CAP Theorem:

- **Consistency**: Every read from the database gets the latest (and correct) piece of data or an error
- **Availability**: Every request is received and a response is given -- without a guarantee that the data is the latest update
- **Partition Tolerance**: The system continues to work regardless of losing network connectivity between nodes

Additional Resource:

You can also check out this Wikipedia page on CAP theorem

Commonly Asked Questions:

Is Eventual Consistency the opposite of what is promised by SQL database per ACID principle?

Much has been written about how *Consistency* is interpreted in the ACID principle and the CAP theorem. Consistency in the ACID principle refers to the requirement that only transactions that abide by constraints and database rules are written into the database, otherwise the database keeps previous state. In other words, the data should be correct across all rows and tables. However, consistency in the CAP theorem refers to every read from the database getting the latest piece of data or an error.

To learn more, here are some links you may find useful:

- Discussion about ACID vs. CAP
- Wikipedia

Which of this combination is desirable Consistency and Availability, Consistency and Partition, or Availability and Partition for a production system?

There is no such thing as Consistency and Partition, you can only have Consistency and Availability or Availability and Partition. Remember, relational and non-relational databases do different things, and that's why most companies have both types of database systems.

Does Cassandra meet just Availability and Partition in the CAP theorem?

According to CAP theorem, a database can actually only guarantee two out of the three in CAP. So supporting Availability and Partition makes sense, since Availability and Partition are the biggest requirements.

If Apache Cassandra is not built for consistency, won't the analytics pipeline break?

If I am trying to do analysis, such as getting a trend overtime, e.g., how many friends does John have on Twitter, and if you have one less person counted because of "eventual consistency" (the data may not be up-to-date in all locations), that's OK. In theory, that can be an issue but only if you are not constantly updating. If the pipeline pulls data from one node and it has not been updated, then you won't get it. Remember, it is about Apache Cassandra about **Eventual Consistency**.