1. Data that does not necessarily comply with normal relational database constraints (like strict datatypes) but still contains some markers to identify certain segments. (Like XML with tags or JSON with key-value pairs)
2. You can flatter it using SQL or other tools like pythons. Keys may be extracted and transformed into columns, so keys become the names of columns and arrays the records
3. It is a Javascript Object notation which basically a python dictionary. It is widespread since it is convenient to read and it is transferable between back and front end easily(the reason why it was developed)
4. XML(extensible markup language) - all info coded using tags. Every value needs close and open tag, this thing is horrible to even look at. Both JSON and XML used for semi structured data, JSON just prettier to llok, functionally they almost the same
5. Open storage file format. It is used for analytics since it column - oriented, thus, the aggregation becomes much faster, the disk pointer does not need to jump.
6. Row - oriented store together records of data, and the column ones store together on disk columns. So if u need do operations with rows more frequently (Transactions for example) pick row oriented, and for analytics - column.
7. Because JSON allows to represent hierarchical structure (inside nested objects) CSV unable to do this. Moreover, on back side it is frequently the case that Java Screpт is used so JSON is more convenient
8. It still kinda loads it, but only temporary, so the query runs, and then engine uploads the data into memory but then it gets evaporated. The engine treads them as virtual table, it doesnt upload the data per se, just metadata