*Carga de la base de datos con PostgreSQL*

## **Instalación**

Para realizar la instalación de este sistema de gestión de bases de datos relacional orientado a objetos y libre, primero hemos procedido a descargar el paquete de instalación a través de la página de la compañía **EnterpriseDB**, que proporciona software y servicios basados ​​en la base de datos de código abiertoPostgreSQL:

<https://www.enterprisedb.com/downloads/postgres-postgresql-downloads>. En nuestro caso, hemos procedido a instalar la versión 9.6.6 para el sistema operativo Linux x86-64.

## **Conexión a la base de datos**

Para conectarnos a la base de datos PostgreSQL, hemos procedido a utilizar el programa terminal interactivo proporcionado por PostgreSQL denominado **psql**.

### **Base de datos con los datos originales**

Antes que nada, hemos creado una base de datos con los datos originales que nos proporciona la competición KDD Cup 2017. Para poder construirla, primero nos hemos conectado a la base de datos de esta forma:

*$ psql tfgdatosoriginales*

Para cargar los datos originales de la competición en la base de datos PostgreSQL, hemos desarrollado un script que realiza la carga de todas estas tablas de forma directa. Para ejecutarlo, en el terminal se escribe lo siguiente:

*$ \i [ruta\_del\_script]*

#### Tabla *road\_links (“links\_table3.csv”)*

El esquema de la tabla original proporcionado por la competición es la siguiente:

|  |  |  |
| --- | --- | --- |
| Campo | Tipo | Descripción |
| ***link\_id*** | string (char(3)) | Identificador del enlace |
| ***length*** | float | Longitud del enlace en metros |
| ***width*** | float | Anchura del enlace en metros |
| ***lanes*** | int | Número de carriles |
| ***in\_top*** | string (varchar(7)) | Este atributo contiene los enlaces entrantes al enlace actual, separados por comas |
| ***out\_top*** | string (varchar(7)) | Este atributo contiene los enlaces salientes del enlace actual, separados por comas |
| ***lane\_width*** | float | Anchura de cada uno de los carriles del enlace en metros |

#### Tabla *vehicle\_routes (“routes\_table4.csv”)*

## El esquema de la tabla original proporcionado por la competición es la siguiente:

|  |  |  |
| --- | --- | --- |
| Campo | Tipo | Descripción |
| ***intersection\_id*** | string (char(1)) | Identificador de la intersección |
| ***tollgate\_id*** | string (char(1)) | Identificador de la barrera de peaje |
| ***link\_seq*** | string (varchar(47)) | Secuencia de enlaces que conforman la ruta desde la intersección hasta la barrera de peaje |

La red de carreteras utilizada en la competición es un grafo dirigido formado por enlaces de carreteras interconectados. Una ruta en la red está representada por una secuencia de enlaces. Para cada enlace de la carretera, el tráfico de vehículos proviene de uno o más "enlaces viales entrantes" y entra en uno o más "enlaces viales salientes”.

#### Tabla *vehicle\_trajectories\_training (“trajectories\_table 5\_training.csv.csv”)*

El esquema de la tabla original proporcionado por la competición es la siguiente:

|  |  |  |
| --- | --- | --- |
| Campo | Tipo | Descripción |
| ***intersection\_id*** | string (char(1)) |  |
| ***tollgate\_id*** | string (char(1)) |  |
| ***vehicle\_id*** | string (varchar(30)) |  |
| ***starting\_time*** | datetime (timestamp) |  |
| ***travel\_seq*** | string (varchar(400)) |  |
| ***travel\_time*** | float |  |