# Sentiment analysis on Students Expectations about College Placements

**Submitted by** 

HARISH G D

2327422

# SUBJECT: TEXT AND SOCIAL MEDIA ANALYTICS

**ETE Assignment** 

Guided by Prof. ROSEWINE JOY



CHRIST SCHOOL OF BUSINESS AND MANAGEMENT
BENGALURU
DECEMBER-2024

#### INTRODUCTION

This project explores **students' perspectives** on college placement processes, analyzing their **expectations, concerns, and sentiments** through data-driven insights. The goal is to access the:

#### 1. Preparation:

- To evaluate students' readiness for placements.
- To assess their self-perception of skills and the effectiveness of collegeorganized initiatives.
- o To analyze the accessibility of resources provided for placement preparation.

# 2. Support:

- o To understand the type and level of support received from the institution, faculty, and peers.
- o To identify gaps in support that hinder placement navigation.

# 3. Career Aspirations:

- o To measure the alignment of placement opportunities with students' personal and professional goals.
- o To evaluate students' satisfaction with the available placement options.

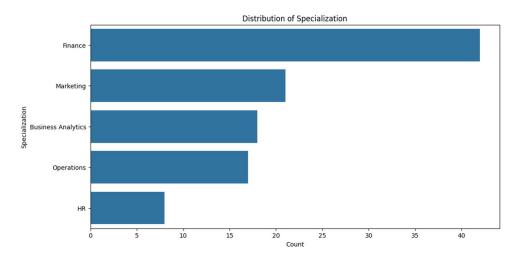
# **Key Questions Explored in the Survey:**

- 1. What are your primary expectations from the college placement process?
- 2. How well-prepared do you feel for the placement process?
- 3. What kind of support or resources from the college would enhance your placement readiness?
- 4. What are your major concerns or challenges regarding the placement process?
- 5. How well do you perceive the alignment between placement opportunities and your career goals?
- 6. What innovative ideas or practices would you recommend to make the placement process more student-centered?

#### **OBJECTIVES:**

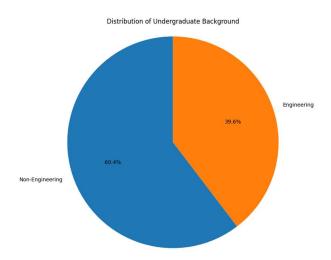
- 1. To identify key expectations students have from college placement processes.
- 2. To analyze the main concerns students experience during placements.
- 3. To assess sentiment distribution (positive, neutral, negative) across feedback.
- 4. To provide actionable insights and recommendations for improving placement strategies.

#### **EDA**



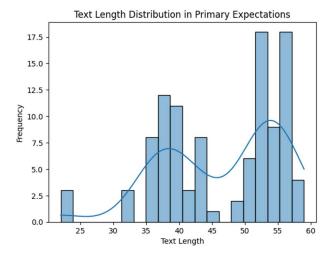
#### **Overall Observations**

- 1. **Dominance of Finance:** Finance is the most popular specialization with a count of around 42.
- 2. **Marketing and Business Analytics:** Both Marketing and Business Analytics have a similar number of respondents, with Marketing slightly ahead.
- 3. **Operations and HR:** Operations and HR have a lower number of respondents compared to the other specializations.



# **Overall Observations**

- 1. **Non-Engineering Dominance:** Non-engineering backgrounds make up the majority with 60.4% of the students.
- 2. **Engineering Representation:** Engineering backgrounds account for 39.6% of the students.



Two Distinct Groups: The bimodal nature might indicate that there are two different types of primary expectations, each with its own typical text length. For example, one group might be concise and direct, while the other might be more detailed and elaborate.

Right Skewness: The right skewness suggests that a majority of the expectations are shorter in length, while a smaller proportion are longer. This could be due to various reasons, such as time constraints, writing style, or the nature of the expectations themselves.

#### **METHODOLOGY:**

# PREPROCESSING:

1. Converting each column into a document, Hence we will have 6 documents for Topic modelling

```
Document for question: Primary Expectations
Content: Increased focus on niche roles in emerging industries More guidance on industry-specific interview preparation Higher CTC offers from ro
Document for question: Preparedness
Content: Partially ready but struggling with aptitude tests Prepared but could use more practice sessions Feeling ready but need more company-spond
Document for question: Support Needed
Content: Hands-on workshops for technical and non-technical skills Mock interviews with feedback from professionals Regular updates on placement
Document for question: Concerns
Content: Roles offered do not align with long-term career goals High level of competition during placement drives High level of competition durin
Document for question: Alignment with Career Goals
Content: Not aligned due to lack of opportunities in my field Neutral, as I am still exploring different career paths Not aligned due to lack of
Document for question: Innovative Ideas
Content: Expand outreach to startups and mid-sized firms Provide real-time feedback on interview performance Provide real-time feedback on interv
```

2. Converting text from upper case to lower case



3. Counting the No of characters of each document before the further analysis

```
Document for question: Primary Expectations
Length: 5030 characters

Document for question: Preparedness
Length: 5595 characters

Document for question: Support Needed
Length: 5505 characters

Document for question: Concerns
Length: 5344 characters

Document for question: Alignment with Career Goals
Length: 6080 characters

Document for question: Innovative Ideas
Length: 5633 characters
```

4. Performing RegEx Tokenization

```
Tokens for question: Primary Expectations Sample tokens: ['increased', 'focus', 'on', 'niche', 'roles', 'in', 'emerging', 'industries', 'more', 'guidance', 'on', 'industry', 'specific', Tokens for question: Preparedness Sample tokens: ['partially', 'ready', 'but', 'struggling', 'with', 'aptitude', 'tests', 'prepared', 'but', 'could', 'use', 'more', 'practice', Tokens for question: Support Needed Sample tokens: ['hands', 'on', 'workshops', 'for', 'technical', 'and', 'non', 'technical', 'skills', 'mock', 'interviews', 'with', 'feedback', Tokens for question: Concerns Sample tokens: ['roles', 'offered', 'do', 'not', 'align', 'with', 'long', 'term', 'career', 'goals', 'high', 'level', 'of', 'competition', 'dur. Tokens for question: Alignment with Career Goals Sample tokens: ['not', 'aligned', 'due', 'to', 'lack', 'of', 'opportunities', 'in', 'my', 'field', 'neutral', 'as', 'i', 'am', 'still', 'explor Tokens for question: Innovative Ideas Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'outreach', 'to', 'startups', 'and', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'on', 'interview Sample tokens: ['expand', 'to', 'to'
```

5. Importing NLTK for stop word removal to enhance the process time and the lemmatization is been done to bring the word's root format before analysing the frequency of the words in the documents.

```
Cleaned tokens for question: Primary Expectations
Sample tokens: ['increased', 'focus', 'niche', 'roles', 'emerging', 'industries', 'guidance', 'industry', 'specific', 'interview', 'preparation'
Cleaned tokens for question: Preparedness
Sample tokens: ['partially', 'ready', 'struggling', 'aptitude', 'tests', 'prepared', 'could', 'use', 'practice', 'sessions', 'feeling', 'ready',
Cleaned tokens for question: Support Needed
Sample tokens: ['hands', 'workshops', 'technical', 'non', 'technical', 'skills', 'mock', 'interviews', 'feedback', 'professionals', 'regular', '
Cleaned tokens for question: Concerns
Sample tokens: ['roles', 'offered', 'align', 'long', 'term', 'career', 'goals', 'high', 'level', 'competition', 'placement', 'drives', 'high', '
Cleaned tokens for question: Alignment with Career Goals
Sample tokens: ['aligned', 'due', 'lack', 'opportunities', 'field', 'neutral', 'still', 'exploring', 'different', 'career', 'paths', 'aligned',
Cleaned tokens for question: Innovative Ideas
Sample tokens: ['expand', 'outreach', 'startups', 'mid', 'sized', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'interview', 'performance', 'paths', 'goals', 'goals', 'goals', 'firms', 'provide', 'real', 'time', 'feedback', 'goals', 'go
```

# 6. Performing word cloud for the documents based on the frequency of the words:

# **Word Cloud Document 1: Primary Expectations:**



- 1. **Emphasis on Opportunities:** Words like "opportunities," "emerging industries," and "international" are prominently displayed, suggesting that students are looking for exposure to diverse career paths and global opportunities.
- 2. **Focus on Guidance and Support:** Words like "guidance," "networking sessions," and "recruiter" indicate that students value mentorship and support from industry professionals and alumni.
- 3. **Desire for Quality Placements:** Words like "reputed firms," "higher offers," and "better placement" highlight students' expectations of securing high-quality placements with reputable companies.

# **Word Cloud 2 (Preparedness)**



- Confidence and Preparedness: Words like "confident," "ready," and "prepared" are
  prominent, suggesting that students generally feel positive about their placement
  readiness.
- Focus on Technical Skills: Words like "technical skills," "aptitude," and "tests" are highlighted, indicating that technical skills and aptitude are perceived as crucial for placements.

• Concerns and Uncertainties: Words like "unprepared," "struggling," and "unclear" are also present, suggesting that some students may have concerns or uncertainties about specific aspects of the placement process.

# **Word Cloud 3 (Support Needed):**



- 1. **Technical Skills Emphasis:** Words like "technical skills," "technical," and "aptitude tests" are prominent, indicating that technical skills are a major focus for students.
- 2. Workshops and Preparation: Words like "workshops," "preparation," and "placement preparation" suggest that students value hands-on training and preparation for placements. Words like "industry," "guidance," and "mentoring programs" indicate that students seek exposure to industry trends and guidance from professionals.
- 3. **Access and Resources:** Words like "access," "resources," and "online resources" highlight the importance of access to resources and tools for placement preparation.

#### Word Cloud 4 (Concerns):



- 1. **Preparation and Competition:** Words like "preparation," "competition," and "challenges" are prominent, indicating that students perceive the placement process as competitive and demanding significant preparation.
- 2. **Industry and Domain:** Words like "industry trends," "domain," and "specific roles" suggest that students are aware of industry trends and are seeking placements that align with their specific domain knowledge and career goals.
- 3. **Company and Role-Specific:** Words like "company specific," "roles offered," and "company selection" highlight the importance of targeting specific companies and roles that align with their interests and aspirations.
- 4. **Uncertainty and Guidance:** Words like "uncertainty," "limited guidance," and "eligibility criteria" indicate that students may feel uncertain about the placement process and seek guidance and clarity regarding eligibility criteria and selection processes.

# **Word Cloud 5 (Alignment with Carrer Goals):**



- 1. **Exploration and Alignment:** Words like "exploring," "aligned," and "match" are prominent, suggesting that students are actively exploring career paths and seeking alignment between their interests and available opportunities.
- 2. Career Paths and Industries: Words like "career paths," "industries," and "domain" indicate that students are considering various career paths and industries.
- 3. **Niche and Specific Roles:** Words like "niche," "specific roles," and "match interests" suggest that students are looking for specific roles that align with their interests and skills.
- 4. **Uncertainty and Misalignment:** Words like "limited," "slightly misaligned," and "neutral" indicate that some students may feel uncertain about their career paths or may perceive a mismatch between their interests and available opportunities.

# **Word Cloud 6 (Innovative Suggestions):**



- 1. **Mentorship and Guidance:** Words like "mentorship," "guidance," and "recruiters" are prominent, indicating that students value mentorship and guidance from industry professionals and alumni.
- 2. **Career Development:** Words like "career programs," "career development," and "skill development" suggest that students are focused on developing their skills and preparing for their careers.
- 3. **Industry Exposure and Networking:** Words like "industry," "companies," and "alumni networks" indicate that students seek exposure to industry trends and networking opportunities with professionals.
- 4. **Experiential Learning:** Words like "hackathons," "simulations," and "competitions" suggest that students value experiential learning opportunities to develop practical skills and gain real-world experience.

# Word cloud of entire responses

# Word Cloud For The Entire Reviews | Companies | Compa

#### **Sentiment Analysis:**

The inputs gathered from students across various questions have been combined to gain a comprehensive understanding of their perspectives. This collective analysis focuses on key aspects such as preparation for placements, institutional support, alignment with career aspirations, and the challenges faced during the placement process. By synthesizing these responses, we aim to uncover common concerns and experiences, providing valuable insights into student sentiment and areas requiring improvement in the placement process.

# **VADER (Rule Based Sentiment Analysis)**

The sentiment analysis reveals that the responses have a predominantly neutral tone (86.1%), indicating that most individuals shared descriptive or factual inputs without strong emotional expressions. However, the positive sentiment (5%) is significant and much higher than the negative sentiment (8%), suggesting an optimistic or favorable outlook in the responses.

The compound score of-0.2263, being highly Negative, reinforces this overall Negative sentiment, reflecting an optimistic perception despite any concerns or challenges mentioned in the feedback. This indicates that while respondents acknowledge issues, they may also see potential solutions or maintain a constructive perspective

#### Polarity score of each response:

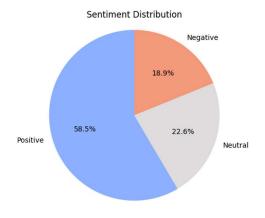
	ID	neg	neu	pos	compound	cleaned_text
0	0	0.116	0.544	0.340	0.8360	increased focus niche role emerging industry p
1	1	0.086	0.856	0.058	-0.2263	guidance industry specific interview preparati
2	2	0.064	0.748	0.188	0.5423	higher offer reputed firm feeling ready need c
3	3	0.180	0.526	0.294	0.6957	opportunity international company fully confid
4	4	0.000	0.638	0.362	0.9231	better placement opportunity tier company well
101	101	0.065	0.726	0.208	0.7096	pathway internship leading full time role conf
102	102	0.000	0.686	0.314	0.8225	workshop specialized industry skill confident
103	103	0.051	0.678	0.271	0.7964	better recruiter college coordination technica
104	104	0.056	0.538	0.406	0.9231	improved recruiter communication strategy conf
105	105	0.000	0.912	0.088	0.4215	timely update employer expectation gaining pre

# Finding out the sentiments for each response

	cleaned_text	ID	Sentiment	
10	networking session alumnus industry expert unp	10	Positive	
11	guidance industry specific interview preparati	11	Neutral	
12	opportunity align closely career aspiration un	12	Neutral	
13	networking session alumnus industry expert fee	13	Negative	
14	transparency placement process moderately prep	14	Neutral	
101	pathway internship leading full time role conf	101	Positive	
102	workshop specialized industry skill confident	102	Positive	
103	better recruiter college coordination technica	103	Positive	
104	improved recruiter communication strategy conf	104	Positive	
105	timely update employer expectation gaining pre	105	Neutral	

The responses highlight significant concerns among students about the college placement process, including challenges with preparation, alignment of opportunities with career aspirations, and insufficient support. Despite these concerns, many students share constructive suggestions, such as enhanced skill-building programs, better career guidance, and innovative placement practices. The feedback underscores the need for targeted interventions to address these challenges and improve the overall placement experience. Leveraging these insights can help create a more student-centered, efficient, and supportive placement process, fostering greater satisfaction and better outcomes.

#### **Distribution of Sentiments**



- **Overall Positive Sentiment:** The dominance of positive sentiment suggests that the overall tone of the data is generally positive.
- **Mixed Sentiments:** The presence of negative and neutral sentiments indicates that there is a range of opinions or perspectives within the data..

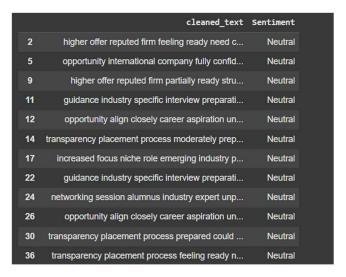
# **Printing Positive sentiments**



#### Printing the negative sentence:

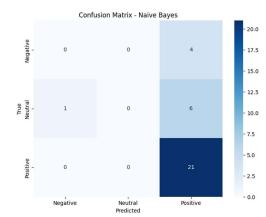


#### **Printing the neutral sentences:**



# Performing Statistic based algorithm (Naïve Bayes)

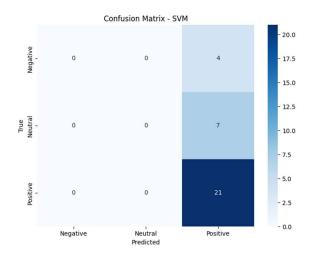
Accuracy: 0.66							
Classification Report:							
	precision	recall	f1-score	support			
Negative	0.00	0.00	0.00	4			
Neutral	0.00	0.00	0.00	7			
Positive	0.68	1.00	0.81	21			
accuracy			0.66	32			
,	A 22	0.33	0.00	32			
macro avg	0.23						
weighted avg	0.44	0.66	0.53	32			



The Naive Bayes model achieved an accuracy of 66%, primarily performing well in predicting the **Positive class** (Precision: 68%, Recall: 100%). However, it failed to classify the **Negative** and **Neutral** classes correctly, resulting in **0% precision, recall, and F1-score** for these categories. The model demonstrates a strong bias toward the Positive class, likely due to class imbalance or insufficient representation of Negative and Neutral samples.

# Performing Statistic based algorithm (SVM)

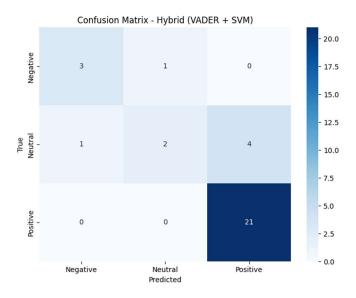
Accuracy: 0.66							
Classification Report: precision recall f1-score support							
200 200	precision	Lecari	11-20016	support			
Negative	0.00	0.00	0.00	4			
Neutral	0.00	0.00	0.00	7			
Positive	0.66	1.00	0.79	21			
accuracy			0.66	32			
macro avg	0.22	0.33	0.26	32			
weighted avg	0.43	0.66	0.52	32			



The SVM model achieved an accuracy of 66%, performing well in predicting the Positive class (Precision: 66%, Recall: 100%). However, it failed to classify any instances of the Negative and Neutral classes, resulting in 0% precision, recall, and F1-score for these categories. The confusion matrix confirms a strong bias toward the Positive class, likely due to class imbalance. To improve, techniques like class weighting, resampling, or hyperparameter tuning should be considered.

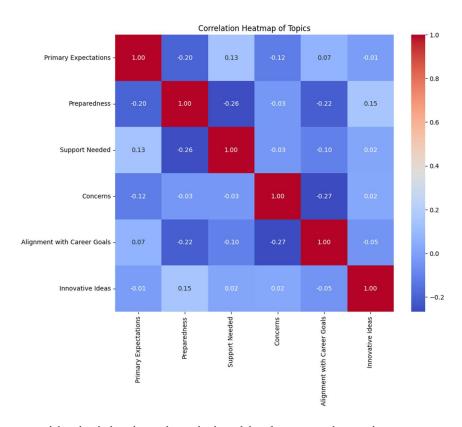
# Performing Hybrid (VADER + SVM) Model:

Accuracy: 0.81					
Hybrid Model Classification Report:					
precision	recall	f1-score	support		
0.75	0.75	0.75	4		
0.67	0.29	0.40	7		
0.84	1.00	0.91	21		
			20 MF 400 A		
		0.81	32		
0.75	0.68	0.69	32		
0.79	0.81	0.78	32		
	Classification precision  0.75  0.67  0.84	Classification Report precision recall  0.75 0.75 0.67 0.29 0.84 1.00 0.75 0.68	Classification Report: precision recall f1-score  0.75 0.75 0.75  0.67 0.29 0.40  0.84 1.00 0.91  0.81  0.75 0.68 0.69		



The **Hybrid Model (VADER + SVM)** achieved an accuracy of **81%**, significantly improving performance compared to individual models. It effectively predicted **Positive** sentiments (Precision: 84%, Recall: 100%, F1-Score: 0.91). The **Negative** class was well-classified with **75% precision and recall**. However, the model struggled with the **Neutral** class, achieving **67% precision** but only **29% recall**, indicating misclassifications. Overall, the hybrid model demonstrates strong performance, particularly for Positive and Negative sentiments, though further refinements could help improve Neutral class predictions.

# **Correlation Analysis of sentiments:**



The heatmap provides insights into the relationships between the topics:

# 1. Key Negative Correlations:

- Preparedness vs. Support Needed (-0.26):
  - o Individuals who are more prepared tend to need less support. This suggests that improving preparedness could reduce dependency on external help.
- Alignment with Career Goals vs. Concerns (-0.27):
  - When career goals are well-aligned, concerns tend to decrease. This indicates a strong link between personal alignment and reduced anxiety or doubts.

# 2. Key Positive Correlations:

• Preparedness vs. Innovative Ideas (0.15):

 Being prepared slightly enhances the ability to generate innovative ideas, hinting at a small but positive link between confidence and creativity.

# • Primary Expectations vs. Support Needed (0.13):

o Those with clearly defined expectations may require moderate levels of support, highlighting the role of clarity in shaping support needs.

#### 3. Weak or Insignificant Correlations:

Most relationships, such as Concerns vs. Innovative Ideas (0.02) or Primary
 Expectations vs. Alignment with Career Goals (0.07), show negligible interaction.

 This implies these variables operate independently in most contexts.

#### 4. Actionable Takeaways:

# • Focus on Preparedness:

Since preparedness reduces support needs and boosts innovation, organizations or individuals should prioritize improving readiness through training or resources.

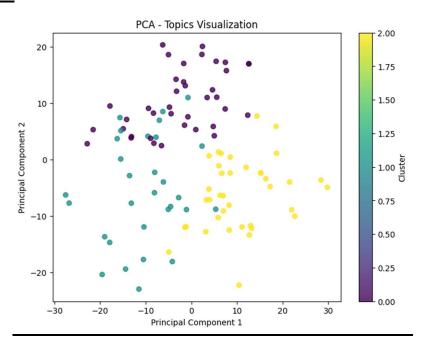
# Address Misalignment with Career Goals:

 Helping individuals align with their career goals can reduce concerns and improve overall satisfaction and focus.

# • Define Expectations Clearly:

 Clear expectations can guide support strategies, ensuring resources are allocated effectively.

#### **PCA Analysis:**



The PCA visualization and clustering provide valuable insights into the relationships among the topics derived from the correlation matrix:

#### 1. Identified Clusters

The topics group into three distinct clusters, reflecting patterns in their relationships:

# Cluster 1 (Yellow):

Topics in this cluster may be strongly associated with preparedness and innovative ideas, as these tend to align with proactive and positive characteristics.

Suggests a group of individuals or scenarios where forward-thinking and readiness dominate.

#### Cluster 2 (Teal):

Likely includes topics like support needed and possibly weaker correlations from the matrix. These may represent individuals or scenarios requiring assistance or experiencing variability in performance.

# Cluster 3 (Purple):

Appears to group topics like alignment with career goals and concerns, which are more inward-focused. These may reflect personal or subjective attributes such as satisfaction or anxieties.

# 2. Cluster Separation

The distance between clusters indicates distinct patterns:

- Cluster 1 (Yellow) is relatively distant, suggesting topics within this group are highly self-sufficient and outwardly driven.
- Cluster 2 (Teal) overlaps with both clusters slightly, indicating topics here are transitional or influence multiple groups.
- Cluster 3 (Purple) is more concentrated, showing closely related topics around alignment and concerns.

#### 3. Key Relationships

The clustering highlights:

- Preparedness and innovation are strongly linked and independent of concerns.
- Concerns and career alignment may oppose topics like preparedness, as seen in the correlation heatmap.
- Support needed forms a middle ground, interacting with multiple topics.

#### 4. Practical Implications

Cluster-Specific Strategies:

• For Cluster 1 (Preparedness & Innovation): Promote innovation workshops or proactive strategies to build on strengths.

- For Cluster 2 (Support Needed): Provide tailored support, such as mentoring or resources, to meet diverse needs.
- For Cluster 3 (Concerns & Career Alignment): Focus on career planning and addressing personal challenges to reduce concerns.

#### Conclusion

This analysis provides a detailed understanding of students' expectations, sentiments, and challenges regarding college placements. Key findings include:

- 1. **Positive Sentiment Dominance:** Students exhibit an optimistic outlook toward placements, with concerns mainly stemming from preparation gaps, alignment with career goals, and insufficient support.
- 2. Clusters of Interest: PCA reveals three clusters, emphasizing preparedness and innovation, support needs, and career alignment with concerns, each with distinct characteristics.
- 3. **Correlations in Sentiment:** Preparedness reduces support needs and promotes innovation, while career misalignment heightens concerns.
- 4. **Feedback on Improvements:** Students value experiential learning, mentorship, and industry exposure, with emphasis on career-aligned opportunities and practical skill-building.

# **Business Implications**

#### 1. Enhancing Preparedness:

- Action: Invest in technical training, mock interviews, and industry-aligned skill workshops to boost students' confidence and reduce dependency on external support.
- o **Impact:** Better-prepared students will approach placements with greater confidence, increasing placement rates and satisfaction.

# 2. Targeted Support Programs:

- o **Action:** Provide tailored mentoring, peer learning sessions, and real-world project opportunities for students in Cluster 2 (Support Needed).
- o **Impact:** Personalized support can improve placement readiness and bridge gaps in skills and confidence.

#### 3. Alignment with Career Goals:

- Action: Offer personalized career counseling and sessions with industry professionals to help students identify and pursue roles aligned with their aspirations.
- o **Impact:** Reduced student concerns and increased satisfaction through clearer career paths and role alignment.

# 4. Focus on Experiential Learning:

- Action: Introduce hackathons, competitions, and internship programs that mirror real-world challenges. Encourage active alumni involvement and company partnerships to expose students to emerging industries.
- o **Impact:** Practical learning experiences will enhance skills, foster innovation, and improve employability.

# 5. Strengthen Feedback Loops:

- **Action:** Establish regular channels for students to share concerns and innovative ideas regarding placement processes.
- o **Impact:** Dynamic feedback can guide resource allocation and strategy development, making the placement process more student-centric.

# 6. Data-Driven Decision-Making:

- o **Action:** Leverage insights from sentiment and cluster analyses to refine placement strategies and identify at-risk groups for targeted interventions.
- o **Impact:** Optimized placement processes will improve outcomes, strengthen institutional reputation, and attract better recruitment opportunities.