

The Devil Wears Data: The Codebook

year

definition	The business year at which the brand operates
type of variable	numerical (yyyy)
range / possible values	2018 – 2024 (year)
data collection process	Data was collected on the brand's official website.

season

definition	The season at which the brand releases the collection, complements the year variable
type of variable	categorical
range / possible values	<ul style="list-style-type: none">- ss: Spring/Summer- fw: Fall/Winter
data collection process	Data was collected on the brand's official website & fashion week official schedules in London, Paris, NYC, Milan.

homebase

definition	The city at which the brand is based, e.g. London, Paris, Milan, Barcelona, etc.
type of variable	categorical (city)
range / possible values	Any city in the world, most likely Paris/London/New York/Milan.
data collection	Data was collected based on the brand's official website.

process	
----------------	--

parent

definition	The parent company that owns the brand, one of the large firms in luxurious fashion oligopolies
type of variable	categorical (brand)
range / possible values	<ul style="list-style-type: none"> - LVMH - Burberry - Kering - Independent
data collection process	Data was collected based on the brand's official website.

parent

definition	The parent company that owns the brand, one of the large firms in luxurious fashion oligopolies
type of variable	categorical (brand)
range / possible values	<ul style="list-style-type: none"> - LVMH - Burberry - Kering - Independent
data collection process	Data was collected based on the brand's official website.

director

definition	The brand's creative director at a given season of fashion
type of variable	categorical (creator name)
range / possible values	Names of any creative directors. e.g. Daniel Lee, Jonathan Anderson, Sean McGirr, Sarah Burton, Matthew Blazer, etc.

data collection process	Data was collected based on the brand's official website and Vogue magazines/press releases across many years from 2018 till now.
--------------------------------	---

director_years

definition	The number of years the creator works in the luxury fashion space prior to joining the house
type of variable	numerical (integer)
range / possible values	Non-negative integer from 0 to +inf
data collection process	Data was collected based on the brand's official website and Vogue magazines/press releases across many years from 2018 till now.

director_previous_house

definition	The number of luxury fashion houses the creator works in the luxury fashion space prior to joining the house
type of variable	numerical (integer)
range / possible values	Non-negative integer from 0 to +inf
data collection process	Data was collected based on the brand's official website and Vogue magazines/press releases across many years from 2018 till now.

newyork

definition	Whether the brand participates in the season's New York Fashion Week
type of variable	boolean
range / possible values	1 if yes, 0 if no

data collection process	Data was collected based on the official schedule of New York Fashion Week.
--------------------------------	---

london

definition	Whether the brand participates in the season's London Fashion Week
type of variable	boolean
range / possible values	1 if yes, 0 if no
data collection process	Data was collected based on the official schedule of London Fashion Week (related to British Fashion Council)

paris

definition	Whether the brand participates in the season's Paris Fashion Week
type of variable	boolean
range / possible values	1 if yes, 0 if no
data collection process	Data was collected based on the official schedule of Paris Fashion Week.

milan

definition	Whether the brand participates in the season's Milan Fashion Week
type of variable	boolean
range / possible values	1 if yes, 0 if no
data collection	Data was collected based on the official schedule of Milan Fashion Week.

process	
----------------	--

met

definition	Whether the brand participates in the season's Met Gala
type of variable	boolean
range / possible values	1 if yes, 0 if no
data collection process	Data was collected based on the reports from Vogue, L'Officiel, and Harper's Bazaar

employees_num

definition	The number of employees in the brand
type of variable	numerical (integer)
range / possible values	Non-negative integer from 0 to +inf
data collection process	Data was collected based on the brand's official website and the brand's official business release.

google_search_sentiment

definition	The sentiment score of the brand found through universal Google search
type of variable	numerical (float)
range / possible values	Real value from -1 to 1 (range of sentiment polarity score), with -1 indicating most negative sentiment and 1 indicating the most positive sentiment
data collection process	Data was collected based on a sentiment analysis run through a custom Google search engine based on Google Search API

fashion_magazine_sentiment

definition	The sentiment score of the brand found through fashion magazine posts, also searched on Google
type of variable	numerical (float)
range / possible values	Real value from -1 to 1 (range of sentiment polarity score), with -1 indicating most negative sentiment and 1 indicating the most positive sentiment
data collection process	Data was collected based on a sentiment analysis run through a custom Google search engine based on Google Search API, but now limiting the search to just fashion magazines, including Vogue, Harper's Bazaar, Business of Fashion, Elle, L'Officiel USA.