



MENU

ANALYSIS

CONTACT

NETWORK ANALYSIS

ML&DATA SCIENCE NETWORK ANALYSIS

Data Source: Kaggle

IACOBONI SEBASTIAN **MEDIA DIGITALĂ**, ANUL 3

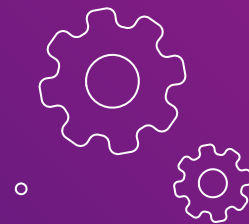


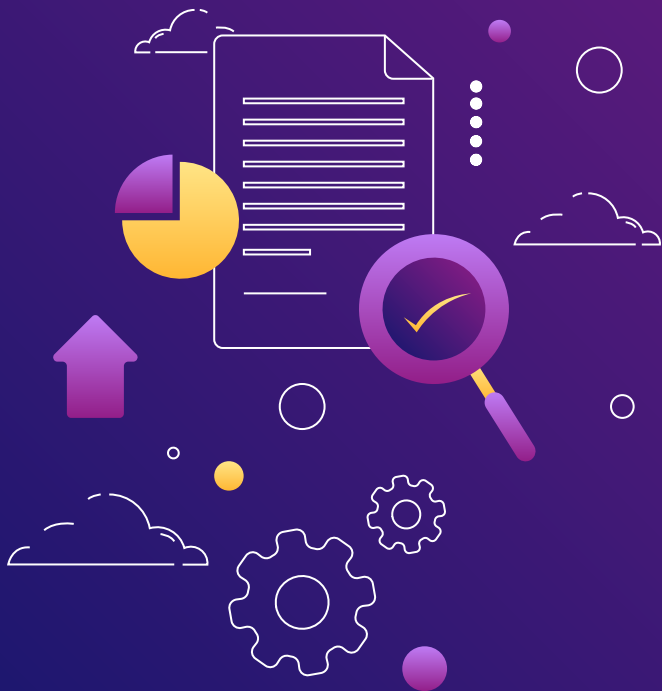


IMPORTANȚA REȚELEI

Am analizat chestionarul anual oferit de cei de la **Kaggle**, unde adună informații despre persoanele care le folosesc platforma. Acest lucru **este relevant** pentru oamenii care deja **lucrează** în domeniu, sau cei care **vor să înceapă** acest lucru, pentru că le oferă un insight legat de domeniu și oamenii aceia, **insight-uri precum:**

- Date demografice;
- Platformele pentru cursurile folosite;
- Programare, de cât timp, în ce, ce instrumente folosesc;
- Job-ul în industrie, salariu, vechime.





COLECTAREA DATELOR

Datele sunt raportate anual din 2017 până în prezent de cei de la Kaggle.





CALITATEA DATELOR



ALEGEREA PARTICIPANȚILOR

A fost trimis un mail tuturor oamenilor care au optat pentru asta pe Kaggle.

2 MINUTE

Răspunsurile care au durat mai puțin de 2 minute au fost șterse.

15 ÎNTREBĂRI

Rezultatele care răspundeau doar la 15 întrebări au fost șterse.





NODURI ȘI CONEXIUNI



NODURILE (NODES)

Reprezintă oamenii care au răspuns la întrebările din chestionar.

AND

CONEXIUNI (EDGES)

Relația dintre nodes și variantele de răspunsuri (limbajele de programare/ environment.



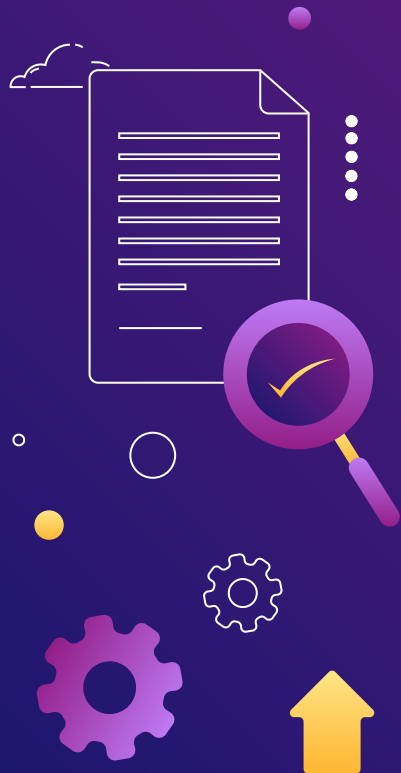


MENU

ANALYSIS

CONTACT

NETWORK ANALYSIS



24012 NODES

59200 EDGES





MENU

ANALYSIS

CONTACT

NETWORK ANALYSIS

DATA SCIENCE PROGRAMMING LANGUAGES



[MENU](#)[ANALYSIS](#)[CONTACT](#)

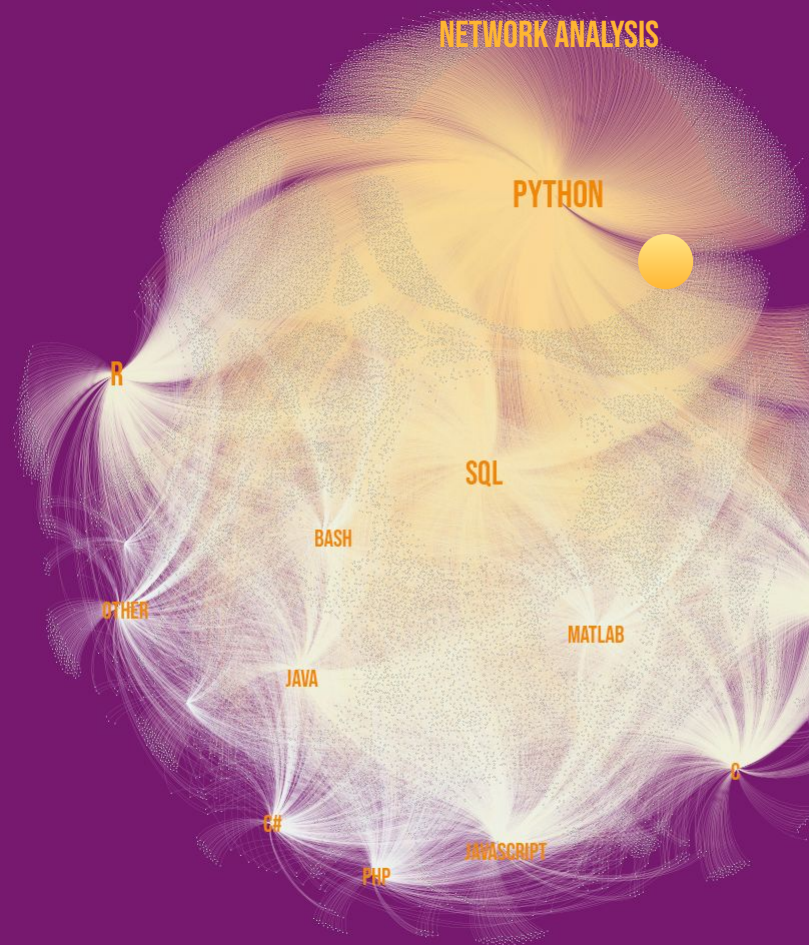
IN- DEGREE: 2

În medie o persoană care programează, folosește aproximativ 2 limbaje de programare.

Popularitatea limbajelor:

- Python
- SQL
- R

NETWORK ANALYSIS



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

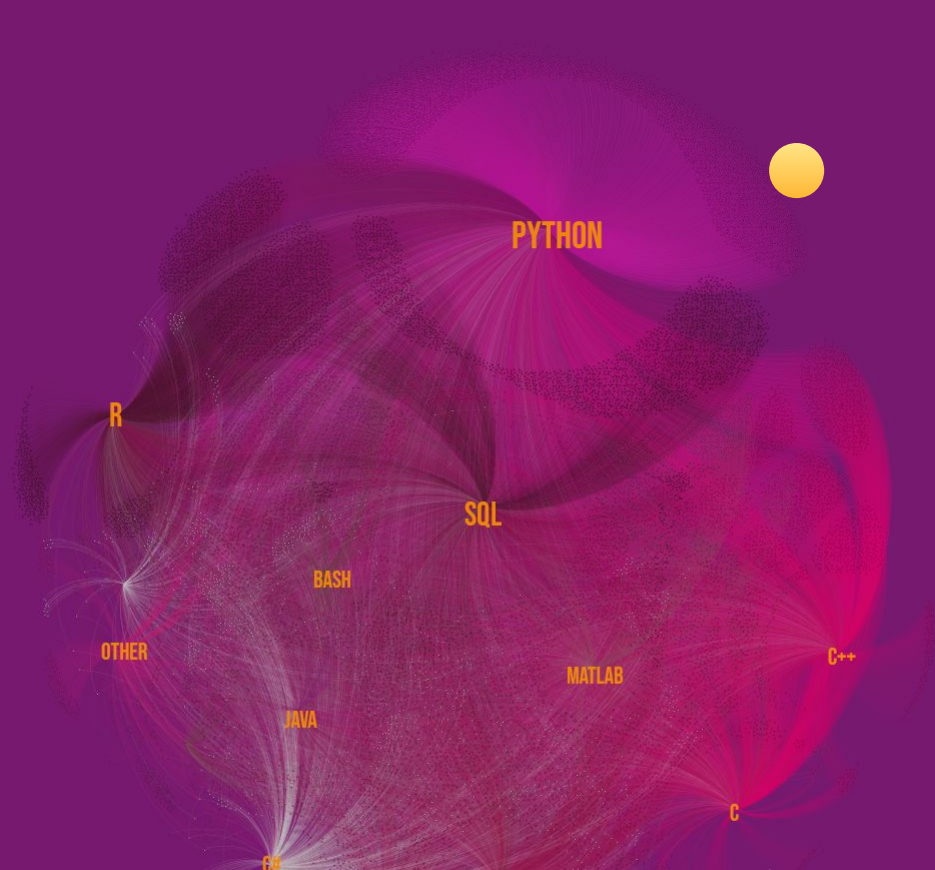
MODULARITY: 10

Culorile au fost ajustate pentru a indica numărul de comunități și membrii firecăria



Cele mai importante comunități.

- Python
- SQL
- R



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

BETWEENNESS

CENTRALITY

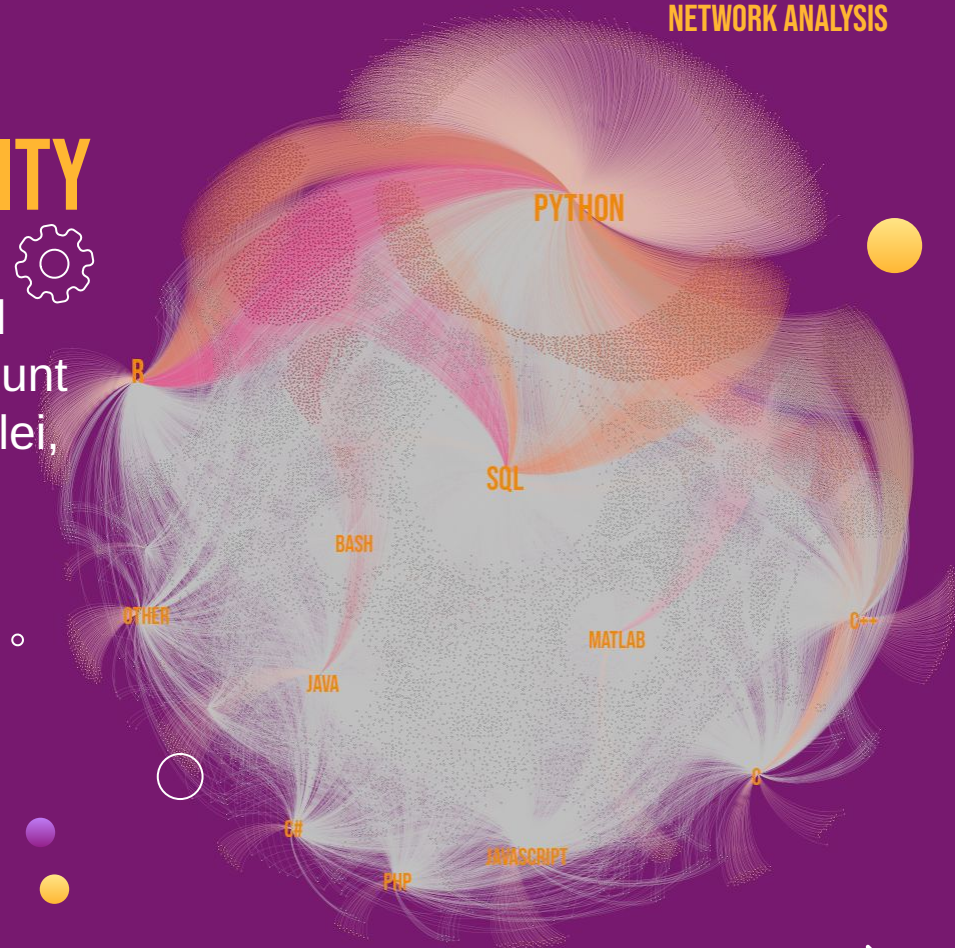


Măsoară numărul de conexiuni al unui limbaj de programare care sunt intermediare cu alte părți ale rețelei, care altfel ar fi fost deconectate



Cele mai multe conexiuni

- Python
- SQL
- R



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

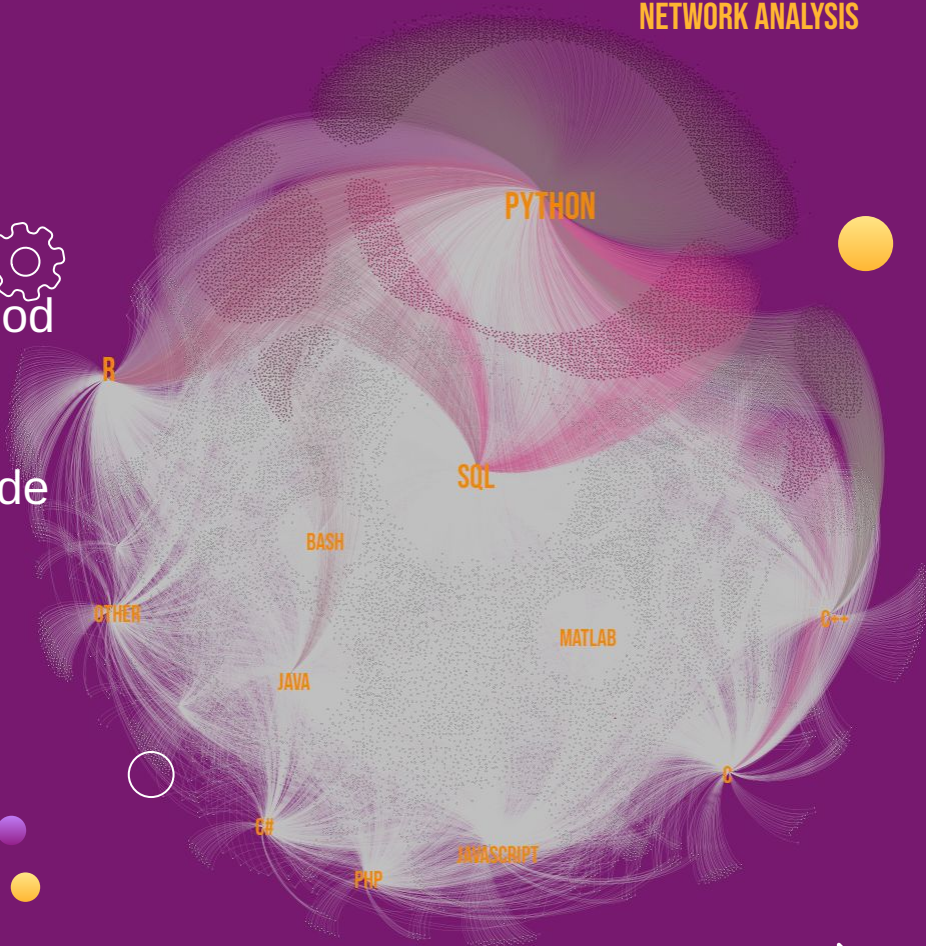
CLOSENESS

CENTRALITY

Măsoară cât de aproape este un nod față de celelalte noduri din rețea. Asta înseamnă că Python este cel mai ușor de conectat cu alt limbaj de programare.

Cel mai conectat nod.

- **Python**





MENU

ANALYSIS

CONTACT

NETWORK ANALYSIS



DATA SCIENCE

INTEGRATED DEVELOPMENT ENVIRONMENTS



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

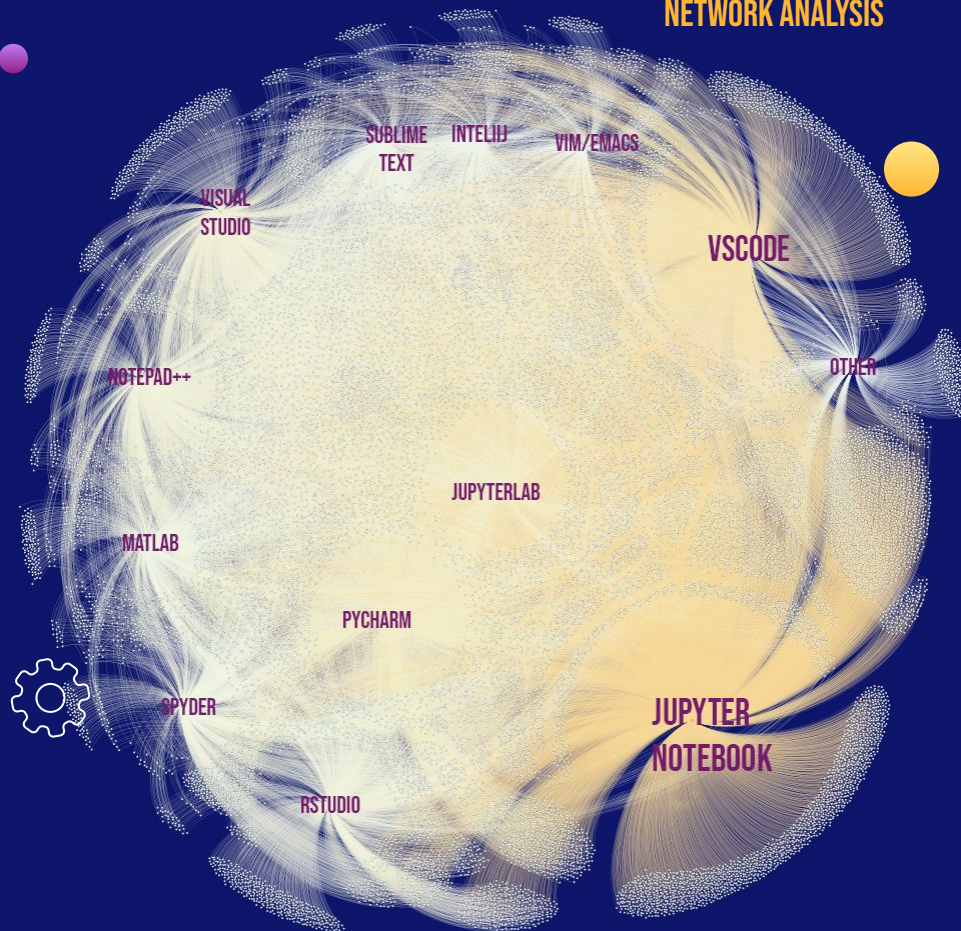
IN- DEGREE: 2

În medie o persoană care este în acest domeniu, folosește două development environments



Popularitatea mediilor:

- Jupyter Notebook
- VSCODE
- PYCHARM



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

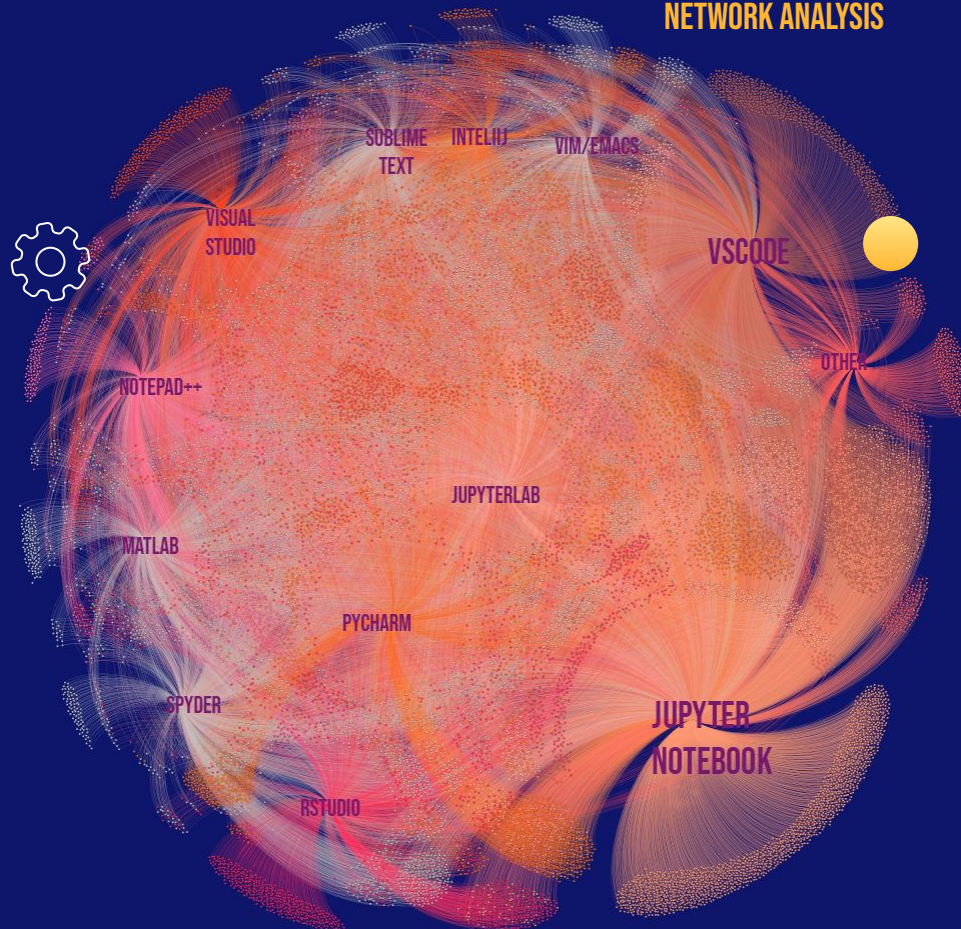
MODULARITY: 12

Culorile au fost ajustate pentru a indica numărul de comunități și membrii firecăria



Cele mai importante comunități.

- Jupyter Notebook
- VSCODE
- PYCHARM



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

BETWEENNESS

CENTRALITY

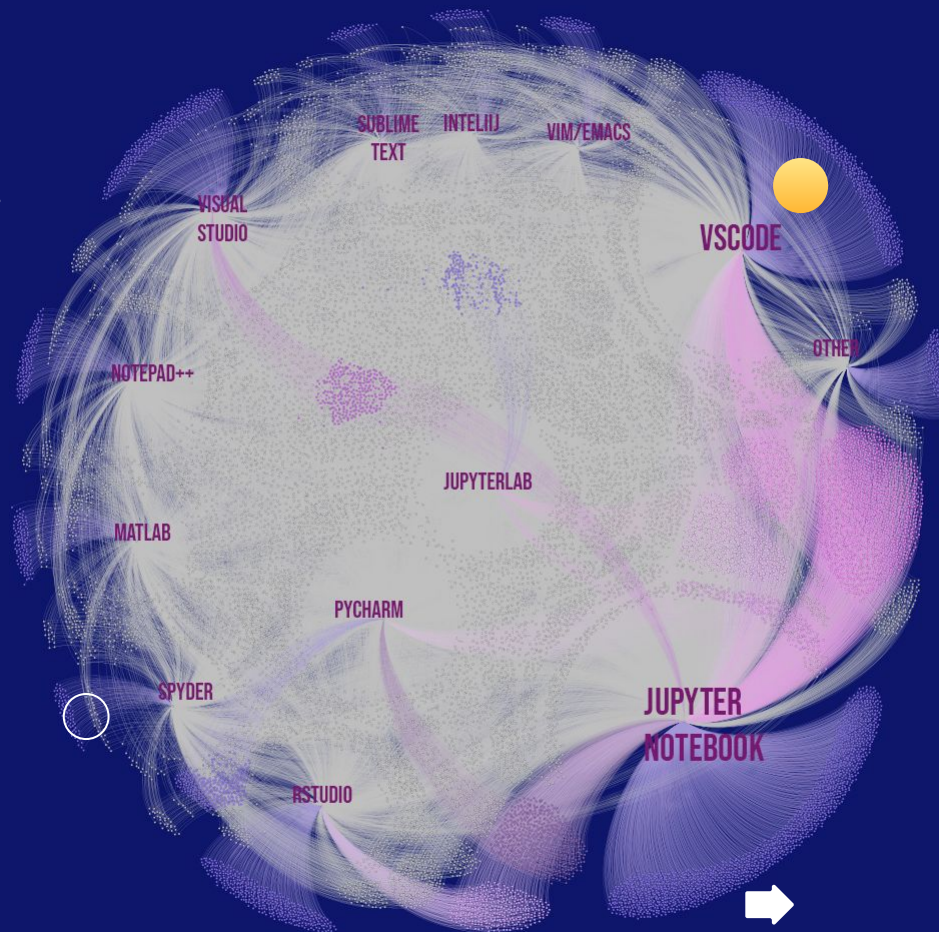


Măsoară numărul de conexiuni al unui mediu de dezvoltare, care sunt intermediare cu alte părți ale rețelei, care altfel ar fi fost deconectate



Cele mai multe conexiuni

- Jupyter Notebook
- VSCODE
- Visual Code



STUDIO

VSCODE

OTHER

NOTEPAD++

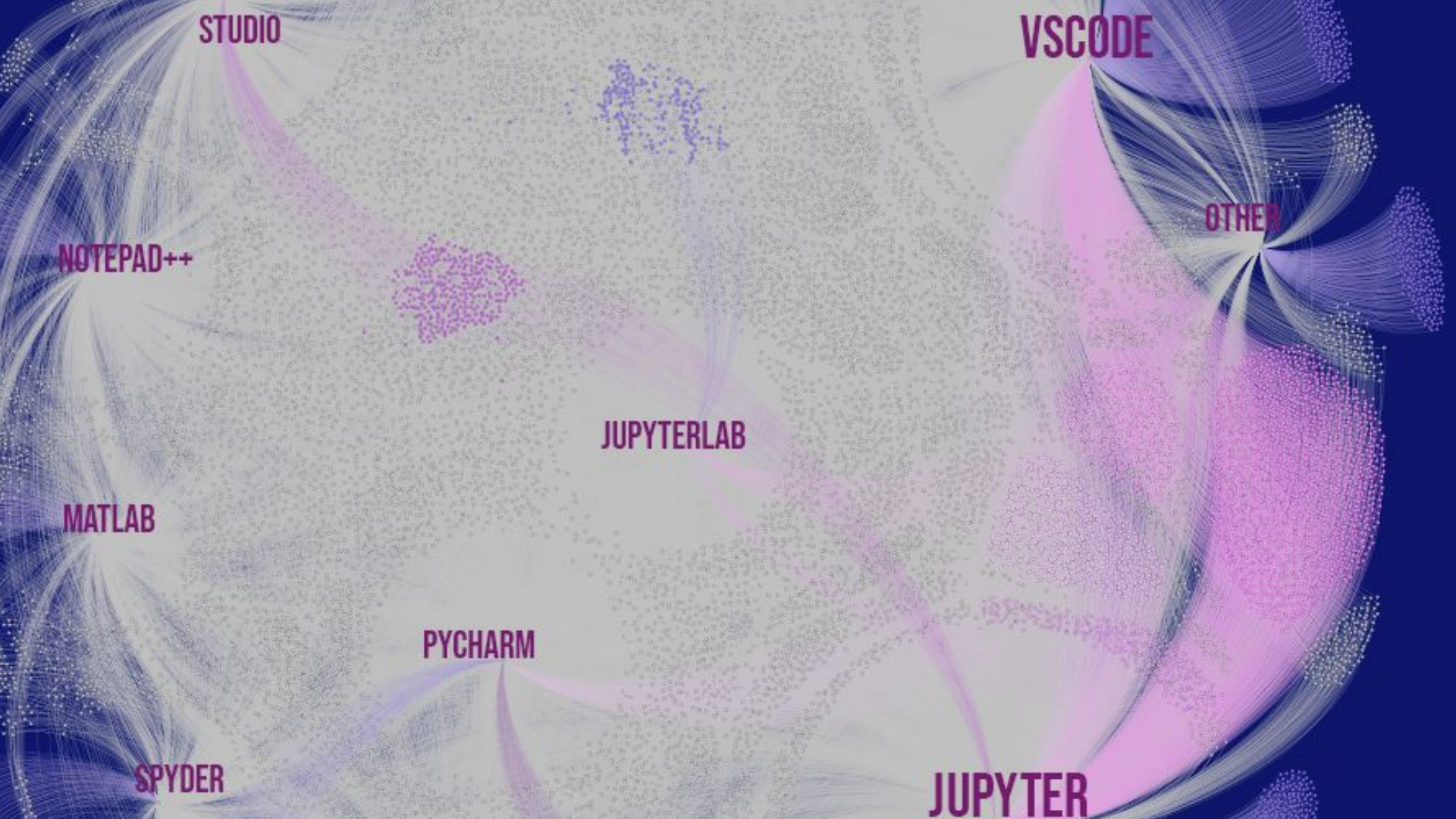
JUPYTERLAB

MATLAB

PYCHARM

SPYDER

JUPYTER



[MENU](#)[ANALYSIS](#)[CONTACT](#)[NETWORK ANALYSIS](#)

CLOSENESS

CENTRALITY

Măsoară cât de aproape este un nod față de celelalte noduri din rețea. Asta înseamnă că Jupyter Notebook este cel mai ușor de conectat cu alt mediu de dezvoltare.

Cel mai conectat nod.

- **Jupyter Notebook**





IMPORTANȚA ȘTIINȚEI REȚELELOR

Adunarea de acest tip de date și analizarea, ne ajută pentru a înțelege mai bine sisteme și industrii. Conexiunile pe care le putem face într-o rețea, ne pot oferi foarte multe răspunsuri. În cazul rețelei mele de date pot face o analiză unde să spun:

- Pe ce tip de tehnologie e recomandat să lucrezi, cât de apropiate sunt tehnologiile între ele



RECOMANDĂRI



Cred că e nevoie ca pe viitor, pentru un asemenea tip de bază de date, să se culeagă date din mai multe surse, pentru răspunsuri cât mai sigure.

Eșantionarea persoanelor să fie mai bine făcută, pentru că de exemplu. Foarte multe persoane sunt bărbați din India.

Existența unor întrebări de natură personală.

