620 - 52292 David K. Cab Corp.		1
2823 - 73307 Lee Express Inc		1
6057 - 24657 Richard Addo		1
6742 - 83735 Tasha ride inc		1
1085 - 72312 N and W Cab Co		1
3591 - 63480 Chuks Cab		1
0118 - 42111 Godfrey S.Awir		1
6574 - Babylon Express Inc.		1

3094 - 24059 G.L.B. Cab Co		1
5874 - 73628 Sergey Cab Corp.		1
6743 - 78771 Luhak Corp		1
5074 - 54002 Ahzmi Inc		1
3623 - 72222 Arrington Enterprises		1
4053 - 40193 Adwar H. Nikola		1
Chicago Star Taxicab		1
3011 - 66308 JBL Cab Inc.		1
(51 rows)		

La solicitud que te ayudó a resolver la tarea.

```
SELECT company_name, COUNT (cab_id) AS cnt FROM cabs GROUP BY company_r
```

## Ejercicio 3

La aplicación de taxis calcula el coeficiente del costo del viaje. Si el clima es bueno, el valor del coeficiente es 1. Si llueve o hay tormentas en el exterior, el coeficiente aumenta a 2. El equipo tiene una hipótesis de que hay un error en el cálculo del coeficiente. Para revisar el cálculo del coeficiente, el equipo necesita una selección de datos: el área de desarrollo puede comparar el coeficiente con los datos en los logs y corregir el bug. Tu tarea es obtener una selección.

Para hacerlo:

1. Obtén una descripción de las condiciones meteorológicas de la tabla `weather_records` para cada hora.
2. Divide todas las horas en dos grupos a través del operador `CASE`: Está `Bad` ("mal") si el campo `description` contiene las palabras "rain" (lluvia) o "storm" (tormenta); `Good` ("bien"), para todas las demás horas.
3. Pon el nombre `weather_conditions` al campo resultante.

La tabla resultante debe tener dos campos: fecha y hora (`ts`) y `weather_conditions`.

Haz una selección para el periodo entre 11-05-2017 12:00 a. m. a 11-06-2017 12:00 a. m.

Esto es lo que debes adjuntar en la respuesta:

1. La tabla resultante con los datos para el periodo especificado.

2. La solicitud que ayudó a resolver el problema.

### Resultados del ejercicio 3:

La tabla con los datos para el periodo especificado.

```
*          ts          | weather_conditions
-----+-----
2017-11-05 00:00:00 | Good
2017-11-05 01:00:00 | Bad
2017-11-05 02:00:00 | Good
2017-11-05 03:00:00 | Good
2017-11-05 04:00:00 | Bad
2017-11-05 05:00:00 | Bad
2017-11-05 06:00:00 | Good
2017-11-05 07:00:00 | Good
2017-11-05 08:00:00 | Good
2017-11-05 09:00:00 | Good
2017-11-05 10:00:00 | Good
2017-11-05 11:00:00 | Good
2017-11-05 12:00:00 | Good
2017-11-05 13:00:00 | Good
2017-11-05 14:00:00 | Bad
2017-11-05 15:00:00 | Good
2017-11-05 16:00:00 | Bad
2017-11-05 17:00:00 | Good
2017-11-05 18:00:00 | Bad
2017-11-05 19:00:00 | Bad
2017-11-05 20:00:00 | Bad
2017-11-05 21:00:00 | Good
2017-11-05 22:00:00 | Good
2017-11-05 23:00:00 | Good
2017-11-06 00:00:00 | Good
(25 rows)
```

La solicitud que utilizaste para resolver la tarea.

```
SELECT ts, CASE WHEN description LIKE '%rain%' OR description LIKE '%st
```

## Ejercicio 4

Tras actualizar el software, la compañía de taxis comienza a reportar que la ganancia que reciben no coincide con los datos que proporciona la aplicación. El equipo de desarrollo sugiere que el problema puede estar en los datos sobre el número de viajes.

Para determinar si hay un bug, debes obtener la selección del número de viajes de cada compañía de taxi para los días 15 y 16 de noviembre de 2017.

1. Devuelve el campo `company_name`. Nombra `trips_amount` (cantidad de viajes) al campo con el número de viajes y devuélvelo.
2. Organiza en orden descendente los resultados obtenidos en el campo `trips_amount`.

Pista: para resolver el problema, conecta las tablas de taxis y viajes.

Aplica funciones de agregación con agrupamiento. No olvides escribir la construcción con una condición.

Esto es lo que debes adjuntar en la respuesta:

1. La tabla resultante con los datos para el periodo especificado.
2. La solicitud que ayudó a resolver el problema.

### Resultados del ejercicio 4

La tabla con datos para un período específico.

company_name	trips_amount
Flash Cab	19502
Taxi Affiliation Services	11370
Medallion Leasin	10347
Yellow Cab	9861
Taxi Affiliation Service Yellow	9270
Chicago Carriage Cab Corp	9161
City Service	8424
Sun Taxi	7684
Star North Management LLC	7421
Blue Ribbon Taxi Association Inc.	5925
Choice Taxi Association	4998
Globe Taxi	4378
Dispatch Taxi Affiliation	3339

Nova Taxi Affiliation Llc		3169
Patriot Taxi DbA Peace Taxi Associat		2231
Checker Taxi Affiliation		2213
Blue Diamond		2067
Chicago Medallion Management		1948
24 Seven Taxi		1774
Chicago Medallion Leasing INC		1600
Checker Taxi		1482
American United		1398
Chicago Independents		1292
KOAM Taxi Association		1252
Chicago Taxicab		1010
Top Cab Affiliation		973
Gold Coast Taxi		426
Service Taxi Association		402
5 Star Taxi		310
303 Taxi		250
Setare Inc		229
American United Taxi Affiliation		210
Metro Jet Taxi A		146
Leonard Cab Co		145
Norshore Cab		127
6742 - 83735 Tasha ride inc		39
3591 - 63480 Chuks Cab		37
1469 - 64126 Omar Jada		35
6743 - 78771 Luhak Corp		33
0118 - 42111 Godfrey S.Awir		33
6574 - Babylon Express Inc.		31
Chicago Star Taxicab		29
1085 - 72312 N and W Cab Co		29
2809 - 95474 C & D Cab Co Inc.		29
2092 - 61288 Sbeih company		27
3011 - 66308 JBL Cab Inc.		25
3620 - 52292 David K. Cab Corp.		21
4615 - 83503 Tyrone Henderson		21
3623 - 72222 Arrington Enterprises		20

La solicitud que utilizaste para resolver la tarea.



```
SELECT company_name, COUNT (trip_id) AS trips_amount FROM cabs INNER JOIN
```