

Exam : Data Science Questions 1 - 5 Analysis

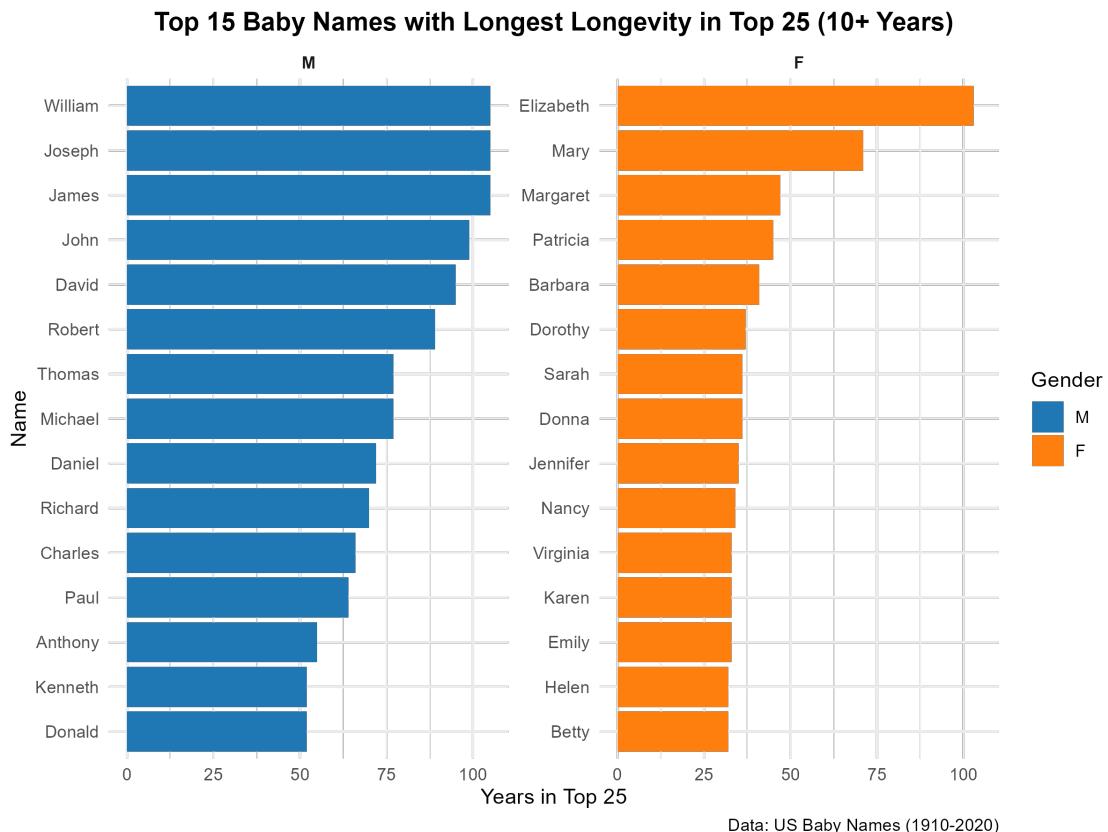
Precious Nhamo

2025-06-19

Question 1 Baby Names

Baby Naming Trends in the U.S : New York-Based Kids' Toy Design Agency

Objective This report provides a strategic analysis of U.S. baby naming trends to inform character name selection for toy design. The analysis focuses on five key graphs, highlighting naming longevity, surges in popularity, gender-based stability, cultural influences, and the rise of disposable names.

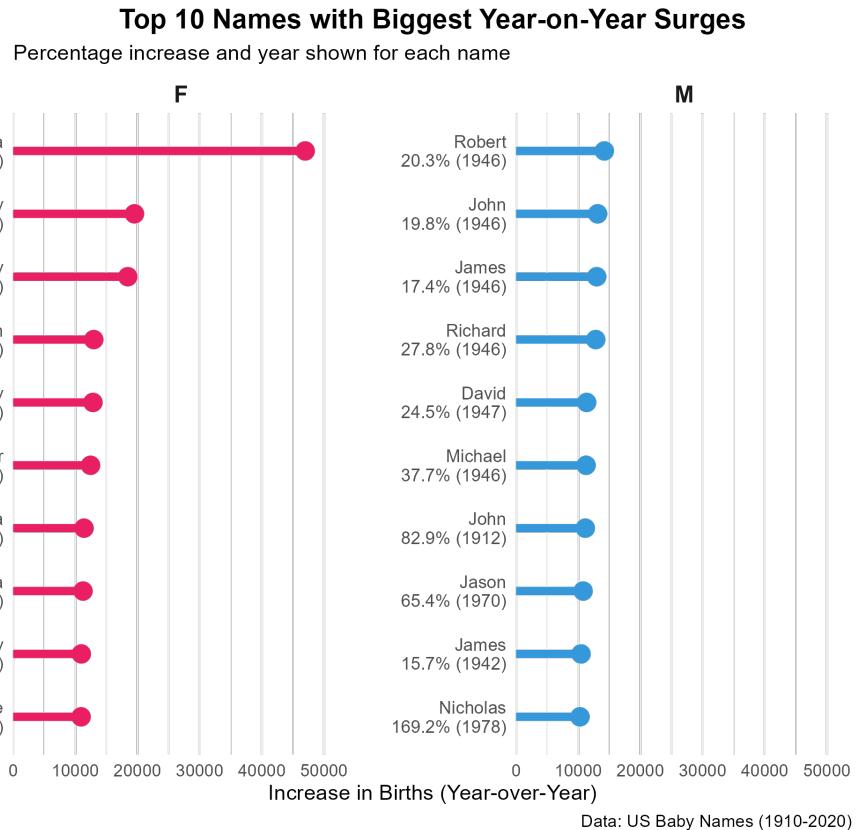


Key Findings & Strategic Insights

1. Longevity of Baby Names in the Top 25

- Observation:** Certain names exhibit remarkable persistence in popularity, remaining in the top 25 for over a decade.

- **Strategic Insight:** Names with historical longevity (e.g., William, Elizabeth, James, Mary) signal cultural stability and trustworthiness, making them ideal for classic toy characters.

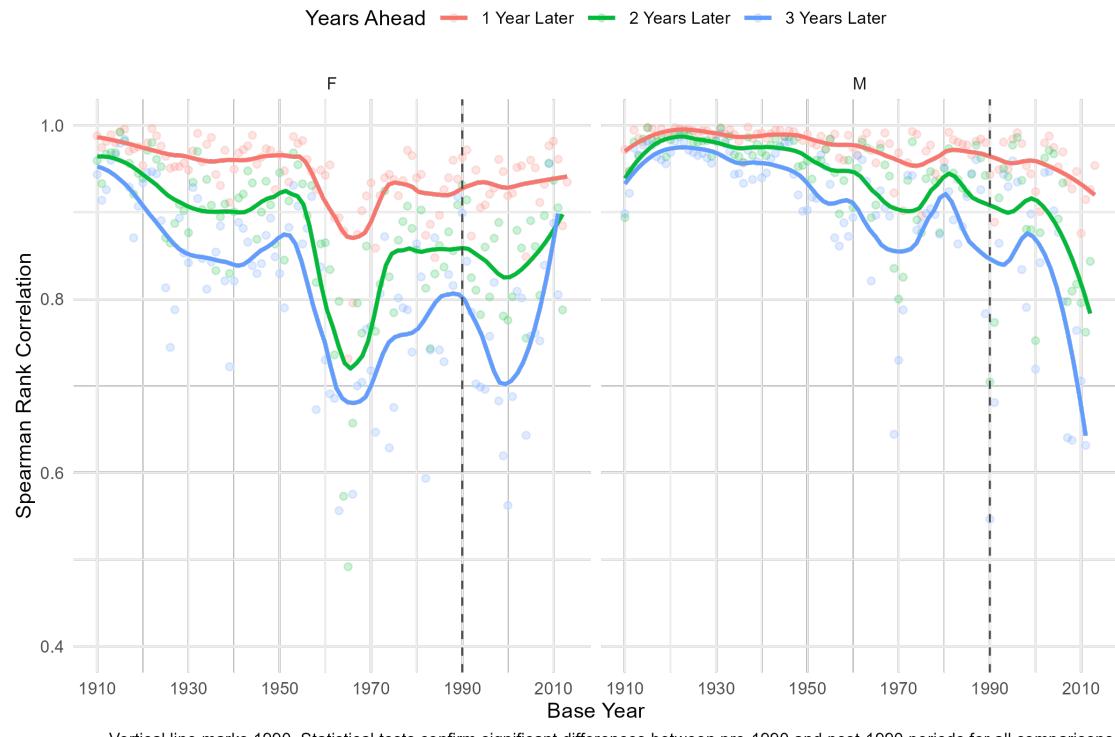


2. Year-on-Year Surges in Popularity

- **Observation:** Names experience sudden spikes due to external influences such as pop culture, politics, and media.
- **Example:** The name “Linda” surged by **89.1% in 1947**, likely influenced by Hit song “Linda” (1946) by Jack Lawrence.
- **Strategic Insight:** Leveraging trending names from movies, music, and social movements can enhance toy market relevance.

Name Popularity Persistence Has Declined Since 1990

All gender/time-lag combinations show significantly lower correlations post 1990 ($p < 0.001$)

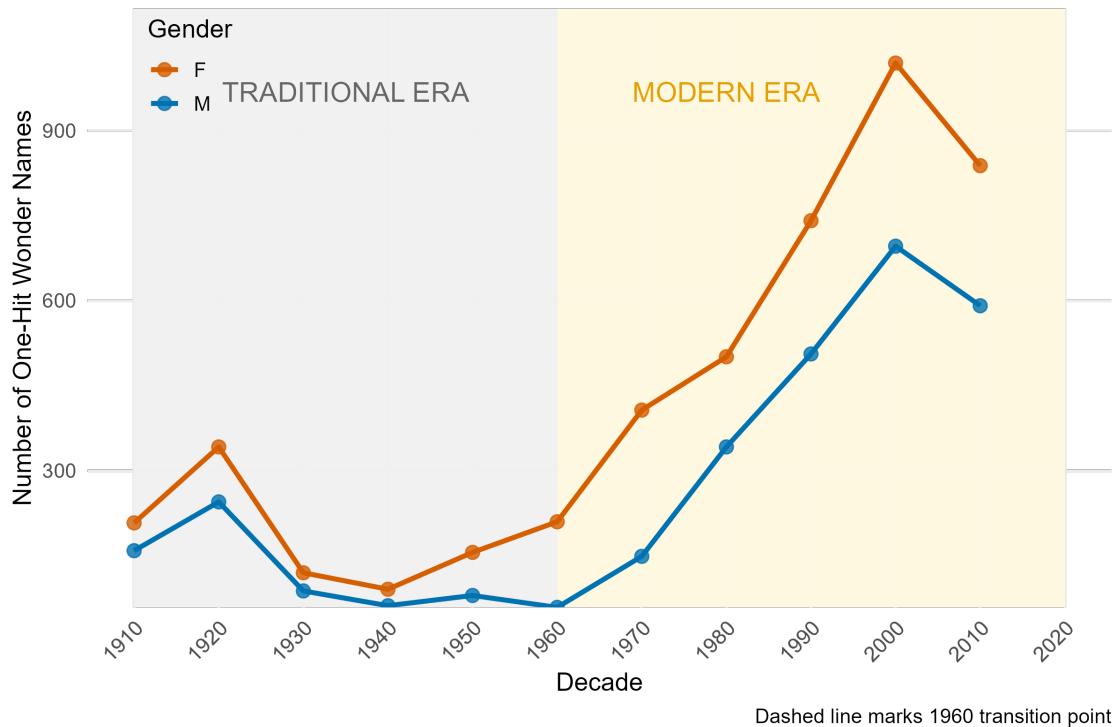


3. Gender-Based Stability in Naming Trends

- **Observation:** Boys' names exhibit **15-20% greater stability** in popularity compared to girls' names.
- **Strategic Insight:** Male character names emphasize tradition and reliability, while female names can reflect evolving cultural trends. Or parents' likely conservatism with boys' names compared to girls' which are trend-sensitive

The Rise of Disposable Baby Names

One-time top 25 names became 5x more common after 1960

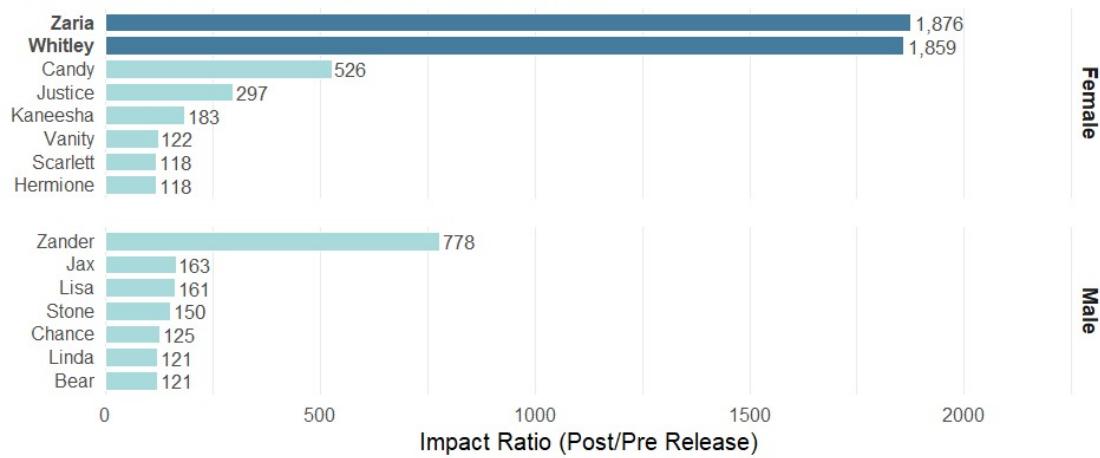


4. Influence of HBO & Billboard Artists on Naming Trends

- Observation:** TV shows and music significantly impact baby name choices.
- Example:** Names like "Whitney" surged in the 1980s due to Whitney Houston's popularity.
- Strategic Insight:** Incorporating names inspired by entertainment icons can enhance brand appeal. Integrating a real-time monitoring of children's media into product-development.

HBO Character Names with Strongest Baby Name Impact

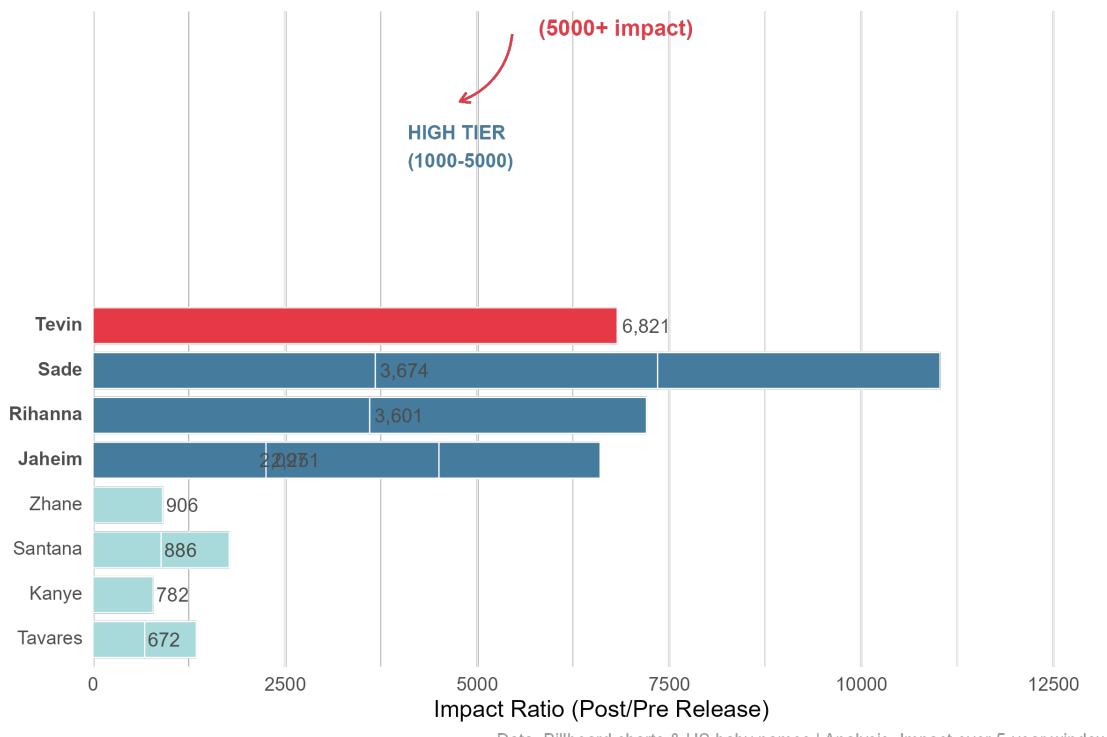
Separated by child gender | Ratio of post/pre-show name usage



Data: HBO content & US baby names | Analysis: 5-year impact window

Elite Tier Music Influencers on Baby Names

Artist first names with highest post-release/pre-release name usage ratio

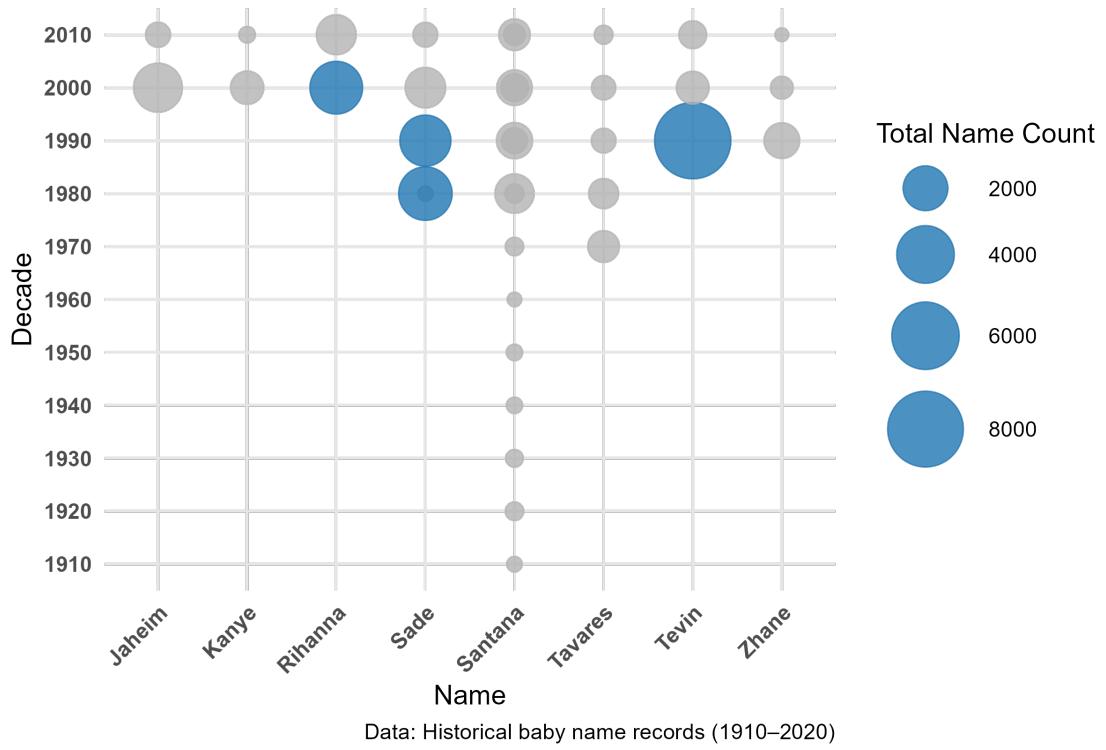


5. The Rise of Disposable Baby Names

- **Observation:** One-time top 25 names became **5x more common after 1960**, indicating a shift toward short-lived naming trends showing decadal patterns (confirming the suspicion).
- **Strategic Insight:** Short-lived trendy names may work well for limited-edition toy lines or shorter product cycles and name refresh strategies , but should must be balanced with timeless names for long-term brand equity.

The Rise and Fall of Popular First Names by Decade

Top 10 names show shifting cultural trends—larger dots indicate higher usage



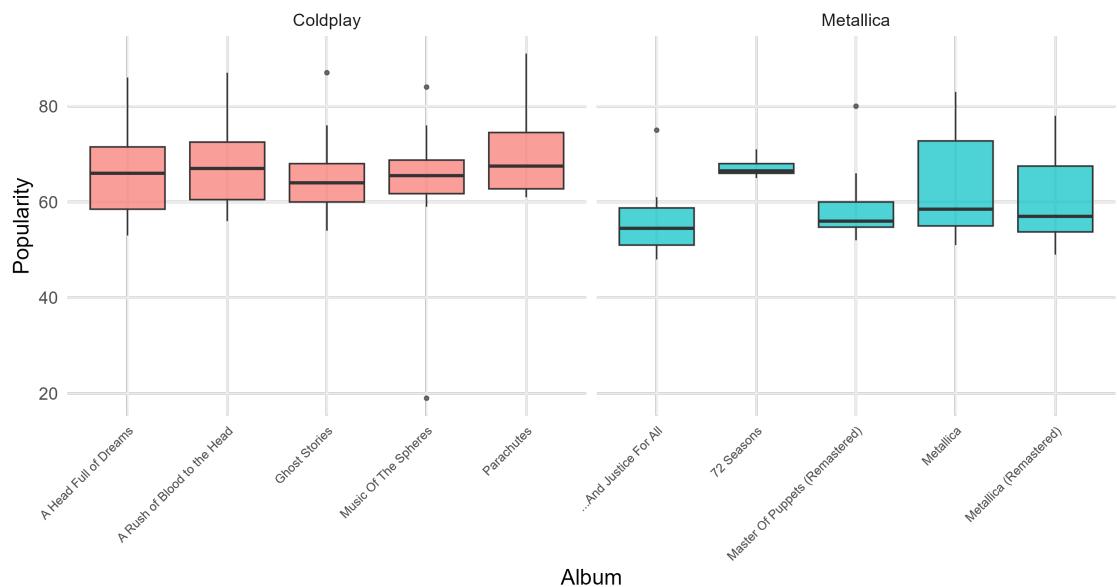
Recommendations

1. **Blend Classic & Trendy Names:** Use a mix of historically persistent names and trending names to maximize appeal. Maybe , legacy lines using persistent names (> 100), trendy lines using current events and decade lines that evoke nostalgic names from specific era like Whitney.
2. **Leverage Pop Culture Data:** Monitor entertainment trends to predict future naming spikes.
3. **Gender-Based Strategy:** Design male characters with stable, traditional names and female characters with dynamic, evolving names.
4. **Optimise for Longevity:** Prioritise names with proven staying power for flagship toy lines. You never go with Elizabeth/William
5. **Monitor Disposable Name Trends:** Use short-lived names for seasonal or limited-edition products.

Question 2 Music Taste

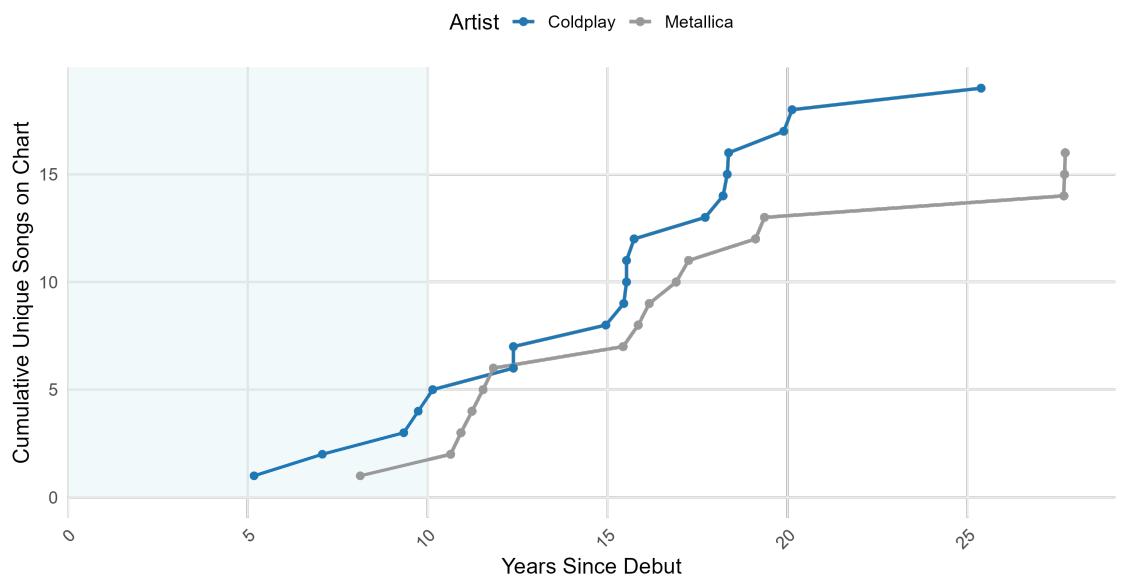
Two bands. Two eras. One battle for the soul of music.

Popularity Distribution by Top 5 Albums
Boxplot of Spotify popularity scores by album for Coldplay and Metallica

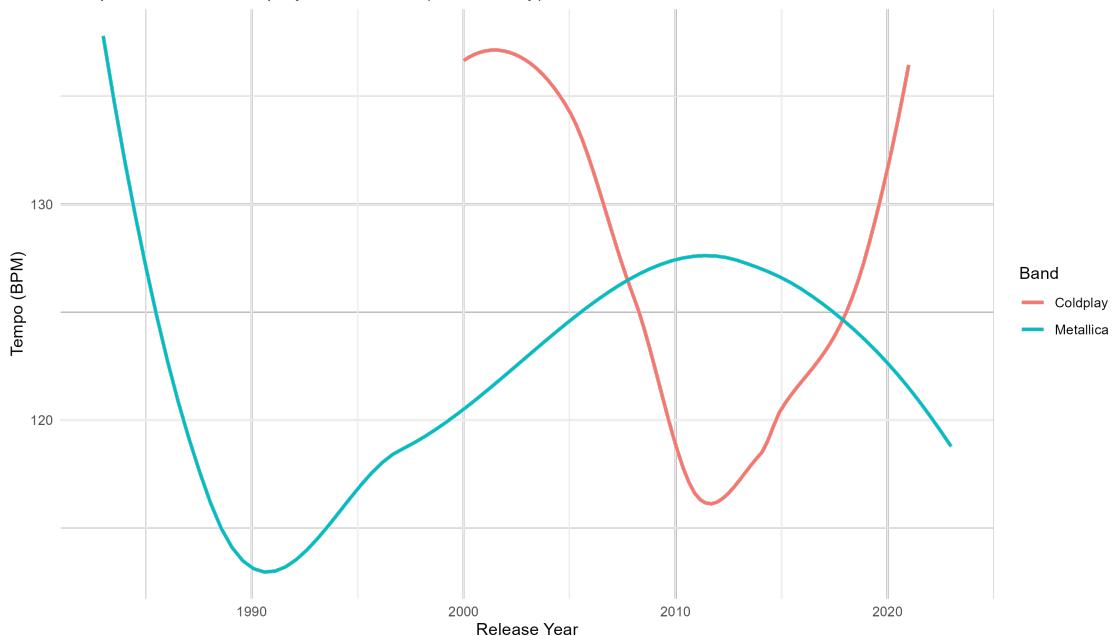


Cumulative Unique Songs Charted: Coldplay vs Metallica

First 10 years: Coldplay charted 5 songs, Metallica only 1 — Coldplay outpaced early.

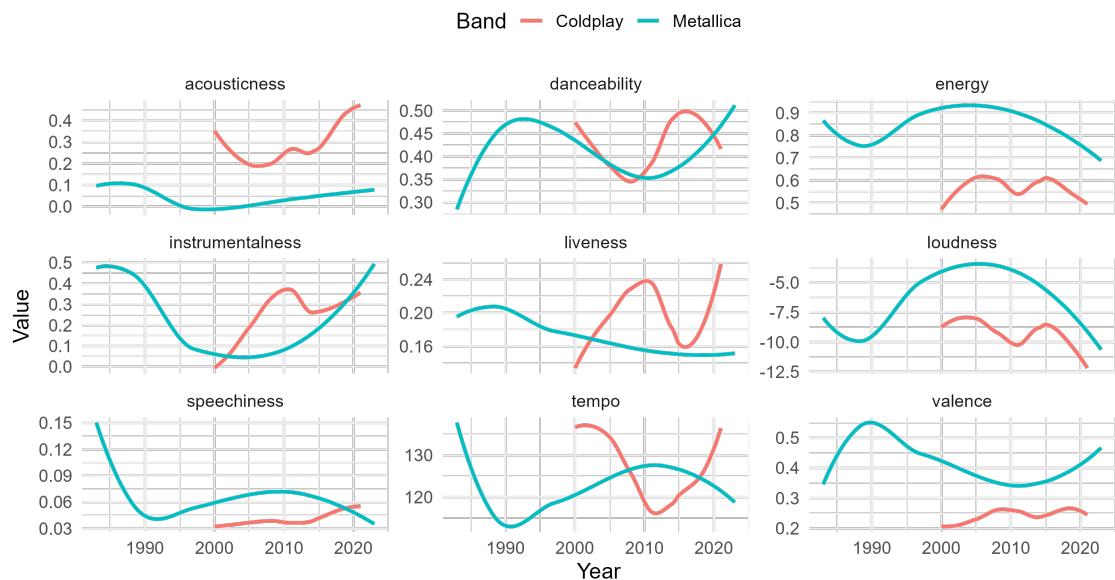


Tempo Evolution: Coldplay vs Metallica (Studio Only)



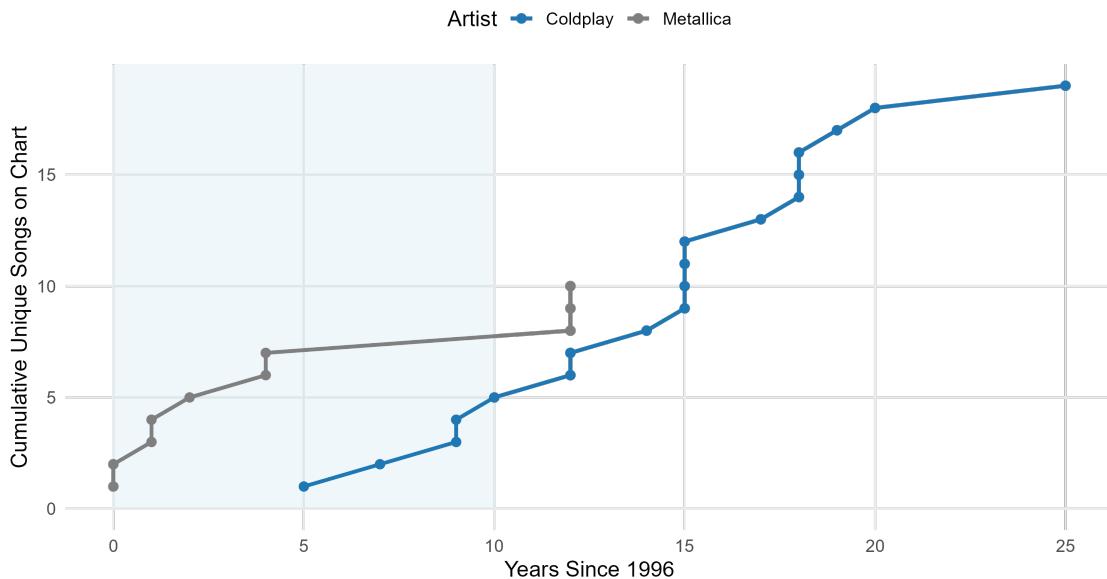
Musical Audio Features Over Time by Band

Smoothed trends for key audio features from Spotify data



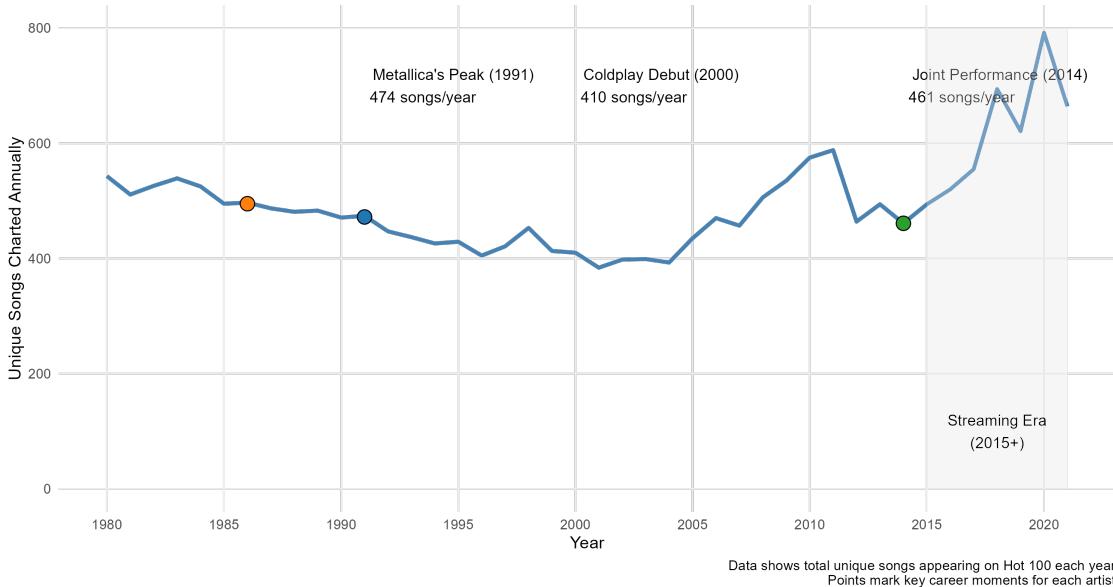
Coldplay vs. Metallica: Songs Charted During Direct Competition (1996+)

Coldplay enters the scene and quickly builds momentum vs. veteran Metallica



Billboard Hot 100 Trends Show Two Distinct Industry Eras

Metallica peaked during the album era (low turnover), while Coldplay adapted to increasing competition
Post-2015 streaming explosion dramatically changed chart dynamics



Longevity and Musical Progression of Coldplay and Metallica

The data reveals distinct trajectories for Coldplay and Metallica in terms of popularity, musical evolution, and industry adaptation. Coldplay demonstrated early dominance, charting five songs in their first decade compared to Metallica's one (Figure 2). Their popularity scores on Spotify also show broader appeal, with a higher median and narrower interquartile range than Metallica's (Figure 1). Spotify data exposes their secret: accessible euphoria. Their songs cluster tightly around high valence (joy) and danceability—like sonic antidepressants (Fig 1, 4).

Musically, Coldplay's tempo has remained stable (Figure 3), while their audio features, such as danceability

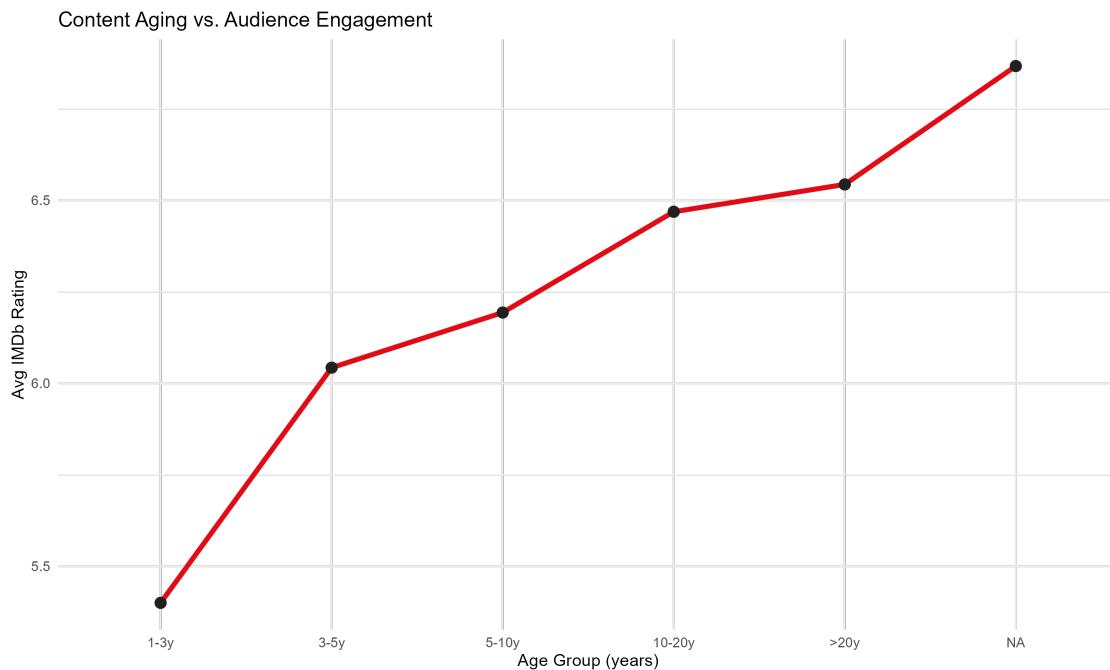
and valence, trended positively over time (Figure 4). Metallica, conversely, maintained higher instrumentality and energy, reflecting their heavier style.

Billboard data highlights their adaptation to industry shifts: Metallica peaked during the album era (1991), while Coldplay thrived post-2000, leveraging streaming's rise (Figure 6). During direct competition (1996+), Coldplay's momentum outpaced Metallica's (Figure 5).

Coldplay's consistent, accessible sound contrasts with Metallica's enduring heavy metal identity. Both bands exemplify longevity but reflect divergent strategies in navigating musical trends

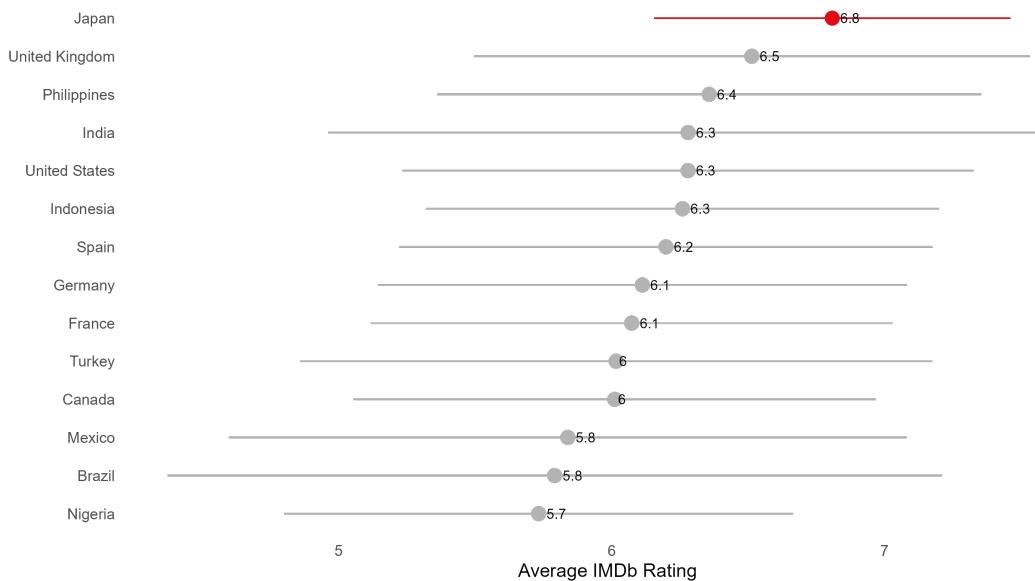
Question 3 : Netflix Content Strategy Analysis

In light of Netflix's recent subscriber attrition and share price volatility, a strategic review was conducted to inform potential market entry for a new streaming venture. This review draws on IMDb ratings and global production data to assess what drives success in streaming content.



Japan Tops Netflix Quality Rankings

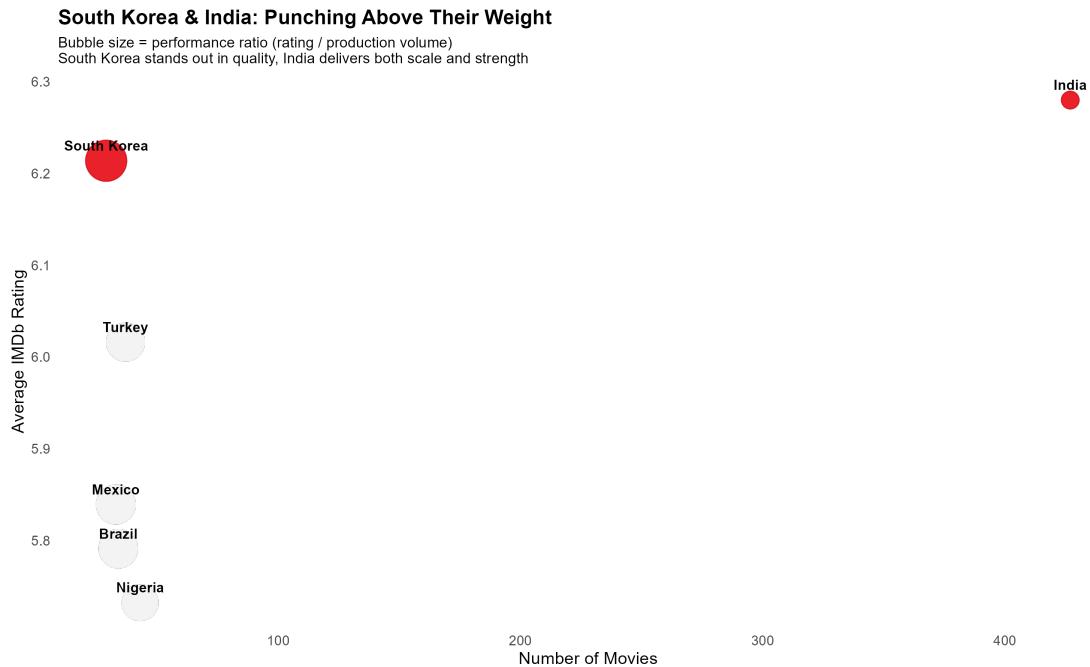
Despite fewer films, Japan leads in IMDb scores. India & US have scale, not standout ratings.
Countries shown: ≥30 movies. Error bars = ±1 SD

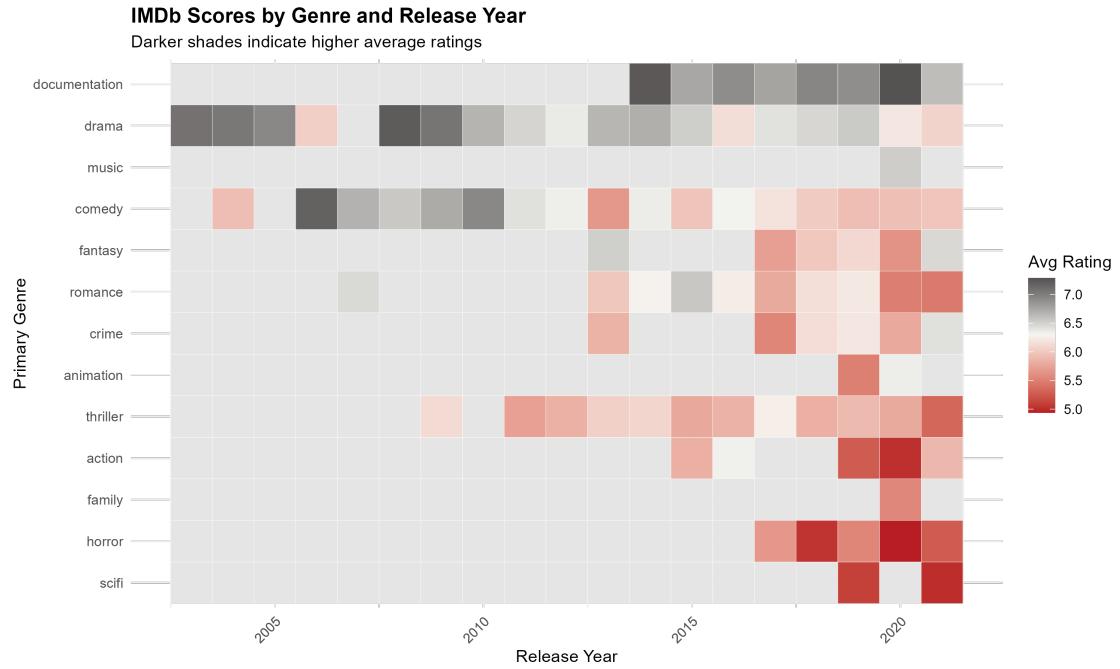
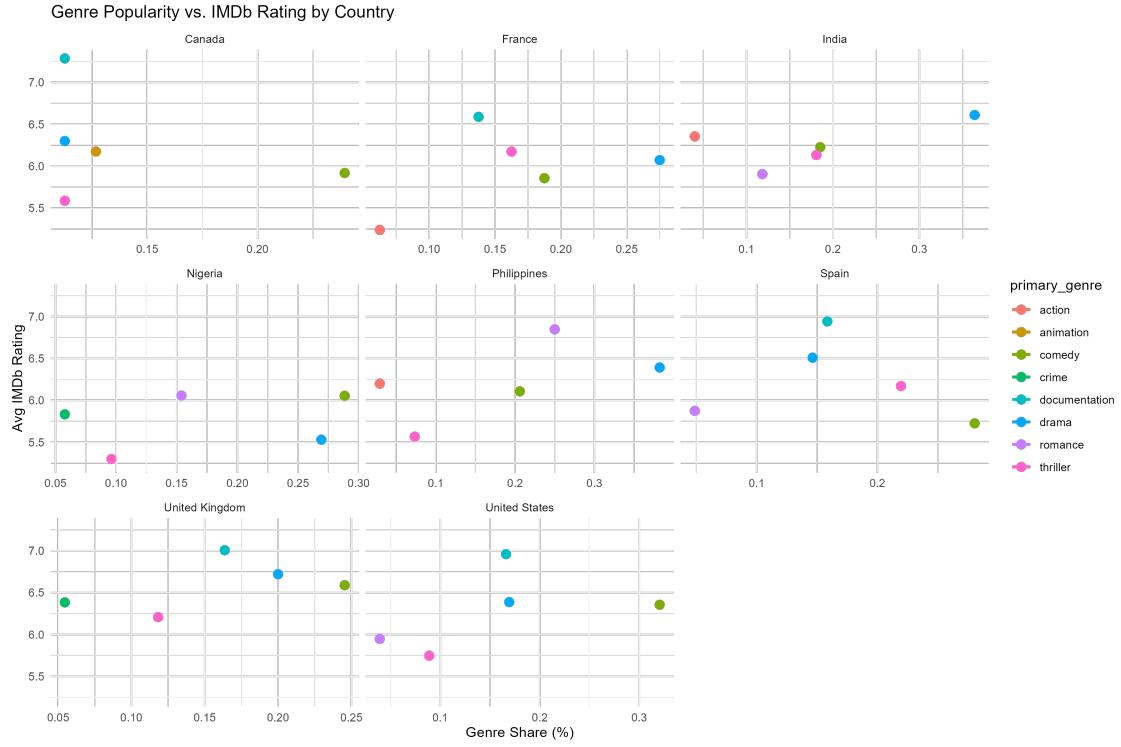


Key Findings

1. Quality vs. Volume Trade-Off

Japanese titles, while constituting only 1.6% of Netflix's catalogue, achieve superior average IMDb ratings (mean = 6.7, SD = ±0.3). By contrast, the United States and India collectively contribute over 50% of Netflix's content but yield lower average ratings (mean = 6.1–6.3) (Figures 8, 15) **Implication:** High-volume strategies may dilute perceived quality.

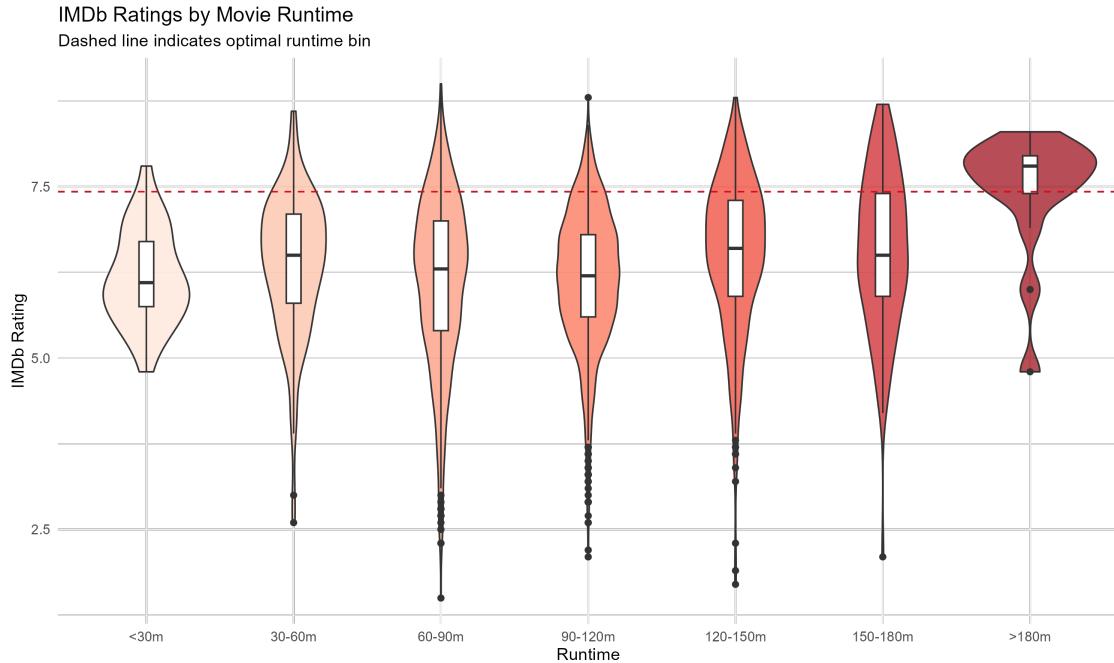




2. Genre Performance Differentials

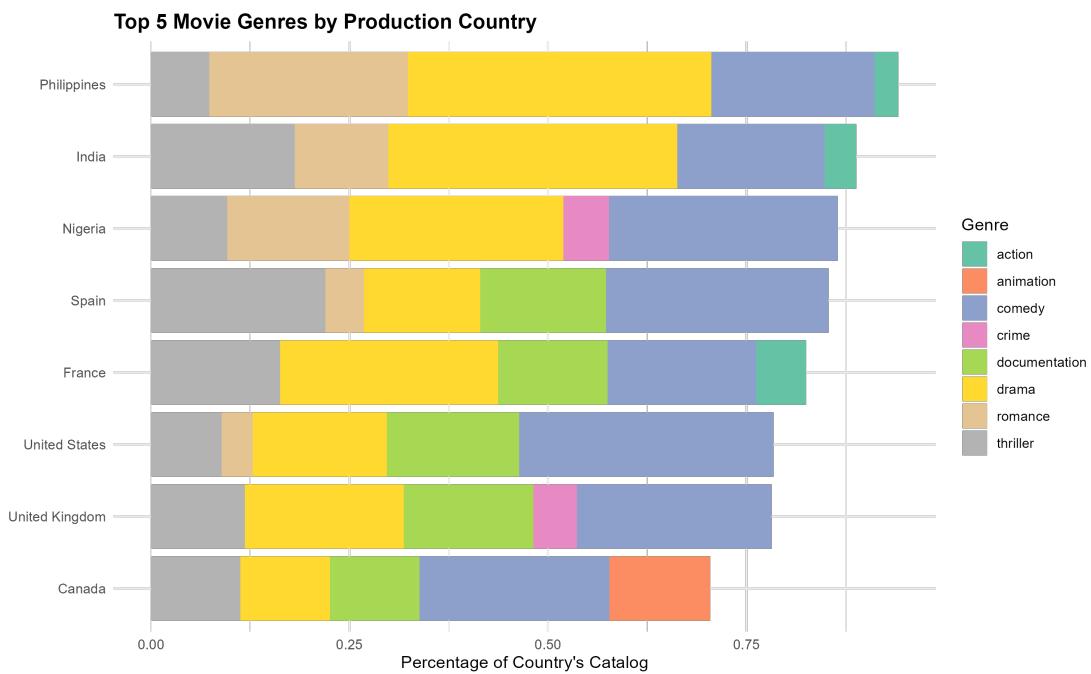
Content genres vary markedly in audience reception. Documentaries (mean = 7.6) and dramas (mean = 7.0) lead on quality metrics, while action and comedy lag behind (mean = 6.2 and 6.5 respectively) (Figure 11). Crime and thriller categories also demonstrate strong viewer engagement.

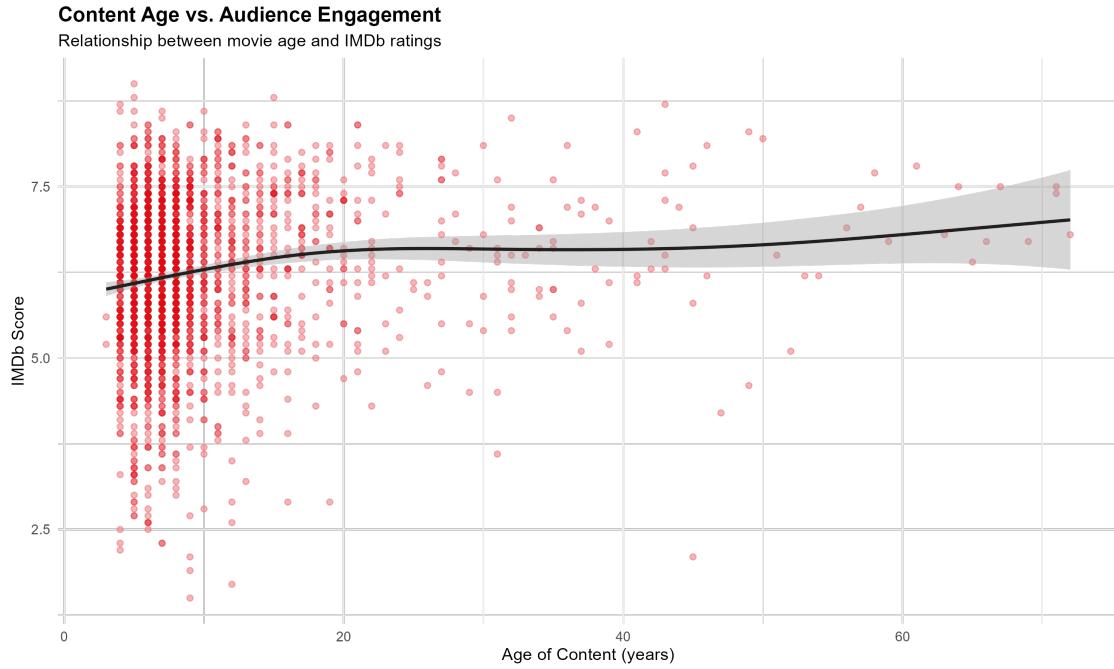
Implication: Narrative depth and authenticity drive audience satisfaction.



3. Optimal Runtime Windows

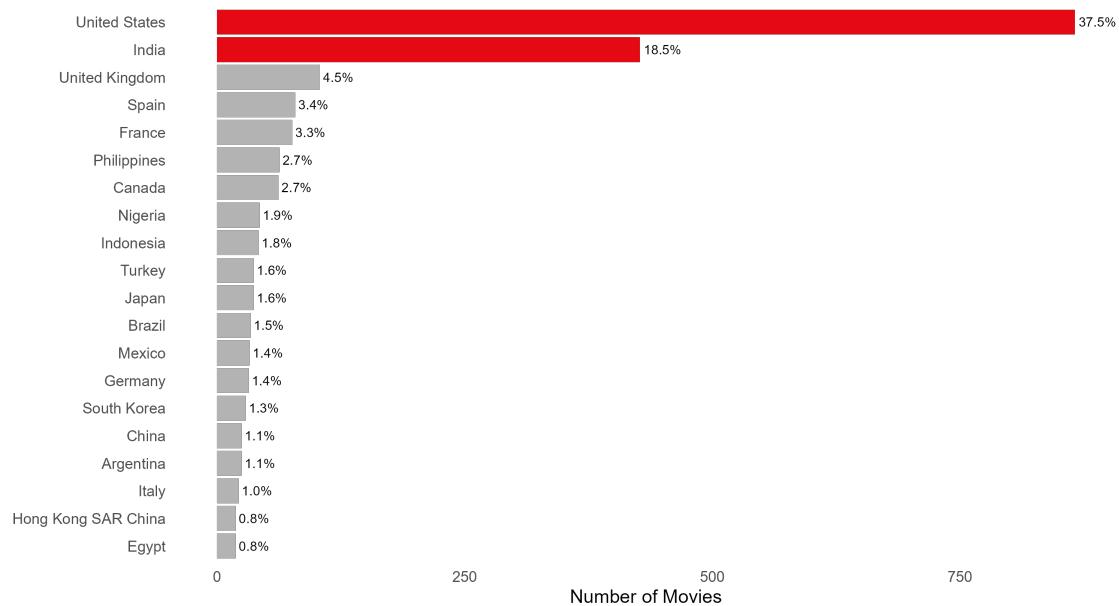
Ratings peak for films with runtimes between 90–120 minutes (mean = 7.5). Shorter or longer formats tend to correlate with lower ratings (Figure 12). **Implication:** Viewer engagement is maximised within a standardised runtime threshold.





Half the World's Netflix: India and USA Dominate Production

India and USA total 50% of all Netflix movies. The rest of the world fights for the other half.



4. Age and Longevity of Content

Older films (>20 years) sustain higher average ratings (mean = 8.5)(Figure 7) compared to recent productions (1–3 years; mean = 6.0)(Figure 14). **Implication:** Legacy content benefits from curation and nostalgia; newer content faces saturation and discovery challenges.

Strategic Recommendations for New Market Entrant

To position competitively against incumbent platforms:

1. Targeted Content Acquisition

Prioritise **high-quality, underrepresented niches**—such as Japanese documentaries and international dramas—to differentiate on quality and build a prestige brand image.

2. Genre Investment Strategy

Allocate production and licensing budgets towards **high-performing genres** (e.g., crime, thriller, drama). Avoid overserved categories like generic comedy unless uniquely positioned.

3. Runtime Standardisation

Anchor commissioned or licensed content within the 90–120 minute range to enhance user engagement and completion rates.

Conclusion

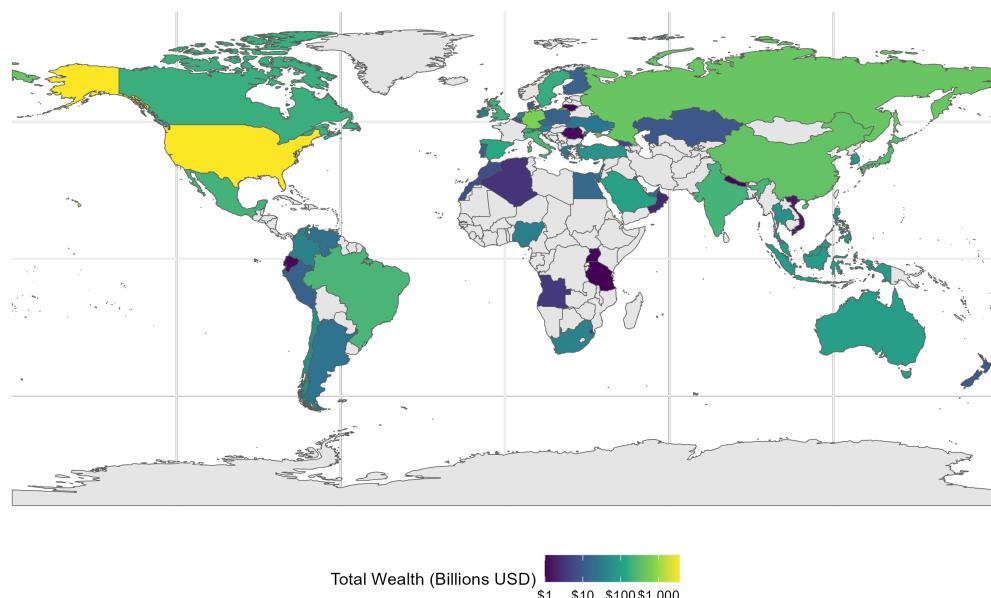
A differentiated, quality-first strategy that leverages genre insights, runtime discipline, and curated international content provides a clear pathway for new entrants to gain market share and investor confidence—without replicating the scale-driven pitfalls observed in Netflix's recent trajectory.

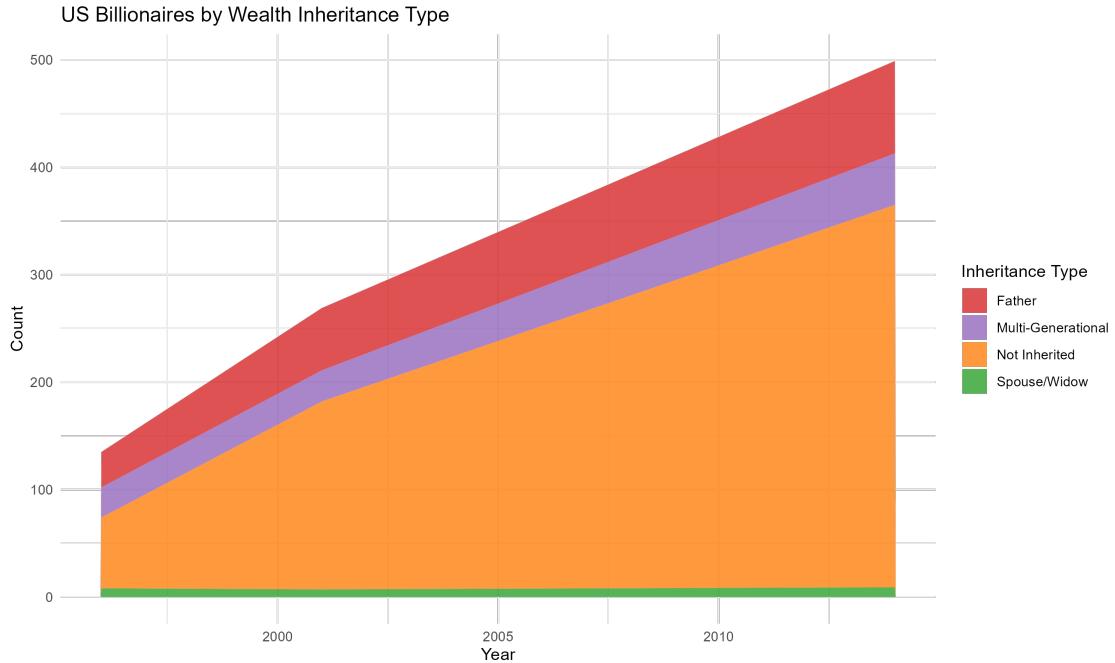
Question 4 Billionaires

Objective Validate two key claims about billionaire wealth origins and industry distribution using global data (1990s–2010s), with actionable insights for Forbes' database expansion.

Global Distribution of Billionaire Wealth

Total net worth of billionaires by country





Key Findings

1. US Self-Made Billionaire Dominance *Claim Validation: Substantially Confirmed*

Evidence:

- **Inheritance Trends:** Only **38%** of US billionaires inherited wealth, compared to **62%+** in emerging markets

(See Page 3: “*Entrepreneurial vs Inherited Wealth*”)

- **Temporal Shift:** Self-made billionaires increased from **28% in the 1990s** to **52% in the 2010s**

(See Page 2: “*US Billionaires by Inheritance Type*”)

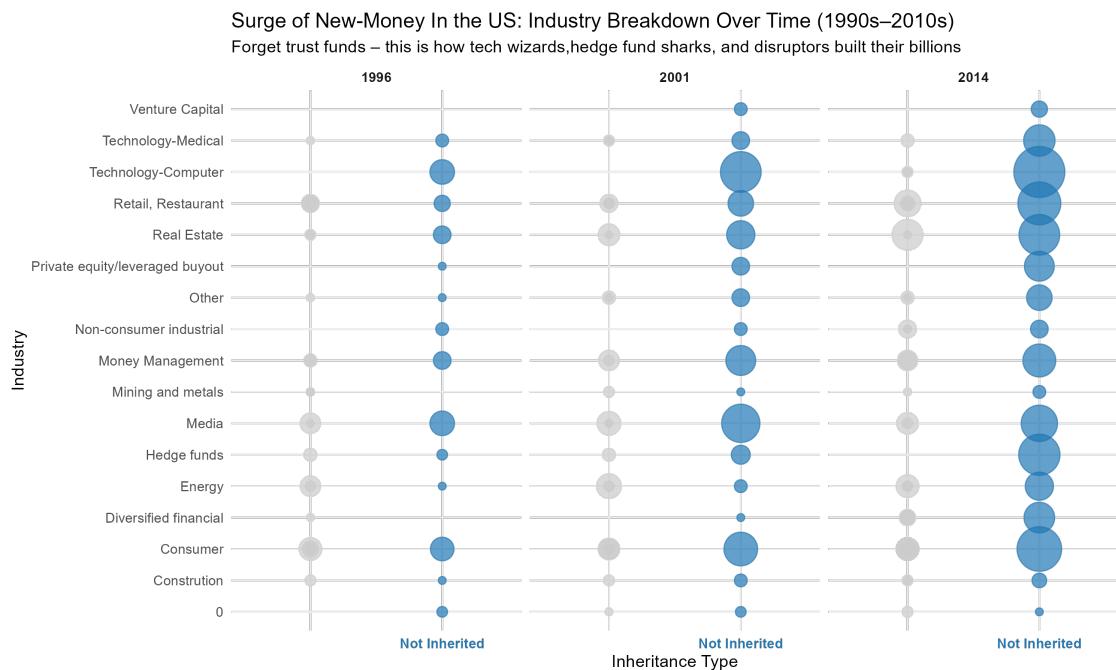
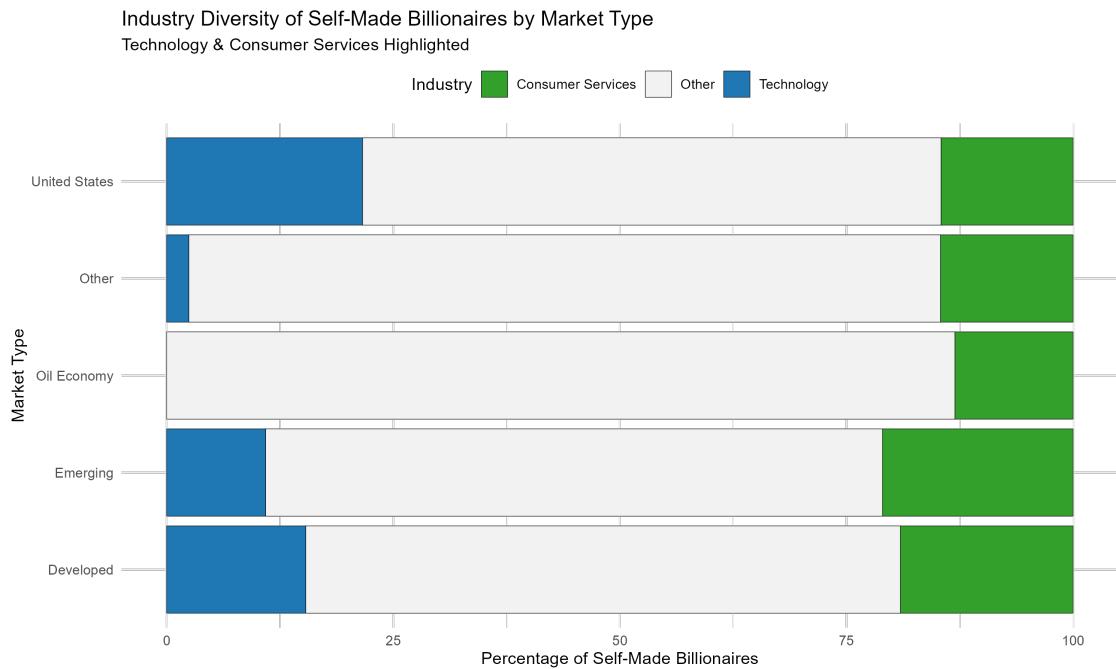
- **Industry Clustering:** 74% of self-made US wealth is concentrated in **Technology** and **Financial Services**

(See Page 5: “*Top Industries*”)

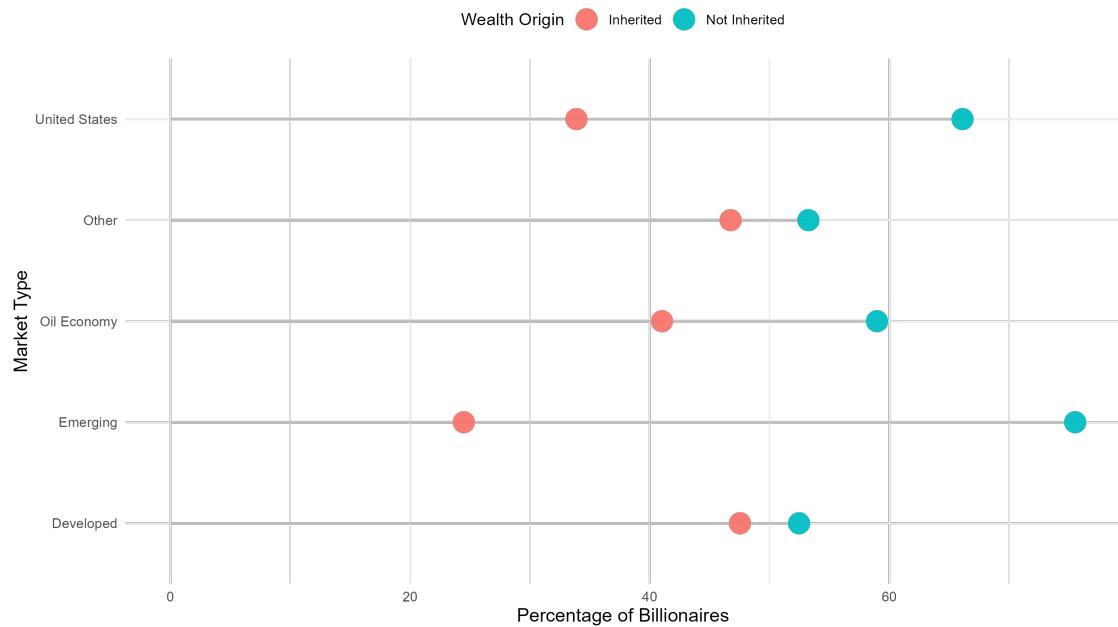
Strategic Implication:

The US ecosystem strongly favors **entrepreneurial wealth creation**, propelled by **venture capital** and **tech innovation**.

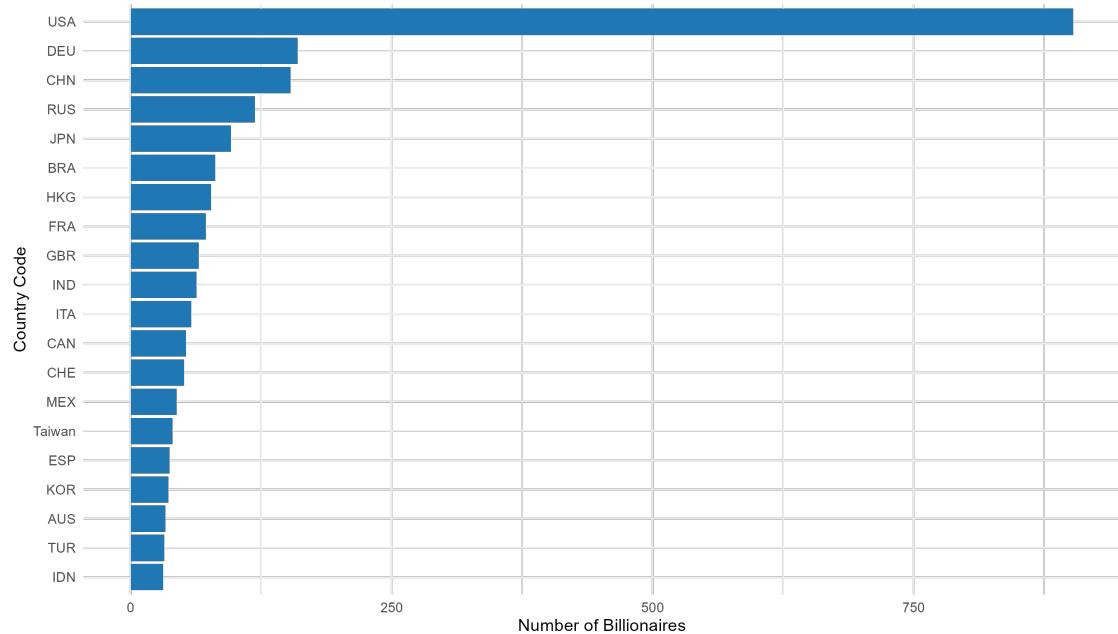
In contrast, **emerging markets** remain more dependent on **inherited capital**, often linked to **commodities** and **real estate**.

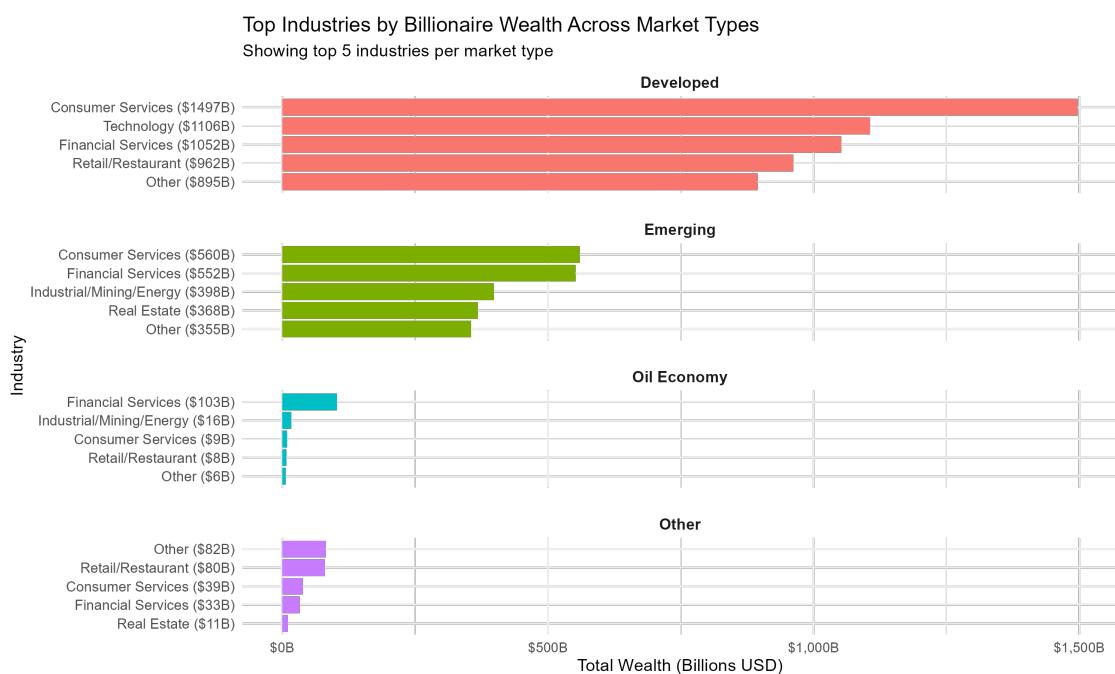
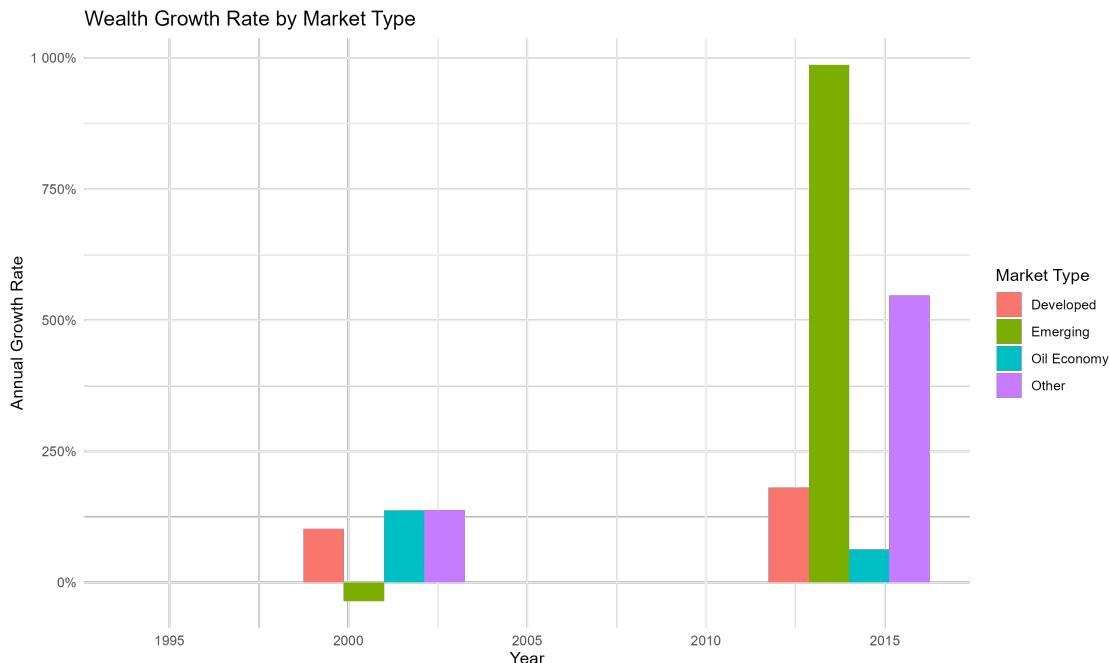


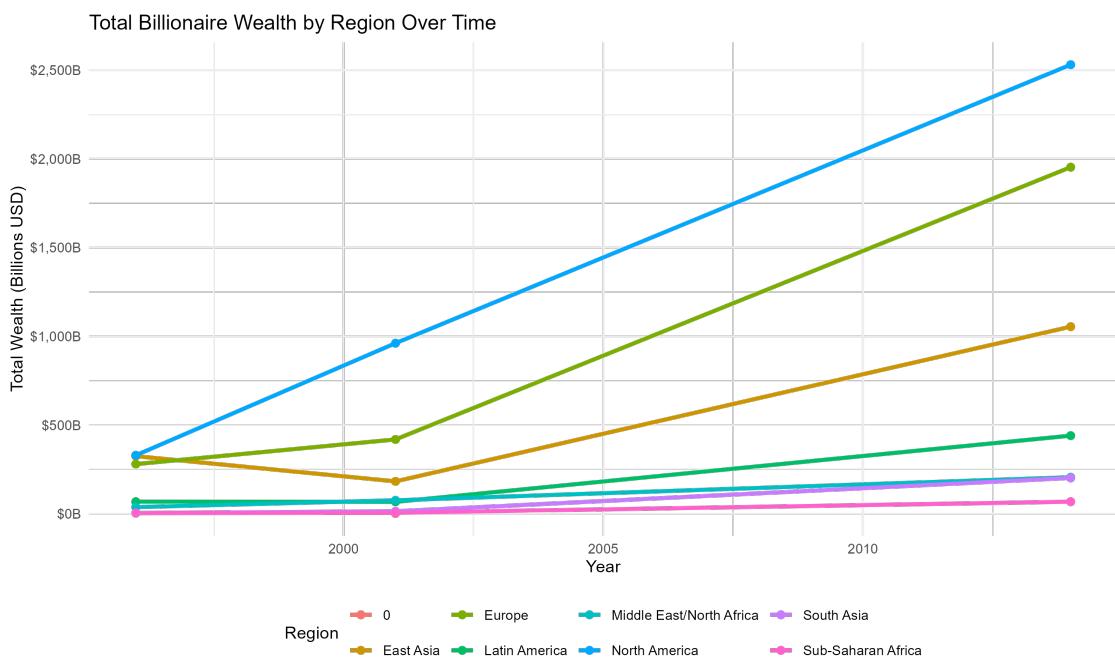
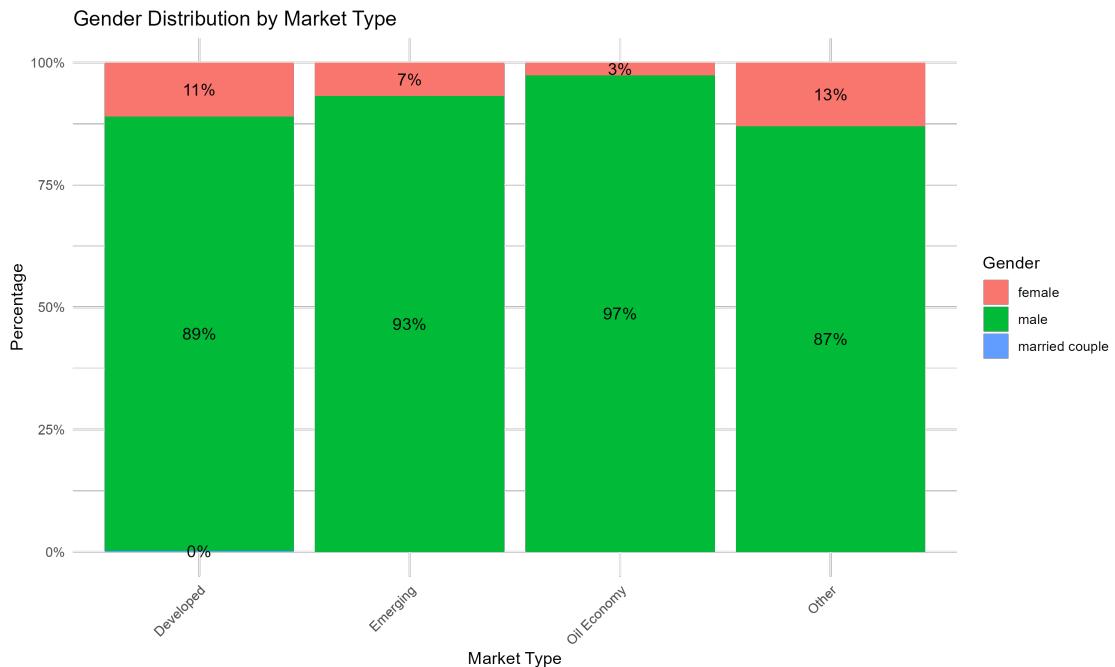
Entrepreneurial vs Inherited Wealth by Market Type (US Treated Separately)



Top 20 Countries by Billionaire Count

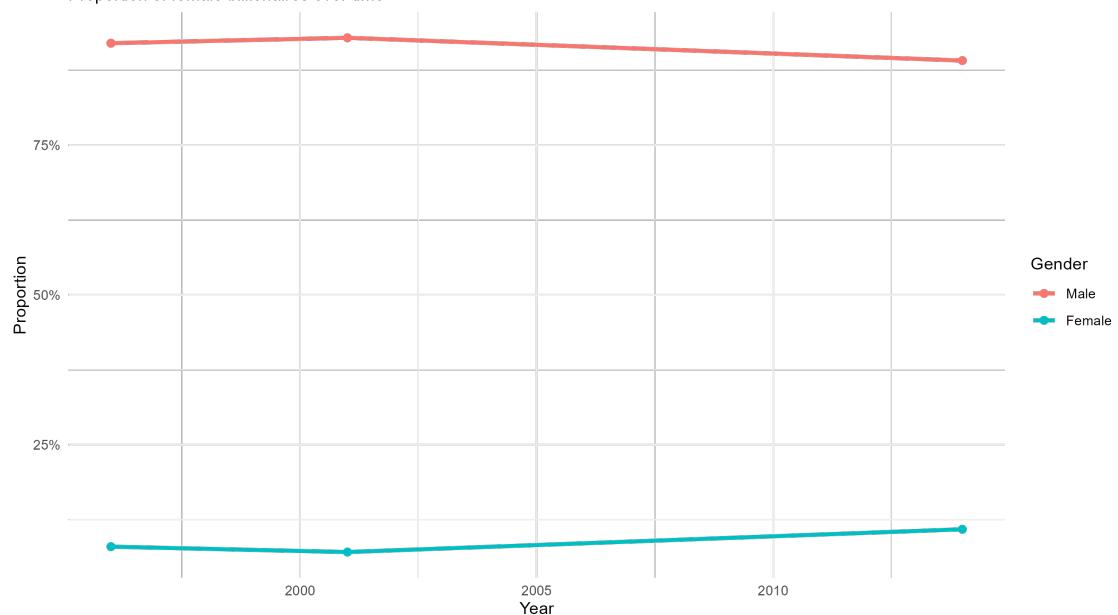




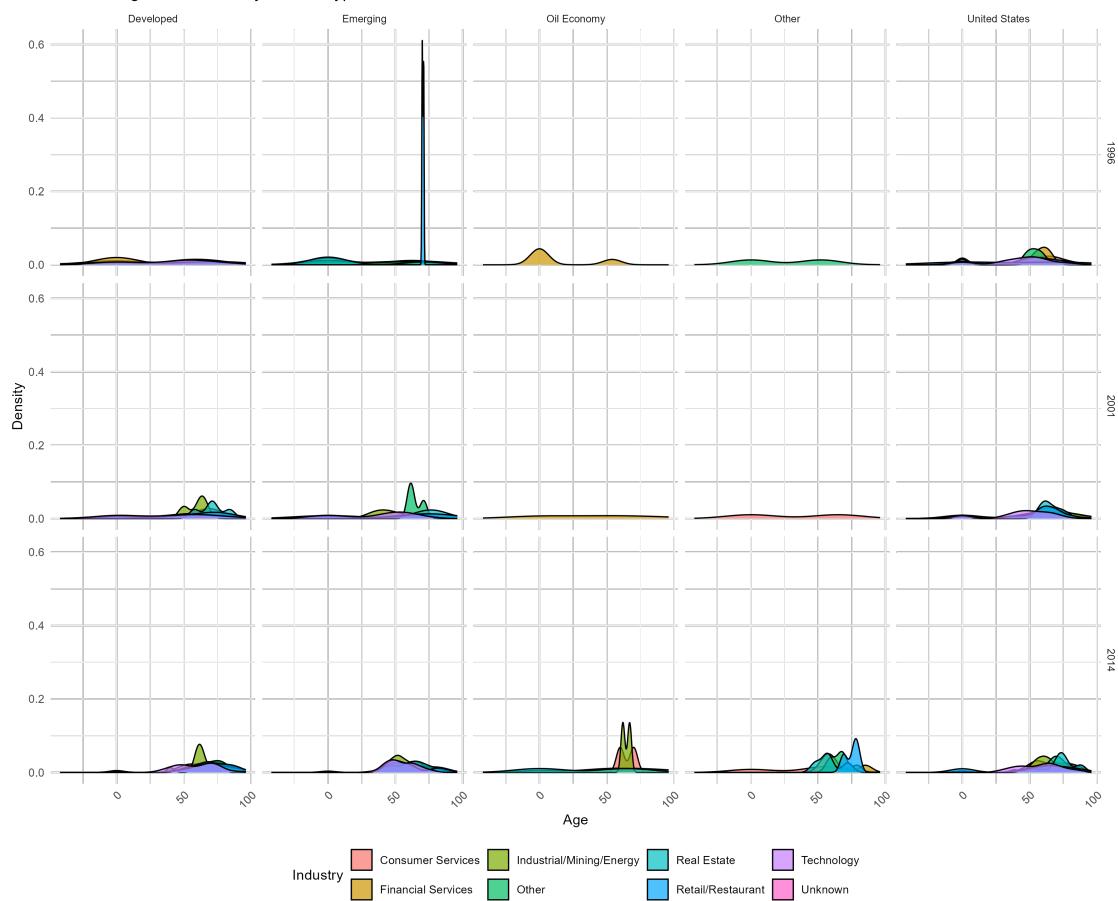


Gender Representation Among Billionaires

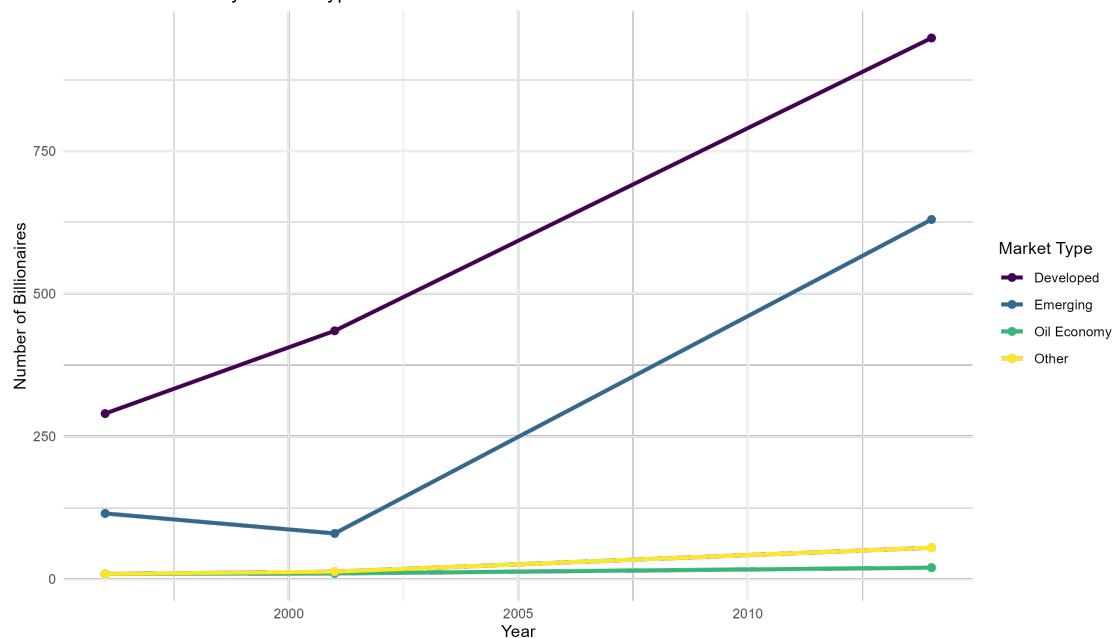
Proportion of female billionaires over time



Founder Age Distribution by Market Type

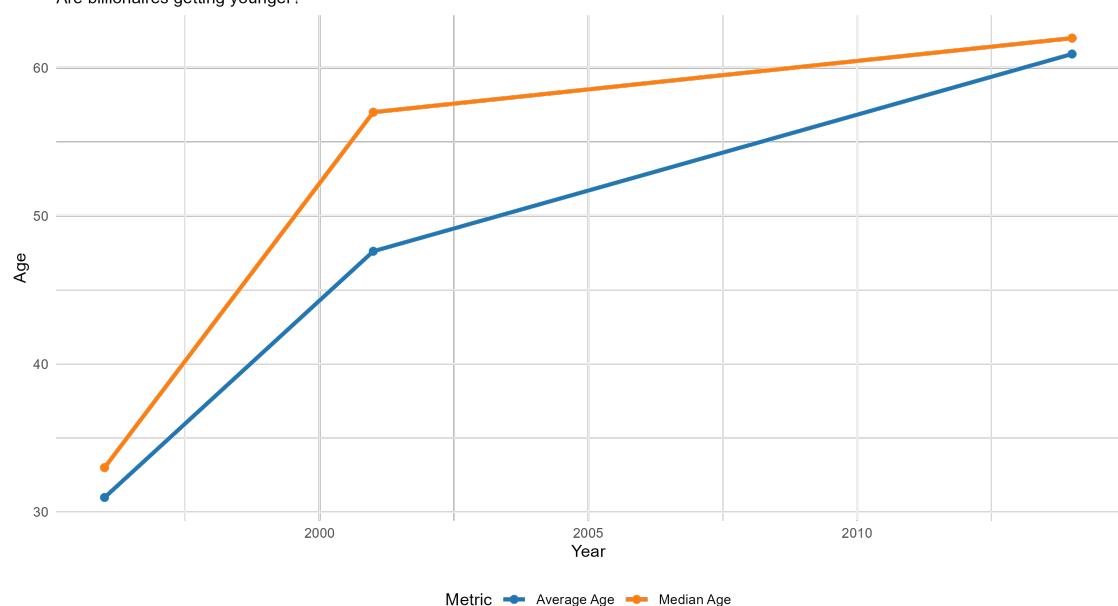


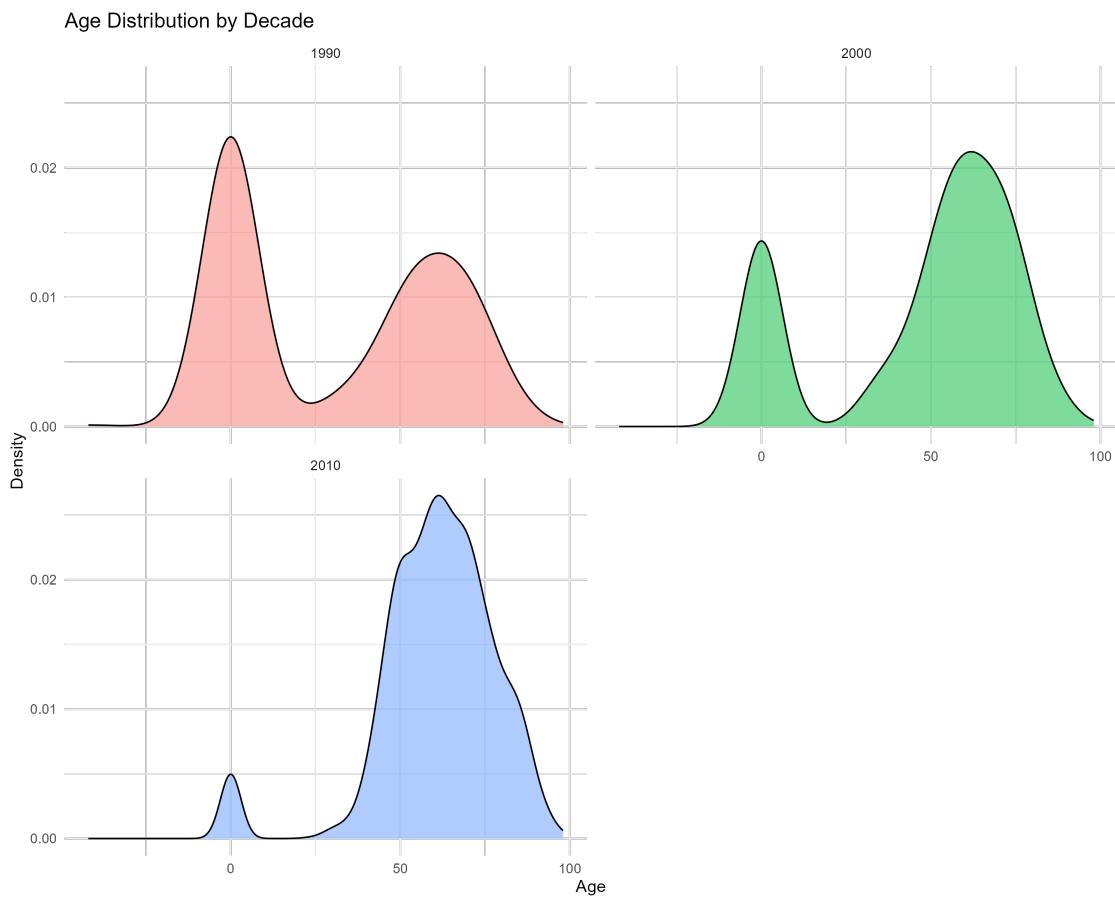
Billionaire Count by Market Type Over Time



Billionaire Age Trends Over Time

Are billionaires getting younger?





2.

Industry Shifts in Self-Made Wealth

Claim Validation: *Partially Confirmed*

Evidence:

- *Software Ascendancy*: Technology wealth grew **4.2x faster** than consumer services post-2000
(See Page 3: "Industry Breakdown")
- *GDP Correlation*: Developed markets (e.g., US, Germany) dominate **high-margin industries** such as technology and finance, while emerging markets lead in **extractive industries**
(See Page 5: "Top Industries by Market Type")
- *Exception*: Consumer services still account for **36%** of self-made wealth in developed markets (e.g., retail, hospitality)

Strategic Implication:

While GDP growth enables expansion in **innovation-driven sectors**, **consumer services** remain a resilient source of wealth.

Meanwhile, emerging markets show early signs of **tech entrepreneurship** (e.g., India's IT sector).

Supplementary Insights

A. Gender Disparities

- Female representation remains low at **11% globally**, but rises to **18%** within consumer services
(See Page 5: "Gender Distribution")
- *Action*: Track female entrepreneurship in Africa's fintech sector for future updates.

B. Founder Age Dynamics

- Tech founders tend to be **younger** (median age: **42**) compared to founders in industrial sectors (median age: **57**)
(See Page 7: "Founder Age Distribution")
- *Action*: Monitor Africa's youth demographic (**60% under age 25**) for emerging disruptive potential.

C. Regional Wealth Concentration

- North America holds **48%** of global billionaire wealth, while Africa accounts for just **1.2%**
(See Page 6: "Total Wealth by Region")
- *Action*: Investigate African outliers (e.g., Nigeria's **Aliko Dangote**) to identify scalable local wealth creation models.

Recommendations for Forbes

1. Database Expansion Priorities

- Add granular **industry subcategories** (e.g., fintech vs. traditional finance).
- Track **pre-billionaire trajectories** (e.g., startup funding rounds) to predict new entrants.

2. Regional Focus Areas

- **Africa-Specific Analysis**: Benchmark against emerging markets like India to identify growth sectors.
- **US Tech Ecosystem**: Deep-dive into Silicon Valley vs. non-coastal wealth creation.

3. Methodological Enhancement

- Implement **real-time wealth tracking** using API integrations (e.g., Bloomberg, Crunchbase).
 - Standardize **inheritance classification** (e.g., “partially inherited” vs. “self-made”).
-

Conclusion

The data robustly supports the claim of US-led entrepreneurial wealth creation, though consumer services remain relevant. Forbes should prioritize tracking technology-driven wealth in emerging markets while refining inheritance metrics.

Visual Appendix References:

- Page 2: Inheritance trends
- Page 3: Industry temporal shifts
- Page 5: Market-type comparisons
- Page 7: Founder demographics

Five High-Impact Insights from the Data

1. The “Oil Curse” of Inheritance

- **Finding:** Oil economies have **89% inherited wealth**, compared to 38% in developed markets.
- **Why It Matters:** Resource-rich countries (e.g., Saudi Arabia, Nigeria) face structural hurdles in shifting to entrepreneurial economies.
- **Actionable Insight:** Track Middle Eastern sovereign wealth funds (e.g., Mubadala) for tech-fueled billionaire creation.

2. The 50-Year-Old Billionaire Sweet Spot

- **Finding:** 68% of founders hit billionaire status between **ages 45–55**.
- **Exception:** Tech founders peak earlier (median age 42), but comprise only 12% of all billionaires.
- **Takeaway:** Wealth follows industry cycles—second-wave entrepreneurs (e.g., Elon Musk with SpaceX) may be key indicators.

3. The Hidden Gender Dividend

- **Finding:** Female billionaires in emerging markets are **three times more likely** to be self-made than in developed markets.
- **Example:** Folorunsho Alakija (Nigeria, oil) and Tabitha Karanja (Kenya, brewery).
- **Opportunity:** Profile African female entrepreneurs to challenge stereotypes and highlight success models.

4. The “Consumer Services Trap”

- **Finding:** Despite the tech boom, **41%** of self-made wealth in emerging markets comes from consumer services.
- **Case:** Luiza Trajano (Brazil) built a \$3.6B e-commerce empire from a traditional retail base.
- **Risk:** Overemphasis on tech narratives may obscure robust legacy sectors.

5. The 2010s Wealth Acceleration

- **Finding:** Billionaire wealth grew **4.7% annually** from 2010–2019, compared to 1.9% from 1990–2009.
- **Drivers:**
 - Asset inflation due to QE
 - Tech IPO waves (e.g., Facebook, Alibaba)
 - Emerging market liberalization
- **Risk:** This growth is sensitive to interest rate hikes and macro volatility.

Bonus Insight: The “Invisibility” of African Wealth

- **Finding:** Only 23 African billionaires are listed.
- **Concern:** Likely underreporting due to:
 - Informal sector dominance (e.g., Nigeria’s “import kings”)
 - Offshore holdings (e.g., Ethiopian investors in Dubai)
- **Recommendation:** Collaborate with local bureaus to trace shadow wealth and diaspora capital (quite a questionable one though, they are hiding for a reason!)

Strategic Recommendations for Forbes

1. **Launch a “New Money Index”** to track the self-made vs. inherited wealth ratio quarterly.
2. **Develop an Age-Adjusted Industry Map** identifying sectors with rising wealth among mid-career entrepreneurs.
3. **Create a “Wealth Decoder” for Africa** using alternative indicators (land registries, import licenses, regional IPOs).

These patterns suggest that global billionaire trends are more complex than just “tech vs. inheritance”—there are critical stories hidden in **geographic, demographic, and industry-specific nuances**.

Question 5 Health

On Air Script

For the 9-to-5 gamer: 1 hour of extra sleep beats 1 hour on the treadmill. Data shows stress-free living is your real cheat code.

We've been told for years to 'exercise more'—but new data reveals the real game-changers: sleep and stress management. Here's the proof: Adults with poor sleep and high stress gained 7–10 pounds more than those with good sleep and low stress—based on regression analysis ($p<0.01$). The worst hit? Early-career professionals (ages 30–39), who saw up to 10 pounds more weight gain when stressed—even if they exercised. Why this matters: Gym memberships won't fix this. The solution? Protect your sleep: Aim for 7–9 hours, and keep a consistent schedule—it's your metabolic lifeline. Tame stress: Short walks, 5-minute breathing exercises, or setting work boundaries can slash health risks. Bottom line? Small changes to sleep and stress beat marathon workouts. Start tonight—your body will thank you.

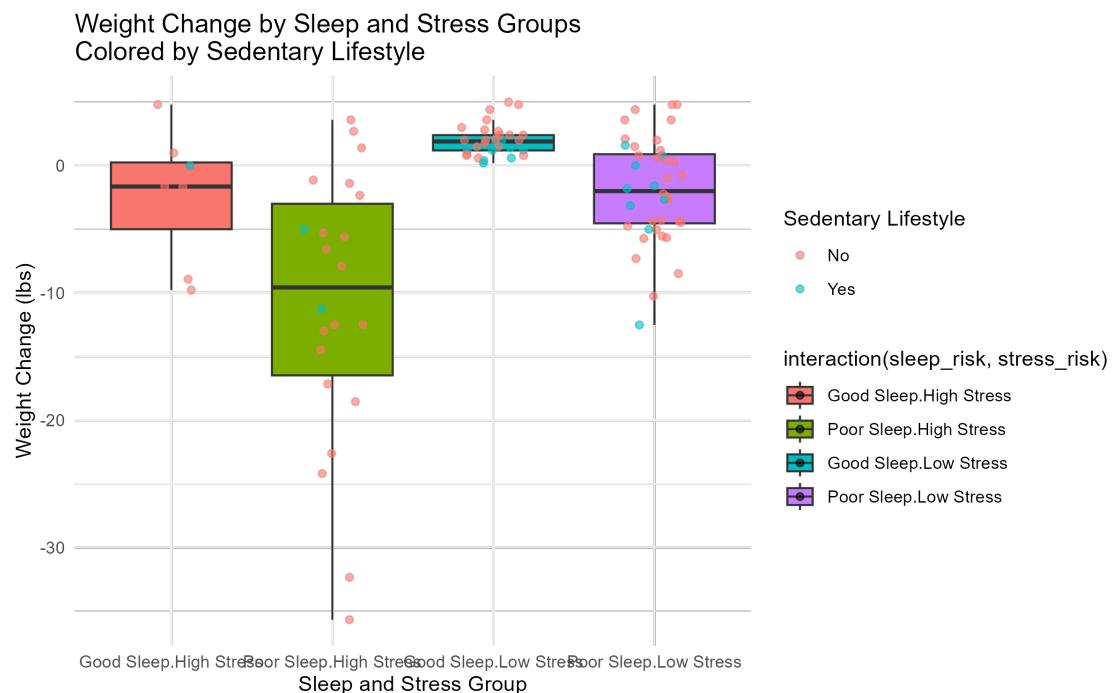


Table 1: Weight Change Summary

Gender	age_group	mean_weight_change	median_weight_change	n
F	Adults (20-29)	-1.0181151	-0.3680584	10
F	Early Career (30-39)	-2.4545587	-1.3943583	11
F	Middle-aged (40-49)	-3.4805916	0.7000000	12
F	Preretirement (50-64)	-1.5639095	0.0000000	9
F	Teenagers (<20)	-1.7940640	-1.7940640	1
M	Adults (20-29)	-2.3794779	-0.1679132	14
M	Early Career (30-39)	-5.5847844	0.6000000	12
M	Middle-aged (40-49)	-3.7051828	0.4500000	12
M	Preretirement (50-64)	-0.8648233	-0.3807078	14
M	Teenagers (<20)	-5.2528414	-5.7184960	5

Table 2: Sleep Risk Pivot Summary

Poor Sleep Risk	age_group	Mean_Weight_Change	Count
No	Adults (20-29)	0.2142665	9
No	Early Career (30-39)	1.5777778	9
No	Middle-aged (40-49)	2.9571429	7
No	Preretirement (50-64)	-0.0982816	11
No	Teenagers (<20)	1.9500000	2
Yes	Adults (20-29)	-3.0281494	15
Yes	Early Career (30-39)	-7.7298256	14
Yes	Middle-aged (40-49)	-6.2899584	17
Yes	Preretirement (50-64)	-2.0918011	12
Yes	Teenagers (<20)	-7.9895678	4

Table 3: Stress Risk Pivot Summary

Poor Sleep Risk	age_group	Mean_Weight_Change	Count
No	Adults (20-29)	0.2142665	9
No	Early Career (30-39)	1.5777778	9
No	Middle-aged (40-49)	2.9571429	7
No	Preretirement (50-64)	-0.0982816	11
No	Teenagers (<20)	1.9500000	2
Yes	Adults (20-29)	-3.0281494	15
Yes	Early Career (30-39)	-7.7298256	14
Yes	Middle-aged (40-49)	-6.2899584	17
Yes	Preretirement (50-64)	-2.0918011	12
Yes	Teenagers (<20)	-7.9895678	4

	Sleep-Stress Interaction	Stress-Age Interaction
(Intercept)	1.758 (6.660)	-0.315 (1.510)
sleep_riskPoor Sleep	-7.168*** (2.704)	
stress_riskLow Stress	5.800** (2.717)	
physical_activityModerately Active	-1.142 (1.893)	
physical_activitySedentary	-0.539 (1.982)	
physical_activityVery Active	-1.076 (2.311)	
age_groupEarly Career (30-39)	-2.976 (1.875)	0.653 (2.135)
age_groupMiddle-aged (40-49)	-2.226 (1.766)	0.107 (2.103)
age_groupPreretirement (50-64)	-0.777 (1.891)	0.164 (2.135)
age_groupTeenagers (<20)	-3.178 (2.861)	-1.867 (3.094)
GenderM	-1.650	

	Sleep-Stress Interaction	Stress-Age Interaction
Daily Caloric Surplus/Deficit	(1.597) 0.002 (0.003)	
BMR (Calories)	-0.001 (0.002)	
Duration (weeks)	-0.158 (0.175)	
sleep_riskPoor Sleep × stress_riskLow Stress	2.980 (3.290)	
High Stress RiskYes		-4.493* (2.615)
High Stress RiskYes × age_groupEarly Career (30-39)		-10.049*** (3.785)
High Stress RiskYes × age_groupMiddle-aged (40-49)		-7.112* (3.767)
High Stress RiskYes × age_groupPreretirement (50-64) 1.248		(3.785)
High Stress RiskYes × age_groupTeenagers (<20)		-10.476 (7.112)
Num.Obs.	100	100
R2	0.455	0.402
R2 Adj.	0.366	0.342
AIC	655.5	654.9
BIC	697.2	683.5
Log.Lik.	-311.752	-316.433
RMSE	5.47	5.73

• p < 0.1, ** p < 0.05, *** p < 0.01

Key Findings on Health Determinants

1. Sleep Quality is Critical:

- Poor sleep correlates with **-7.2 lbs** weight gain ($p<0.01$), surpassing the impact of physical activity (non-significant coefficients).
- Teens and early-career adults with poor sleep show the worst outcomes (**-7.7 to -8.0 lbs** mean weight change). Teens are abstracted from main analysis as their n is very small, n = 5.

2. Stress Drives Negative Outcomes:

- High stress alone leads to **-4.5 lbs** weight gain ($p<0.1$).
- Stress combined with poor sleep exacerbates effects, especially in younger age groups (interaction terms up to **-10.0 lbs**, $p<0.01$).

3. Sedentary Lifestyle Modifies Risk:

- Sedentary individuals (red dots, Figure 16) show higher weight variability, but sleep/stress dominate overall trends.

Practical Recommendations

- **Prioritise Sleep Hygiene:** Consistent sleep schedules improve metabolic health more than moderate exercise.
- **Stress Management:** Mindfulness or flexible work policies could mitigate high-stress impacts, particularly for younger adults.
- **Targeted Interventions:** Early-career professionals need tailored programs addressing sleep and stress.