

# 🌟 Team Formation

# Milestone 1: Team Formation

Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

Team Members:

1. Derin Adeleke - [derinadeleke@college.harvard.edu](mailto:derinadeleke@college.harvard.edu)
2. Cindy Sun - [csun@g.harvard.edu](mailto:csun@g.harvard.edu)
3. Patience Madumera - [pmadumera@college.harvard.edu](mailto:pmadumera@college.harvard.edu)

Project Ideas:

1. Analyze how luxury pricing shifts over time and what styles dominate—giving insight into resource decisions around inventory and markdowns.
  - a. Dataset # 1 - [Mr. Porter and Net-A-Porter](#)
  - b. Dataset # 2 - [ZARA US retail products dataset](#) - PAID dataset but largely meets our requirements
2. Simulate inventory allocation decisions—e.g. how to distribute limited stock across stores to optimize sales.
  - a. Dataset - [Visuelle 2.0](#)
3. Explore how fashion platforms allocate “recommendation slots”—a scarce resource—to boost sales or engagement.
  - a. Dataset - [Open Bandit Dataset](#)
4. Makeup businesses: Investigate selling trend/customer need/satisfaction
  - a. [Sephora data 2020](#) (just for inspiration b/c maybe too old)
  - b. [Sephora data 2023](#)
5. Other great datasets - not useful for our main goal but could be relevant
  - a. [Women's E-Commerce Clothing Reviews](#) - categorical/string variable
  - b. [Labubu Status & facts](#)- PAID statista analysis, documented Labubu sales until 2025

 Team Agreement

# Milestone 2: Team Agreement

Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

This agreement establishes the expectations, responsibilities, and procedures for our team throughout the semester-long project. All members acknowledge and agree to the following:

## **1. Roles and Responsibilities:**

- **Rotating Lead:** Each week, one team member takes the lead on the milestone. The lead is responsible for:
  - Keeping track of deadlines and ensuring the submission is made on time.
  - Coordinating tasks between members for that week.
  - Making sure the milestone is complete and polished.
- **Shared Core Responsibilities:**
  - **Data:** Each member takes responsibility for analyzing a section of the dataset. We all contribute to sourcing, cleaning, and preparing the data.
  - **Analysis:** Everyone participates in exploring the data and generating insights. Different sections or methods can be divided among members.
  - **Visualization:** All members contribute sketches and ideas; one person may implement, but everyone reviews for clarity and effectiveness.
  - **Final Deliverables:** The final website, process book, and screencast video are a joint responsibility. We will review them together before submission.

*All members share in project management tasks such as scheduling meetings, communicating with the instructor, and coordinating deadlines.*

## **2. Communication Protocols:**

Our primary platform of communication is our iMessage Group chat for quick updates. Weekly check in meetings will be scheduled for **Sunday at 3pm**. With interchanges between Zoom and In person meetings. If problems arise, additional meetings should be scheduled with the agreement of the group members. We aim to reply to group messages within 24 hours.

### **3. Working Arrangements:**

During meetings all team members will come prepared with updates and blockers on assigned work. Internal deadlines will aim to be set 2-3 days before official due dates to allow room for review. Everyone is expected to contribute actively and respect each others time.

### **4. Work Allocation:**

Tasks will be divided fairly based on each person's role and current workload. Shared tasks such as writing the report or presentation prep, will be split equally. If workload feels unbalanced, members will raise concerns early so adjustments can be made.

### **5. Accountability:**

If someone cannot meet a deadline they must inform the team at least 48 hours. If individual expectations are not met by team members:

1. First the issue will be discussed with thin the group to find solutions
2. If problems persist, the team will escalate to the instructor for support

Everyone is responsible for the overall success of the project, and not just their role

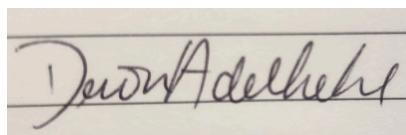
### **6. Signatures**

By signing below, all team members agree to the expectations outlined above.

#### **Name/ Signature / Date**

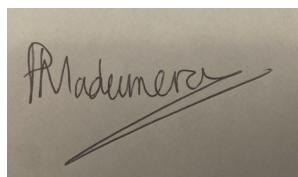
Aderinsola Adeleke

Sept 14th, 2025



Patience Madumera

Sept 14, 2025



Cindy Sun

Sept 14, 2025

Cindy Sun



# Project Proposal

# Milestone 3: Project Proposal

Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

## Instructions:

### 1. Basic Information

- **Project Title:** Give your project an informative and concise title.
- **Team Members:** Make sure the names and e-mail addresses of all team members are up to date.
- **Team Name:** Choose a unique name for your team.

### 2. Abstract

- Provide a paragraph describing your project's main goals, the data you will use, and the type of visualization(s) you plan to create.

### 3. Background and Motivation

- Explain why you chose this project. Discuss your motivation and any background, coursework, or research interests that inspired your team's decision.
- Describe any prior work, examples, or sources that inspired your project. This may include research papers, websites, or specific visualizations (including those discussed in class).

### 4. Data

- Outline plans where and how you will obtain your data. Briefly describe potential data source(s) and provide URLs or citations where appropriate. If data collection or preprocessing is required, briefly outline your plan.

# A Comparative Visualization of Consumer Purchasing Patterns in Luxury Markets

## Team Members

- Cindy Sun, [csun@g.harvard.edu](mailto:csun@g.harvard.edu)
- Patience Madumera, [pmadumera@college.harvard.edu](mailto:pmadumera@college.harvard.edu)
- Derin Adeleke, [derinadeleke@college.harvard.edu](mailto:derinadeleke@college.harvard.edu)

## Team Name

- The Labubu Ladies

## Abstract

This project explores how consumer purchasing patterns and pricing strategies vary across three distinct luxury companies: Hermès, Gucci, and Coach. By analyzing product listings, resale market data, and company financials, we aim to visualize the differences in pricing strategies, category sales, resale value, and geographic demand. Our visualizations will include comparative dashboards, price distribution charts, resale vs. retail scatterplots, and geographic heatmaps. The goal is to highlight how different business models within the luxury industry influence consumer decisions and long-term value perception.

## Background and Motivation

We chose this project because luxury companies provide a compelling case study of how consumer purchasing decisions intersect with pricing strategy, exclusivity, and brand identity. By comparing Hermès, Gucci, and Coach, three companies that differ in history, product category, and marketing strategy, we can explore how different models of luxury shape consumer behavior.

Our motivation is twofold. First, luxury markets are uniquely positioned at the intersection of economics and culture: luxury goods are not merely products, but symbols of status, heritage, and identity. Second, the diversity among Hermès (heritage luxury), Gucci (trend-driven fashion), and Coach (accessible luxury) allows us to study how these dynamics vary across tiers of the industry, providing a richer comparative framework.

Prior research informed our decision to focus on these brands. Studies of Hermès highlight how exclusivity and craftsmanship sustain demand even during economic downturns, positioning products like the Birkin and Kelly handbags as cultural symbols and even financial investments [1,2]. In contrast, Gucci has been analyzed as a case of brand reinvention, successfully

attracting Millennials and Gen Z through collaborations, digital marketing, and localized strategies in key markets like China [3].

Scholarly literature also emphasizes cross-cutting themes that guide our project. For example, history and heritage play a crucial role in shaping perceptions of authenticity and value, as consumers trust brands with deep cultural roots [5]. Pricing in luxury goes beyond cost-recovery: it becomes a form of symbolic communication, where prestige pricing and scarcity strategies reinforce exclusivity and desirability [6,7]. Research in consumer psychology further suggests that brand personality and message framing strongly influence how consumers perceive value, particularly in the luxury sector [8].

We see this project as an opportunity to apply lessons from CS1710 to real-world questions of luxury consumption, combining rigorous data analysis with compelling visual storytelling.

## References

1. Li, X. (2022). *Analysis of the Exclusivity of Hermès Handbags*. BCP Business & Management EDMI.
2. Zong, Z. (2024). *Analysis of Hermes' Marketing Strategy: Evolution, Brand Positioning, and Global Impact*. SHS Web of Conferences.
3. Wang, R. (2025). *From Traditional Luxury to the Youthful Trend: Analysis of Gucci's Brand Reinvention and Marketing Strategies in the Chinese Market*. SHS Web of Conferences.
4. Dryl, W. (2022). *History as a Factor Creating the Value of a Luxury Brand*. Annales Universitatis Mariae Curie-Skłodowska, Sectio H – Oeconomia.
5. Sup de Luxe (2024). *How to Establish a Luxury Pricing Strategy*. Sup de Luxe Institute.
6. Jedidi, H., & Heil, O. (2020). *Prestige Pricing in Markets for Luxury & Premium Products*. Scientific Journal of Research and Reviews.
7. Kim, D. H., & Sung, Y. (2013). *Gucci versus Old Navy: Interplay of Brand Personality and Regulatory Focus in Advertising Persuasion*. *Psychology & Marketing*, 30(12), 1076–1087.

## Data

We have 4 main areas we are trying to explore:

1. Pricing Strategies
  - a. H:
  - b. G: [Net a Porter](#)
  - c. C:
2. Consumer Purchasing Patterns
  - a. H: Google Trends (compare search terms: birkin bag vs. gucci sneakers vs. “Advanced Night Repair”)
  - b. G: GT
  - c. C: GT

3. Resale Value and Longevity
  - a. H: [Vesitaire Collective Dataset](#)
  - b. G: [Vesitaire Collective Dataset](#)
  - c. C:
4. Geographic Demand
  - a. Annual Reports/ Statista / Google Trends

| Company              | Category/Angle  | Possible Datasets   |
|----------------------|---|---|
| Hermes               | Heritage luxury, leather goods, handbags, timeless pieces | <ul style="list-style-type: none"> <li>- <a href="#">Annual Report</a></li> <li>- <a href="#">Statista Data</a></li> <li>- <a href="#">Google Trends</a></li> </ul>   |
| Gucci (Kering Group) | Trend-driven fashion, collabs, seasonal drops             | <ul style="list-style-type: none"> <li>- <a href="#">Annual Report (Kering Group)</a></li> <li>- Google</li> </ul>  |
| Coach                | Accessible Luxury / Contemporary Luxury                   | <ul style="list-style-type: none"> <li>- <a href="https://www.sec.gov/Archives/e_dgar/data/1116132/000095012305006229/y09049e10vq.htm">https://www.sec.gov/Archives/e_dgar/data/1116132/000095012305006229/y09049e10vq.htm</a></li> </ul> |

1. <https://www.supdeluxe.com/en/luxury-news/how-do-you-establish-luxury-pricing-strategy>(qualitative)
2. [https://www.shs-conferences.org/articles/shsconf/pdf/2024/27/shsconf\\_icdeba2024\\_02022.pdf](https://www.shs-conferences.org/articles/shsconf/pdf/2024/27/shsconf_icdeba2024_02022.pdf)
3. <https://irispublishers.com/sjrr/fulltext/prestige-pricing-in-markets-for-luxury-premium-products.ID.000542.php> [Hermes]
4. [https://www.researchgate.net/publication/362150158\\_Analysis\\_of\\_the\\_Exclusivity\\_of\\_Hermes\\_Handbags](https://www.researchgate.net/publication/362150158_Analysis_of_the_Exclusivity_of_Hermes_Handbags) [Hermes]
5. [https://www.researchgate.net/publication/394477307\\_From\\_Traditional\\_Luxury\\_to\\_the\\_Youthful\\_Trend\\_Analysis\\_of\\_Gucci's\\_Brand\\_Reinvention\\_and\\_Marketing\\_Strategies\\_in\\_the\\_Chinese\\_Market](https://www.researchgate.net/publication/394477307_From_Traditional_Luxury_to_the_Youthful_Trend_Analysis_of_Gucci's_Brand_Reinvention_and_Marketing_Strategies_in_the_Chinese_Market) [Gucci]
6. [https://hollis.harvard.edu/permalink/f/1mdq5o5/TN\\_cdi\\_istex\\_primary\\_ark\\_67375\\_WNG\\_7R3KTV6H\\_5](https://hollis.harvard.edu/permalink/f/1mdq5o5/TN_cdi_istex_primary_ark_67375_WNG_7R3KTV6H_5) [Gucci]

# Data sets

| Company              | Category/Angle  | Possible Datasets   |
|----------------------|---|---|
| Hermes               | Heritage luxury, leather goods, handbags, timeless pieces | - <a href="#">Annual Report</a><br>- <a href="#">Statista Data</a><br>- <a href="#">Google Trends</a> |
| Gucci (Kering Group) | Trend-driven fashion, collabs, seasonal drops             | - <a href="#">Annual Report (Kering Group)</a><br>- Google  |
| Estee Lauder         | Luxury beauty & skincare, marketing-heavy                 | - <a href="#">Annual Report</a><br>-  |
| Louis Vuitton (LVMH) | Diversified luxury (bags, RTW, accessories)               |   |
| Rolex                | Watches, status signaling, long waitlists                 |   |
| Chanel               | High fashion, beauty, perfume, “classic luxury”           | - <a href="#">Annual Report (Kering Group)</a>  |
|                      |   |   |
|                      |   |   |
|                      |   |   |

#### Overall Luxury Fashion Dataset:

- <https://www.kaggle.com/datasets/justinpakzad/net-a-portermr-porter-fashion-dataset>
- <https://crawlfeeds.com/datasets/luxury-fashion-products-dataset-from-mytheresa> (\$185 )

# Idea #1

Main Question:

How do consumer purchasing patterns and pricing strategies differ across heritage luxury, trend-driven luxury, and beauty luxury companies?

Sub Questions:

1. Pricing Strategies

- Do heritage brands (e.g., Hermès) avoid markdowns, while trend-driven brands (e.g., Gucci) rely on seasonal discounts, and beauty brands (e.g., Estée Lauder) use promotions/bundles?
- How do average prices compare across product categories?

2. Consumer Purchasing Patterns

- Which product categories dominate consumer demand (handbags vs RTW vs skincare)?
- Do consumers prioritize timeless “classics” vs seasonal “trends”?

3. Resale Value & Longevity

- Which brands' products retain or appreciate in value on resale markets (Hermès vs Gucci vs beauty)?
- How does this reflect consumer behavior and brand strategy?

4. Geographic Demand

- Which regions drive sales growth (Asia, US, Europe)?
- Do different categories resonate more strongly by region?

## Idea #2

Main Question:

How does the introduction of “hype” products (ex. Labubu-like trend-driven fashion products) impact sales and consumer behavior in heritage luxury brands?

Sub Questions:

1. Sales Impact

- Do limited-edition or hype-inspired products increase overall sales for heritage brands (e.g., Hermès, Louis Vuitton)?
- Do these products attract new consumer segments, or mainly encourage existing customers to purchase more?
- How does exposure to social media content (hype products) affect consumers' willingness to buy trend-driven luxury vs heritage luxury products?

2. Consumer Demographics & Behavior

- Which age groups or social media-following consumers are most attracted to these products?
- Do these products encourage impulse purchases or long-term brand loyalty?



Map

# Milestone 4: Map

Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

## Instructions:

**As a team**, your discussion should focus on the following questions:

1. Who is your audience? Come up with **at least three** options and pick one target audience.
  - business students and researchers interested in consumer behavior, pricing strategy, and brand positioning in the luxury sector
  - entrepreneurs and emerging brands hoping to learn from established players' pricing and branding strategies
  - general audiences curious about the luxury market
2. Describe your target audience in more detail. What do they know? What are their interests? What visualization literacy do they have? At what level of detail will you present information to them?
3. What questions about your data will be interesting for your audience? Come up with a list of interesting questions that your audience may have about your data. The more, the better, but your team should come up with **at least ten questions**.
4. What data do you have or plan to obtain? Briefly describe the data you envision to use and the respective data types (categorical, ordinal, or quantitative) in your process book. It's OK if you are unsure about the data type for some attributes - you can simply describe them (e.g., geographic location).

You may want to use a virtual whiteboard to facilitate your discussion. We recommend [Miro](#)

[Links to an external site.](#)

. You can sign up for an educational account following the instructions here: [Miro education plan sign-up](#)

[Links to an external site.](#)

**Individually** at home, you should then

1. Sketch **at least three** initial visualizations **per team member** that answer questions you came up with earlier. You may use AI-based tools (e.g., Claude), traditional tools (e.g., Excel), or pen and paper.
2. Upload the images into your process book. Make sure that the titles (i.e., the questions) are visible. Put the name of each team member before the screenshots that they uploaded.
3. After your sketches, write a brief one-paragraph reflection: How do the questions addressed in your sketches compare with the team's original questions? Why do they differ (if they do)? Are some questions better than others? Why? If you stayed with the original questions, why?

Please document your discussion and answers in your process book. Be comprehensive and add sufficient details. At this point, we do not want you to rule out anything - use your inspiration, generate many ideas, and diverge. Please include both text and any screenshots of your virtual whiteboard (if you used one) in your process book.

# Part 1 - Team

## Chosen Target Audience

Our target audience is **luxury industry professionals**, including brand strategists, marketers, consultants, pricing analysts at luxury retailers (e.g., Neiman Marcus, Saks Fifth Avenue), and both emerging and current retailers.

- Interested in understanding consumer demand, pricing strategies, and brand positioning across different tiers of the luxury market.
- Have strong product knowledge such as seasonality, MSRP/markdown structures, and sell-through rates.
- Focused on maximizing margins and selecting the right products by brand, category, and region.
- Visualization literacy: medium to high. Comfortable with dashboards, box plots, scatterplots, and geographic maps.
- Prefer clarity, storytelling, and actionable insights over raw statistics or overly complex visuals.

## Other Target Audiences Considered

### **Business Students / Academics (Harvard peers, professors)**

- Interested in learning how data visualization can reveal insights about economics, culture, and luxury markets.
- Visualization literacy: high. They can handle complex dashboards.

### **General Consumers / Fashion Enthusiasts**

- Curious about how luxury brands differ in value retention, popularity, and pricing.
- Visualization literacy: medium. Need more accessible and visually engaging charts.

### **Entrepreneurs and Emerging Brands**

- Looking for patterns they can adapt to smaller-scale or emerging businesses
- Visualization literacy: high. They can handle complex dashboards, infographics, and clear comparisons

## Possible Questions

1. How do average handbags/shoes/accessories prices compare across Hermès, Gucci, and Coach?
2. Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand?
3. Which categories are most protected from markdowns?
4. Do new product launches start at higher price points over time?
5. Is there a seasonal pattern to markdowns?
6. Do certain categories (bags, RTW, accessories) dominate across all three brands, or only some?
7. Which styles stay in collections for many years, and which disappear quickly?
8. How does product variety differ (number of SKUs per season) across brands?
9. What role do limited editions and collaborations play in Gucci's product mix?
10. How much of their original price do Hermès bags retain on resale?
11. How fast do Gucci bags lose value compared to Hermès and Coach?
12. Which region contributes the most to Hermès' sales? To Gucci? To Coach?
13. How does search interest (Google Trends) differ by region?
14. How often do each brand's flagship products trend in Google searches regionally?
15. How does consumer interest in these brands fluctuate over time—are they seasonal, trend-driven, or stable?
16. What drives revenue growth across these brands?

## Data

### Core Sources

- Retail/Product Listings
  - [Net a Porter/Mr Porter](#)
  - Brand sites (webscraping if possible)
    - [Gucci](#) - webscrape possible for product name, price, color
    - [Hermes](#) - blocked webscrape - google account will be blocked once tried to webscrape
    - [Coach](#) - webscrape possible for product name, price (color might also be scrapable - needs more efforts)
    - [saks fifth avenue](#) - webscrape not possible for brand name, product name, price (original and sale price), best seller indicator (Yes - if the product is marked as "BEST SELLER")
    - [Nordstrom](#) - webscrape possible for brand name (need API), product name, price (original and sale price), color
- Resale Data
  - [Vestiaire Collective \(dataset\)](#)
  -

- Company Filings: Annual reports for category/region breakdowns.
  - [Hermes](#)
  - [Kering \(Gucci\)](#)
  - [Coach \(Tapestry\)](#)
  - <https://tapestry.gcs-web.com/static-files/06a073d6-5ec4-4c12-af7c-f6ac954cd8d8>
  -
- Signals: [Google trends](#) for brand/style search interest by region and time
- Basic

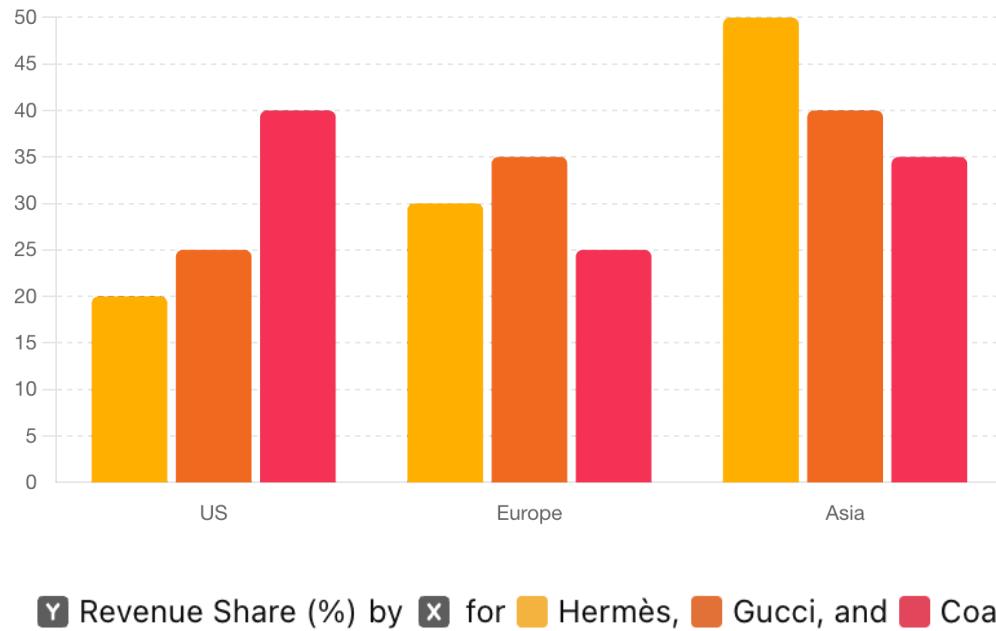
## Attributes and data types

| Attribute  | Type                  |
|--|-----------------------|
| Brand, Category (bags, ready to wear clothes), Subcategory , (tote, mini skirt, dress, knit) | Categorical           |
| Product  | Categorical           |
| Season, Year   | Temporal, Ordinal     |
| Retail Price (MSRP)  | Quantitative          |
| Discount% / Markdown depth   | Quantitative          |
| “Classic/Trend” tag  | Binary/Ordinal        |
| Resale price   | Quantitative          |
| Resale date/Time since launch  | Temporal Quantitative |
| Region/Country   | Categorical           |
| Revenue share by category  | Quantitive            |
| Customer Rating  | Quantitive            |
| Customer Review  | Categorical (strings) |

## Part 2 - Individual Sketches

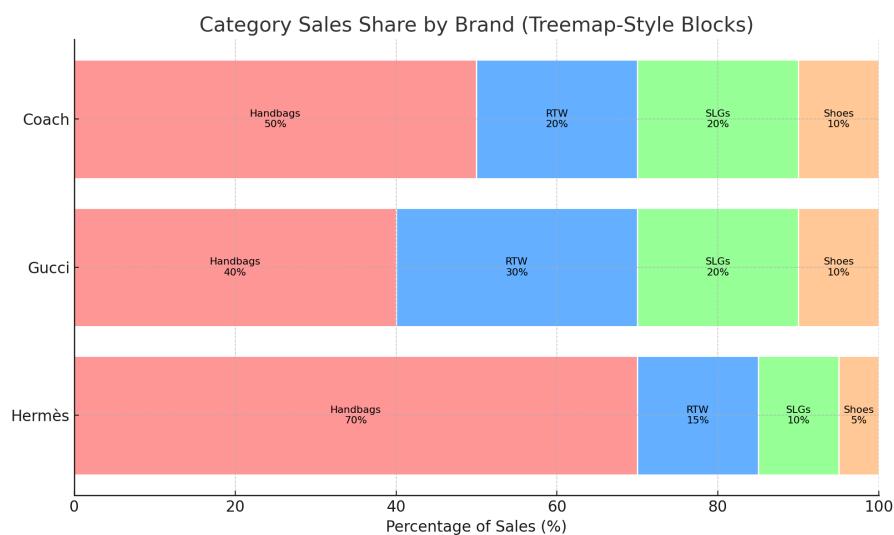
Derin

Question 1: Which regions drive the most sales growth?

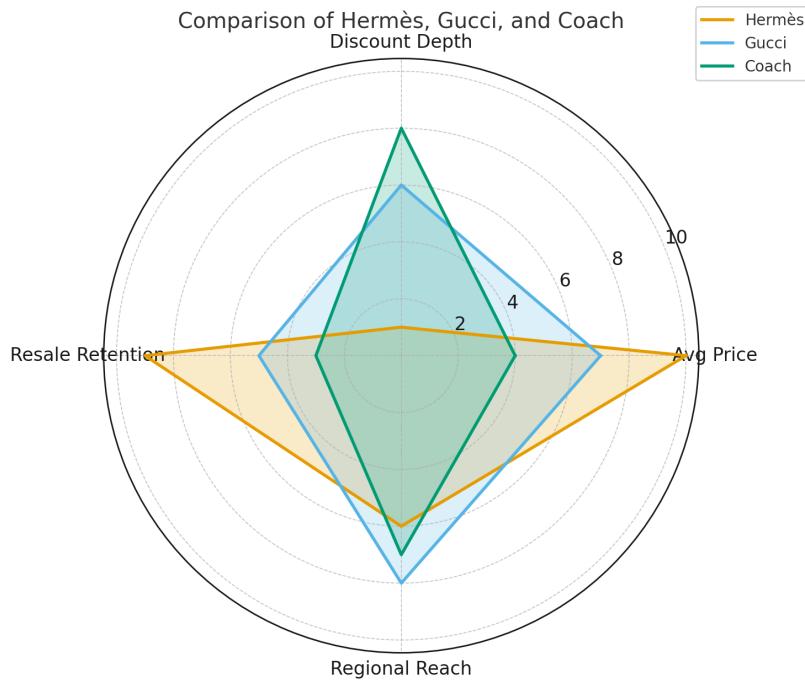


Y Revenue Share (%) by X for Hermès, Gucci, and Coach

Question 2: What product categories dominate sales for each brand?



Question 3: How do Hermès, Gucci, and Coach compare across key dimensions: average price, discount depth, resale retention, and regional reach?

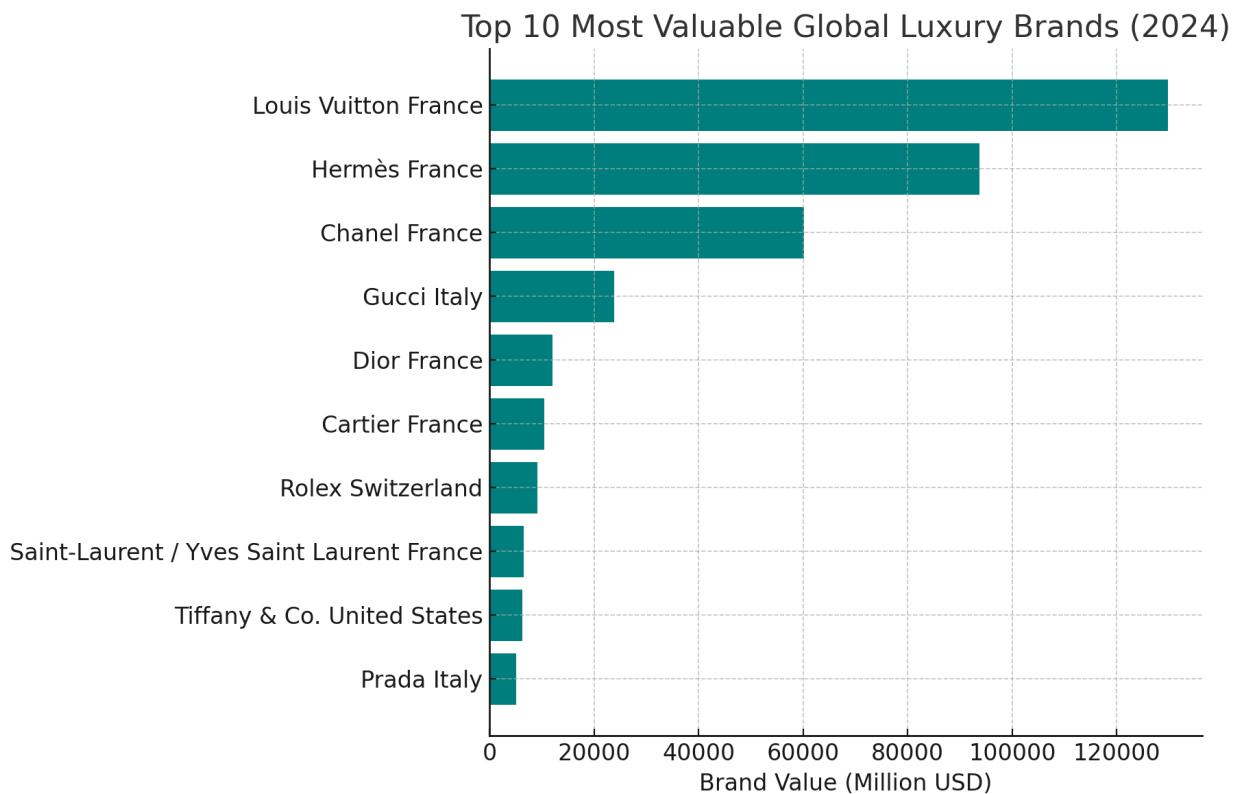


#### Reflection:

I decided to create my sketches using ChatGPT visualizations. Admittedly, while coming up with these sketches what came to mind the most was how I wanted our final project to be as engaging and visually stimulating as possible. I started with the first question using a bar chart, which mirrors one of the questions exactly. It felt somewhat basic, though technically still useful for the target audience. I liked the treemap one more because it communicates proportion in a more intuitive and visually interesting way, it makes the dominance of a category stand out quickly and doesn't feel as flat as a simple bar. My favorite was the last spider (radar) plot. I came up with a slightly different question for this one, and I think it looks great. Depending on which dimensions we choose, it could turn into a fascinating and visually appealing visualization. That said, I'm not entirely sure how helpful it would be for the target audience since it communicates so much at once. It might be more suited for a high-level presentation where the goal is to compare brands broadly rather than to dig into specific numbers. One idea that came to mind is that adding an extra dimension of time, so the graphs could change or animate season by season, would make even the simplest visualizations, like the bar chart for regional sales, much more powerful. For example, showing how regional sales shift over time or how category dominance evolves would give the visuals both depth and narrative flow.

# Patience

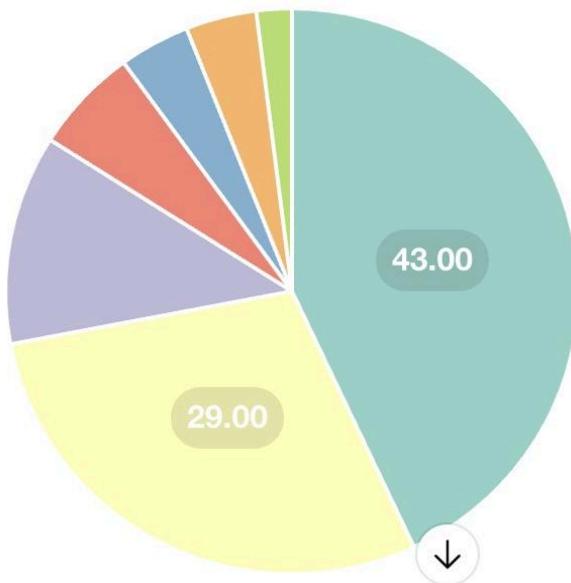
1. How does Hermès compare to other top luxury brands worldwide?



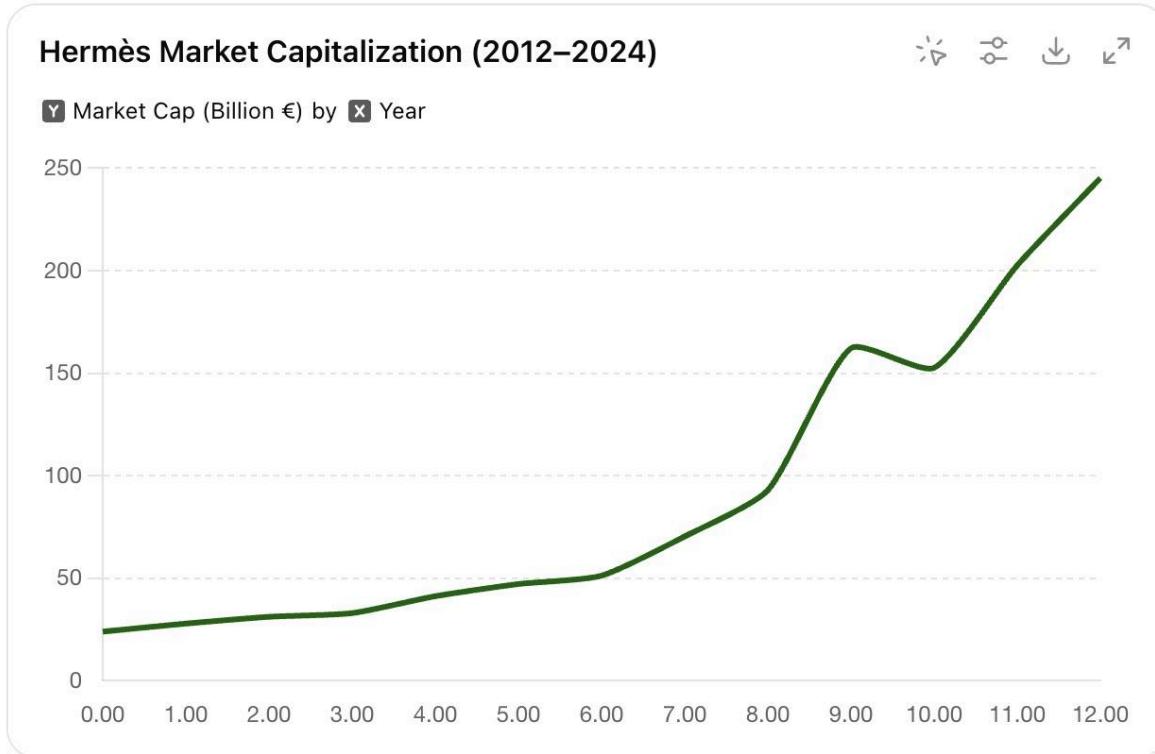
2. Which product categories drive the most revenue for Hermès?

### Hermès Revenue Breakdown By Segment (2024)

For █ Leather goods & saddlery, █ Ready-to-wear & accessories, █ Other Hermès sectors,  
█ Silk and textiles, █ Perfume and beauty, █ Watches, and █ Other products



### 3. How has Hermès grown financially over time?



#### Reflection:

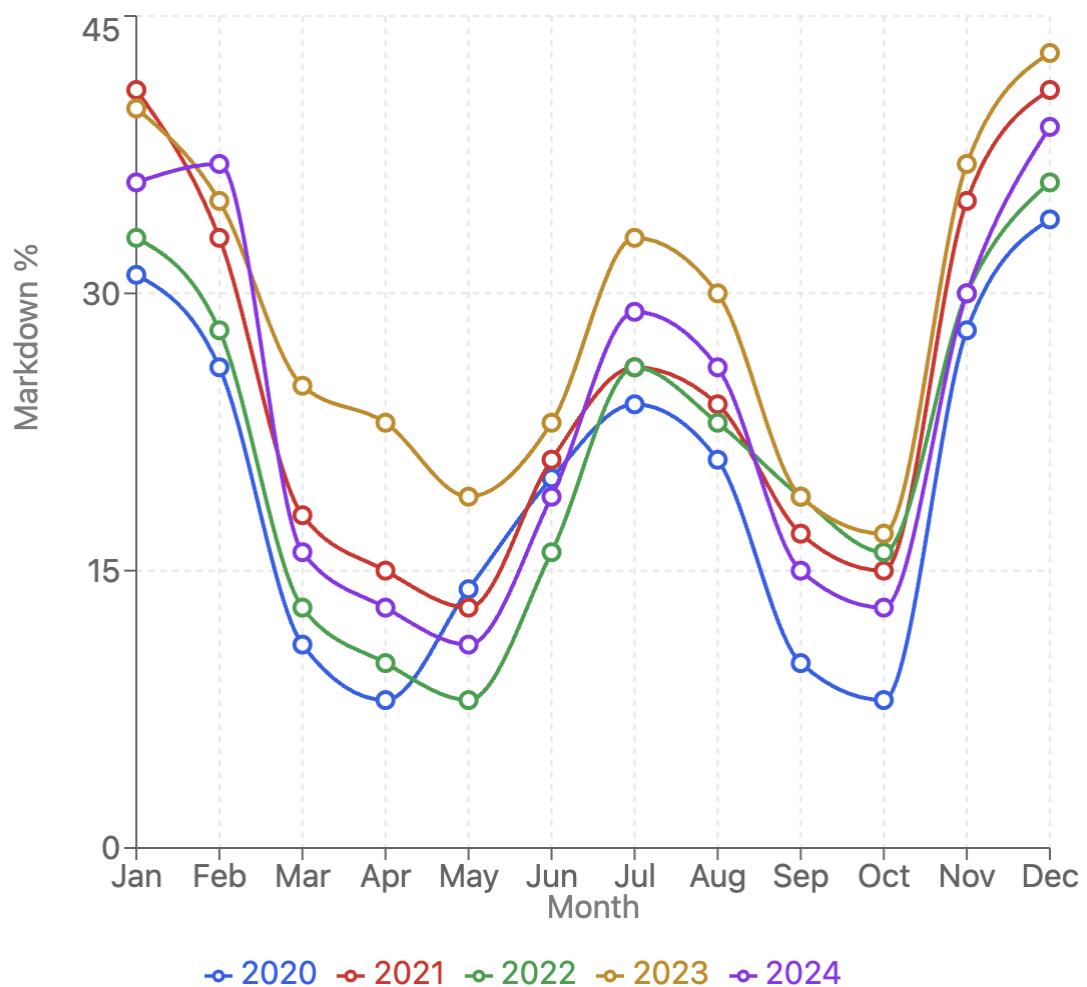
I sketched these three visualizations because they directly connect to several of the questions our team raised about Hermès' brand positioning, product strategy, and competitive standing. The market capitalization line chart answers questions about Hermès' growth over time and how investor confidence reflects its performance. The revenue breakdown pie chart addresses which product categories contribute most to Hermès' business, helping to explore consumer behavior and category dominance. The global brand value bar chart situates Hermès within the wider luxury landscape, answering questions about how it compares to other leading brands like Louis Vuitton, Gucci, and Dior. Together, these sketches translate our team's broader discussion into clear visuals that illustrate financial growth, product strengths, and competitive rank—key areas of interest for our chosen audience of business students and researchers focused on consumer behavior and brand strategy.

Cindy

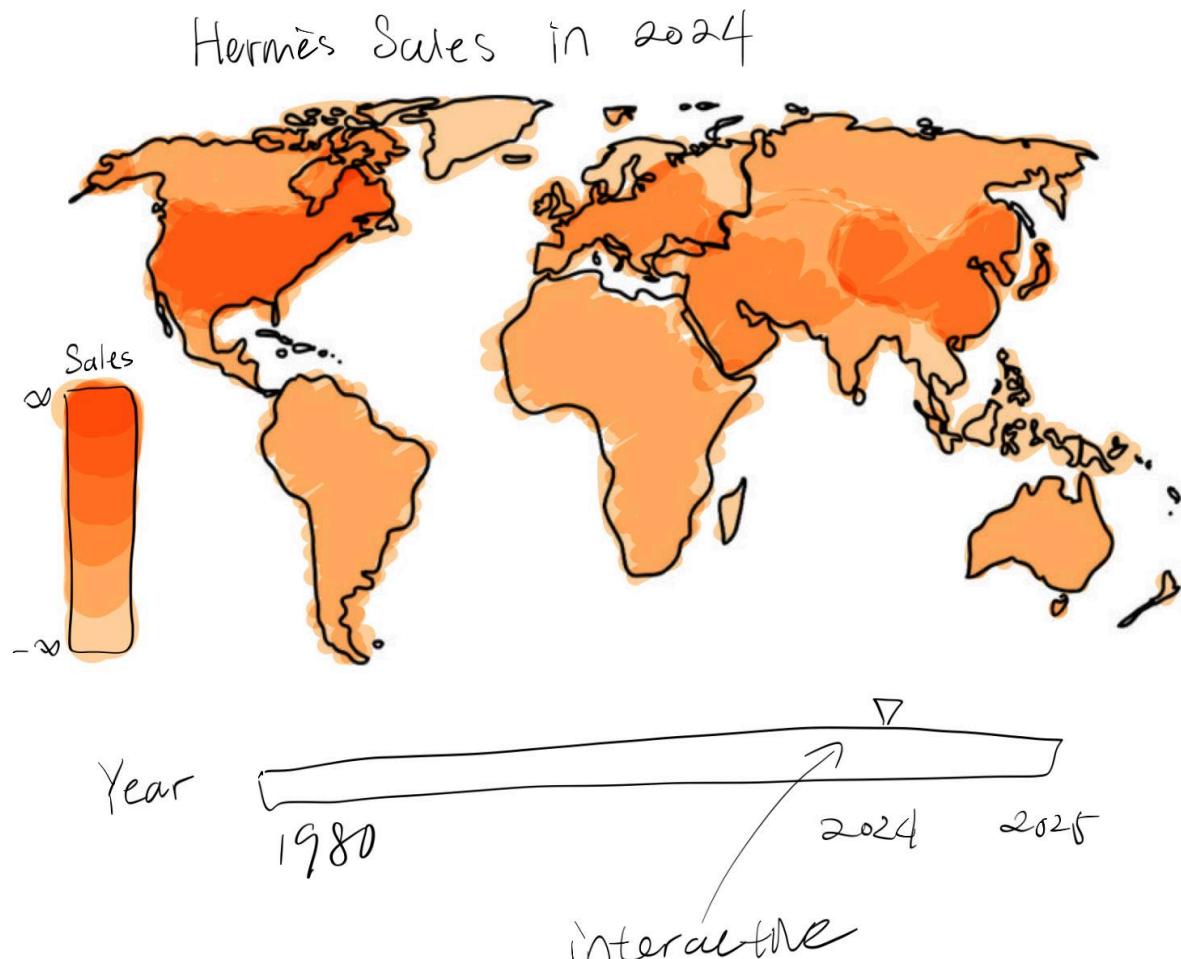
Question 1: Is there a seasonal pattern to markdowns?

## Seasonal Markdown Patterns Analysis

Monthly markdown percentages across 5 years (2020-2024)



Question 2: Which region contributes the most to Hermès' sales? To Gucci? To Coach?  
(Hermes for example)



Question 3: Which styles stay in collections for many years, and which disappear quickly? (Hermès for example)

## Hermès Collection Longevity

*Years each style remained in the collection (1980-2025)*

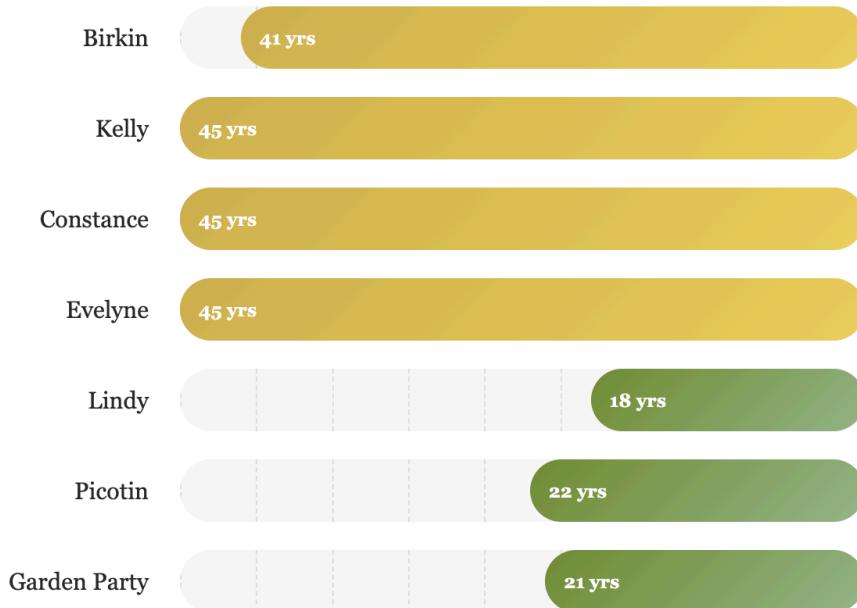
Iconic (25+ years)

Long-lasting (15-24 years)

Moderate (8-14 years)

Short-lived (1-7 years)

1980      1990      2000      2010      2020      2025



### Reflection

In my sketches, I focused on creating visualizations that directly answered the team's original questions in a clear and obvious way. For example, the seasonal markdown pattern chart highlighted differences across years and seasons, making trends immediately visible. The hand-drawn heat map of sales by country directly addressed the geographic sales question, allowing audiences to grasp brand differences at a glance. I also want to make the timeline (Year) interactively to audiences so they can explore the sales data across years. Lastly, the timeline of Hermès' styles made it easy to see which designs were enduring versus short-lived. These visualizations stayed true to the original questions but reframed them in ways that made the answers more visually accessible and impactful, showing that well-scoped, concrete questions often lead to stronger, more interpretable visuals. I stayed with the original questions because they were already strong, clear, and directly connected to our interest.



# Data

# Milestone 5: Data

Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

## Overall Datasets

| Dataset/URL  | Format               | Description / Usage Notes   |
|--|----------------------|---|
| <a href="#">Net a Porter/Mr Porter</a> [Derin]   | .csv                 | Prices/product descriptions, brand, and category info from two fashion retailers. Net-a-Porter offers a vast array of women's luxury fashion items, while Mr Porter caters to men's designer fashion.   |
| <a href="#">Vestiaire Collective (dataset)</a> [Derin]   | .csv                 | Academic Use This dataset contains product listings from Vestiaire, an online marketplace for buying and selling pre-owned luxury fashion items   |
| Google Trends [Derin]<br>-  Regional Breakdown ...<br>-  Compartitive Timeline ...   | .csv<br>exp<br>orts  | Public Use; Search Interest by brand and region   |
| Brand Reports: [Cindy]<br><br><a href="#">Hermes</a> - 2025 H1 + Q1Q2<br><a href="#">Hermes</a> - 2023 H1 + Q1Q2<br><a href="#">Hermes</a> - 2022 H1 + Q1Q2<br><a href="#">Hermes</a> - 2024 annual + Q4<br><a href="#">Hermes</a> - 2023 annual + Q4<br><a href="#">Hermes</a> - 2022 annual + Q4<br><a href="#">Hermes</a> - 2021 annual + Q4<br><a href="#">Hermes</a> - 2024 Q3<br><a href="#">Hermes</a> - 2022 Q3<br><a href="#">Statista</a><br>- <a href="#">Google sheet</a><br><br><a href="#">Kering Publications</a> | .pdf<br>.xlsx<br>web | Original data in published annual revenue report, manually categorized into google sheets.<br><br>Note: Coach financial reports differs in <ul style="list-style-type: none"><li>- Units of money (millions of dollars, not millions of euros)</li><li>- Amount (Net-Sales instead of Revenue)</li><li>- Fiscal year (not normal years)</li></ul> |

|  |      |  |
|--|------|--|
| <p><a href="#">Kering (Gucci)</a> - 2025 H1<br/> <a href="#">Kering (Gucci)</a> - 2024 Q1-4<br/> <a href="#">Kering (Gucci)</a> - 2023 Q1-4<br/> <a href="#">Kering (Gucci)</a> - 2022 Q1-4<br/> <a href="#">Kering (Gucci)</a> - 2021 Q1-4<br/> <a href="#">Kering (Gucci)</a> - full 2024 annual<br/> <a href="#">Kering (Gucci)</a> - full 2023 annual<br/> <a href="#">Kering (Gucci)</a> - full 2022 annual<br/> <a href="#">Kering (Gucci)</a> - full 2021 annual<br/> <a href="#">Kering (Gucci)</a> - full 2020 annual<br/> <a href="#">Statista</a><br/> - <a href="#">Google sheet</a></p> <p>Coach data from:<br/> <a href="https://tapestry.gcs-web.com/presentations-and-financial-reports">https://tapestry.gcs-web.com/presentations-and-financial-reports</a></p> <p><a href="#">Coach (Tapestry)</a> - 24-25 fiscal year (html)<br/> <a href="#">Coach (Tapestry)</a> - 23-24 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 22-23 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 21-22 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 19-20 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 17-18 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 15-16 fiscal year<br/> <a href="#">Coach (Tapestry)</a> - 12-13 fiscal year<br/> - <a href="#">Google sheet</a></p> <p><a href="#">Combining all google sheets together</a></p> |      |  |
| <a href="https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id1309471%20vestiaire-collective-app-downloads-per-month-worldwide-2021-2025%20(1).xlsx?d=wd2201814acc442f8a3d35e61a4abdd4a&amp;csf=1&amp;web=1&amp;e=oLocY5(Patience)">https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id1309471%20vestiaire-collective-app-downloads-per-month-worldwide-2021-2025%20(1).xlsx?d=wd2201814acc442f8a3d35e61a4abdd4a&amp;csf=1&amp;web=1&amp;e=oLocY5(Patience)</a>  | .xls | <p><b>Overview of second hand luxury goods market</b></p> <p>Monthly downloads of the Vestiaire Collective app worldwide from January 2021 to March 2025</p> |

|  |       |   |
|--|-------|---|
| <a href="https://www.accio.com/i/b2c?products_id=US_849506333%2CUS_-420876817%2CUS_1925891346%2CUS_US_93153193%2CUS_93821291%2CUS_1759843410%2CUS_-1246107556%2CUS_-421021975%2CUS_93262159%2CUS_92245319&amp;search=top%20selling%20coach%20bags&amp;tab=top&amp;title=top%20selling%20coach%20bags&amp;cluster_id=(Patience)"><u>https://www.accio.com/i/b2c?products_id=US_849506333%2CUS_-420876817%2CUS_1925891346%2CUS_US_93153193%2CUS_93821291%2CUS_1759843410%2CUS_-1246107556%2CUS_-421021975%2CUS_93262159%2CUS_92245319&amp;search=top%20selling%20coach%20bags&amp;tab=top&amp;title=top%20selling%20coach%20bags&amp;cluster_id=(Patience)</u></a> | Table | Top selling coach bags  |
| <a href="https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id245922_revenue-breakdown-of-hermes-worldwide-2024-by-segment%20(1).xlsx?d=w3a2bd1c20e8c499eb164bff7cb6fd4aa&amp;csf=1&amp;web=1&amp;e=4EuUFL"><u>https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id245922_revenue-breakdown-of-hermes-worldwide-2024-by-segment%20(1).xlsx?d=w3a2bd1c20e8c499eb164bff7cb6fd4aa&amp;csf=1&amp;web=1&amp;e=4EuUFL</u>(Patience)</a>   | .xls  | Revenue breakdown of hermes by product name                                   |
| <a href="https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id245920_revenue-breakdown-of-hermes-worldwide-2024-by-geographical-region.xlsx?d=wc6f77ca60363471e8fa47f7f91616c81&amp;csf=1&amp;web=1&amp;e=dtFO4d"><u>https://hu-my.sharepoint.com/:x/r/personal/pmadumera_college_harvard_edu/Documents/statistic_id245920_revenue-breakdown-of-hermes-worldwide-2024-by-geographical-region.xlsx?d=wc6f77ca60363471e8fa47f7f91616c81&amp;csf=1&amp;web=1&amp;e=dtFO4d</u>(Patience)</a>   | .xls  | Revenue breakdown of Hermès worldwide 2024, by geographical region            |
| - Marketing mix of coach :<br><a href="https://www.marketing91.com/marketing-mix-of-coach/#google_vignette"><u>https://www.marketing91.com/marketing-mix-of-coach/#google_vignette</u>(Patience)</a>   |       | In depth analysis of the 4Ps of coach( product, pricing, place and promotions |
|  |       |   |

## Dataset : Versatile app downloads(Patience)

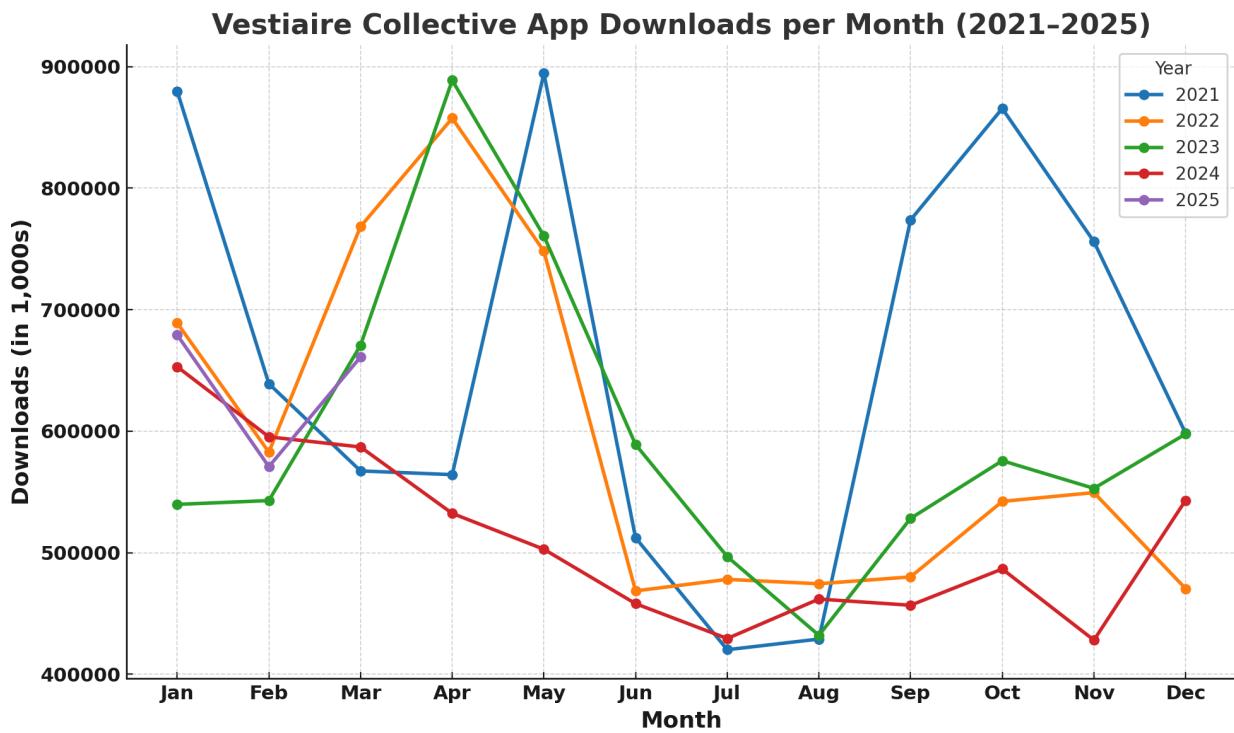
Rows : 11, 5 columns

The mobile app Vestiaire Collective, a C2C e-commerce platform for buying and selling second-hand designer apparel, was downloaded approximately 660,000 times in March 2025. The app's popularity peaked in May 2021, when downloads reached nearly 900,000.

| Variable           | Type     | Description  |
|--------------------|----------|--|
| Month              | Nominal  | Month of the year  |
| Year               | Ordinal  | Year of observation                                      |
| Downloads in 1000s | Discrete | Number of app downloads worldwide, measured in thousands |

## Data Wrangling and Cleaning

- Replaced missing values (“–”) with “Missing.”
- Ensured all download figures were numeric and consistently formatted.
- Verified units were in thousands (1,000s) across all years.
- Checked month order and naming consistency.
- Noted that 2025 data only covers January–March.
- Reviewed for outliers and kept table structure uniform.

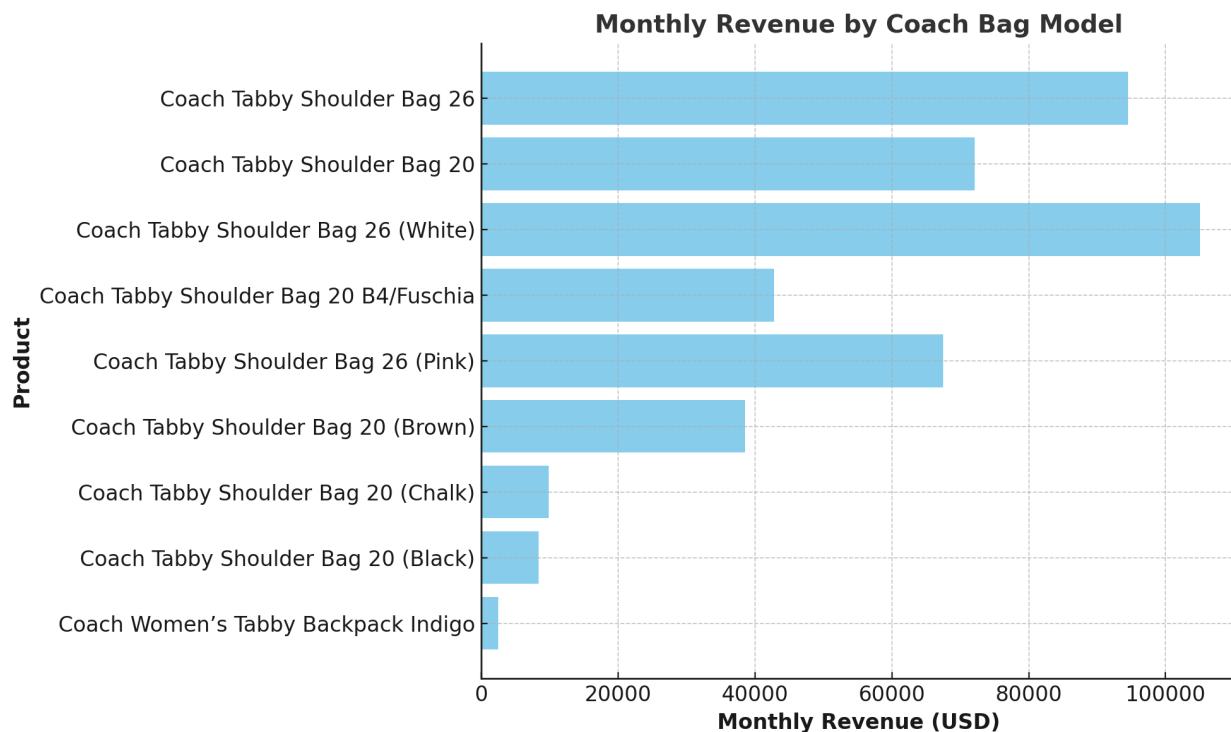


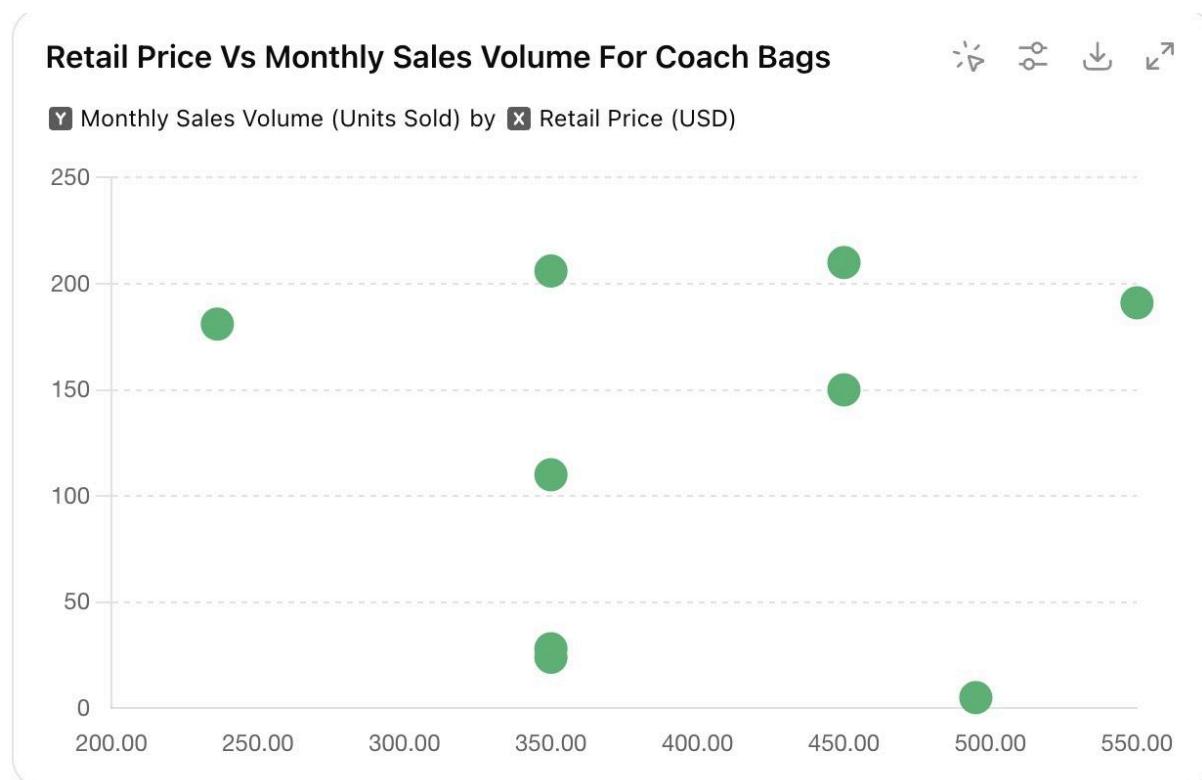
## Data set: Top Selling Coach bags(Accio)

**Rows:** 10, **Columns:** 5

The dataset provides an overview of the best-selling **Coach Tabby and Women's Tabby bags** listed on Amazon. It includes each product's retail price, estimated monthly revenue, monthly sales volume, and sales trend (month-over-month growth or decline). The data highlights which Coach bag models perform best in terms of sales and popularity.

| Variable              | Type       | Description   |
|-----------------------|------------|---|
| Product Name          | Nominal    | Name and model of the Coach bag                           |
| Retail Price (USD)    | Continuous | Selling price of the product in U.S. dollars              |
| Monthly Revenue (USD) | Continuous | Total estimated monthly earnings per product              |
| Monthly Sales Volume  | Discrete   | Number of units sold per month                            |
| Sales Trend (MoM)     | Continuous | Percentage change in sales compared to the previous month |





### Dataset :Revenue breakdown of hermes by product name

This dataset presents the **revenue share of Hermès worldwide in 2024**, categorized by product segment. It highlights the contribution of each division—such as leather goods, fashion and accessories, silk and textiles, perfumes, and watches—to the company's total revenue. In 2024, the **watches segment accounted for approximately four percent** of Hermès' global Revenue.

| Variable          | Type       | Description   |
|-------------------|------------|---|
| Product Segment   | Nominal    | Category of Hermès products (e.g., Leather Goods, Watches, Perfume, Ready-to-Wear, Accessories) |
| Revenue Share (%) | Continuous | Percentage of total Hermès global revenue contributed by each product segment                   |
| Year              | Ordinal    | Year of observation (2024)  |

Removed extra header rows so the dataset starts directly with the variable names (*Product name* and *Contribution to revenue*).

Standardized column names for consistency (e.g., changed “Product name” → Product\_Segment, “Contribution to revenue” → Revenue\_Share\_%).

Removed redundant text (“in %”) from data cells and column headers to ensure numeric formatting.

Converted revenue share values from text to numeric (e.g., “43” → 43.0).

Checked for missing or duplicate entries and confirmed all product segments are unique.

Verified that total revenue share sums to 100%, confirming data completeness.

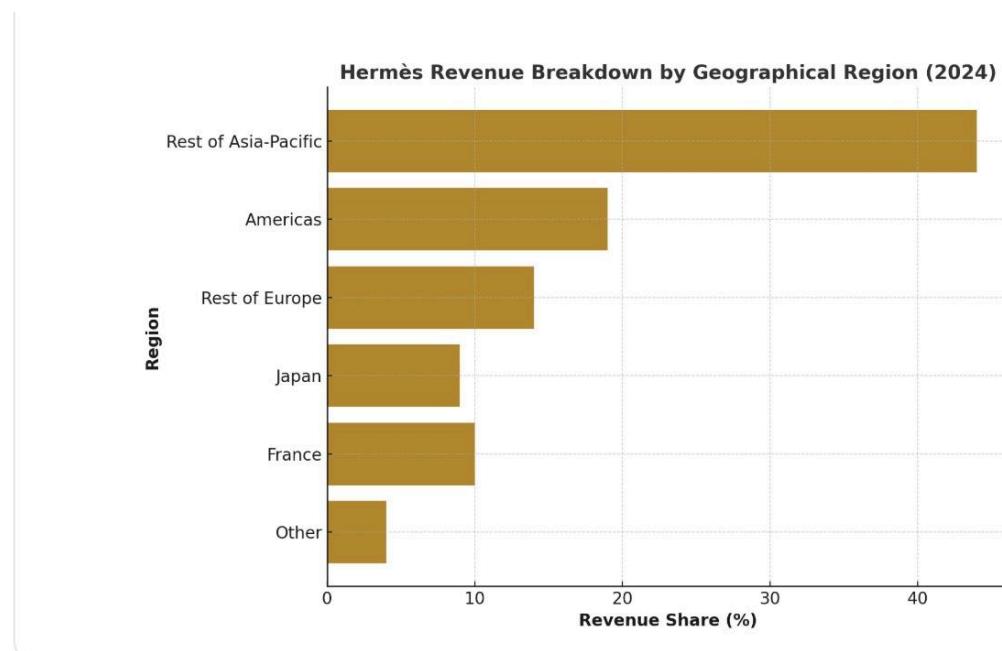
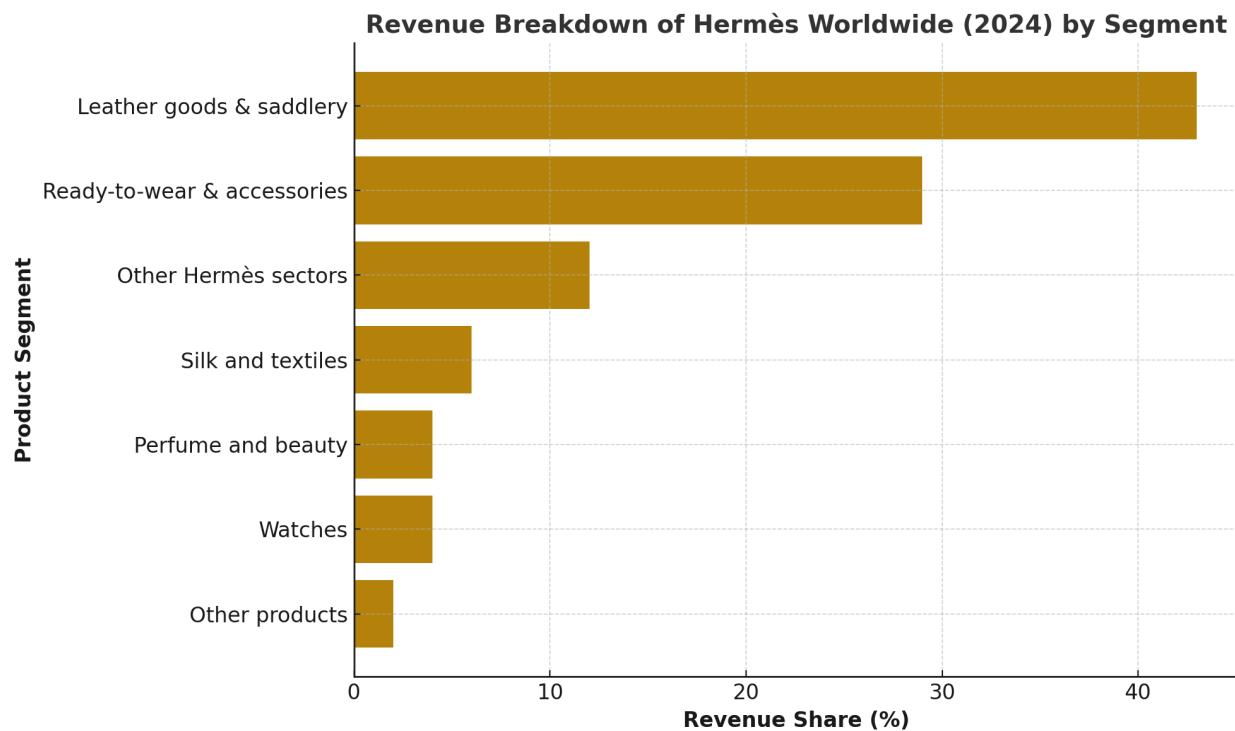
## Dataset: Revenue Breakdown of Hermès Worldwide (2024), by Geographical Region

**Rows:** 6    **Columns:** 2

This dataset presents the **revenue distribution of Hermès worldwide in 2024**, categorized by major geographical regions. The data highlights that the **Rest of Asia-Pacific** is Hermès's most significant market, contributing **44%** of global revenue. The **Americas** and **Rest of Europe** follow, with **19%** and **14%**, respectively, while **Japan** and **France** contribute **9%** and **10%**, and **Other regions** make up **4%**. This breakdown illustrates Hermès's strong presence in the Asia-Pacific region and its global diversification across markets.

### Variables

| Variable          | Type       | Description   |
|-------------------|------------|---|
| Region            | Nominal    | Name of the geographical region contributing to Hermès's revenue (e.g., Asia-Pacific, Americas, France) |
| Revenue Share (%) | Continuous | Percentage of total Hermès global revenue generated from each region                                    |
| Year              | Ordinal    | Year of observation (2024)  |
| Company Name      | Nominal    | Name ↓ of the company — Hermès International S.A.   |



## Dataset: Marketing Mix of Coach (2025)

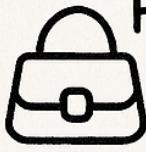
This dataset provides a **qualitative analysis** of the **4Ps (Product, Price, Place, and Promotion)** that define Coach's marketing strategy. It is a **text-based dataset** compiled from secondary sources and written descriptions rather than numerical data. The content outlines how Coach positions itself as an *affordable luxury brand*, emphasizing high-quality craftsmanship, competitive yet accessible pricing, a strong global retail network, and distinctive promotional campaigns.

### Key variables

|                  |  |
|------------------|--|
| <b>Product</b>   | Details the variety of products offered by Coach, including handbags, footwear, apparel, fragrances, and accessories, highlighting craftsmanship and durability. |
| <b>Price</b>     | Explains Coach's mid-premium pricing strategy within the accessible luxury segment, balancing exclusivity and affordability.                                     |
| <b>Place</b>     | Describes the brand's global distribution strategy through flagship stores, boutiques, licensed retailers, and online platforms.                                 |
| <b>Promotion</b> | Outlines the company's marketing approach, including celebrity endorsements, digital campaigns, fashion shows, and seasonal discounts.                           |

## PRODUCT

- Handbags
- Footwear
- Apparel, Perfume



## PRICE

- Premium pricing
- Accessible lux.
- Occasional sale



Marketing  
Mix of  
Coach

## PLACE

- Flagship stores
- E-commerce site
- Boutiques



## PROMOTION

- Celebrity ads
- Fashion shows
- Social media



# Net-a-Porter / Mr Porter [Derin ]

## Data Inventory

### Net-a-Porter Dataset

- Rows: 23,161 | Columns: 4

| Variable    | Type     | Description                               |
|-------------|----------|---|
| brand       | Nominal  | Brand name                                |
| description | Nominal  | Description of the product                |
| price_usd   | Discrete | Retail price of the product in US dollars |
| Type        | Nominal  | Category of the product                   |

### Mr. Porter Dataset

- Rows: 20,347 | Columns: 4

| Variable    | Type     | Description                               |
|-------------|----------|---|
| brand       | Nominal  | Brand name                                |
| description | Nominal  | Description of the product                |
| price_usd   | Discrete | Retail price of the product in US dollars |
| Type        | Nominal  | Category of the product                   |

## Data Cleaning and Exploration

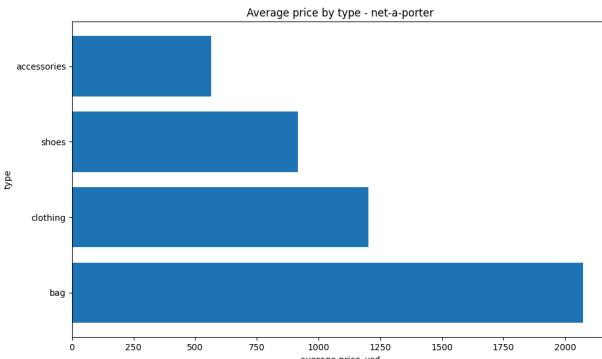
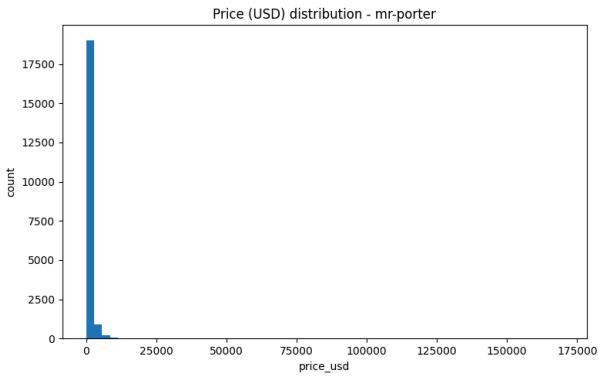
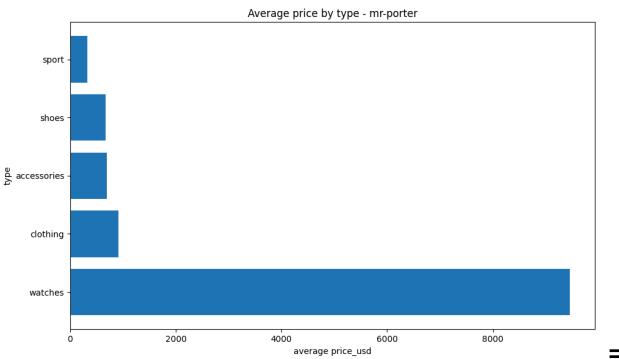
### Cleaning steps

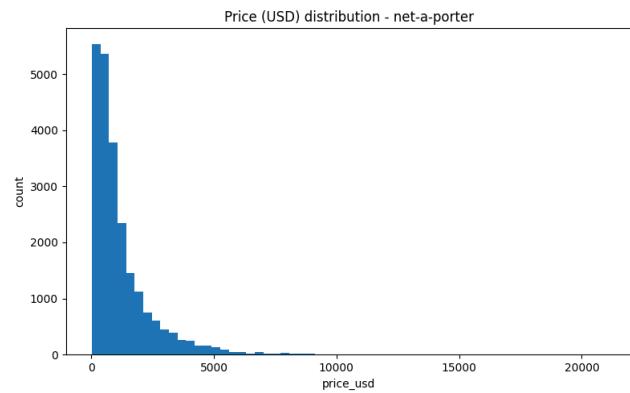
- Read CSV using pandas.
- Strip leading/trailing whitespace from text columns.
- Title-case brand names and strip whitespace.
- Coerce `price\_usd` to numeric (strip currency symbols and commas).
- Standardize `type` using a small mapping: `footwear` to `Shoes`, `rtw` to `Ready-to-Wear`, `bags` to `Bags`.
- Drop exact duplicate rows. (2 from Net-a-porter 3 from Mr. Porter)

## Findings

- Net-A-Porter: ~23k rows; mean price  $\approx$  1,197.80; median  $\approx$  780; no missing values in columns.
- Mr Porter: ~20k rows; mean price  $\approx$  1,013.14; median  $\approx$  460; no missing values in columns.
- Merged 99th percentile threshold: 7,200 USD; 432 items above that threshold in the merged data.

## Visualizations





## Vesitaire Collective [Derin]

### Data Inventory

Rows: 900,514 | Columns : 36

Below are the 15 that are most applicable to our project.

| Variable              | Type                     | Description  |
|-----------------------|--------------------------|--|
| product_id            | Nominal                  | Id of the product  |
| product_type          | Nominal                  | Type of product  |
| product_name          | Nominal                  | Name of product  |
| product_description   | Nominal                  | Description of Product   |
| product_keywords      | Nominal                  | Keywords in product listing  |
| product_gender_target | Nominal                  | Gender target of the product   |
| product_category      | Nominal                  | Cateogry of the product (Men Clothing, Women clothing, Accessories, Shoes,etc. ) |
| Brand_name            | Nominal                  | Name of the brand  |
| product_season        | Ordinal                  | Season og the product (Autumn/WInter, Spring/Summer)                             |
| Product_condition     | Ordinal                  | Condition of the product (very good, never worn, etc)                            |
| product_material      | Nominal                  | The product material (leather, cotton, etc.)                                     |
| product_color         | Nominal                  | Color of the product (black, blue, yellow, etc.)                                 |
| price_usd             | Quantitative             | Price of the product in USD  |
| seller_price          | Quantitative             | Seller price of the product  |
| Seller_country        | Categorical (Geographic) | Location of seller   |

### Data Cleaning and Exploration

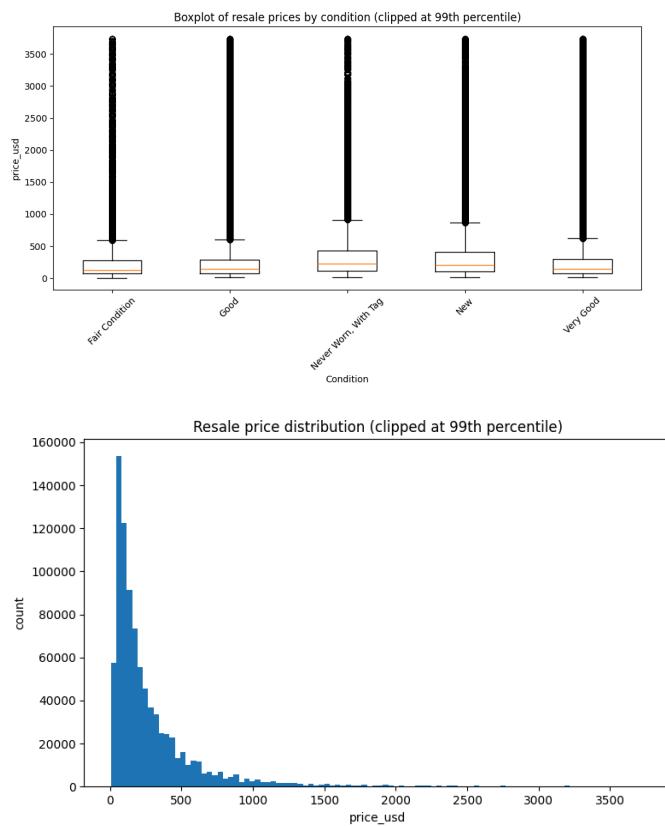
- `price\_usd` stats: count ≈ 900k, mean ≈ 387, min ≈ 6.13, max ≈ 632,610

- - Many auxiliary columns have missing values (notably `usually\_ships\_within`) and `original\_price\_usd` is not consistently present.
- Aggregates (`vestiaire\_aggregates.py`)
- Number of unique brands: ~8,884
- Unique categories detected: 6 (based on heuristic mapping)
- Top 10 brands by average resale price are watch/jewelry heavy (e.g., Richard Mille, Patek Philippe, Graff)

Cleaning steps:

- Remove missing/zero prices
- Drop exact duplicates
- Title-case and strip brand names
- Map categories heuristically to Bags/Shoes/Ready-to-Wear/Other
- Remove items above 99th percentile

## Visualization



# Google Trends [Derin ]

## Data Inventory

### Regional break down

- Rows: 250 | Columns : 4
- 

| Variable           | Type         | Description        |
|--------------------|--------------|--------------------|
| Week               | Temporal     |                    |
| Coach (Worldwide)  | Quantitative | Amount of Interest |
| Hermes (Worldwide) | Quantitative | Amount of Interest |
| Gucci (Worldwide)  | Quantitative | Amount of Interest |

### Comparatie Timeline

- Rows: 261 | Columns: 4

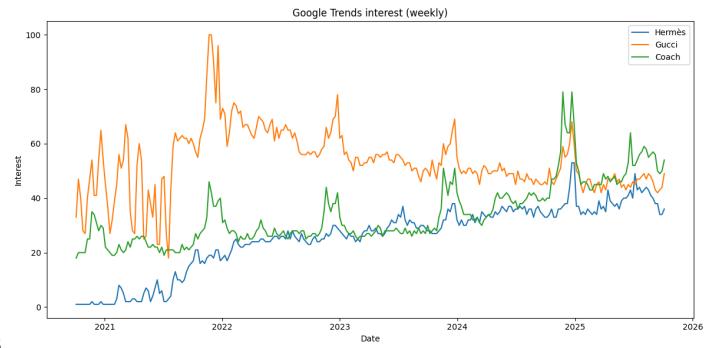
| Variable | Type         | Description         |
|----------|--------------|---------------------|
| Country  | Geographic   | Country of interest |
| Coach    | Quantitative | Percent Interest    |
| Hermes   | Quantitative | Percent Interest    |
| Gucci    | Quantitative | Percent Interest    |

### Cleaning & processing

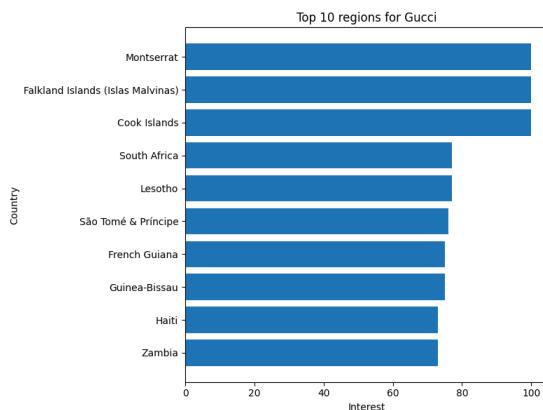
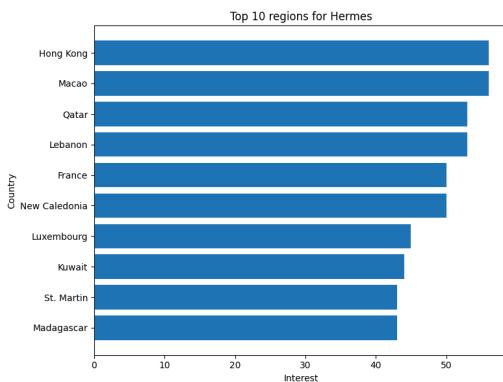
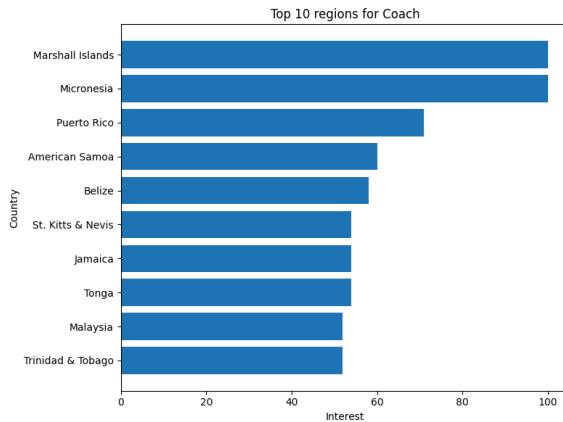
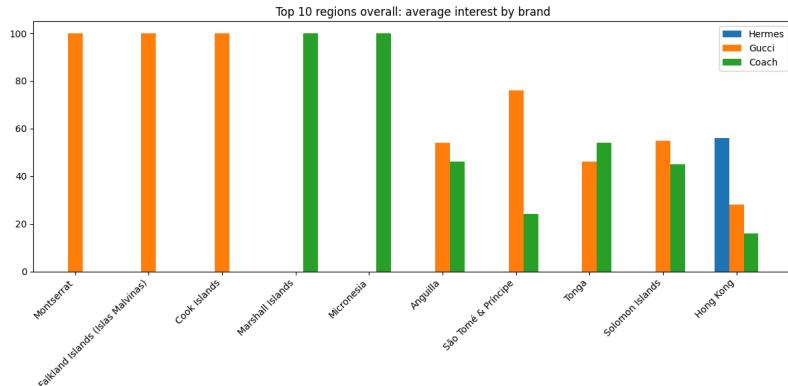
- Didn't really need any cleaning
- Parse dates in the first column; coerce brand columns to numeric.

### Findings

- Average interest values (from current run): Hermès ≈ 25.5, Gucci ≈ 53.3, Coach ≈ 34.1.



S



# Brand Reports [Cindy]

Data Wrangling and Cleaning Process:

1. Manually recording data from official brand reports (annual, quarterly reports) into google spreadsheets
2. Manually cleaning-up spreadsheets by making each observation unique (ie. similar to pivot\_longer() function in R)
3. In spreadsheets, force numerical variables such as percentages% into decimals, so R would recognize the data in correct type
4. Remove duplicated rows and columns
5. Rename column names to cleaned names (ie. "Published Growth" into "published\_growth")
6. Perform EDA by summarize stats table and visualizations

## [Hermes Google Sheet](#)

We will mainly look at sheets:

Hermes Revenue (in millions of Euro) by Geography

Hermes Revenue (in millions of Euro) by Sector

Variable Types:

| 🕒 Hermes_Financials_Hermes_Revenue_in_millions_of_Euro_by_Geography |   | 108 obs. of 6 variables  |
|---|---|--|
| \$ revenue  | : | num [1:108] 740 1088 1828 815 3574 ...                                   |
| \$ year   | : | num [1:108] 2025 2025 2025 2025 2025 ...                                 |
| \$ period   | : | chr [1:108] "H1" "H1" "H1" "H1" ...                                      |
| \$ region   | : | chr [1:108] "France" "Europe (excl. France)" "Total Europe" "Japan" ...  |
| \$ published_growth   | : | num [1:108] 0.087 0.121 0.107 0.176 0.015 0.042 0.095 0.163 0.071 NA ... |
| \$ at_constant_exchange_rates                                       | : | num [1:108] 0.087 0.126 0.11 0.16 0.03 0.051 0.117 0.172 0.081 NA ...    |

| 🕒 Hermes_Financials_Hermes_Revenue_in_millions_of_Euro_by_Sector |   | 96 obs. of 6 variables  |
|--|---|---|
| \$ revenue   | : | num [1:96] 3578 2255 447 1056 248 ...                                       |
| \$ year  | : | num [1:96] 2025 2025 2025 2025 2025 ...                                     |
| \$ period  | : | chr [1:96] "H1" "H1" "H1" "H1" ...  |
| \$ sector  | : | chr [1:96] "Leather Goods and Saddlery" "Ready-to-wear and Accessories" ... |
| \$ published_growth  | : | num [1:96] 0.113 0.043 0.026 0.092 -0.041 -0.089 0.072 0.071 NA NA ...      |
| \$ at_constant_exchange_rates                                    | : | num [1:96] 0.124 0.055 0.035 0.103 -0.038 -0.079 0.075 0.081 NA NA ...      |

Summary stats tables:

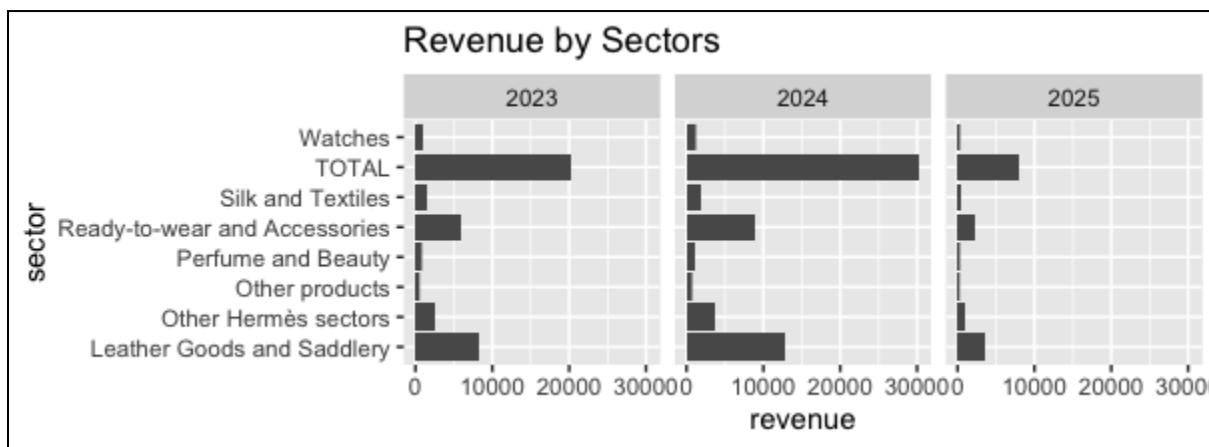
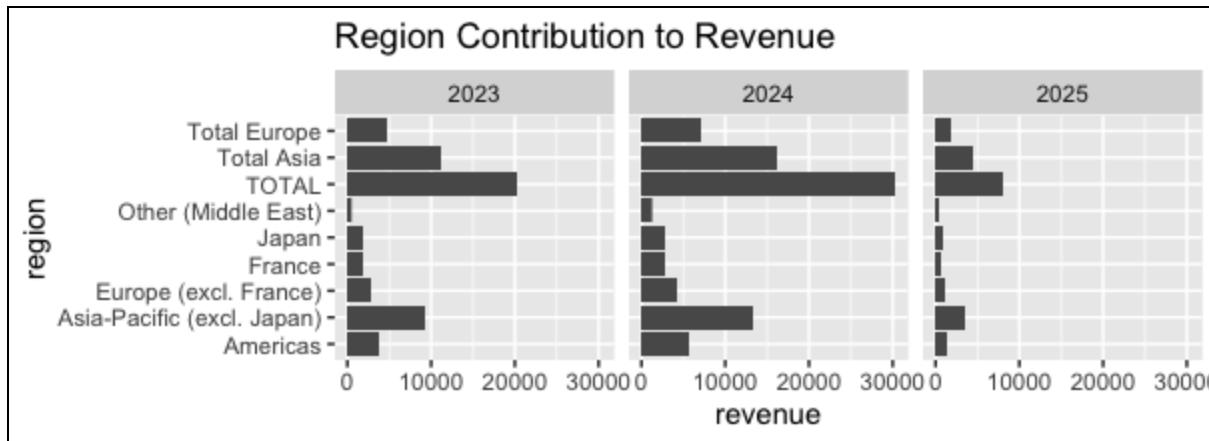
```
> summary(Hermes_Financials_Hermes_Revenue_in_millions_of_Euro_by_Geography)
      revenue           year        period       region
Min.    : 72.0   Min.   :2023   Length:108   Length:108
1st Qu.: 438.2  1st Qu.:2024  Class :character  Class :character
Median  : 959.5 Median  :2024  Mode   :character  Mode   :character
Mean    : 1904.5 Mean   :2024
3rd Qu.: 2034.5 3rd Qu.:2024
Max.    :15170.0 Max.   :2025

published_growth at_constant_exchange_rates
Min.    :0.0010  Min.   :0.01000
1st Qu.:0.0760  1st Qu.:0.08925
Median  :0.1285  Median  :0.12900
Mean    :0.1716  Mean   :0.18050
3rd Qu.:0.1688  3rd Qu.:0.17200
Max.    :1.2520  Max.   :1.23200
NA's    :54      NA's   :54
```

```
> summary(Hermes_Financials_Hermes_Revenue_in_millions_of_Euro_by_Sector)
      revenue           year        period       sector
Min.    : 74.0   Min.   :2023   Length:96   Length:96
1st Qu.: 163.8  1st Qu.:2024  Class :character  Class :character
Median  : 498.5 Median  :2024  Mode   :character  Mode   :character
Mean    : 1543.1 Mean   :2024
3rd Qu.: 1657.0 3rd Qu.:2024
Max.    :15170.0 Max.   :2025

published_growth at_constant_exchange_rates
Min.    :-0.18900  Min.   :-0.18200
1st Qu.: 0.01525  1st Qu.: 0.03275
Median  : 0.08000  Median  : 0.08400
Mean    : 0.06669  Mean   : 0.07508
3rd Qu.: 0.12700  3rd Qu.: 0.14725
Max.    : 0.21700  Max.   : 0.21500
NA's    :48      NA's   :48
```

Simple EDA Visualizations:



[Gucci Google Sheet](#)

Variable Types:

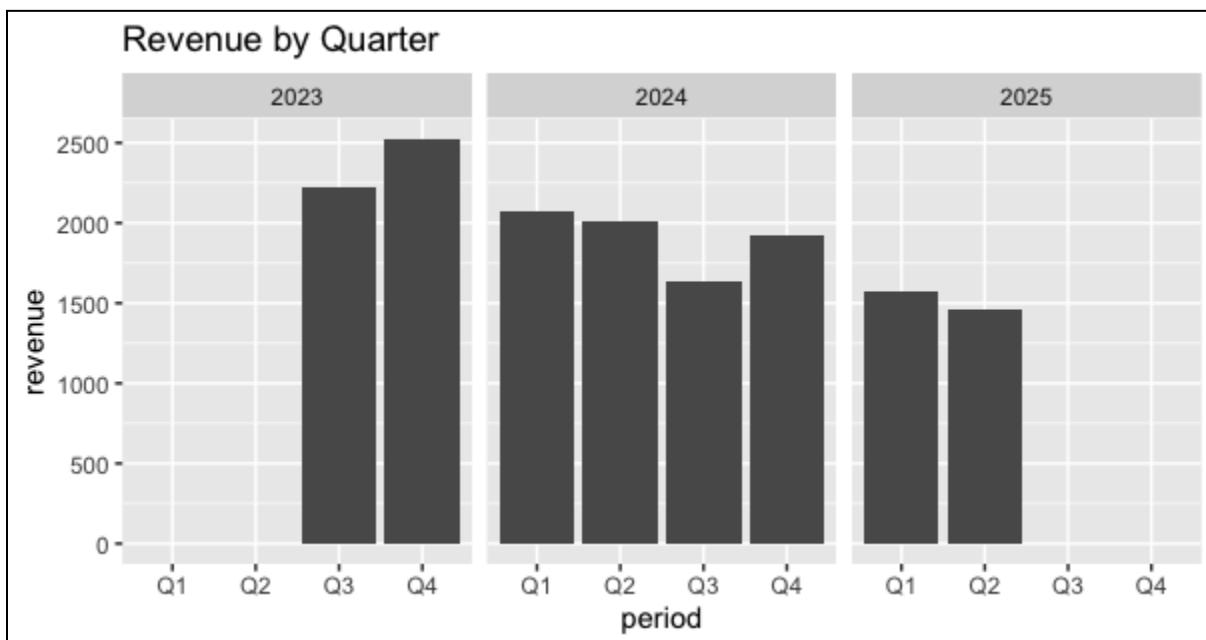
|                      |   |  |
|----------------------|---|--|
| Gucci_Financials     | 16 obs. of 7 variables  |  |
| \$ amount            | : num [1:16] 3027 4085 1571 2079 1456 ...                     |  |
| \$ year              | : num [1:16] 2025 2024 2025 2024 2025 ...                     |  |
| \$ period            | : chr [1:16] "H1" "H1" "Q1" "Q1" ...                          |  |
| \$ category          | : chr [1:16] "Revenue" "Revenue" "Revenue" "Revenue" ...      |  |
| \$ reported_change   | : num [1:16] -0.26 NA -0.24 -0.21 -0.27 -0.2 NA NA -0.24 N... |  |
| \$ comparable_change | : num [1:16] -0.25 NA -0.25 -0.18 -0.25 -0.19 NA NA -0.24 ... |  |
| \$ change            | : num [1:16] NA NA NA NA NA NA -0.52 NA NA NA ...             |  |

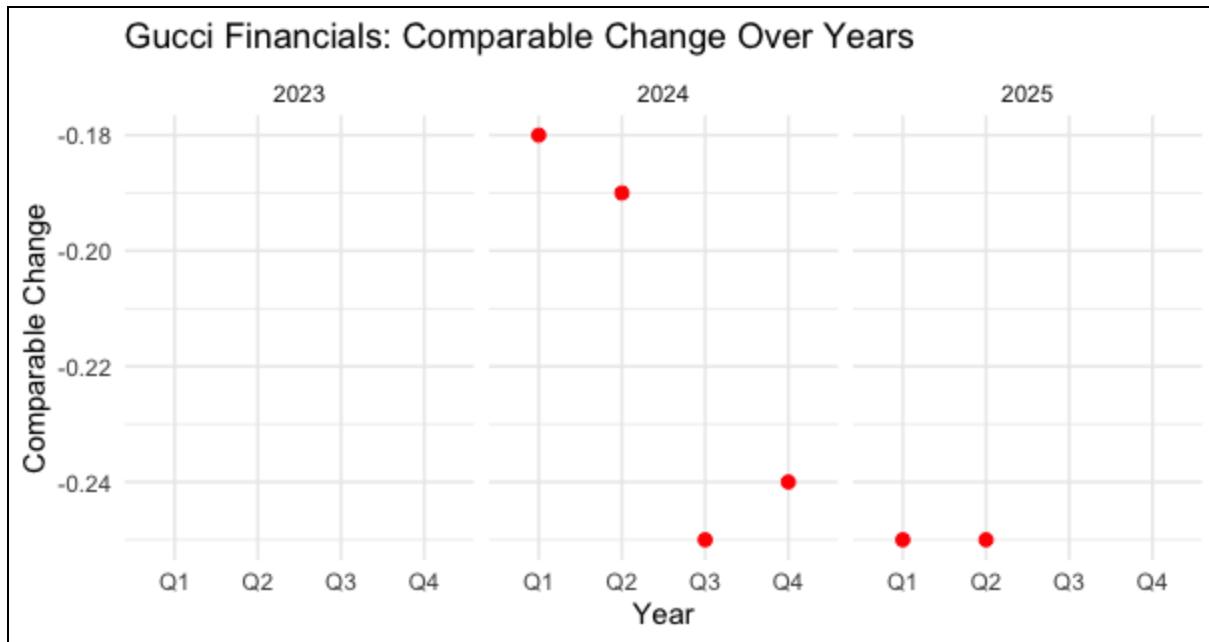
Summary stats table:

```
> summary(Gucci_Financials)
   amount          year       period      category
Min.   : 486   Min.   :2023   Length:16    Length:16
1st Qu.:1596  1st Qu.:2024  Class :character  Class :character
Median  :2042  Median  :2024  Mode   :character  Mode   :character
Mean    :2901  Mean    :2024
3rd Qu.:3086  3rd Qu.:2024
Max.   :9873   Max.   :2025

  reported_change  comparable_change     change
Min.   :-0.2700    Min.   :-0.2500    Min.   :-0.5200
1st Qu.:-0.2600    1st Qu.:-0.2500    1st Qu.:-0.5175
Median :-0.2400    Median :-0.2450    Median :-0.5150
Mean   :-0.2387    Mean   :-0.2275    Mean   :-0.5150
3rd Qu.:-0.2250    3rd Qu.:-0.2050    3rd Qu.:-0.5125
Max.   :-0.2000    Max.   :-0.1800    Max.   :-0.5100
NA's    :8          NA's    :8          NA's    :14
```

Simple EDA Visualizations:





### [Coach Google Sheet](#)

We will mainly look at sheets:

Coach Sales by Category

Coach Sales by Region

Variable Types:

|  |   |  |
|--|---|--|
| Coach_Financials_Coach_Sales_by_Category | 30 obs. of 4 variables                    |  |
| \$ category                              | : chr [1:30] "Handbags" "Accessories" ... |  |
| \$ fiscal_year_ended                     | : chr [1:30] "June 28, 2025" "June 28,... |  |
| \$ amount                                | : num [1:30] 3223 1540 342 493 5598 ...   |  |
| \$ percentage_of_total_segment_net_sales | : num [1:30] 0.576 0.275 0.061 0.088 1... |  |

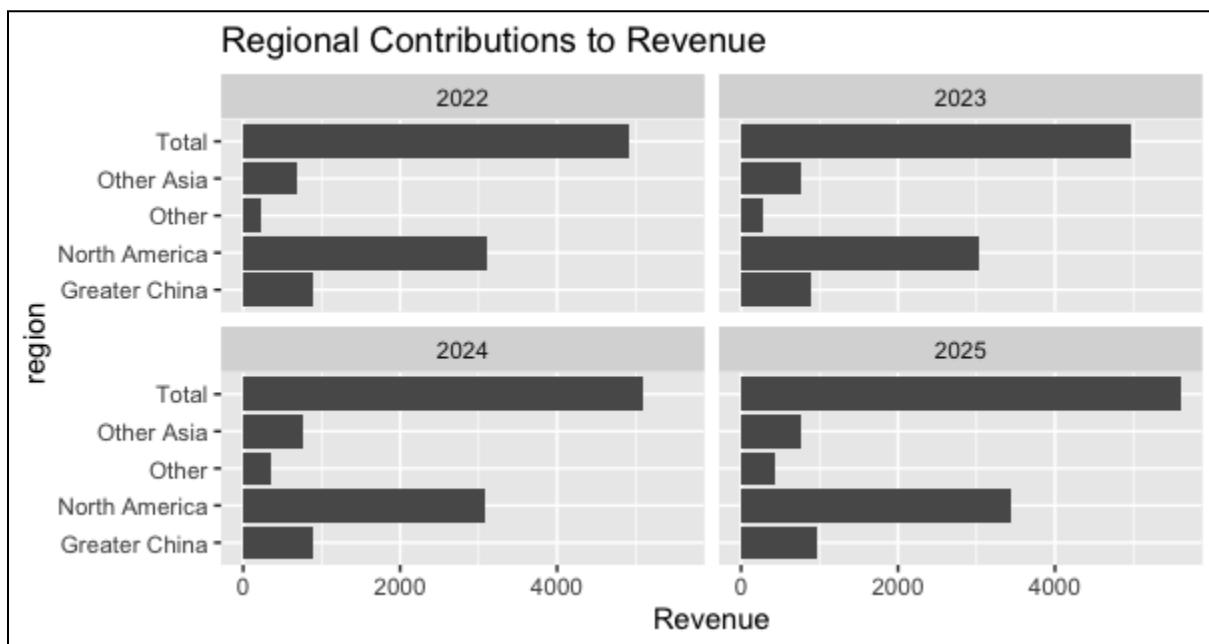
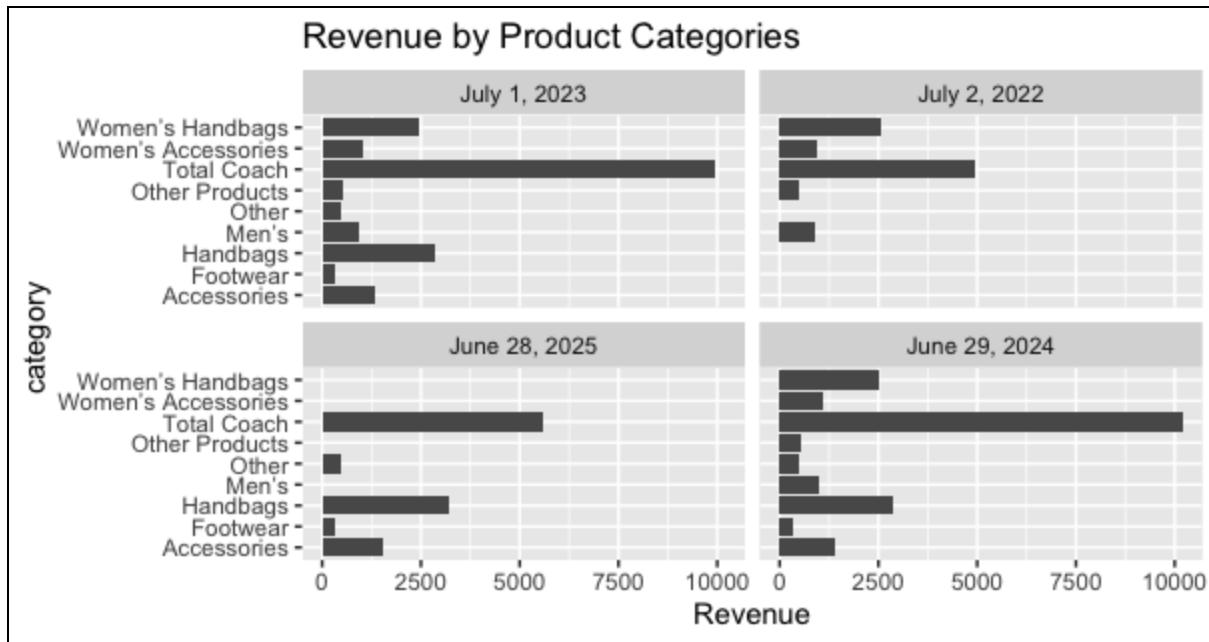
|  |   |  |
|--|---|--|
| Coach_Financials_Coach_Sales_by_Region | 20 obs. of 3 variables  |  |
| \$ region                              | : chr [1:20] "North America" "Greater China" "Other Asia" "Othe..." |  |
| \$ fiscal _year                        | : num [1:20] 2025 2025 2025 2025 2025 ...                           |  |
| \$ net_sales                           | : num [1:20] 3430 970 770 429 5598 ...                              |  |

Summary stats tables:

```
> summary(Coach_Financials_Coach_Sales_by_Category)
  category      fiscal_year_ended      amount
Length:30      Length:30          Min.   : 311.5
Class :character Class :character 1st Qu.: 536.5
Mode  :character Mode  :character Median  :1202.7
                           Mean   :2042.1
                           3rd Qu.:2879.2
                           Max.   :5598.5
percentage_of_total_segment_net_sales
Min.   :0.06100
1st Qu.:0.09375
Median :0.21450
Mean   :0.34967
3rd Qu.:0.57225
Max.   :1.00000
```

```
> summary(Coach_Financials_Coach_Sales_by_Region)
  region      fiscal _year      net_sales
Length:20      Min.   :2022      Min.   : 235.0
Class :character 1st Qu.:2023    1st Qu.: 737.5
Mode  :character Median :2024    Median : 899.4
                           Mean   :2024    Mean   :2057.6
                           3rd Qu.:2024    3rd Qu.:3184.5
                           Max.   :2025    Max.   :5598.5
```

Simple EDA Visualizations:



## Reflection

- How do average handbags/shoes/accessories prices compare across Hermès, Gucci, and Coach?
- Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand?
- Which categories are most protected from markdowns?
- Do new product launches start at higher price points over time?
- Is there a seasonal pattern to markdowns?
- Do certain categories (bags, RTW, accessories) dominate across all three brands, or only some?
- Which styles stay in collections for many years, and which disappear quickly?
  - Notes: not sure about because I don't know if we have that much data
- How does product variety differ (number of SKUs per season) across brands?
- What role do limited editions and collaborations play in Gucci's product mix?
- How much of their original price do Hermès bags retain on resale?
- How fast do Gucci bags lose value compared to Hermès and Coach?
- Which region contributes the most to Hermès' sales? To Gucci? To Coach?
- How does search interest (Google Trends) differ by region?
- How often do each brand's flagship products trend in Google searches regionally?
- How does consumer interest in these brands fluctuate over time—are they seasonal, trend-driven, or stable?
- What drives revenue growth across these brands?

## Milestone 5: Data Overview

When we first mapped out our project, we envisioned combining live web-scraped data from multiple luxury retailers to answer detailed questions about pricing dynamics, markdowns, and product variety across Hermès, Gucci, and Coach. However, access restrictions and anti-scraping protections made that infeasible. As a result, our final data collection pivoted toward publicly available, academic, and secondary sources, combining both quantitative and qualitative datasets to ensure a well-rounded analysis of the global luxury market.

The Net-a-Porter and Mr Porter CSVs provided a strong baseline for retail pricing data. These datasets include over 43,000 rows combined, with detailed information on brand, price, and product type. They capture a wide range of women's and men's luxury items, giving us reliable insight into contemporary retail price levels although they primarily represent Gucci, leaving gaps for Hermès and Coach.

To fill these gaps, we incorporated the Vestiaire Collective resale dataset ( $\approx 900k$  rows), which offers extensive data on pre-owned luxury fashion products. This dataset captures consumer resale behavior, allowing us to compare secondhand market value retention and category

dominance. We complemented this with Google Trends exports, which show five-year patterns of regional and temporal search interest for Hermès, Gucci, and Coach. This data provides an indirect view of consumer attention and geographic demand.

Additionally, we collected brand report datasets from official financial filings for Hermès, Kering (Gucci), and Tapestry (Coach). These spreadsheets provide annual and quarterly revenue data, segmented by region and product category, allowing for comparisons in financial performance, growth rates, and global reach.

To further contextualize Hermès's performance, we used two supplementary datasets:

- Revenue Breakdown by Product Segment (2024) detailing the share of total revenue from categories such as leather goods, ready-to-wear, silk, perfume, and watches.
- Revenue Breakdown by Geographical Region (2024) showing regional contributions to Hermès's global revenue, with Asia-Pacific (44%) emerging as the dominant market.

For Coach, we expanded our scope to include market-specific data:

- The Top-Selling Coach Bags dataset from Accio, which provides product-level metrics like retail price, monthly revenue, sales volume, and growth trend, identifying key models driving brand performance.
- The Marketing Mix of Coach (2025) dataset, a qualitative source analyzing the brand's strategic positioning through its 4Ps: Product, Price, Place, and Promotion. This text-based dataset helps contextualize quantitative pricing and sales data by revealing how Coach's brand identity, pricing strategy, and promotional activities shape its performance in the accessible luxury segment.

Finally, the Vestiaire Collective App Downloads (2021–2025) dataset connects directly to consumer engagement trends, quantifying monthly app usage as a proxy for overall activity in the secondhand luxury market.

## AI Interaction Examples

I have two CSVs called `data_raw/netaporter.csv` and `data_raw/mrporter.csv`. Each has four columns: brand, description, price\_usd, and type.

Write Python code using pandas to:

- load both files
- show the number of rows and columns
- list the column data types
- count missing values
- calculate min, max, mean, and median price per dataset.

I'll create a small todo plan, then add a Python script that reads the two CSVs and prints the requested summaries. After creating the file I'll run it to validate the output. Expected outcome: a new script `scripts/describe_datasets.py` and its runtime output shown here.

I'll first write the todo list (plan) and mark the first task in-progress. Then I'll add the script. Finally I'll run it.

✓ Created 3 todos

Now I'll create the Python script file `scripts/describe_datasets.py`. Expected outcome: file added to

Write pandas code to create:  
a histogram of price\_usd for each dataset  
a bar chart of average price by type  
a table showing top 10 brands by number of products

I'll update the todo list to start the work of extending the script (add plots and top-10 table), then patch `describe_datasets.py` to add plotting and table features, and finally run the script to validate outputs. Expected outcome: updated script, saved plot files under `output/`, and printed/saved top-10 brand tables.

✓ Starting (1/3) Update script for plots & table

Now I'll update `describe_datasets.py` to add plotting and top-1

Find any outliers or extreme values.

Compute the 99th percentile of price\_usd and list all products above it.

Also show rows with missing brand or type values.

I'll add a todo, create a new script `scripts/find_outliers.py` that reads the merged cleaned CSV, computes the 99th percentile and lists products above it, and finds rows missing `brand` or `type`, then run it and save results to `output`. Expected outcome: new script, printed summary, and two CSV outputs (`price_above_99th.csv`, `missing_brand_or_type.csv`).

✓ Starting (1/3) Add outlier detection script

Now I'll create `scripts/find_outliers.py`. It will load `merged_clean.csv`, compute 99th percentile excluding NaNs, rows above it, and save CSVs. Then I'll run the script.

✓ Created `find_outliers.py`



Sketch

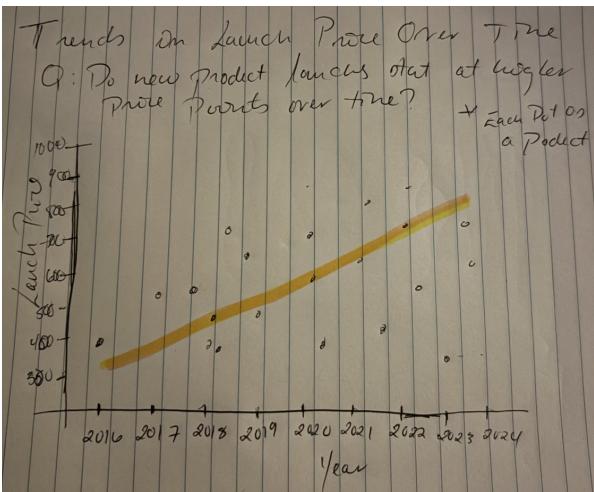
# Milestone 6: Sketch

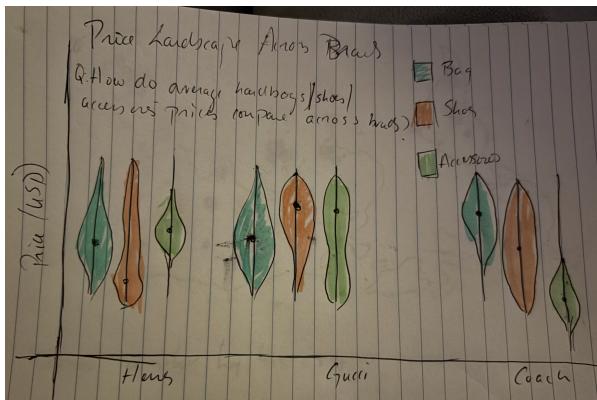
Aderinsola Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

Derin Adeleke

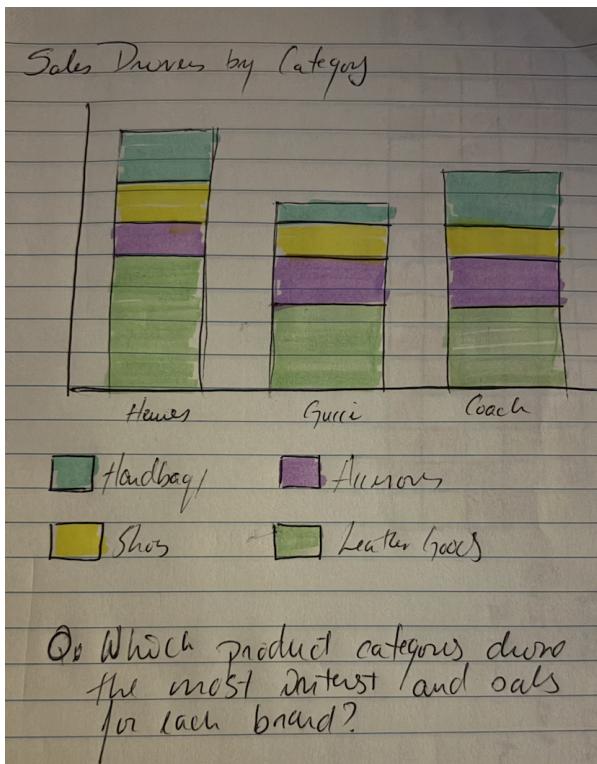
| Sketch  | Details   |
|---|---|
|   | <p>Title: Regional Popularity Map<br/>Question: How does search interest vary by region in country<br/>Key: Purple - Hermes   Orange - Gucci  Green - Coach</p> <p>Sketch number: 1<br/>Question ID: 13</p> |
|  | <p>Title: trends in launch price Over time<br/>Question: Do new product launches start at higher price points over time?<br/>*Every dot is a product</p> <p>Sketch number: 2<br/>Question ID: 4</p>         |



Title: Price Landscape Across Brands  
 Q: How do average handbags/shoes/accessories prices compare across brands?

Sketch number: 3

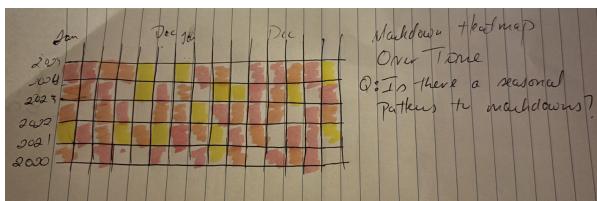
Question ID: 1



Title: Sales Drivers by Category  
 Question: which product categories drive the most interest and sales for each brand?

Sketch number: 4

Question ID: 2



Title: Markdown Heatmap Over time  
 Question: Is there a seasonal pattern to markdowns?

\*Each brand would get a heatmap

Sketch number: 5

Question ID: 5

Cindy Sun

### 3 Main Questions I Want to Address in Sketches (Using Brand Reports Datasets):

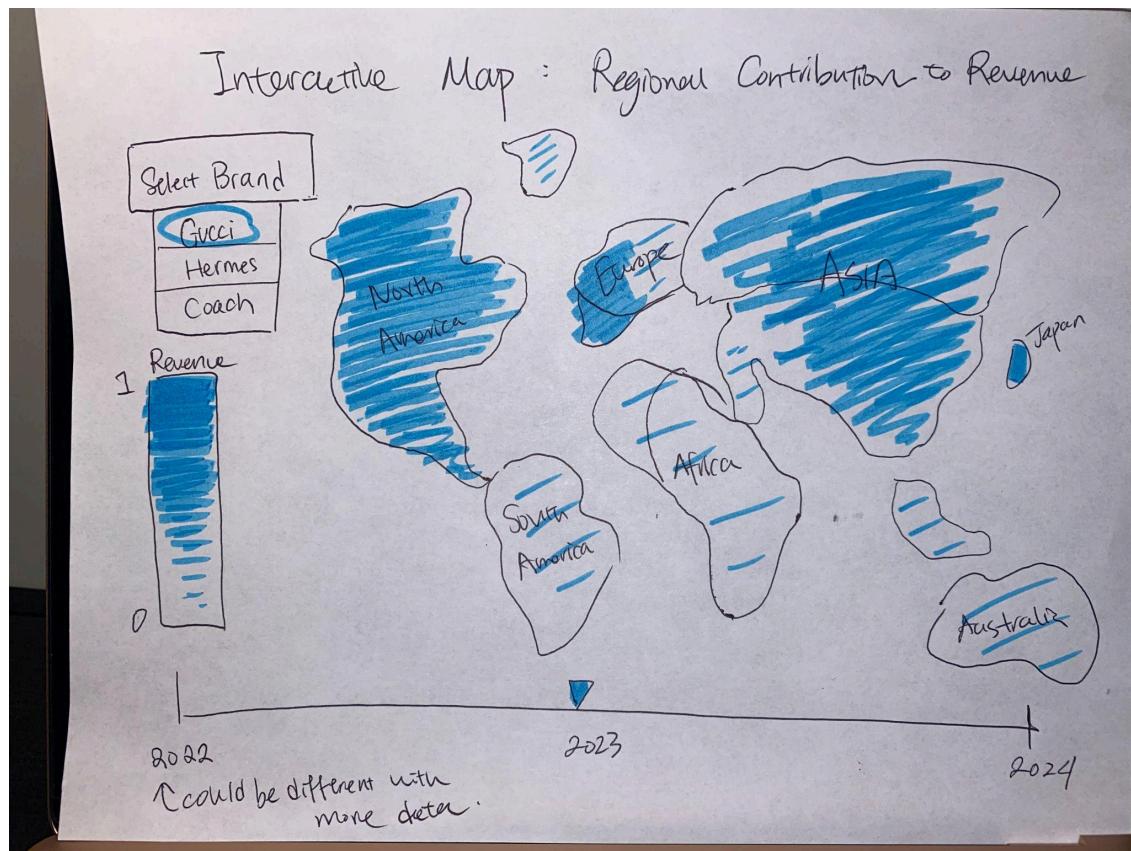
- Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand?
- Which region contributes the most to Hermès' sales? To Gucci? To Coach?
- Do certain categories (bags, RTW, accessories) dominate across all three brands, or only some?

Sketch number: 6

Question ID: 12

Question: Which region contributes the most to Hermès' sales? To Gucci? To Coach?

Note: This is one example map for Gucci (2023). The audience will have choices to select the brand and the year they are interested.

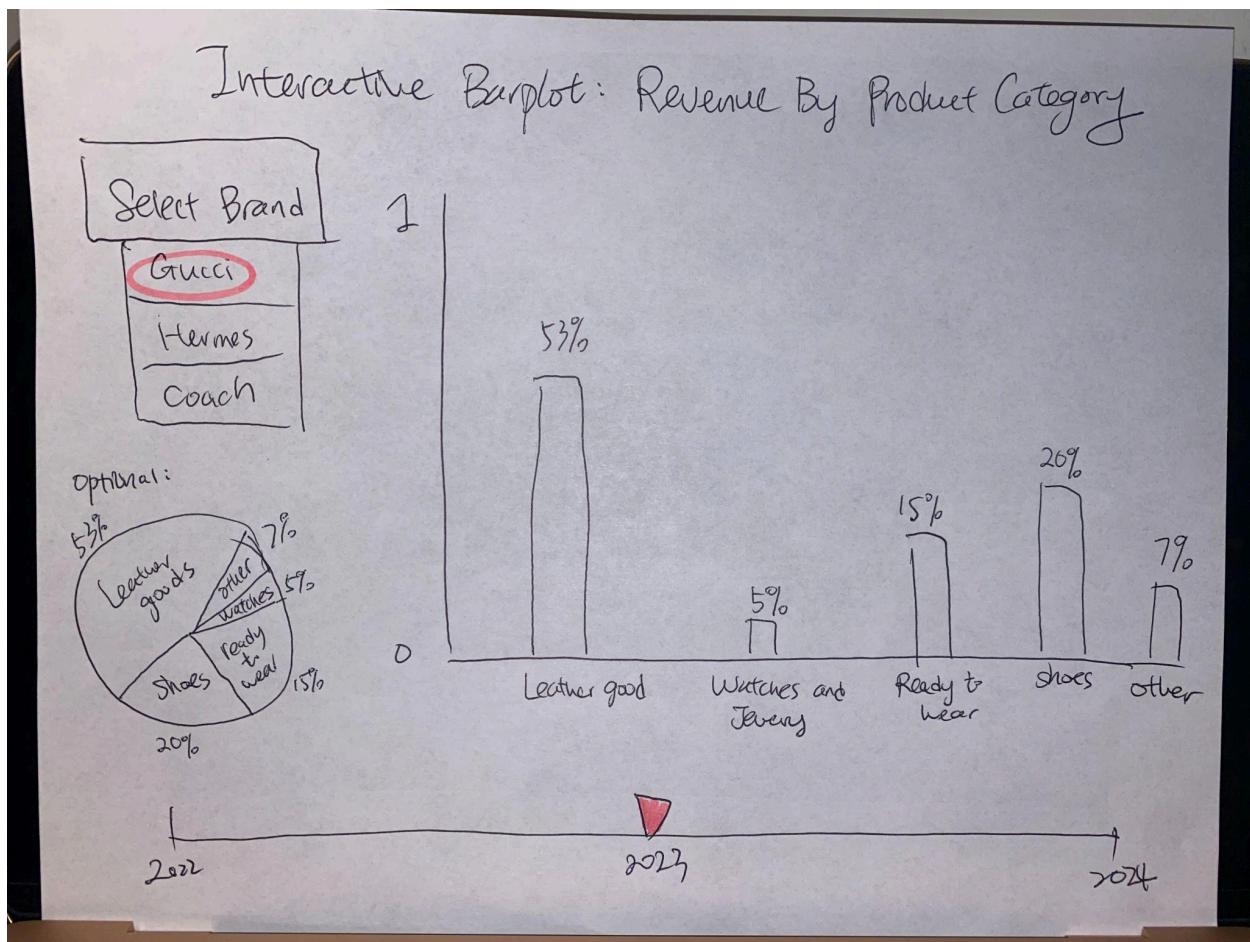


Sketch number: 7

Question ID: 2

Question: Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand?

This is one example map for Gucci (2023). The audience will have choices to select the brand and the year they are interested.

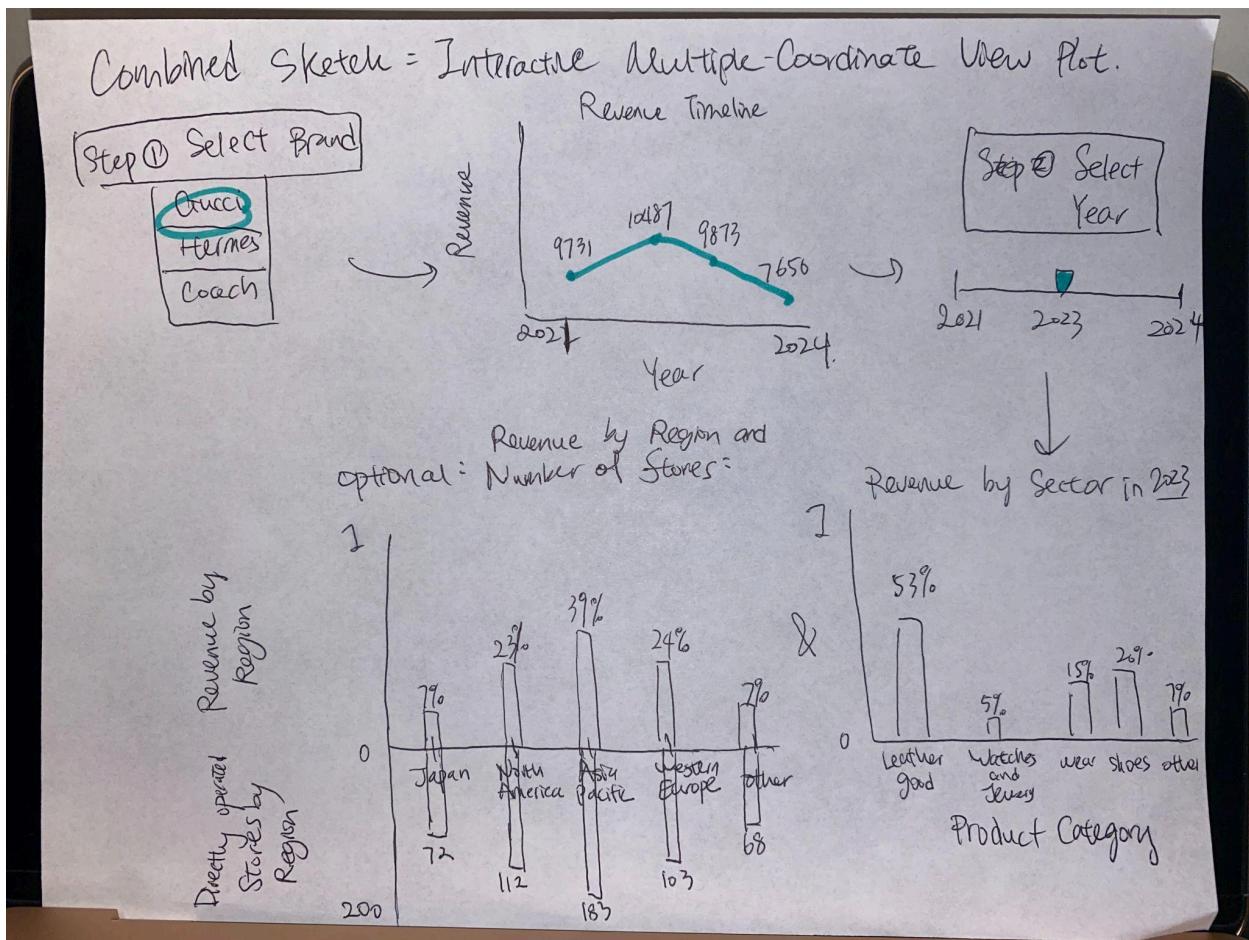


Sketch number: 8

Question ID: 2,12

Questions: Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand? Which region contributes the most to Hermès' sales? To Gucci? To Coach?

Note: Again, this is an interactive map. The audience will have choices to select the brand and the year they are interested. But instead of just answering one question (as the previous 2 sketches do), this interactive plot answers both.

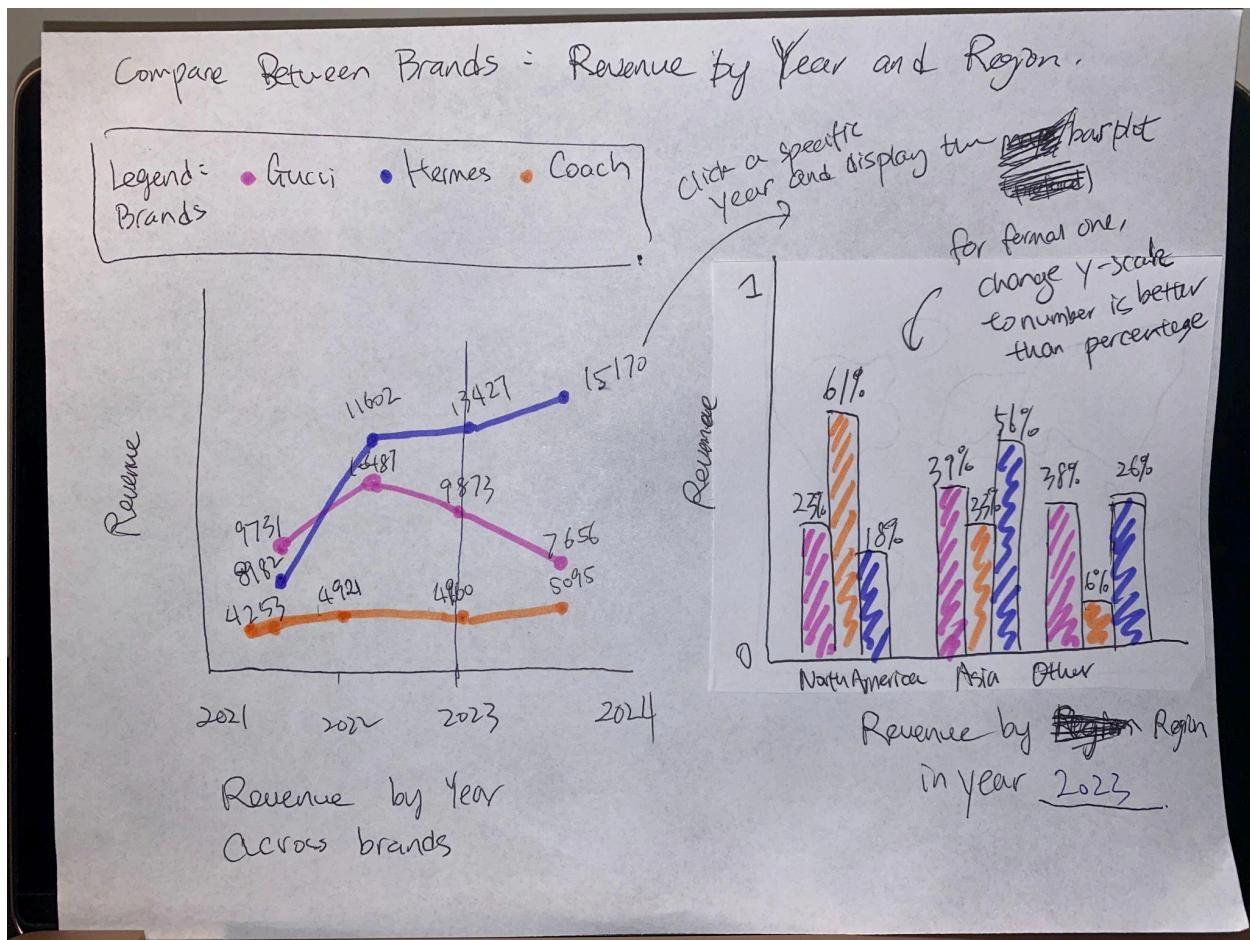


Sketch number: 9

Question ID: 12

Questions: Which region contributes the most to Hermès' sales? To Gucci? To Coach?

Note: Different from sketch 1, this plot allows straightforward comparison on regional contributions to revenue between brands.

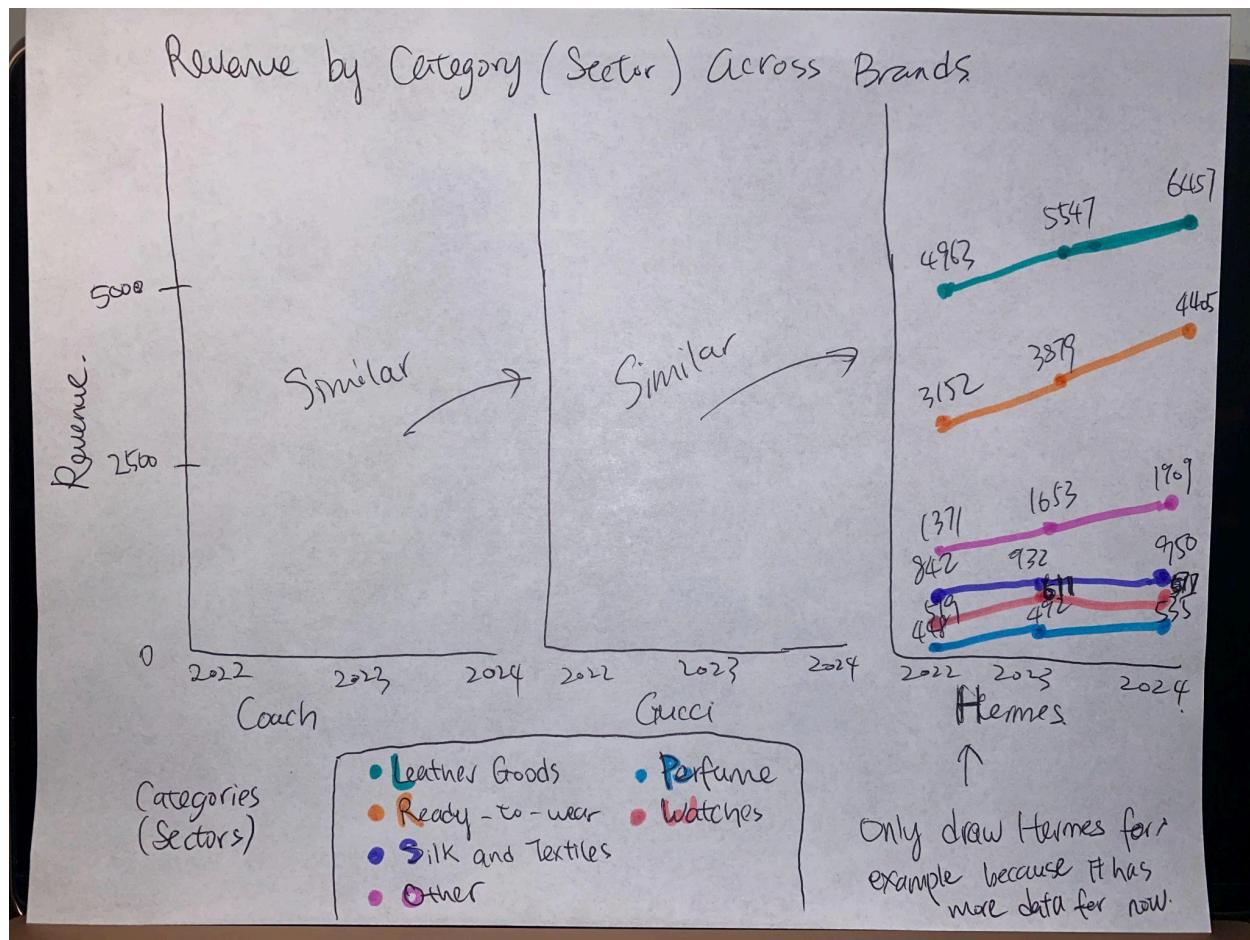


Sketch number: 10

Question ID: 2, 6

Questions: Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand? Do certain categories (bags, RTW, accessories) dominate across all three brands, or only some?

Note: Because I am still scraping data from brands' annual reports, I only created the line plot for Hermes for demonstration. The line graphs for Coach and Gucci will look similar — I just need to collect more data on them.



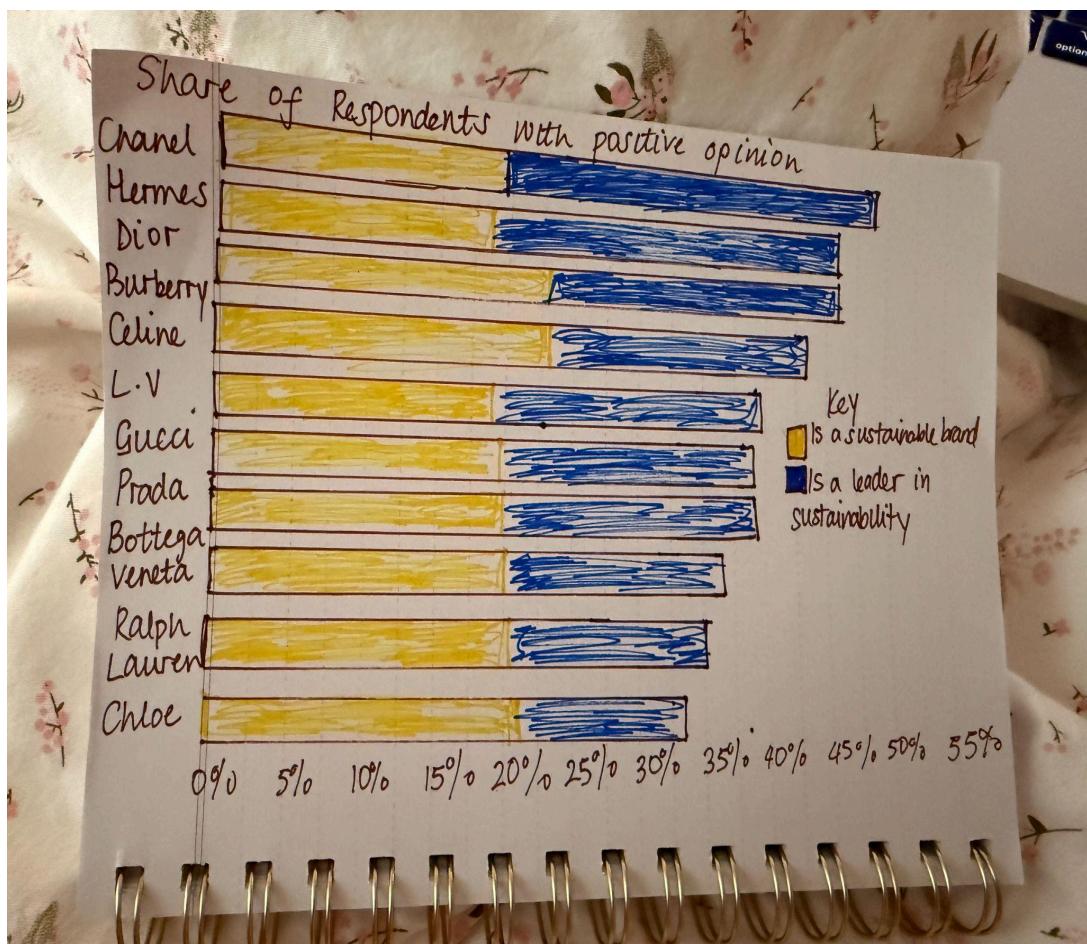
# Patience Madumera

Sketch number: 11

Question ID: 17

**Question:** What factors affect consumer interest in these brands?

**Title:** Leading luxury brands consumers view as sustainable worldwide in 2022



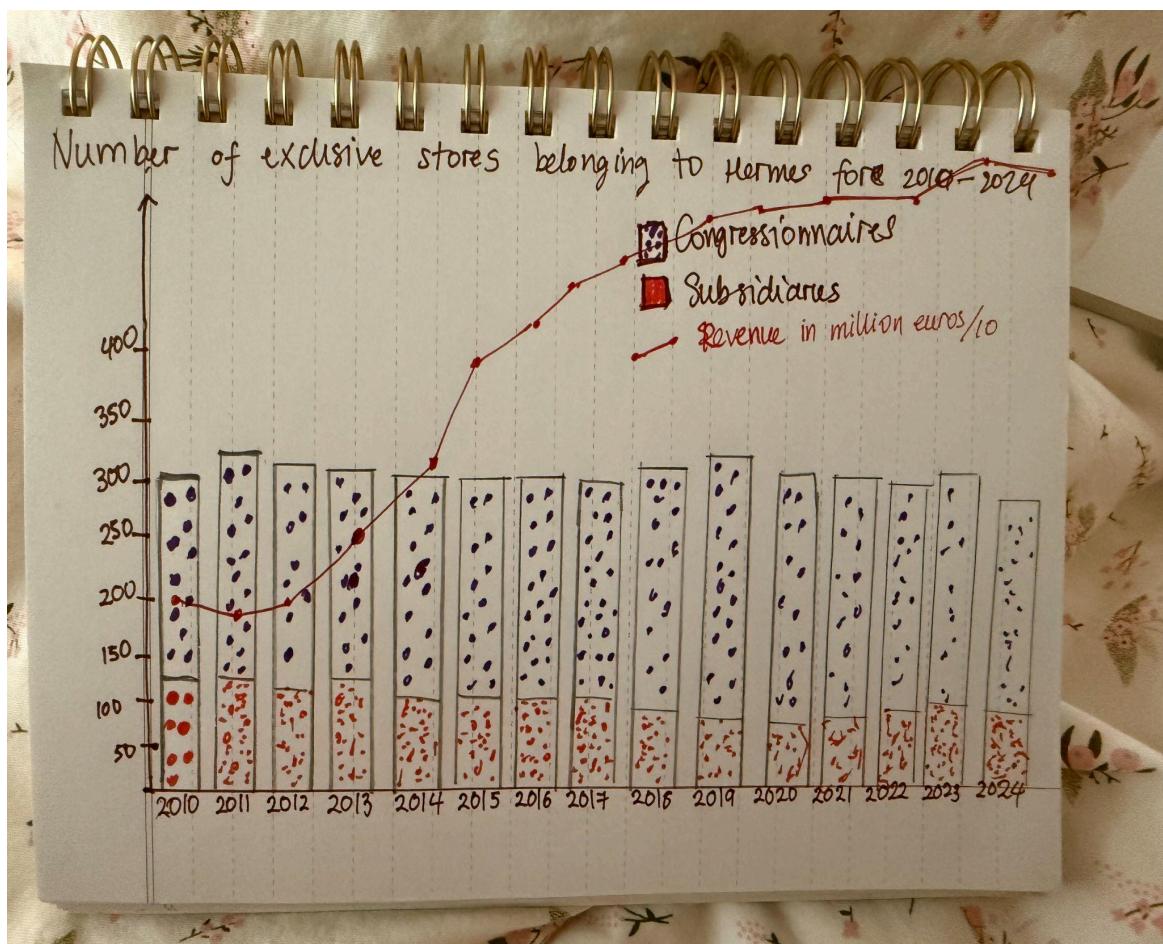
Sketch number: 12

Question ID:

**Question:** What contributes most to Hermès' sales(pricing/ quantity)? and What drives revenue growth?; it shows how expanding stores affects income. Ldk if we already had the question but I thought it was important to explore for all stores so we would have 3 of the charts

The bar chart show the number of stores belonging to hermes each year and the superimposed line graph shows revenue growth so it's just showing the relationship between quantity(number of stores) and revenue trends

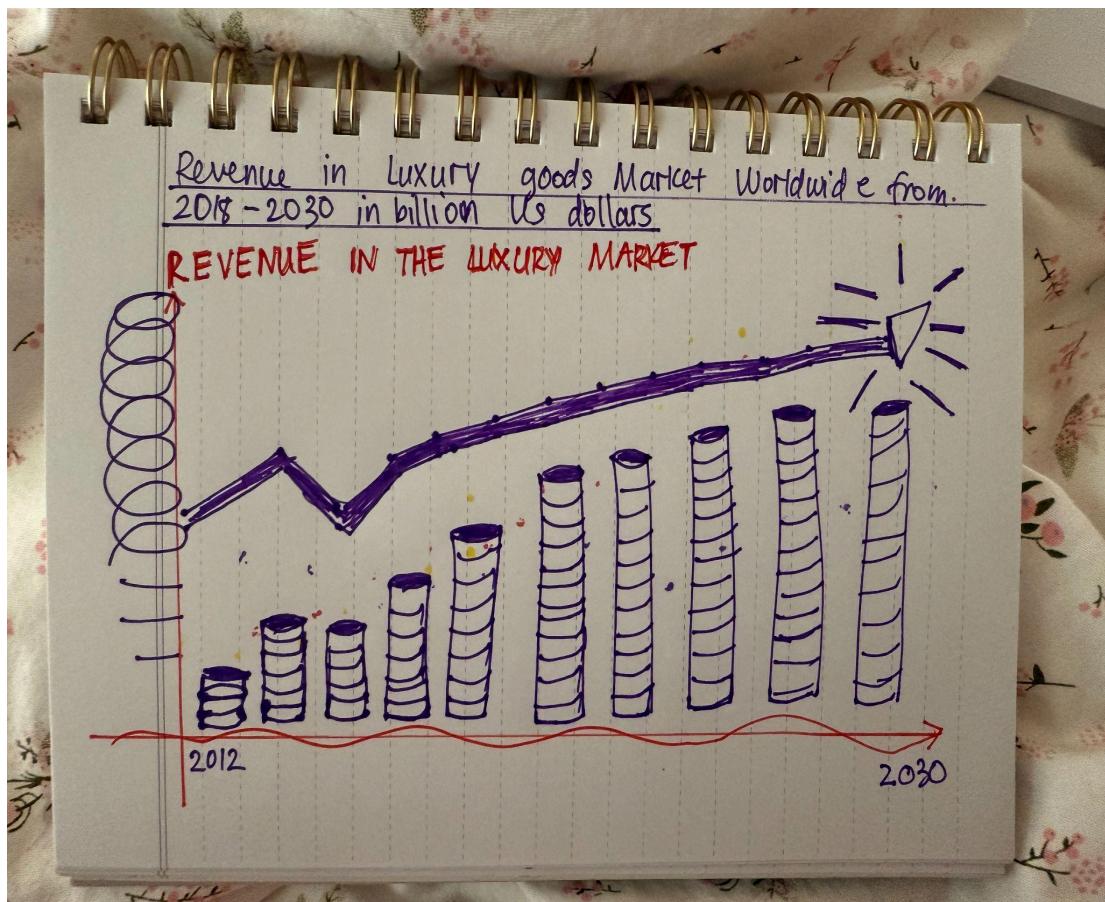
**Title :** Number of exclusive stores belonging to Hermes from 2010 to 2024



Sketch number: 13

Question ID: 16

**Question :** What drives revenue growth across these brands? it provides the macro backdrop showing how the whole sector expands



Sketch number: 14

Question ID: 12

**Question :** Which region contributes most to each brand's sales? and How does consumer interest differ by region?; it visualizes geographic expansion

**Title :** New Luxury store openings by Region in 2024

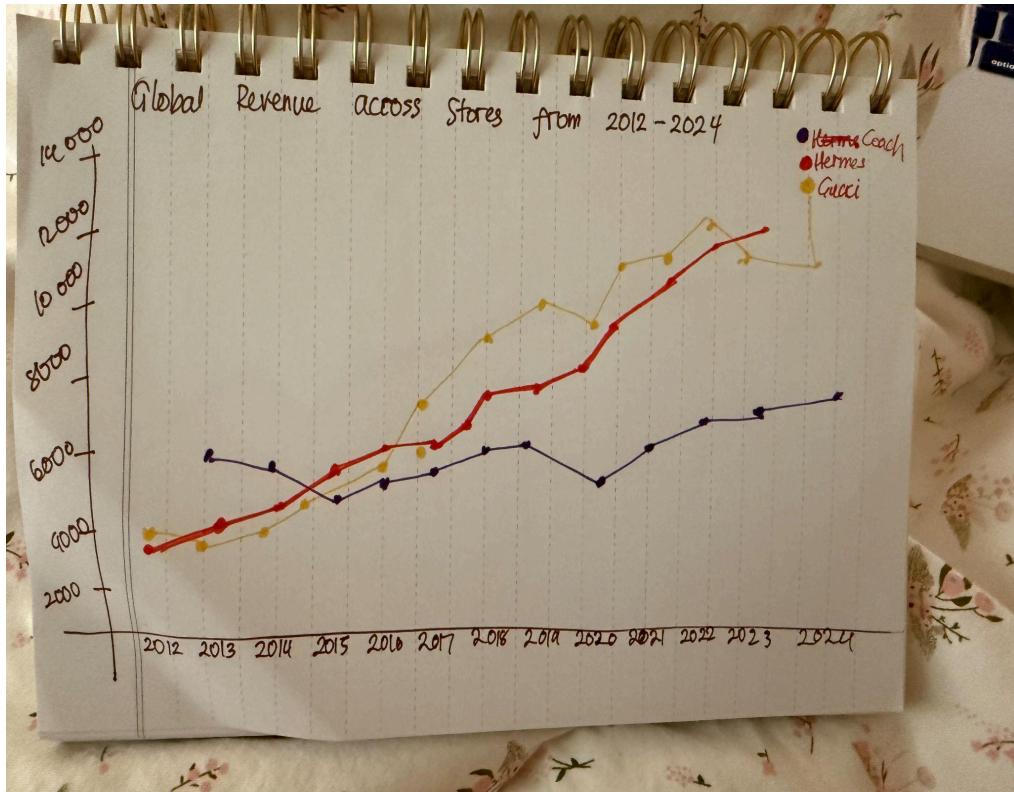


Sketch number: 15

Question ID: 16

**Question :** Which brand grows fastest? and What drives revenue growth?; it visualizes each brand's trajectory.

**Title:** Global Revenue Across Gucci, Hermes and Coach from 2012 to 2024





Decide

# Milestone 7: Decide

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

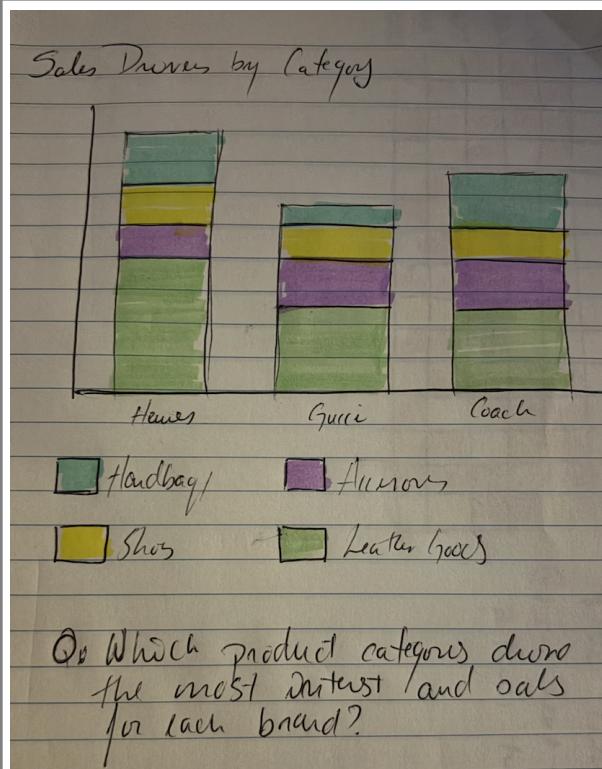
## Questions Recap (Milestone 4 Map):

1. How do average handbags/shoes/accessories prices compare across Hermès, Gucci, and Coach?
2. Which product categories (handbags, shoes, apparel) drive the most interest and sales for each brand?
3. Which categories are most protected from markdowns?
4. Do new product launches start at higher price points over time?
5. Is there a seasonal pattern to markdowns?
6. Do certain categories (bags, RTW, accessories) dominate across all three brands, or only some?
7. Which styles stay in collections for many years, and which disappear quickly?
8. How does product variety differ (number of SKUs per season) across brands?
9. What role do limited editions and collaborations play in Gucci's product mix?
10. How much of their original price do Hermès bags retain on resale?
11. How fast do Gucci bags lose value compared to Hermès and Coach?
12. Which region contributes the most to Hermès' sales? To Gucci? To Coach?
13. How does search interest (Google Trends) differ by region?
14. How often do each brand's flagship products trend in Google searches regionally?
15. How does consumer interest in these brands fluctuate over time—are they seasonal, trend-driven, or stable?
16. What drives revenue growth across these brands?
17. What factors affect consumer interest in these brands?

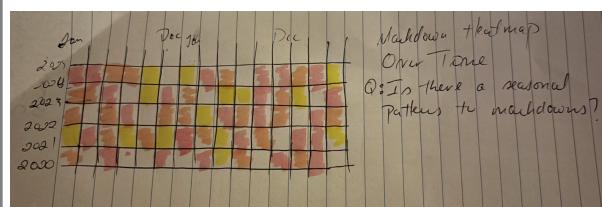
**Visualizations from Milestone 6 are labeled from 1-15.**

## Affinity Diagramming

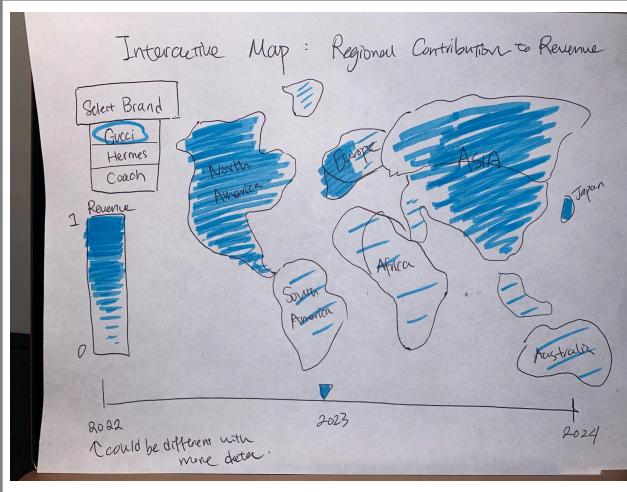
| Sketch  | Sketch ID          | Question ID        | Author      |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
|---|--------------------|--------------------|-------------|-------------|------|-----|------|-----|-------|-----|------|-----|-------|-----|------|-----|------|-----|------|-----|---|---|----|
| <p>Regional Popularity Map<br/>Q: How does search interest vary by region or country</p>  | 1                  | 13                 | DA          |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| <p>Trends in Launch Price Over Time<br/>Q: Do new product launches start at higher price points over time?<br/>+ Launch Price of a Product</p> <table border="1"> <caption>Estimated data points from Trends in Launch Price Over Time chart</caption> <thead> <tr> <th>Year</th> <th>Launch Price (USD)</th> </tr> </thead> <tbody> <tr><td>2016</td><td>350</td></tr> <tr><td>2017</td><td>400</td></tr> <tr><td>2018</td><td>450</td></tr> <tr><td>2019</td><td>500</td></tr> <tr><td>2020</td><td>550</td></tr> <tr><td>2021</td><td>600</td></tr> <tr><td>2022</td><td>650</td></tr> <tr><td>2023</td><td>700</td></tr> <tr><td>2024</td><td>750</td></tr> </tbody> </table> | Year               | Launch Price (USD) | 2016        | 350         | 2017 | 400 | 2018 | 450 | 2019  | 500 | 2020 | 550 | 2021  | 600 | 2022 | 650 | 2023 | 700 | 2024 | 750 | 2 | 4 | DA |
| Year  | Launch Price (USD) |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2016  | 350                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2017  | 400                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2018  | 450                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2019  | 500                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2020  | 550                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2021  | 600                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2022  | 650                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2023  | 700                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| 2024  | 750                |                    |             |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| <p>Price Landscape Across Brands<br/>Q: How do average handbags/shoes/goods prices compare across brands?</p> <table border="1"> <caption>Estimated data points from Price Landscape Across Brands chart</caption> <thead> <tr> <th>Brand</th> <th>Handbags (USD)</th> <th>Shoes (USD)</th> <th>Goods (USD)</th> </tr> </thead> <tbody> <tr><td>Toms</td><td>100</td><td>120</td><td>150</td></tr> <tr><td>Gucci</td><td>300</td><td>350</td><td>380</td></tr> <tr><td>Coach</td><td>250</td><td>280</td><td>300</td></tr> </tbody> </table>  | Brand              | Handbags (USD)     | Shoes (USD) | Goods (USD) | Toms | 100 | 120  | 150 | Gucci | 300 | 350  | 380 | Coach | 250 | 280  | 300 | 3    | 1   | DA   |     |   |   |    |
| Brand   | Handbags (USD)     | Shoes (USD)        | Goods (USD) |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| Toms  | 100                | 120                | 150         |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| Gucci   | 300                | 350                | 380         |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |
| Coach   | 250                | 280                | 300         |             |      |     |      |     |       |     |      |     |       |     |      |     |      |     |      |     |   |   |    |



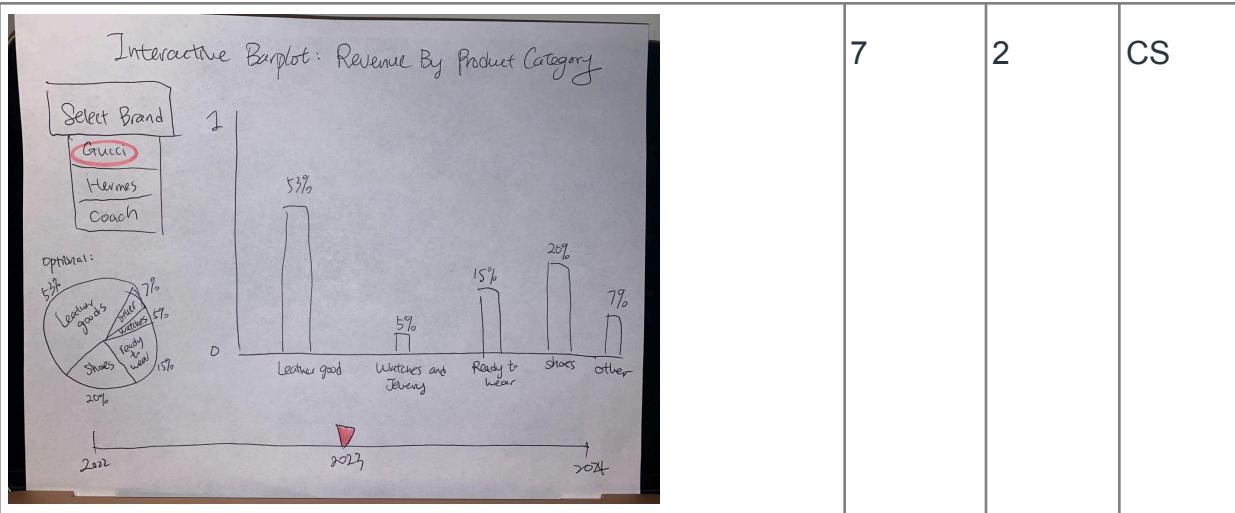
4 2 DA



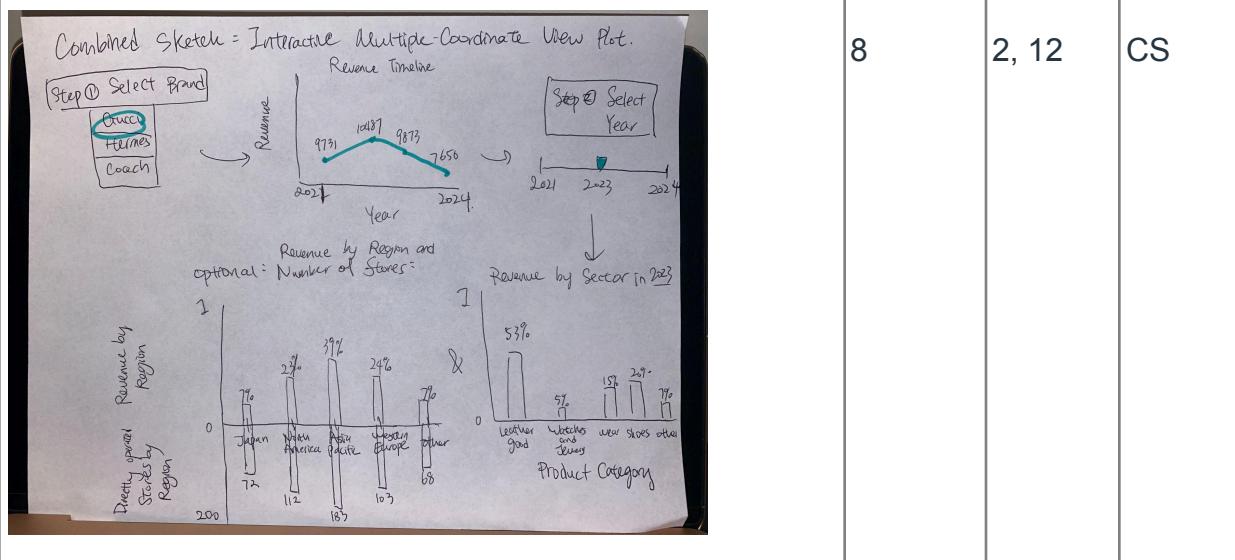
5 5 DA



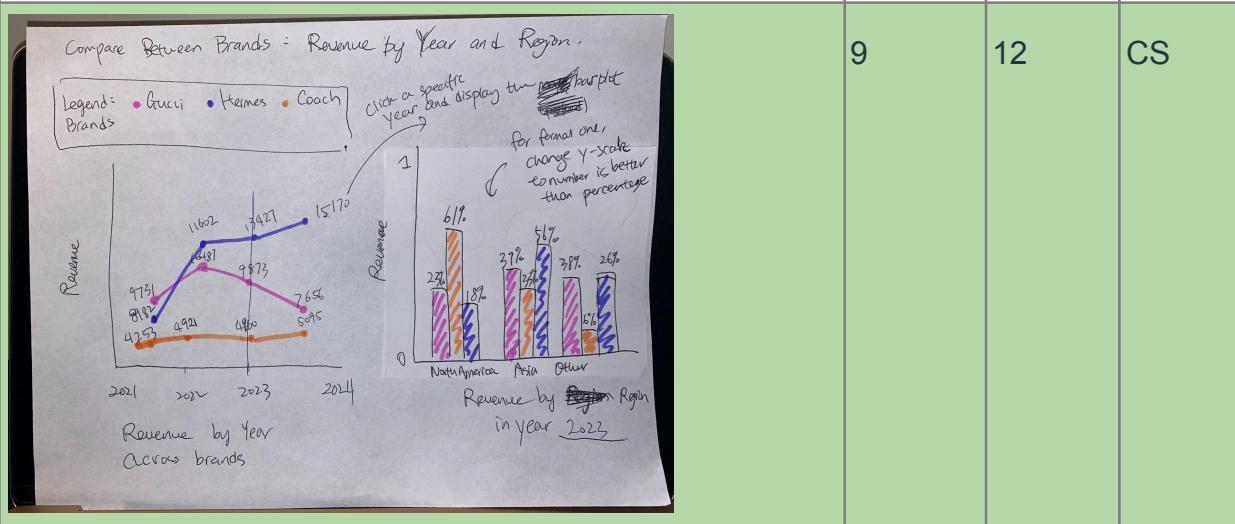
6 12 CS



7 2 CS



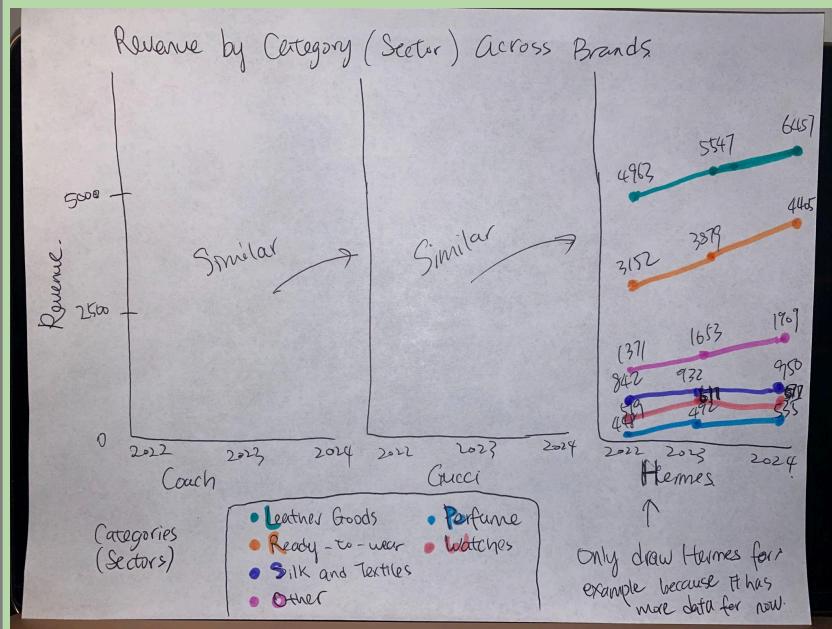
8 2, 12 CS



9 12 CS

Dataset used (or will be using):

- [Hermes Google Sheet \(Region\)](#)
- [Coach Google Sheet \(Region\)](#)
- [Gucci Google Sheet \(Region\)](#)



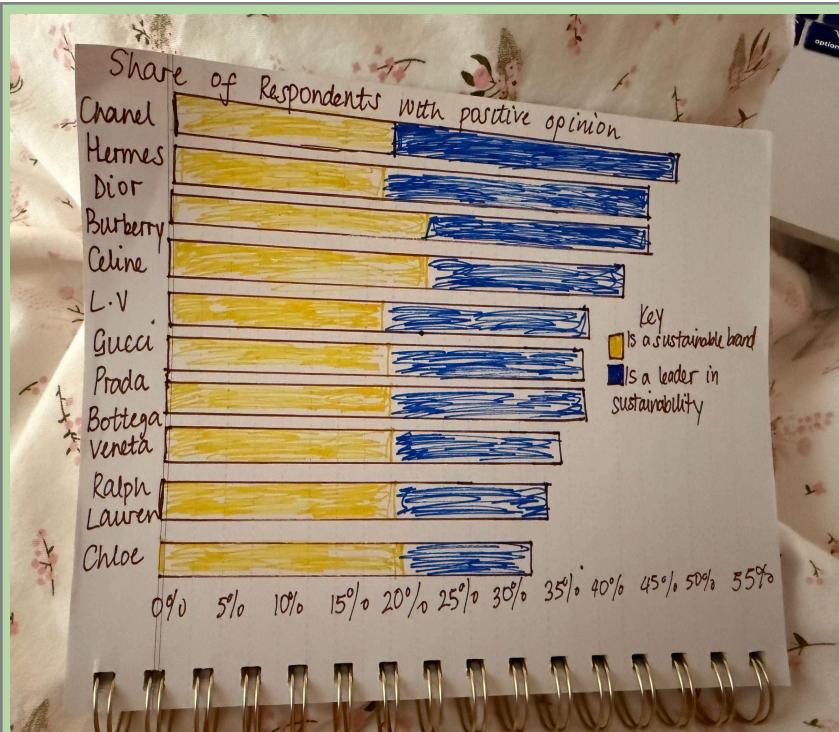
10

2,6

CS

Dataset used (or will be using):

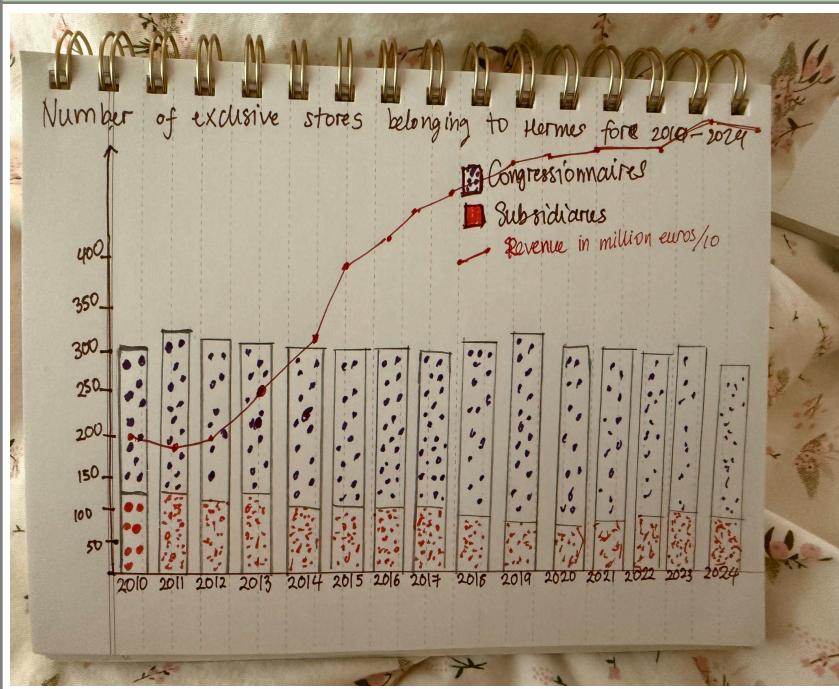
- [Hermes Google Sheet \(Category\)](#)
- [Coach Google Sheet \(Category\)](#)
- [Gucci Google Sheet \(Category\)](#)



11

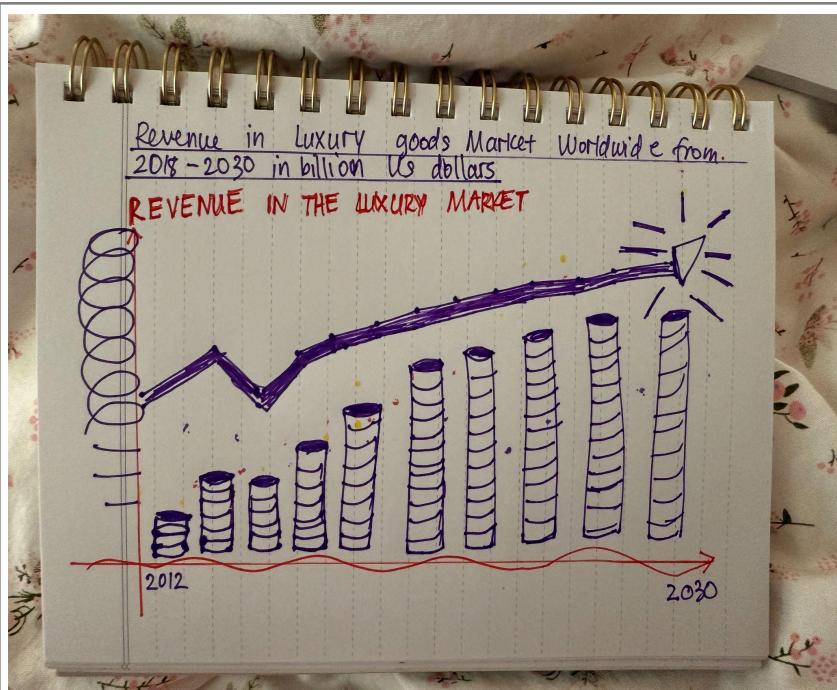
17

PM



12

PM



13

16

PM



14

12

PM



15

16

PM

## Reasoning

We chose sketches 1, 9, 10, 11, and 15 because they best capture the main questions we want to explore and offer a good mix of clarity and creativity. Sketch 1 focuses on regional patterns, helping us visualize how interest and revenue vary by location. Sketches 9 and 10 both highlight brand performance across categories and regions, showing differences in product focus and overall growth across years. We selected Sketch 11 because it adds a unique dimension by visualizing consumer perception and sustainability, which goes beyond just numbers. Finally, Sketch 15 presents overall revenue trends in a clear and cohesive way, helping us connect all the insights together. As a group, we felt these five visualizations were the most effective and engaging to implement in D3, balancing both storytelling and novelty.

Arrange from most relevant (specific) to least relevant (specific):  
Sketches 9 (15), 10, 1, 11



# Storyboard

# Milestone 8: Storyboard

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

## Story Board Link:

[https://miro.com/welcomeonboard/bFRMbIJMOEFJb3Ixd2FFeFVDNHo4TXhiSy9ReTJldXISL0hWYml1NCszdkEvVTJMMWI0cm1WaG1WMmdEdm5nN1iajhpdUUxckxmVVBlleW9ONHBsdHZkK0ZZQTJKSjNaTTQrVnNxVWdJRVFOL1BWVnpBcXpuRzJRL041dIFFUFd3VHhHVHd5UWtSM1BidUtUYmxycDRnPT0hdjE=?share\\_link\\_id=760327674290](https://miro.com/welcomeonboard/bFRMbIJMOEFJb3Ixd2FFeFVDNHo4TXhiSy9ReTJldXISL0hWYml1NCszdkEvVTJMMWI0cm1WaG1WMmdEdm5nN1iajhpdUUxckxmVVBlleW9ONHBsdHZkK0ZZQTJKSjNaTTQrVnNxVWdJRVFOL1BWVnpBcXpuRzJRL041dIFFUFd3VHhHVHd5UWtSM1BidUtUYmxycDRnPT0hdjE=?share_link_id=760327674290)

Frame 1

Luxury is not monolithic. Hermès, Gucci, and Coach represent distinct models of prestige that differ in pricing, consumer engagement, and global reach.

**Hook:**

Did you know that Hermès' annual revenue is \$1 billion last year, while Gucci brought in nearly \$3 billion, and Coach just over \$8 billion? These brands all represent luxury, but they have very different approaches to maintaining their status. Despite one difference in size and revenue, now is "better" than the other. Each brand follows a distinct model of luxury, which can make many brands.

Visualisation 1: Line graph showing the growth/decline of the 3 different luxury brands over time using revenue.

**Rising Insights:**

- Observation 1: Google Trends data shows interest across various regions rather than just the US. Hermès is most popular in Asia, Gucci in Europe, and Coach in North America. This indicates a global reach and appeal.
- Observation 2: Market research shows that luxury consumers are more likely to purchase products from multiple brands. This suggests a desire for variety and a lack of loyalty to a single brand.
- Observation 3: Overlapping price distribution graphs for each brand show that prices are relatively consistent across brands.

**Main Message:**

Visualizations & maps showing how each of the three brands represent different areas of luxury. Using data to support the argument, we see that Gucci is most expensive, most exclusive, and most sought after. Coach is most accessible.

**Solution:**

The diversity among Hermès, Gucci, and Coach underscores how "luxury" has evolved into a fluid concept. Consumers today engage with luxury in ways that were unimaginable just a few years ago. The takeaway: Luxury isn't a price tag or an experience that resonates differently across different cultures. Instead, it's the values that his company are the ones that endure.

## Individual Exploration

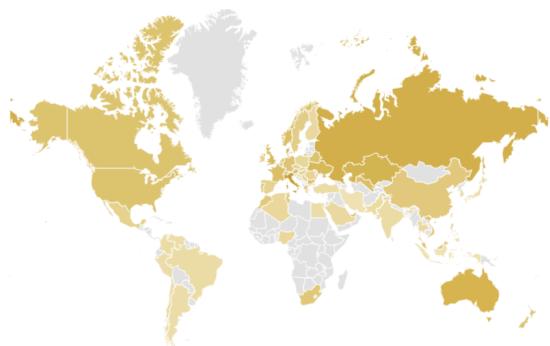
Derin:

1. Countries in which each brand shows the highest search interest by percentage.
  - a. Hermes - Hong Kong; great focus on

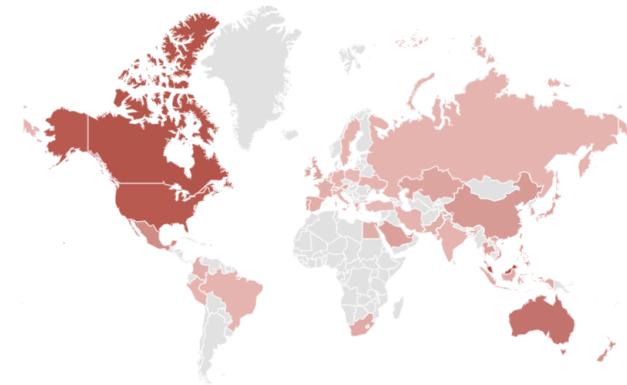


i.

- b. Gucci - Italy; but interest appears to be generally the same world wide

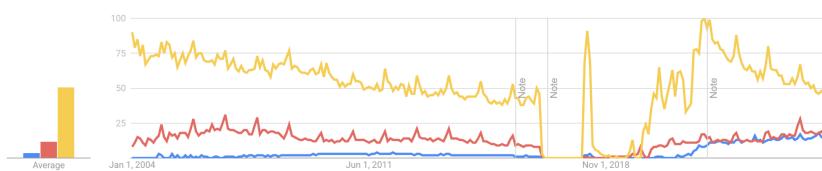


i.  
c. Coach-



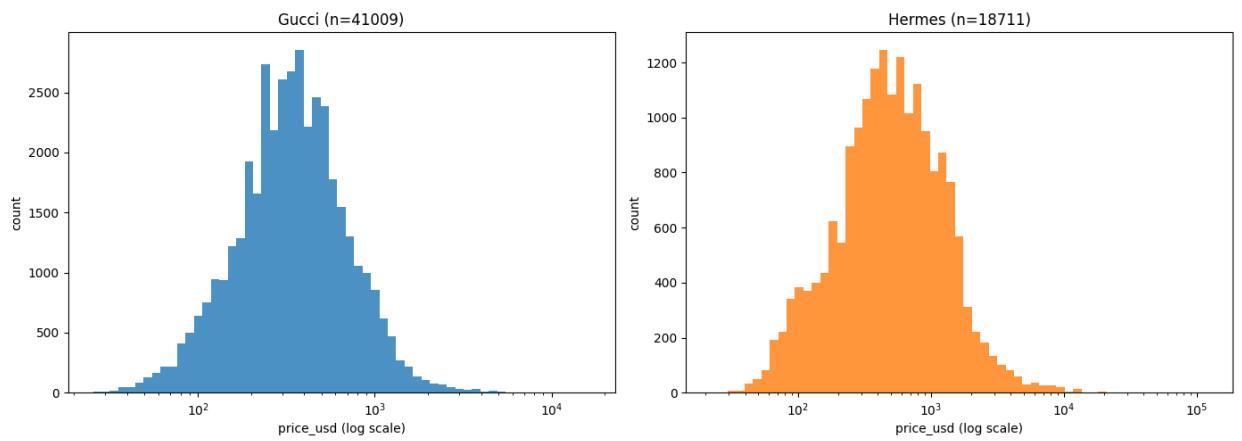
i.

2. Since 2004 - Gucci has maintained the highest search interest of all the brands



a.

3. Gucci has a higher average price across resale and a similar price distribution  
 - Dataset: <https://www.kaggle.com/datasets/justinpakzad/vestiaire-fashion-dataset>

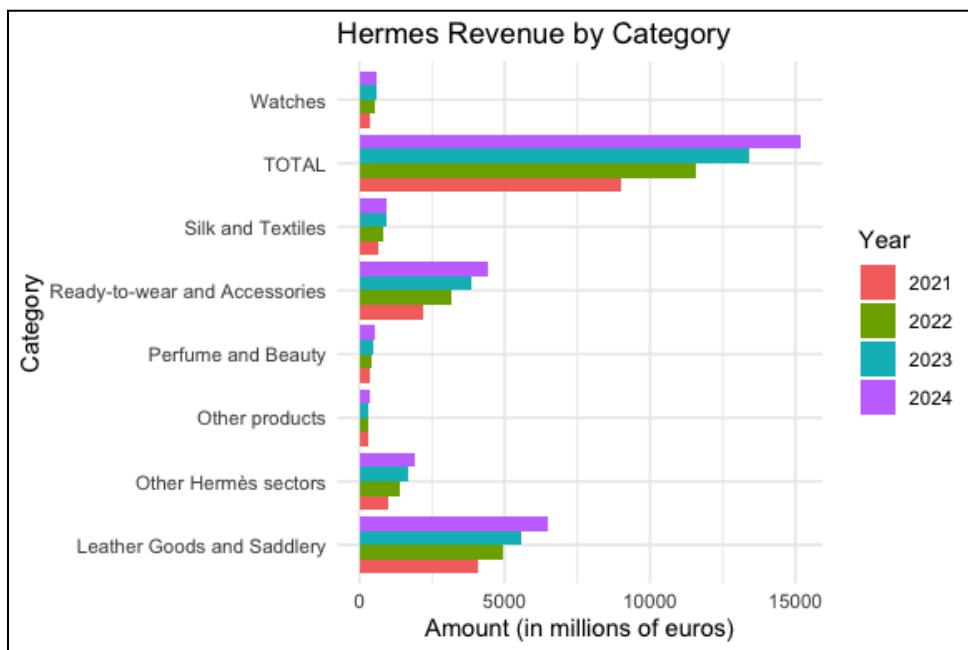


Cindy:

1. Sales by product category

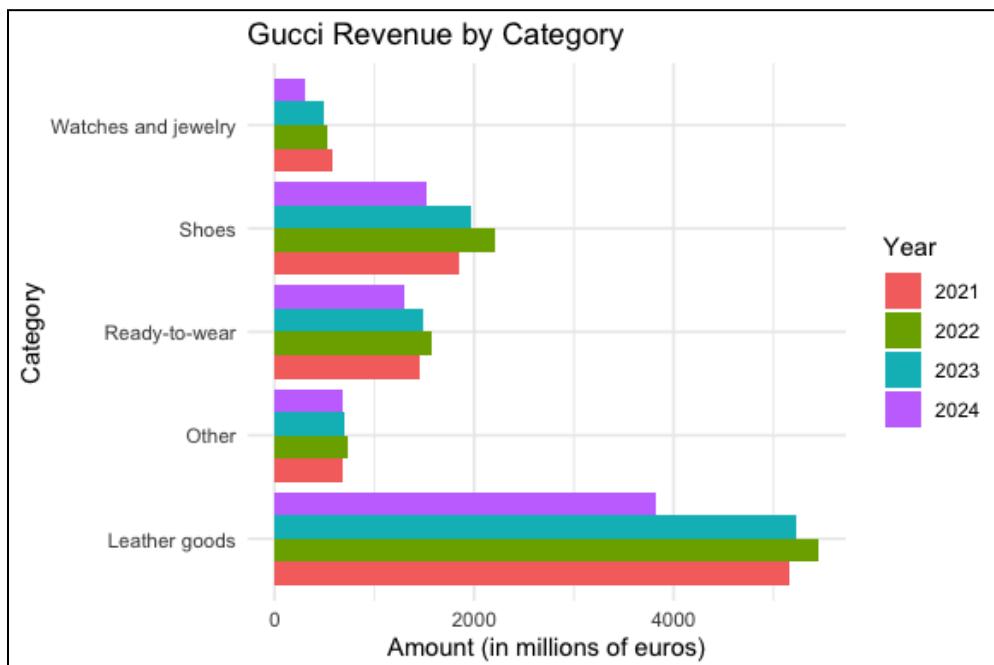
- a. Hermes: Hermes Revenue by Category (Sector) plot below shows that the most revenue came from the “Leather Goods and Saddlery” sector. Revenue increases each year.

i. Dataset: [hermes\\_category](#)



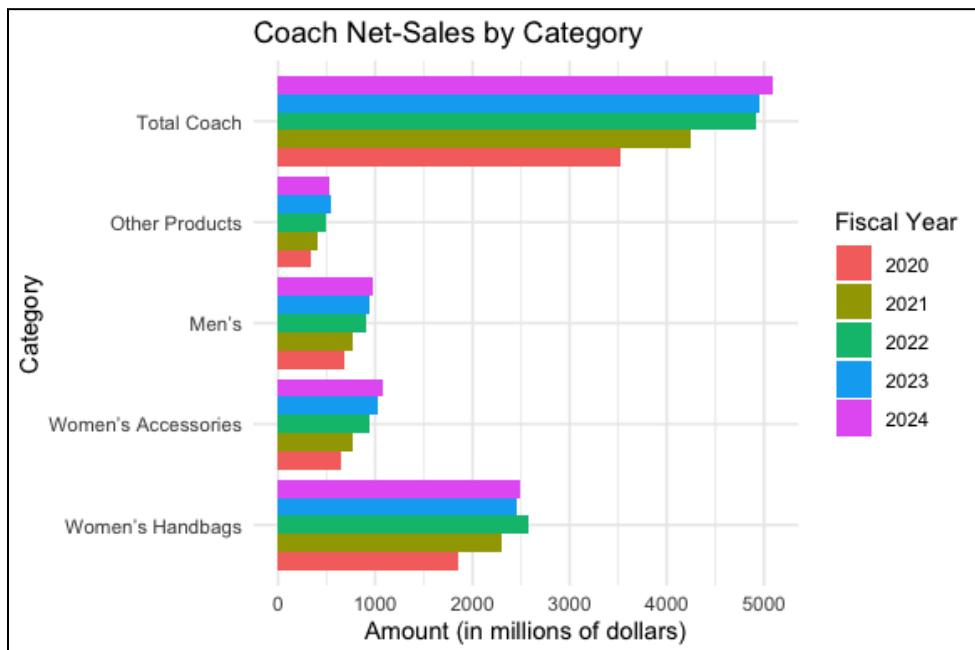
- b. Gucci: The most revenue came from the “Leather Goods” category as well. Revenue decreases by years.

i. Dataset: [gucci\\_category](#)



- c. Coach: Most Net-Sales came from Women's Handbags. Net sales increased by years.

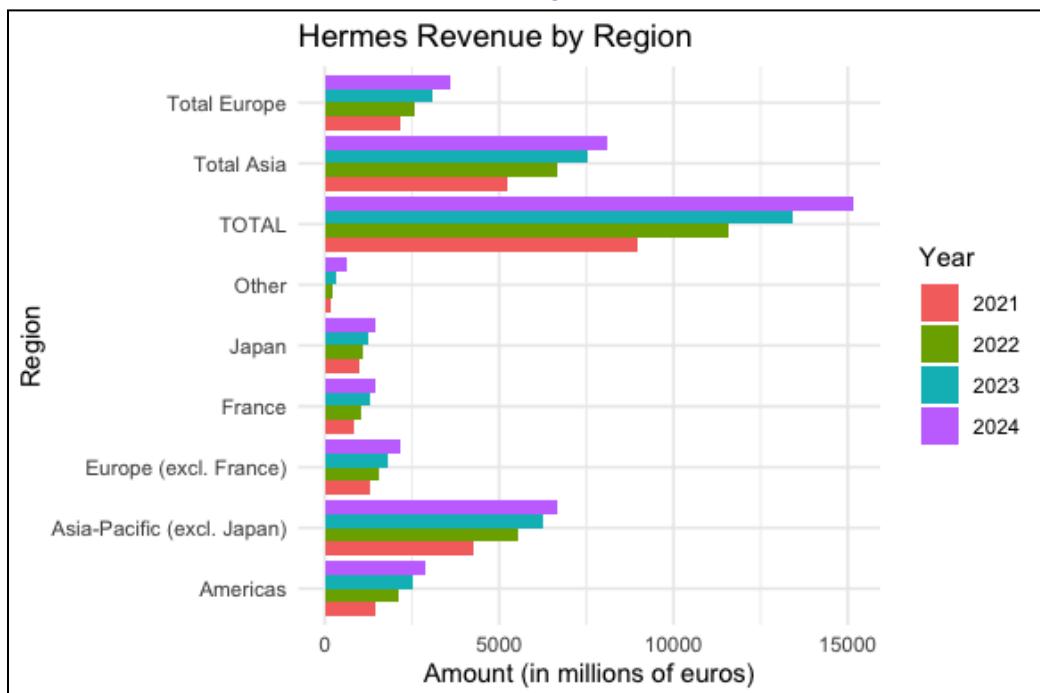
i. Dataset: [coach\\_category](#)



## 2. Sales by region

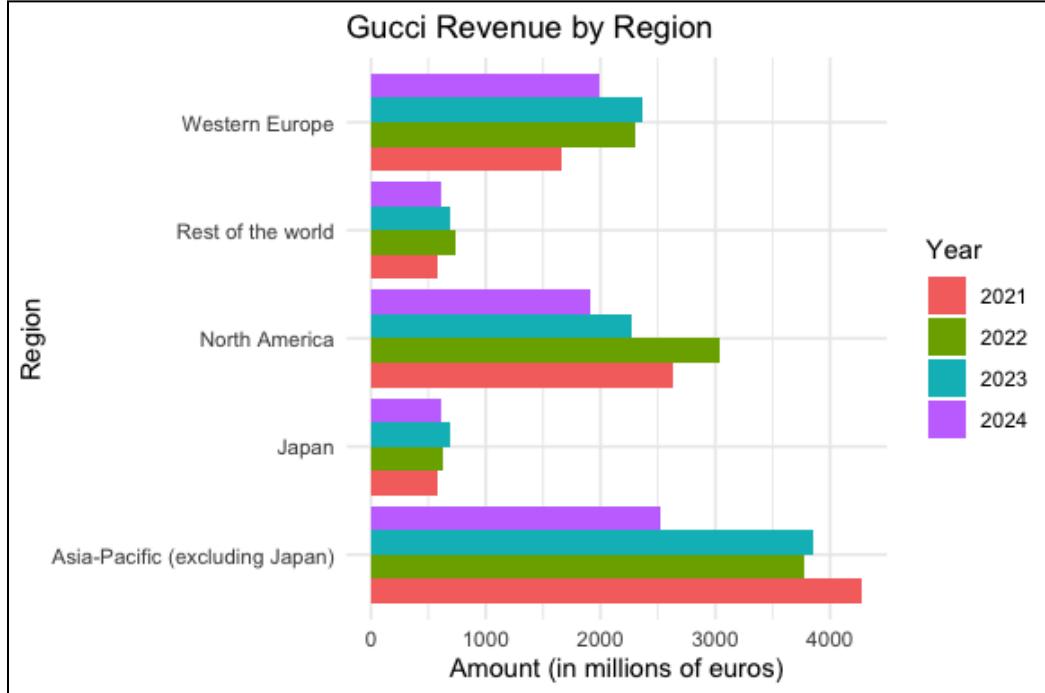
- a. Hermes: Captures a revenue increasing trend similarly as the category plot. Most revenue came from Asia and Europe

i. Dataset: [hermes\\_region](#)



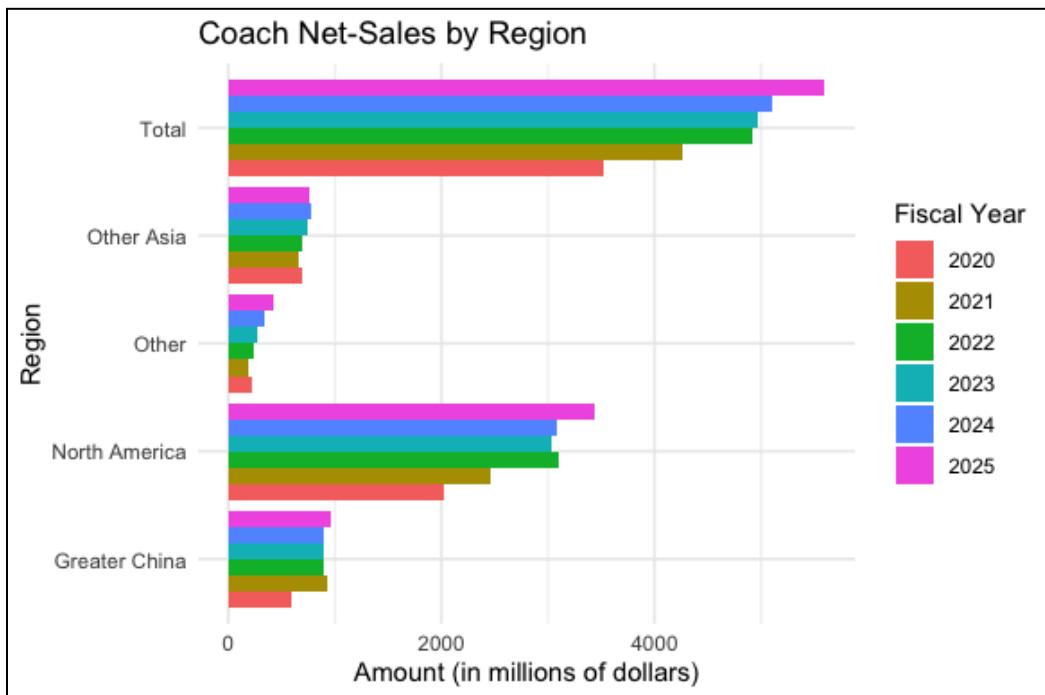
- b. Gucci: Captures a revenue decreasing trend similarly as the category plot. Most revenue came from Asia

i. Dataset: [gucci\\_region](#)



- c. Coach: Captures an increasing trend similarly as the category plot. Most revenue came from North America

i. Dataset: [coach\\_region](#)



Patience:

1. Average Resale Price by Brand and Category (Vestiaire Dataset):  
<https://drive.google.com/file/d/1IIFDGXFITBG8nZ7Cst0sg7DB3Nj4KH0z/view?usp=sharing>

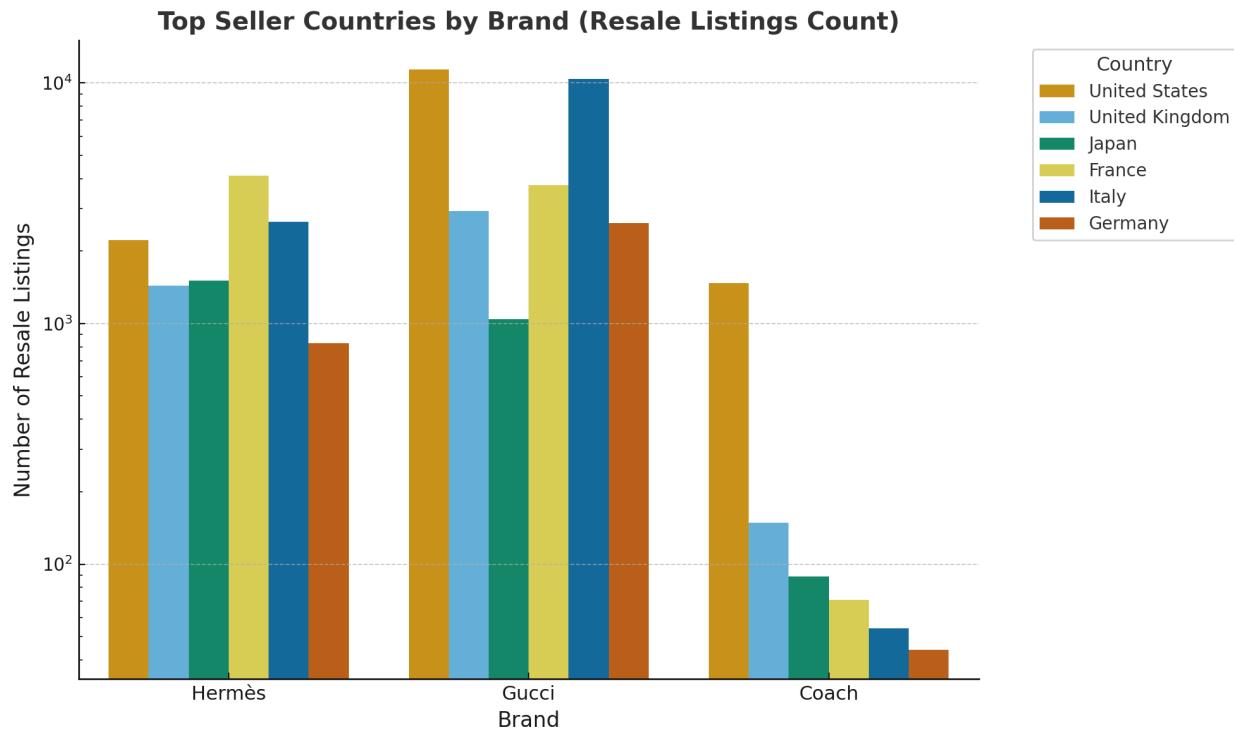


2.

When I looked at the resale data, I saw a clear price hierarchy. Hermès consistently sells for the most, with its accessories and apparel far above Gucci and Coach. Gucci sits in the middle, holding some value but still below Hermès. Coach has the lowest resale prices across every category.

2.

[https://drive.google.com/file/d/12HOykEC1\\_sFPI7WXRhD5M9VoSj17HaN/view?usp=sharing](https://drive.google.com/file/d/12HOykEC1_sFPI7WXRhD5M9VoSj17HaN/view?usp=sharing)  
Feb 2024-oct 2025



When I looked at where most resale listings come from, the pattern was clear. **Coach** has most of its resale activity in the United States, which matches its retail sales data. Almost 60% of Coach's total sales come from North America, so it makes sense that most resale listings are U.S.-based.

**Gucci** and **Hermès** have resale activity spread across the U.S., U.K., France, and Japan. This lines up with their global sales footprint. Both brands rely heavily on Asia and Europe for new sales, and their resale markets show that same international reach.

Comparing resale listings to new sales helps us see how each brand's customer base behaves after purchase. Coach's resale market looks more local, while Gucci and Hermès have resale networks that move across countries. It shows that global brands also create global resale ecosystems.

[https://drive.google.com/file/d/1Y5a\\_2pYMj27j4W8KmLMjAatdZD6rTDf9/view?usp=sharing](https://drive.google.com/file/d/1Y5a_2pYMj27j4W8KmLMjAatdZD6rTDf9/view?usp=sharing)

Top selling coach bags

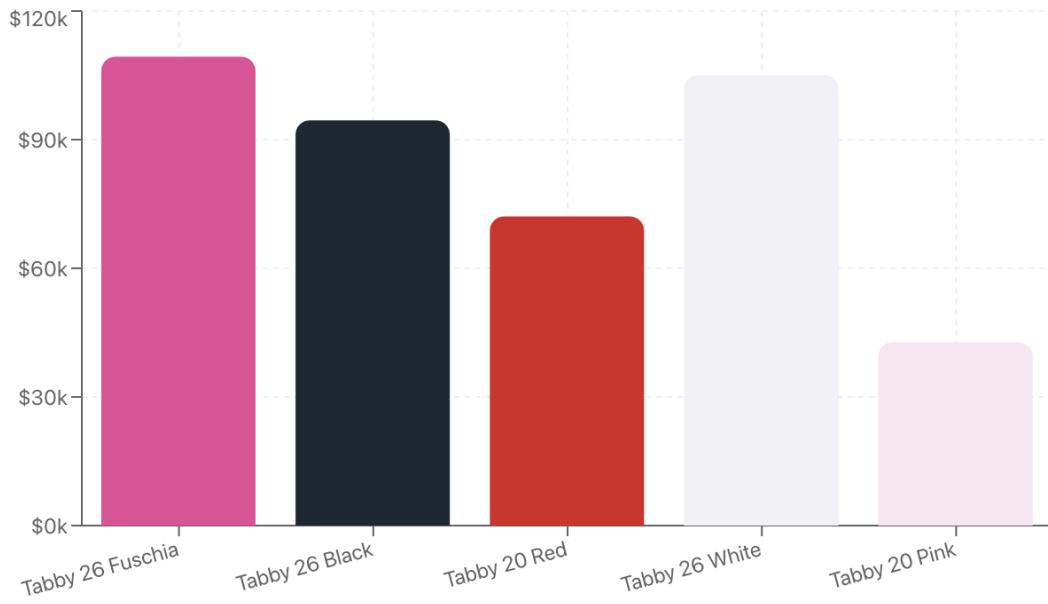
Revenue Analysis

Sales Volume

Growth Trends

Price vs Volume

## Monthly Revenue by Product



Total Revenue

**\$423,737.72**

Claude cont...

Patience (Ideas for storyboard)  
Title: Luxury Brand Popularity

The main message could be that luxury brand popularity is not uniform. Each brand leads in a different dimension of influence.

Luxury brand popularity is not uniform. Gucci leads in visibility and cultural relevance, Hermès in exclusivity and long-term value, and Coach in accessibility and market growth. Each represents a distinct form of influence, showing that luxury success depends on visibility, prestige, or reach rather than a single definition of popularity.

We can tell our story by starting off with an overview of the luxury brand market then show the following insights. I guess today we can just pick one (I am not sure which insight). The idea being that our main message would be like luxury brand popularity is multifaceted. I think any creative message that is different from my suggestion is fine as long as it is not too broad.

## Insights

- **Search Popularity Over Time**
  - Comparing search activity reveals which brands capture the most public attention.
  - This data shows *how visibility defines modern luxury*, with Gucci consistently leading global interest.
  - It introduces the idea that online engagement is a key signal of brand influence today.
- **Financial Performance by Category**
  - Revenue data demonstrates how attention translates into sales success.
  - Hermès' steady growth in leather goods shows that exclusivity and craftsmanship create lasting profitability.
  - Gucci's slight decline in recent years suggests that high visibility does not always guarantee continued financial dominance.
  - Coach's growth in handbags highlights how accessible luxury can drive consistent commercial performance.
- **Financial Performance by Region**
  - Regional data helps explain where each brand's popularity is strongest and why.
  - Hermès and Gucci generate most of their revenue from Europe and Asia, underscoring their reliance on traditional luxury markets.
  - Coach's concentration in North America shows its strength as a more attainable brand appealing to local, everyday luxury consumers.
- **Resale Price Distribution**
  - Resale data connects popularity to long-term consumer perception and value.
  - Hermès achieves the highest resale prices, proving its reputation for rarity and investment value.
  - Gucci maintains moderate resale prices, consistent with its trend-driven appeal.

- Coach ranks lowest in resale value, reflecting its position as an entry-level luxury brand rather than a collectible one.
- **Overall Purpose**
  - These insights collectively show that popularity in the luxury industry is multifaceted.
  - True understanding comes from combining visibility, financial strength, and value retention to reveal each brand's distinct place in the market.

# 1 Prototype I

# Milestone 9: Prototype I

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

**All members actively participated in this milestone.**

**Prototype I Dashboard:** <http://127.0.0.1:5500/dashboard.html> (if the link did not open, please open the html file named dashboard.html in VSCode using Live Server)

## Data Pipeline

We completed scraping and cleaning for five real datasets used in our dashboard:

- region\_category\_clean\_data.csv – Brand-category-year revenue data; converted all currencies to euros, fixed typos, and added 2020 values.
- geoMap.csv – Google Trends search interest by country; cleaned headers, parsed %, standardized region names.
- luxury\_revenue.csv – Brand-level financial summaries for consistency checks.
- vestiaire\_cleaneddataset.csv – Cleaned Vestiaire Collective listings with retail and resale prices. (not yet used in this milestone)
- resale\_prices.csv – Computed average markdowns and resale ratios by brand and category.

These datasets feed into our D3 visualizations: revenue facets, global map, price distributions, and radar charts.

## Functional (incomplete) D3 Visualizations

### Luxury Brand Revenues Over Time

- Dataset: luxury\_revenue.csv
- Description: Multi-line chart comparing Hermès, Gucci, and Coach revenues from 2020–2024.

### 2024 Revenue Comparison (Bar Chart)

- Dataset: luxury\_revenue.csv
- Description: Bar chart comparing 2024 total revenues for each brand.

### Revenue by Category Over Time (Right Panel)

- Dataset: region\_category\_clean\_data.csv
- Description: Line chart showing category-level revenue trends for a selected brand (e.g., Gucci).

### Faceted Category Trends (Interactive Line + Bar Charts)

- Dataset: region\_category\_clean\_data.csv
- Description: Three brand panels (Hermès, Gucci, Coach).
- Feature: Click year → updates bar chart dynamically.

### Choropleth Map — Global Search Interest

- Dataset: geoMap (2).csv + World GeoJSON
- Description: Interactive world map showing Google Trends interest per brand.

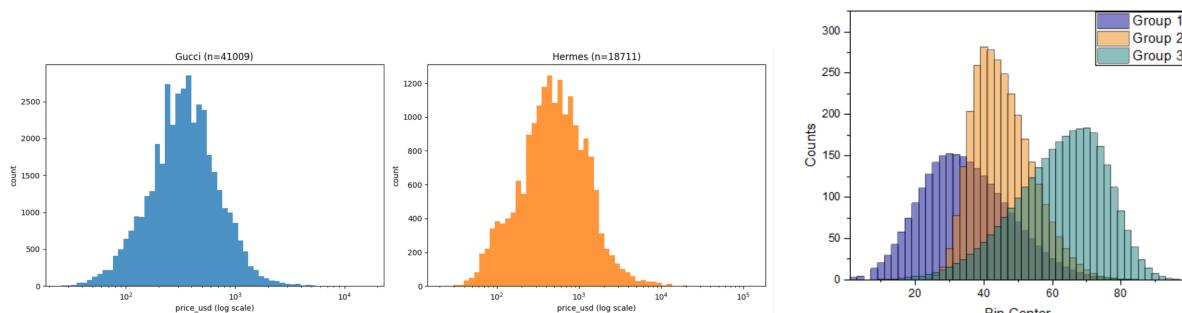
### Resale Prices by Brand & Category (Stacked Bar Chart)

- Dataset: resale\_prices.csv
- Description: Final “Additional Visualization” on the right.

## Visualizations To Do

### Overlapping Price Distribution

- Idea: Compare retail (original) and resale (Vestiaire) prices through dual kernel density plots to illustrate luxury value retention and pricing gaps.



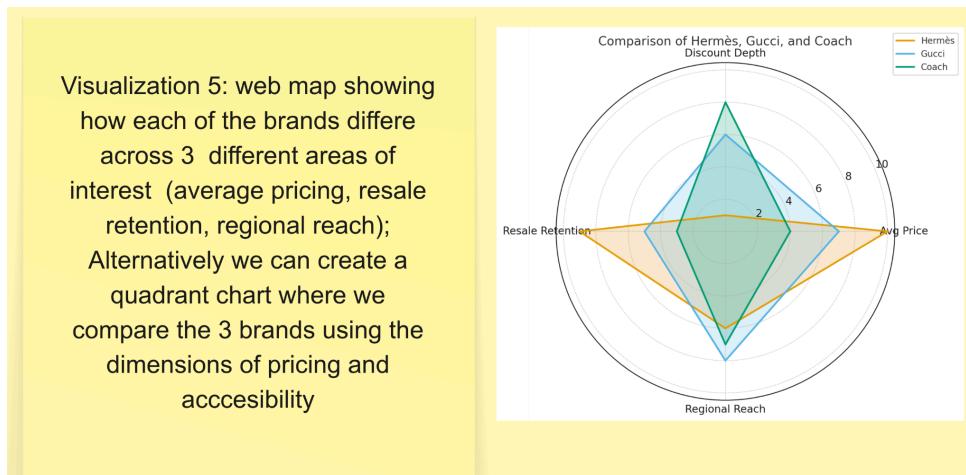
- Currently looks like:



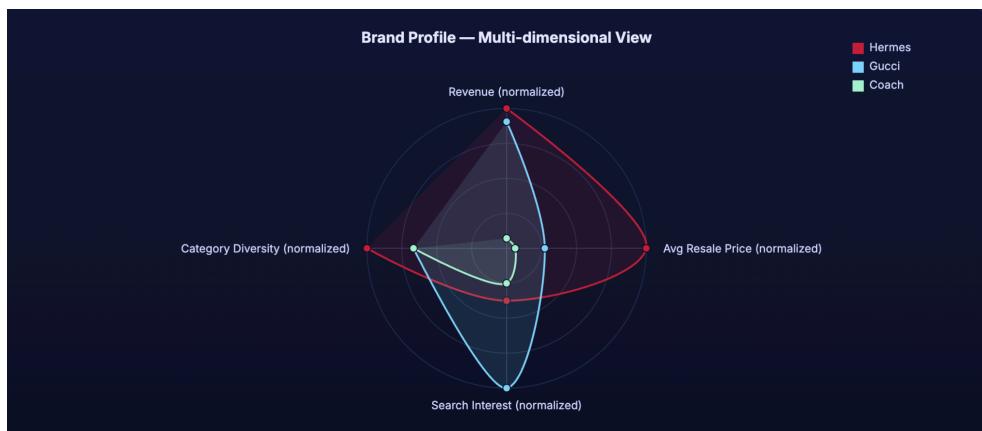
Our original idea was to create a price distribution visualization with two views. One with the original prices of items and the second the same item as its listed on vestiaire. We've been able to create one for the price before vestiaire but the one for the seller price is still in development.

### Radar Chart

- Idea: Show a multi-dimensional brand profile using normalized metrics (revenue, resale value, category diversity, and search interest).



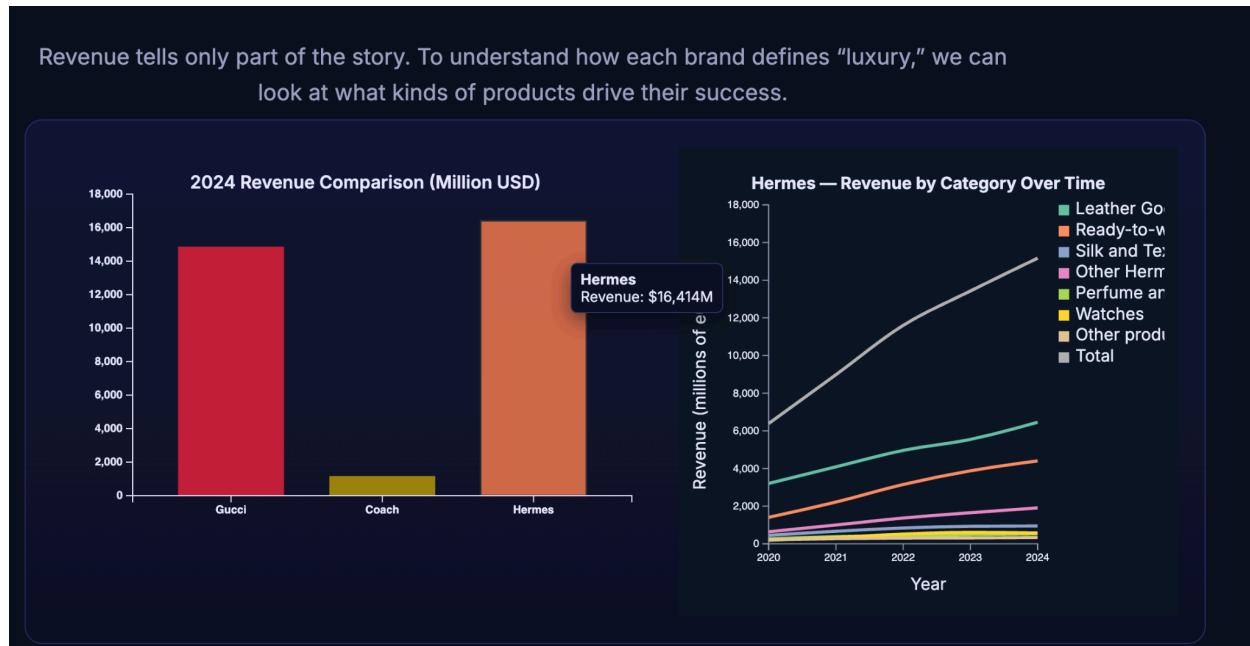
- Currently looks like:



We've succeeded in creating a radar chart that visualized the different dimensions that each of the three brands can be measured on. Right now we are just exploring if there are any other characteristics we would prefer to highlight from a story telling point of view.

## 2024 Revenue Comparison (Bar Chart)

- Currently looks like:



This visualization is partially overlapped with the next visualization: Faceted Category Trends (Interactive Line + Bar Charts). We might adjust it by improving the spacing and alignment between the two sections, ensuring that the bar chart serves as a clear entry point to the category-level trends. We might also link user interaction — hovering or clicking on a brand's bar — to automatically highlight that brand's category trends in the visualization below.

## *Timeline of each company's history*

- Currently not implemented, but will work on it for the next milestone

The timeline will serve to provide additional context about each company's background and evolution, helping viewers understand how their histories shape their current brand identities. By highlighting distinctions such as heritage-driven luxury versus trend-based modernity, the timeline aims to deepen audience insight into why these companies differ in market positioning and brand perception.

## 2 Prototype II

# Milestone 10: Prototype II

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

Since we shifted our HTML layout into a slide HTML format, we created a Google Slides mockup to illustrate our intended output structure. Our final project will follow a framework like this:  CS1710 Mockup 1

<https://docs.google.com/presentation/d/15M5tR2-QPnUffXyJcaApr3XXKMrwdKGZm7SFvRRUi0/edit?usp=sharing>

Repository link: <https://github.com/DerinAdeleke/CS171-Final-Project>

**All members actively participated in this milestone.**

**Prototype II Dashboard:** Please open the html file named presentation.html in VSCode using Live Server

## 3 Prototype III

# Milestone 11: Prototype III

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

Link to the full process book:

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

Since we shifted our HTML layout into a slide HTML format, we created a Google Slides mockup to illustrate our intended output structure. Our final project will follow a framework like this:  CS1710 Mockup 2

[https://docs.google.com/presentation/d/1mluQF0HjiJ6HuJCT1gICl3Bl0hEdzkckqAl1odJEPfg/edit?slide=id.g3a4631d2131\\_0\\_2#slide=id.g3a4631d2131\\_0\\_2](https://docs.google.com/presentation/d/1mluQF0HjiJ6HuJCT1gICl3Bl0hEdzkckqAl1odJEPfg/edit?slide=id.g3a4631d2131_0_2#slide=id.g3a4631d2131_0_2)

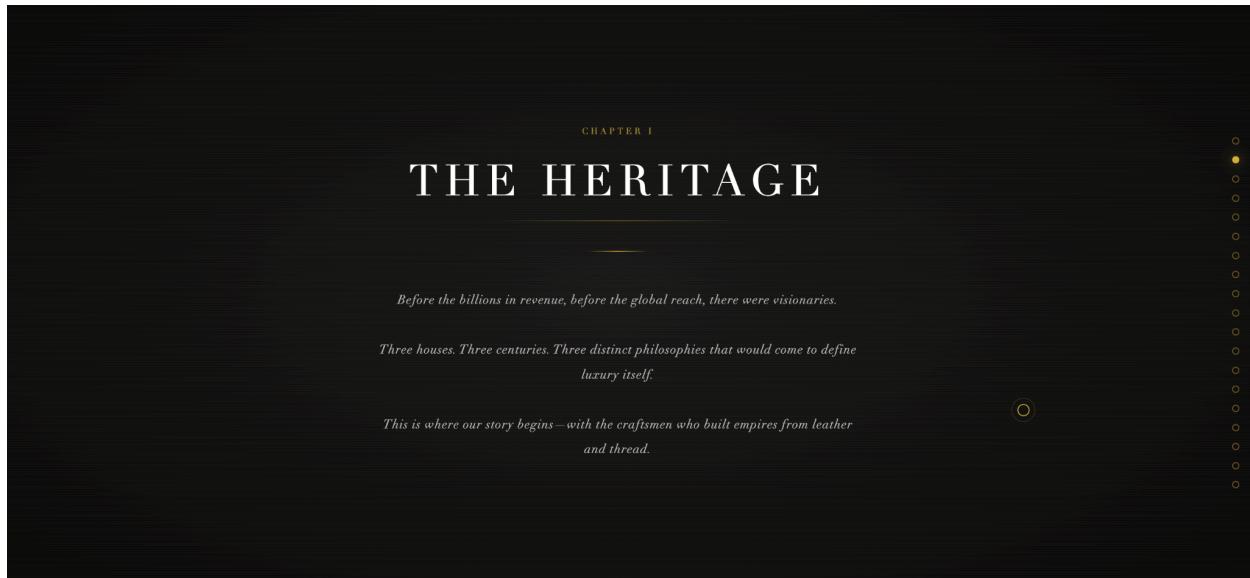
Repository link: <https://github.com/DerinAdeleke/CS171-Final-Project>

**All members actively participated in this milestone.**

**Prototype III Dashboard:** Please open the html file named index.html in VSCode using Live Server.

**Scroll down to see screenshots of the current dashboard:**

**Screenshots of current dashboard:**



# HERITAGE & EVOLUTION

*A journey through centuries of luxury craftsmanship*

ALL BRANDS    HERMÈS (1837)    GUCCI (1921)    COACH (1941)



## THE BIG THREE

*Hermès, Gucci, and Coach define distinct segments of luxury*

### HERMÈS

\$16.4B

2024 REVENUE

*Ultra-luxury positioning with exceptional craftsmanship and timeless appeal*

### GUCCI

\$14.9B

2024 REVENUE

*Bold contemporary luxury with mass appeal and cultural relevance*

### COACH

\$1.2B

2024 REVENUE

*Accessible luxury with American heritage and practical elegance*

## WHAT HISTORY REVEALS



### 188 Years

HERMÈS JOURNEY

From harness maker to haute couture, the oldest of our three houses proves that longevity breeds luxury. Nearly two centuries of uncompromising craftsmanship.



### Post-War Boom

THREE DISTINCT ERAS

Hermès emerged from aristocratic Europe (1837), Gucci from post-WWII Italy (1921), and Coach from American pragmatism (1941). Three continents, three philosophies.



### Iconic Moments

PRODUCT INNOVATION

The Birkin (1984), Jackie Bag (1961), and Coach's bonded leather (1941)—each breakthrough redefined what luxury could be. Innovation, not just tradition, drives legacy.

## CHAPTER II

# THE MARKET TODAY

*From heritage to hypergrowth.*

*Today, luxury is no longer confined to Parisian ateliers or Manhattan boutiques. It's a \$42+ billion global phenomenon, driven by digital transformation and emerging markets.*

*The data reveals an industry in unprecedented expansion—and unexpected evolution.*

# MARKET OVERVIEW

The luxury fashion market spans decades of heritage and billions in revenue

\$42.6B

TOTAL REVENUE 2024

42K+

RESALE ITEMS

13 Years

HISTORICAL DATA



# GLOBAL TRAFFIC HUB

Top 10 countries driving online luxury market engagement in 2024

FASTEST GROWING  
Turkey +38.8%

MARKET LEADER  
USA 19.8%

COMBINED TRAFFIC  
72.8%

Japan  
Traffic Share: 7.5%  
YoY Growth: +24.2%  
Total Visits: 300M  
Region: Asia

## THE NEW GEOGRAPHY OF DESIRE



CHAPTER III

## THE PRODUCTS

*Icons are not born – they are meticulously crafted.*

*From the Birkin to the Marmont, certain pieces transcend mere accessories to become cultural symbols. These are the products that define desire, drive revenue, and create lifelong brand devotion.*

*Let's explore what the world is buying.*

## BEST SELLERS

A curated selection of iconic pieces from each house

ALL PRODUCTS

COACH

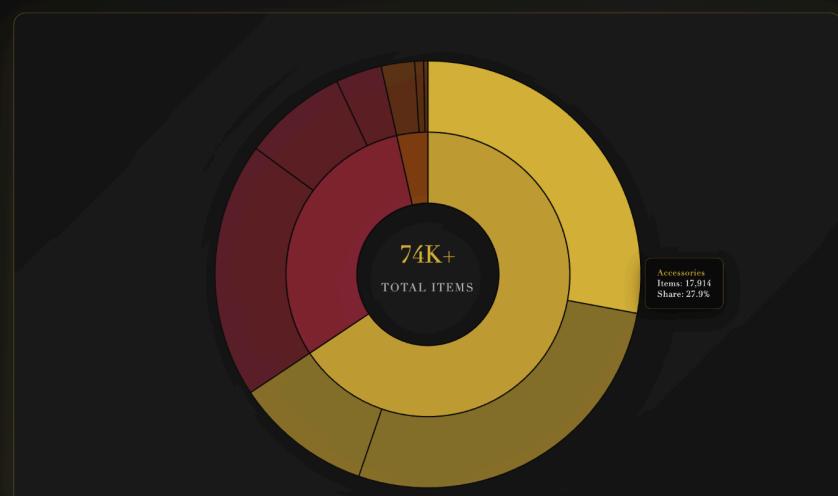
GUCCI

HERMÈS



## MARKET DISTRIBUTION

Category breakdown reveals strategic positioning



## WHAT 74,000 ITEMS TELL US



CHAPTER IV

## THE RESALE REVOLUTION

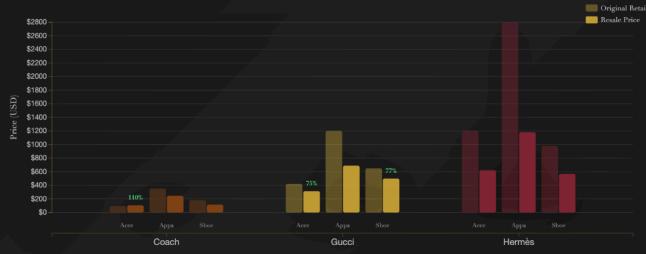
*A new economy is emerging—one where pre-loved is prestigious.*

*The secondary market isn't just about sustainability; it's a powerful indicator of brand equity. When a Birkin sells for more than retail decades later, it reveals something profound about value, desire, and legacy.*

*Here's what the resale market tells us about these brands.*

## RESALE MARKET

*Secondary market reveals brand value retention and desirability*

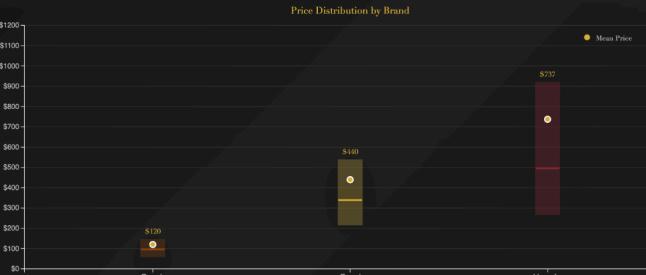


*Hermès items retain exceptional value, often exceeding 70% of original retail price*

[ALL CATEGORIES](#) [ACCESSORIES](#) [APPAREL](#) [SHOES](#)

## THE VESTIAIRE MARKETPLACE

*Real-time pricing data from 62,000+ luxury resale items*



[PRICE DISTRIBUTION](#) [BY CONDITION](#) [MARKET VOLUME](#)

# THE VALUE OF LEGACY



## ACKNOWLEDGEMENTS

*With gratitude to those who made this project possible*

### CS1710 Course Staff

Special thanks to the teaching assistants and staff of CS1710 for their guidance, support, and invaluable feedback throughout this project.

### Data Sources

This visualization was made possible through publicly available data from:

Vestiaire Collective - Luxury Resale Market Data  
Brand Annual Reports & Financial Disclosures  
Google Trends - Search Interest Data  
Historical Brand Archives & Documentation

### Tools & Technologies

Built with D3.js, HTML5, CSS3, and a passion for data storytelling

# OUR TEAM

*Creators of this luxury data experience*

**Derin Adeleke**

[derinadeleke@college.harvard.edu](mailto:derinadeleke@college.harvard.edu)

**Cindy Sun**

[csun@g.harvard.edu](mailto:csun@g.harvard.edu)

**Patience Madumera**

[pmadumera@college.harvard.edu](mailto:pmadumera@college.harvard.edu)

HARVARD UNIVERSITY • CS1710 • FALL 2025



## Think Aloud Results

## Think Aloud Study 1

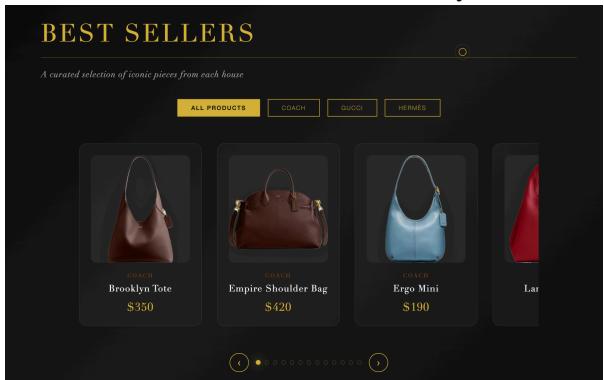
For each think-aloud session, create a new copy of the feedback table below in your process book. During the session, one team member should carefully listen to the tester and document their observations and feedback in the "Notes" column. Ensure that all sections of the table are completed for each tester. If your group has 3 or more members, you must conduct at least two independent think-aloud sessions and document feedback for both.

|   | Notes (To be filled by project leads)   |
|---|---|
| Tester Name   | Audrey Yang<br>atyang@college.harvard.edu   |
| Describe any usability issues or confusion the tester encountered while using the prototype.        | <ul style="list-style-type: none"> <li>Scrolling too fast. Was not able to interact with slide 3 fully (timeline)</li> </ul>  |
| Was the tester able to understand the main message of the data story? (e.g., Yes/No + why/why not?) | <ul style="list-style-type: none"> <li>Not fully but they got the idea. Asked why Coach is being grouped with the other 2 brands that were not doing the same financially. May not have gotten the idea that we were trying to define luxury (personal opinion)</li> </ul>                    |
| What parts of the interface or visualization did the tester find most engaging or effective?        | <ul style="list-style-type: none"> <li>Loved the fonts</li> </ul>   |
| What parts did the tester find confusing or less effective?   | <ul style="list-style-type: none"> <li>Wanted to know why we chose the specific three brands maybe put that first</li> <li>The global traffic hub is hard to interpret immediately</li> </ul> <ul style="list-style-type: none"> <li>Market overview numbers at the top questioned</li> </ul> |
| Did the tester encounter any inconsistencies in design, data, or narrative?                         | <ul style="list-style-type: none"> <li>Found the data on the what 74000 items tell us inconsistent with the next slide</li> </ul>   |

|   |  |
|---|--|
|   |  |
| Were there any unexpected interactions or insights that emerged during the session? | <ul style="list-style-type: none"> <li>Global traffic hub slide was hard to interact with</li> </ul>   |
| What specific improvements or changes did the tester suggest for the prototype?     | <ul style="list-style-type: none"> <li>Alignment of the pie chart needs to be adjusted</li> <li>Adjust the font for bar chart : resale market</li> <li>Best sellers slide could make use of a feature that scrolls back and forth</li> <li>Dots on the best sellers slide are confusing and the tester off a bit (suggesting more spacing on the chart on that slide)</li> </ul> |
| Did the tester suggest any additional insights or visualizations to include?        | <ul style="list-style-type: none"> <li>Thought we needed more visualisations for slide 4 since it's a visualisation class. Too much text with no visual insights.</li> <li>Add more images in the history. The tester found it too text heavy. Would love to see more of the iconic bags</li> </ul>  |
| General observations or comments from the tester.                                   | <ul style="list-style-type: none"> <li>Font, liked storytelling. Thought colors did not have enough contrast</li> <li>Found slide with new geographies helpful</li> <li>Cool images!</li> </ul>  |

## Think Aloud Study 2

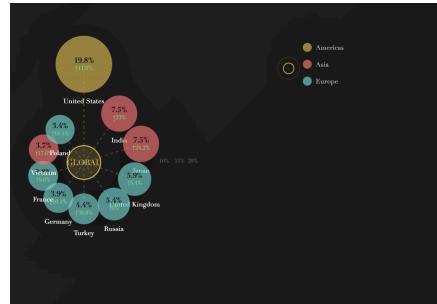
For each think-aloud session, create a new copy of the feedback table below in your process book. During the session, one team member should carefully listen to the tester and document their observations and feedback in the "Notes" column. Ensure that all sections of the table are completed for each tester. If your group has 3 or more members, you must conduct at least two independent think-aloud sessions and document feedback for both.

|   | Notes (To be filled by project leads)   |
|---|---|
| Tester Name   | Kirthi Chigurupati<br>kchigurupati@college.harvard.edu  |
| Describe any usability issues or confusion the tester encountered while using the prototype.        | <ul style="list-style-type: none"> <li>- Scrolling feature design (scroll but not click though) - prefer click than scroll</li> </ul>   |
| Was the tester able to understand the main message of the data story? (e.g., Yes/No + why/why not?) | <ul style="list-style-type: none"> <li>- How “top 3” brands evolves overtime (not necessarily correct) - need definition/explanation for choosing 3 brands</li> </ul>   |
| What parts of the interface or visualization did the tester find most engaging or effective?        | <ul style="list-style-type: none"> <li>- Love the overall dark theme/style with the font</li> </ul>   |
| What parts did the tester find confusing or less effective?   | <ul style="list-style-type: none"> <li>- The title “the big three” are misleading (consider add description for brands or change the title “the big three”)</li> <li>- Market overview boxes - thought might be clickable but in fact not</li> <li>- Not sure about the takeaway from best sellers:</li> </ul>  |
| Did the tester encounter any inconsistencies  | <ul style="list-style-type: none"> <li>- No clear distinction between sections (the</li> </ul>  |

in design, data, or narrative?

tester suggests different background to diff slides to indicate change of info)

- Change the bubble chart into geo map:



- Fun facts but could be linked to bubble chart/map:



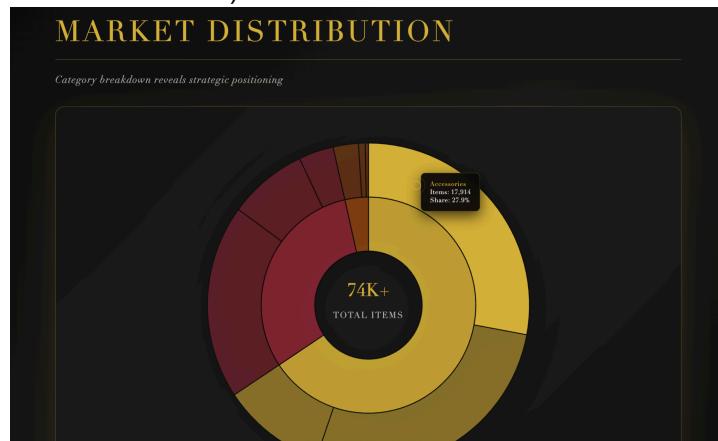
- The “what history reveals” page might tight to the timeline (suggesting combining two slides)
- “Market Distribution” Pie chart misleading because looks like revenue/market share instead of number of sold items

Were there any unexpected interactions or insights that emerged during the session?

- Timeline instructions needed (like scroll/ click on the points) because the tester did not scroll to see the whole timeline.

What specific improvements or changes did the tester suggest for the prototype?

- For the pie chart: Maybe Interact with the previous slide (any change in the past years? ie. Does the number of items decrease / increase?):



- Duplicated functionality (the clicking buttons on top of the timeline duplicates with the color dots below the timeline)

## HERITAGE & EVOLUTION



- Add pictures for timeline popups
- For “resale market” slide: Do not use short abbreviations for “acce”, “appa”, “shoe” (instead use “accessories”, “apparel”, and “shoes”):

## RESALE MARKET



Did the tester suggest any additional insights or visualizations to include?

- The tester suggests combining some info slides with the visualization slides together (map and timeline)
- Might be interesting to explain why the fashion market matters (ex. How much US gdp is fashion)

General observations or comments from the tester.

- Great visualizations
- Unclear story
- Unclear interaction features

Conclusion:

**Based on the results of your ‘think aloud’ study, what would you improve in your data story?**

Testers expressed confusion about the rationale behind selecting the three brands and did not always grasp the intended takeaway, which suggests that our narrative needs clearer motivation and context early on. Several interactive components (especially scrolling and the timeline) were difficult to navigate and lacked instructions. We also received feedback about visual density, layout alignment, and insufficient contrast.

**Are there any additional insights and visualizations you would use? Would you amplify or change your message? Did your narrative work? Did the tester get your takeaways?**

The narrative resonated only partially with testers, indicating that while the theme was compelling, the main message was not consistently communicated. We learned that certain sections relied too heavily on text. Testers suggest we add more images of iconic products, and clearer geographic mapping. We will improve our narrative by better explaining how we made the choice of brands, highlighting how their histories and consumer trends reveal what “luxury” means today, and ensuring each slide contributes directly to this takeaway.

**Decide as a team which of these improvements you will implement and write down your decisions and why you made them in your process book as a numbered list.**

We decided to prioritize improvements that increase clarity of narrative and ease of interaction.

1. Introduce an early slide explaining why Hermès, Gucci, and Coach were chosen (done)
2. reduce text-heavy sections( add pictures to the timeline/history) (done)
3. Improve interaction by adding instructions, reducing overscrolling, and eliminating duplicated filtering features on the timeline
4. Adjust chart spacing, alignments, and labels where testers noted confusion (e.g., pie chart interpretation, best sellers dot chart, market overview design)
5. Apply stronger visual contrast and clearer section transitions (done)
6. Fix the cursor (done)
7. Modularizing our code by splitting files into separate html, css, and javascript files.
8. Switching the global traffic diagram into an interactive map of global traffic or search interest.
9. Experiment with contrast between charts and background to see how it turns out
10. Rename misleading titles and make story clear (done)

There will be more changes in the future..



# Final Project

# Milestone 13: Final Project

Aderinsola (Derin) Adeleke, Patience Madumera, Cindy Sun

## **Link to the full process book:**

[https://docs.google.com/document/d/1I2M\\_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing](https://docs.google.com/document/d/1I2M_CEjiH24KI8ZEnrl-yQEhh3kB92lwM-DrJ0pMR4s/edit?usp=sharing)

## **Video Link:**

<https://youtu.be/OUaekgi8gls>

## **Project Website:**

<https://derinadeleke.github.io/CS171-Final-Project/>

## Final Edits

Here's a comprehensive list of the final edits made to our data story.

### **Dashboard & Interactivity Enhancements**

#### 1. Vestiaire Marketplace Dashboard Transformation

- Converted from button-switched views to a 4-panel linked dashboard (2x2 grid)
- Implemented cross-filtering: interactions in one panel update all others
- Added expand/fullscreen feature for each panel ([] icon buttons)
- Modal system with 900×600px enlarged views and backdrop blur

#### 2. Scale Fixes for Data Visualization

- Fixed price distribution box plots using 95th percentile scaling instead of max
- Fixed price range histogram using same quantile approach
- Prevents outlier compression (items over \$10k were squashing the visualization)

#### 3. Timeline Slider UX Improvement (Market Overview)

- Added click-to-position functionality (was drag-only before)
- Increased handle sizes from 7px to 8px radius
- Added visual feedback on drag (handles grow to 10px, color brightens)
- Much better for trackpad users

#### 4. Global Dominance Map Refinements

- Removed redundant legend and title overlay
- Converted filter buttons to serve as color-coded legend

- Disabled scroll-to-zoom (was disruptive), kept drag and +/- buttons
- Applied brand-specific colors to button borders

## 5. Category Performance Analysis Optimization

- Reduced dimensions by 10-20% across all layouts
- Tightened margins and gaps (35px → 25px, 20px → 15px)
- Smaller font sizes (titles 14px, axes 9-10px)
- Eliminated slide scrolling completely

## 6. Build Your Closet Advanced Features

- Expanded product database from ~23 to 55 items (15 Coach, 20 Gucci, 20 Hermès)
- Removed 5-item limit, now up to 15 items
- Implemented two-pass knapsack algorithm for budget optimization
- Added budget utilization indicator with color coding
- Fixed scrolling: only recommendations scroll, not entire slide
- Custom gold-themed scrollbars

## 7. Slide Instructions & Guidance

- Added three horizontally-aligned animated instructions on title slide:
  - Scroll to Explore
  - Hover to Learn More
  - Click Lightbulbs for Insights
- All use consistent pulse animation
- Updated subtitles on visualization slides with interaction instructions
- Fixed "Did You Know?" slide to prevent internal scrolling

## **Code Structure & Architecture Changes**

Original: Single index.html file with embedded CSS and JavaScript (likely 1000+ lines)

New Structure:

### 1. Separation of Concerns

- index.html (484 lines) - Clean HTML structure, script imports only
- style.css (2,811 lines) - All styling and animations
- `slides/\*.js` (20 modular files) - Each slide self-contained

### 2. Modular Slide System

- Each slide is an IIFE (Immediately Invoked Function Expression)
- Slides inject their own HTML into `scroll-container`
- Enables easy maintenance, testing, and reordering

### 3. Slide Files Created

```

slide1.js - Title slide  
slide1\_5.js - "Did You Know?" intro  
slide2.js - Chapter I divider  
slide3.js - Brand timeline  
slide4.js - Our Three Brands  
slide5.js - Chapter II divider  
slide6.js - Market Overview chapter  
slide7.js - Market Overview visualization  
slide7\_5.js - Revenue comparison (hidden)  
slide8.js - Global Traffic Hub (hidden)  
slide8\_5.js - Global Brand Dominance map  
slide9.js - Chapter III divider  
slide10.js - Category chapter  
slide11.js - Best Sellers carousel  
slide12\_5.js - Category Performance Analysis  
slide13.js - Chapter IV divider  
slide14.js - Resale chapter  
slide15.js - Resale overview (hidden)  
slide16.js - Vestiaire Marketplace Dashboard  
slide17.js - Chapter V divider  
slide17\_5.js - Build Your Closet chapter  
slide18\_5.js - Build Your Luxury Closet tool  
slide18.js - Team intro  
slide19.js - Team members  
slide20.js - Conclusion message  
```

#### 4. Global Constants Extraction

- `window.brandColors` defined in index.html (accessible to all slides)
- Ensures color consistency across all visualizations
- Single source of truth for brand identity

#### 5. CSS Organization

- Base styles (reset, body, containers)
- Slide-specific backgrounds and layouts
- Component styles (buttons, cards, tooltips)
- Animation keyframes
- Responsive media queries
- Visualization-specific styles
- Custom scrollbar styling

### Data Management Improvements

#### 1. CSV Data Loading

- All external data loaded via D3.csv() asynchronously
- Promise-based data fetching in index.html
- Data passed to visualization functions as needed

## 2. Hardcoded Constants Strategy

- `productDatabase` (55 items) in slide18\_5.js for recommendation engine
- `timelineData` (30 events) in slide3.js for historical visualization
- Strategic decision: these are curated content, not raw data

## Code Quality Enhancements

### 1. Function Encapsulation

- Each visualization has dedicated creation function
- Clear separation: `createWorldMap()`, `createPriceDistribution()`, etc.
- Filter state management with objects and Sets

### 2. Event Handling

- Centralized filter state in dashboard
- `updateAllViews()` orchestration function
- Proper event listener cleanup

### 3. Performance Optimizations

- Quantile-based scaling reduces computation on large datasets
- Efficient data filtering with Sets
- Debounced updates where appropriate

### 4. Accessibility & Overflow Management

- Slide-specific overflow rules (`slide-17-5`, `intro-slide`)
- Nested overflow control (hidden on parents, auto on children)
- Custom scrollbar styling maintains luxury aesthetic
- Proper z-index layering for modals (10000+)