CUSTOMER CHURN ANALYSIS

IE 7275

By,
Krithik Sai Sreenish Gopinath
Ullal Atul Nayak
Jay Thakkar

OBJECTIVE

Welcome to our music streaming service! As a company, we pride ourselves on having a wealth of user data, and we're always looking for ways to use that data to accomplish our **2** main goals: **user retention** and **increased profitability**. To achieve these goals, we're pursuing three key strategies:

1. Personalizing the user experience
2. Improving customer engagement and service
3. Optimizing payment plans and pricing.

By pursuing these strategies, we hope to build a better experience for our users, while also increasing our bottom line.

MAIN DATASET

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	msno	city	bd	gender	registered_via	registration_init_time	payment_method_id	payment_plan_days	plan_list_price	actual_amount_paid	is_auto_renew	transaction_date	membership_expire_date	is_cancel	date
2	+tJonkh+O1CA796Fm5X60UMOtB6POHAwPjbTRVl/EuU=	1	0.33945	male	7	2011-09-14 00:00:00	1	0.066666667	0.072147651	0.072147651	1	2017-03-27 00:00:00	2017-04-28 00:00:00	0	2017-03-25 00:00:00
3	l0yFvqMoNkM8ZNHb617e1RBzlS/YRKemHO7Wj13EtA0=	13	0.917431	male	9	2011-09-18 00:00:00	0.967741935	0.066666667	0.083333333	0.083333333	1	2017-03-27 00:00:00	2017-04-26 00:00:00	0	2017-03-29 00:00:00
4	GqYHRxlZChiZvB1uzR410wcQzuxqZNZci4AzOTzkAao=	5	0.587156	male	9	2011-09-27 00:00:00	0.935483871	0.066666667	0.083333333	0.083333333	1	2017-02-28 00:00:00	2017-05-01 00:00:00	0	2017-03-16 00:00:00
5	den0Kb2s4BV47zV+tSC1u0W07M7BOMq+fnrGj+9ax0l=	14	0.577982	female	9	2011-10-05 00:00:00	0.870967742	0.066666667	0.083333333	0.083333333	1	2017-03-04 00:00:00	2017-04-03 00:00:00	0	2017-03-15 00:00:00
6	m0h0+lJIok7k6BzAoOnq+x+mB1tbi9shGF/egYLbed4=	4	0.541284	female	7	2014-10-20 00:00:00	0.838709677	0.066666667	0.100671141	0.100671141	1	2017-03-16 00:00:00	2017-04-15 00:00:00	0	2017-03-30 00:00:00
7	7oCpSPj4+SdVJLfDRc+tN5HLp2R1DRozaWsTpbLTz5g=	15	0.53211	female	3	2014-10-20 00:00:00	0.967741935	0.066666667	0.083333333	0.083333333	1	2017-03-24 00:00:00	2017-05-27 00:00:00	0	2017-03-04 00:00:00
8	vWYxNeXA8NwsiQQr7Beq5XQgCcj0y9N0YxEN2A2iqxI=	1	0.33945	male	9	2014-10-21 00:00:00	0.903225806	0.066666667	0.083333333	0.083333333	0	2017-03-30 00:00:00	2017-04-29 00:00:00	0	2017-03-04 00:00:00
9	fkLgfIOX0bWM9/BQQChOCDzoos23szsckxPvxrBbtmY=	4	0.522936	female	9	2014-10-29 00:00:00	0.903225806	0.066666667	0.083333333	0.083333333	0	2017-03-01 00:00:00	2017-03-31 00:00:00	0	2017-03-04 00:00:00
10	mIsuyquVwPXmJhd178ipGqdXJ6W7RSuDhd0fN+Pgn98=	13	0.522936	female	3	2014-11-01 00:00:00	0.903225806	0.066666667	0.083333333	0.083333333	0	2017-03-13 00:00:00	2017-04-12 00:00:00	0	2017-03-06 00:00:00
11	0uAjNEzX6IY+lZ+PYv9W31mZjHxUWkbFz9LL8m9G46k=	15	0.550459	female	7	2014-11-03 00:00:00	1	0.066666667	0.072147651	0.072147651	1	2017-03-27 00:00:00	2017-04-27 00:00:00	0	2017-03-12 00:00:00
12	+lkl5vyTsFd0RQlcpqvsrYx/6xy8lHu68XNB0p3bJZM=	14	0.559633	female	9	2014-11-04 00:00:00	0.967741935	0.066666667	0.083333333	0.083333333	1	2017-03-06 00:00:00	2017-04-09 00:00:00	0	2017-03-19 00:00:00
13	TW8bEcKDh38eaQ7nMZRioBtfZXnKyPtgLiRQhz2bII8=	1	0.33945	male	7	2014-11-07 00:00:00	1	0.066666667	0.055369128	0.055369128	1	2017-03-05 00:00:00	2017-04-05 00:00:00	0	2017-03-08 00:00:00
14	6GDUJe0yLoCvTtY9N+KnYKWN+M/z5KI5etSmrOANm/0=	13	0.53211	female	3	2014-11-10 00:00:00	0.903225806	0.066666667	0.083333333	0.083333333	0	2017-03-15 00:00:00	2017-04-14 00:00:00	0	2017-03-22 00:00:00

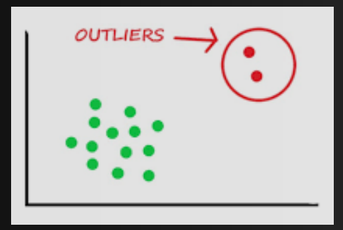
	O	P	Q	R	S	T	U	V	W
1	l date	num_25	num_50	num_75	num_985	num_100	num_unq	total_secs	is_churn
2	0 2017-03-25 00:00:00	0.580482024	0.613147193	0	0.630929754	0.6484145	0.637422634	0.835485393	0
3	0 2017-03-29 00:00:00	0	0	0	0	0.736677936	0.624196351	0.878507556	0
4	0 2017-03-16 00:00:00	0.75	0.773705614	0.630929754	0.630929754	0.881151974	0.8671124	0.937484834	0
5	0 2017-03-15 00:00:00	0.25	0	0	0	0.568700379	0.485719171	0.797626899	0
6	0 2017-03-30 00:00:00	0.580482024	0.613147193	0	0	0.717595065	0.650033514	0.869693357	0
7	0 2017-03-04 00:00:00	0.92510993	0	0.630929754	0.630929754	0.696775094	0.743000702	0.862286981	1
8	0 2017-03-04 00:00:00	0.5	0.386852807	0	0	0.502860922	0.248392701	0.752676954	0
9	0 2017-03-04 00:00:00	0.25	0.386852807	0	0.630929754	0.502860922	0.248392701	0.775217341	0
10	0 2017-03-06 00:00:00	0.25	0	0	0	0.317269918	0.248392701	0.675000584	0
11	0 2017-03-12 00:00:00	0.580482024	0.773705614	0.630929754	1	0.958852213	0.8671124	0.971417609	0
12	0 2017-03-19 00:00:00	0.580482024	0.386852807	0	0	0.661495881	0.662083721	0.834898656	0
13	0 2017-03-08 00:00:00	0.25	0	0.630929754	0.630929754	0.807049459	0.790963968	0.912542235	0
14	0 2017-03-22 00:00:00	0.580482024	0	1	0	0.785908728	0.775899814	0.900437341	0

PERSONALIZING USER EXPERIENCE

OBJECTIVE

From this music streaming dataset, we aim to develop strategies for each customer segment to personalize their UX and user engagement, to increase user retention and profits.

DATA PRE-PROCESSING



OUTLIER REMOVAL

PCA, ENCODING (CAT)

DATE FORMAT CONVERSION

STANDARDIZE

FEATURE ENGINEERING

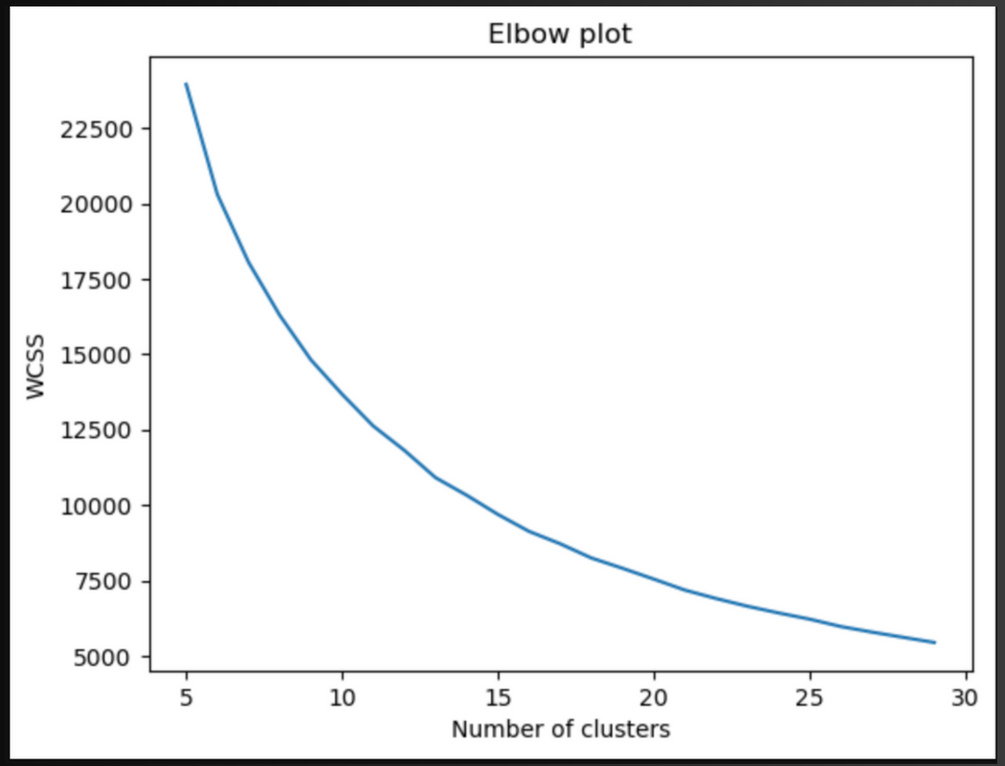
AVG SEC PER SONG

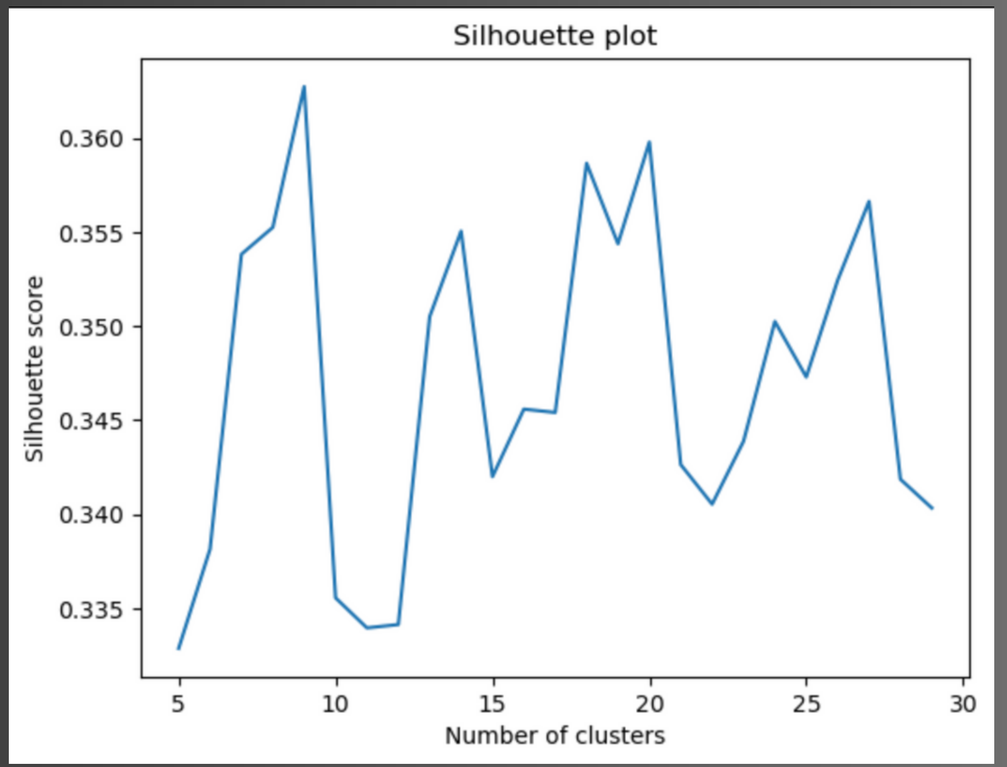
Average of total seconds played per unique song in the entire tenure of the user on the app.

LENGTH OF MEMBERSHIP

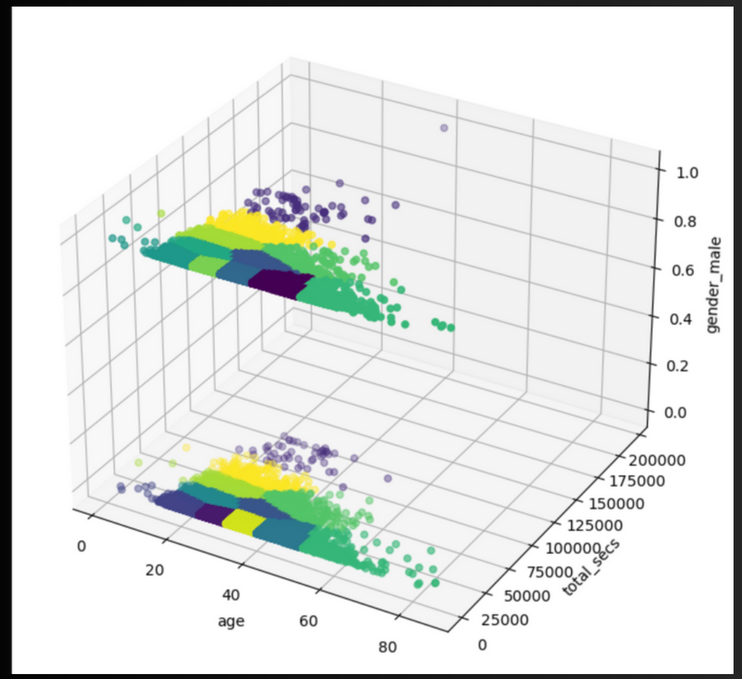
Number of days it has been since the user started using the app

CHOOSING K – NUMBER OF CLUSTERS

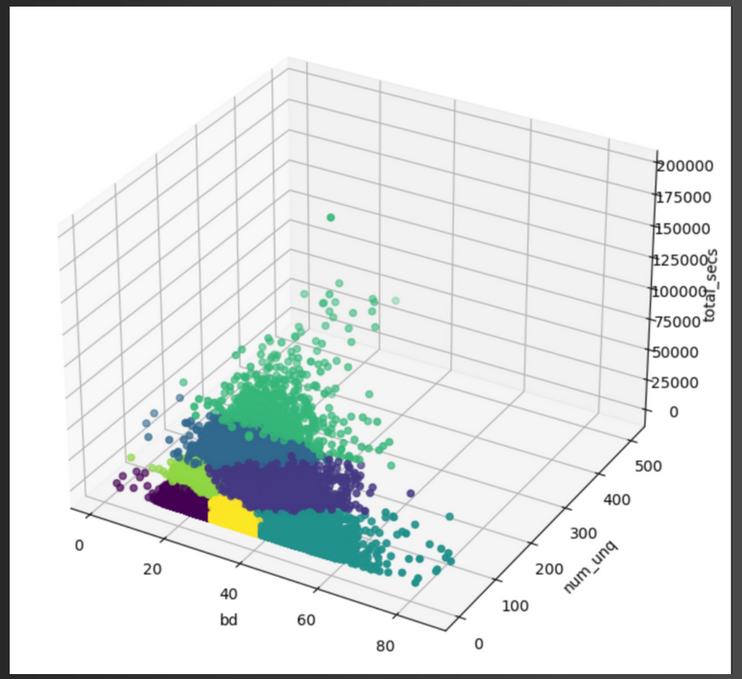




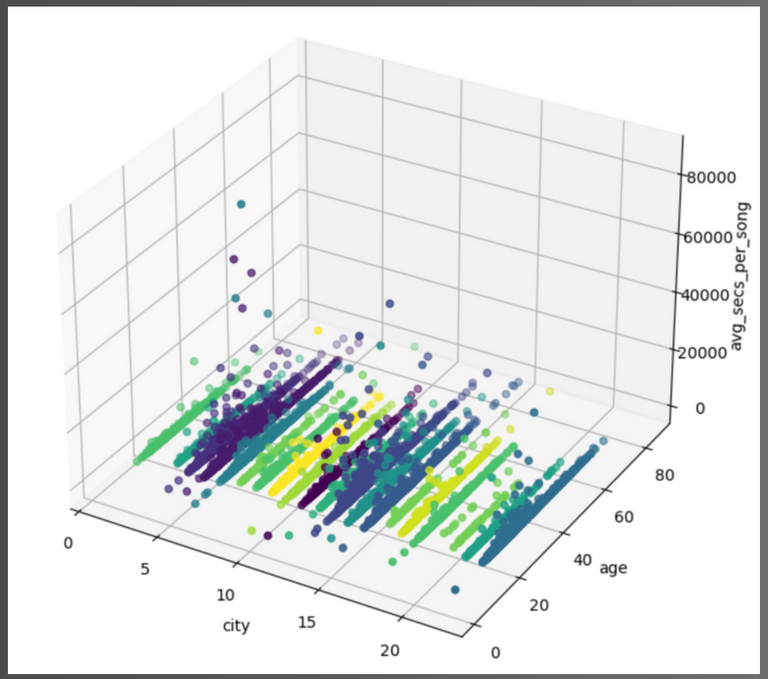
CLUSTERS BASED ON VARIOUS DEMOGRAPHICS



Activity - Age - Gender



Activity - Age - Type of Activity



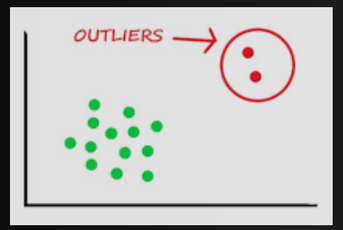
Activity - Age - City

IMPROVING CUSTOMER SERVICE AND USER ENGAGEMENT

OBJECTIVE

From this music streaming dataset, we aim to develop strategies for each customer segment to enhance customer service and user engagement, ultimately increasing user retention and profits.

DATA PRE-PROCESSING



OUTLIER REMOVAL

- MEDIAN IMPUTATION (CONT.)
- MODE IMPUTATION (CAT.)

STANDARDIZE

DATE FORMAT CONVERSION

SKEWNESS REDUCTION

UNDER SAMPLING (FOR EQUAL CHURN)

FEATURE ENGINEERING

TENURE

Number of days the user has been on the app.

AVG DAILY SONGS

Average of total no of songs (completely) played every day in the entire tenure of the user on the app.

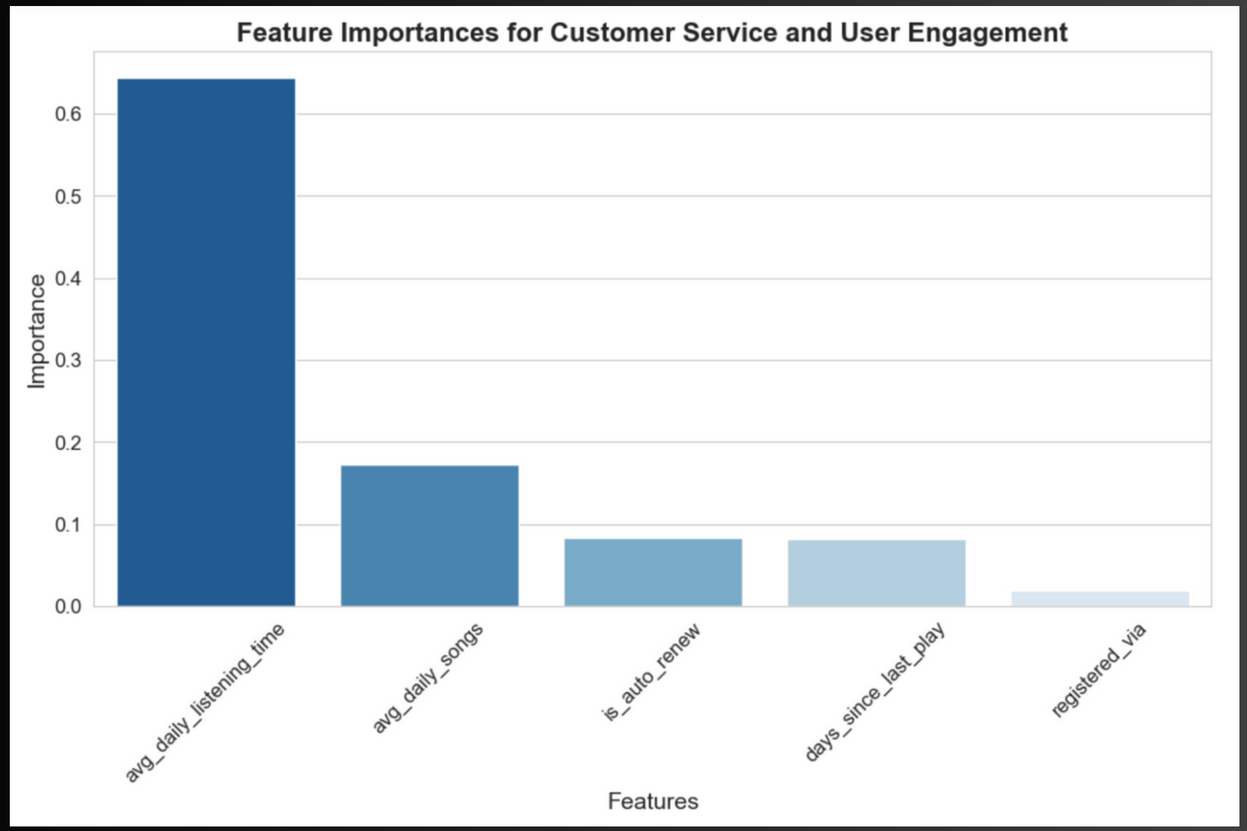
AVG DAILY LISTENING TIME

Average daily time spent by the user on the app in the entire tenure

DAYS SINCE LAST PLAY

Number of days it has been since the user last used the app.

FEATURES IMPORTANCE FOR USER ENGAGEMENT

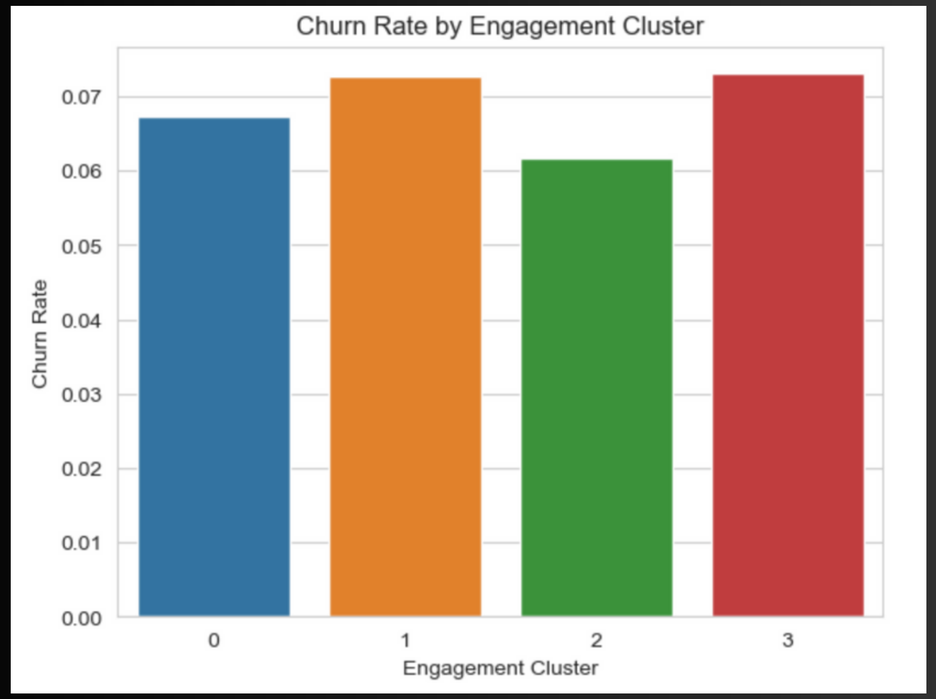


We have applied decision tree to train a model using the features that determine payment behaviors to predict churn and have used that trained model to extract the feature importance. Below is the feature importance in decreasing order.

Performed

- Correlation Analysis
- Decision Tree

BUSINESS RECOMMENDATIONS



We performed clustering analysis.
Overall, these clusters are quite similar in terms of their demographic and registration characteristics

Since average daily listening time is the most important feature. We should provide personalized content and gamify their experience. So we have come up with such recommendations particular to clusters which would help reduce churn and increase user engagement.

FEATURE ENGINEERED COLUMNS AND STRATEGY

W	X	Y	Z	AA	AB	AC	AD
is_churn	tenure	days_since_last_transaction	avg_daily_songs	avg_daily_listening_time	days_since_last_play	cluster	engagement_strategy
0	32	2203	0.020262953	0.026108919	6	2	Since this cluster has a much lower average days_since_last_transaction, it may be beneficial to focus on retaining users by providing high-quality content and personalized recommend
0	30	2203	0.024555931	0.029283585	2	2	Since this cluster has a much lower average days_since_last_transaction, it may be beneficial to focus on retaining users by providing high-quality content and personalized recommend
0	62	2230	0.014212129	0.015120723	15	3	Since this cluster has a higher proportion of female users, consider tailoring promotional offers or content to appeal to this demographic. Consider implementing social features, such a
0	30	2226	0.018956679	0.026587563	16	3	Since this cluster has a higher proportion of female users, consider tailoring promotional offers or content to appeal to this demographic. Consider implementing social features, such a
0	30	2214	0.023919836	0.028989779	1	2	Since this cluster has a much lower average days_since_last_transaction, it may be beneficial to focus on retaining users by providing high-quality content and personalized recommend
1	64	2206	0.010887111	0.013473234	27	1	Similar to Cluster 0, it is important to send reminders or offers to encourage users to make more transactions. Consider implementing a loyalty or rewards program to incentivize users t
0	30	2200	0.016762031	0.025089232	27	1	Similar to Cluster 0, it is important to send reminders or offers to encourage users to make more transactions. Consider implementing a loyalty or rewards program to incentivize users t
0	30	2229	0.016762031	0.025840578	27	1	Similar to Cluster 0, it is important to send reminders or offers to encourage users to make more transactions. Consider implementing a loyalty or rewards program to incentivize users t
0	30	2217	0.010575664	0.022500019	25	1	Similar to Cluster 0, it is important to send reminders or offers to encourage users to make more transactions. Consider implementing a loyalty or rewards program to incentivize users t
0	31	2203	0.030930717	0.031336052	19	3	Since this cluster has a higher proportion of female users, consider tailoring promotional offers or content to appeal to this demographic. Consider implementing social features, such a
0	34	2224	0.019455761	0.024555843	12	0	Since this cluster has a higher average days_since_last_transaction, it is important to send reminders or offers to encourage users to make more transactions. Consider providing persona
0	31	2225	0.026033854	0.029436846	23	1	Similar to Cluster 0, it is important to send reminders or offers to encourage users to make more transactions. Consider implementing a loyalty or rewards program to incentivize users t
0	30	2215	0.026196958	0.030014578	9	0	Since this cluster has a higher average days_since_last_transaction, it is important to send reminders or offers to encourage users to make more transactions. Consider providing persona

OPTIMIZING PAYMENT PLANS AND PRICING

OBJECTIVE

By analyzing our user data, we aim to identify the payment plans that are already working and use this information to promote those plans to our users, ultimately increasing user retention and profits.

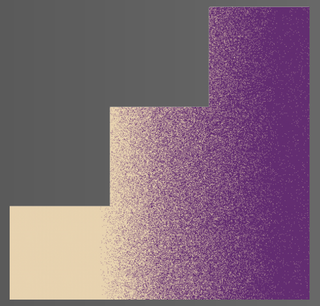
FEATURE ENGINEERING

AVERAGE PAYMENT AMOUNT

Mean of payment amounts of each user

PAYMENT FREQUENCY

Number of payments made by each user within a specific month



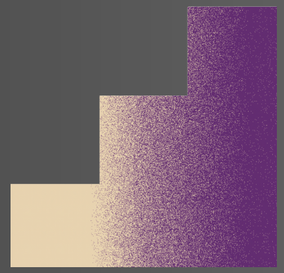
NUMBER OF PLAN CHANGES

Count the number of times a user changed their plan

FEATURES INDICATING PAYMENT BEHAVIORS

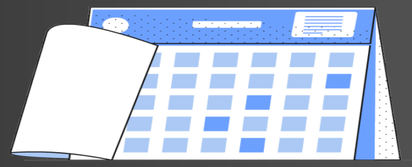
AVERAGE PAYMENT AMOUNT

PAYMENT FREQUENCY



NUMBER OF PLAN CHANGES

PLAN LIST PRICE



PAYMENT PLAN DAYS

FEATURE IMPORTANCE

We have applied **logistic regression** to train a model using the features that determine payment behaviors to predict churn and have used that trained model to extract the feature importance. Below is the feature importance in decreasing order.

PLAN LIST PRICE

PAYMENT PLAN DAYS

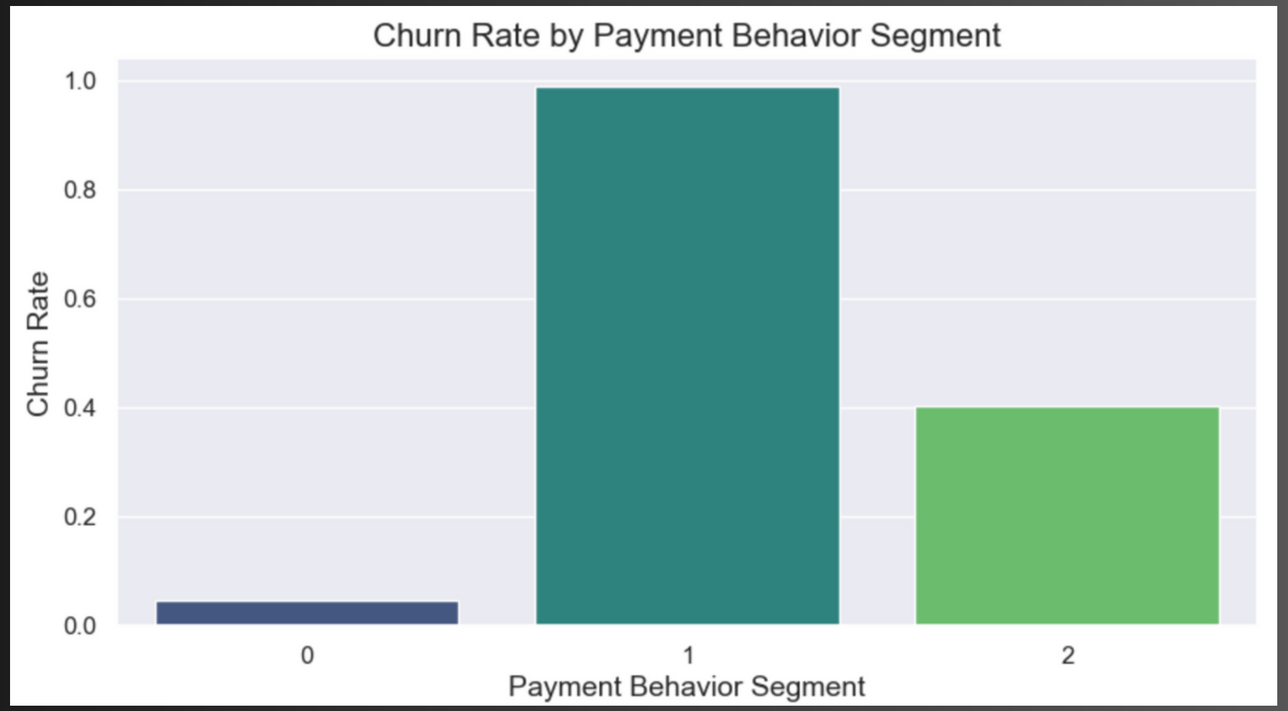
NUMBER OF PLAN CHANGES

AVERAGE PAYMENT AMOUNT

MOST IMPORTANT FEATURE

PLAN LIST PRICE

PAYMENT BEHAVIOR SEGMENTS



payment_behavior_segment	average_payment_amount	number_of_plan_changes	payment_plan_days	payment_method_id	plan_list_price
0	149	0.00000	30	41	149
1	1788	0.99661	410	32	1788
2	149	0.99963	30	41	149

OPTIMAL PAYMENT PLAN & PRICING

30-DAY PLAN

$149

THANK YOU