

Apache Spark



KEEPCODING
Tech School

Antonio Gutiérrez López
gutierrez.lopez.ant@gmail.com



Spark GraphX



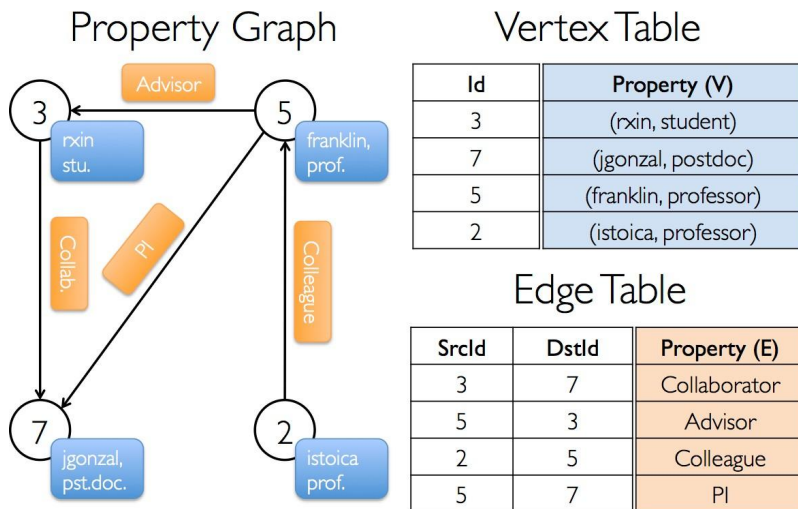
KEEPCODING
Tech School

Antonio Gutiérrez López
gutierrez.lopez.ant@gmail.com



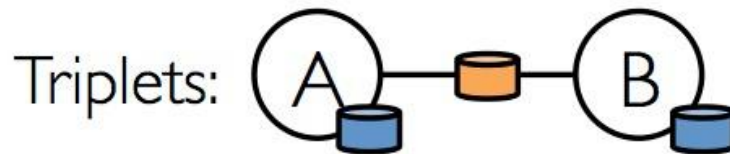
Spark GraphX

- GraphX es una librería en Spark para trabajar con gráficos de manera distribuida.
- Un grafo se compone de vertex y edge.



Spark GraphX

- Spark GraphX nos permite trabajar con:



- **Vertices:** Vertices existentes
- **Edges:** Relaciones existentes
- **Triplets:** Relación entre dos vértices y su edge.

Spark GraphX

- Spark GraphX nos proporciona una serie de algoritmos para trabajar:
 - **PageRank:** Mide la importancia de cada vertex en el grafo. Considerando el número de relaciones que tiene cada vértice, directas o indirectas
 - **Connected Components:** Obtiene los componentes conectados.
 - **Triangle Counting:** Obtiene los componentes que están conectados formando un triángulo.

| Exercise 1

School GraphX



KEEPCODING

Tech School

Madrid | Barcelona | Bogotá

Datos de contacto