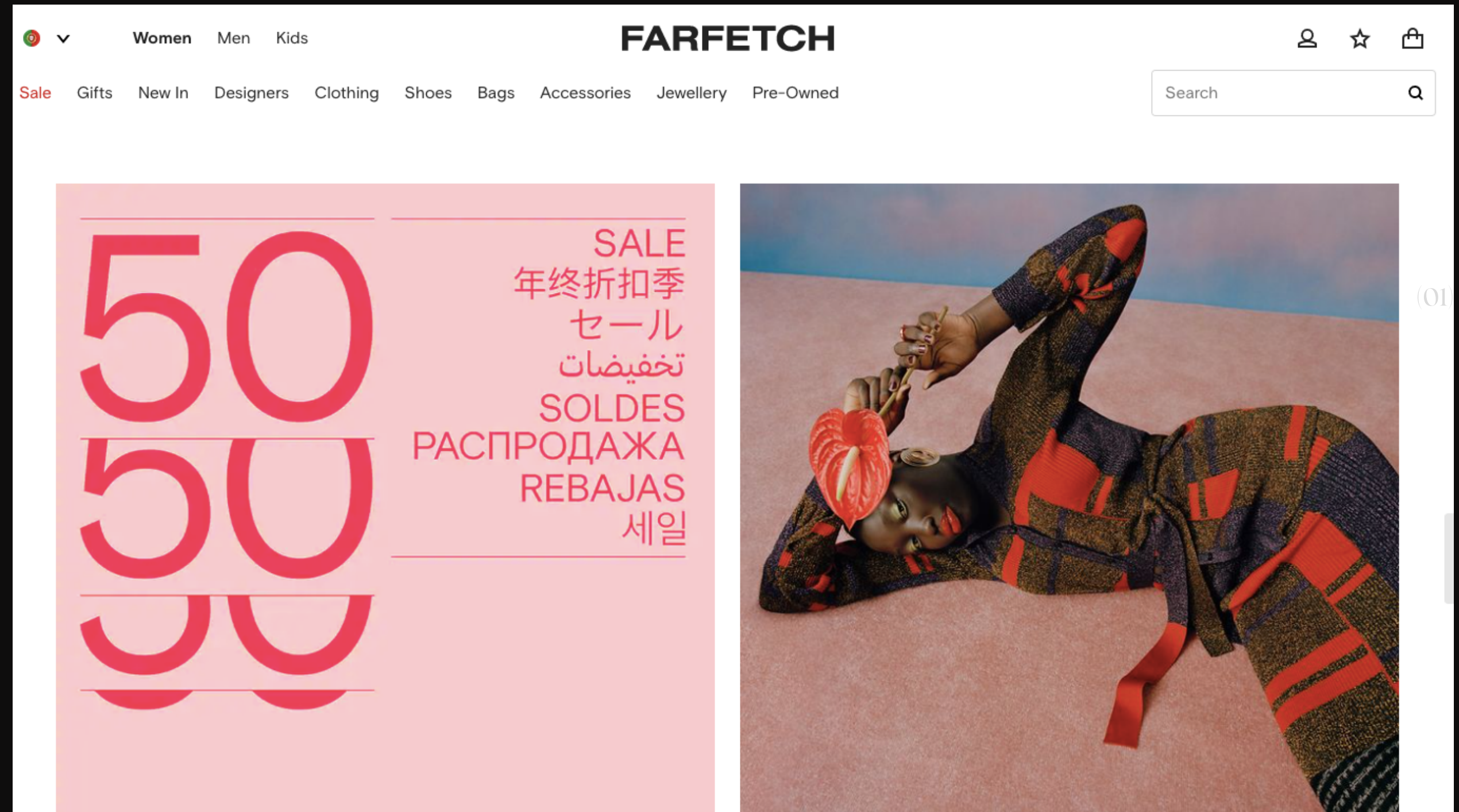


FINAL PROJECT

DECEMBER 4TH 2020

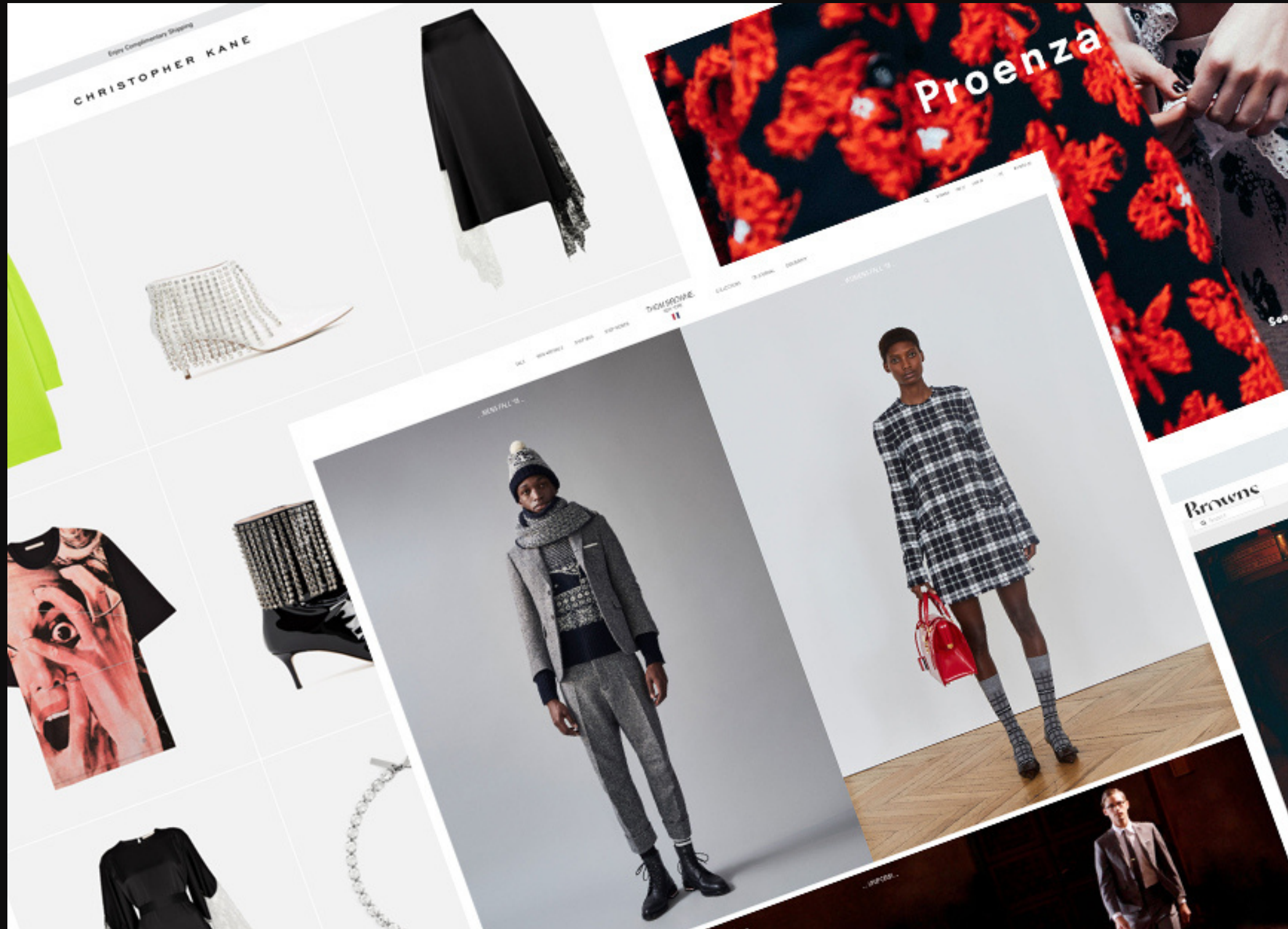
Farfetch Discount Predictor



WOMEN AND
ACCESSORIES

“Why luxury
goods from
Farfetch have
discounts?”

CLASSIFICATION PREDICTIVE
MODEL



Agenda

INTRODUCTION

PROJECT OVERVIEW

DATA ANALYSIS

PREDICTIVE MODEL

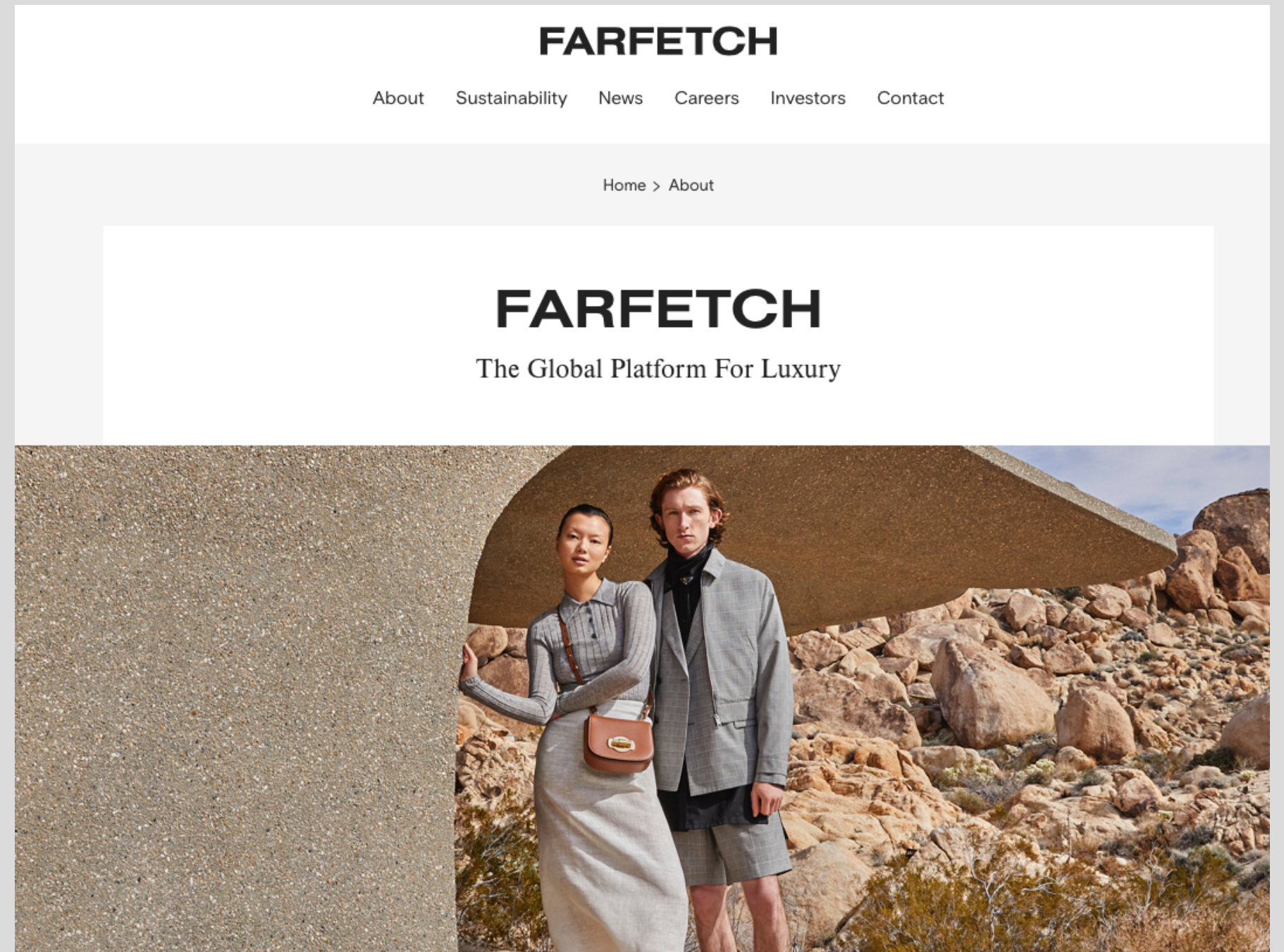
PROBLEMS

SOLUTIONS AND

RECOMMENDATIONS

INTRODUCTION

"FARFETCH is the leading global technology platform for the luxury fashion industry. E-commerce Marketplace connects customers in over 190 countries with items from more than 50 countries and over 1,200 of the world's best brands (...)"



Project Overview

GitHub
Repository

Location of the Project with all files used in the creation of the new model

Jupyter Notebook

Open-source web application that combines exploratory data analysis and data modelling

Tableau Software

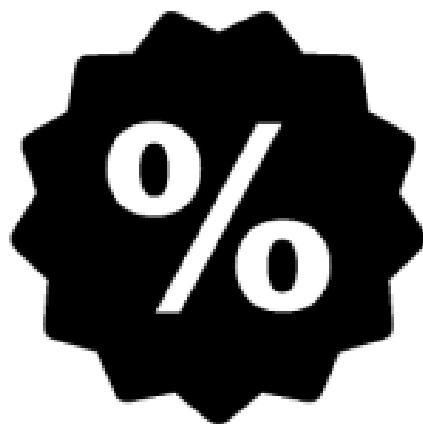
Data visualisation tool used for data analysis and creation of new dashboard

Data Analysis

Key reasons affecting the discount on a product?

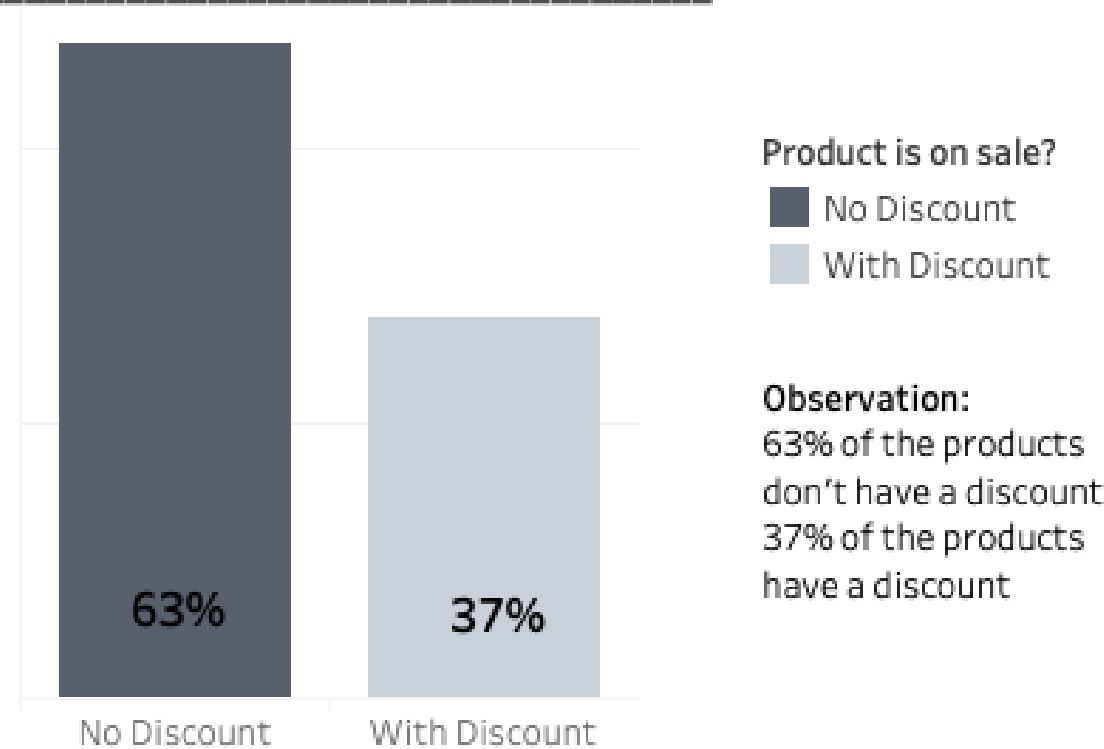
Part 1

Finding data patterns to explain which are the products from Farfetch that have a discount.

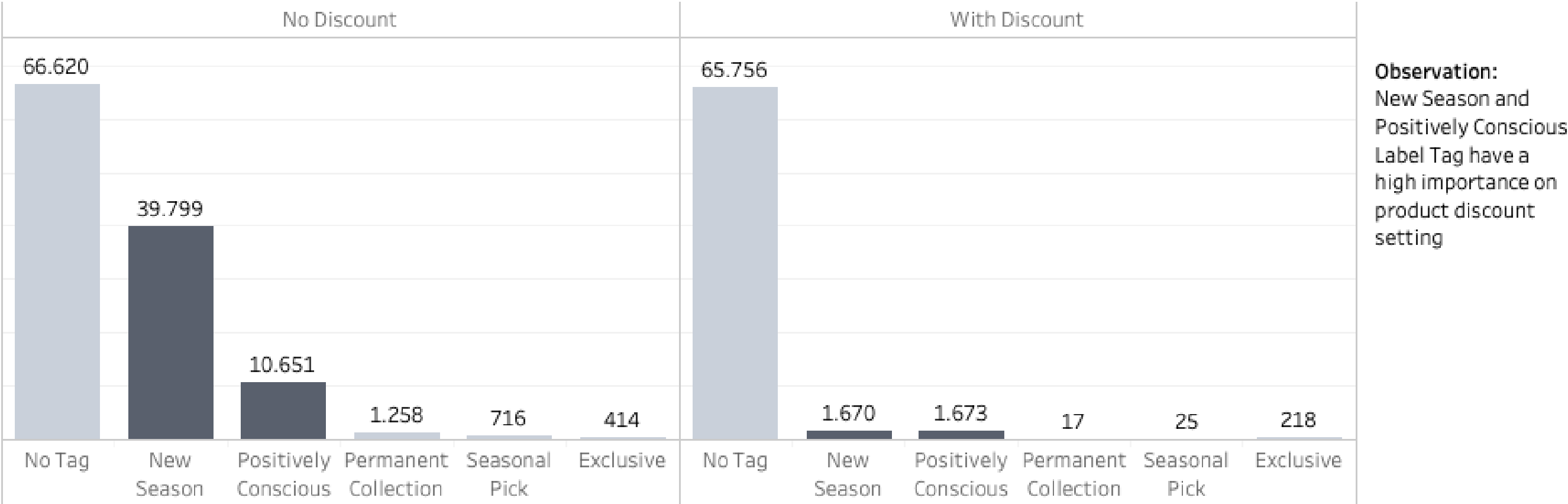


FARFETCH

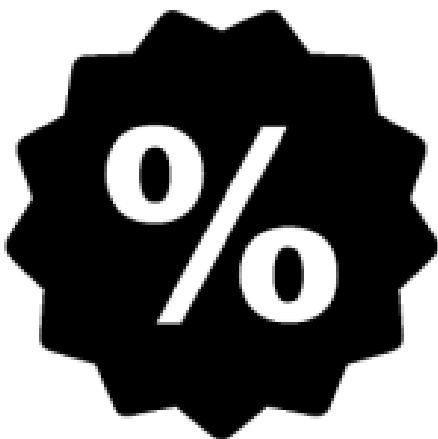
How many products are on sale?



Merchandise vs Products on sale



Key reasons affecting the discount on a product?
Part 2



89%

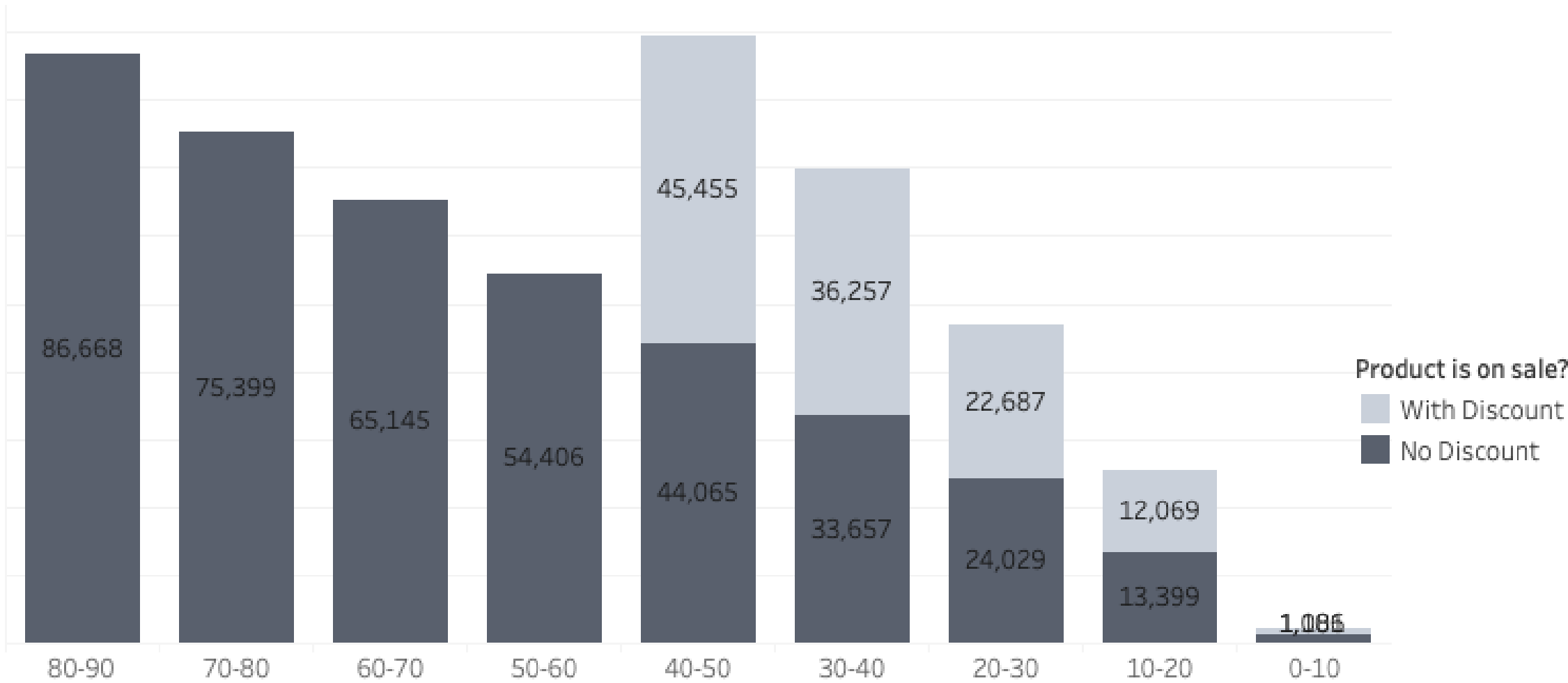
of the items with discount have less than 20 units available in stock

FARFETCH

How many products are on sale?

| | With Discount | No Discount |
|-----------------------|---------------|-------------|
| Number of total items | 69,359 | 119,458 |
| Maximum price | 46,870 | 89,724 |
| Minimum price | 24 | 2 |
| Average price | 1,049 | 1,336 |

Price Range vs Product on sale



PREDICTIVE MODEL

FIRST MODEL



Logistics Regression
Model

Result

Accuracy %: 0.75
Kappa Score : 0.50

SECOND MODEL



SMOTE
(oversampling)

Result

Accuracy %: 0.77
Kappa Score : 0.53

THIRD MODEL



Decision Tree
Classifier

Result

Accuracy %: 0.85
Kappa Score : 0.52

FOURTH MODEL



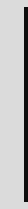
Ada Boost
Classifier

Result

Accuracy %: 0.80
Kappa Score : 0.58

PREDICTIVE MODEL

FIFTH MODEL



Random Forest Classifier

Result

Accuracy %: 0.88

Kappa Score :
0.74

Problems

Irrelevant Features

Number of Features

Diversified Dataset

Solutions & Recommendations

Creation of a item category

Creation of range of prices feature

Additional features related with
item characteristics

Testing new model predictors and
check feature importance

Thank you!

For questions contact:

Phone Number

123-456-789

Email Address

catarinabatista@ironhack.com