# 2.1

a)  
  
Observăm diferențe semnificative în aranjarea cuvintelor între cele două imagini. Spre exemplu, cuvântul „mysterious” este situat în partea dreaptă sus în prima imagine, în timp ce în a doua apare în partea stângă, semn că distribuția s-a modificat considerabil.  
  
Totuși, cele două vizualizări au și elemente comune. Conțin aceleași cuvinte precum „book”, „story”, „movie” și altele, iar termeni similari ca sens – de exemplu „huge”, „large” și „massive” – sunt grupați în ambele cazuri. Asta demonstrează că ambele metode surprind într-o anumită măsură relațiile semantice dintre cuvinte.  
  
Aceste diferențe apar deoarece GloVe utilizează o metodă mai avansată de analiză, integrând atât statistici globale ale co-aparițiilor, cât și contextul local. În plus, metodele de reducere a dimensiunii pot produce reprezentări 2D diferite în funcție de tehnica folosită.  
  
GloVe nu se bazează doar pe numărul de apariții, ci învață reprezentări vectoriale ce reflectă importanța relațiilor dintre cuvinte. Mai mult, este antrenat pe corpuri de text foarte mari și beneficiază de ajustări fine ale parametrilor, în contrast cu o simplă matrice de co-apariție, care poate proveni dintr-un text restrâns și neprocesat.

b)  
  
Principalele diferențe dintre reprezentarea GloVe și cea derivată dintr-o matrice de co-apariție se explică prin câteva aspecte fundamentale:  
  
- Metoda de învățare: GloVe merge dincolo de simplele frecvențe de co-apariție, oferind o abordare complexă care captează atât tendințele globale cât și informația contextuală.  
- Obiectivul algoritmului: GloVe folosește o funcție de optimizare ponderată care favorizează relațiile relevante dintre cuvinte, în timp ce o matrice brută tratează toate co-aparițiile la fel.  
- Reducerea dimensională: Proiecțiile 2D sunt generate prin metode precum PCA, t-SNE sau UMAP, fiecare evidențiind alte tipare.  
- Calitatea datelor: GloVe e antrenat pe corpusuri vaste și bine pregătite, în timp ce matricea simplă poate proveni din texte limitate, fără procesări suplimentare.  
- Normalizare: Vectorii GloVe sunt ponderați inteligent pentru a echilibra cuvintele frecvente și cele rare, oferind astfel o reprezentare semantică mai echilibrată.

# 2.5

Codul prezentat folosește aritmetica vectorială pentru a rezolva o analogie semantică. Formula utilizată în fundal este:  
  
x = g - m + w  
  
unde:  
- m reprezintă vectorul pentru „man”  
- g este vectorul pentru „grandfather”  
- w este vectorul pentru „woman”  
- x este vectorul care corespunde răspunsului, probabil „grandmother”  
  
Intuiția din spatele acestei formule este următoarea: plecăm de la „man” și aplicăm transformarea care îl duce la „grandfather”. Apoi, aplicăm aceeași transformare asupra „woman” pentru a obține rezultatul așteptat.  
  
Această abordare valorifică o proprietate importantă a reprezentărilor vectoriale ale cuvintelor – aceea că relațiile semantice pot fi surprinse prin operații matematice simple. Într-o reprezentare bidimensională, vectorul de la „man” la „grandfather” simbolizează o relație familială specifică, pe care o replicăm pentru „woman” pentru a ajunge la „grandmother”.

# 2.9

Vectorii de cuvinte pot încorpora părtiniri (bias) din cauza tiparelor statistice prezente în datele pe care sunt antrenați. Deoarece modelele precum GloVe sau Word2Vec sunt instruite pe cantități uriașe de text scris de oameni (articole, cărți, pagini web), ele reflectă inevitabil stereotipurile și inegalitățile sociale existente în acele texte.  
  
Un exemplu notabil provine dintr-un studiu realizat în 2016 de Bolukbasi et al., intitulat „Man is to Computer Programmer as Woman is to Homemaker”. Autorii au descoperit că embedding-urile antrenate pe corpusul Google News evidențiau puternice stereotipuri de gen. De exemplu, analogia „man : computer\_programmer :: woman : x” returna „homemaker” pentru x, semnalând o asociere învățată din texte în care bărbații apăreau mai frecvent în contexte tehnice, iar femeile în roluri domestice.  
  
Aceste tendințe nu sunt introduse manual, ci apar ca rezultat al modelării statistice automate asupra datelor.  
  
Reducerea părtinirii prin post-procesare  
  
O metodă eficientă de diminuare a acestor părtiniri în vectorii pre-antrenați constă în identificarea direcțiilor din spațiul vectorial care encodează aceste asocieri (de gen, rasă, vârstă etc.) și în neutralizarea lor. Bazându-se pe același studiu, cercetătorii au propus o tehnică practică de „debiasing”.  
  
Aceștia au determinat direcția de gen prin diferența dintre perechi de cuvinte de gen opus (precum „he-she”, „man-woman”), apoi au ajustat vectorii cuvintelor neutre (ex: profesii), eliminând componentele de gen, dar menținând aceste diferențe în cuvintele inerent genizate.  
  
După aplicarea acestei metode, analogii precum „man : doctor :: woman : nurse” erau corectate pentru a produce rezultate mai echilibrate. Microsoft a implementat tehnici similare pentru a asigura neutralitatea în căutări și în procesele lor de procesare a textului.