

THE BUSINESS GAME

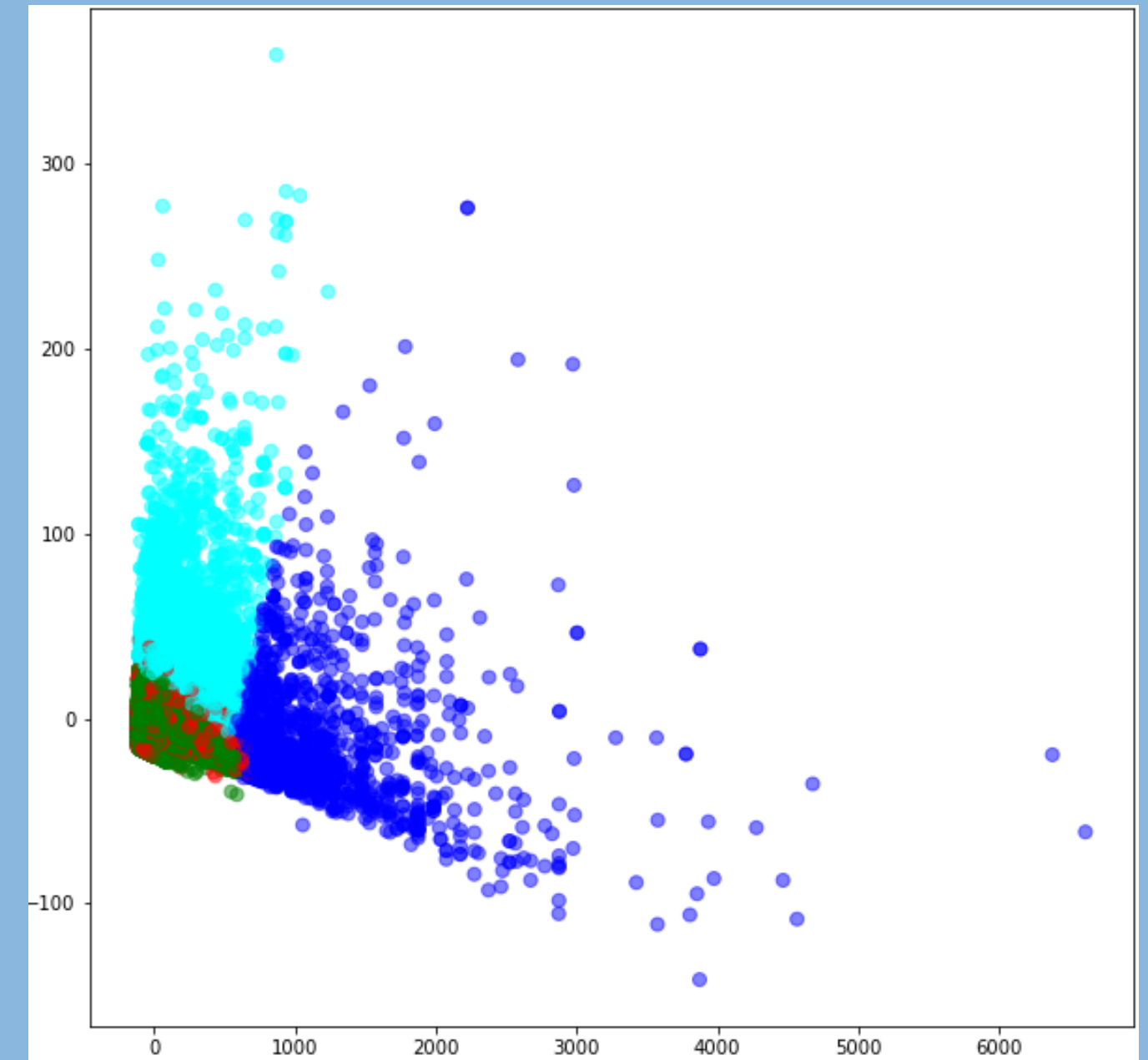
NTT DATA X LUISS



**Lorenzo Baglini
Edoardo Borriello
Giacomo Flores**

CLUSTERING

- Clusters based on Price, Shipping cost, Review score
- Identified 4 as the optimal number of clusters using the elbow method
- Analysis of the cluster to implement the best market strategy for every group



MARKETING STRATEGY

Cluster 0

- The one with more observations
- Rating value of 4 and 5
- Price from 0.85€ to 760€
- Shipping cost up to 44.88€

Cluster 1

- Rating values ranging from 1 to 3
- Price and shipping cost similar to cluster 0

Cluster 2

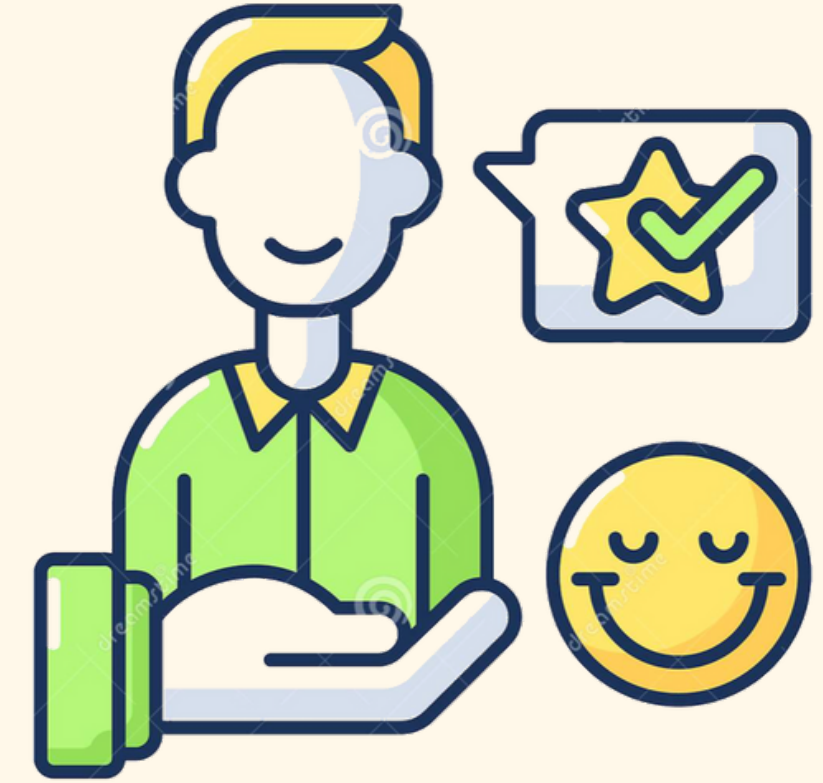
- All rating values (from 1 to 5)
- Price little higher (from 9.9€ to 1149€)
- High shipping cost (from 29.71€ to 409.68€)

Cluster 3

- All rating values (from 1 to 5)
- Price starts from 709.9€ and goes up to 6735€
- All possible shipping cost

SATISFIED CUSTOMERS

Cluster 0



- **Continue to offer the same services**
- **Discount campaigns to avoid dropout**
- **Advertising and recommendation of products similar to the ones they already purchased**
- **Act promptly in case of negative reviews or experiences**
- **Measure your customer satisfaction regularly**

UNSATISFIED CUSTOMERS

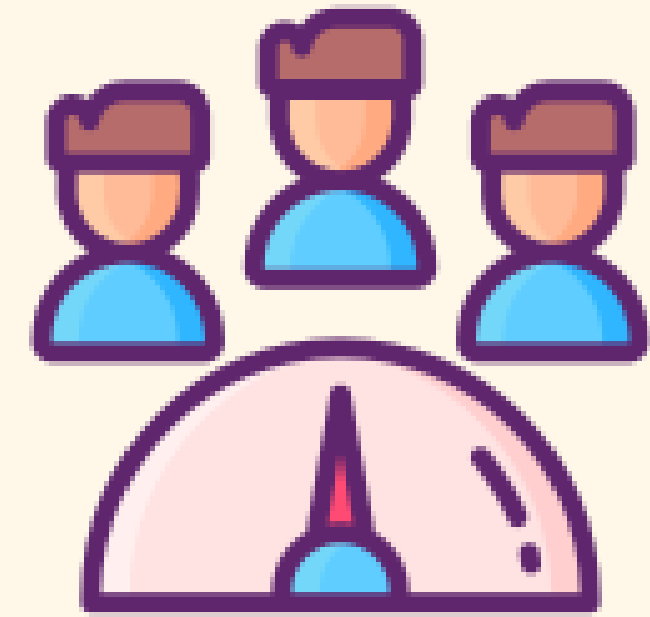
Cluster 1



- **Analyze negative comments and reviews to identify the problems**
- **Discount campaigns to attract the costumer again on the website**
- **Advertising and recommendation of products that other costumers bought instead of the one they did not liked**
- **Ask for feedback on what went bad to show interest in their problems**
- **Provide instant response with live chat**

AVERAGE CUSTOMERS

Cluster 2



- Continue with similar strategies for products and categories they already purchased and liked
- Analyze negative feedbacks and reviews to understand preferences
- Try to reinforce brand image to retain the costumer
- Provide instant response with live and bot chat to avoid website abandonment
- Try to identify one-time customers to avoid overspending

HIGH SPENDERS

Cluster 3



- **Very similar to the average customers**
- **Focus on negative experiences to try to retain the costumer**
- **Continue with similar strategies for products and categories they already purchased and liked**
- **Provide instant response with live and bot chat to avoid website abandonment**

RECOMMENDATION SYSTEM

- Prediction using Surprise python library
- Use Singular Value Decomposition (SVD) algorithm
- Find best recommendations for every user ID selected
- Print distribution of ratings of any product

```
# find best product for every customer (uid)
algo.fit(data.build_full_trainset())
my_recs = []
for iid in unique_ids:
    my_recs.append((iid, algo.predict(uid='24f12460aad399ba18f4ed2c2fbab65d', iid=iid).est))
pd.DataFrame(my_recs, columns=['iid', 'est']).sort_values('est', ascending=False).head(10)
```

	iid	est
2361	06bf70b6e1d67d96308235ef350edc61	5.000000
731	777d2e438a1b645f3aec9bd57e92672c	5.000000
11919	3215010238fcd9cab6ba7d2b81a6973d	5.000000
7731	a04f52ded97b5530e8783e3c002b90f0	5.000000
4220	45e967683e7292b195609137fadaf2fe	4.999547
1005	e5ae72c62ebfa708624f5029d609b160	4.993086
10407	90f97298579cd20412fdcc9b7a2d4b6b	4.980601
4190	c0350d6ac413eda4641bf92ab687f1b5	4.980283
3120	a0abcee0132a5aed003d98e459b37698	4.966959
330	ed2067a9c1f79553088a3c67b99a9f97	4.955853



DELIVERY TIME PREDICTION

Prediction done using Polynomial Regression

Computed distance between seller and customer city using the geographic coordinates provided

Computed the size of every product using the measurements provided

Used this two data together with the estimated shipping time to train the model

Predicted the real shipping time for every order in the dataset



**THANKS FOR
YOUR
ATTENTION**