INFORMATION MANAGEMENT PROJECT NARRATIVE:

SCRIPTING THE WWE

Group No xx:

Ali Sakhi Khan - ask3467

Joseph Bridges -

Jahnavi Angati - ja54632

Antra Sinha - as224589

Cassie Ren - clr3745

1. Project Selection:

WWE has iconic figures that have been the face of the program, drawing fans and ensuring its success. Imagining alternatives, we ponder how different choices could have altered WWE's narrative. By leveraging Data Science, we aim to align WWE's character development with fan perceptions, enhancing engagement and strengthening the bond between fans and wrestlers.

2. Data Acquisition:

Our research consisted of collecting wrestler information and audience comments from TheSmackdownHotel and WrestleTalk's YouTube channel, respectively, which is known for its engaged community, providing a broad range of opinions, independent of specific events. Our successful scraping of over 10,000 comments from this channel offered a rich perspective on audience sentiments.

3. Data Sampling and Cleaning:

We analyzed 10,000 comments from WrestleTalk's year-long comment section, using the nltk package in Python to remove stop words and VADER for sentiment analysis. Each comment received a sentiment score, resulting in a new 'Sentiment_Words' column categorizing comments as positive, negative, or neutral based on the score's sign (Table 2.1). This enhances our analysis and provides deeper insights into audience reactions. The Smackdown Hotel, a source known for its well-organized and detailed content, provided wrester data, which was exceptionally clean and free of any missing values (Table 2.2).

4. Tools and Technology:

3.1 PYTHON:

Utilizing user-generated YouTube comments, we performed a lift analysis to assess co-mentions between wrestlers (Figure 3.1.1) and another between wrestlers and wrestling moves, emphasizing pairs with lift values above 1, indicating significant co-mention frequency. The final table highlighted five unique cases where wrestlers were notably associated with moves they don't typically perform (Figure 3.1.2)

3.2 NEO4J:

To represent the WWE narrative structure based on lift analysis relationships, Neo4J code was algorithmically generated. Wrestlers were established as nodes, and for the first lift analysis, the highest lift counterparts were matched with a 'WITH' relation (Figure 3.2.1). In the second analysis, wrestlers were matched based on move associations, and relationships were formed with move names. The resulting Neo4j graph reflects the audience's perception of wrestlers'

connections in the WWE narrative (Figure 3.2.2). Wrestlers with numerous connections likely play central roles, while those with fewer connections are perceived as less relevant to the audience's narrative understanding.

3.3 MONGODB:

Utilizing MongoDB, we analyzed wrestler popularity, uncovering the most celebrated moves and wrestlers. This data-driven approach highlighted trends and preferences, informing marketing and event strategies in professional wrestling. The codes for each section are attached in the appendix below.

3.4 SNOWFLAKE:

A WWE database was organized into three schemas: Comments, WWE_Authors, and WWE_Info. The Comments schema contained Wrestletalk's comments in the Comments_Sentiments table, while the Authors table held author information with Author_ID and Author_Name columns. This structured setup ensures clarity for author-comment relationships and facilitates future scalability. The WWE_INFO schema included the WWE_INFO table and WWE_WRESTLER_ATTRIBUTES, capturing words associated with wrestlers. Code snippets of each section are attached in the appendix.

5. Research Question:

How can we identify under and over-performing wrestlers and adapt the WWE narrative to include under-performing wrestlers?

To identify under and over-performing wrestlers and adapt the WWE narrative, our models leverage Python and Neo4j to determine the central and peripheral wrestlers according to audience perception. MongoDB aggregates wrestler information, extracting attributes linked to success. Snowflake provides sentiment scores for attributes, indicating whether a wrestler is better suited for a star or heel role. This holistic approach guides storyline decisions, allowing us to spotlight peripheral wrestlers and optimize narrative dynamics.

6. Real Life Application:

This model can be applied to other marginally-scripted reality shows, such as House Hunters. For example, analysis of reviews regarding previous house remodels, architecture styles, and budgets can identify the popular features according to viewers, allowing HH to script their future episodes and seasons towards different segments of their audience.

Appendix:

Comment_No_Stopword	Sentiment	Sentiment_Words
new information came out. Please bring back Adam. Can't believe lied.	0.5399	Positive
sure anything going different Kairi's last stint WWE. NXT little nothing show mains tag champion title everyone knows WWE care about. Beyond injured 3x WWE proclaimed "best world" Nia Jax. probably going end getting r	0.8979	Positive
John Cena needed do, mention believe going TBS part Wipeout Show I'm mistaken.	-0.3612	Negative
Holy crap Luke Sunday 😲🎉😲❤ love	0.3818	Positive
Add Asuka.	0.0	Neutral
Wow David Beckham trolling ideas… Maybe cena posted pic Beckham bc married spice girl, John really teasing becoming Cena spice spice girls	0.4951	Positive
Banks/Mone injury, backstage, PR nightmare. hurt, major heat backstage getting mad fans signings. really hope resign disappeares obscurity.	-0.8353	Negative
Osprey WWE severely limits ability great matches. There's maybe 2 people WWE could good match him. options literally anywhere else much more, much better, much brighter.	0.9017	Positive
writer strike ended September. Cena leaving soon.	-0.128	Negative
4:24 Wrestling fans cringey lol!	0.4753	Positive
Kurt angle beats long Paul 2024?	0.0	Neutral
Cena tna Aew Internet leak real 😮	-0.34	Negative
Including herself; Sasha Banks injured people Nia Jax Charlotte.	-0.4019	Negative
Osprey go wherever desires. long get buried like Ricochete WWE see go there. But, higher ceiling AEW see go there. positive negative both.	0.1531	Positive
3:57 You're insane Luke Owen ultimately it's whether wants come back	-0.4019	Negative
Luke see bash Berlin main event ilja Gunther raise Dunn Dragunov Bate Gunther 4 way	0.0	Neutral
Osprey would ruined wwe. Dont throw career away kid. like 4 5 wrestlers wwe could actually hang osprey. would sub mid best wwe	0.5574	Positive
like Mercedes Mone back WWE. would rather see upcoming NXT women make Main roster shuffle better.	0.6597	Positive
guys wanna fly Australia, love come!!!	0.7249	Positive
think WWE Crown Jewel 2023? Let know comments… Get Survivor Series watchalong tickets 👉https://www.eventbrite.com/e/wrestletalks-wwe-survivor-series-2023-party-tickets-750689170517?aff=oddtdtcreator Fantasy	0.8876	Positive
Honestly...Crown Jewel upto mark all.... average ppv.	0.3612	Positive
Yeah feel way. seems convenient Esco	0.296	Positive
Glad I'm one can't stand current iteration Seth Rollins, point every match title run disappointment lose. fan "character." wanted sing along, I'd watch Cocomelon nephew.	-0.1779	Negative
Screw scammer Logan Paul!	-0.1759	Negative
John cena deserved way better ending solo. Derp. wtf wwe.	-0.2263	Negative
feel bad Bayley seems finding success Damage CTRL! it's sad LA Knight clearly main event match lost anyway.	-0.7177	Negative
Saturday Night RAW	0.0	Neutral
think roman going drop belt crown jewel, feel like derails L Knight. PRAY trying rock beat roman wrestlemania belt put belt someone else wrestle. Roman Bruno....not defending every 30 days, keeping interference, gets old fa	0.5622	Positive
point, Roman's persistent hoarding getting stale faster anyone could imagined about.... someone says "hE HeeL sUpPoseD pLaY fOU!" Honestly, Roman selfish reign past decade so... No-one looks better coming Roman's	0.6369	Positive

Table 2.1: Snippet of a larger table highlighting the sentiment score of each comment

Real Name	Ring Name	Gender	Birth Place	Finisher	Age_yrs
Theresa Serrano	Zoey Stark	Female	Salt Lake City, Utah	Z-360	29
Thea Megan Trinidad Bdgen	Zelina Vega	Female	Queens, New York	Code Red	32
Daniel Sean Vidot	Xyon Quinn	Male	Brisbane, Queensland, Australia	Swinging Samoan Drop	33
Zhao Xia	Xia Li	Female	Chongqing, China	Spinning Heel Kick	35
Austin Watson	Xavier Woods	Male	Columbus, Georgia	Limit Break	37
Barry Young	Wolfgang	Male	Glasgow, Scotland, United Kingdom	Caber Toss Suplex	36
Deveon Everhart Aiken	Wes Lee	Male	Dayton, Ohio	Kardiak Kick / Jocay-le	34
Karen Yu	Wendy Choo	Female	Queens, New York	Spring Roll	31
Stuart Alexander Bennett	Wade Barrett	Male	Penwortham, Lancashire, England	Bull Hammer	43
Cal Bloom	Von Wagner	Male	Osseo, Minnesota	Double Underhook Brainbuster	29
Vincent Kennedy McMahon	Vince McMahon	Male	Pinehurst, North Carolina	McMahon Stunner	78
Rinku Singh Rajput	Veer	Male	Gopiganj Bhadohi, Uttar Pradesh, India	Cervical Clutch	35
Sarah Rowe ne Bridges	Valhalla	Female	Louisville, Kentucky	Kentucky Knee	30
Rita Reis	Valentina Feroz	Female	Brazil	Alley-oop into Crossbody - with Yulisa León	28
Tyler Bate	Tyler Bate	Male	Netherton, Dudley, England	Tyler Driver '97	26
Patricia Anne Stratigias	Trish Stratus	Female	Toronto, Ontario, Canada	Stratusfaction	47
Paul Michael Levesque	Triple H	Male	Greenwich, Connecticut	Pedigree	54
Matrick Belton	Trick Williams	Male	Columbia, South Carolina	The Flash Knee	29
Joseph Ariola	Tony DAngelo	Male	Oak Park, Illinois	Forget About It	28
Tommaso Whitney	Tommaso Ciampa	Male	Boston, Massachusetts	Project Ciampa	38
Jessica Woynilko	Tiffany Stratton	Female	Prior Lake, Minnesota	Prettiest Moonsault Ever	24
Maddie Knisely	Thea Hail	Female	Pittsburgh, Pennsylvania	Kimura Lock	20

Table 2.2: Scraped Roster Information from The Smackdown Hotel

New Name	Highest Lift	Second Highest
zoeystark	tatumpaxley	trickwilliams
zelinavega	eltonprince	damonkemp
xyonquinn	odysseyjones	damonkemp
xiali	eltonprince	damonkemp
xavierwoods	maxxinedupri	lyravalkyria
...	...	...
alexabliss	gigidolin	lyravalkyria
albafyre	sanga	solruca
akiratozawa	robertroode	meikosatomura
adampearce	kitwilson	kianajames
ajstyles	ludwigkaiser	drewgulak

183 rows × 2 columns

Figure 3.1.1: Highest Two Lift Wrestlers for Each Wrestler

```
spear          braunstrowman
caterpillar          valhalla
insaneelbow            themiz
piledriver        adampearce
silence         nikkitalyons
```

Figure 3.1.2: The Five Move-Wrestler Pairs from Lift Analysis

```
CREATE (damonkemp)-[:WITH]->(andrechase)
MATCH (amarimiller:Wrestler {name: 'Amari Miller'}), (damonkemp:Wrestler {name: 'Damon Kemp'})
CREATE (amarimiller)-[:WITH]->(damonkemp)
CREATE (damonkemp)-[:WITH]->(amarimiller)
MATCH (alexabliss:Wrestler {name: 'Alexa Bliss'}), (gigidolin:Wrestler {name: 'Gigi Dolin'})
CREATE (alexabliss)-[:WITH]->(gigidolin)
CREATE (gigidolin)-[:WITH]->(alexabliss)
MATCH (albafyre:Wrestler {name: 'Alba Fyre'}), (sanga:Wrestler {name: 'Sanga'})
CREATE (albafyre)-[:WITH]->(sanga)
CREATE (sanga)-[:WITH]->(albafyre)
MATCH (akiratozawa:Wrestler {name: 'Akira Tozawa'}), (robertroode:Wrestler {name: 'Robert Roode'})
CREATE (akiratozawa)-[:WITH]->(robertroode)
CREATE (robertroode)-[:WITH]->(akiratozawa)
MATCH (adampearce:Wrestler {name: 'Adam Pearce'}), (kitwilson:Wrestler {name: 'Kit Wilson'})
CREATE (adampearce)-[:WITH]->(kitwilson)
CREATE (kitwilson)-[:WITH]->(adampearce)
MATCH (ajstyles:Wrestler {name: 'AJ Styles'}), (ludwigkaiser:Wrestler {name: 'Ludwig Kaiser'})
CREATE (ajstyles)-[:WITH]->(ludwigkaiser)
CREATE (ludwigkaiser)-[:WITH]->(ajstyles)
```

Figure 3.2.1: Neo4j Code Generated in Python

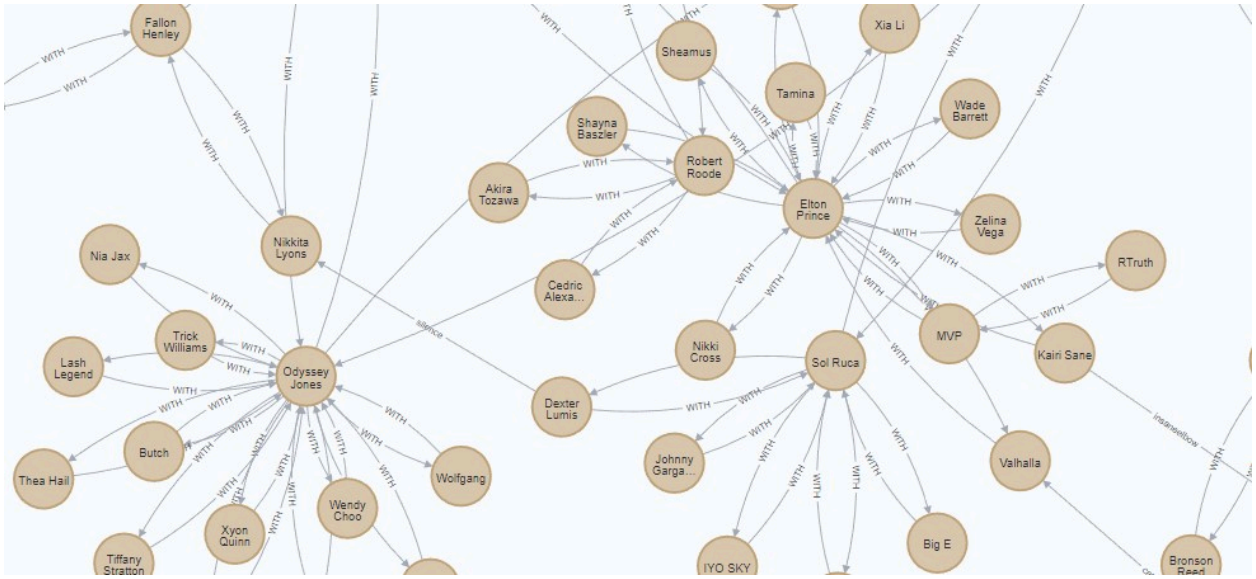

Figure 3.2.2: A Section of the Resulting Neo4j Graph with Both Moves and 'WITH' Relationships

```
WRESTLERINFO> db.wrestlerInfo.aggregate([
...   {
...         $group: {
...         _id: "$Ring Name", // Group by the wrestler ring name
...         count: { $sum: 1 } // Count occurrences
...         }
...   },
...   {
...         $sort: { count: -1 } // Sort by count in descending order
...   }
... ])
[
  { _id: 'Nikki Cross', count: 2 },
  { _id: 'Thea Hail', count: 2 },
  { _id: 'Meiko Satomura', count: 2 },
  { _id: 'Ivy Nile', count: 2 },
  { _id: 'Tatsumi Fujinami', count: 2 },
  { _id: 'RTruth', count: 2 },
  { _id: 'Tatum Paxley', count: 2 },
  { _id: 'Boa', count: 2 },
  { _id: 'Blair Davenport', count: 2 },
  { _id: 'Tyler Bate', count: 2 },
  { _id: 'Wade Barrett', count: 2 },
  { _id: 'Scrypts', count: 2 },
  { _id: 'Roxanne Perez', count: 2 },
  { _id: 'Sanga', count: 2 },
  { _id: 'Brutus Creed', count: 2 },
  { _id: 'Vince McMahon', count: 2 },
  { _id: 'Brock Lesnar', count: 2 },
  { _id: 'Dijak', count: 2 },
  { _id: 'Asuka', count: 2 },
  { _id: 'Zelina Vega', count: 2 }
]
```

Figure 3.3.1: Code Snippet to find Most Frequent Wrestler Name

```
WRESTLERINFO> db.wrestlerInfo.aggregate([
...   {
...         $group: {
...         _id: "$Finisher", // Group by the finisher move
...         count: { $sum: 1 } // Count occurrences of each finisher move
...         }
...   },
...   {
...         $sort: { count: -1 } // Sort by count in descending order
...   }
... ])
[
  { _id: 'Running Powerslam', count: 4 },
  { _id: 'Samoan Splash', count: 4 },
  { _id: 'Code Red', count: 4 },
  { _id: 'Spinning Heel Kick', count: 4 },
  { _id: '630° Senton', count: 4 },
  { _id: 619, count: 4 },
  { _id: 'The Final Prayer', count: 2 },
  { _id: 'Devlin Side / Devil Inside', count: 2 },
  { _id: 'Stunner', count: 2 },
  { _id: "Gallows Pole / Hangman's Noose", count: 2 },
  { _id: 'McMahon Stunner', count: 2 },
  { _id: 'Diving Split Legdrop', count: 2 },
  { _id: 'Pop-Up Kneeling Powerbomb', count: 2 },
  { _id: 'Rocket Kick', count: 2 },
  { _id: 'Khallas', count: 2 },
  { _id: 'Cervical Clutch', count: 2 },
  { _id: 'Caber Toss Suplex', count: 2 },
  { _id: 'Big Top Pop', count: 2 },
  { _id: 'Death Rowe', count: 2 },
  { _id: 'Flawless Victory', count: 2 }
]
```

Figure 3.3.2: Code Snippet to find Most Mentioned Finisher

```
WRESTLERINFO> db.wrestlerInfo.aggregate([
...   {
...         $group: {
...         _id: "$Gender", // Group by the gender field
...         numberOfWrestlers: { $sum: 1 } // Count the number of wrestlers per gender
...         }
...   },
...   {
...         $sort: { numberOfWrestlers: -1 } // Sort by the count in descending order
...   }
... ])
[
  { _id: 'Male', numberOfWrestlers: 238 },
  { _id: 'Female', numberOfWrestlers: 128 }
]
```

Figure 3.3.3: Code Snippet to find Wrestler Popularity by Gender

```
WRESTLERINFO> db.wrestlerInfo.aggregate([
...   {
...         $group: {
...         _id: "$Birth Place",
...         count: { $sum: 1 }
...         }
...   },
...   { $sort: { count: -1 } }
... ])
[
  { _id: 'Glasgow, Scotland', count: 10 },
  { _id: 'Los Angeles, California', count: 6 },
  { _id: 'Knoxville, Tennessee', count: 6 },
  { _id: 'Chicago, Illinois', count: 4 },
  { _id: 'Queens, New York', count: 4 },
  { _id: 'Lagos, Nigeria', count: 4 },
  { _id: 'Sydney, New South Wales, Australia', count: 4 },
  { _id: 'Sacramento, California', count: 4 },
  { _id: 'Cleveland, Ohio', count: 4 },
  { _id: 'Tampa, Florida', count: 4 },
  { _id: 'Charlotte, North Carolina', count: 4 },
  { _id: 'Calgary, Alberta, Canada', count: 4 },
  { _id: 'San Francisco, California', count: 4 },
  { _id: 'Pittsburgh, Pennsylvania', count: 4 },
  { _id: 'Monterrey, Nuevo León, Mexico', count: 2 },
  { _id: 'Limerick, Republic of Ireland', count: 2 },
  { _id: 'Penwortham, Lancashire, England', count: 2 },
  { _id: 'Edgewater, Illinois', count: 2 },
  { _id: 'Tokyo, Japan', count: 2 },
  { _id: 'Springfield, Missouri', count: 2 }
]
```

Figure 3.3.1: Code Snippet to find Geographic Distribution

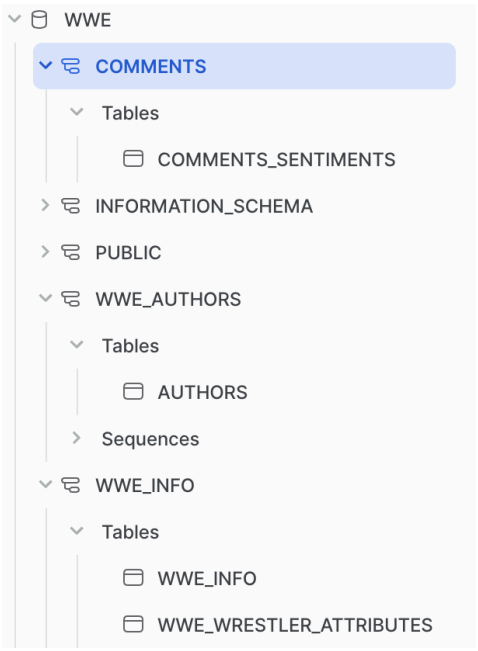

Figure 3.4.1: WWE Database on Snowflake

```sql
-- Frequency of wrestler names
SELECT
  w.First_Name,
  w.Last_Name,
  COUNT(*) AS Name_Count
FROM
  WWE_INFO w
JOIN
  COMMENTS.COMMENTS_SENTIMENTS c ON
    LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.First_Name), '%')
    OR LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.Last_Name), '%')
GROUP BY
  w.First_Name, w.Last_Name
ORDER BY
  Name_Count DESC;
```

Figure 3.4.2: SQL code to find Wrestler mentions

```sql
SELECT
  w.First_Name,
  w.Last_Name,
  c.Sentiment_words,
  COUNT(*) AS Name_Sentiment_Count
FROM
  WWE_INFO w
JOIN
  COMMENTS.COMMENTS_SENTIMENTS c ON
    LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.First_Name), '%')
    OR LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.Last_Name), '%')
GROUP BY
  w.First_Name, w.Last_Name, c.Sentiment_words
ORDER BY
  Name_Sentiment_Count DESC;
```

Figure 3.4.3: SQL code showing wrestler sentiments based on the comments

```sql
SELECT
  w.First_Name,
  w.Last_Name,
  COUNT(*) AS Name_Sentiment_Count,
  SUM(CASE WHEN c.Sentiment_words = 'Positive' THEN 1 ELSE 0 END) AS Positive_Count,
  SUM(CASE WHEN c.Sentiment_words = 'Negative' THEN 1 ELSE 0 END) AS Negative_Count,
  SUM(CASE WHEN c.Sentiment_words = 'Neutral' THEN 1 ELSE 0 END) AS Neutral_Count,
  IFNULL(SUM(CASE WHEN c.Sentiment_words = 'Positive' THEN 1 ELSE 0 END) / NULLIF(SUM(CASE WHEN c.Sentiment_words = 'Negative'
THEN 1 ELSE 0 END), 0), 0) AS Positive_Negative_Ratio
FROM
  WWE_INFO w
JOIN (
  SELECT DISTINCT
    COMMENT_NO_STOPWORD,
    Sentiment_words
  FROM
    wwe.COMMENTS.COMMENTS_SENTIMENTS
) c ON
  LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.First_Name), '%')
  OR LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.Last_Name), '%')
GROUP BY
  w.First_Name, w.Last_Name
ORDER BY
  Positive_negative_ratio DESC;
```

Figure 3.4.4: SQL code showing positive to negative sentiment ratio

```sql
-- WWE Wrestler attributes
CREATE TABLE WWE_Wrestler_Attributes (
    Adjective VARCHAR(255)
);


INSERT INTO WWE_Wrestler_Attributes (Adjective) VALUES
('Charismatic'), ('Aggressive'), ('Athletic'), ('Bold'), ('Brave'), ('Cunning'), ('Daring'), ('Determined'), ('Dynamic'),
('Energetic'),
('Fearless'), ('Fierce'), ('Gritty'), ('Heroic'), ('Intense'), ('Jubilant'), ('Lively'), ('Mighty'), ('Nimble'), ('Outstanding'),
('Passionate'), ('Powerful'), ('Agile'), ('Radiant'), ('Resilient'), ('Robust'), ('Stalwart'), ('Swift'), ('Tenacious'),
('Triumphant'),
('Unyielding'), ('Vibrant'), ('Wise'), ('Zealous'), ('Awe-inspiring'), ('Breathtaking'), ('Clever'), ('Dazzling'), ('Epic'),
('Formidable'),
('Gallant'), ('Harmonious'), ('Invincible'), ('Jovial'), ('Kinetic'), ('Luminous'), ('Majestic'), ('Noble'), ('Optimistic'),
('Prestigious'),
('Quirky'), ('Rambunctious'), ('Spirited'), ('Titanic'), ('Unbeatable'), ('Vivacious'), ('Wholesome'), ('Xtraordinary'),
('Youthful'), ('Zestful'),
('Adaptable'), ('Bold-hearted'), ('Capable'), ('Determined'), ('Eloquent'), ('Fearless'), ('Generous'), ('Innovative'),
('Joyful'),
('Keen'), ('Loyal'), ('Magnificent'), ('Nurturing'), ('Optimistic'), ('Passionate'), ('Quixotic'), ('Radiant'), ('Steadfast'),
('Trustworthy'),
('Unyielding'), ('Valiant'), ('Wholesome'), ('Xenial'), ('Youthful'), ('Zealous');
```

Figure 3.4.5: SQL code creating the attribute table

```sql
SELECT
  a.Adjective,
  w.FIRST_NAME,
  w.LAST_NAME,
  c.SENTIMENT_WORDS,
  COUNT(*) AS Frequency
FROM
  wwe.wwe_info.WWE_Wrestler_Attributes a
JOIN
  wwe.comments.COMMENTS_SENTIMENTS c ON LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(a.Adjective), '%')
LEFT JOIN
  wwe.wwe_info.WWE_INFO w ON LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.FIRST_NAME), '%') OR
LOWER(c.COMMENT_NO_STOPWORD) LIKE CONCAT('%', LOWER(w.LAST_NAME), '%')
GROUP BY
  a.Adjective,
  w.FIRST_NAME,
  w.LAST_NAME,
  c.SENTIMENT_WORDS
ORDER BY
  Frequency DESC;
```

Figure 3.4.6: SQL code showing attribute association