

KaggleV2-May-2016

Prepare

Explore

Predict

Deployments

Reports

Enable Support

KaggleV2-May-2016.csv

110,527 rows, 16 columns

Data Cleaning and Preprocessing X

Calculate Appointment Duration X

Age Group Categorization X

Clean

Merge

Generate dashboard

Download

PatientId	AppointmentID	Gender	ScheduledDay	AppointmentDay	Age	Neighbourhood	Scholarship	Hipertension
Number (Integer)	ID	Category	Date	Date	Number (Integer)	Category	Category	Category
	110,527 unique values							
29872499824296	5642903	F	2016-04-29 18:38:08	2016-04-29 00:00:00	62	JARDIM DA PENHA	0	1
558997776694438	5642503	M	2016-04-29 16:08:27	2016-04-29 00:00:00	56	JARDIM DA PENHA	0	0
4262962299951	5642549	F	2016-04-29 16:19:04	2016-04-29 00:00:00	62	Other	0	0
867951213174	5642828	F	2016-04-29 17:29:31	2016-04-29 00:00:00	8	Other	0	0
8841186448183	5642494	F	2016-04-29 16:07:23	2016-04-29 00:00:00	56	JARDIM DA PENHA	0	1
95985133231274	5626772	F	2016-04-27 08:36:51	2016-04-29 00:00:00	76	Other	0	1
733688164476661	5630279	F	2016-04-27 15:05:12	2016-04-29 00:00:00	23	Other	0	0
3449833394123	5630575	F	2016-04-27 15:39:58	2016-04-29 00:00:00	39	Other	0	0
56394729949972	5638447	F	2016-04-29 08:02:16	2016-04-29 00:00:00	21	ANDORINHAS	0	0
78124564369297	5629123	F	2016-04-27 12:48:25	2016-04-29 00:00:00	19	Other	0	0
734536231958495	5630213	F				TINA	0	0

What do you want to do with your dataset?

Appointment Duration

Age Group

Number (Integer)

Category

Adult 55.9%

Senior 17.9%

Other 26.2%

-1	Senior
-1	Adult
-1	Senior
-1	Child
-1	Adult
1	Senior
1	Adult
1	Adult
-1	Adult
1	Adult

[Back to Training Types](#)

Predict

Predict Fields

Select which numerical or categorical fields to predict and optionally ignore.

PredictIgnore (2)

Search fields...

☐ Scholarship

☐ Hipertension

☐ Diabetes

☐ Alcoholism

☐ Handcap

☐ SMS_received

☒ No-show

☐ Appointment Duration

Prediction Column
No-show CATEGORY

Outcome of Interest
Yes Edit

[Expand Insights](#)

Classification Summary

Below is a breakdown of how well the model predicted your outcome of interest

Overall Accuracy

79.1%

17,491 out of 22,106 rows tested are classified correctly

[See accuracy details](#)



Performance for Yes

1.8x better than baseline

Rows predicted to be Yes are 1.83x more likely to be Yes than the baseline rate.

[Add to Report](#)

Top Fields

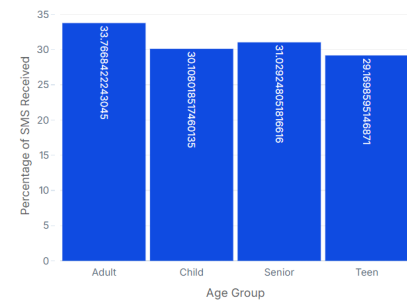
Fields ranked by their contribution to the prediction results



[Share](#) [Preview](#)

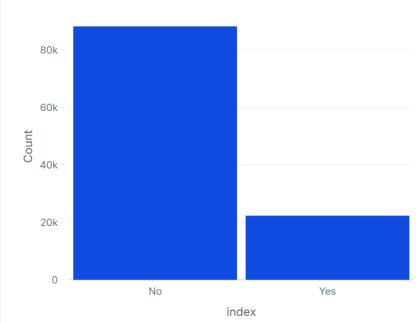
Effectiveness of SMS Reminders by Age Group

Effectiveness of SMS reminders varies by age group, with the percentage of SMS received plotted on the y-axis against different age...



No Show Distribution

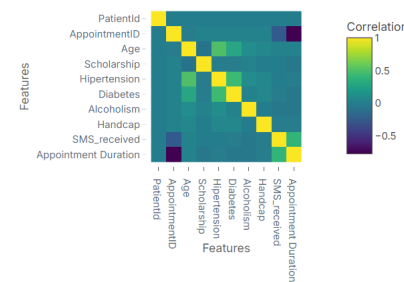
This bar chart illustrates the distribution of No Show responses (Yes/No).



[Share](#) [Preview](#)

Correlation Matrix of Features

Correlation matrix displaying the strength and direction of relationships between different features, with features on both the x-axis and y-axis.



Box Plot of Wait Time by No Show Status

This box plot shows the distribution of appointment wait times for patients who did and did not show up for their appointments.

