Your boss enjoyed hearing how you characterized the differences between online transactional processing and online analytical processing systems. He then asked you to research how ETL moves data from online transactional processing systems to online analytical processing systems this week.

You should return and report with a 3–5 paragraph report that explains how ETL processes work and whether or not they consistently rely on a single programming language or method.

* Database Repository Types
  + [Online Transaction Processing (OLTP) Systems (Links to an external site.)](https://en.wikipedia.org/wiki/Online_transaction_processing)
  + [Online Analytical Processing (OLAP) Systems (Links to an external site.)](https://en.wikipedia.org/wiki/Online_analytical_processing)
* Extract, Transform, and Load (ETL) Process:
  + [ETL Defined (Links to an external site.)](https://en.wikipedia.org/wiki/Extract,_transform,_load)
* Data Sources:
  + [Relationship Databases (Links to an external site.)](https://en.wikipedia.org/wiki/Relational_database)
  + [XML Files (Links to an external site.)](https://en.wikipedia.org/wiki/XML)
  + [JSON Files (Links to an external site.)](https://en.wikipedia.org/wiki/JSON)
  + [Flat Files (Links to an external site.)](https://en.wikipedia.org/wiki/Flat-file_database)
* Environments:
  + [Shell Environments](https://en.wikipedia.org/wiki/Shell_(computing))
  + [Unix Shell (Links to an external site.)](https://en.wikipedia.org/wiki/Unix_shell)
  + [Bash Shell (Links to an external site.)](https://en.wikipedia.org/wiki/Bash_(Unix_shell))
  + [PowerShell (Links to an external site.)](https://en.wikipedia.org/wiki/PowerShell)
  + [AWS CLI (Command Line Interface) (Links to an external site.)](https://docs.aws.amazon.com/cli/latest/reference/)
* Programming Languages:
  + [Python (Links to an external site.)](https://en.wikipedia.org/wiki/Python_(programming_language))
  + [Java (Links to an external site.)](https://en.wikipedia.org/wiki/Java_(programming_language))
  + [JavaScript (Links to an external site.)](https://en.wikipedia.org/wiki/JavaScript)
  + [XSLT (Links to an external site.)](https://en.wikipedia.org/wiki/XSLT)
  + [XQuery](https://en.wikipedia.org/wiki/XQuery)

One of the challenges that we often face when analyzing data is that our data is stored in multiple databases and other data sources. The ETL (Extract, Transform, Load) process helps us overcome that challenge because it “provides a method of moving the data from various sources into a data warehouse, as data sources change the Data Warehouse will automatically update”. (1) Not only does it bring that data into one place, it also keeps that data up to date.

For example, if you were a company that has data that is always needing to change and update, then an OLTP system would work best to maintain your data. But when you want to analyze that data, instead of expecting your users to have the skills to join all those tables together to ask their questions in your OLTP database, the better option would be to use ETL to transform the data and put it into an OLAP database. OLAP makes it much easier for users to analyze data.

ETL has three parts to it. It first extracts the data from various data sources, including files and databases. It then transforms the data. This includes cleaning the data, removing duplicates, changing formats, filtering out certain data, aggregating the data, and more. Finally, it loads the data into the database warehouse of your choosing.

The ETL process can be complicated depending on the data sources, environments, and programming languages that are used. There are many ways to go about extracting, transforming, and loading the data. While a lot of data are in complicated relationship databases, some data may also be in flat files (like Excel or csv files). The benefit of ETL is that it can convert data stored in these different ways and then put the data together into one data warehouse for whatever purposes you need.

ETL is not an actual software that you use. It is a process that can work with many different types of environments, programming languages, and software. Think of environments as an area to run code for different types of languages as long as the programming language is set up and installed on your machine. Some of the types of the environments you can use are Unix Shell, Bash Shell, Powershell, and more. Some of the different programming languages that you can use for ETL are Python, SQL, Java, and more. Each type of programming language and environment has its benefits and drawbacks.

* + 1. https://www.guru99.com/etl-extract-load-process.html