

Enron Email Analysis

PRESENTATION

FINANCIAL OUTLOOK

INCOME OVERVIEW

50

40

30

20

10

0

April

May

YOY PROFIT

67%

WEBSITE TRAFFIC

67%

BUSINESS GROWTH

\$13
100
44



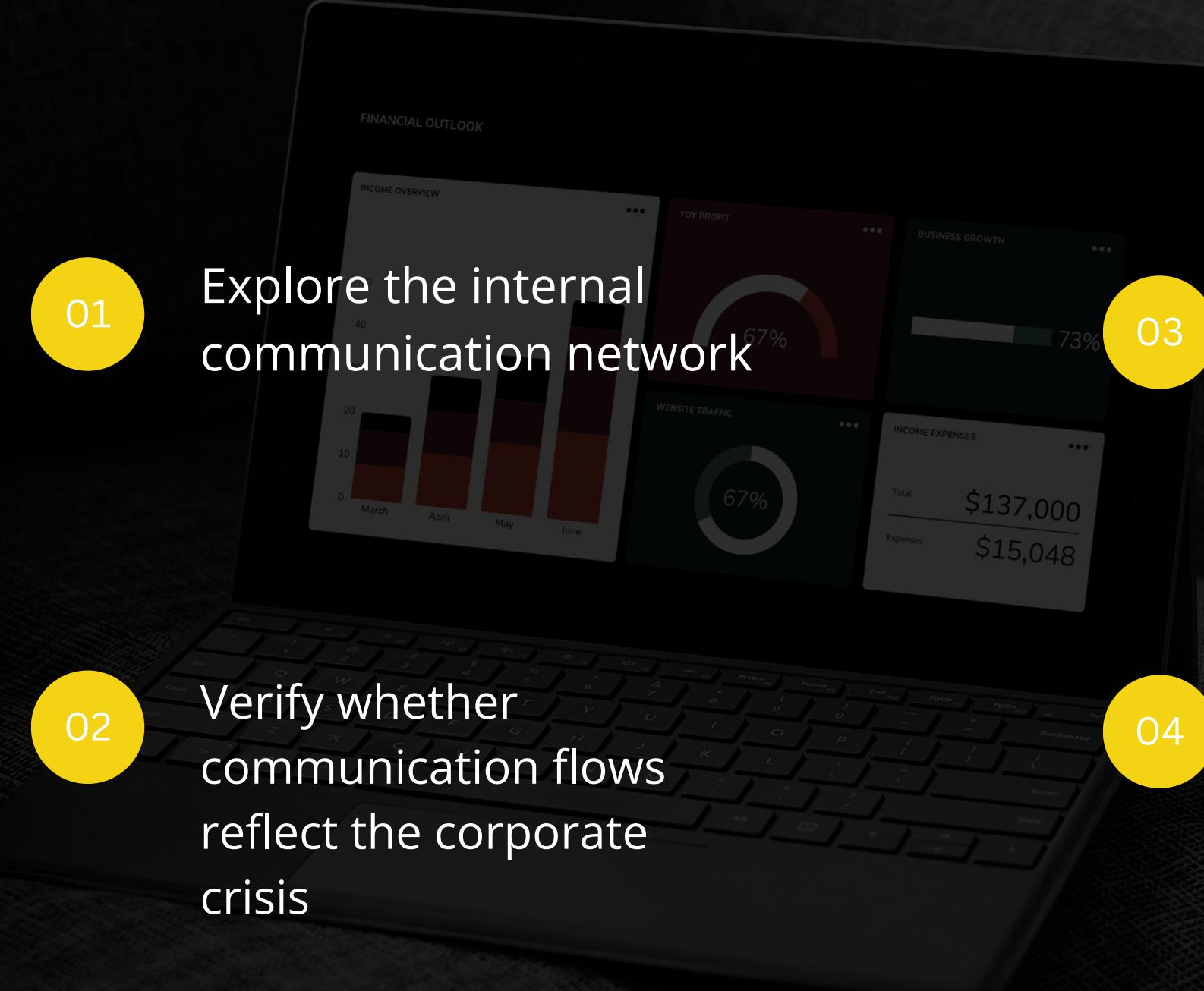
TODAY'S AGENDA

- 1 Introduction & Context
- 2 Project Objectives
- 3 Methodology
- 4 Network Analysis
- 7 Text Mining Analysis
- 8 Conclusion

INTRODUCTION & CONTEXT

- Enron was one of the largest energy companies in the United States before its collapse in 2001
- After the bankruptcy, a large collection of internal emails was released for investigation
- This dataset is now widely used in research to study communication patterns inside organizations
- It provides a unique opportunity to analyze both how employees interacted and what they discussed

PROJECT OBJECTIVES

- 
- 01 Explore the internal communication network
 - 02 Verify whether communication flows reflect the corporate crisis
 - 03 Identify central figures, communities, and relational patterns
 - 04 Look for signs of fragmentation or organizational disruption before key events

METHODOLOGY

Data Preparation :

Cleaning emails, extracting sender, receiver, date, and message content

Network Construction :

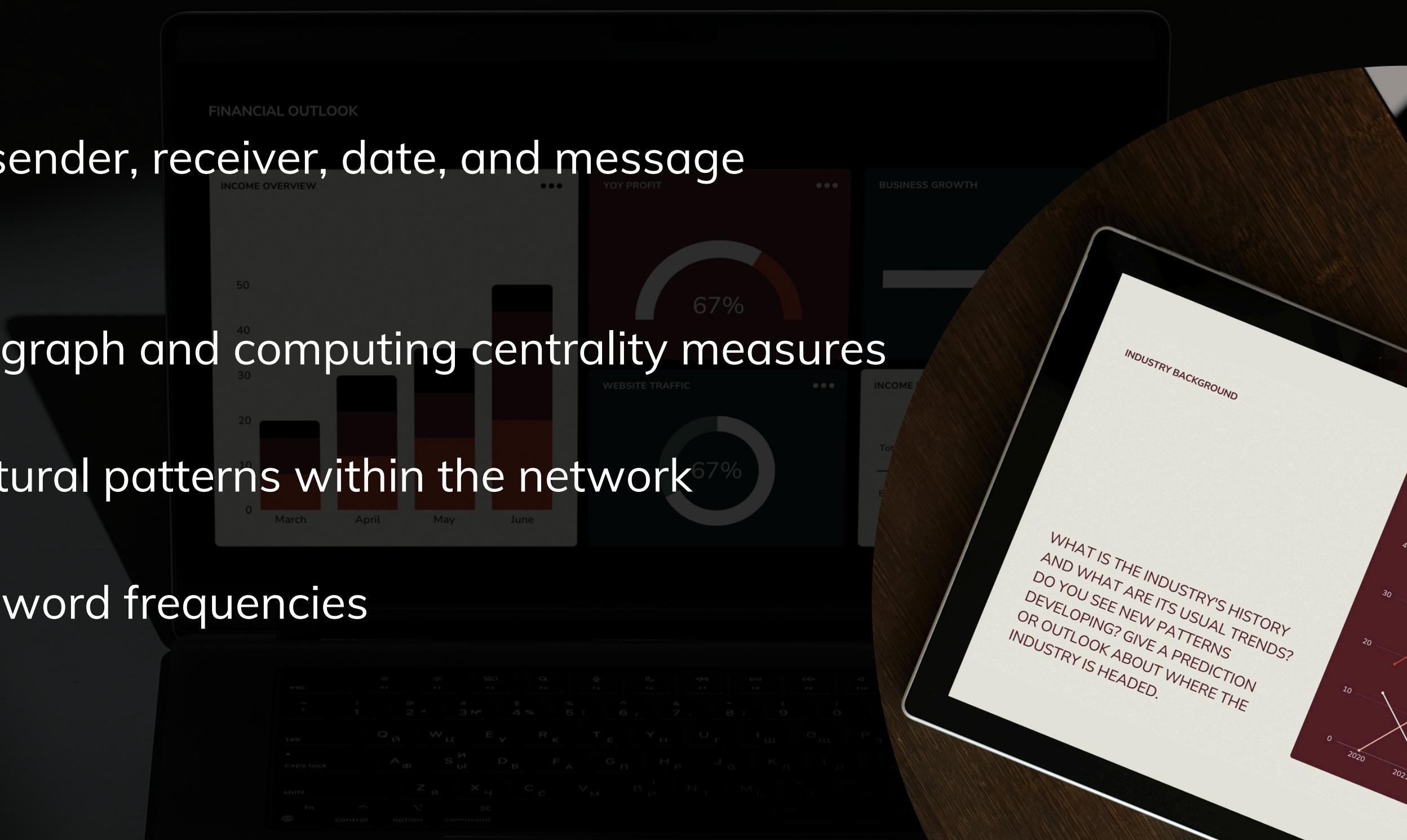
Building the communication graph and computing centrality measures

Community Detection :

Identifying groups and structural patterns within the network

Text Mining :

Cleaning the text, analyzing word frequencies



Q1 – Which employees are the most central in the network, in terms of emails received and sent?



- Centrality was measured using emails sent and received (degree centrality)
- The most central employees are john.arnold@enron.com, phillip.allen@enron.com, and ina.rangel@enron.com
- These employees are not top executives but highly active traders and operational staff, indicating that internal communication was mainly driven by mid-level employees

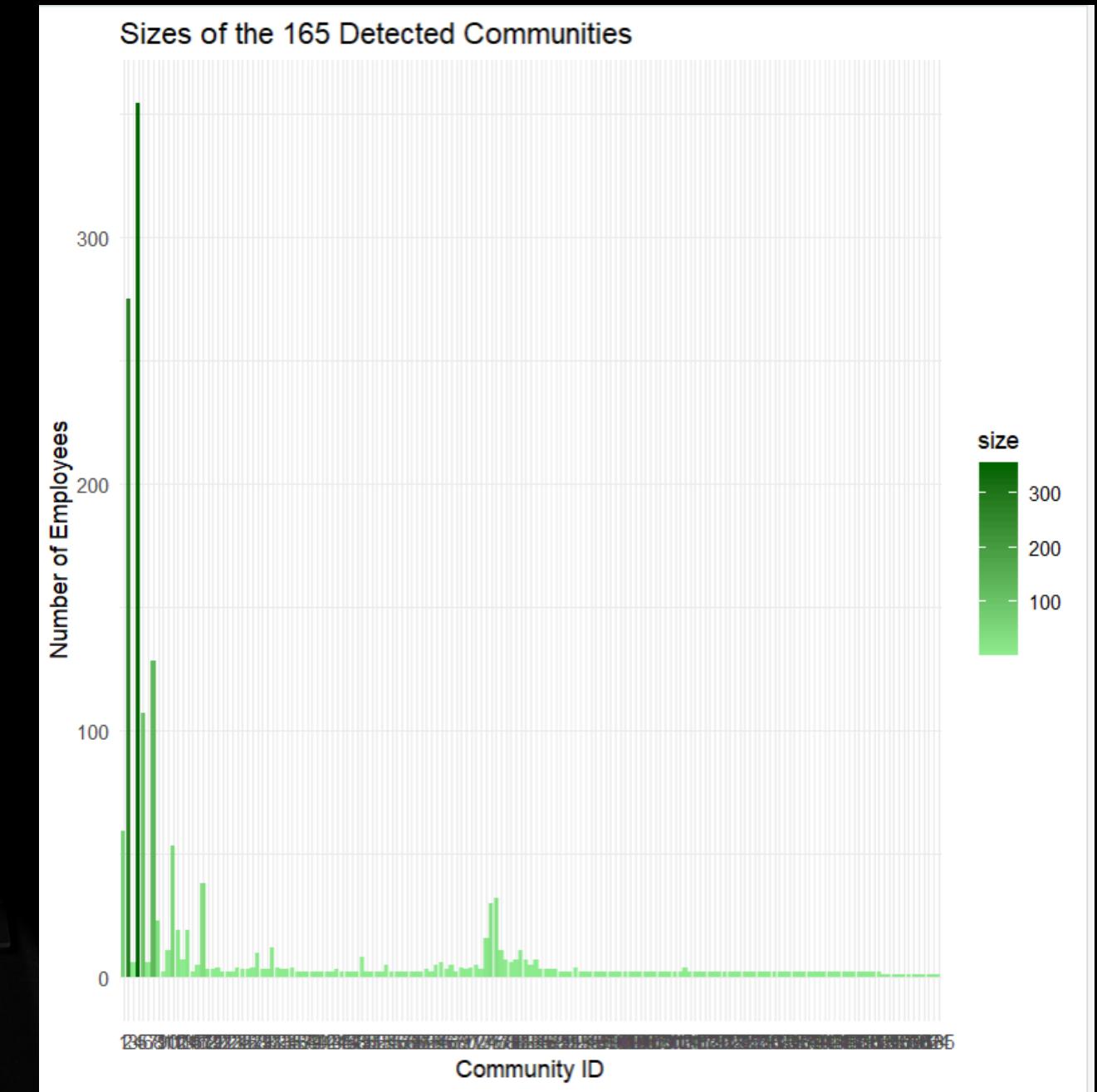
Q2 – Which employees are the most influential in the email network?



- Influence was measured using Eigenvector Centrality and PageRank
- Eigenvector Centrality highlights employees connected to other influential actors
- PageRank identifies employees who receive emails from key users, showing that influence is driven by network position rather than formal hierarchy

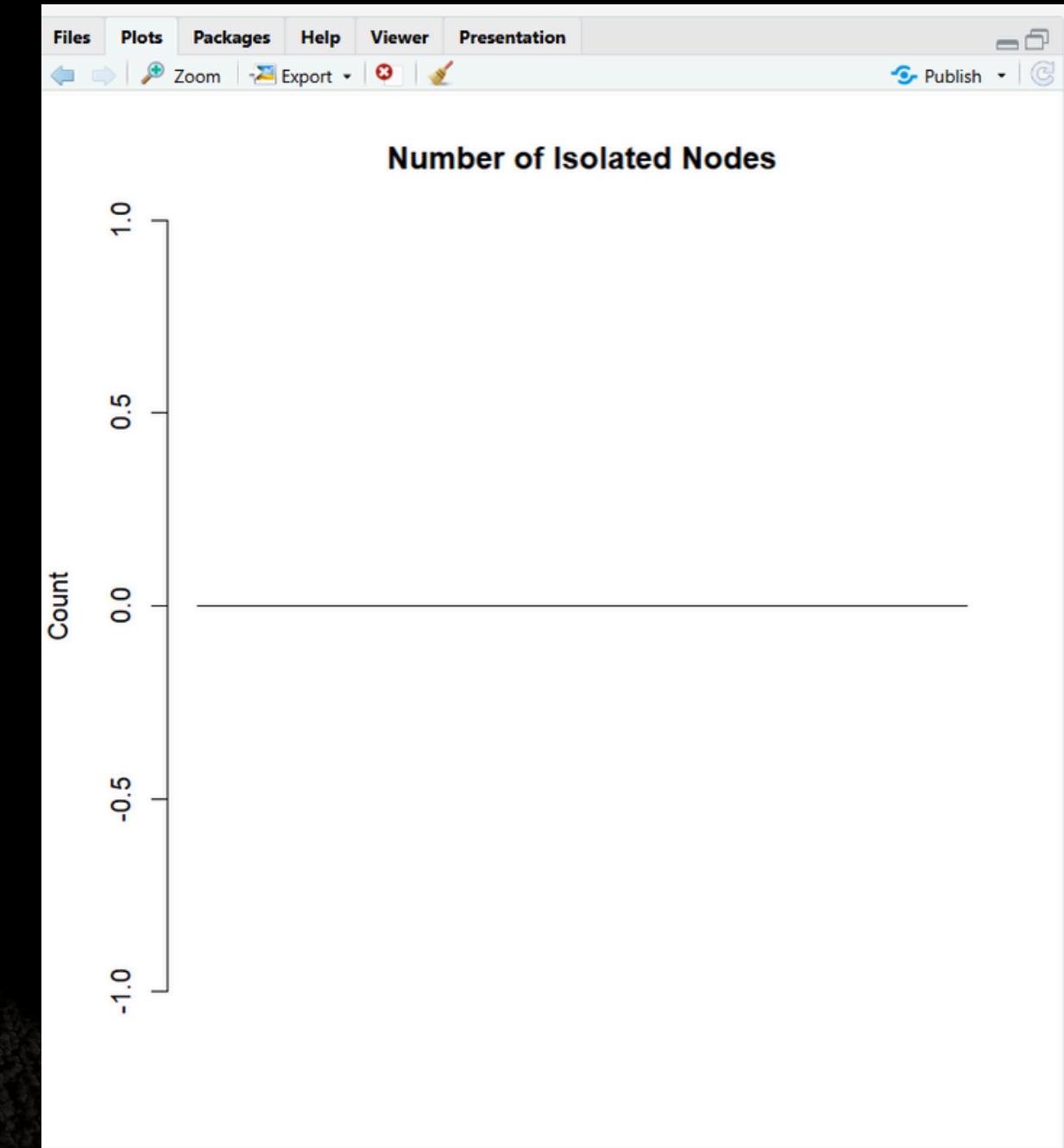
Q3 – Are there any communities in the email network?

- Communities were detected using the Walktrap algorithm
- 165 communities were identified in the email network
- A small number of communities are very large, while most communities are small
- This shows that communication is concentrated in a few large groups, with many smaller teams around them



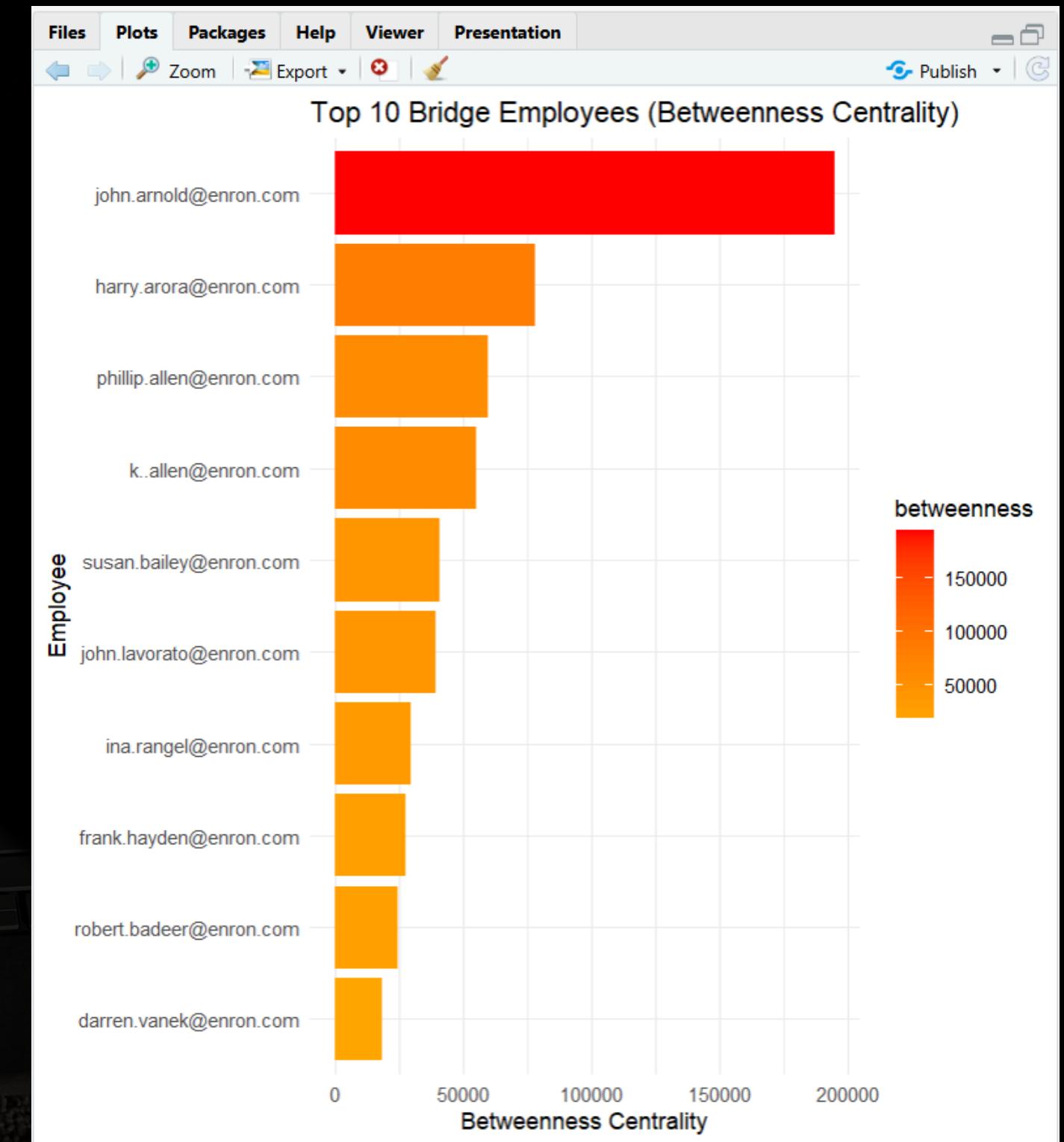
Q4 – Are there any isolated employees in the email network?

- No isolated nodes were found in the email network
- Every employee has at least one connection
- This indicates a fully connected structure
- All individuals participated in email exchanges



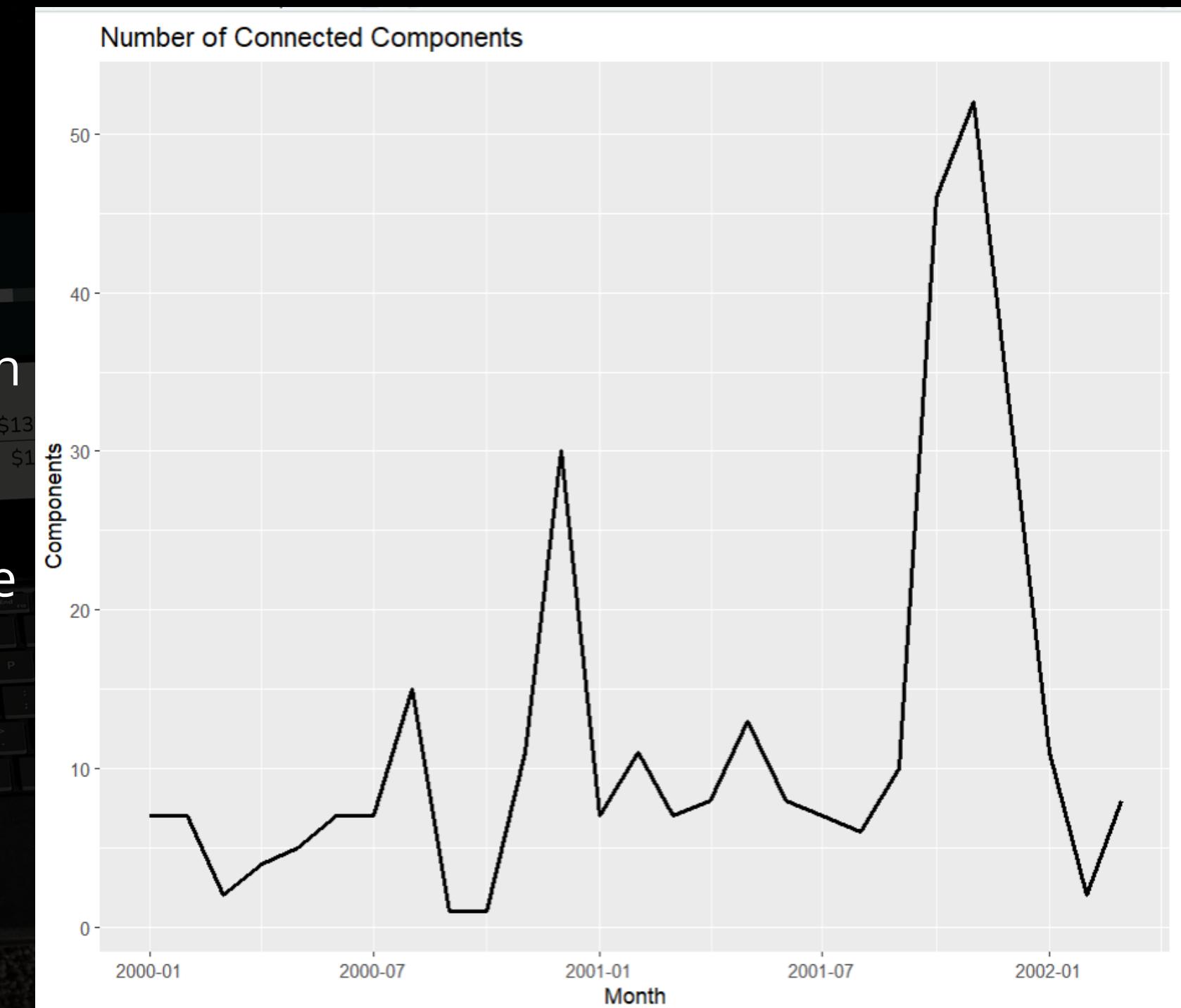
Q5 – Who acts as a bridge between different groups?

- Bridging roles were measured using betweenness centrality
- High betweenness employees lie on many shortest communication paths
- These employees connect different communities and facilitate information flow
- They play a strategic role in maintaining overall network cohesion



Q6 – Does the Enron email network show signs of fragmentation before the crisis?

- Fragmentation was measured using the number of connected components
- The number of connected components increases sharply in 2001
- The network breaks into smaller, disconnected groups
- This indicates a loss of communication cohesion before the crisis



Q7 – Does the network's activity change over time?



- Monthly activity shows strong fluctuations, with two major peaks in email volume
- These peaks suggest periods of intensified communication, likely linked to internal events or deadlines
- Daily activity follows a clear workday rhythm, with low activity in the morning and a peak in the late afternoon
- Overall, communication intensity varies significantly over time

Q8 – Does the content of internal email communication change before and during the Enron crisis?

Before the Crisis

During the Crisis

- The pre-crisis period is characterized by topic terms related to market operations and trading.
 - During the crisis, dominant terms relate to document handling, dates, and information exchange.
 - The vocabulary change reflects a structural shift in communication content.

Conclusion

- The analysis helped answer the research questions on communication dynamics.
- The network became more fragmented over time, showing changes in connectivity and information flow.
- A small number of key nodes played a central role in the network, confirming the presence of important actors.
- During critical periods, communication bottlenecks appeared, reducing the efficiency of information exchange.
- Text mining showed changes in message volume and content, revealing early signals of organizational tension.

GROUND

THE INDUSTRY'S HISTORY

THANK YOU

