LeBron James: A Legacy of Dominance

Motivation: Growing up, LeBron James has been a constant presence in my life as a basketball fan and I simply don't know the NBA without him. He entered the NBA the same year I was born, and for as long as I can remember, he has been a dominant force in the league. His journey from a high school prodigy to the greatest player of all time has been nothing short of inspirational. This project is a tribute to his incredible career and a personal exploration of the data that highlights his dominance as he's still defying time and longevity today.

Background: LeBron James has redefined what it means to be a basketball legend. From his early days in Akron, Ohio, to his multiple championships and MVP awards, his ability to consistently perform at the highest level, year after year, is a remarkable feat that deserves to be celebrated and analyzed. This project aims to highlight just how great he is compared to the rest of his peers, as well as the many achievements that have solidified LeBron's place in basketball history.

Step 1: Ask

- What is the actual skill gap between LeBron James and the average NBA player?
- What are the key metrics that showcase LeBron's performance?
- How has his performance evolved over the years?
- Is there any indication of his output on the court slowing down or regressing?

Step 2: Prepare

- Datasets Used:
    - NBA Player Stats (1950-2017)
    - NBA Player Stats (1996-2023)
    - https://www.basketball-reference.com/
- The first objective was to find a dataset that I could use to get all of LeBron's statistics during his career from 2004-Present. Unfortunately, there wasn't a dataset I could find on Kaggle that had stats from his entire career so I had to improvise and use two separate datasets. One offered statistics of all NBA players from 1950 to 2017 and another one had statistics from 1996 to 2023. I decided that I would combine them in SQL by using the 1st dataset for seasons 2004 to 2017 and the 2nd dataset for seasons 2018 to 2023.
    - I used basketball reference to pull any missing stats and fill any null values that were present throughout the project.

Step 3: Process

- **Data Cleaning and Manipulation:**
  - Tools: SQL and Excel for data cleaning and manipulation
  - Steps: Combining datasets, removing duplicates, handling missing values, organizing and transforming data into a suitable format to make it more useful for analysis and visualization.
- Combining the 2 datasets

```
--Combine the datasets using UNION ALL with shared columns, calculated averages, True Shooting Percentage, and additional columns
SELECT player,
    age,
    tm              AS team,
    Round((pts / g), 2) AS pts,
    Round((trb / g), 2) AS reb,
    Round((ast / g), 2) AS ast,
    `ts%`           AS ts_pct,
    ws,
    NULL AS net_rating,
    per,
    `efg%`,
    cast(year AS string) AS year,
    NULL            AS season
FROM   `lebron-455307.nbastats.stats1`
UNION ALL
SELECT player,
    age,
    team,
    round(pts, 2) AS pts,
    round(reb, 2) AS reb,
    round(ast, 2) AS ast,
    ts_pct,
    NULL AS ws,
    net_rating,
    NULL AS per,
    NULL AS `efg%`,
    NULL AS year,
    season
FROM   `lebron-455307.nbastats.stats2`
```

- Save the combined the dataset as a new table

```
-- Save the combined dataset to a new table
CREATE OR REPLACE TABLE `lebron-455307.nbastats.nba_1950_2023_shared` AS
SELECT player,
    age,
    team,
    pts,
    reb,
    ast,
    ts_pct,
    ws,
    net_rating,
    per,
    `efg%`,
    year,
    season
FROM  (
        SELECT player,
            age,
            tm                  AS team,
            round((pts / g), 2) AS pts,
            round((trb / g), 2) AS reb,
            round((ast / g), 2) AS ast,
            `ts%`               AS ts_pct,
            ws,
            NULL AS net_rating,
            per,
            `efg%`,
            cast(year AS string) AS year,
            NULL            AS season
        FROM  `lebron-455307.nbastats.stats1`
        UNION ALL
        SELECT player,
            age,
            team,
            round(pts, 2) AS pts,
            round(reb, 2) AS reb,
            round(ast, 2) AS ast,
            ts_pct,
            NULL AS ws,
```

```
        net_rating,
        NULL AS per,
        NULL AS `efg%`,
        NULL AS year,
        season
    FROM   `lebron-455307.nbastats.stats2` )
```

- After I made this new table, I had every player's name, age, team, pts average, reb average, ast average, true shooting %, win shares, net rating, PER, eFG%, and year of play from 1950 all the way up to 2023.
- The first task was to pull LeBron's stats to show his career progression and also compare him to the average player in the NBA.
- I started with points and did the same for rebounds, assists, and PER by just substituting "Points" with the other statistics.

```
-- Combine LeBron's average points per game and league average points per game for each season from 2004 to 2023

WITH lebron_points AS (

  SELECT

    COALESCE(CAST(Year AS STRING), Season) AS Season,

    ROUND(AVG(Pts), 2) AS LeBron_Avg_Points

  FROM `lebron-455307.nbastats.nba_1950_2023_shared`

  WHERE Player = 'LeBron James' AND (CAST(Year AS INT64) BETWEEN 2004 AND 2023
OR Season BETWEEN '2004' AND '2023')

  GROUP BY Season

),

league_points AS (

  SELECT

    COALESCE(CAST(Year AS STRING), Season) AS Season,

    ROUND(AVG(Pts), 2) AS League_Avg_Points

  FROM `lebron-455307.nbastats.nba_1950_2023_shared`
```

```
 WHERE CAST(Year AS INT64) BETWEEN 2004 AND 2023 OR Season BETWEEN '2004' AND '2023'

 GROUP BY Season

)

-- Select from both subqueries

SELECT

 Season,

 LeBron_Avg_Points,

 League_Avg_Points

FROM lebron_points

FULL OUTER JOIN league_points USING (Season)

ORDER BY Season
```

- Unfortunately, for the PER query, I was missing both LeBron's averages and the league averages data from 2018 to 2023, so I had to manually add that from Basketball Reference after I imported the table into Excel.
- Next, for LeBron's accolades by year I decided to strictly focus on his NBA accolades even though he does have multiple gold medals from the olympics and other awards like Sports Illustrated Sportsperson of the Year. I focused on a handful of awards including MVPs, Finals MVPs, All-NBA selections, All-Defensive Team selections, All-Star selections, his Rookie of the Year award, and of course the number of times he's made the NBA finals and the number of times he's won. I imported all of his accolades into Excel from Basketball Reference and organized it into an easily readable table sorted by year.
- Lastly, I exported these new tables to Excel for final formatting

Step 4: Analyze

- Stats & Trends
    - LeBron James is miles ahead of his peers in every single statistical category
        - He has doubled or almost doubled the average players PER in every season of his career outside his rookie season and 2023 (still 1.5x the average)

- After his rookie season, he has never averaged fewer than 25 points per game, fewer than 6 rebounds per game, or fewer than 6 assists per game!
- LeBron James has more championships than 25 NBA franchises!
- He has more 50-points games than single-digit scoring games in his career
- He is on a current streak of 1,290 straight games scoring 10+ points!
- He has finished top-10 in MVP voting in all but 3 seasons
- LeBron was first in PER for 6 straight seasons, an accomplishment only achieved by Wilt Chamberlain and Michael Jordan

Step 5: Share

- I created 4 visualizations in Tableau to highlight LeBron's dominance throughout his long career, and used these 4 visualizations to make a very insightful dashboard including some of his background, nicknames, and images to make it more visually appealing.

See Next Page for Dashboard

LeBron James: A Legacy of Dominance

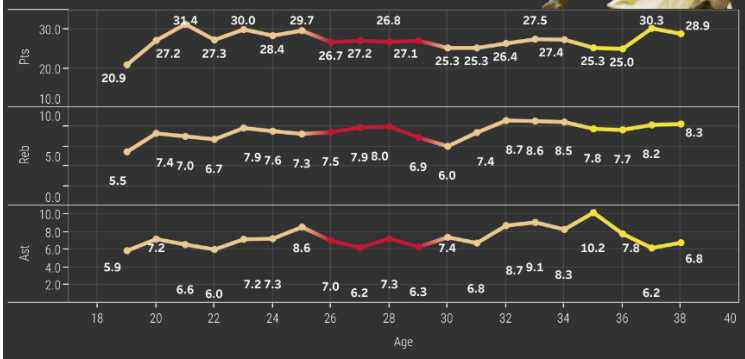
"The Chosen One"

Anointed as the 'Chosen One' from a young age, LeBron James has lived up to and exceeded the colossal expectations placed upon him

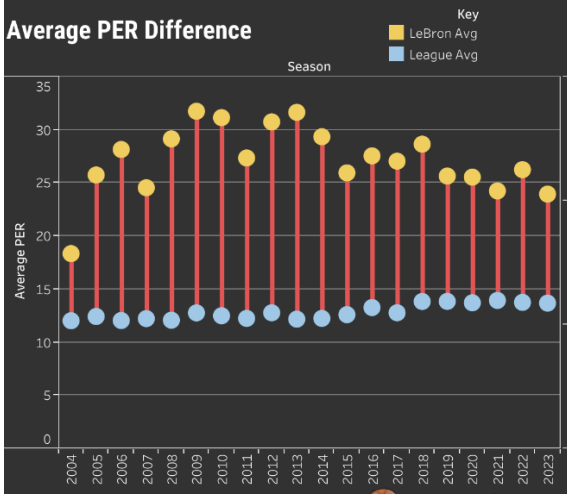
"The Kid From Akron"

From humble beginnings in Akron, Ohio to the pinnacle of basketball greatness, LeBron James' journey is a testament to hard work, perseverance, and an unyielding passion for the game

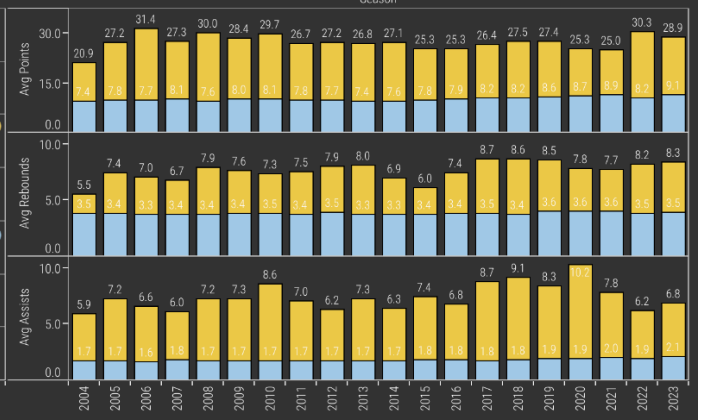
Pts, Reb, Ast Progression



Average PER Difference



LeBron Vs The Average

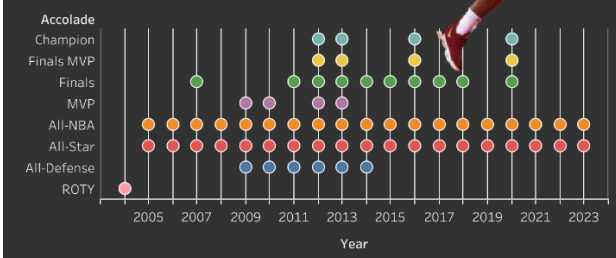


LeBron James' explosive athleticism and gravity-defying dunks have become iconic in the world of basketball. Whether it's a coast-to-coast sprint ending in a thunderous slam or a high-flying alley-oop, his ability to dominate the airspace is unmatched

"King James"

Crowned as one of the greatest to ever play the game and the king of longevity, LeBron James has redefined what it means to be a basketball lege..

LeBron's Accolades By Year



A visualization of the dominance of LeBron James
By Vito Mkrtychyan

Step 6: Act

Key Takeaways:

LeBron James continues to defy the odds and redefine what is possible in professional basketball. There is no indication anywhere of a decline or regression in the product that LeBron James puts out onto the court on a night-to-night basis. Even at the age of 40, when most players have long retired, his output is simply unprecedented. His ability to maintain peak performance levels, adapt his game, and lead his team is a testament to his dedication, training, and basketball IQ.

Insights and Actions:

- Sustained Excellence: LeBron's career longevity and sustained excellence provide valuable insights into the importance of physical fitness, adaptability, and mental toughness. These attributes are crucial for athletes aiming for long-term success.
- Leadership and Impact: Beyond his individual statistics, LeBron's leadership on and off the court has had a profound impact on his teams and the broader basketball community. His influence extends beyond the game, inspiring future generations of players.
- Data-Driven Decisions: The analysis of LeBron's career statistics highlights the power of data in sports. Teams, coaches, and analysts can leverage similar data-driven approaches to make informed decisions, optimize player performance, and develop effective strategies.
- Fan Engagement: By visualizing and sharing LeBron's achievements, fans gain a deeper appreciation for his contributions to the sport. This enhances fan engagement and fosters a greater connection between players and their supporters.

What I've Learned

Throughout the process of creating this project, I've gained valuable insights and refined my skills to enhance my understanding of data analysis and visualization. Here are some of my personal takeaways:

1. Importance of Data Preparation:

- The initial steps of collecting, cleaning, and transforming data are crucial for ensuring accurate and meaningful analysis. I learned how to effectively use SQL to handle these tasks, which is essential for any data-driven project.

2. Power of Visualization:

- Visualizing data in Tableau allowed me to present complex information in an engaging and appealing way. I discovered how different chart types and interactive elements can highlight key insights and tell a compelling story.

3. Storytelling with Data:

- Combining data with narrative elements, such as quotes and images, can significantly enhance the impact of a project. This approach not only makes the data more relatable but also helps to convey the significance of the findings.

4. Analytical Thinking:

- Following the six-step data analysis process (ask, prepare, process, analyze, share, act) provided a structured framework for my work. This methodical approach ensured that each step was thoughtfully executed and contributed to the overall success of the project.

5. Appreciation for LeBron's Career:

- Delving into LeBron James' career statistics gave me a deeper appreciation for his achievements and the factors that contribute to his sustained excellence. It reinforced the importance of dedication, adaptability, and leadership in achieving long-term success.

6. Practical Application of Skills:

- This project demonstrated the real-world relevance of these skills in sports analytics and beyond.

By reflecting on these lessons, I feel more confident in my ability to tackle future data projects and continue exploring the fascinating world of sports analytics.