githubPROJECT NUMBER: 27

TEAM: 108

Ensuring product availability, avoiding the out of stock

Business problem

Ensuring product availability on the shelf is essential for today's retail sector, so much so that it is considered a measure of retail performance. Retail is a highly competitive industry, and it is imperative to ensure that products are on the shelf when the customer is buying them. Consumers still prefer to shop in retail stores, but low shelf availability and a high number of out-of-stock events can affect their experience. Customers go to the retail outlet to find the products they want; if they can't find what they are looking for on the shelves, they will choose a substitute product, leave empty-handed, or go to a competitor to satisfy their needs. In each case, the retailer, the manufacturer and the customer lose.

Business Impact

Out-of-stocks is a major problem in retailing, as it leads to lost sales and reduced customer loyalty, because the term "out-of-stocks" is used to describe a situation where a consumer cannot find the product on the shelf at the time, he or she wants to buy it. In addition, the shortage of products on the shelves may affect consumers' choices in the future.

Data

The following table has the name and description of the five datasets received by Eficacia:

TABLE	DESCRIPTION	COLUMNS	COLUMN'S DESCRIPTION
		CK.Key2	Master registry key
		Agotados. Agotado	String (binary) if the product is out of stock (1) or not out of stock (0)
		Agotados.MarcasFoco	String, Brand associated with the product
AGOTADOS	Detail of out-of- stock products	Agotados.IDProducto	String, Code associated with the product
	Stock products	Agotados.IDSucursal	String, Code associated with the point of sale
		Agotados.IDPersona	String, Code associated with the person who took the record
		Agotados.Fecha	Quantitative, Date on which the record was taken

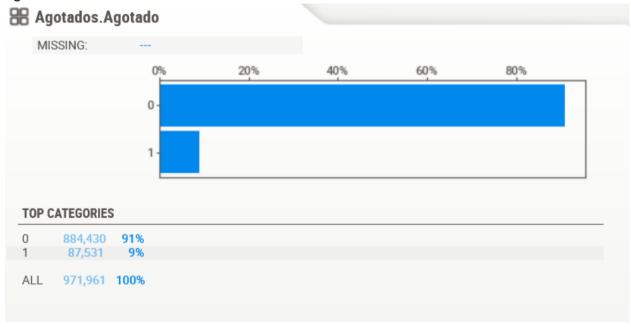
		Agotados.IDRegistro	String, Registration code
		Producto.ldProducto	String, Product Code
		Producto.Linea	String, Line to which the product belongs
PRODUCTOS	Products Detail	Producto.Categoria	String, Category to which the product belongs within a line
		Producto.ldMarca	String, Code associated with the product brand
		Exhibiciones.Id_Registro	String, Registration code
		Exhibiciones.Id_Pdv	String, Point of sale code
		Exhibiciones.Id_Usuario	String, Code associated with the person who took the record
	Datail aftha	Exhibiciones.Fecha	Quantitative, Date on which the record was taken
EXHIBICIONES	Detail of the exhibitions carried out at	Exhibiciones.TipoExhibicion	String, Type or main classification of the exhibits
LATIBLEIONES	the points of sale	Exhibiciones.SubTipoExhibicion	String, Secondary classification of exhibits
	Said	Exhibiciones.Categoria	String, Category to which the exhibition belongs
		Exhibiciones.MarcaFoco	String, Brand associated with the product
		Exhibiciones.Clasificacion	String, Tertiary classification of exhibits
		CK.Key2	String, Primary registry key
		Precios.IDSucursal	String, Code of the point of sale or branch
	Detail of the	Precios.MarcasFoco	String, Brand associated with the product
PRECIOS	prices of the different	Precios.Precio	Integer, Price of the product at the time of registration
	products	Precios.IDPersona	String, Code associated with the person who took the record
		Precios.Fecha	Quantitative, Date on which the record was taken
		Precios.IDProducto	String, Product Code
		SK.PuntoVenta	String, Code of the point of sale or branch
	Detail of the	PuntoVenta.NombreCiudad	String, City where the point of sale is located
PUNTOS DE VENTA	points of sale where the products are	PuntoVenta.NombreDepto	String, Department in which the point of sale is located
	found	PuntoVenta.Canal	String, Sales channel to which the point of sale belongs
		PuntoVenta.Regional	String, Regional of the company where the point of sale is located

The information in these datasets, although it is uniform, does not have null values and tries to be detailed in certain aspects, is insufficient to carry out a more in-depth analysis due to the lack of total stock values in order to make a comparison with respect to the movement of inventories and in turn to be able to compare them with a table that relates the sales data; if we had this data, we could make a more useful approximation to a model of inventory movement and management of products without stock.

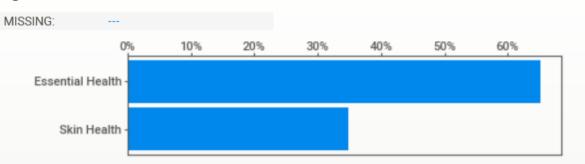
Basic EDA

EDA básico

Se recibieron 5 data set en formato csv, se renombraron de la siguiente forma modelos_agotados, modelos_agotados_exhibiciones, modelos_agotados_precios, modelos_agotados_productos, modelos_agotados_pventas agotados



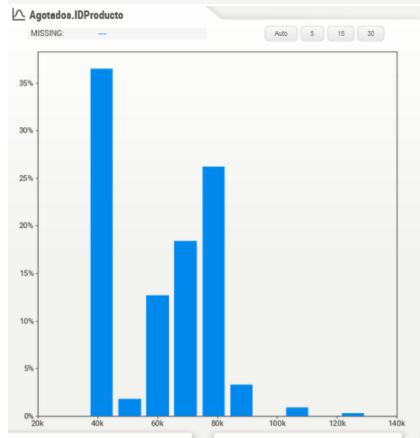
□ Agotados.MarcasFoco



TOP CATEGORIES

Essential Health	633,590	65%
Skin Health	338,371	35%

ALL 971,961 100%



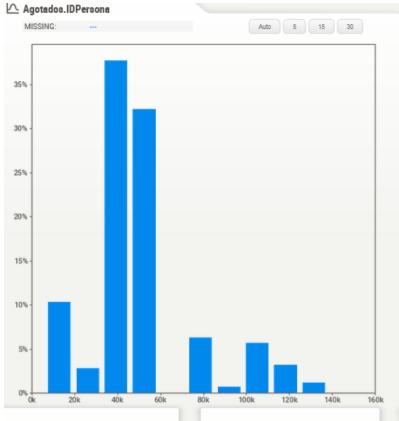
Agotados.IDRegistro	0.08
Agotados.IDSucursal	0.02
Agotados.IDPersona	0.02
CATEGORICAL ASSO	CIATIONS

MOST F	REQUEN	T VALUES	
36876	15,462	1.6%	
36841	15,025	1.5%	
36842	15,009	1.5%	
36875	14,983	1.5%	
37081	14,736	1.5%	
63463	14,700	1.5%	
79518	14,423	1.5%	
36766	14,393	1.5%	
63530	14,301	1.5%	
63502	14,280	1.5%	
79521	14,200	1.5%	
79516	14,082	1.4%	
36880	14,080	1.4%	
37078	13,883	1.4%	
70527	12 292	1.4%	

SMALL	EST VALU	JES	
36766	14,393	1.5%	
36767	11,246	1.2%	
36770	1,898	0.2%	
36775	332	<0.1%	
36780	1,845	0.2%	
36790	7,092	0.7%	
36793	2,361	0.2%	
36801	10,312	1.1%	
36807	1,557	0.2%	
36816	7,271	0.7%	
36830	449	< 0.1%	
36831	482	< 0.1%	
36838	249	< 0.1%	
36841	15,025	1.5%	
36842	15,009	1.5%	

130014	2	<0.1%	
130013	3	<0.1%	
130012	21	<0.1%	
130010	15	<0.1%	
130009	40	<0.1%	
130005	16	< 0.1%	
130004	87	< 0.1%	
130003	113	<0.1%	
130002	119	<0.1%	
130000	78	<0.1%	
129998	4	<0.1%	
129997	24	< 0.1%	
128861	68	<0.1%	
128860	87	<0.1%	
128857	53	< 0.1%	





Agotados.IDRegistro	0.37
Agotados.IDSucursal	0.03
Agotados.IDProducto	0.02
	IATION
CATEGORICAL ASSOC	
	0.04 0.03

MOST FREQUENT VALUES

6418	17,074	1.8%	
11526	16,611	1.7%	
34611	16,389	1.7%	
6435	15,288	1.6%	
51955	15,227	1.6%	
24469	14,412	1.5%	
34641	12,818	1.3%	
34620	12,664	1.3%	
10622	11,826	1.2%	
34606	11,745	1.2%	
40200	11,569	1.2%	
38789	11,458	1.2%	
34926	11,406	1.2%	
51959	11,194	1.2%	
51057	10.988	1.1%	

SMALLEST VALUES

			_
6408	1,463	0.2%	
6409	202	<0.1%	
6411	18	<0.1%	
6412	929	<0.1%	
6414	5,902	0.6%	
6418	17,074	1.8%	
6421	57	<0.1%	
6424	1,092	0.1%	
6435	15,288	1.6%	
6440	6,200	0.6%	
6903	1,299	0.1%	
7024	8,378	0.9%	
7863	3,976	0.4%	
10622	11,826	1.2%	
11283	10,124	1.0%	

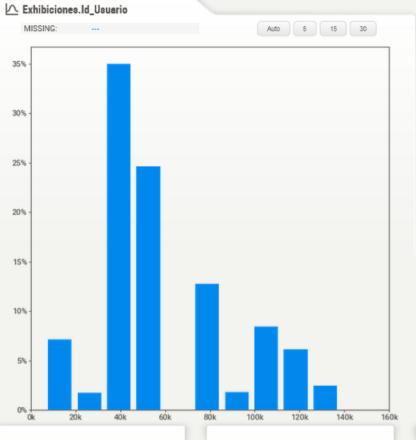
LABOTO	- 3/41 117	-0	
LARGES ¹	VALUE	:5	
138070	126	<0.1%	
138069	44	<0.1%	
137856	103	<0.1%	
137855	82	<0.1%	
137854	217	<0.1%	
137853	27	<0.1%	
137852	65	<0.1%	
136055	193	<0.1%	
135744	281	<0.1%	
134652	94	<0.1%	
130416	790	<0.1%	
130414	429	<0.1%	
126471	507	<0.1%	
126470	1,511	0.2%	
126469	1,501	0.2%	





Exhibiciones





Exhibiciones.ld_Pdv	0.00
ATEGORICAL ASSOCIA	TIONS
ORRELATION RATIO, 0 to 1)	
Exhibiciones.SubTipoExhib.	_0.13
Exhibiciones.TipoExhibicion	0.09
Exhibiciones.Clasificacion	0.07
Exhibiciones.Categoria	0.05
Exhibiciones.MarcaFoco	0.02

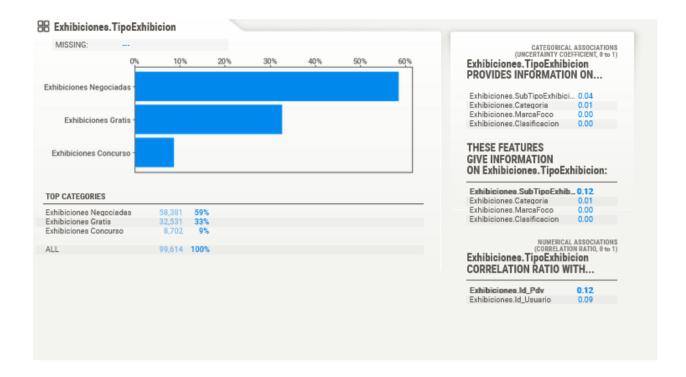
MOST FR	EQUEN	T VALUES	
43076	3,226	3.2%	
45401	2,707	2.7%	
47389	2,657	2.7%	
78636	2,286	2.3%	
42958	2,126	2.1%	
51057	1,990	2.0%	
34641	1,869	1.9%	
11526	1,641	1.6%	
115689	1,568	1.6%	
53332	1,439	1.4%	
34638	1,382	1.4%	
38789	1,285	1.3%	
24469	1,282	1.3%	
82231	1,280	1.3%	
51955	1,276	1.3%	

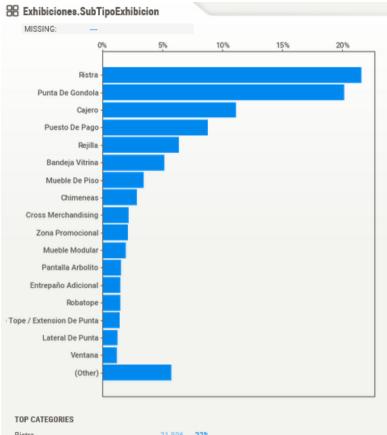
,	SMALLEST	T VALU	JES	
	6409	18	<0.1%	
	6411	238	0.2%	
	6412	218	0.2%	
	6414	229	0.2%	
	6418	1,271	1.3%	
	6421	53	<0.1%	
	6424	144	0.1%	
	6435	924	0.9%	
	6440	136	0.1%	
	6903	39	<0.1%	
	7024	263	0.3%	
	7863	210	0.2%	
	10622	1,051	1.1%	
	11283	635	0.6%	
	11526	1,641	1.6%	

LARGEST	VALUE	S	
138070	47	<0.1%	
138069	14	<0.1%	
137856	83	<0.1%	
137855	66	<0.1%	
137854	102	0.1%	
137853	9	<0.1%	
137852	19	<0.1%	
136055	12	<0.1%	
135744	22	<0.1%	
134652	65	<0.1%	
130416	336	0.3%	
130414	55	<0.1%	
126471	166	0.2%	
126470	138	0.1%	
126469	81	<0.1%	

Exhibiciones.Fecha

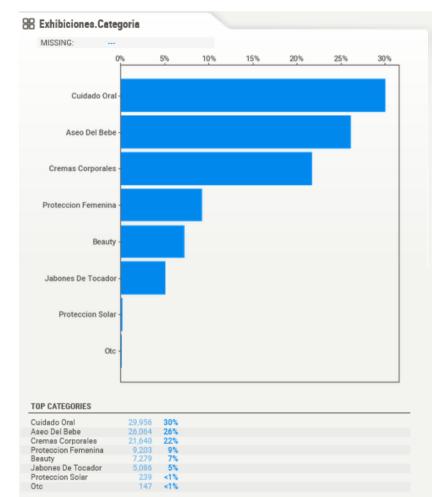
	MISS	ING:	
	794	<1%	2021-03-03
	769	<1%	2021-03-13
	755	<1%	2021-03-02
	731	<1%	2021-04-08
	716	<1%	2021-04-07
	705	<1%	2021-03-12
	664	<1%	2021-03-04 2021-04-05
	639 636	<1%	2021-04-06
	613	<1%	2021-04-00
	602	<1%	2021-05-19
	556	<1%	2021-03-19
	552	<1%	2021-03-06
	546	<1%	2021-05-18
	543	<1%	2021-04-09
	542	<1%	2021-02-02
	541	<1%	2021-04-10
	541	<1%	2021-02-24
	537	<1%	2021-03-10
	531	<1%	2021-03-01
	499	<1%	2020-11-06
	476	<1%	2020-10-06
	473	<1%	2020-11-05
	471	<1%	2021-03-09
	468	<1%	2020-11-04
	459	<1%	2021-03-11
	448	<1%	2020-11-07
	444	<1%	2021-03-19
	441	<1%	2020-10-07
	438	<1%	2021-04-20
	433	<1%	2020-12-03
	429	<1%	2021-03-18
	426	<1%	2021-03-16
	408	<1%	2022-03-02
	407	<1%	2021-03-08
	405	<1%	2021-04-26
	404	<1% <1%	2020-12-02
	397	<1%	2021-12-07 2021-04-12
	397	<1%	2021-03-24
	389 389	<1%	2021-03-24
	388	<1%	2021-05-20
	379	<1%	2021-04-23
	378	<1%	2021-12-14
	376	<1%	2021-04-13
	371	<1%	2020-09-05
	371	<1%	2020-09-03
	366	<1%	2020-10-01
	364	<1%	2021-12-09
	364	<1%	2020-09-02
	362	<1%	2022-01-12
	362	<1%	2020-12-05
	361	<1%	2021-12-02
	356	<1%	2021-04-15
	354	<1%	2021-02-03
	348	<1%	2021-11-05
	346	<1%	2020-12-11
	344	<1%	2020-12-09
	344	<1%	2020-10-02
	343	<1%	2021-12-10
71	,123	71%	(Other)





CATEGORICAL ASSOCIATIONS (UNCERTAINTY COEFFICIENT, 0 to 1) Exhibiciones. SubTipoExhibicion PROVIDES INFORMATION ON... Exhibiciones. Clasificacion 0.12 Exhibiciones. Clasificacion 0.12 Exhibiciones. MarcaFoco 0.06 THESE FEATURES GIVE INFORMATION ON Exhibiciones. Clasificacion 0.32 Exhibiciones. Clasificacion 0.32 Exhibiciones. Clasificacion 0.07 Exhibiciones. TipoExhibicion 0.04 Exhibiciones. MarcaFoco 0.02 NUMERICAL ASSOCIATIONS (CORRELATION PATIO, 0 to 1) Exhibiciones. SubTipoExhibicion CORRELATION RATIO WITH... Exhibiciones.Id_Pdv 0.39 Exhibiciones.Id_Pdv 0.39 Exhibiciones.Id_Pdv 0.39 Exhibiciones.Id_Pdv 0.39 Exhibiciones.Id_Usuario 0.13

Ristra	21,506	22%	
Punta De Gondola	20,088	20%	
Cajero	11,109	11%	
Puesto De Pago	8,758	9%	
Rejilla	6,371	6%	
Bandeja Vitrina	5,135	5%	
Mueble De Piso	3,414	3%	
Chimeneas	2,857	3%	
Cross Merchandising	2,174	2%	
Zona Promocional	2,087	2%	
Mueble Modular	1,935	2%	
Pantalla Arbolito	1,566	2%	
Entrepaño Adicional	1,505	2%	
Robatope	1,479	1%	
Tope De Tope / Extension De Punta	1,431	1%	
Lateral De Punta	1,277	1%	
Ventana	1,215	1%	
(Other)	5,707	6%	
ALL	99,614	100%	

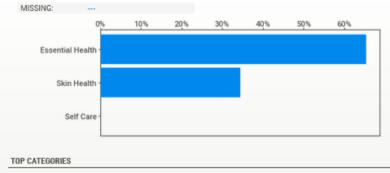


CATEGORICAL ASSOCIATIONS (UNICERTAINTY COEFFICIENT, 0 to 1) Exhibiciones. Categoria PROVIDES INFORMATION ON	ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria	N Exhibiciones. Categoria: Exhibiciones. MarcaFoco 0.40 Exhibiciones. SubTipoExhib. 0.11 Exhibiciones. Clasificacion 0.01 Exhibiciones. TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) EXHIBICIONES. Categoria CORRELATION RATIO WITH	Exhibiciones. Categoria PROVIDES INFORMATIO Exhibiciones. MarcaFoco Exhibiciones. Clasificacion Exhibiciones. Clasificacion Exhibiciones. TipoExhibicion Exhibiciones. TipoExhibicion	000 ON 1.00
	Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib. 0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS	Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 CHESE FEATURES EVENTORMATION EXHIBICIONES.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib_0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 EXHIBICIONES.TIPOEXHIBICION RATIO, 0 to 1) EXHIBICIONES.CATEGORIA EXHIBICION		
	THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib. 0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria	CHESE FEATURES SIVE INFORMATION ON Exhibiciones. Categoria: Exhibiciones. MarcaFoco Exhibiciones. SubTipoExhib. O.01 Exhibiciones. Clasificacion O.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones. Categoria CORRELATION RATIO WITH Exhibiciones. Id_Pdv O.23		
Exhibiciones.SubTipoExhibici 0.07	GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 6 to 1) Exhibiciones.Categoria	Exhibiciones.MarcaFoco 0.40 Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib. 0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria CORRELATION RATIO WITH Exhibiciones.Id_Pdv 0.23	Exhibiciones.TipoExhibicion	0.01
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION	Exhibiciones Clasificacion Exhibiciones TipoExhibicion NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones Categoria	Exhibiciones.Clasificacion 0.01 Exhibicionee.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria CORRELATION RATIO WITH Exhibiciones.Id_Pdv 0.23	ON Exhibiciones.Categ	oria:
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40	Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria	NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria CORRELATION RATIO WITH Exhibiciones.Id_Pdv 0.23	Exhibiciones.MarcaFoco	0.40
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib_0.11	NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones. Categoria	NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) ixhibiciones.Categoria CORRELATION RATIO WITH Exhibiciones.Id_Pdv 0.23	Exhibiciones.MarcaFoco Exhibiciones.SubTipoExhi	0.40 b0.11
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib 0.11 Exhibiciones.Clasificacion 0.01			Exhibiciones.MarcaFoco Exhibiciones.SubTipoExhi Exhibiciones.Clasificacion	0.40 b_0.11 0.01
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib 0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria CORRELATION RATIO WITH		EXTINUICIONES.IU_USUBIIO U.U3	Exhibiciones.MarcaFoco Exhibiciones.SubTipoExhi Exhibiciones.Clasificacion Exhibiciones.TipoExhibicion NUMERIC. (CORRELAT EXHIBICIONES.Categoria CORRELATION RATIO V	0.40 b_ 0.11 0.01 0.01 0.01 AL ASSOCIATIONS TION RATIO, 0 to 1)
Exhibiciones.SubTipoExhibici 0.07 Exhibiciones.Clasificacion 0.02 Exhibiciones.TipoExhibicion 0.01 THESE FEATURES GIVE INFORMATION ON Exhibiciones.Categoria: Exhibiciones.MarcaFoco 0.40 Exhibiciones.SubTipoExhib 0.11 Exhibiciones.Clasificacion 0.01 Exhibiciones.TipoExhibicion 0.01 NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Exhibiciones.Categoria CORRELATION RATIO WITH Exhibiciones.Id_Pdv 0.23			Exhibiciones.MarcaFoco Exhibiciones.SubTipoExhi Exhibiciones.Clasificacion Exhibiciones.TipoExhibicion NUMERIC. (CORRELAT EXHIBICIONES.Categoria CORRELATION RATIO V Exhibiciones.Id_Pdv	0.40 b0.11 0.01 0.01 0.01 AL ASSOCIATIONS TON RATIO, 0 to 1) VITH 0.23

⊞ Exhibiciones.MarcaFoco

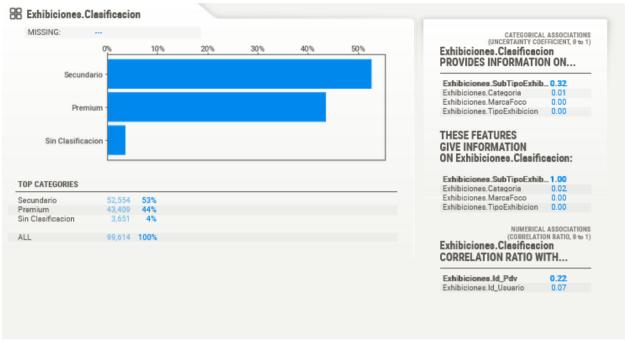
99,614 100%

ALL

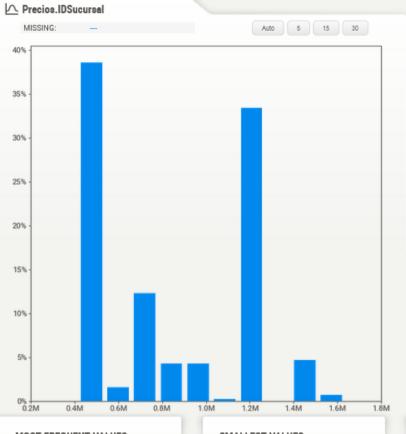


TOP CATEGORIES		
Essential Health	65,223	65%
Skin Health	34,244	34%
Self Care	147	<1%
ALL	99,614	100%

Exhibiciones.Cut Exhibiciones.Sut	0.40
Exhibiciones.Tip	0.02
Exhibiciones.Cla	0.00
Exhibiciones.C	 1.00
Exhibiciones.Sul Exhibiciones.Tip	0.06
Exhibiciones.Cla	0.00



Precios



NUMERICAL ASSOCIATIONS (PEARSON, -1 to 1) **-0.35 -0.15** -0.03 Precios.IDProducto Precios.IDPersona CATEGORICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1) Precios.MarcasFoco

MOST FREQUENT VALUES

1231899	8,916	0.9%	
1231945	8,881	0.9%	
1231942	8,509	0.9%	
427198	8,442	0.8%	
426637	8,381	0.8%	
1232260	8,276	0.8%	
1440741	8,011	0.8%	
1231664	7,770	0.8%	
1231895	7,742	0.8%	
1231908	7,606	0.8%	
427066	7,600	0.8%	
426732	7,406	0.7%	
1231820	7,397	0.7%	
1232259	7,315	0.7%	
427063	7,313	0.7%	

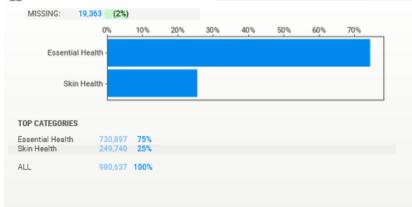
SMALLEST VALUES

416743	1,404	0.1%
416745	937	<0.1%
416746	2,394	0.2%
416750	2,696	0.3%
416751	2,353	0.2%
416755	1,591	0.2%
416799	1,768	0.2%
416801	2,095	0.2%
416910	2,095	0.2%
416914	1,957	0.2%
416928	1,794	0.2%
416979	895	<0.1%
416992	1,571	0.2%
417056	2,325	0.2%
447464	0.400	0.00

LARGEST VALUES

1633895	1,000	0.1%
1599946	2,170	0.2%
1599945	2,183	0.2%
1599944	1,702	0.2%
1566110	3	< 0.1%
1479107	99	<0.1%
1471240	2,685	0.3%
1471237	165	<0.1%
1442145	1,627	0.2%
1440803	1,982	0.2%
1440802	3,221	0.3%
1440801	3,148	0.3%
1440800	2,753	0.3%
1440760	378	<0.1%
1440757	368	< 0.1%

₽ Precios.MarcasFoco



CATEGORICAL ASSOCIATIONS (UNCERTAINTY COEFFICIENT, 0 to 1)

Precios.MarcasFoco PROVIDES INFORMATION ON...

THESE FEATURES **GIVE INFORMATION** ON Precios.MarcasFoco:

NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1)

Precios.MarcasFoco

CORRELATION RATIO WITH...

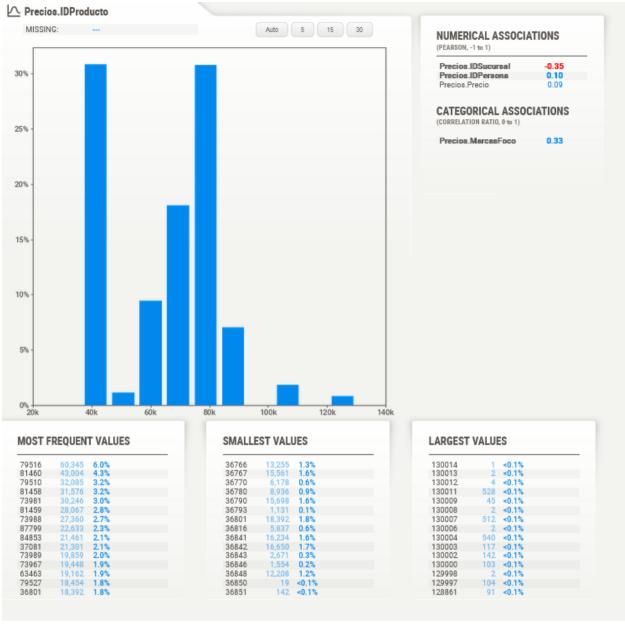
Precios Precio Precios IDProducto	0.69
Precios.IDSucursal	0.33
Precios.IDPersona	0.06



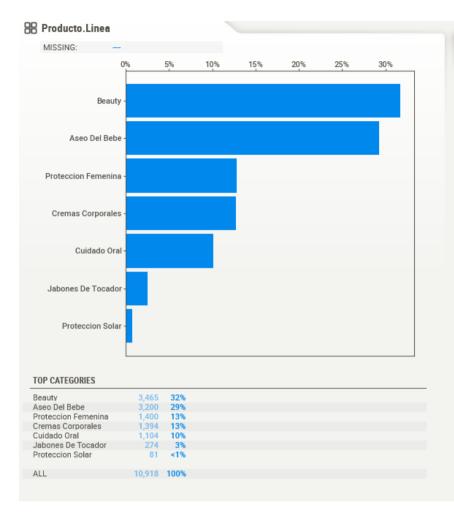


Precios.Fecha

_			
	MISS	ING:	
8	761	<1%	2021-09-09
	482	<1%	2021-12-07
	152	<1%	2022-01-26
	980	<1%	2021-05-18
7	191	<1%	2021-10-08
- 7	,068	<1%	2021-11-10
	,024	<1%	2021-02-02
- 7	,020	<1%	2021-02-03
	,998	<1%	2021-12-14
6	,955	<1%	2021-11-04
	,844	<1%	2021-09-14
	,709	<1%	2021-08-06
	,550	<1%	2021-03-11
6	,520	<1%	2021-11-26
	,485	<1%	2022-03-08
	,398	<1%	2021-02-04
	,332	<1%	2021-10-14
	,190	<1%	2021-04-12 2022-01-12
	,145	<1%	2021-04-09
	,079	<1%	2021-11-09
	.072	<1%	2022-02-25
	.058	<1%	2022-03-02
	,974	<1%	2021-11-08
	891	<1%	2021-03-10
	879	<1%	2021-11-05
	871	<1%	2022-01-13
	.814	<1%	2022-02-07
5	776	<1%	2021-04-06
- 5	,761	<1%	2021-12-15
5	,664	<1%	2021-12-06
	,544	<1%	2022-03-07
	,411	<1%	2021-02-12
	,394	<1%	2021-11-12
	,350	<1%	2021-10-26
	,348	<1%	2022-02-03
	,308	<1%	2021-03-09
	,305	<1%	2021-12-09
	.271 .253	<1% <1%	2021-04-08 2021-07-13
	237	<1%	2021-04-05
	204	<1%	2021-12-13
	141	<1%	2021-12-02
	126	<1%	2022-02-08
	084	<1%	2021-06-11
	.073	<1%	2021-03-08
	055	<1%	2022-01-07
- 5	,011	<1%	2021-11-02
5	,011	<1%	2021-09-08
4	,983	<1%	2021-11-11
	,895	<1%	2021-07-06
	,853	<1%	2021-10-13
	,852	<1%	2021-03-06
	,801	<1%	2021-08-10
	,784	<1%	2022-02-04
	,737	<1%	2021-12-16
	736	<1%	2022-01-25
	710	<1%	2021-03-03
	,708 .704	<1%	2022-02-22
	.335	<1% 65%	2021-10-15 (Other)
040	,008	00/6	(Oulet)



Productos



CATEGORICAL ASSOCIATIONS (UNCERTAINTY COEFFICIENT, 0 to 1)

0.99

Producto.Linea

Producto.Categoria

THESE FEATURES
GIVE INFORMATION

ON Producto.Linea: Producto.Categoria

Producto.ldMarca

PROVIDES INFORMATION ON...

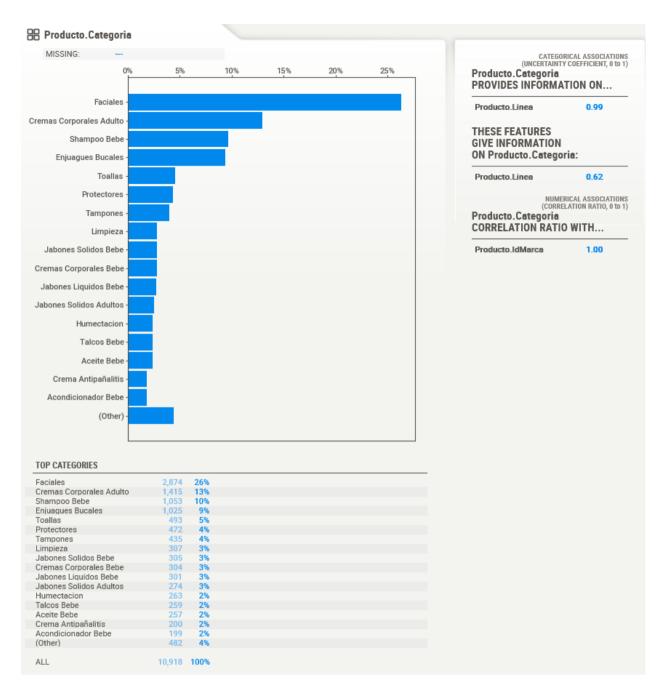
NUMERICAL ASSOCIATIONS (CORRELATION RATIO, 0 to 1)

Producto.Linea

CORDEL ATION:

CORRELATION RATIO WITH...

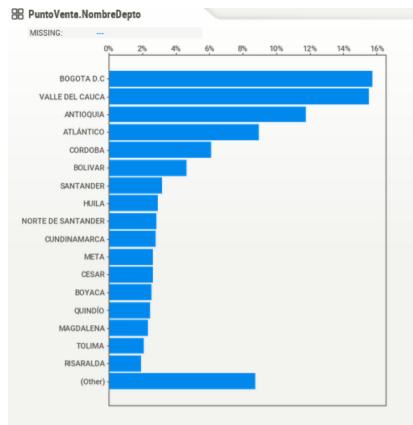




Puntos de venta

PuntoVenta.NombreCiudad

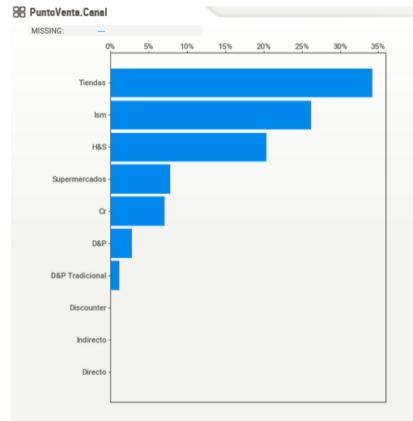
			a.iioiiibi coidada		
	MIS	SING:			
			D		
	,507	16%	Bogotá		
- 1	,196	13%	Santiago De Cali		
	637 589	7% 6%	Medellin Barranguilla		
	378	4%	Cartagena		
	320	3%	Montería		
	220	2%	Villavicencio		
	192	2%	Santa Marta		
	184	2%	Netva		
	181	2%	Valledupar		
	177	2%	Bucaramanga		
	169	2%	Cúcuta		
	167	2%	Armenia		
	142	1%	lbagué		
	135	1%	Soledad		
	130	1%	Popayán		
	114	1%	Tunja		
	111	1%	Pereira		
	108	1%	Rionegro		
	84	<1%	Manizales		
	77	<1%	Ocaña		
	69	<1%	Pasto		
	63	<1%	Pueblo Nuevo		
	55	<1%	Tulúa		
	52 49	<1%	Dosquebradas		
	48	<1%	Sogamoso Sincelejo		
	44	<1%	Candelaria		
	44	<1%	Yopal		
	44	<1%	Guarne		
	41	<1%	Florencia		
	41	<1%	Palmira		
	40	<1%	Duitama		
	40	<1%	Tierralta		
	40	<1%	Cereté		
	38	<1%	Galapa		
	37	<1%	Aguachica		
	36	<1%	Mocoa		
	35	<1%	Magangue		
	34	<1%	Soacha		
	33	<1%	Baranoa		
	33	<1%	Planeta Rica		
	32	<1%	Pitalito		
	32	<1%	Palmar De Varela		
	31	<1%	La Dorada Floridablanca		
	30	<1%	Clénaga De Oro		
	30	<1%	Carmen De Viboral		
	30	<1%	Bello		
	30	<1%	Girardot		
	28	<1%	Marinilla		
	27	<1%	Arauca		
	27	<1%	Envigado		
	27	<1%	Córdoba		
	25	<1%	Cartago		
	24	<1%	Chía		
	24	<1%	Guadalajara De Buga		
	24	<1%	Villanueva		
	22	<1%	İtagui		
	21	<1%	Jamundí		
1	,336	14%	(Other)		



PuntoVenta.Non PROVIDES INFO		
PuntoVenta.Regio PuntoVenta.Canal		0.76 0.13
THESE FEATURE GIVE INFORMAT ON PuntoVenta.	ION	eDepto:
GIVE INFORMAT ON PuntoVenta.	10N Nombi	
GIVE INFORMAT	10N Nombi	eDepto: 0.59 0.07

TOP CATEGORIES

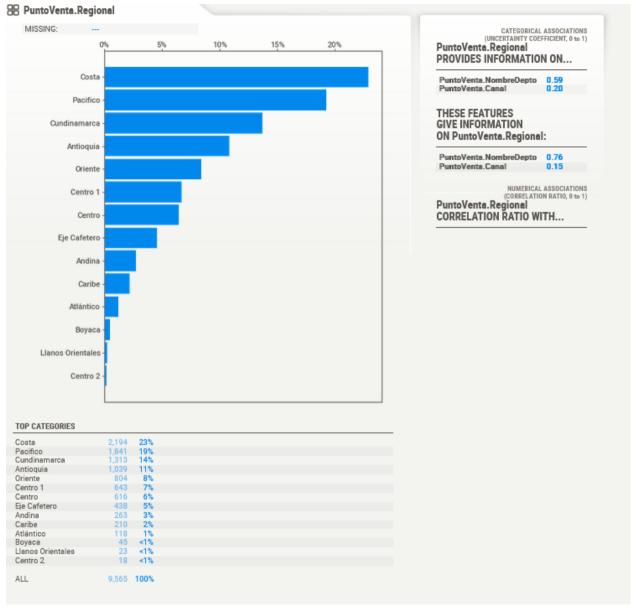
TOP CATEGORIES		
BOGOTA D.C	1,507	16%
VALLE DEL CAUCA	1,488	16%
ANTIOQUIA	1,127	12%
ATLÁNTICO	856	9%
CORDOBA	586	6%
BOLIVAR	444	5%
SANTANDER	303	3%
HUILA	282	3%
NORTE DE SANTANDER	273	3%
CUNDINAMARCA	267	3%
META.	252	3%
CESAR	251	3%
BOYACA	245	3%
QUINDÍO	237	2%
MAGDALENA	225	2%
TOLIMA	202	2%
RISARALDA	184	2%
(Other)	836	9%
ALL	9,565	100%



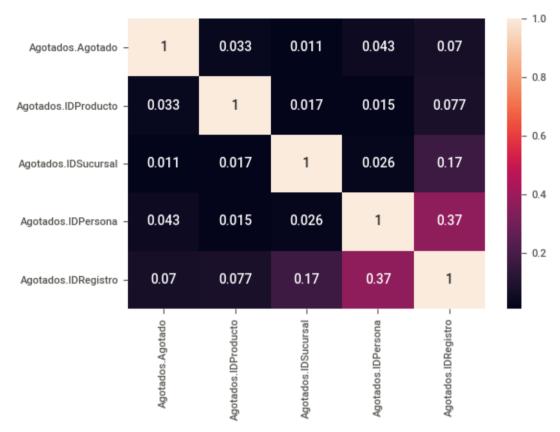
PuntoVenta.Regi	onal	0.15
PuntoVenta.Nomb		0.07
PuntoVenta.Regi		0.20
PuntoVenta.Nom	breDepto	0.13

TOP CATEGORIES

Tiendas 3,276 34% Iam 2,510 26% H&S 1,955 20% Supermercados 750 8% Cr 681 7% D&P 274 3% D&P Tradicional 110 1% Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
lem 2,510 26% H&S 1,955 20% Supermercados 750 8% Cr 681 7% D&P 274 3% D&P Tradicional 11 1% Discounter 7 <1%
H&S 1,955 20% Supermercados 750 8% Cr 681 7% D&P 274 3% D&P Tradicional 110 1% Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
Cr 681 7% D&P 274 3% D&P Tradicional 110 1% Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
D&P 274 3% D&P Tradicional 110 1% Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
D&P Tradicional 110 1% Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
Discounter 7 <1% Indirecto 1 <1% Directo 1 <1%
Indirecto 1 <1% Directo 1 <1%
Directo 1 <1%
ALL 9,565 100%



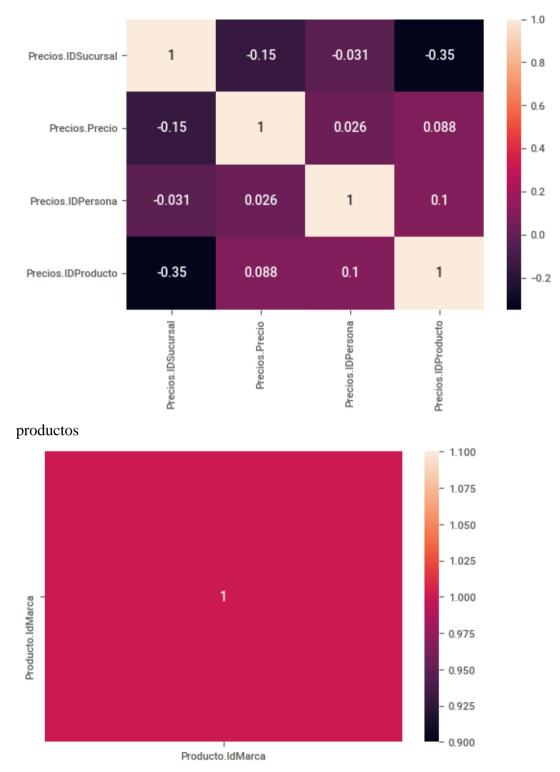
Matrices de correlación Agotados



Exhibiciones



Precios



Puntos de venta no hay correlación entre variables

Methods and Models

Visualization

We are going to use the following types of plots.

- Line charts to visualize time series data like product prices, unemployment rates, CPI, weather, etc.
- Scatter plots to represent graphically dependencies between variables like stockouts and sales or between prices and CPI and unemployment rates, or sales and prices.
- Choropleth maps to visualize how the amount of stockout events depend on geographical location (City, Department, Region) and how these change with time.
- Pie charts to understand the composition of the total amount of stockout events by product category.

•

Models

We can use time series models like Exponential smoothing or ARIMA to forecast variables like prices and sales.

Use methods like X11, STL, SEATS to extract components (trend, seasonality, cyclic) from time series.

We could use clustering algorithms like K-means, DBSCAN, OPTICS to find groups of product and branch pairs with similar sales and stockouts patterns, and use regression algorithms like linear regression, SVM, Decision tree Regressor, k-NN to predict stockouts in terms of the other variables.

Libraries

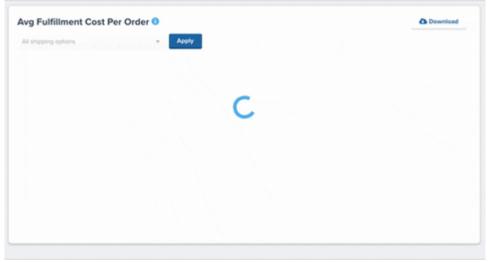
For plotting we can use matplotlib, seaborn, folium, plotly, for times series analysis statsmodels, sktime, darts, for clustering sklearn and pyclustering, for regression and classification statsmodels and sklearn.

Interface

To accomplish the objective of solving the customer's needs, the final front-end product will be an Analytics page with visualizations of the historical stock out of products and a short forecast based on statistics predictive models. The end user can click on a particular mobility dataset which contains the data provided by the client and some parameters that could be adjusted to generate changes in the forecast (ex. postponed time-based causes, retarded time-based causes, quantity-based causes, all this depending on the data).

The next figures are example of the interface in an early stage. We expect the interface to meet the following characteristics: Simplicity (The fewer decorative and/or ornamental elements are part of an interface, the easier it is to identify the functional elements), Clarity. Consistency.





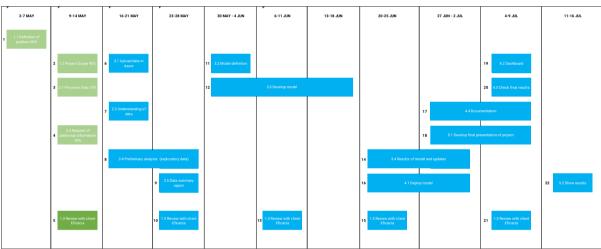
Milestones

Text.

Timeline

The development of the project is shown in the following WBS (work breakdown structure) with its dictionary and the timeline related to it.





STAGE	DESCRIPTION			
1. Business Understanding				
1.1 Definition of problem 90%	Identify the problem raised by effectiveness with respect to your client.			
1.2 Project Scope 90%	Define the scope of the project according to the needs of the client.			
1.3 Review with client	Periodic reviews with the end customer in order to validate and consolidate the advanced stages.			
2. Understanding & Data Analysis				
2.1 Provision Data	Collection of information from both direct and external sources.			
2.2 Understanding of data	Understanding, comprehension, breakdown and assimilation of data.			
2.3 Request of additional information 50%	Request additional information from Efficacy to complement the datasets and narrow down the problem and the solution			
2.4 Preliminary analysis (exploratory data)	First stage of analysis of the data provided and collected to solve the requirement.			
2.5 Data summary report	Report with the results of the collected data.			
3. Modelling				
3.1 Upload data in Azure	Push the collected data to the Azure repository.			
3.2 Model definition	Definition of models and methods to be used for development.			
3.3 Develop model	Development of the defined model.			
3.4 Results of model and updates	Results obtained from the execution of the model and updates to it.			
4. Deployment				
4.1 Deploy model	Deployment of the model developed in the environment designed in Azure.			
4.2 Dashboard	Generation of panels and boards for displaying information.			
4.3 Check final results	Final reviews of the results.			
4.4 Documentation	Documentation corresponding to the project in its entirety.			
5. Results presentation & Acceptance				
5.1 Develop final presentation of project	Design and development of the project presentation.			
5.2 Show results	Final delivery of the project.			

	PROJECT START	lun, 02/	05/2022
TAREA	RESPONSABLE	INICIO	FIN
PHASE 1. BUSINESS UNDERSTANDING			
1.1 Definition of problem		02/05/2022	07/05/2022
1.2 Project Scope		09/05/2022	14/05/2022
1.3 Review with client (1st Session)		12/05/2022	12/05/2022
1.3 Review with client (2nd Session)		26/05/2022	26/05/2022
1.3 Review with client (3rd Session)		09/06/2022	09/06/2022
1.3 Review with client (4th Session)		23/06/2022	23/06/2022
1.3 Review with client (5th Session)		07/07/2022	07/07/2022
1.3 Review with client (6th Session)		21/07/2022	21/07/2022
PHASE 2. UNDERSTANDING AND DATA ANALYSIS			
2.1 Provision Data		09/05/2022	14/05/2022
2.2 Understanding of data		16/05/2022	21/05/2022
2.3 Request of additional information		09/05/2022	14/05/2022
2.4 Preliminary analysis (exploratory data)		16/05/2022	28/05/2022
2.5 Data summary report		23/05/2022	28/05/2022
PHASE 3. MODELLING			
3.1 Upload data in Azure		16/05/2022	21/05/2022
3.2 Model definition		30/05/2022	04/06/2022
3.3 Develop model		30/05/2022	18/06/2022
3.4 Results of model and updates		20/06/2022	02/07/2022
PHASE 4. DEPLOYMENT			
4.1 Deploy model		20/06/2022	02/07/2022
4.2 Dashboard		04/07/2022	09/07/2022
4.3 Check final results		04/07/2022	09/07/2022
4.4 Documentation		27/06/2022	09/07/2022
PHASE 5. RESULTS, PRESENTATION AND ACCEPTANCE			
5.1 Develop final presentation of project		27/06/2022	09/07/2022
5.2 Show results		16/07/2022	16/07/2022

Corcens

The primary concerns with our project are (1) that some of our team members do not know anything about data analysis, and (2) that we have no data.