**Discuss the differences between Hadoop, Relational Databases and NoSQL Databases (Hint: ACID vs. BASE)**

While the other two are actual databases, Hadoop is an ecosystem where files are distributed and stored, which allows parallel computing. There is no structure specified for the data in Hadoop environment. The system is very good with batch processing. File sizes are large and parallel computing available, therefore it is a great solution for large and distributed tasks. On the other hand, it is limiting speed since parallel processing requires a lot of relational links. In this sense, NoSQL DBs are great solutions while they provide speed, due to not having to join tables, along with big data option, however, they are not as consistent, data are not normalized and its integrity is in question. From another perspective, RDBMS are great solutions due to its ACID form, because they are highly consistent and also has a large ecosystem. Some downsides of these databases are that the data relations are subjected to constraints and structured forms are not easy (especially time-wise) to obtain.

**Why do you think the Relational Databases will persist in the near future?**

Simple because they are industry-wise extensive. They are old and have a great and established ecosystem. When it comes to online transactions and analytics, they provide the most reliable and consistent solution. RDBMs are very simple in structure, data retrieval is rather easy, data integrity is the best among the alternatives, and it works with normalized data. One other advantage of RDBMs are the security of the data. This management system provides superior data security comparing to its alternatives.

**Why is important that you know SQL as a data scientist or data engineer?**

Most of the industry is still dominated by the SQL systems. It is almost impossible for a data scientist or a data engineer not to come across to SQL based data. When the most necessary element of these job titles (i.e. data) has been provided through SQL based databases, it is a must to know SQL, so that job and tasks can be carried out. By its nature, this type of database will be around for some time since it has advantages from some aspects comparing to the alternatives. There will be companies needing SQL based solutions and, most enterprise data remain a great fit for an RDBMS (89% market share - Gartner)1.

1  Presented by James Serra, P. Relational Databases vs Non-Relational Databases vs Hadoop. Presented at 24 Hours of Pass.