

The problem that I am solving is helping the New England Patriots determine statistical insights from a large dataset including all 2014 NFL plays. Typical NFL stats include a quarterback's completion rate and total yards, but there are more patterns in the data that could help the Patriots and are only accessible to a data scientist. I imagine that I am hired as a consultant by the Patriots to work with their NFL plays dataset and determine values such as Tom Brady's completion rate to each of his top receivers. I need a database system that is fast, can manage a large amount of data (14+ MB), and provides the capability to do complex queries.

The technology that I utilize to solve my problem and also explain more about in this report is VoltDB. VoltDB database is a new system that runs a database from memory rather than disk. It processes more quickly than disk-based systems and is the first database designed to operate solely on memory. It gaining popularity and customers include well-known enterprises such as HP and Nokia. I downloaded the enterprise trial version onto my Mac and immediately noticed the advantages of memory-based data. Each query or process includes a stop-watch illustrating how quickly the database functions such as loading the NFL dataset in only a few seconds. VoltDB is also user-friendly and designed for MySQL-type users to convert to VoltDB. It features an SQL executable file sqlcmd to access the database with .sql files and similarly has a local-host console and Java API as well. My project begins with directing a .sql file to the sqlcmd executable and then does the data processing through a Java program and Java API.

Overall, my report shows my solution from installation to graphically displaying the results in R.

There are many benefits and advantages to using VoltDB and if you are interested in new generation databases, I recommend downloading VoltDB and trying the steps I outline.

VoltDB is a cutting edge database, has recently acquired \$23+ million in funding, and

Ryan Ballenger
CSCIE63 Final Project Summary
Topic: VoltDB

represents the evolution happening in this field and the high quality products available to all users.