

Luxury Handbag Predictor



Ketevana Barabadze • General Assembly Data Science Immersive

Luxury Handbags Are Now the Fastest Growing Collectible



Facts

Designer handbags are have become extremely valuable assets.

They have surpassed art, classic cars, and rare whiskey in terms of their long-term value.

source : artmarketresearch.com

The global handbag market size was valued at \$47.0 billion USD in 2018.

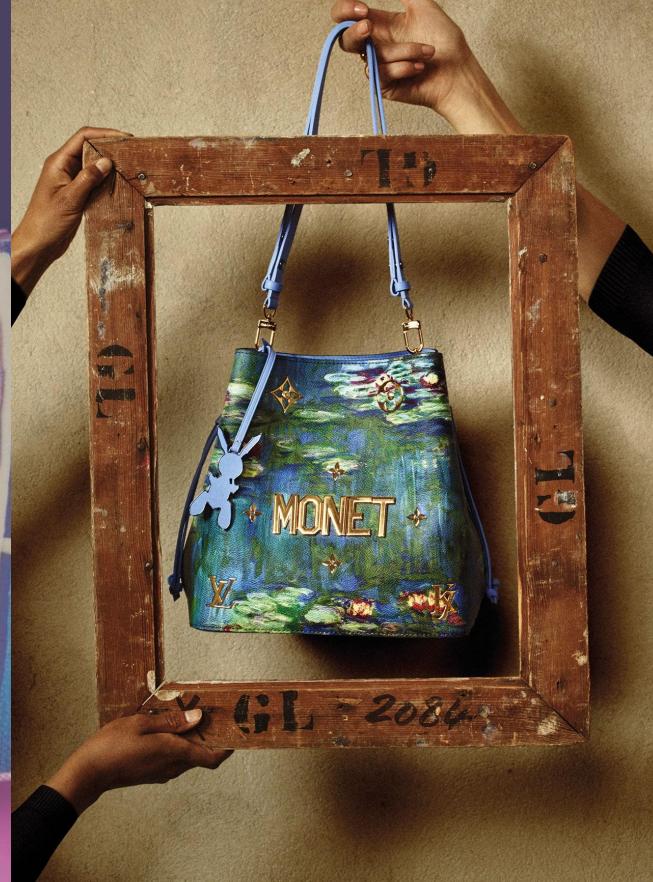
source : grandviewresearch.com

This year alone, over 3,700 designer handbags were sold at auctions around the globe generating more than \$32 million.

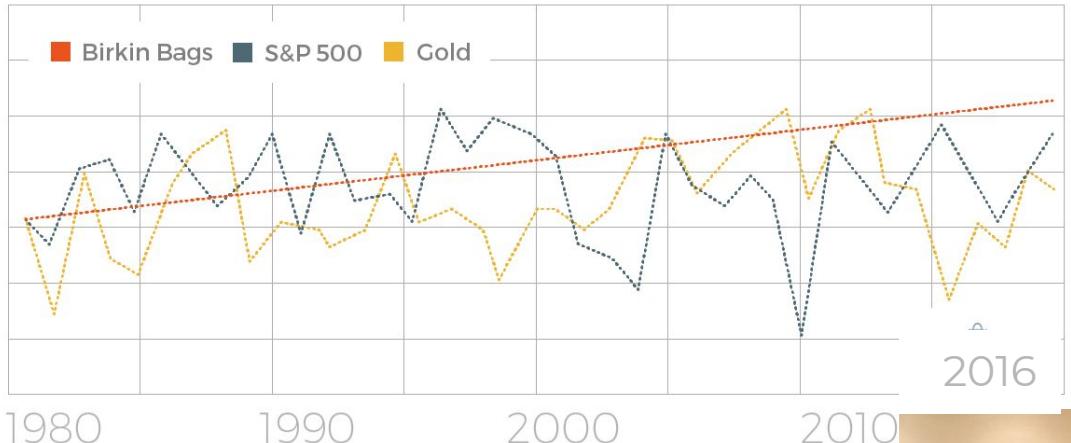
AMR found that Birkin bags increased their value on average 42% since 2019

Kelly, increased in value —129% over ten years.

Vestiaire Collective revealed that the Birkin is the most in demand bag right now and it increased in sales of +185% sin 2020.



Year on Year Percentage Increase for S&P 500, Gold & Birkin Bags



- Over the last 35 years, the Birkin has increased in value over 500% and has never decreased in value even during economic difficulties.
- Annually, the Birkin increases in value by an average of 14.2%.



CHRISTIE'S		LOT 3687			
HKD	港幣	160,000	RMB	人民幣	128,056
USD	美元	20,736	SGD	星幣	27,056
GBP	英鎊	13,312	NTD	台幣	644,800
JPY	日元	2,451,744	CHF	瑞郎	20,240
EUR	歐羅	16,832	THB	泰銖	682,512

A smaller image of the same white Hermès Birkin bag shown in the first image, displayed on a screen or projection during the auction. The bag is centered against a black background.

Overall Approach

- Web Scraping and Cleaning the Data
- Exploratory Data Analysis
- Defining the Target
- Modelling
- Interpreting results



Data

- Web Scraping from Vestiaire Collective webpage
- Added additional column of Mean Retail Price taken from official brand websites and public magazines

A screenshot of a Louis Vuitton Speedy Cloth Handbag listing on the Vestiaire Collective website. The bag is shown in a dark brown monogram canvas with red cherry blossom print. It features two gold-toned leather handles and a front flap closure. The listing includes a price of £3,960.70, payment terms, and seller information. The page also shows social sharing options and a detailed description section.

Details



LV SPEEDY 25 cerise
Lock & key included
Minor stains on handle
Tarnished and plating wear

Packaging: Dustbag, Card or certificate

Details
- Online since: 25/03/2021
- Categories : Women
- Category: Bags
- Sub-category: Handbags
- Designer: Louis Vuitton
- Model: Speedy
- Condition: Very good condition [More infos](#)
- Material: Cloth
- Colour: Brown
- Style:Vintage
- Location:Philippines, from the seller Maegan
- Reference: 15172658

Measurements
-Width: 10 in
-Height: 12 in
-Depth: 6 in

Sold with
-Dustbag
-Card or certificate

This item is offered by an individual seller. Its price has been suggested by its seller.

A screenshot of the Vestiaire Collective homepage. The top navigation bar includes links for New In, Designers, Women, Men, We Love, Vintage, Bags, Watches & Jewellery, Lifestyle, Express Delivery, and Direct Shipping. The main search bar is set to "Search by brand, article...". Below the search bar, there are several filter buttons: Chanel, Hermès, Louis Vuitton, Women, Bags, and Handbags. A "Create an alert" button and a "Clear filters" link are also present. The main content area displays a grid of handbag items. One item is highlighted: a Louis Vuitton Viva Cité cloth handbag with a price of £46,752,505.91 and a mean retail price of \$18,270,000. Other visible items include a Hermès Birkin 35 crocodile handbag and a Hermès Birkin 30 crocodile handbag.

Unprocessed Data

Data Cleaning

```
bags = pd.read_csv('preview.csv')
bags.head()
executed in 97ms, finished 14:49:03 2021-04-25
```

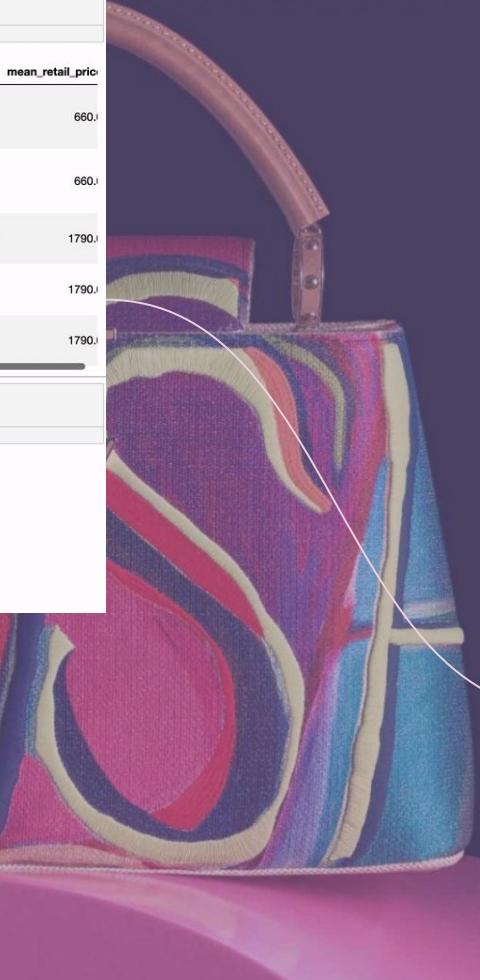
	title	price_range	price	discount_price	n_likes	limited	model	condition	m_model	material	color	m_material	sold_out	description	mean_retail_price	
Pochette Accessoire cloth clutch bag		£1,080 £776	£776	NaN	53.0	1	NaN	Never worn	Model: Pochette Accessoire	NaN	Color: Brown	Material: Cloth	1.0	NaN	660.0	
Pochette Accessoire cloth clutch bag		£495 £495	£495	NaN	Nan	42.0	1	NaN	Never worn	Model: Pochette Accessoire	NaN	Color: Brown	Material: Cloth	1.0	NaN	660.0
Speedy cloth handbag		£1,480 £1480	£1480	NaN	Nan	52.0	1	NaN	Never worn	Model: Speedy	NaN	Color: Pink	Material: Cloth	1.0	NaN	1790.0
Speedy cloth handbag		£1,880 £1880	£1880	NaN	Nan	59.0	1	NaN	Very good condition	Model: Speedy	NaN	Color: White	Material: Cloth	1.0	NaN	1790.0
Speedy cloth handbag		£1,850 £1,680 £1,680	£1,680	NaN	Nan	60.0	1	NaN	Very good condition	Model: Speedy	NaN	Color: Red	Material: Cloth	1.0	NaN	1790.0

```
# defining information about our data, checking types and non-null values
bags.info()
```

```
executed in 19ms, finished 14:49:03 2021-04-25
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12916 entries, 0 to 12915
Data columns (total 16 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   brand            12308 non-null   object  
 1   title             12916 non-null   object  
 2   price_range       12916 non-null   object  
 3   price              0 non-null      float64 
 4   discount_price    0 non-null      float64 
```

- Listings with null values dropped except description column
- Dropped resale prices with < £200 (typos / not real data points)
- Dummified or Binarized most of the features in next steps of preparation for modelling



Clean Data

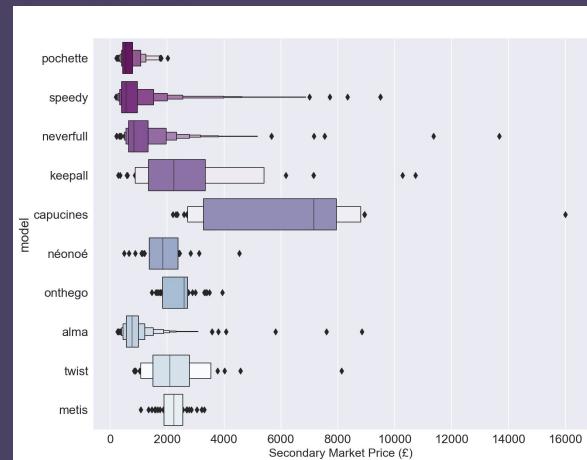
```
print(handbags.shape, '\n')
handbags.info()

executed in 19ms, finished 21:33:07 2021-05-05
(12774, 14)

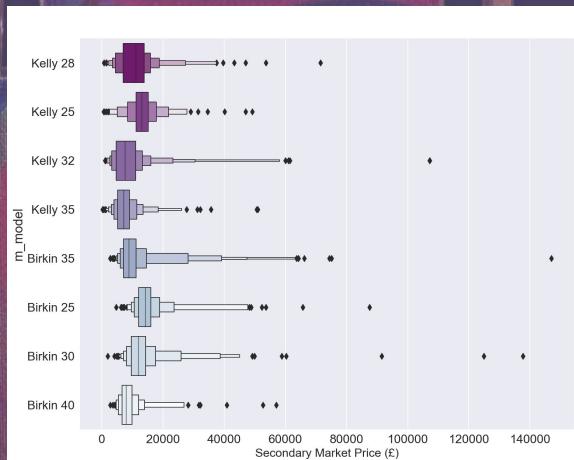
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12774 entries, 0 to 12773
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   brand            12774 non-null   object  
 1   title             12774 non-null   object  
 2   price             12774 non-null   float64 
 3   final_resale_price 12774 non-null   float64 
 4   n_likes           12774 non-null   float64 
 5   limited            12774 non-null   int64  
 6   model              12774 non-null   object  
 7   condition          12774 non-null   object  
 8   m_model             12774 non-null   object  
 9   material            12774 non-null   object  
 10  color               12774 non-null   object  
 11  sold_out            12774 non-null   float64 
 12  description         5122 non-null    object  
 13  mean_retail_price   12774 non-null   float64 
dtypes: float64(5), int64(1), object(8)
memory usage: 1.4+ MB
```

Resale Price Distribution

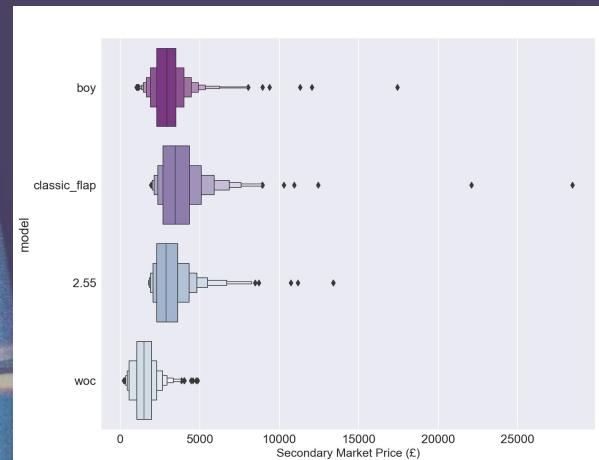
Louis Vuitton



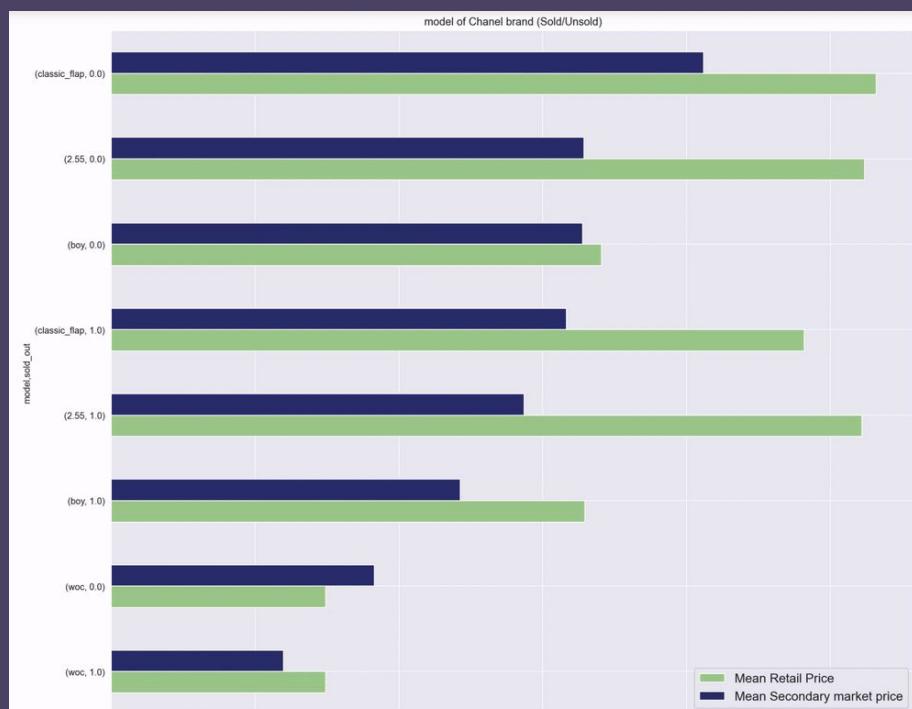
Hermès



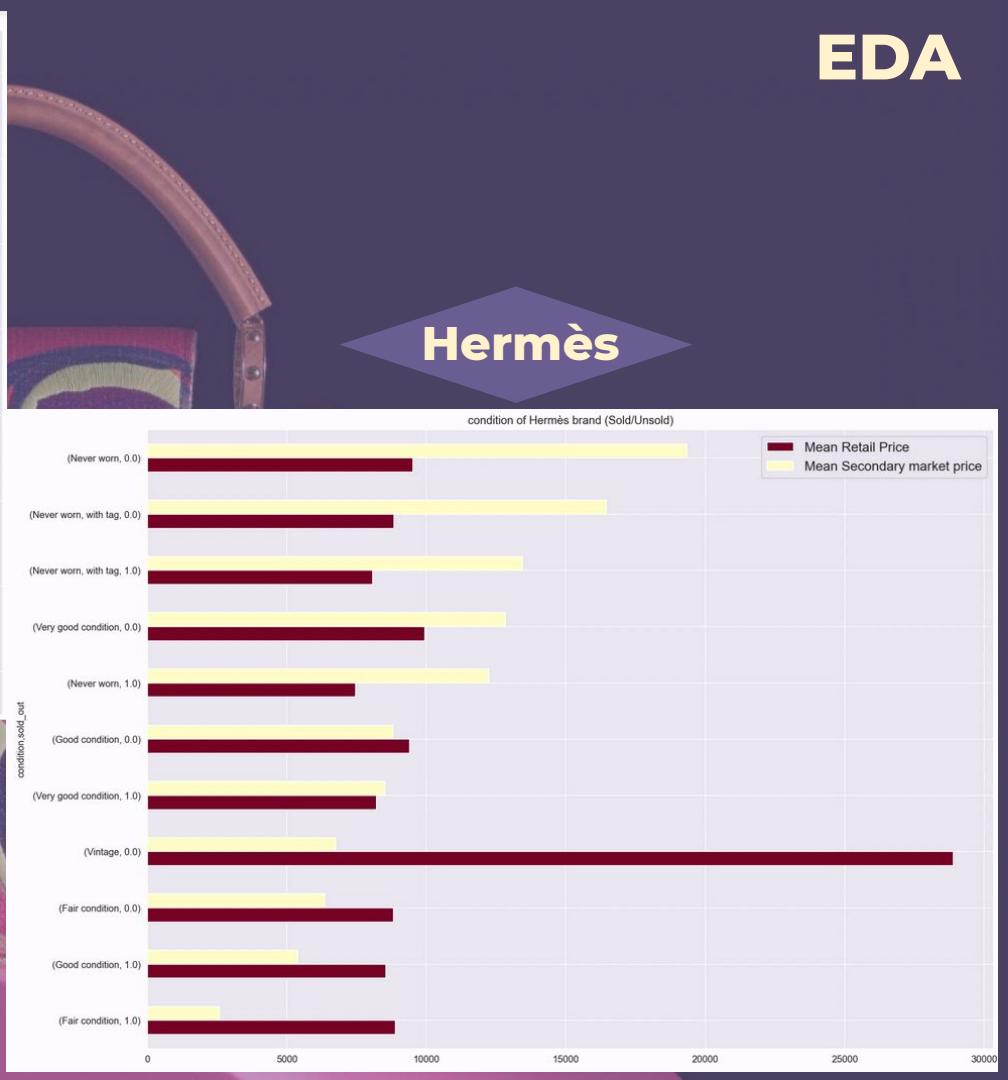
Chanel



model of Chanel brand (Sold/Unsold)



Chanel



Hermès

1.13 Impact of the bag being limited edition

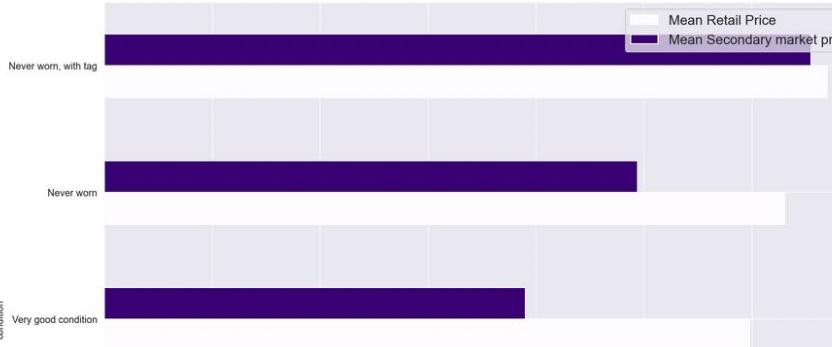
More than 1250 observations belongs to limited edition in Louis Vuitton brand.

Capucines and keepall are standing out, but most of the models has a positive return except twist, simple speedy and alma, marginal is pochette. Never worn and very good condition seems more valuable , along with unique materials and leather materials . Suede and exotic materials drop a lot in value. Profitable colors are Khaki, green, black, orange, whereas grey, camel, metallic doesn't make good returns.

```
for x in handbags.describe(include='object').columns:
    if x != 'description' and x != 'title' and x != 'brand' and x != 'model':
        handbags[(handbags.brand == 'Louis Vuitton') & (handbags.limited == 0)].groupby([x])[["mean_retail_price", "final_resale_price"]].mean().sort_values(by='final_resale_price').plot(kind='barh', figsize=(14, 12),
        title=x+" of Louis Vuitton brand (Non Limited)", cmap = 'Purples').legend(['Mean Retail Price', 'Mean Secondary market price'],
        prop={'size': 14})
```

executed in 1.75s, finished 20:59:01 2021-05-04

condition of Louis Vuitton brand (Non Limited)



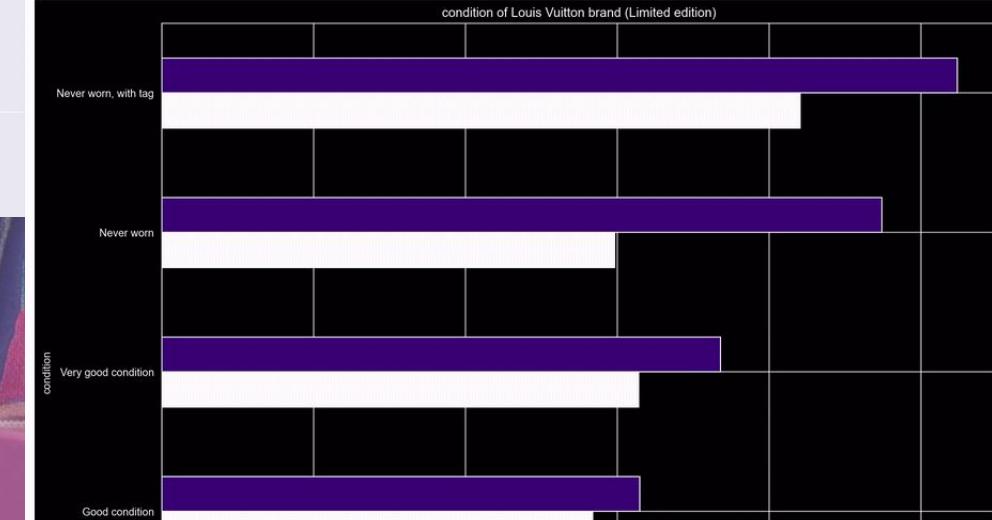
LV Non Limited

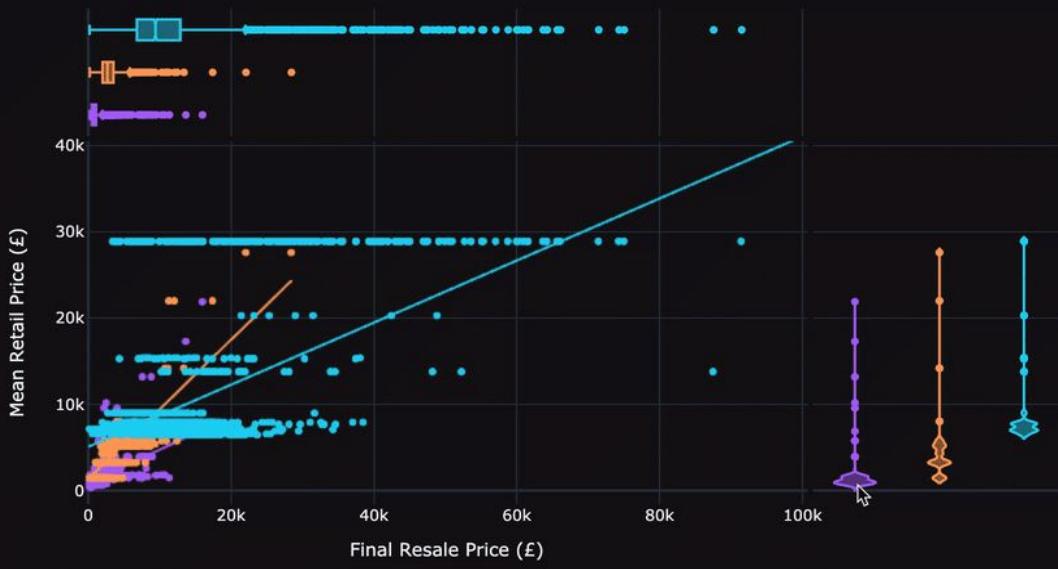
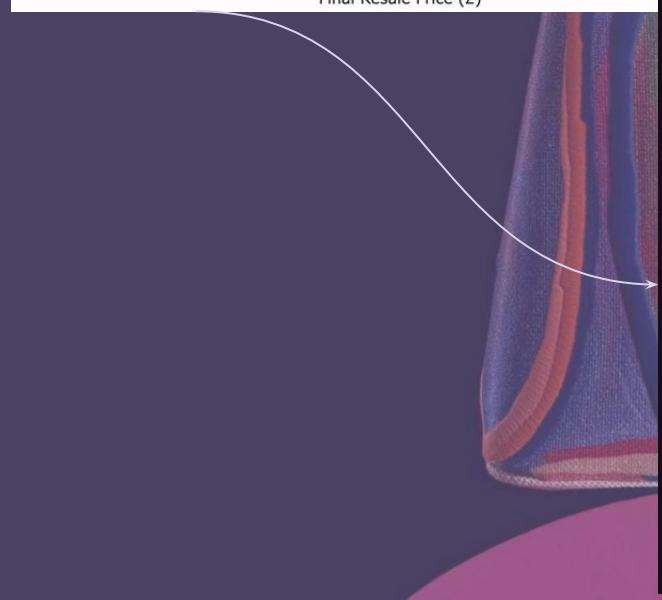
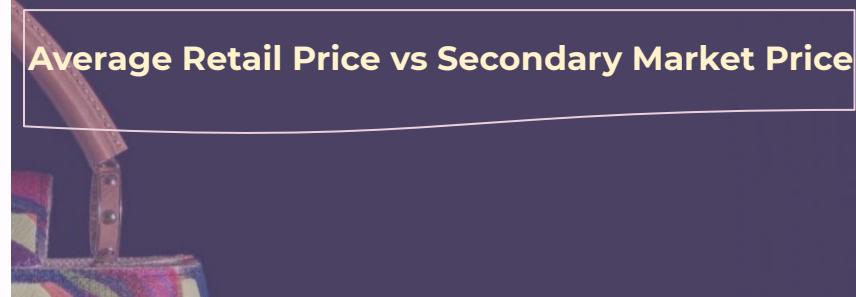
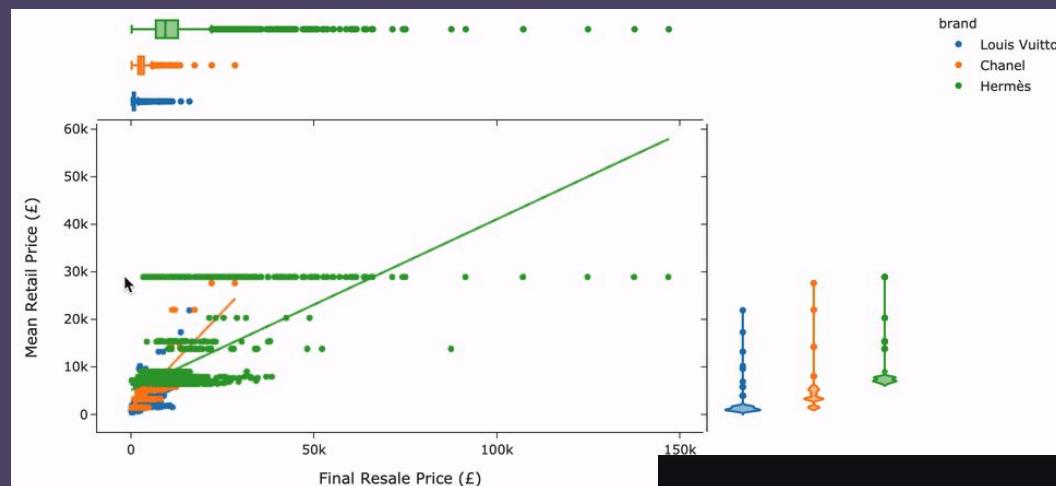


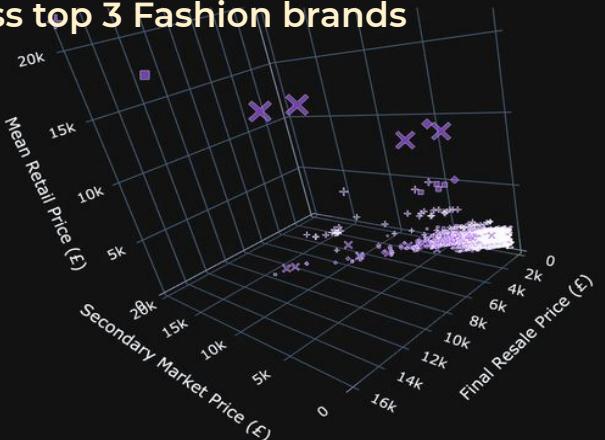
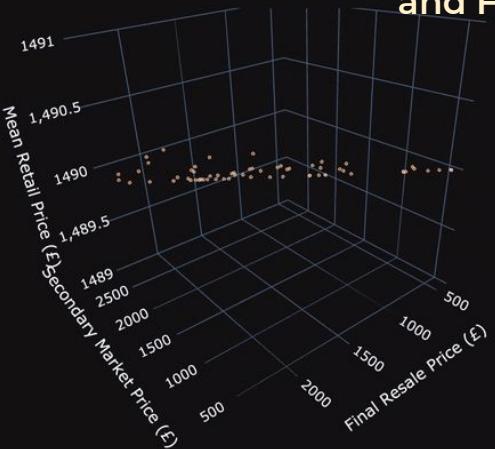
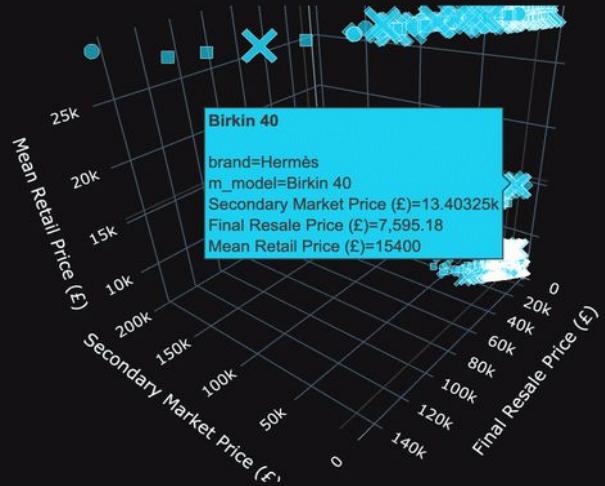
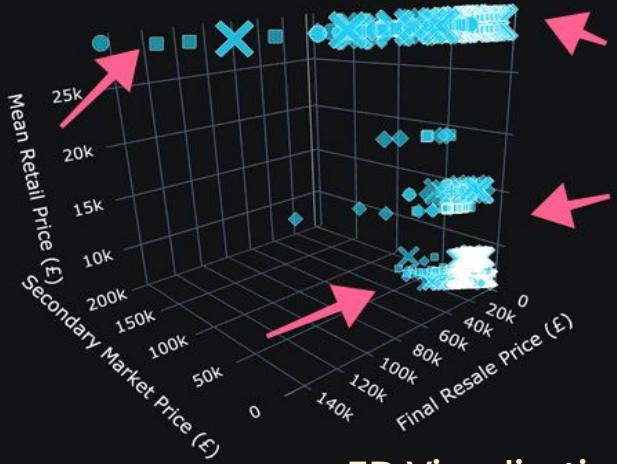
```
for x in handbags.describe(include='object').columns:
    if x != 'description' and x != 'title' and x != 'brand' and x != 'model':
        handbags[(handbags.brand == 'Louis Vuitton') & (handbags.limited == 1)].groupby([x])[["mean_retail_price", "final_resale_price"]].mean().sort_values(by='final_resale_price').plot(kind='barh', figsize=(14, 12),
        title=x+" of Louis Vuitton brand (Limited edition)", cmap = 'Purples').legend(['Mean Retail Price', 'Mean Secondary market price'],
        prop={'size': 14})
```

executed in 1.60s, finished 21:09:32 2021-05-04

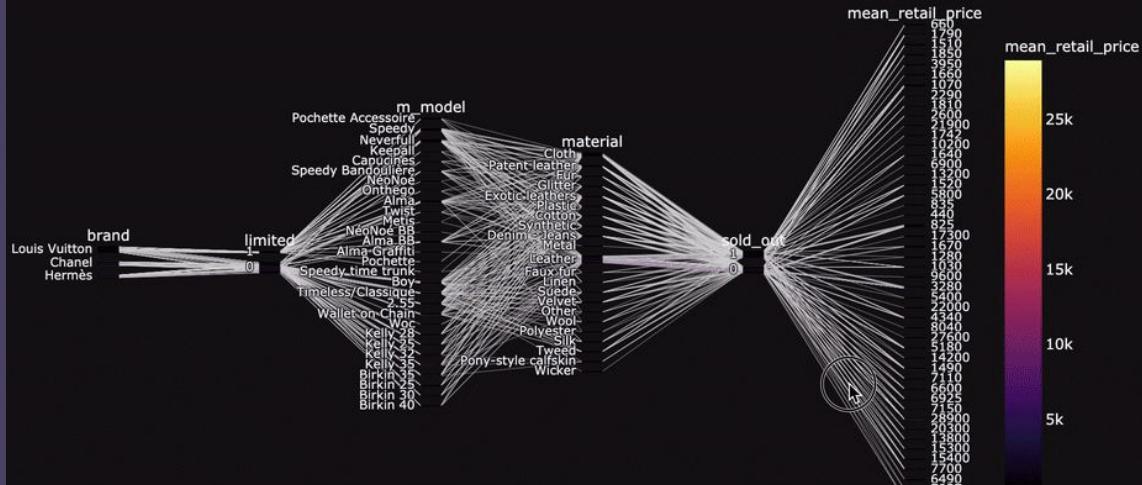
Louis Vuitton Limited Edition



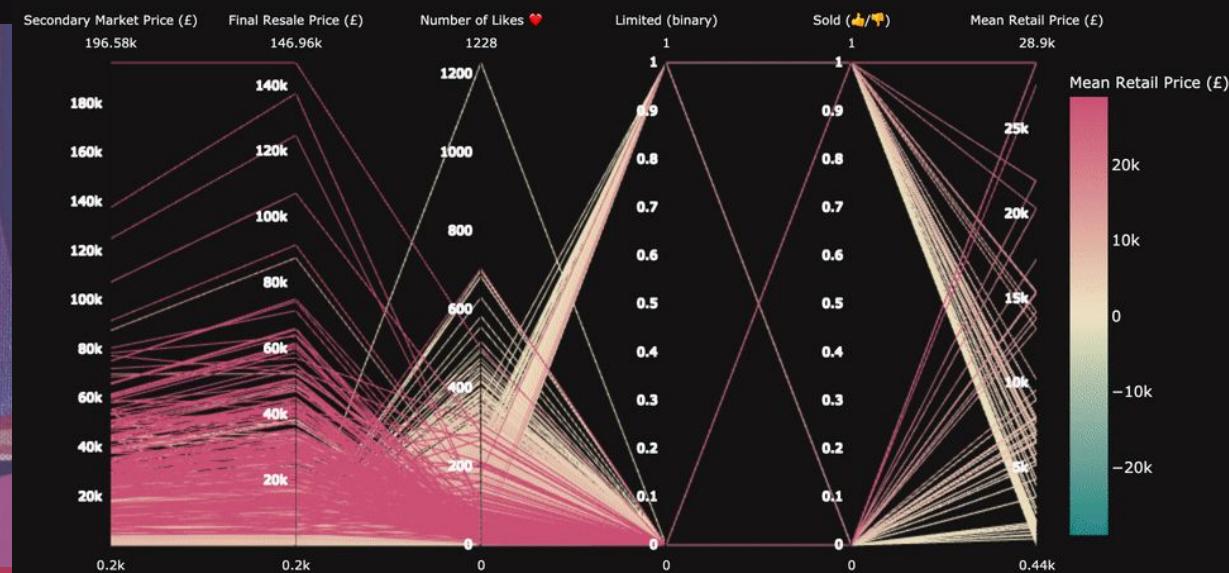




- brand, m_model
- Louis Vuitton, Pochette Acce
 - Louis Vuitton, Speedy
 - Louis Vuitton, Neverfull
 - Louis Vuitton, Keepall
 - Louis Vuitton, Capucines
 - Louis Vuitton, NéoNoé
 - Louis Vuitton, Onthego
 - Louis Vuitton, Alma
 - Louis Vuitton, Twist
 - Louis Vuitton, Metis
 - Louis Vuitton, NéoNoé BB
 - Louis Vuitton, Alma BB
 - Louis Vuitton, Alma Graffiti
 - Louis Vuitton, Pochette
 - Louis Vuitton, Speedy time t
 - Chanel, Boy
 - Chanel, Timeless/Classique
 - Chanel, 2.55
 - Chanel, Wallet on Chain
 - Chanel, Woc
 - Hermès, Kelly 28
 - Hermès, Kelly 25
 - Hermès, Kelly 32
 - Hermès, Kelly 35
 - Hermès, Birkin 35
 - Hermès, Birkin 25
 - Hermès, Birkin 30
 - Hermès, Birkin 40



Features Network



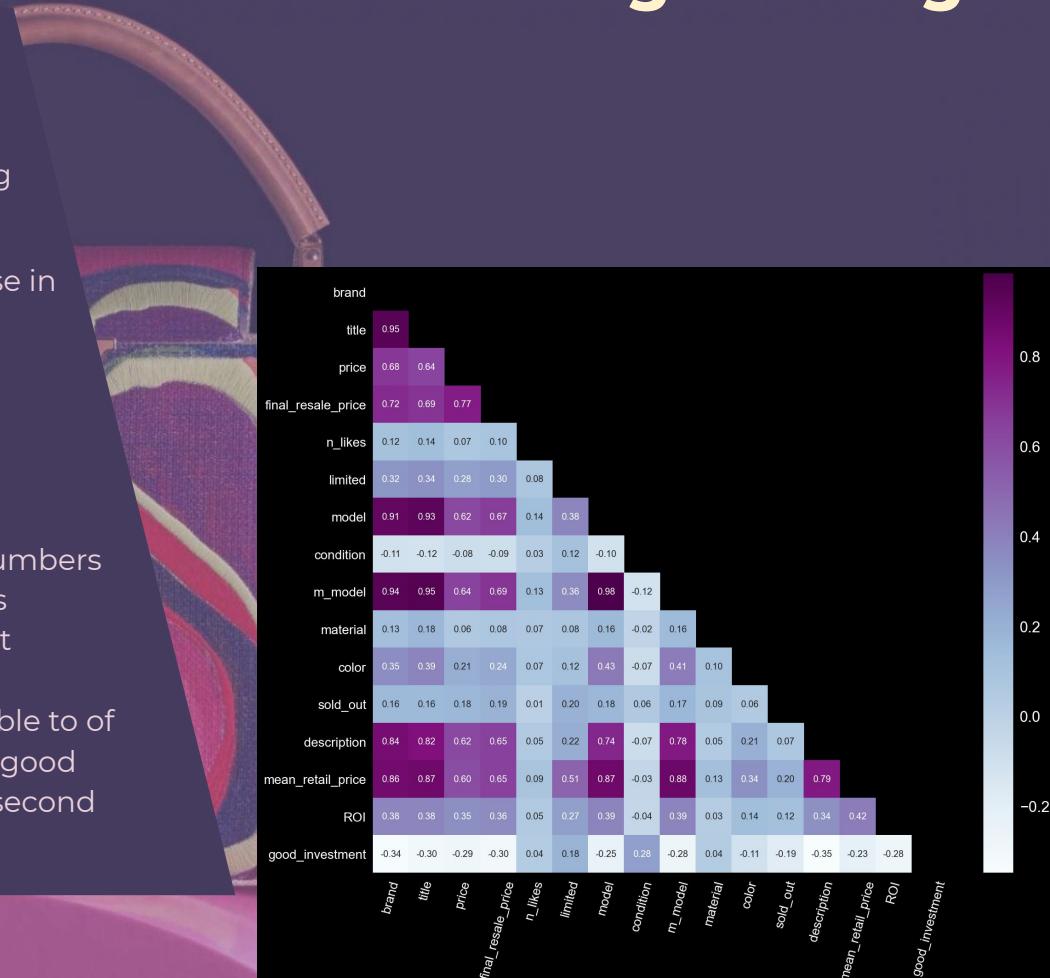
Defining the target

Audience

- As a seller, I would like to predict the Secondary market price based on bag features.
- As a buyer, I would like to predict increase in value and determine whether it is an investment purchase.

Targets

- Regression models with the continuous numbers of resale value of designer handbags ('final_resale_price') as the first target
- Classification models with categorical variable to defining whether designer handbag is a good investment or not ('good_investment') as second target



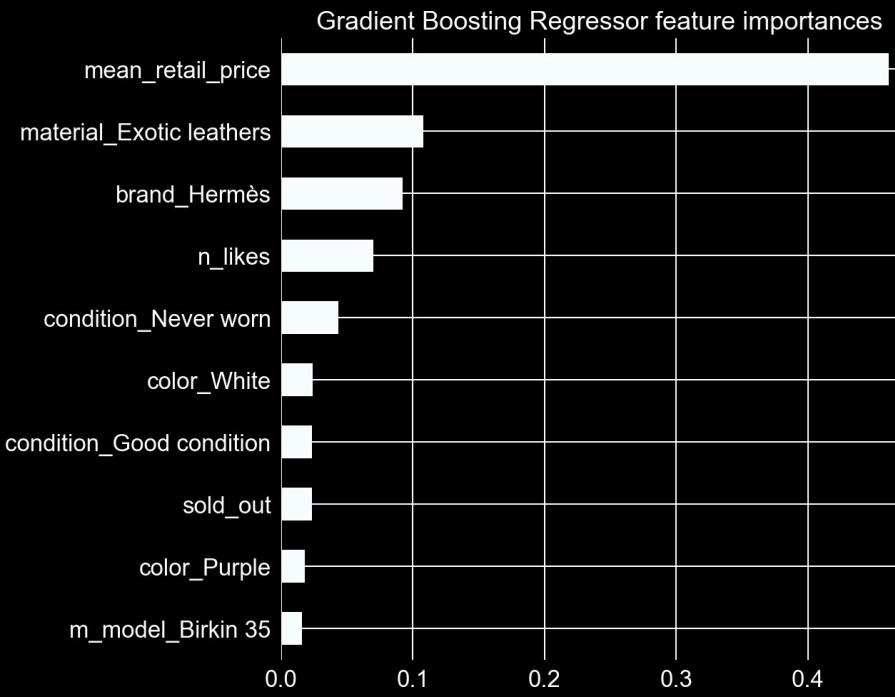
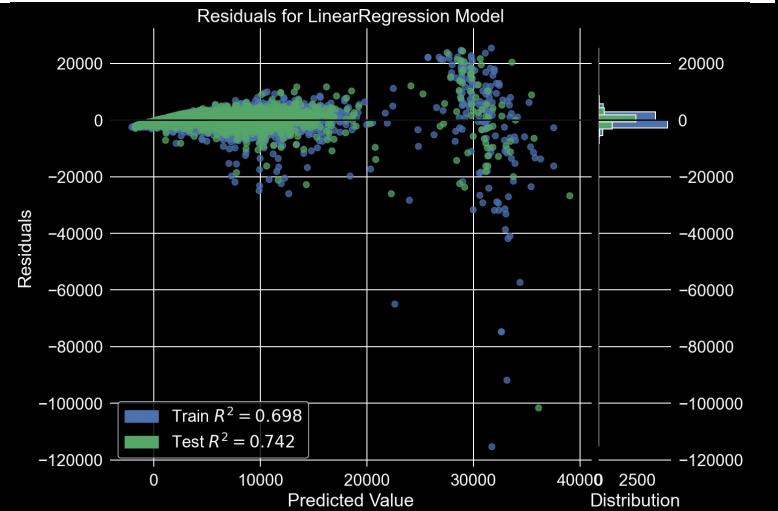
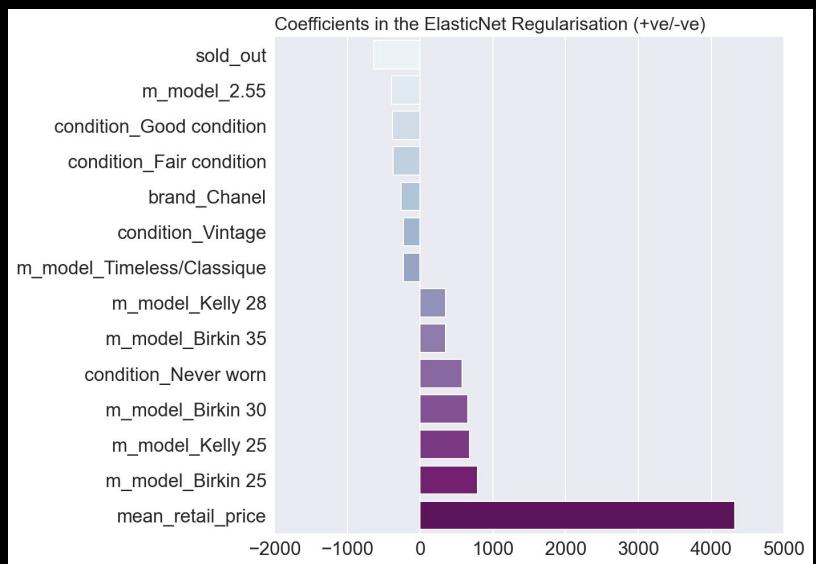
Modelling with sklearn

List of Regression models I applied, along with their highest cross validated r-squared scores on the training data:

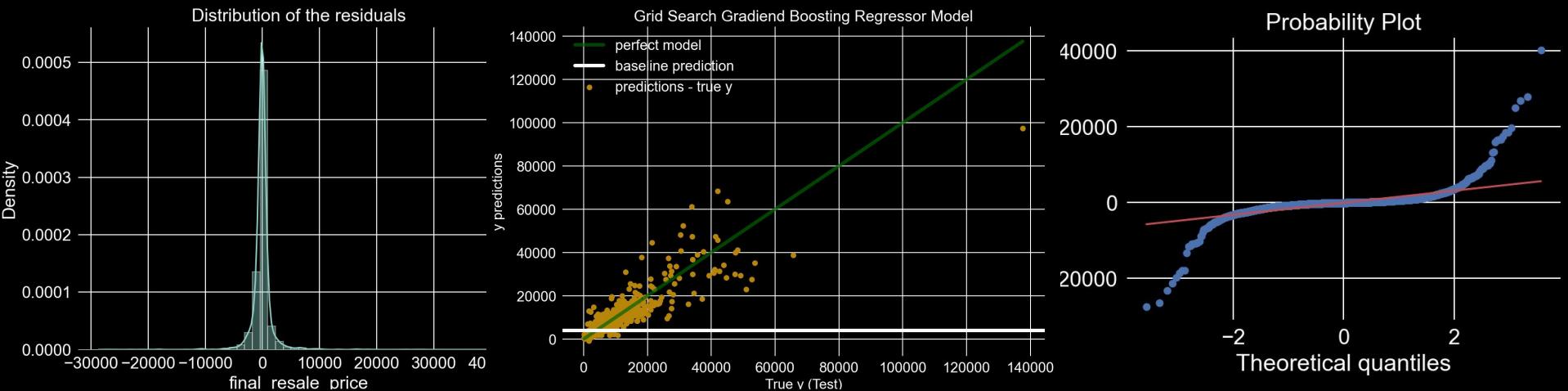
	Method	R2_Mean_CV_train	R2_test	RMSE_test
0	LinearRegression	0.68668	0.74243	2994.02
1	LR_Lasso	0.68719	0.74342	2988.27
2	LR_Ridge	0.68685	0.74258	2993.16
3	LR_ElasticNet	0.68722	0.74322	2989.49
4	GradientBoostingRegressor	0.73204	0.85844	2219.62
5	RandomForestRegressor	0.78004	0.82971	2434.49



Modelling



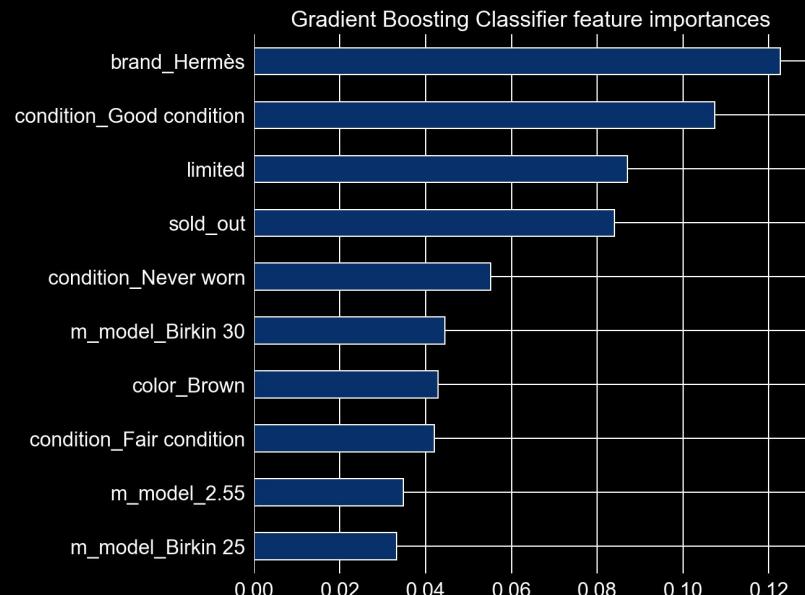
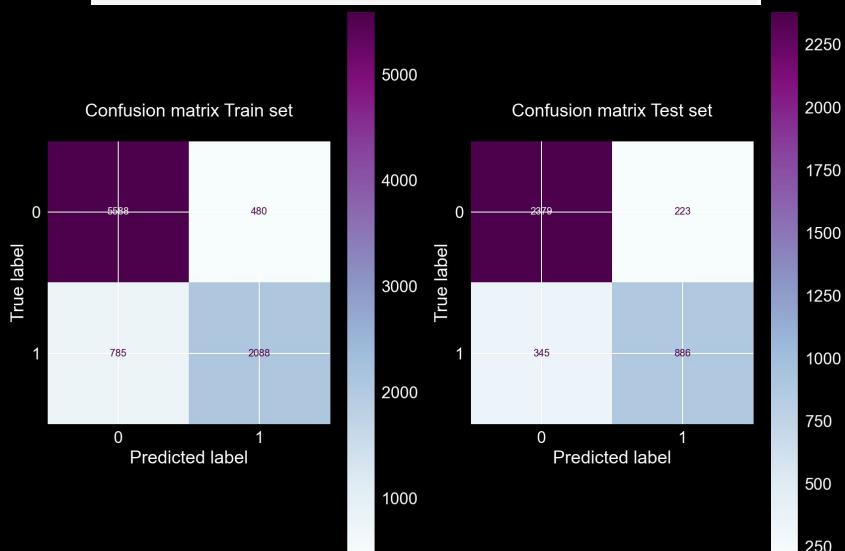
Modelling



Modelling

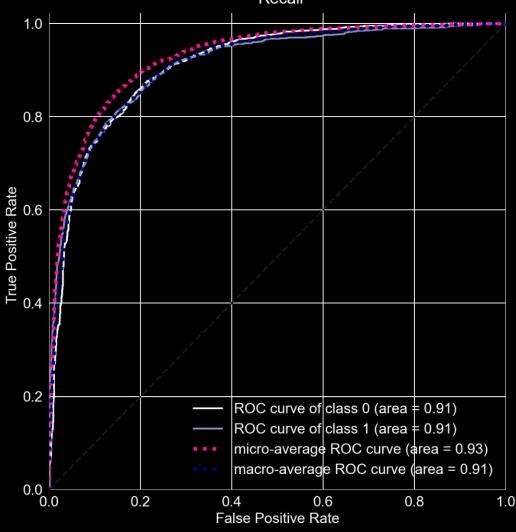
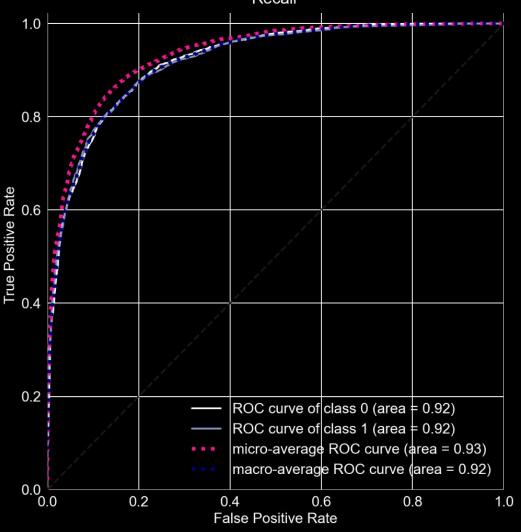
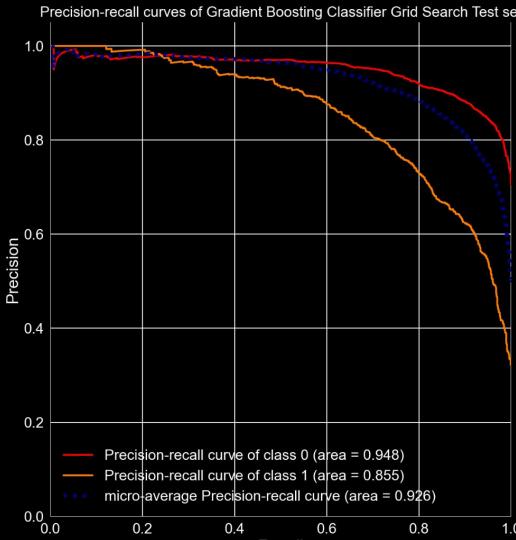
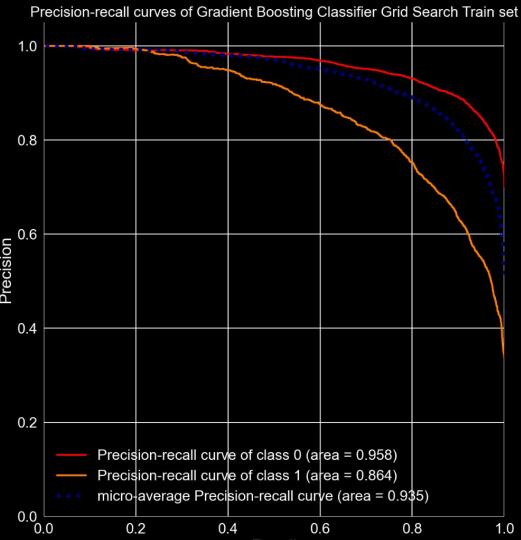
List of Classification models I applied, along with their highest cross validated r-squared scores on the training data and r-squared scores on unseen data :

	Method	R2_Mean_CV_train	R2_test
0	LogisticRegression	0.83838	0.84659
1	LogisticRegression_GridS	0.83805	0.84685
2	DecisionTreeClassifier_GridS	0.83626	0.83407
3	RandomForestClassifier_GridS	0.84163	0.84607
4	GradientBoostingClassifier_GridS	0.84498	0.85181



	precision	recall	f1-score	support
0	0.87	0.91	0.89	2602
1	0.80	0.72	0.76	1231
accuracy				0.85
macro avg	0.84	0.82	0.83	3833
weighted avg	0.85	0.85	0.85	3833

Modelling



Findings / Conclusions

Key points to make your purchase
a good investment:

Limited Editions

Patience is a Virtue

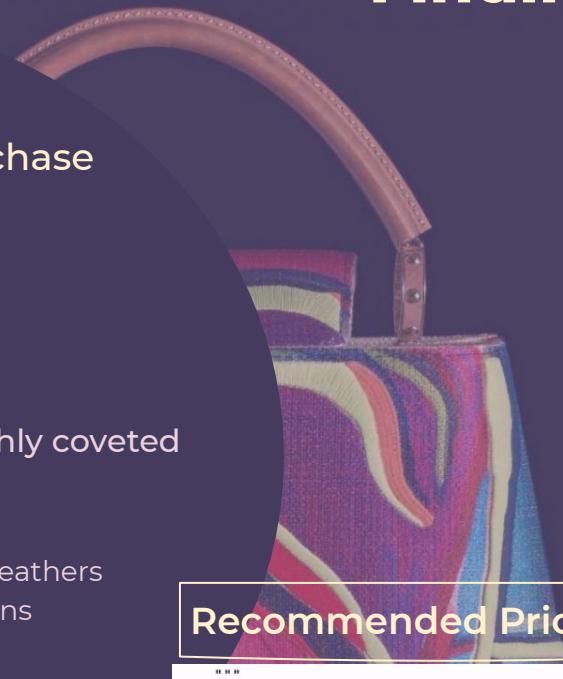
Maintain the Condition

Anything that is hard to get is highly coveted

Invest in the holy trinity:

- Hermes iconic bags with exotic leathers
- Louis Vuitton bags limited editions
- Chanel Woc / Classic flaps

THE HOLY
TRINITY



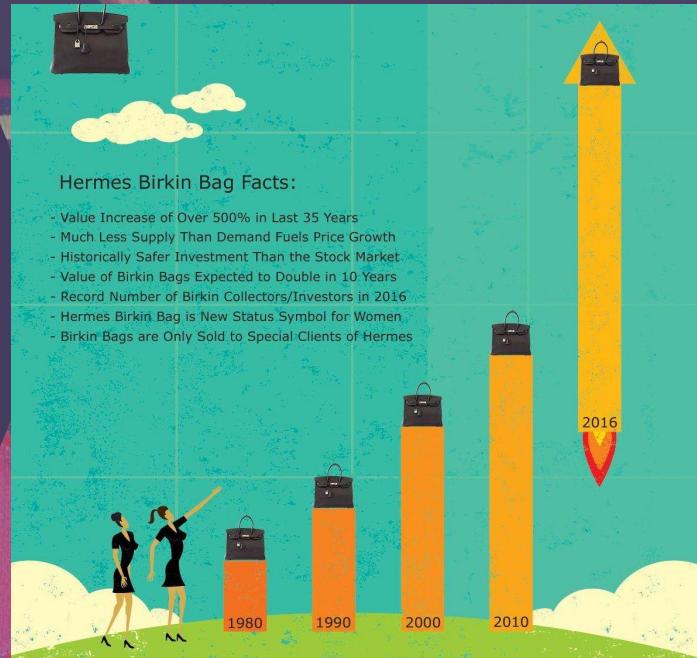
Recommended Price for your handbag for Resale

```
"""
Function: takes a dataframe and categorical features, then generates the name of the bag,
it's approximate evaluation at current time and it gives a price range with minimum and maximum values
based on the criteria's for this exact bag.
"""
```

```
return_resale_price(handbags, 'Chanel', '2.55', 'Cloth', 'Black')|  
executed in 2ms, finished 22:05:16 2021-05-05
```

Limitations / Future work

Time data for each handbag



Any Questions or Comments?

Thank You

