

HOTEL CUSTOMER SEGMENTATION BUSINESS CASE

BY OPTIMADATACONSULTING



Company Consultants

Ana Carolina Ottavi, Carolina Bezerra, Duarte Girão, João Pólvora e Luca Loureiro

BUSINESS SITUATION

- **Hotel H** belongs to an Hotel Chain C that operated 4 hotels until 2015;
- It is located in **Lisbon, Portugal**;
- There is an expansion plan with new acquisitions in mind;
- There is a **new marketing manager** who discredits the current marketing segmentation used;

KEY PROBLEMS

- The current marketing segmentation only reflects **one** customer characteristic, **sales origin**;
- Do not incorporate **behavioral, demographic** and **geographic characteristics**;
- Without a **sophisticated marketing strategy**, it becomes more difficult to:
 - Retain the actual customers
 - Capture new ones;

DEFINITION OF BUSINESS GOALS

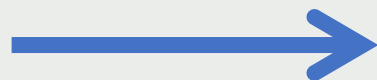
- Identify the **main characteristics** and **behaviours** to include in customers segmentation;
- Provide a **new segmentation profile solution**;
- Suggest business measures to **find** and **retain customers**;
- Provide **updated data periodically** to the marketing department;

Data Understanding

Categorical



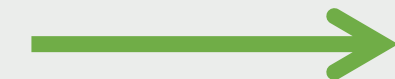
Numerical



Categorical

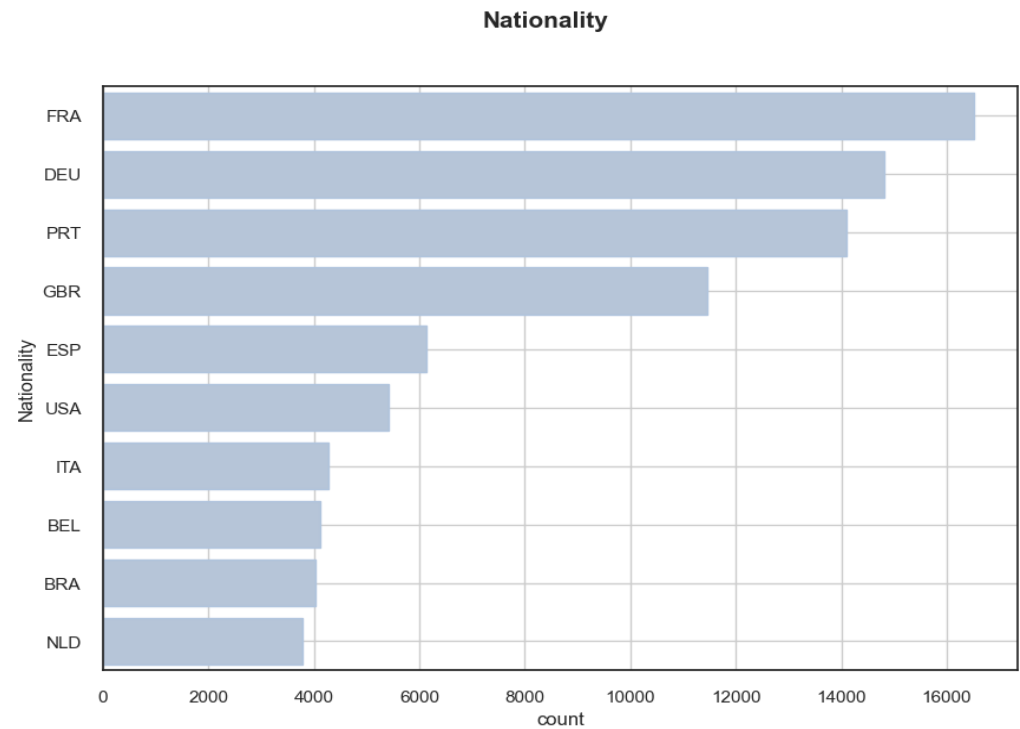


Special
Requests

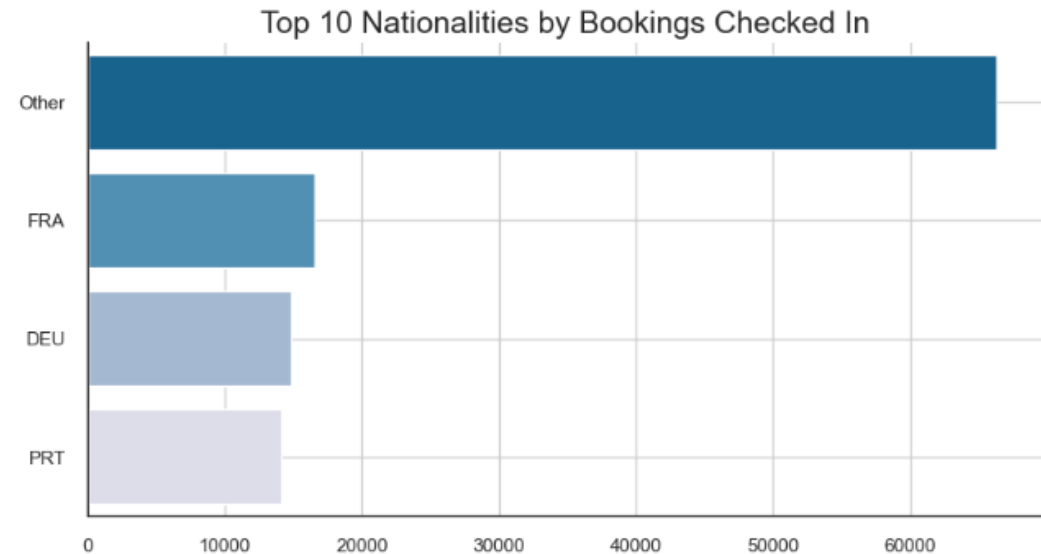


	count	unique		top	freq	mean	std	min	25%	50%
ID	111733.0	NaN		NaN	NaN	55867.0	32254.683151	1.0	27934.0	55867.0
Nationality	111733	199		FRA	16516	NaN	NaN	NaN	NaN	NaN
Age	107561.0	NaN		NaN	NaN	45.639191	17.244952	-10.0	33.0	47.0
DaysSinceCreation	111733.0	NaN		NaN	NaN	595.026599	374.657382	36.0	288.0	522.0
NameHash	111733	107584	0x15A713CE687991691A18F6CDC56ABE24979C73CF5D51...		75	NaN	NaN	NaN	NaN	NaN
DocIDHash	110732	103480	0xE3B0C44298FC1C149AFBF4C8996FB92427AE41E4649B...		3032	NaN	NaN	NaN	NaN	NaN
AverageLeadTime	111733.0	NaN		NaN	NaN	60.833147	85.11532	-1.0	0.0	21.0
LodgingRevenue	111733.0	NaN		NaN	NaN	283.851283	379.131556	0.0	0.0	208.0
OtherRevenue	111733.0	NaN		NaN	NaN	64.682802	123.580715	0.0	0.0	31.0
BookingsCanceled	111733.0	NaN		NaN	NaN	0.002282	0.080631	0.0	0.0	0.0
BookingsNo Showed	111733.0	NaN		NaN	NaN	0.0006	0.028217	0.0	0.0	0.0
BookingsCheckedIn	111733.0	NaN		NaN	NaN	0.737607	0.730889	0.0	0.0	1.0
PersonsNights	111733.0	NaN		NaN	NaN	4.328318	4.630739	0.0	0.0	4.0
RoomNights	111733.0	NaN		NaN	NaN	2.203825	2.301637	0.0	0.0	2.0
DistributionChannel	111733	4	Travel Agent/Operator		91019	NaN	NaN	NaN	NaN	NaN
MarketSegment	111733	7	Other		63680	NaN	NaN	NaN	NaN	NaN
SRHighFloor	111733.0	NaN		NaN	NaN	0.042512	0.201755	0.0	0.0	0.0
SRLowFloor	111733.0	NaN		NaN	NaN	0.001307	0.036125	0.0	0.0	0.0
SRAccessibleRoom	111733.0	NaN		NaN	NaN	0.000224	0.014957	0.0	0.0	0.0
SRMediumFloor	111733.0	NaN		NaN	NaN	0.00077	0.027733	0.0	0.0	0.0
SRBathtub	111733.0	NaN		NaN	NaN	0.003132	0.055881	0.0	0.0	0.0
SRShower	111733.0	NaN		NaN	NaN	0.001629	0.040327	0.0	0.0	0.0
SRCrib	111733.0	NaN		NaN	NaN	0.016181	0.126173	0.0	0.0	0.0
SRKing SizeBed	111733.0	NaN		NaN	NaN	0.363268	0.480943	0.0	0.0	0.0
SRTwinBed	111733.0	NaN		NaN	NaN	0.156811	0.363624	0.0	0.0	0.0
SRNearElevator	111733.0	NaN		NaN	NaN	0.000331	0.018195	0.0	0.0	0.0
SRAwayFromElevator	111733.0	NaN		NaN	NaN	0.003598	0.059874	0.0	0.0	0.0
SRNoAlcoholInMiniBar	111733.0	NaN		NaN	NaN	0.000197	0.014031	0.0	0.0	0.0
SRQuietRoom	111733.0	NaN		NaN	NaN	0.087718	0.282886	0.0	0.0	0.0

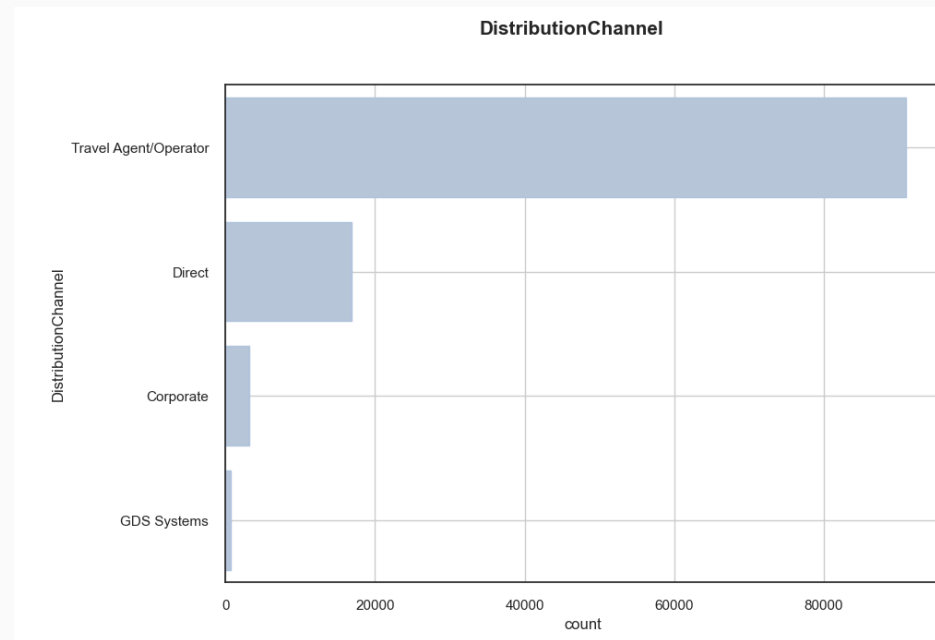
Where are our
customers from?



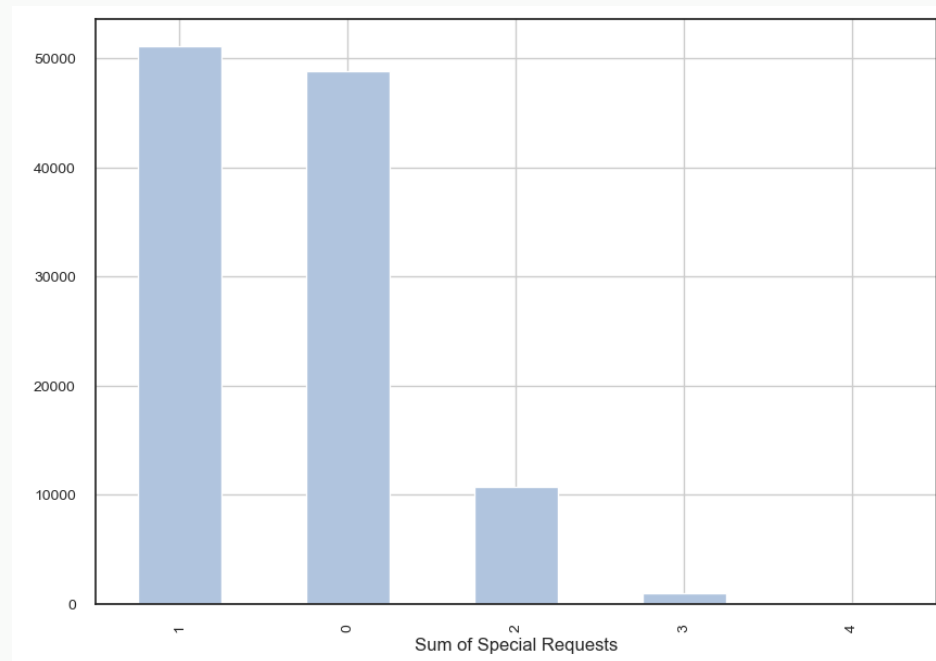
A new look on
nationalities



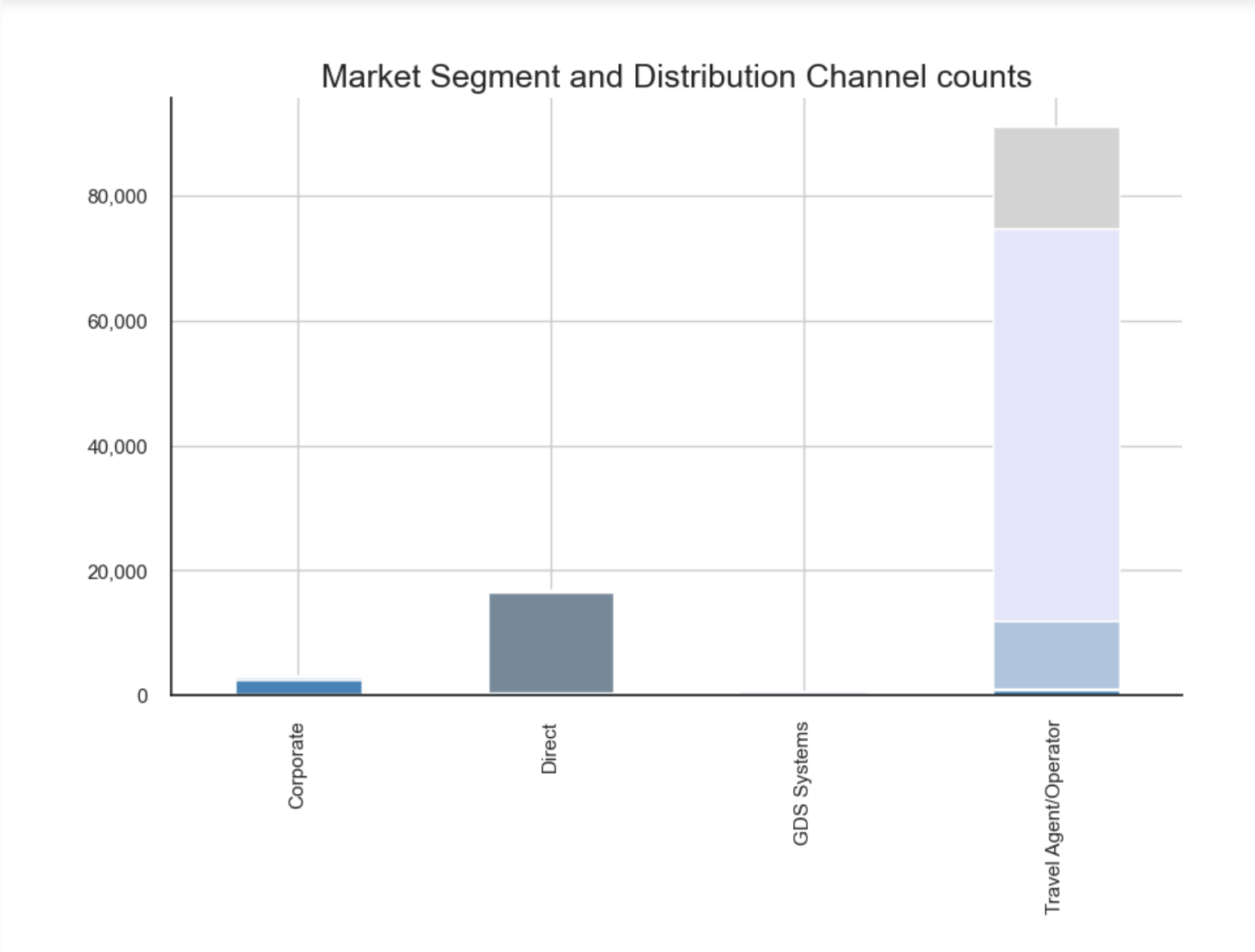
**What's the main
distribution
channel?**



**Do our costumers
request many
special requests?**



**Does the Distribution
Channel corresponds
to the actual Market
Segment?**



Data Preparation

Selecting the main
features

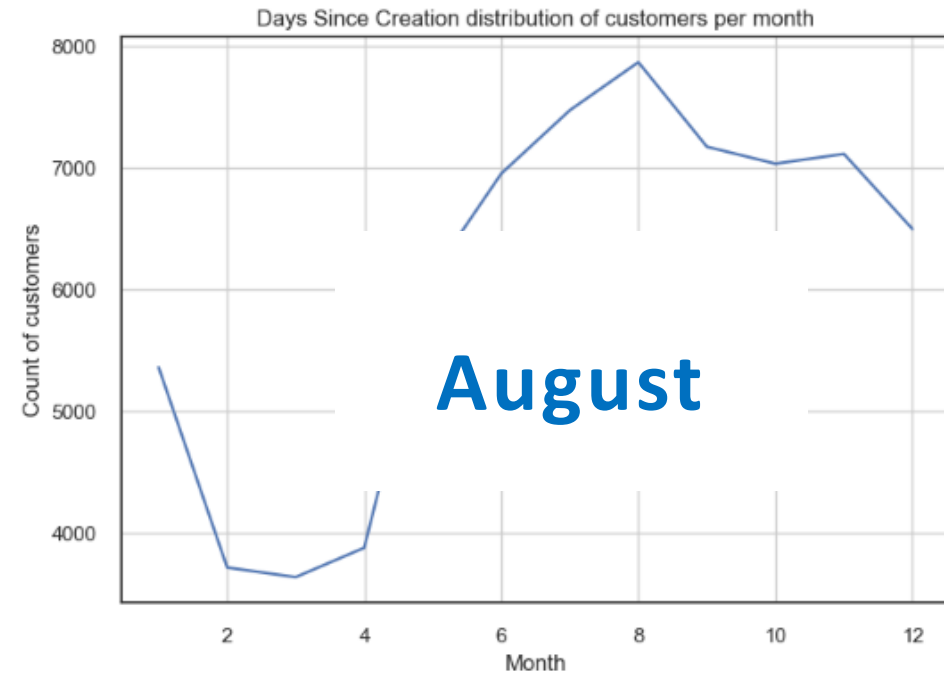
- ID
 - NameHash
 - DocIDHash
 - ForceCreati
 - Cancr
 -
 - Bo
 - Mar
 - SRHigh
 - SRLo
 - SR
 - Fl
 - b
 - ver
 - SRCrib
 - SRNearElevator
 - SRAwayFromElevator
 - SRNoAlcoholInMiniBar
 - RoomNights
- 

Selecting the main
features

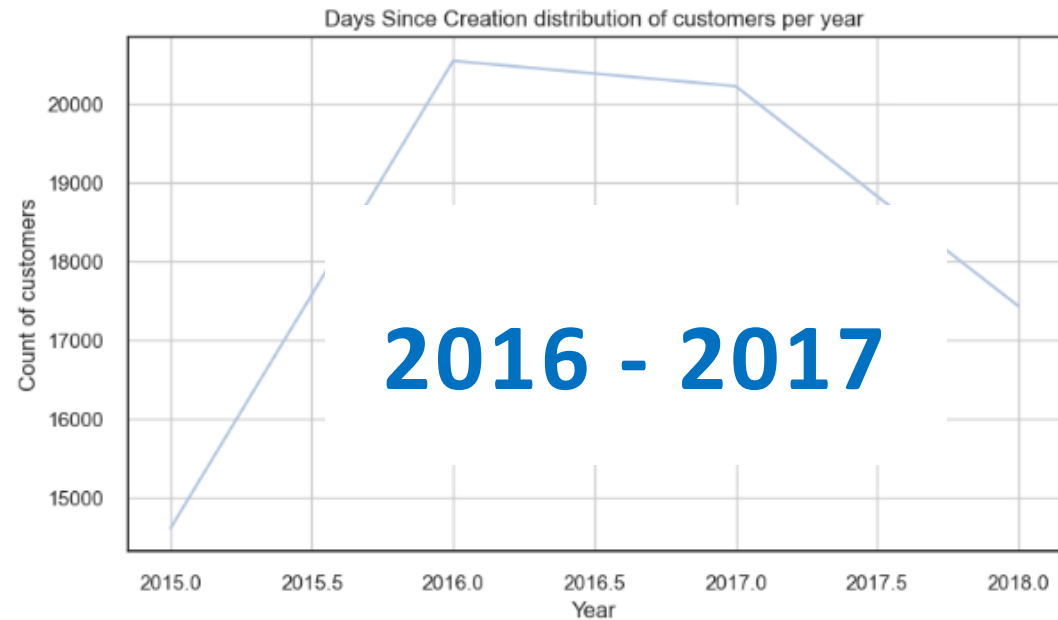
- **Age_gap**
- **Total Revenue** = LodgingRevenue + OtherRevenue
- **AvgRevenueperYear** = TotalRevenue * 365 / DaysSinceCreation
- **Revenueperbooking** = TotalRevenue / BookingsCheckedIn




When does our
customers prefer
to visit our
customers?



In which year did
we acquired most
of our customers?




How is
AvgRevenueperYear
and
Revenueperbooking
distributed per **Age**
category?




Age_gap	AvgRevenueperYear	Revenueperbooking
20-30	6.016751e+06	4.215943e+06
31-40	6.824128e+06	6.186154e+06
41-50	1.017139e+07	9.077642e+06
51-60	8.835185e+06	8.172351e+06
60+	7.661190e+06	6.609524e+06
<20	1.865500e+06	1.209239e+06

How is
DistributionChannel
distributed per **Age**
category?



DistributionChannel	Corporate	Direct	GDS Systems	Travel Agent/Operator
Age_gap				
20-30	173	1071	56	7107
31-40	392	1995	144	11172
41-50	798	3067	200	14810
51-60	430	2101	114	13483
60+	260	1396	23	12545
<20	8	285	1	1181

How are
DaysSinceCreation
and **AverageLeadTime**
distributed per **Age**
category?



	DaysSinceCreation	AverageLeadTime
Age_gap		
20-30	5286148	597080
31-40	9357763	918271
41-50	13527295	1416316
51-60	11382205	1403641
60+	9892178	1620683
<20	910801	148616

Final variables used
for segmentation



- : Nationality_PRT int64
- Nationality_FRA int64
- Nationality_Other int64
- Nationality_DEU int64
- AverageLeadTime int64
- DistributionChannel_Corporate int64
- DistributionChannel_Travel Agent/Operator int64
- DistributionChannel_Direct int64
- DistributionChannel_GDS Systems int64
- SRKingSizeBed int64
- SRTwinBed int64
- SRQuietRoom int64
- AvgRevenueperYear float64
- Revenueperbooking float64
- Age_gap_51-60 int64
- Age_gap_41-50 int64
- Age_gap_60+ int64
- Age_gap_31-40 int64
- Age_gap_20-30 int64
- Age_gap_<20 int64
- dtype: object

MODELING

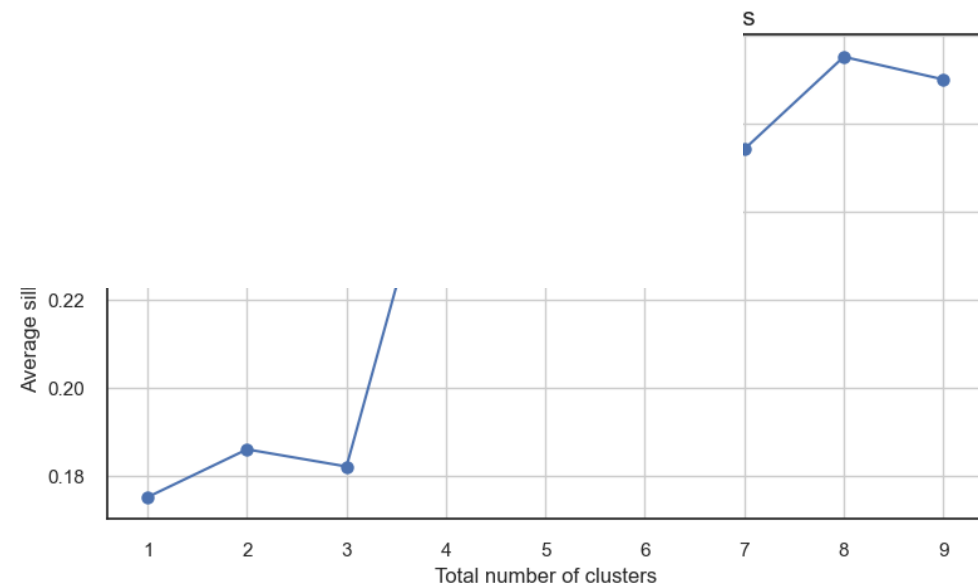
KMeans

Inertia



$K = 5$

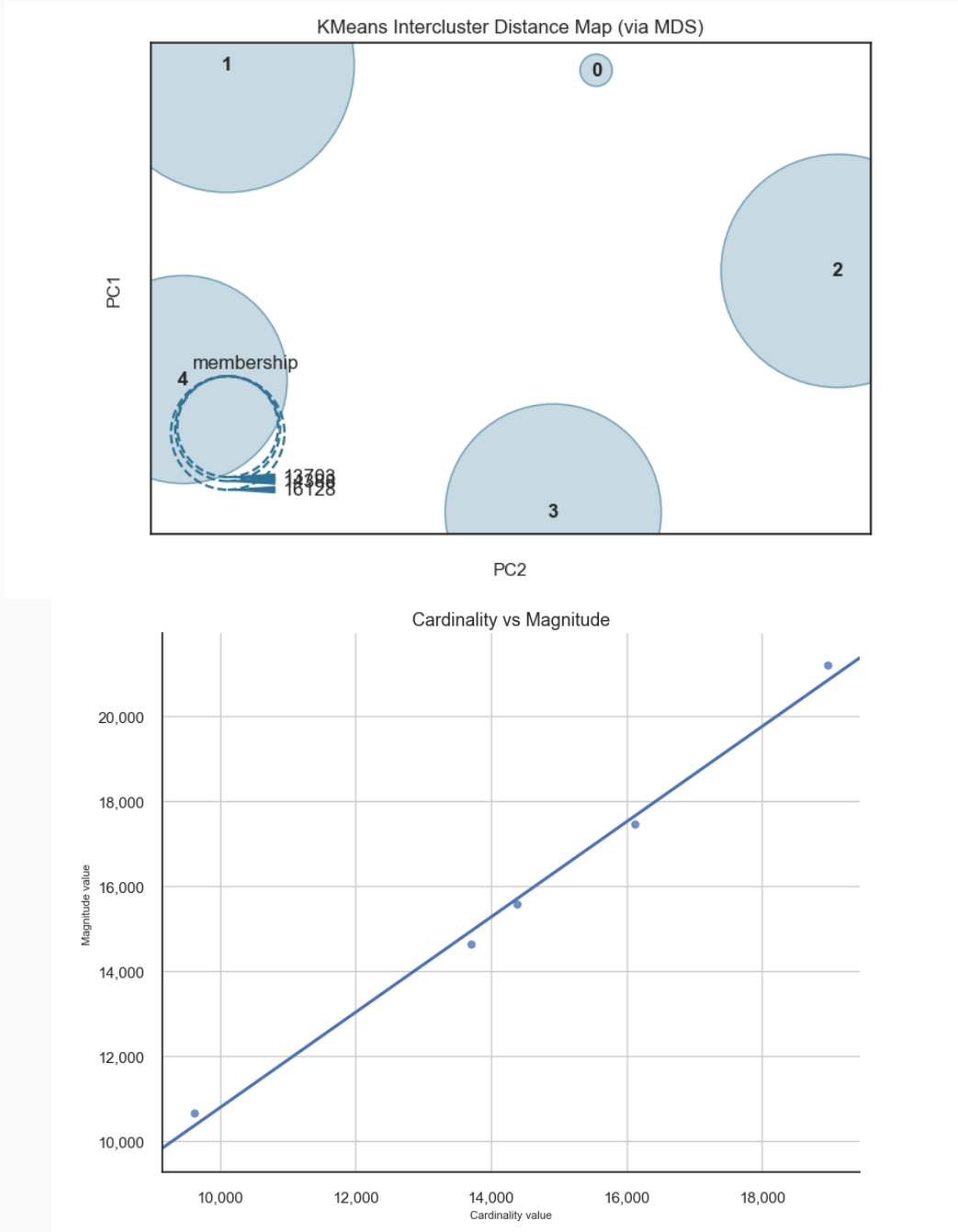
Silhouette



Taking a visual look at
the **clusters**
distribution



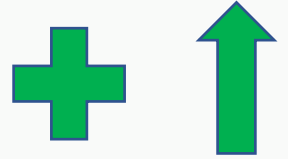
Cardinality vs
Magnitude have a
POSITIVE
CORRELATION



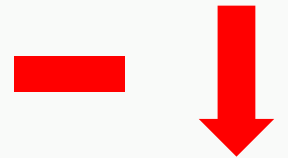
Cluster Characterization & Marketing Strategies Approaches

Features:

- *Age gap: 20-30*
- *Distribution Channel: Travel Agent/Operator*
- *French Nationality*
- *AvgRevenueperYear*
- *Revenueperbooking.*



- *AverageLeadTime, SRQuietRoom*



Marketing Strategies:

- Use social media influencers
- Focus on younger people with **acquisitive power**
- Focus on **French market**
- Promote mainly the high **luxurious products** and **services**



Cluster 0

“Young Customers with
more Wealth”



Cluster 1
“Portuguese Clients”



Features:

- *Age gap: 41-50*
- *Portuguese Nationality*
- *Distribution Channel: Corporate and Direct*



Marketing Strategies:

- Create a media plan mainly by television and radio
- Create a **loyalty program**, recognizing their preferences through special discounts and offers
- Focus on **Portuguese market**

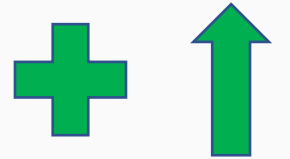


Cluster 2
“Potential and Mature
Clients”



Features:

- **Age gap: 51-60**
- Highest value of **nationality** is **Others**
- **Distribution Channel: GDS Systems**
- **SRQuietRoom**



Marketing Strategies:

- Create a **loyalty campaign** to retain, with a points system, providing **extra bookings for usage**
- Create parallel discounts with **dinners** and **outdoor activities**, for this age category

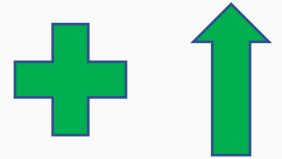


Cluster 3
“Advanced Booking
Customers”



Features:

- *Age gap: 60+*
- *German Nationality*
- *AverageLeadTime*
- *SRTwinBed*
- *Distribution Channel: Travel Agent/Operator*



Marketing Strategies:

- Create a partnership with travel agencies to organize trips with tailored experiences and activities;
- Design a campaign to promote opportunities to travel with buddies



Cluster 4

“Low cost travelers”

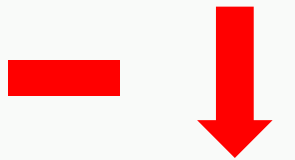


Features:

- *Age gap: 31-40*
- Highest value of *nationality* is *Others*
- *Distribution Channel: Direct (the second highest value)*
- *SRKingSizeBed*



- *AvgRevenueperYear*
- *Revenueperbooking*



Marketing Strategies:

- Promote offers with discounts to in the less **busy period** of the year throughout the **main distribution channels**;
- Create a tailored offer for middle age and sensible to prices customers.

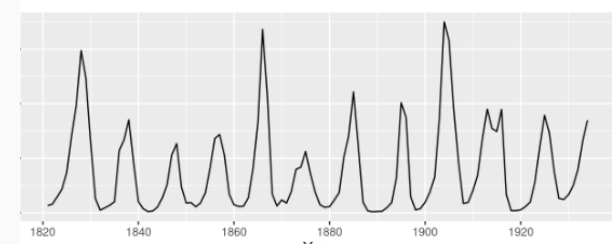


Deployment and Maintenance Plans

**Execute the
Model Quarterly**



- *To capture **seasonal behaviours***
- *Results can be accessed by **Google Sheets***
- ***Easier communication** between departments*



Google Sheets

Internal Chatbot



- *Developed using a **Telegram bot** used by a **decision maker***
- *Can provide access to the main features already predefined*
- ***Fast implementation, accessible** and with **low costs***



Final Report



- *All data information can be revisited*
- *New results and insights are generated*
- ***Useful for decision makers** to take more accurate and conscious financial decisions*



Conclusions

Most **profitable**
clusters



Cluster 0
*"Young Customers
with more Wealth"*

Cluster 2
*"Potential and
Mature Clients"*

Cluster 3
*"Advanced Booking
Customers"*

More **frequent**
clusters



Cluster 1
"Portuguese Clients"

Final suggestion



- *Long-term main suggestion: Cluster 1*
- *Also invest in **profitable** clusters 0, 2 and 3*
- *Cluster 4 has low profitables and low frequency. **Not a priority!***



Thank you!