Bridging Language and Data: Transforming Agricultural Curricula for Big Data Analytics Through Linguistic Insights

Zhihong Xu[1]*, Jaehyun Ahn[2], Shuai Ma[1], Anjorin Ezekiel Adeyemi[3], Fahmida Husain Choudhury[1], Xiting Zhuang[4], Rafael Landaverde, Gary J. Wingenbach

1. Texas A&M University; 2. University of Florida; 3. Fort Hays State University 4. North Dakota State University

Introduction

- Data-driven future
  - The Food, Agriculture, Natural Resources, and Human Sciences (FANH) discipline rely on advanced analytics.
- Skills gap and curriculum need
  - Workforce lacks strong data analytics training.
  - Embed analytics into ag education for better outcomes.
- Learner diversity
  - Students vary by demographics, experience, and usage

Research Questions

RQ1: Are there different groups between the users?
RQ2: How do overall word usage patterns differ between Groups?
RQ3: Are there notable patterns between groups subsets regarding tool and resource usage, activities and topics, additional knowledge and skills for curriculum development?

Literature Review

- Power of data analytics:
  💡 Enhances decision-making, resource use, and innovation (Marvin et al., 2017)
- Skills gap:
  ⚠️ Many graduates lack statistical, visualization, and advanced analytical skills (Kumar & Singh, 2019).
- Curriculum gap:
  🧱 Traditional ag education emphasizes science foundations but often omits data analytics (Braganza et al., 2017)

Methods

- Participants
  👩‍🎓👨‍🎓: Texas A&M alumni (n = 535 valid responses).
- Survey design
  📝: identified 8 core data analytics knowledge/skill areas (Ariyachandra, 2020; Guttmann, 2018; Jones, 2020; World Economic Forum, 2018; Zin et al., 2022).
- Data collection
  Demographics (educational level, gender, age, work experience, workplace data analytics usage).
  Open-ended responses (tools/resources, competencies, additional skills).
- Data Analysis
  Cluster analysis (demographics → learner groups).
  Text analysis (linguistic differences across groups).
- Software tools 💻: Python (clustering) + RStudio (text analysis).

TEXAS A&M UNIVERSITY
Agricultural Leadership, Education & Communications

Results

❖ Clustering results

Group 0 → Older, less formal education, more work experience, higher usage of big data & ML focus.

Group 1 → Younger, stronger in foundational concepts, data infrastructure, visualization, & advanced analytics.

❖ Word usage patterns

Group 0: "Platforms, formulas, business, training, roles"
→ value foundational learning & business applications.

Group 1: "Analysis, Excel, Python, PowerBI, Code"
→ strong emphasis on software/tools & technical proficiency.

❖ Subgroup differences 🔍

1. Tools/resources

Group 0 → leadership & management, YouTube (older users).

Group 1 → technical tools (Excel, Python, PowerBI).

2. Curriculum preferences

Group 0 → foundational/abstract elements.

Group 1 → big data & technical skills.

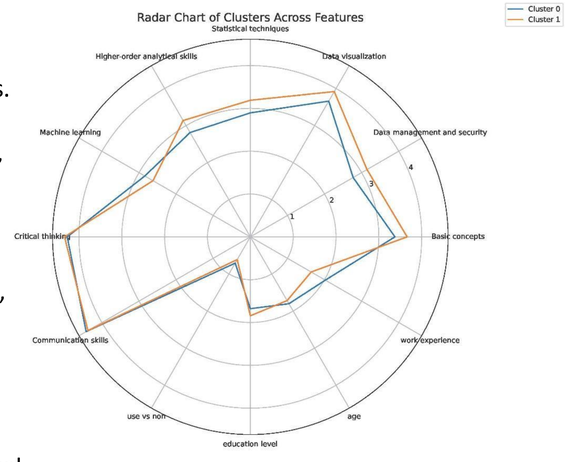

Figure 3. Radar chart of cluster characteristics

Figure 4. Word Cloud (B0 vs. B1)

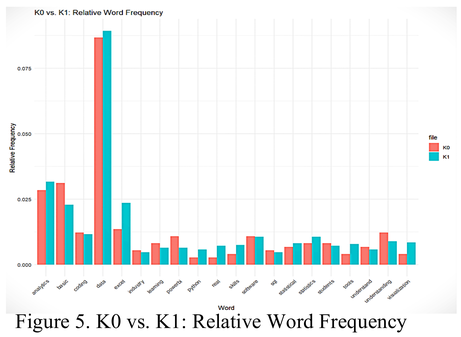

Figure 5. K0 vs. K1: Relative Word Frequency

TEXAS A&M UNIVERSITY

D-Lab in Agriculture
Email: d_lab@ag.tamu.edu

Results

3. Supplemental skills

Group 0 → communication & teamwork.

Group 1 → analysis & research skills.

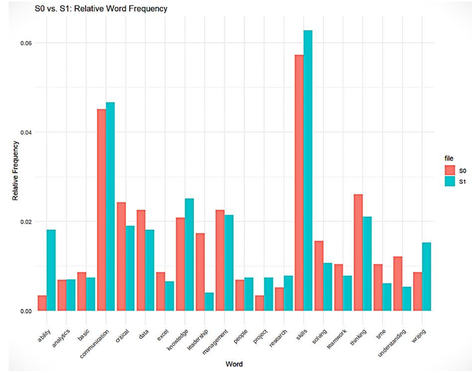

Figure 6. S0 vs. S1: Relative Word Frequency

Implications

- **Industry relevance**
  Align with emerging needs
- **Curriculum design**
  Differentiated & tailored curriculum
- **Future research**
  Qualitative approaches
- **Broader application**
  Extend to other disciplines

References