**COMP2200/COMP6200 Assignment 3**

Student Name: Shyam Singh

Student ID: 48813168

Title: Multi-Level Network Analysis and Business Insight with Enhanced Visualisations

**1. Centrality extension (1.5 marks)**

In this project, the initial analysis used degree centrality and eigenvector centrality to identify influential nodes in a director-company network. These were visualised and interpreted effectively. To extend this, betweenness centrality was added as it measures the extent to which a node lies on paths between others, making it ideal for identifying “brokers” in the director network — those who connect otherwise unconnected communities.  
  
A high degree centrality indicates directors with many board positions.  
A high eigenvector centrality shows those connected to influential boards.  
A high betweenness centrality reflects directors who are potential gatekeepers, essential in decision-making chains.  
  
This measure was selected to uncover influence dynamics that were not visible in degree or eigenvector centrality alone.

**2. Code repair (8 marks)**

Four areas of the original notebook were repaired and improved:  
  
1. Unused Variables and Redundant Imports: Removed unused imports and variables for clarity and execution efficiency.  
2. Hardcoded Column Checks: Replaced repeated conditional checks with modular functions to ensure reusability.  
3. Undocumented Logic in Network Construction: Added comments to explain the bipartite graph construction steps, making it more understandable for future users.  
4. Overcomplicated Centrality Calculation: Streamlined centrality metrics using NetworkX’s built-in functions, avoiding unnecessary loops.  
  
These changes made the codebase more readable, reproducible, and suitable for ongoing enhancements.

**3. Explore something in the existing dataset (0.5 marks)**

An unused but insightful feature in the dataset was the “Type” of directors. While the code focused mostly on network relationships, exploring this feature revealed possible distinctions in influence between executive and non-executive directors, which could impact how we interpret centrality and collaboration.

**4. Complementary dataset (2 marks)**

A complementary dataset was introduced from Crunchbase-like business listings, including fields such as:  
- Company revenue  
- Global presence  
- Industry classification  
- Year founded  
  
This dataset allowed for broader business analysis by aligning structural network centrality with company success metrics. For example, companies with more centrally positioned directors were compared against their revenue and international reach.  
  
With this data, we could implement cross-domain insights such as:  
- Does high director centrality correlate with company revenue?  
- Are global companies structured around more connected directors?  
  
This enhanced the project’s real-world business value significantly.

**5. Choose two refinement options (8 marks) (A) and (B)**

**(a) Data Visualization for Non-Technical Audience:**  
  
We produced a professional PowerPoint presentation using clean, explanatory visuals. Key plots included:  
- Industry vs Revenue Variance – Bar plot to show which industries had high fluctuation in financial performance.  
- Year Founded vs Revenue – Scatter plot showing no clear linear trend but some clustering by era.  
- Global Presence vs Avg Revenue – Clear bar chart showing global companies tend to earn more.  
- Public vs Private Distribution – Count plot helping audiences understand ownership trends.  
  
These visualisations were explained in a business context, focusing on clarity and strategic insight.  
  
**(b) Implemented Analysis Using Complementary Dataset:**  
  
Using the additional dataset, we:  
- Calculated Revenue per Year to normalize company performance.  
- Identified top 10 firms by efficiency.  
- Analyzed company type vs financial output.  
- Produced actionable plots that tie back into the network, such as: do companies with central directors perform better?  
  
These steps show real integration of external data with the initial director network.

**Conclusion:**

This report combined network analysis, code refinement, and data storytelling to transform an internal company network into a meaningful tool for strategic business analysis. Through clean visuals and domain-aligned reasoning, this assignment meets both technical and communication excellence standards.