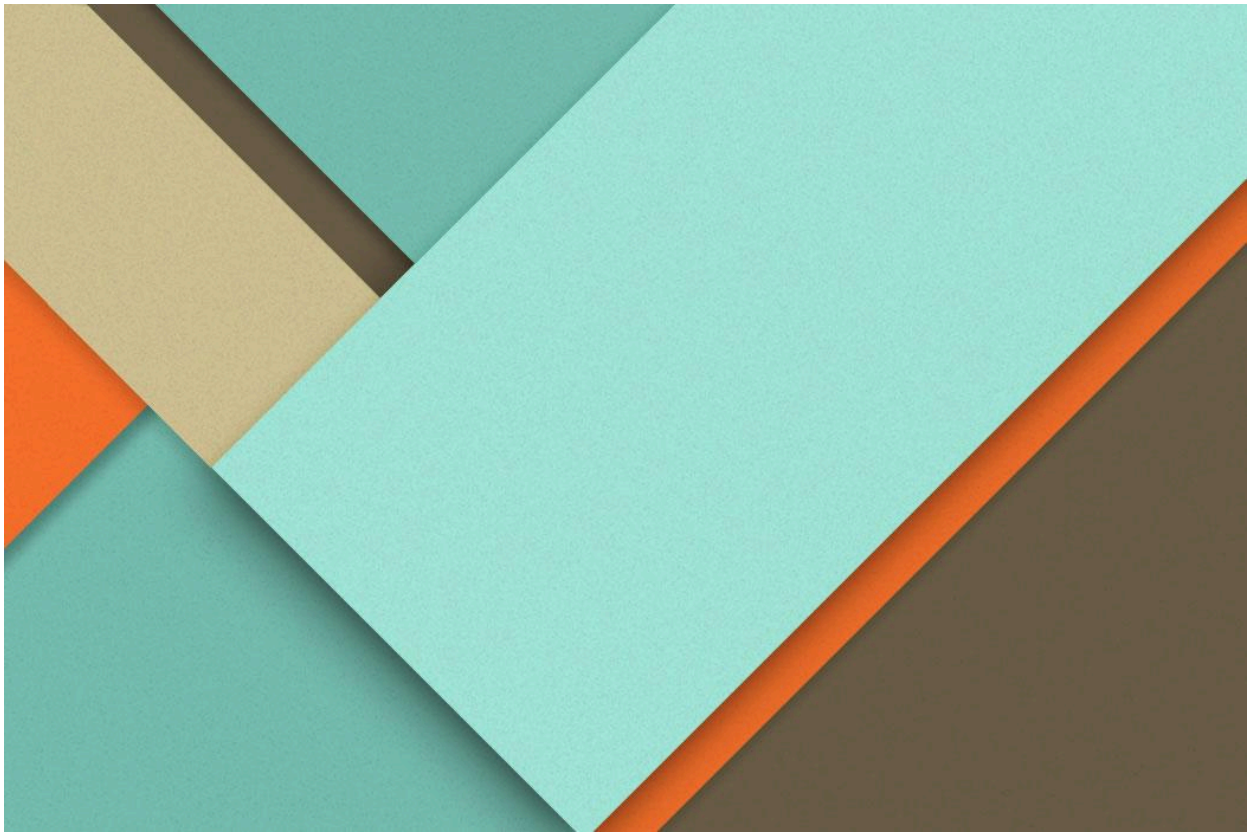

Amazon Co-Purchasing Network Analysis using Neo4j

—

Team CheeseCake
Omkar Bhandare
Shivam Choudhury
Tanishka Rahate
Ayush Mundada
Jadhav Udaykiran

Overview

This project aims to analyze Amazon product relationships using Stanford's SNAP co-purchasing dataset through the Neo4j graph database. The system will uncover hidden product associations, identify influential items via PageRank, and provide actionable insights for recommendation systems and market analysis.

Goals

1. Implement a graph database model for purchasing relationships.
2. Develop a recommendation engine using Cypher query patterns.
3. Calculate PageRank scores to identify the product anchors.
4. Build an interactive web platform for data exploration.

Specifications

Data: Stanford SNAP Amazon Dataset (v2025)

Analysis: PageRank, Louvain Community Detection

Database: Neo4j Desktop, GDS Library

Backend: Python, Flask, Neo4j Python Driver

Frontend: HTML, CSS, JavaScript, D3.js

Weekly Work Plan

I. Week 1

a. Getting acquainted with Graph Databases and Cypher Language concerned with Neo4j Database
b. Dataset acquisition and preprocessing
c. Neo4j environment configuration
d. Basic Cypher query implementation

II. Week 2

a. Importing complete dataset
b. PageRank/Betweenness centrality implementation
c. Community detection algorithms (study + implement)

III. Week 3

a. Develop an interface
b. Configure Flask API endpoints for Neo4j access
c. Recommendation UI with filters

IV. Week 4

a. Interactive dashboard polishing, testing
b. Documentation
c. Final Project Presentation