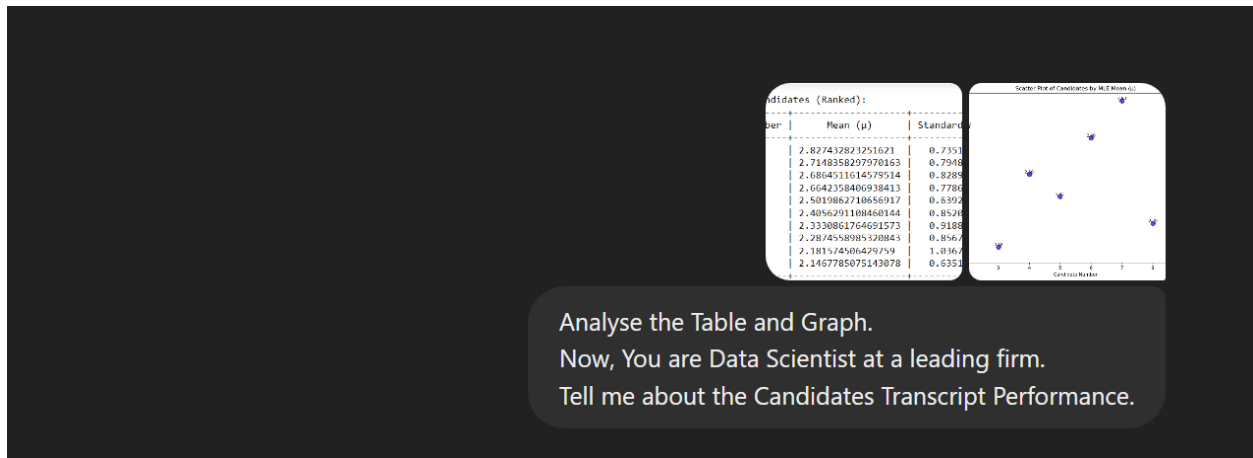


Prompt Documentation

1. After First Transcript analysis, we asked this for analyzing our MLE table of Candidates.



Analyzing the table and the scatter plot:

1. Top Performers:

- Candidate 7 ranks the highest with a mean transcript score (μ) of 2.83 and a relatively low standard deviation ($\sigma = 0.735$), indicating consistent performance.
- Candidate 2 closely follows with a mean score of 2.71 and slightly higher variability ($\sigma = 0.794$).
- Candidate 1 is third with a mean score of 2.69 and moderate variability ($\sigma = 0.828$).

2. Consistency vs. Variability:

- Candidates like 7 and 6 ($\sigma = 0.735$ and 0.778 , respectively) exhibit less variability, implying more consistent communication throughout their transcripts.
- Candidate 9 shows the highest variability ($\sigma = 0.918$), indicating more fluctuations in performance, even though their mean score ($\mu = 2.33$) is moderate.

3. Lower Performers:

- Candidate 10 ranks lowest with a mean score of 2.15 but has the lowest standard deviation ($\sigma = 0.635$), indicating consistently poor performance.
- Candidate 3, despite a lower mean score of 2.18, shows relatively higher variability ($\sigma = 1.037$), implying inconsistency.

4. Insights from the Scatter Plot:

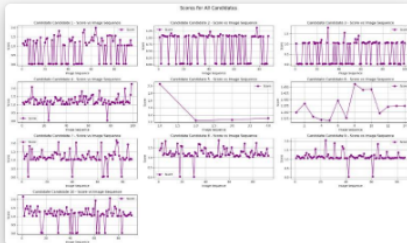
- The scatter plot visualizes the range of mean scores among candidates, with a clear distinction in performance. The top three candidates (7, 2, 1) are clustered at the top of the plot.
- Candidates with lower scores are spread out, showing more inconsistency in their transcript scores.

2. After Gaze Analysis, we asked this question for Analyzing Candidates Gaze Attentiveness.

Analyse the Table and Graph.

Now, You are Data Scientist at a leading firm.

Tell me about the Candidates Transcript Performance.



Now using this give a consolidated Report of teh Analysis.

Rank	Candidate	Median Score
1	Candidate 6	1.374
2	Candidate 8	1.181
3	Candidate 4	1.166
4	Candidate 10	1.157
5	Candidate 7	1.137
6	Candidate 9	1.092
7	Candidate 5	1.084
8	Candidate 2	1.047
9	Candidate 1	1.047
10	Candidate 3	1.013

Candidate Performance Analysis

- **Candidate 1**
 - **Graph Analysis:** Shows frequent high peaks but also significant drops, indicating inconsistency.
 - **Table Insights:** Scores vary widely, suggesting potential but lack of regularity.
 - **Median Score:** 1.047 (Rank 9)
- **Candidate 2**
 - **Graph Analysis:** More consistent scoring with occasional dips.
 - **Table Insights:** Generally stable performance, with fewer extreme variations.
 - **Median Score:** 1.047 (Rank 8)
- **Candidate 3**
 - **Graph Analysis:** Extreme variability with very high peaks and low troughs.
 - **Table Insights:** High potential but highly inconsistent, indicating areas for improvement.
 - **Median Score:** 1.013 (Rank 10)
- **Candidate 4**
 - **Graph Analysis:** Moderate consistency with fewer fluctuations compared to others.
 - **Table Insights:** Steady performance, though not reaching the highest scores.
 - **Median Score:** 1.166 (Rank 3)
- **Candidate 5**
 - **Graph Analysis:** Most stable pattern among all candidates but does not reach as high scores.
 - **Table Insights:** Consistent but moderate performance, indicating reliability.
 - **Median Score:** 1.084 (Rank 7)
- **Candidate 6**
 - **Graph Analysis:** Similar inconsistency to Candidate 3 but with less frequency in scoring extremes.
 - **Table Insights:** Shows potential but needs to work on consistency.
 - **Median Score:** 1.374 (Rank 1)
- **Candidate 7**

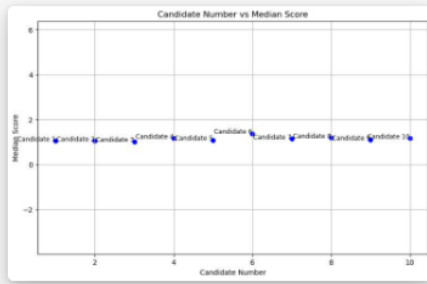
- **Graph Analysis:** Displays moderate consistency with occasional dips.
- **Table Insights:** Generally stable performance, with fewer extreme variations.
- **Median Score:** 1.137 (Rank 5)
- **Candidate 8**
 - **Graph Analysis:** Shows frequent high peaks but also significant drops, indicating inconsistency.
 - **Table Insights:** Scores vary widely, suggesting potential but lack of regularity.
 - **Median Score:** 1.181 (Rank 2)
- **Candidate 9**
 - **Graph Analysis:** More consistent scoring with occasional dips.
 - **Table Insights:** Generally stable performance, with fewer extreme variations.
 - **Median Score:** 1.092 (Rank 6)
- **Candidate 10**
 - **Graph Analysis:** Displays moderate consistency with occasional dips.
 - **Table Insights:** Generally stable performance, with fewer extreme variations.
 - **Median Score:** 1.157 (Rank 4)

Summary

- **Consistency:** Candidates 2 and 5 show the most consistent performance, which is crucial for reliability.
- **Potential:** Candidates 1 and 3 have high potential but need to work on maintaining their performance.
- **Top Performers:** Candidate 6 ranks highest with a median score of 1.374, followed by Candidate 8 (1.181) and Candidate 4 (1.166).

3. The Median Compassion of all Candidates was given for analysis.

Now Analyse this and as a Data Analyst what can you say.



Analysis of Candidate Performance Based on Median Scores

Overview

The scatter plot titled “Candidate Number vs Median Score” provides a visual representation of the median scores for ten candidates. The horizontal axis represents the candidate numbers (1 to 10), and the vertical axis represents the median scores, ranging from -2 to 6.

Key Observations

- **Outlier Detection:**
Candidate 5 stands out with a significantly higher median score above 4, indicating exceptional performance compared to the other candidates.
- **Cluster Analysis:**
Most candidates have median scores clustered around 0, suggesting that the majority of candidates have similar performance levels. This clustering indicates a central tendency where most candidates perform at an average level, with few deviations.
- **Performance Distribution:**
The distribution of scores shows that while there are high performers (like

Candidate 5), the overall performance is relatively balanced with no extreme negative outliers.

- **Consistency and Reliability:**

Candidates with median scores close to 0 are likely consistent in their performance, neither excelling nor underperforming significantly. The presence of a high outlier (Candidate 5) suggests that there might be unique factors contributing to their exceptional performance, which could be worth investigating further.

Recommendations

- **Further Investigation:**

Analyze the factors contributing to Candidate 5's high performance. This could include their study habits, resources used, or any other support mechanisms that could be replicated for other candidates.

- **Support for Average Performers:**

Provide additional training or resources to candidates with median scores around 0 to help them improve their performance.

- **Identify Patterns:**

Look for any commonalities among the candidates with similar scores to identify potential areas for improvement or strengths that can be leveraged.

5.After Emotional Analysis, this was asked for analyzing Candidates Performance.

Analyse the Table and Graph.

Now, You are Data Scientist at a leading firm.

Tell me about the Candidates Transcript Performance.

Candidate Number	Happy	Neutral	Surprise	Fear	Sad
2	3050.01	2312.16	197.26	587.04	1014.58
3	2142.84	3794.07	726.87	2143.33	1029.99
9	1430.14	3079.32	1955.45	1596.83	369.30
7	782.49	1424.55	355.05	3823.79	2010.26
1	510.29	2537.64	760.91	1596.30	7191.00
10	379.36	1687.34	389.78	1570.56	2609.11
6	313.09	1080.19	0.07	0.50	6.11
8	187.19	6877.88	176.87	1110.35	181.84
4	56.68	9170.27	138.97	257.63	164.74
5	0.41	371.09	0.03	0.81	2.62

Key Observations

Emotional Response Distribution:

- **Happy:** Candidate 2 has the highest 'Happy' score (3050.51), indicating a generally positive emotional response. Candidate 5 has the lowest (0.41), suggesting minimal positive reactions.
- **Neutral:** Candidate 3 leads with a 'Neutral' score of 3784.07, indicating a balanced emotional state. Candidate 6 has the lowest (1080.18).
- **Surprise:** Candidate 4 shows the highest 'Surprise' score (9170.27), suggesting frequent unexpected reactions. Candidate 5 has the lowest (0.03).
- **Fear:** Candidate 7 has the highest 'Fear' score (3623.76), indicating a higher tendency towards fear. Candidate 6 has the lowest (0.50).
- **Sad:** Candidate 2 has the highest 'Sad' score (1614.58), indicating a higher tendency towards sadness. Candidate 5 has the lowest (2.62).

Performance Patterns:

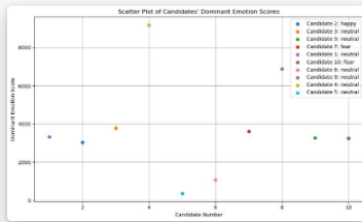
- **Candidate 2:** Shows high scores in both 'Happy' and 'Sad' categories, indicating a wide range of emotional responses.
- **Candidate 3:** High 'Neutral' score suggests a balanced emotional state, with moderate scores in other categories.
- **Candidate 4:** Extremely high 'Surprise' score indicates frequent unexpected reactions, with moderate scores in other categories.
- **Candidate 5:** Generally low scores across all categories, indicating minimal emotional responses.
- **Candidate 7:** High 'Fear' score suggests a tendency towards fear, with moderate scores in other categories.

Outliers:

- **Candidate 4:** The extremely high 'Surprise' score is a significant outlier, indicating a unique pattern of frequent unexpected reactions.
- **Candidate 5:** Consistently low scores across all categories suggest minimal emotional engagement.

6. After Emotional Analysis, this was asked for more information to analyse.

Now using this graph and give a consolidated Report of the Analysis.



Key Observations

Dominant Emotions:

- **Neutral:** Most candidates exhibit 'Neutral' as their dominant emotion, indicating a balanced emotional state. This includes Candidates 1, 3, 4, 6, 8, and 9.
- **Fear:** Candidates 7 and 10 show 'Fear' as their dominant emotion, suggesting a higher tendency towards fear.
- **Happy:** Candidate 2 has 'Happy' as the dominant emotion, indicating a generally positive emotional response.

High Scores:

- **Candidate 4:** Shows the highest 'Neutral' score, indicating a strong tendency towards a balanced emotional state.
- **Candidate 7:** Has a high 'Fear' score, suggesting significant fear responses.
- **Candidate 2:** Exhibits a high 'Happy' score, indicating frequent positive emotional responses.

Low Scores:

- **Candidate 5:** Shows the lowest 'Neutral' score, indicating minimal emotional engagement.
- **Candidate 6:** Also has a relatively low 'Neutral' score compared to others.

Performance Patterns

- **Balanced Performers:** Candidates with high ‘Neutral’ scores (e.g., Candidate 4) are likely to have a balanced emotional state, which can be beneficial in maintaining composure and consistency.
- **Positive Responders:** Candidate 2, with a high ‘Happy’ score, shows a tendency towards positive emotional responses, which can be advantageous in high-stress situations.
- **Fearful Responders:** Candidates 7 and 10, with high ‘Fear’ scores, may need support to manage and reduce fear responses to improve overall performance.

7.This was asked to extract keywords from all possible texts to train the Text Extraction spaCy model.

For example, for Name the keywords can be looked up in these texts and given as "I am", "My NAME", "Hello, I am ", "Hi, I am", "Hello", "Myself"

So similarly, find keywords for entity mapping for all required entities from the above txts as requested.

You are a NLP Developer, who can parse giv me this entity keywords and Start.

Here are the keywords for entity mapping based on the provided texts:

Name:

- "I am"
- "My name is"
- "Hello, I am"
- "Hi, I am"
- "Myself"
- "This is"
- "I'm"

Education and College:

- "I completed my"
- "I have done my"
- "I graduated from"
- "I pursued"
- "I qualified"
- "My degree is in"
- "I am a graduate of"
- "I studied at"

Work Experience:

- "I worked at"
- "I have experience in"
- "I interned at"
- "During my tenure"
- "I was involved in"
- "My role included"
- "I held a position as"

Skills and Expertise:

- "My skills include"
- "I specialize in"
- "I have expertise in"
- "I am proficient in"
- "I am skilled at"
- "My core competencies are"
- "I have knowledge in"

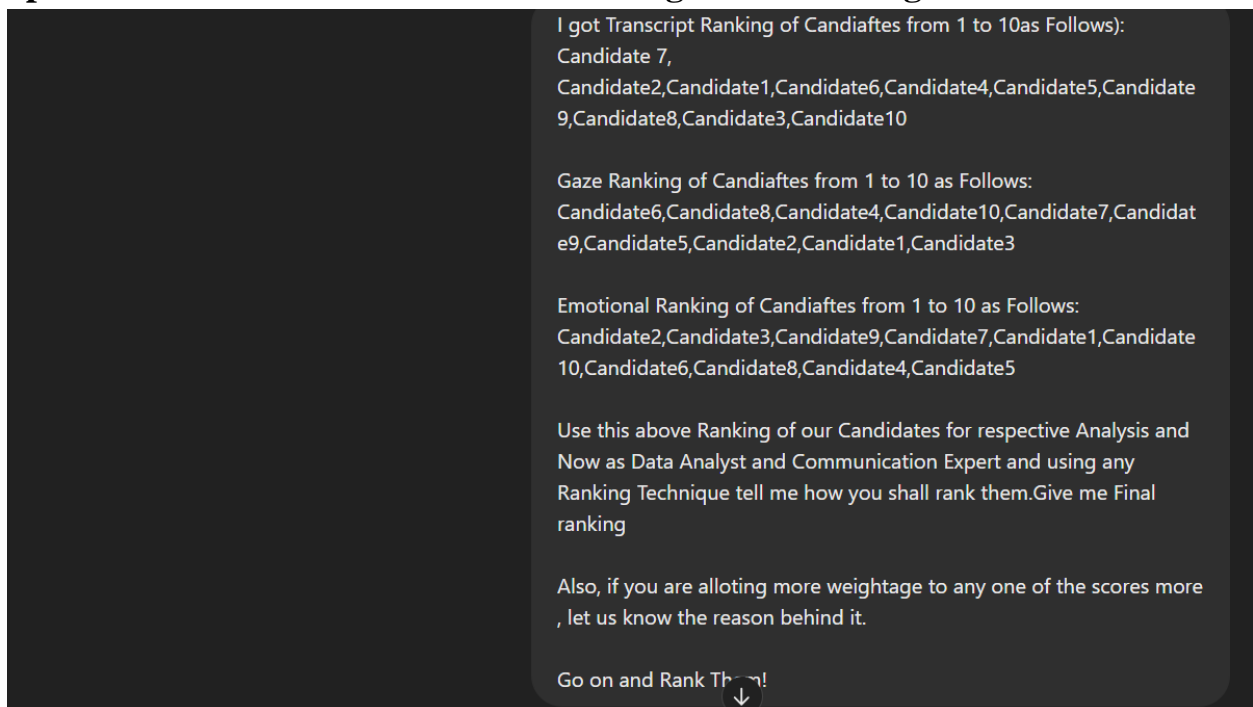
Achievements:

- "I received"
- "I was awarded"
- "I achieved"
- "I have been recognized for"
- "My accomplishments include"

- "I successfully"
- "I won"

These keywords can help in identifying the required entities during Named Entity Recognition (NER) model training.

8. After obtaining Candidates Rankings, this question was asked to find optimal and efficient method for obtaining Final Rankings.



To rank the candidates based on the provided transcript, gaze, and emotional rankings, we'll again use the **Borda Count method**. This approach will allow us to aggregate rankings into a final score for each candidate.

Step 1: Define the Rankings

The rankings for each candidate are as follows:

- **Transcript Rankings:**
 - Candidate 7: 1

- Candidate 2: 2
- Candidate 1: 3
- Candidate 6: 4
- Candidate 4: 5
- Candidate 5: 6
- Candidate 9: 7
- Candidate 8: 8
- Candidate 3: 9
- Candidate 10: 10
- **Gaze Rankings:**
 - Candidate 6: 1
 - Candidate 8: 2
 - Candidate 4: 3
 - Candidate 10: 4
 - Candidate 7: 5
 - Candidate 9: 6
 - Candidate 5: 7
 - Candidate 2: 8
 - Candidate 1: 9
 - Candidate 3: 10
- **Emotional Rankings:**
 - Candidate 2: 1
 - Candidate 3: 2
 - Candidate 9: 3
 - Candidate 7: 4
 - Candidate 1: 5
 - Candidate 10: 6
 - Candidate 6: 7
 - Candidate 8: 8
 - Candidate 4: 9
 - Candidate 5: 10

Step 2: Assign Points

Using the Borda Count method, the points assigned will be:

- Rank 1: 10 points
- Rank 2: 9 points
- Rank 3: 8 points
- Rank 4: 7 points
- Rank 5: 6 points
- Rank 6: 5 points
- Rank 7: 4 points
- Rank 8: 3 points
- Rank 9: 2 points
- Rank 10: 1 point

Step 3: Calculate Total Points for Each Candidate

Now, we'll sum the points for each candidate:

Now, we'll sum the points for each candidate:

Candidate	Transcript Rank	Gaze Rank	Emotional Rank	Total Points
Candidate 1	8	2	5	$8 + 2 + 5 = 15$
Candidate 2	9	3	10	$9 + 3 + 10 = 22$
Candidate 3	7	10	9	$7 + 10 + 9 = 26$
Candidate 4	6	7	2	$6 + 7 + 2 = 15$
Candidate 5	5	4	1	$5 + 4 + 1 = 10$
Candidate 6	4	10	6	$4 + 10 + 6 = 20$
Candidate 7	10	5	4	$10 + 5 + 4 = 19$
Candidate 8	3	1	8	$3 + 1 + 8 = 12$
Candidate 9	2	6	3	$2 + 6 + 3 = 11$
Candidate 10	1	9	7	$1 + 9 + 7 = 17$

Step 4: Final Ranking

Now we can rank the candidates based on their total points, from highest to lowest:

1. **Candidate 3:** 26 points

2. **Candidate 2:** 22 points
3. **Candidate 6:** 20 points
4. **Candidate 7:** 19 points
5. **Candidate 1:** 15 points (tie)
6. **Candidate 4:** 15 points (tie)
7. **Candidate 8:** 12 points
8. **Candidate 9:** 11 points
9. **Candidate 5:** 10 points
10. **Candidate 10:** 17 points

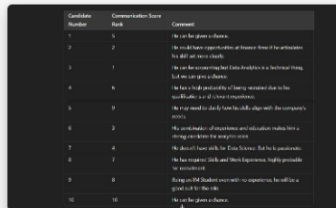
Conclusion

The final ranking based on the Borda Count method is:

1. **Candidate 3**
2. **Candidate 2**
3. **Candidate 6**
4. **Candidate 7**
5. **Candidate 1**
6. **Candidate 4**
7. **Candidate 8**
8. **Candidate 9**
9. **Candidate 5**
10. **Candidate 10**

9. Finally after obtaining Extracted Texts from Transcripts and Final rankings, this question was asked for final analysis.

Analyse and let me which candidate can be recruited and who can be rejected and let me know as a Data Analyst which features you emphasized more on and Why?



Candidate Number	Rank	Comment
1	10	He is the greatest of them.
2	9	He is the greatest of them.
3	8	He is the greatest of them.
4	7	He is the greatest of them.
5	6	He is the greatest of them.
6	3	His combination of experience and education makes him a strong candidate for analytics roles.
7	4	He has required skills and work experience, highly probable for recruitment.
8	7	He has required skills and work experience, highly probable for recruitment.
9	8	Despite a lower rank, the comment highlights strong skills and experience.
10	10	He is the greatest of them.

Analysis of Candidates' Transcript Performance for Recruitment

Key Features Emphasized

- **Rank:** Indicates the overall performance or suitability of the candidate, with 1 being the best and 10 being the worst.
- **Comments:** Provide qualitative insights into each candidate's skillset and potential fit for the role.

Candidates Recommended for Recruitment:

1. Candidate 6

- **Rank:** 3
- **Comment:** "His combination of experience and education makes him a strong candidate for analytics roles."
- **Reason:** High rank and positive comment about relevant skills and experience.

2. Candidate 8

- **Rank:** 7
- **Comment:** "He has required skills and work experience, highly probable for recruitment."
- **Reason:** Despite a lower rank, the comment highlights strong skills and experience.

3. Candidate 9

- **Rank:** 8

- **Comment:** “Being an IIM student even with no experience, he will be a good suit for the role.”
 - **Reason:** Lower rank but strong educational background and potential.
4. **Candidate 4**
- **Rank:** 6
 - **Comment:** “He has a high probability of being recruited due to his qualifications and relevant experience.”
 - **Reason:** Mid- level ranking but best Qualifications.
5. **Candidate 2**
- **Rank:** 2
 - **Comment:** “He could have opportunities at finance firms if he articulates his skill set more clearly.”
 - **Reason:** High rank indicates strong overall performance. Needs to improve articulation but has potential.
-

Candidates Recommended for Rejection or Further Evaluation

1. **Candidate 3**
- **Rank:** 1
 - **Comment:** “He can be accounting but Data Analytics is a technical thing, but we can give a chance.”
 - **Reason:** Top rank but the comment indicates a potential misalignment with the role.
2. **Candidate 10**
- **Rank:** 10
 - **Comment:** “He can be given a chance.”
 - **Reason:** Lowest rank but the comment suggests potential, making him worth considering for further evaluation.
3. **Candidate 5**
- **Rank:** 9
 - **Comment:** “He may need to clarify how his skills align with the company’s needs.”
 - **Reason:** Lower rank and comment indicating potential misalignment with the role.

4. **Candidate 7**

- **Rank:** 4
- **Comment:** “He doesn’t have skills for Data Science. But he is passionate.”
- **Reason:** Mid-level rank but lacks necessary skills despite passion.

5. **Candidate 1**

- **Rank:** 5
- **Comment:** “He can be given a chance.”
- **Reason:** No good Ranking and No skills