

KSE30 Companies Network Analysis

CS-343: Graph Data Science

Habib University

By: Abdullah Junejo (07154) Alizain Merchant (07565)

Progress Report

-

Date: May 5, 2024

1 Introduction

In this project, we explore the intricate and ever-evolving landscape of the Karachi Stock Exchange 30 (KSE30) Index, which represents the top 30 corporations in Pakistan by market capitalization and liquidity. Our goal is to employ graph database technology to dissect and elucidate the complex web of relationships and interdependencies among these premier entities. Utilizing graph databases enables us to represent and probe these relationships with unmatched intuitiveness and efficiency. Specifically, by mapping the KSE30 companies and their principal executives and board members as nodes within a graph database, we seek to shed light on the influence of these individuals within the market, explore their network connections, and analyze overall market trends. For this purpose, we will utilize Neo4j, gathering our data directly from the Pakistan Stock Exchange through the PSX Data Portal (<https://dps.psx.com.pk/>).

2 Data Collection

The Data was manually collected by going into the PSX Data Portal and then KSE30 where we saw different market performers. Upon clicking the company and noting down its information one by one, we were able to put together a .csv file for this (Attached). This task can be furthered on KSE100 as well after successful demonstration of this miniature project.

3 Data Modelling/Import

Here is our initial Graph Model that we will use. However, this may very well be improvised upon in the future!

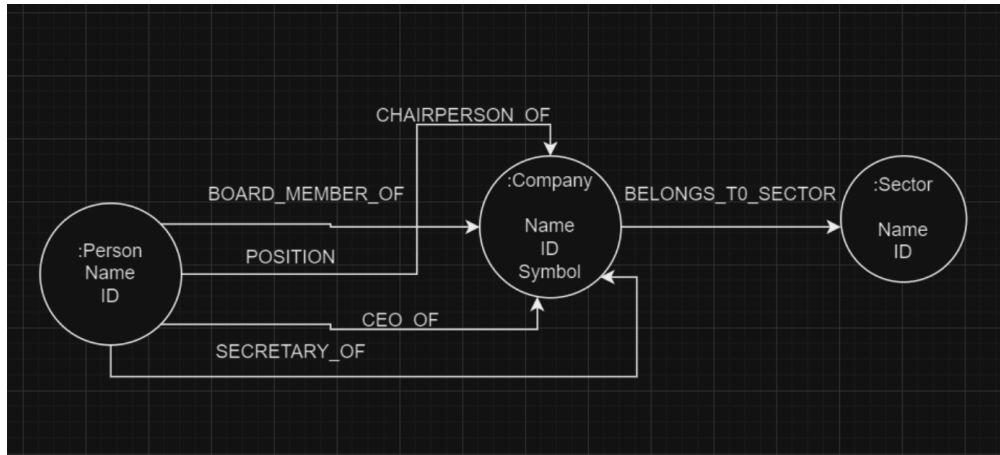


Figure 1: Initial Model

After Creating Project, Installing Graph Data Science/APOC Libraries, dropping the file in import folder, we worked on this script and used it to import data our csv file according to our aforementioned model.

```
1  LOAD CSV WITH HEADERS FROM 'file:///KSE_30_Formatted.csv'
   AS row
2  WITH row
3  WHERE row.Board of Directors IS NOT NULL AND TRIM(row.
   Board of Directors) <> ""
4  UNWIND SPLIT(row.Board of Directors, ";") AS
   directorDetails
5  WITH row, TRIM(directorDetails) AS directorDetails
6  WHERE directorDetails <> ""
7  WITH row,
8       TRIM(SPLIT(directorDetails, "(")[0]) AS DirName,
9       TRIM(SPLIT(SPLIT(directorDetails, "(")[1], ")")[0])
   AS position
10 WHERE row.Year IS NOT NULL AND TRIM(row.Year) <> ""
11 AND row.Company Sector IS NOT NULL AND TRIM(row.
   Company Sector) <> ""
12 AND row.Secretary IS NOT NULL AND TRIM(row.Secretary)
   <> ""
```

```

13      AND row.Chairperson IS NOT NULL AND TRIM(row.
14      Chairperson) <> ""
15      AND row.Company Symbol IS NOT NULL AND TRIM(row.
16      Company Symbol) <> ""
17      AND row.Company Name IS NOT NULL AND TRIM(row.Company
18      Name) <> ""
19      AND row.Company CEO IS NOT NULL AND TRIM(row.Company
20      CEO) <> ""
21      AND position IS NOT NULL AND TRIM(position) <> ""
22
23  MERGE (company:Company {name: row.Company Name, symbol:
24      row.Company Symbol})
25  MERGE (sector:Sector {name: row.Company Sector})
26  MERGE (chairperson:Person {name: row.Chairperson})
27  MERGE (secretary:Person {name: row.Secretary})
28  MERGE (ceo:Person {name: row.Company CEO})
29
30  WITH company, sector, chairperson, secretary, ceo,
31      DirName, position
32  MERGE (director:Person {name: DirName})
33  MERGE (director)-[:BOARD_MEMBER_OF {position: position
34      }]->(company)
35  MERGE (company)-[:CHAIRPERSON_OF]->(chairperson)
36  MERGE (company)-[:SECRETARY_OF]->(secretary)
37  MERGE (company)-[:CEO_OF]->(ceo)
38  MERGE (company)-[:BELONGS_TO_SECTOR]->(sector);

```

After this, it became aparent that there were multiple edges of "Board Member" from the same Person Node to the Same Company Node but having different edge property called position. So we decided to merge them so that future queries are easier and performance is not compromised. The property of position was also in a way merged. Different positions were seperated by a backslash.

```

1  //Step 1: Match and collect positions for each director-
2  company pair
3  MATCH (director:Person)-[r:BOARD_MEMBER_OF]->(company:
4  Company)
5  WITH director, company, COLLECT(DISTINCT r.position) AS
6  positions, COLLECT(r) AS rels
7
8  //Step 2: Delete the existing relationships
9  FOREACH (rel in rels | DELETE rel)
10
11 //Step 3: Create a new single relationship with all

```

```

    positions combined
9  WITH director, company, REDUCE(s = "", position IN positions
    | s + (CASE WHEN s = "" THEN "" ELSE "\\\" END) +
    position) AS combinedPositions
10 MERGE (director)-[newRel:BOARD_MEMBER_OF]->(company)
11 SET newRel.positions = combinedPositions

```

This essentially merged Multiple Board of Director edges into one if they are related to the same company.

4 Graph Statistics

This initial statistic provides a snapshot of the various sectors represented in the KSE30 index, the total number of companies, and detailed counts of key executive roles across all companies. This fundamental data lays the groundwork for deeper analyses, offering a clear view of the breadth and scope of leadership within these influential companies.

```

1  //Statistics 1:
2  MATCH (sector:Sector)
3  WITH COLLECT(sector.name) AS Sector Names
4
5  MATCH (company:Company)
6  WITH Sector Names, COUNT(company) AS Total Number of
    Companies
7
8  MATCH (director:Person)-[:BOARD_MEMBER_OF]->(company)
9  WITH Sector Names, Total Number of Companies, COUNT(DISTINCT
    director) AS Total Number of Board Members
10
11 MATCH (chairperson:Person)<-[:CHAIRPERSON_OF]-(company)
12 WITH Sector Names, Total Number of Companies, Total Number
    of Board Members, COUNT(DISTINCT chairperson) AS Total
    Number of Chairpersons
13
14 MATCH (ceo:Person)<-[:CEO_OF]-(company)
15 WITH Sector Names, Total Number of Companies, Total Number
    of Board Members, Total Number of Chairpersons, COUNT(
    DISTINCT ceo) AS Total Number of CEOs
16
17 MATCH (secretary:Person)<-[:SECRETARY_OF]-(company)
18 RETURN Sector Names, Total Number of Companies, Total Number
    of Board Members, Total Number of Chairpersons, Total

```

Number of CEOs , COUNT(DISTINCT secretary) AS Total Number of Secretaries

This second analysis reveals directors who hold positions across multiple companies (Highest 2 as of now), which could give a general idea about their influence across corporations and potential conflicts of interest. Next query is about tracking chairpersons who currently oversee multiple companies (\exists 1 such chairperson as of now) further sheds light on the leadership dynamics. The third query tries to find the company that sits at the center of the most extensive connections, either through direct executive roles such as secretary, chairperson, CEO or through other types of affiliative relationships such as Board Member. Such a company typically acts as a central node within the network which we hope to unravel and confirm in our future explorations. Next we try to Highlight the sector that includes the most companies within the KSE30 offers insights into which part of the economy is most saturated. This information is crucial for steering investment strategies, creating economic policies (regarding subsidy/taxation), and conducting market analysis, as it reflects the sector's economic significance and its potential for expansion or risk.

```
1 //Statistics 2:
2 // Collect data for board members belonging to multiple
  companies
3 MATCH (director:Person)-[:BOARD_MEMBER_OF]->(company:
  Company)
4 WITH director, COUNT(DISTINCT company) AS Number of
  Companies
5 WHERE Number of Companies > 1
6 WITH COLLECT(director.name) AS Multi Company Board
  Members
7
8 // Collect data for chairpersons of multiple companies
9 MATCH (chairperson:Person)<-[:CHAIRPERSON_OF]-(company:
  Company)
10 WITH chairperson, COUNT(DISTINCT company) AS Number of
  Companies, Multi Company Board Members
11 WHERE Number of Companies > 1
12 WITH COLLECT(chairperson.name) AS Multi Company
  Chairpersons, Multi Company Board Members
13
14 // Identify the company with the most affiliations
15 MATCH (person:Person)-[r]->(company:Company)
16 WITH company, COUNT(DISTINCT person) AS Number of
  Affiliates, Multi Company Board Members, Multi
```

```

17         Company Chairpersons
18 ORDER BY Number of Affiliates DESC
19 WITH COLLECT({companyName: company.name, numAffiliates:
20   Number of Affiliates})[0] AS Company With Most
21   Affiliates, Multi Company Board Members, Multi
22   Company Chairpersons
23
24 // Sector with the most companies
25 MATCH (company:Company)-[:BELONGS_TO_SECTOR]->(sector:
26   Sector)
27 WITH sector, COUNT(DISTINCT company) AS Number of
28   Companies, Company With Most Affiliates, Multi
29   Company Board Members, Multi Company Chairpersons
30 RETURN sector.name AS Sector with Most Companies, Number
31   of Companies,
32   Company With Most Affiliates.companyName AS
33   Company with Most Affiliations,
34   Company With Most Affiliates.numAffiliates AS
35   Number of Affiliated Persons,
36   Multi Company Board Members,
37   Multi Company Chairpersons
38 ORDER BY Number of Companies DESC
39 LIMIT 1

```

These statistics, derived from targeted Cypher queries, do more than just quantify the interconnections within the KSE30. They also provide a deeper qualitative understanding of corporate governance and sectoral prominence, laying a robust foundation for further analysis such as predictive modeling and trend examination. This enriched insight is instrumental in comprehensively understanding the dynamics that shape Pakistan's economic forefront.

5 Creating Further Data

Our data was mostly homogenous clusters of companies and their executives. Only few Board Members were part of 2 companies and only 1 chairperson. This shows that for our future analytics, it could be difficult to find a person to person shortest path. This is why we introduce government as stakeholders. It is a natural and well founded assumption that companies belonging to a sector will be regulated by the relevant government agencies. The company executives will have the meetings with such ministries and would be able to connect with their head, The minister. This could provide us a path to other clusters that was otherwise isolated. The Data was manually collected from Wikipedia and verified from various other official

government web sources.

```
1  //Creation of Ministries
2  CREATE (energy:Ministry {name: "Ministry of Energy"}),(
3      sbp:Ministry {name: "State Bank of Pakistan"}),(
4      itTelecom:Ministry {name: "Ministry of Information
5      Technology and Telecommunication"}),(scienceTech:
6      Ministry {name: "Ministry of Science and Technology"
7      }),(foodSecurity:Ministry {name: "Ministry of
8      National Food Security and Research"}),
9      (industries:Ministry {name: "Ministry of Industries and
10     Production"}),
11     (commerceTextile:Ministry {name: "Ministry of Commerce
12     and Textile Industry"}),
13     (health:Ministry {name: "Ministry of National Health
14     Services, Regulations and Coordination"});
15
16 //Link Ministries to Sectors
17 MATCH (ministry:Ministry), (sector:Sector)
18 WHERE (ministry.name = "Ministry of Energy" AND sector.
19     name IN ["Refinery", "Power Generation & Distribution
20     ", "Oil & Gas Marketing Companies", "Oil & Gas
21     Exploration Companies"]) OR
22     (ministry.name = "Ministry of Information
23     Technology and Telecommunication" AND sector.
24     name = "Technology & Communication") OR
25     (ministry.name = "Ministry of Science and
26     Technology" AND sector.name = "Chemical") OR
27     (ministry.name = "Ministry of National Food
28     Security and Research" AND sector.name = "Food
29     & Personal Care Products") OR
30     (ministry.name = "State Bank of Pakistan" AND
31     sector.name = "Commercial Banks") OR
32     (ministry.name = "Ministry of Industries and
33     Production" AND sector.name IN ["Cement", "
34     Cable & Electrical Goods"]) OR
35     (ministry.name = "Ministry of Commerce and Textile
36     Industry" AND sector.name = "Textile Spinning"
37     ) OR
38     (ministry.name = "Ministry of National Health
39     Services, Regulations and Coordination" AND
40     sector.name = "Pharmaceuticals")
41 MERGE (sector)-[:REGULATED_BY]->(ministry);
```

```

19 //Creation of Minister Nodes and Linking to Ministries
    with Positions
20
21 //Minister of Energy
22 CREATE (musadikMalik:Person {name: "Musadik Masood Malik
    ", position: "Minister"})
23 WITH musadikMalik
24 MATCH (ministryEnergy:Ministry {name: "Ministry of
    Energy"})
25 MERGE (musadikMalik)-[:HEADS]->(ministryEnergy)
26
27 //Minister of Information Technology and
    Telecommunication
28 CREATE (shazaFatima:Person {name: "Shaza Fatima Khawaja"
    , position: "Minister"})
29 WITH shazaFatima
30 MATCH (ministryIT:Ministry {name: "Ministry of
    Information Technology and Telecommunication"})
31 MERGE (shazaFatima)-[:HEADS]->(ministryIT)
32
33 //Minister of Science and Technology
34 CREATE (khalidSiddiqui:Person {name: "Khalid Maqbool
    Siddiqui", position: "Minister"})
35 WITH khalidSiddiqui
36 MATCH (ministryScience:Ministry {name: "Ministry of
    Science and Technology"})
37 MERGE (khalidSiddiqui)-[:HEADS]->(ministryScience)
38
39 //Minister of National Food Security and Research
40 CREATE (ranaTanveer:Person {name: "Rana Tanveer Hussain"
    , position: "Minister"})
41 WITH ranaTanveer
42 MATCH (ministryFoodSecurity:Ministry {name: "Ministry of
    National Food Security and Research"})
43 MERGE (ranaTanveer)-[:HEADS]->(ministryFoodSecurity)
44
45 //Minister of Industries and Production
46 CREATE (ranaTanveerDup:Person {name: "Rana Tanveer
    Hussain", position: "Minister"}) // Duplicate node
    for same person different role
47 WITH ranaTanveerDup
48 MATCH (ministryIndustries:Ministry {name: "Ministry of
    Industries and Production"})

```



```

49  MERGE (ranaTanveerDup)-[:HEADS]->(ministryIndustries)
50
51  //Minister of Commerce and Textile Industry
52  CREATE (jamKamal:Person {name: "Jam Kamal Khan",
53      position: "Minister"})
54  WITH jamKamal
55  MATCH (ministryCommerce:Ministry {name: "Ministry of
56      Commerce and Textile Industry"})
57  MERGE (jamKamal)-[:HEADS]->(ministryCommerce)
58
59  //Health Secretary overseeing Ministry of National
60  Health Services, Regulations and Coordination
61  CREATE (iftikharShallwani:Person {name: "Iftikhar Ali
62      Shallwani", position: "Health Secretary"})
63  WITH iftikharShallwani
64  MATCH (ministryHealth:Ministry {name: "Ministry of
65      National Health Services, Regulations and
66      Coordination"})
67  MERGE (iftikharShallwani)-[:HEADS]->(ministryHealth)
68
69  //Resolving Transport Sector Issue
70
71  MATCH(sector:Sector {name: "Transport"})
72  MERGE(aviationMinistry:Ministry {name: "Ministry of
73      Aviation"})
74  MERGE(maritimeMinistry:Ministry {name: "Ministry of
75      Maritime Affairs"})
76  MERGE(p1:Person{name: "Khawaja Muhammed Asif", position:
77      "Minister"})
78  MERGE(p2:Person{name: "Qaiser Ahmed Sheikh", position: "
79      Minister"})
80  // Establish potentially applicable regulatory
81  relationships
82  MERGE (sector)-[:REGULATED_BY]-(aviationMinistry)-[:
83      HEADS]-(p1)
84  MERGE (sector)-[:REGULATED_BY]-(maritimeMinistry)-[:
85      HEADS]-(p2)
86  return *

```

Now this all long script that is pretty inefficient and has plenty of corrections made to it afterwards is all about creating ministries that regulate sectors. A ministry can regulate many sectors, and a Minister can head multiple ministries (Max 2 in our database). In the case of Ministry of Health which is vacant at

the moment the position is secretary (which for the sake of argument lets assume also works with the same minister in order to make our database relevant). The only problem remains is the Sector: Commercial Banks which is regulated by The State Bank of Pakistan which is headed by Board of Directors. One such director is Finance Secretary who reports to Finance Minister heading the finance ministry. Now we can resolve this issue with the following query:

```

1      //Dealing with SBP
2  MATCH (sbp:Ministry {name: "State Bank of Pakistan"})
3
4  //Creating board members and Secretary, and their link to
   SBP
5  MERGE (jameelAhmad:Person {name: "Mr. Jameel Ahmad",
   position: "Chairperson Governor"})-[:BOARD_MEMBER_OF_SBP
   ]-(sbp)
6  MERGE (aliCheema:Person {name: "Dr. Ali Cheema", position: "
   Non-Executive Director"})-[:BOARD_MEMBER_OF_SBP]-(sbp)
7  MERGE (akbarZaidi:Person {name: "Dr. Syed Akbar Zaidi",
   position: "Non-Executive Director"})-[:
   BOARD_MEMBER_OF_SBP]-(sbp)
8  MERGE (najafKhan:Person {name: "Mr. Najaf Yawar Khan",
   position: "Non-Executive Director"})-[:
   BOARD_MEMBER_OF_SBP]-(sbp)
9  MERGE (fawadAnwar:Person {name: "Mr. Fawad Anwar", position:
   "Non-Executive Director"})-[:BOARD_MEMBER_OF_SBP]-(sbp)
10 MERGE (zahidEbrahim:Person {name: "Mr. Zahid Ebrahim",
   position: "Non-Executive Director"})-[:
   BOARD_MEMBER_OF_SBP]-(sbp)
11 MERGE (mahfoozKhan:Person {name: "Mr. Mahfooz Ali Khan",
   position: "Non-Executive Director"})-[:
   BOARD_MEMBER_OF_SBP]-(sbp)
12 MERGE (aliLatif:Person {name: "Mr. Muhammad Ali Latif",
   position: "Non-Executive Director"})-[:
   BOARD_MEMBER_OF_SBP]-(sbp)
13 MERGE (imdadBosal:Person {name: "Mr. Imdad Ullah Bosal",
   position: "Secretary Finance"})-[:BOARD_MEMBER_OF_SBP]-(
   sbp)
14
15 //Creating Finance Minister, Ministry and its link.
16
17 MERGE (p2:Person{name: "Muhammad Aurangzeb"})
18 MERGE (financeMinistry:Ministry {name: "Ministry for Finance
   and Revenue"})

```

```

19 MERGE (p2)-[:HEADS]->(financeMinistry)
20 Return *
21
22 //Linking Finance Secratary with a unique edge of reports to
    Finance Minister
23 MATCH(p1:Person{name:"Mr. Imdad Ullah Bosal"})
24 MATCH(p2:Person{name: "Muhammad Aurangzeb"})
25 MERGE(p1)-[:REPORTS_TO]->(p2)
26 return *
27
28 //Fixing Property
29 MATCH (muhammadAurangzeb:Person {name: "Muhammad Aurangzeb"
    })
30 SET muhammadAurangzeb.position = "Minister"

```

Now we have a working State Bank Node and its further Heirarchy upto the finance minister. Now we just need to connect it with its relevant sector.

```

1 MATCH(s:Sector{name:"Commercial Banks"})
2 Match(b:Ministry{name:"State Bank of Pakistan"})
3 MERGE(s)-[:REGULATED_BY]-(b)

```

Now we have all the ministers who should clearly know each other, we can simply add them to a common node so that information can easily flow making ease for the company executives.

```

1 //Ending the Heirarchial Structure
2 MATCH (p:Person)-[r:HEADS]-()
3 MERGE (d:Cabinet{name:"The Cabinet of Pakistan"})
4 MERGE (p)-[:REPORTS_TO]-(d)
5 return *

```

This creates a unique labelled node called cabinet which is kind of the point for ministers to share information.

All in all, Our focus extends beyond mere financial metrics to explore the governance structures, regulatory frameworks, and inter-ministerial influences that shape the operational landscapes of these entities. we aim to provide new insights into corporate governance and strategic alignments within Pakistan's economic framework.

