# Abstract

Neo4j is becoming the standard when graph-based data is there. Neo4j emerged to solve the performance problems in the RDBMS (Relational Database Management System) during the early 2000 and with in the first 3 years of development, Team of neo4j had deployed the first production version for the graph-based data. Since 2007 They made neo4j graph database opensource, they received immense response from the graph database adopters. Since then Neo4j is catering various companies and developed various tools around their different business problems. Currently they have 2 categories of product: Graph Database and Graph Data Science. Both of that product has self-managed as well as fully managed products. Where Neo4j Graph Database is self-managed and opensource. Which is used to store the graph-based data. In recent times demand for Business intelligence had also been increased. Where Business Intelligence tools like Tableau and Power BI are used to have insights of the data. However those tools does not able to read data from the graph database like Neo4j directly. These tools required mediatory tools which reads data from Neo4j and provide to BI tools. These are called “BI Connectors”. And the objective of this project is to measure the performance of those connectors as well as to see the difference in the visualization in different BI tools. During the testing, I found that here are some connectors which are easy to setup as well as in Tableau it is easier to read data as compared to the power BI.

# Introduction

Observing over some advancement in the past few decades, data and its related domains are becoming more popular and extremely useful. Along with the advancement of the technology, the applications of the data related field domains had also drastically increased. Alongside, advancement in domains like artificial intelligence and machine learning had exponentially rose the importance of the data.

It can be better understood as considering the situation where personalized content on any website which is only possible by utilizing the user’s watch history and interests. It can also be seen on may social medial platforms like Instagram, Facebook, any many more. It is not only limited to social media platform but e-commerce rise and website-based publicity and advertising had utilized extensively use of the user data.

As the data became important its analysis part had also become crucial. Due to the large available user for any kind of application or specific task, the data collect from it had also been significantly rose. In many applications also currently tabular form of the data is still used but technologies like SQL and database management system had provided immense support to the large data which had ultimately rose Big data domain.

Alongside, visualization of the data is as important as its analysis and collection. For extracting meaningful outcome from the collected data and derive any conclusion involves the visualization of the data. Many tools and software are available for the visualization of the data, but when the size of the data rises they becomes ineffective as very high-end hardware and computing power is needed to handle such big data.

As the advancement of technologies are continuously in an exponential way, understanding the relation between the data and their parameter had also been possible by past few decades, which had led us to introduce us with the graph database which are excellent in providing the graphical representation of the big data as well as relation between the elements of the data.

Technology had been exponentially developing and, in each step, better and quicker performance had been the primary goal. Various technology had been implemented across various applications which are serving their best potential as well as they are still under research. Developing any technology not only provides a better and more efficient way to deal with the dedicated task. Another side of the technological development aspect is the boom in digital technology as well as its implementation in diverse fields of applications. Internet connectivity had not only provided a vast closed network of connectivity over which various web-based applications are deployed and functioning to provide reliable and excellent support to millions of users over it. Observing the broader view of the functionality of the internet and its working, the large number of exchanges of data can be seen. For any activity enclosed into the internet or any kind of digital process data had been generated as well as its related information is also stored in their respective associated files. Over time large amounts of exchange of data over the internet as well as digital processes had been conducted which had built up massive data.

Development of the data science and machine learning provided an excellent scope for the utilization of the data that had a large number of relations with multiple entries. Generally, this type of data created that correlated with much other data is referred to as relational data and with the use of various available tools crucial information can be extracted from it. As the extracted information plays a very vital role in various applications to the company as well as research it becomes very important to extract the correct and authentic piece of information from it so that it can create a positive result when applied for future evaluations. As the amount of data increases the relations between entries also increase as well as their parameters. The primitive method which uses a query-based processing method or any visualization tool becomes ineffective. The solution to this obstacle graph-based database visualization especially for relational databases are used and implemented since 2007. Currently, among all the various tools for graph databases, Neo4j is leading which may include speed, effectiveness, and better performance. Keeping open-source as the development part of the software and tool, growth of the neo4j had been observed exponential with fluence operation in any size of data. Currently, neo4j had been largely deployed in many complex and crucial operations of various platforms across the globe. The majority of the users can be observed for business and scientific purposes. The other side can be of the neo4j can be seen as the large number of computations that shall be required by such tools which makes it bulky in its operation. For the usage of the outcomes, other tools like tableau, Power BI, etc can be used. Various types of connections are available which can integrate the connection between the neoi4j and power bi or tableau which are very efficient as well as effective for the visualization purpose. Commonly called BI connectors, this bridge between neo4j and the visualization tool provides an overall effective operation for managing graph databases. Measuring the performance of various available connectors which are compatible as well as effective in nature concerning our selected graphical database was briefly showcased in this project.

# Literature Review

[A Review of Graph Databases | SpringerLink](https://link.springer.com/chapter/10.1007/978-3-031-20322-0_13)

As graph-based database becomes more popular the tools and applications handling them had become more popular. The effectiveness of the graph database provided an immense processing speed as well as excellent visualization which resulted in maximum utilization of it across various applications and processing tasks. The research community had also been used larger this toll for the handling of graph-based databases or relational databases. Big data had been also exponentially increased the usage of the graph database and its tools. Over 15 years many types of graph databases had been proposed and researched. The author provided a broader view of the timeframe and developments in the graph databases along with their architecture. Categorization of the graph databases based upon many criteria and research had been also showcased. The survey and its various outcomes with results are also showcased by the author in the last part which shall provide a piece of deep information and developments about the various aspects of the graph database. Lastly, a guideline had been proposed by the author from the survey and their outcome which shall be used for the selection of one of the best graph databases among all based upon the provided criteria.

[Using Neo4j database and graph model for analysis of metropolitan railway connections of Silesian Voivodeship in Poland - TASK Quarterly : scientific bulletin of Academic Computer Centre in Gdansk - Tom Vol. 25, No 4 (2021) - Biblioteka Nauki - Yadda (icm.edu.pl)](https://yadda.icm.edu.pl/yadda/element/bwmeta1.element.baztech-ffbd32d4-66ba-4080-a1fa-1d661a4aa317)

With the immense usage of graph database tools and the effective outcome from their visualizations, neo4j had become a popular tool for this purpose. In this research, the author had been showcased the standalone usage of neo4j for the handling of the graph database. The analysis had been conducted based on the railway infrastructure data in Poland. Laboratory environment on the bases of Poland railway infrastructure had been conducted with the relational database and the tool used for computing is neo4j. visualization of the data and obtaining various outcomes have been showcased by the auth briefly in this research which also showcased the capability of the usage of neo4j and its limitations. Various types of design models have been also simulated and measured indications had been used for the outcome indications. Basically, with this research author has provided a detailed overview of the neo4j tool, and its compatibility with the graph database. Using the railway database and demonstrating the simulation environment, the limitation of the graph database and neo4j tools had been also showcased and provided different topologies of the graph and its usage in many research along with its scope have been showcased.

[Supporting Data Types in Neo4j | SpringerLink](https://link.springer.com/chapter/10.1007/978-3-031-15743-1_42)

Over the years many types of databases had been developed in which the latest type shall be graph databases. Relation between the entries and large structure had provided a scope to compute based upon the relation in which neo4j-like tools are very effective as well as fast. This research by auth showcased the capabilities of the neo4j and the types of data supported in them. Support of the data in the tools becomes essential when it becomes opensource as well as its applications were exponentially increasing. Parameters like nodes, data entries, data types, etc had been briefly showcased by the author which shall provide an overview of neo4j tools. Alongside author tried to showcase how assigning data types to each property can be implemented. Checking and extraction of a particular data correctly in neo4j had been also showcased. Change of syntax, creation of nodes, triggers, and many more such concepts associated with neo4j and its usage with the databases are highlighted and demonstrated in this research by the author.

[Towards a framework for tensor ontologies over Neo4j: Representations and operations | IEEE s abd eary Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/8316441/)

This research had been towards an ontology and its various aspects of analysis with the neo4j tools as supporting the application. As ontology is one of the most active research projects filed it shall consist of various domains like history, mathematics, philosophy, etc along with computer science. Entries and information in it shall be also considered as data which had a piece of relevant information and can be processed by the computer for the extraction of the hidden information. From the latest research technique including horn logic knowledge is increased had been observed especially with reasoners and RDF triplets. Moving into graph mining, tools like neo4j, XML parsers, etc had been efficiently processed as they are based upon semantic level instead of combinational level. Multi-layer graphs and namely graphs with many other similar are also becoming more complex as well in length as the data becomes larger. labeled edges, predetermined classes, etc had been recently introduced which shall provide great efficiency to computing which relation to the entries increases. Such complex data had been explained through social network analysis and its details were also showcased along with its descriptions and usage. Many databases which are readable to humans which shall be from the late 1970s and early 1980s apple were manually constructed in 2011 and their usage had been also done in many applications my apple. The usage of neo4j and it's in detail information about the capabilities of ontology use case had been showcased by the author in this research.

[Analysis of film data based on Neo4j | IEEE Conference Publication | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/7960078/)

With the great ability to handle graph databases and relational databases by neo4j, the author in this research provided a demonstration of fil databases and their analysis in the neo4j tool. As the film data contains various entries and information which shall be co-related with the other entries in that same database large number of relation trees can be observed. Ne4j uses graph-based relation mapping which can effectively showcase the various relationships among data and showcased in a visualized and user-friendly way making it one of the most ideal choices by the author. On-film databased neo4j operations were showcased and described by the author in this research paper. Starting with the cypher query language with neo4j is introduced. Followed by an analysis of association key objects in films with complex and multi-connection data in neo4j with the usage of its web interface. Storage of the data and its visualization along with its results and outcome had been lastly conducted. The author showcased a convenient way for the analysis of the film and its data in the neo4j which is one of the most suitable methods and appropriate ways to analysis the film database from which much information can be extracted for any research or usage.