Salient vs Non-Salient

Average Length: 302.05 vs 120.50

Salient scenes are significantly longer than non-salient scenes, with an average length of 302.05 tokens compared to 120.50 tokens for non-salient scenes. This suggests that salient scenes contain more detailed information and are likely to be more complex in their content.

Average Lexical Diversity: 0.60 vs 0.72

Non-salient scenes show higher lexical diversity (0.72) compared to salient scenes (0.60). It could be explained by the fact that salient scenes, being longer, may need to repeat key terms and concepts for emphasis or clarity, thus reducing their lexical diversity score.

Average Named Entities: 22.29 vs 10.36

Salient scenes contain more than twice as many named entities on average (22.29) compared to non-salient scenes (10.36). This indicates that salient scenes are more likely to mention specific people, places, or organizations, which adds to their informational content and potential importance.

Average Sentiment: 0.07 vs 0.02

Salient scenes have a slightly more positive sentiment (0.07) compared to non-salient scenes (0.02). While both scores are close to neutral, this small difference suggests that salient scenes might contain more emotionally engaging or positively framed content.

Average Readability: 87.50 vs 85.11

This indicates that both types of scenes are generally easy to read, with salient scenes being marginally more accessible.

POS Distribution (Salient vs Non-Salient):

VBP: 0.03 vs 0.02
CD: 0.01 vs 0.01
VBD: 0.01 vs 0.01
VBZ: 0.05 vs 0.05
PRP: 0.07 vs 0.06
NNP: 0.14 vs 0.16
JJ: 0.04 vs 0.04

Non-salient scenes have a slightly higher proportion (0.16 vs 0.14)
This could suggest that salient scenes use proper nouns more efficiently or in more meaningful contexts.

Salient scenes use slightly more personal pronouns (0.07 vs 0.06), possibly indicating more focus on character actions or dialogue.

Tense Distribution Analysis:

1. Present Tense:
   o Salient scenes: 78%
   o Non-salient scenes: 80%
2. Past Tense:
   o Salient scenes: 22%
   o Non-salient scenes: 20%

Implications

1. **Information Density:** Salient scenes are packed with detailed information, more named entities, and slightly lower lexical diversity. Salient scenes are more focused on delivering specific, detailed information about particular subjects or events.

2. **Structural Differences**: Salient scenes focus more on actions and specific entities, while non-salient scenes include more general descriptions or explanations.

3. **Accessibility**: Despite being longer and more information-dense, salient scenes maintain high readability, indicating well-presented information.