

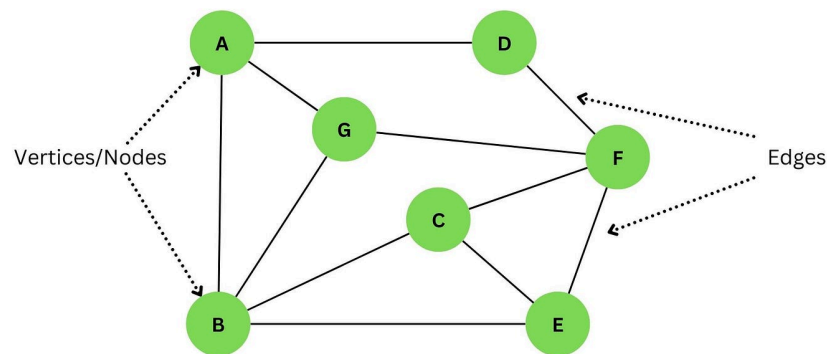
Assignment 05

Penambahan Graf



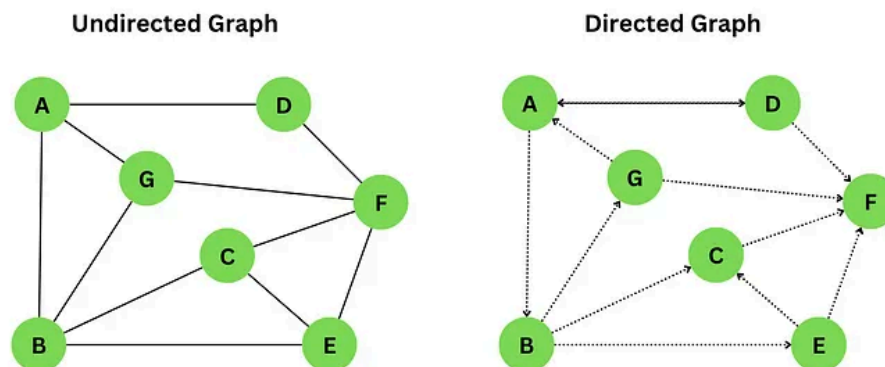
Review Penambahan Graf

Graf adalah struktur data yang terdiri dari simpul (node atau vertices) dan sambungan antara simpul-simpul tersebut (edge). Sebuah graf G dengan sekumpulan vertices (V) dengan sekumpulan edge (E) direpresentasikan sebagai $G(V, E)$.



sumber: <https://medium.com/@nelsonjoseph123/graph-as-a-data-structure-d04db591a0e5>

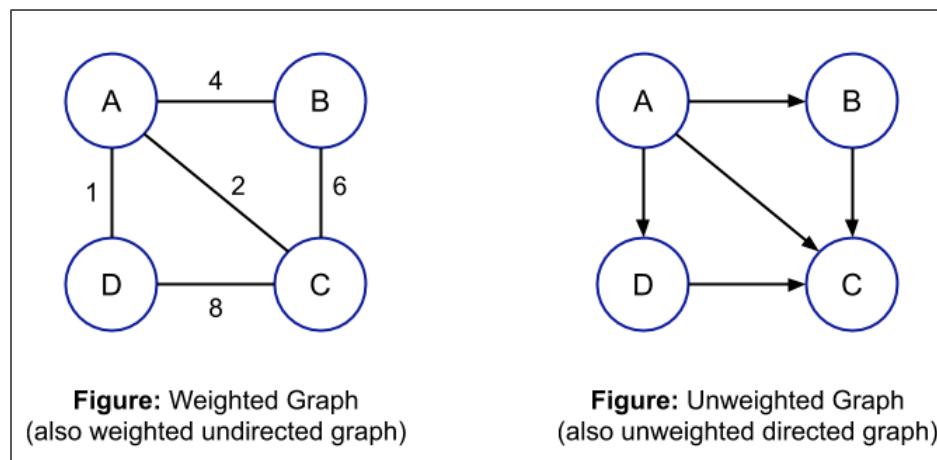
Graf dapat dibagi menjadi dua jenis utama berdasarkan arah sambungan antar simpulnya, yaitu graf berarah (directed graph) dan graf tidak berarah (undirected graph).



sumber: <https://medium.com/@nelsonjoseph123/graph-as-a-data-structure-d04db591a0e5>

Pada graf tidak berarah, sambungan antara dua simpul tidak memiliki arah. Artinya, jika terdapat sambungan (edge) antara simpul A dan simpul B, maka sambungan tersebut bisa dianggap sebagai dua arah, dari A ke B atau dari B ke A. Pada graf berarah, setiap sambungan antara dua simpul memiliki arah yang jelas. Artinya, sambungan tersebut memiliki simpul awal dan simpul tujuan. Jika terdapat sambungan dari simpul A ke simpul B, tidak dapat diasumsikan bahwa terdapat sambungan sebaliknya dari B ke A.

Graf berbobot (Weighted graph) adalah jenis graf di mana setiap edge diberi nilai numerik yang disebut "bobot". Bobot ini mewakili ukuran kuantitatif tertentu yang terkait dengan edge, seperti jarak, biaya, waktu, atau metrik lain yang relevan. Kehadiran bobot menambahkan lapisan informasi tambahan pada graf, yang memungkinkan representasi yang lebih rinci dari hubungan antar node.



sumber:

<https://afteracademy.com/images/introduction-to-graph-in-programming-weighted-unweighted-graph-52bd4ce5100286b8.png>

Penambangan Graf (Graph Mining) berkaitan dengan ekstraksi informasi dan penemuan pola yang terkait dengan struktur graf atau jaringan. Penambangan Graf mencakup berbagai aspek, termasuk analisis ukuran yang digunakan untuk menentukan sejauh mana suatu node penting atau berpengaruh dalam suatu jaringan atau graf (Node Centrality), identifikasi pola subgraf yang sering muncul (Frequent Subgraph Mining), prediksi koneksi yang mungkin terjadi (Link Prediction), serta visualisasi dan pemahaman struktur graf secara keseluruhan (Node Visualization).

Tugas: Node Centrality, Node Visualization, Frequent Subgraph Mining

A. Node Centrality

Node centrality adalah ukuran yang digunakan untuk menentukan sejauh mana suatu node penting atau berpengaruh dalam suatu jaringan atau graf. Ada beberapa metrik node centrality yang umum digunakan:

- Degree centrality
- Closeness centrality
- Betweenness centrality.

Misalnya, dalam jaringan sosial, degree centrality dapat mengukur seberapa banyak teman yang dimiliki oleh seorang individu. Closeness centrality dapat mengukur seberapa cepat informasi dapat menyebar dari suatu node ke semua node lainnya. Sedangkan betweenness centrality dapat menunjukkan seberapa sering suatu node menjadi perantara dalam jalur komunikasi.

B. Node Visualization

Node visualisasi melibatkan representasi grafis dari node dalam suatu jaringan atau graf. Ini membantu pemahaman visual tentang struktur dan hubungan antar-node dalam graf.

C. Frequent Subgraph Mining

Frequent subgraph mining mencari pola berulang atau subgraf yang sering muncul dalam kumpulan data graf. Tujuannya adalah untuk mengidentifikasi struktur yang signifikan dan sering muncul dalam data graf.

Misalkan kita memiliki kumpulan data graf yang mewakili interaksi molekuler dalam suatu senyawa kimia. Frequent subgraph mining dapat membantu mengidentifikasi pola subgraf yang sering muncul, yang mungkin mencerminkan struktur molekuler yang penting atau karakteristik tertentu. Dengan demikian, hal ini dapat digunakan dalam penemuan obat atau analisis struktur molekuler.

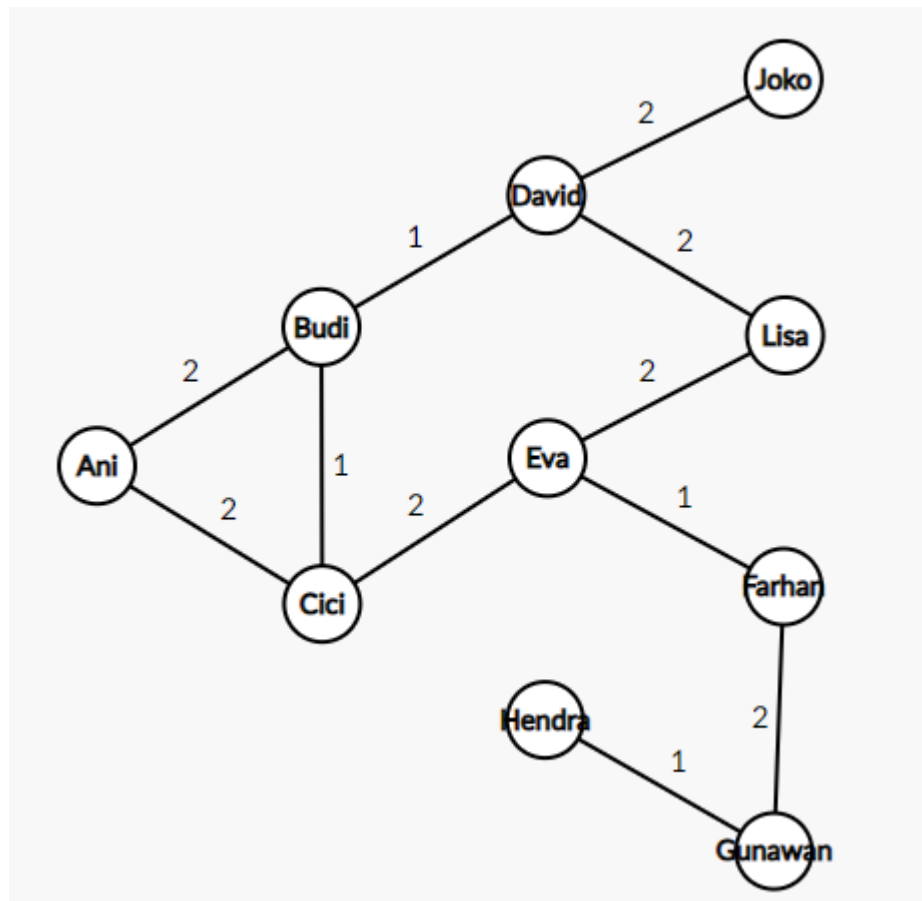
Pada tugas ini, Anda diberikan dataset graph.csv yang berisi node dan node pasangannya, serta weight dari setiap edge. Setiap baris merepresentasikan node yang berhubungan. Silakan download dataset [di sini](#).

Soal:

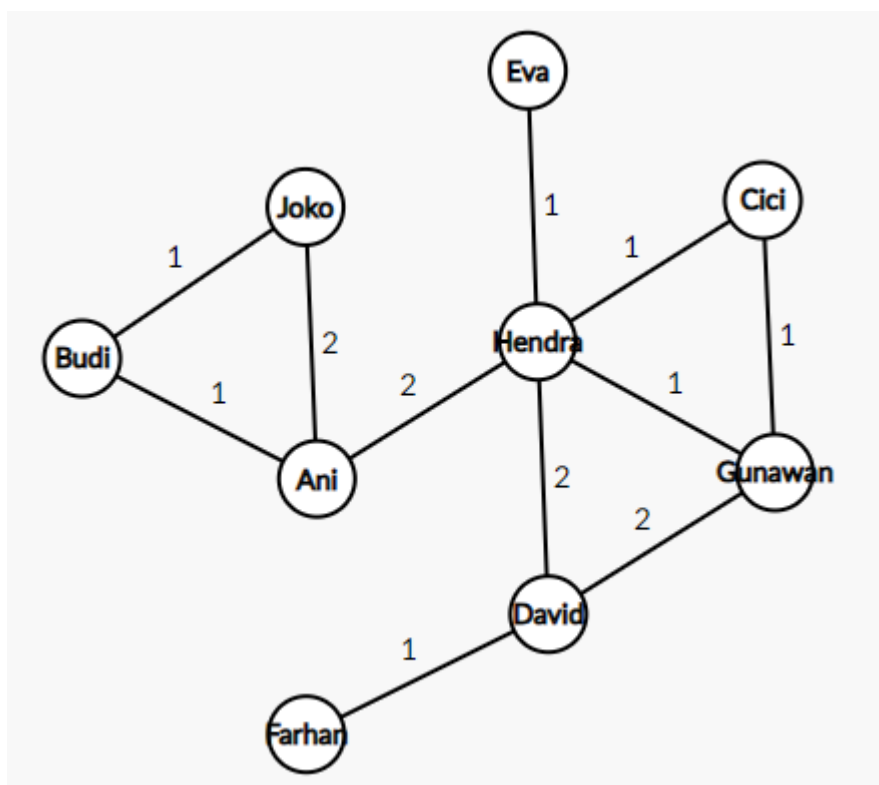
1. Tulislah kode Python untuk menghitung degree centrality untuk semua orang pada data tersebut dengan memberikan keterangan siapa saja yang memiliki hubungan ke orang tersebut! Format output dibebaskan. Jalankan kodenya dan perlihatkan hasilnya!
2. Asumsikan terdapat arah pada data tersebut, dengan arah dari **node1 ke node2**. Tulislah kode Python untuk menghitung in-degree, out-degree, dan total degree centrality untuk setiap orang pada dataset tersebut! Format output dibebaskan. Jalankan kodenya dan perlihatkan hasilnya!
3. Asumsikan terdapat arah pada data tersebut, dengan arah dari **node1 ke node2** dengan bobot **weight**. Farhan ingin mengirimkan pesan ke semua temannya, tetapi ia tidak tahu bagaimana cara yang paling mudah untuk menghubungi seseorang (asumsikan makin besar weight maka makin sulit dihubungi). Tentukan total weight terkecil yang dibutuhkan Farhan untuk mengirimkan pesan ke semua temannya, beserta alurnya! Contoh: Farhan ke Frans: Joko (3) + Gunawan (1) = 4
4. Buatlah visualisasi dari data tersebut dalam bentuk Graph berarah dan berbobot. Silakan melakukan eksplorasi terkait library untuk melakukan visualisasinya!

5. Diberikan beberapa **undirected** graph seperti berikut.

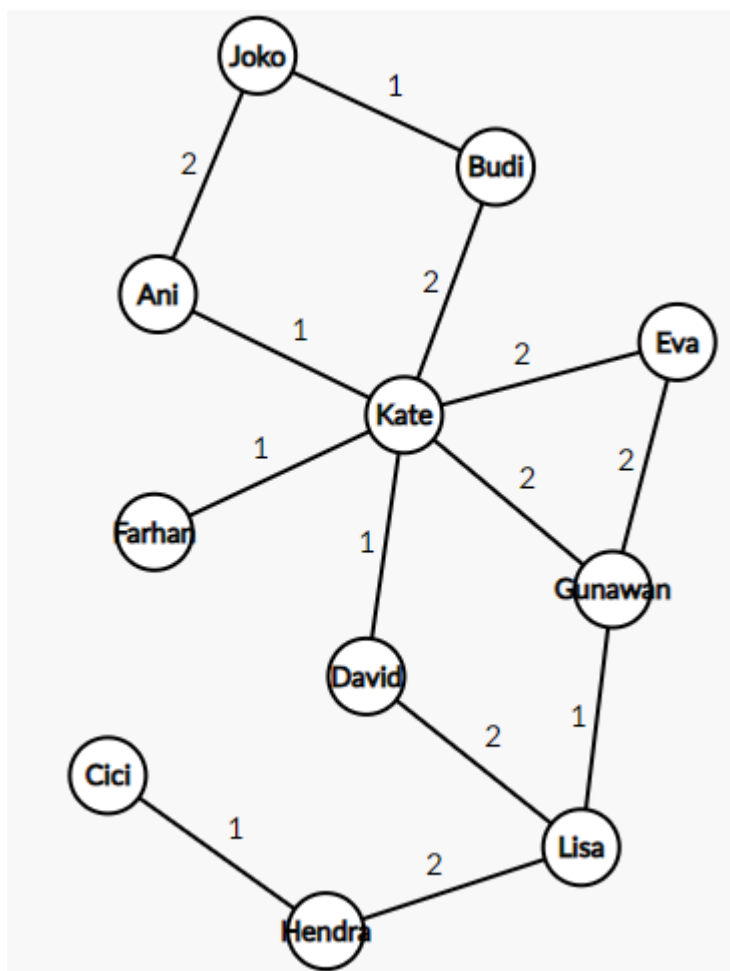
- Graf 1:



- Graf 2:



- Graf 3:



Asumsikan semua orang yang memiliki nama seorang pria dikategorikan sebagai pria, dan semua orang yang memiliki nama seorang wanita dikategorikan sebagai wanita.

Hubungan nodes wanita dengan nodes wanita lainnya yang memiliki weight sama dikategorikan sebagai sebuah struktur atau pola yang sama, begitu juga dengan pria.

Misal:

- Eva — 3 — Lisa dengan Ani — 3 — Cici adalah pola yang sama, yaitu Wanita — 3 — Wanita, maka terdapat total 2 pola untuk jenis ini.
- Hendra — 3 — Gunawan — 2 — Farhan dengan Ani — 3 — Cici — 2 — Lisa adalah pola yang berbeda, karena pola pertama adalah untuk semua pria, dan pola kedua adalah untuk semua wanita.
- Eva — 3 — Gunawan dan Gunawan — 3 — Lisa adalah pola yang sama, karena graph tidak berarah dan polanya adalah pria — 3 — wanita.
- Eva — 3 — Gunawan — 2 — Lisa — 3 — Eva menggambarkan pola yang tertutup (seperti segitiga) dengan kombinasi 1 pria dan 2 wanita dengan bobot 2 untuk pria dengan salah 1 wanita, 3 untuk pria dengan salah 1 wanita, dan 3 untuk wanita dengan wanita.

Tuliskan 5 pola yang paling sering muncul. Contoh : pria — 3 — pria = 10 kali

Silakan tulis semua jawaban pertanyaan dengan jelas dan lengkap secara langsung di sebuah file ipynb. Anda tidak perlu menuliskan soal, cukup menuliskan jawaban setiap nomor.

Untuk soal yang meminta menuliskan kode Python, Anda hanya diperbolehkan menggunakan library Pandas, tetapi tidak boleh digunakan untuk proses perhitungan.

Dilarang keras menyontek. Plagiarisme tidak ditoleransi dan akan dikenai penalti atau nilai akhir E.

Pengurangan nilai akibat keterlambatan pengumpulan tugas akan ditentukan berdasarkan jumlah menit keterlambatan Anda dalam mengumpulkan. Misalnya, apabila terlambat 1 menit, nilai akhir akan dikurangi 1 poin, apabila terlambat 10 menit, nilai akhir akan dikurangi 10 poin, dan seterusnya.

Bobot penilaian:

- Soal 1: 20 poin
- Soal 2: 20 poin
- Soal 3: 20 poin
- Soal 4: 10 poin
- Soal 5: 30 poin

Pengumpulan tugas:

Kumpulkan berkas dengan format penamaan seperti berikut:

Assignment5_[NPM]_[NamaLengkap].ipynb

Contoh:

Assignment5_2006596535_FrancisWilliamSudianto.ipynb