Project

- Each team will use the same data set to study impact of networks.
  - How does (social) networking impact us?

Suggestions

- Process the raw data (1. before and 2. after)
  - Parse a raw of 5 text strings into 5 rows, each of which has one string: Data set D1*
- Cluster D1 into about 10 categories
  - Manually by a few students in your teams as data set D21*
  - Use chatGPT (plus manual if needed) to cluster: by the remaining students in your team, as data set D22*
  - Compare D21 and D22, then consolidate into one data set D2*
  - Summarize differences between g D21 and D22 so we understand more about chatGPT
  - Label clusters using numbers, e.g., 11, 12, …
- Your metrics, evaluation and graphical display
  - Turn the original data set D1 into a set of numbers using the cluster label, this results in data set D3
    - e.g.: CRC (most important 1): 12, food (next important 2): 15, …
  - Devise a metric and use D3 to measure it
  - Plot or show graphically your metric
- Summarize your findings:
  - What are the most important (1), …, and the least important (5) to the class?
  - What differences does the networking make?
  - What are the lessons learned?
  - Provide at least one reference on related work
  - Choose one category as your team like – find out which organization on campus can be reached out to help with the category (website? Person in charge?)

Potential metrics

- You team can choose one from the two slightly different emphasis or invent your own
  - Emphasis 1: Develop an aggregated metric to measure the measure network impact.
  - Emphasis 2: Develop a user based centric to measure the network impact.

In case your team prefers ``physical networks”

- Study practical Wifi by incorporating some physical layer characteristics
  - Find a reference on WiFi data rates currently provided
  - Find a software tool that can measure upload and download rates
  - Choose (a) three locations – two on campus and one off-campus (e.g., at a store); (b) busy and not-busy time
  - Measure the rates at the locations and time periods
  - Graphically show your results
- Submit: data, your plots, and summaries including references and lessons learned

Presentation and submission: One copy per team

- Due: 11:59m 4/25 at canvas. Submit your annotated data sets (D*'s) in .csv format, code, a simple readme about the data sets and how to run your program.

  Also submit your slides in .ppt with file name:

  lastname1_lastname_2…_lastname6.ppt

- Project Presentation (5 slides): 8 min presentation per team 2 mins for questions: Example on annotated data, your approach, results (numerical or plots), summary on findings or lessons learned, your suggestions (e.g., how to reach out or expand the project).

Finally

- Have fun!

- Grade composition (10% for project):
  - Data gathering (3% done already)
  - Technical content (5%)
  - Presentation (2%)