

LinkedIn Network Report

MANISH KUMAR TIWARI

B.tech in Computer Science (Sitare University)

INTRODUCTION

This report analyzes the LinkedIn network of the 2024–28 batch. We collected data from students in CSV format, then used Python to convert it into a network graph. The graph helped us explore connections, relationships, and interesting patterns within the batch, giving insights into how students are linked professionally.

Data Cleaning & Graph Creation

We used Python to:

- - Clean names and remove special characters
- - Ignore students with 0 connections
- - Normalize all names in title case
- - Store the graph as adjacency list (JSON + text format)
- The final graph had 28449 students and 100362 valid connections.

Network Statistics

- Total Students (after cleaning): 126
- Total Connections: 100362
- Average Connections per Student: 7.06
- Most Connected Student:
 - 1. Rohit Malwaiya: 4258 Connections
 - 2. Ravi Rajput: 4070 Connections
 - Manoj Dewda: 3969 Connections
 - Ramraj Nagar: 3713 Connections
 - Nirmal Mewada: 3524 Connections

Random Walk & Path Analysis

- We used a random walk technique to connect two students randomly.
- We calculated the average length of random walks and pruned paths.

- Random Walk Length: 100
- Pruned Path Length: 80
- Mean Length: 100
- Median Length: 79.5
- Standard Deviation: 3.59

Insights & Key takeaways:

- - Most students are well connected, showing strong professional networking.
- - A few students act as bridges between multiple groups.
- - A couple of students had 0 connections, showing they haven't yet built their LinkedIn network.
- - The average connection count shows most people have around 7–10 classmates connected.

This project helped me apply graph theory to real LinkedIn data from our class. I learned to clean data, build a network, and find insights like degrees and connection paths. It improved my Python skills and showed how data science reveals hidden patterns in networks.

CONCLUSION:

This analysis helps understand how our class is connected on LinkedIn. It shows which students are central, and which might need to expand their network. This kind of network analysis can be used for collaboration, team formation, and professional growth tracking.

Future Scope & Innovations:

In the future, this analysis can be expanded by integrating more data fields (like interests, locations, or interactions). A web-based visual interface can allow live exploration of the network. We can even implement algorithms to recommend valuable connections, forming a strong knowledge-sharing community in our class.