

Credit Card Customer Analysis Documentation

Dataset: <https://www.kaggle.com/datasets/whenamancodes/credit-card-customers-prediction>

Performed using AWS S3, Glue, Athena, Quicksight.

1. Dataset found at kaggle.com: *Credit Card Customers Prediction*
 - a. Data Cleaning: Cut Down Dataset to 200 Records from original 17,000 records for simplicity

Credit Card Customers Prediction

Predict Churning Customers



Data

Code (2)

Discussion (2)

About Dataset

A manager at the bank is disturbed with more and more customers leaving their credit card services. They would really appreciate if one could predict for them who is gonna get churned so they can proactively go to the customer to provide them better services and turn customers' decisions in the opposite direction.

I got this dataset from a website with the URL as <https://leaps.analytica.com/home>. I have been using this for a while to get datasets and accordingly work on them to produce fruitful results. The site explains how to solve a particular business problem.

Now, this dataset consists of 10,000 customers mentioning their age, salary, marital_status, credit card limit, credit card category, etc. There are nearly 18 features.

We have only 16.07% of customers who have churned. Thus, it's a bit difficult to train our model to predict churning customers.

Data Dictionary

Usability

10.00

Licence

CC0: Public Domain

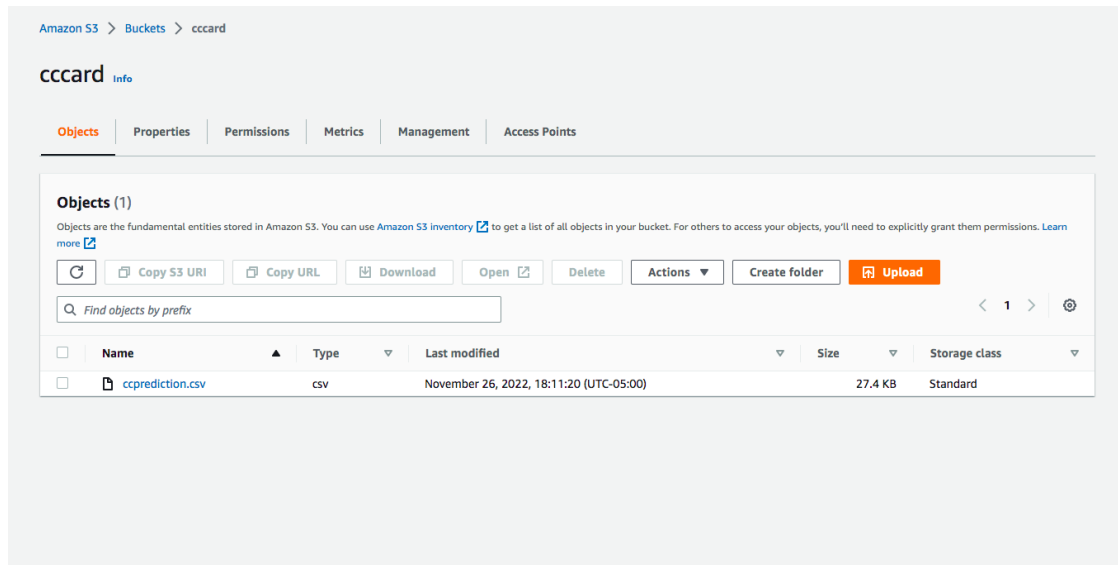
Expected update frequency

Never

CUSTOMER	Attrition_Flag	Customer_Age	Gender	Dependent_cous	Education_Level	Marital_Status	Income_Category	Card_Category	Months_on_book	Total_Relationship_Count	Months_Inactive_12_mon	Contacts_Count_12_mon	Credit_Lim	Total_Revolving_Bal	Avg_Op_Cost_Tx	Total_Acc_Open_12_Mo	Total_Acc_Chg_12_Mo	Total_Acc_Pay_12_Mo	Total_Acc_Inv_12_Mo	Total_Cat_Chg_12_Mo	
70680363	Existing Customer	40 M	M	1	High School	Married	\$60K - \$70K	Blue	39	5	1	3	13991	777	1914	1330	1144	42	1	0	
81871066	Existing Customer	40 F	F	1	Single	1	Over \$100K	Blue	41	6	1	8	8256	894	1341	1391	33	3	0	0	
71390101	Existing Customer	51 M	M	3	Graduate	Married	\$40K - \$50K	Blue	36	4	0	0	3418	0	3418	1887	20	2	0	0	
78991188	Existing Customer	40 F	F	4	High School	Unknown	Less than \$40K	Blue	34	3	4	1	3313	2517	796	1405	1171	20	2	0	
70910056	Existing Customer	40 M	M	1	Uneducated	\$60K - \$80K	Blue	21	5	1	0	0	4716	0	4716	816	39	0	0	0	
71301165	Existing Customer	44 M	M	2	Graduate	Married	\$40K - \$60K	Blue	36	3	1	2	4010	1247	2793	1376	1088	24	0	0	
81041708	Existing Customer	51 M	M	4	Unknown	Married	\$10K - 1	Gold	48	6	1	3	34516	2364	3232	1975	1330	31	0	0	
81840046	Existing Customer	42 M	M	1	High School	Unmarried	\$60K - \$80K	Blue	36	3	2	0	1386	2760	1036	1036	36	0	0	0	
70650006	Existing Customer	37 M	M	3	Uneducated	Single	\$60K - \$80K	Blue	36	5	0	0	22362	1517	1965	3360	1300	24	1	0	
71991056	Existing Customer	49 M	M	2	Graduate	Single	\$60K - \$120K	Blue	36	6	3	0	11656	1677	9679	1524	1441	32	0	0	
70970603	Existing Customer	42 M	M	1	Uneducated	Unknown	\$12K - 1	Blue	31	5	0	2	6746	1467	821	8821	1201	42	1	0	
70910043	Existing Customer	40 M	M	1	Unknown	Married	\$60K - \$80K	Blue	54	2	0	0	3496	708	1403	1014	26	1	0	0	
71930663	Existing Customer	58 M	M	1	College	Single	\$60K - \$120K	Blue	36	3	6	0	11751	0	11751	3387	1039	17	0	0	
81860233	Existing Customer	38 M	M	3	Graduate	Unknown	\$60K - \$80K	Blue	30	5	0	3	8847	1866	681	1163	1311	33	0	0	
71230666	Existing Customer	47 F	F	2	Graduate	Married	Less than \$40K	Blue	48	5	3	2	2438	685	1756	119	1310	29	0	0	
71480655	Existing Customer	44 M	M	4	Unknown	Unknown	\$50K - \$120K	Blue	32	3	0	0	4234	972	2002	1707	1348	27	0	0	
70991738	Existing Customer	44 M	M	4	Post-Graduate	Single	\$60K - \$120K	Blue	36	6	2	3	33847	2362	2855	1768	1871	27	0	0	
70740703	Existing Customer	41 M	M	1	Unknown	Married	\$60K - \$120K	Blue	34	4	0	1	13355	1291	1294	9883	1036	31	0	0	
80618016	Existing Customer	40 M	M	1	High School	Married	\$60K - \$80K	Blue	56	2	0	0	3193	2517	676	1831	1036	30	0	0	
70930763	Existing Customer	40 F	F	2	Graduate	Married	Unknown	Blue	37	6	0	0	14470	1157	1833	9666	1287	21	0	0	
80618036	Existing Customer	40 M	M	1	Doctorate	Divorced	\$60K - \$80K	Blue	42	5	0	0	23979	1800	2170	9306	1176	27	0	0	
70650078	Existing Customer	40 F	F	2	Graduate	Married	Less than \$40K	Blue	49	3	0	0	1458.5	0	1458.5	1587	882	16	0	0	
74473333	Existing Customer	41 M	M	3	High School	Married	\$40K - \$60K	Blue	33	4	0	2	11447	680	3790	1608	801	16	1	0	
81104133	Existing Customer	42 F	F	2	Unknown	Single	Less than \$40K	Blue	36	3	0	2	2462	1580	932	6573	1106	23	0	0	
71071065	Existing Customer	40 M	M	2	Unknown	Married	\$60K - \$120K	Blue	42	4	0	0	12217	0	12217	1025	1110	21	0	0	
71701166	Existing Customer	41 F	F	3	Graduate	Single	Less than \$40K	Blue	28	6	1	2	7787	1689	609	9767	1051	22	0	0	
72048630	Existing Customer	50 M	M	1	High School	Unknown	\$40K - \$60K	Blue	46	4	0	2	14784	1274	15410	6181	1197	23	0	0	
80442403	Existing Customer	40 M	M	1	Unknown	Married	\$60K - \$80K	Blue	56	3	0	0	15216	1010	820	8463	1504	40	0	0	
71801023	Existing Customer	41 F	F	1	Uneducated	Single	Unknown	Blue	34	2	0	0	19100	0	19100	9392	1822	16	1	0	
80661400	Existing Customer	50 M	M	4	High School	Married	\$40K - \$60K	Blue	42	6	0	0	4785	1362	1623	6739	1545	38	0	0	
77394033	Existing Customer	33 M	M	3	Unknown	Married	\$60K - \$120K	Blue	33	3	2	0	3733	1811	94	9377	1038	25	2	0	
71931662	Existing Customer	50 M	M	1	Uneducated	Married	\$60K - \$80K	Blue	45	3	0	0	2451	1695	701	1353	1566	26	0	0	
70001840	Existing Customer	41 M	M	4	Graduate	Married	\$60K - \$80K	Blue	35	2	1	0	8933	2517	6406	1708	1689	24	1	0	
78805445	Existing Customer	50 F	F	2	College	Married	Less than \$40K	Blue	38	5	2	3	2650	1480	1100	175	1411	28	0	0	
70910046	Existing Customer	40 M	M	1	Unknown	Married	\$60K - \$120K	Blue	36	3	0	0	13355	1686	1036	9819	1231	34	0	0	
71310606	Existing Customer	50 F	F	1	College	Single	Less than \$40K	Blue	36	4	2	1	3520	1914	1006	651	1877	43	0	0	
71380233	Existing Customer	50 F	F	3	Graduate	Married	Less than \$40K	Blue	36	6	0	2	3035	2288	737	1724	1877	37	1	0	
78640370	Existing Customer	42 F	F	1	High School	Married	Less than \$40K	Gold	36	3	0	0	15023	0	15023	8860	960	22	0	0	
71910023	Existing Customer	42 F	F	1	Graduate	Unknown	\$60K - \$80K	Blue	36	5	0	0	3672	2786	886	2786	132	1464	28	0	0
70001460	At-Risk Customer	68 F	F	0	Doctorate	Married	Unknown	Blue	56	5	4	0	7862	605	7277	1052	704	16	0	0	
82711393	Existing Customer	40 M	M	3	Graduate	Single	\$60K - \$120K	Blue	41	2	2	0	32426	578	35848	1042	1100	28	0	0	
70001460	Existing Customer	40 M	M	1	Unknown	Married	\$60K - \$80K	Blue	44	4	1	0	6205	2204	6863	1247	26	0	0		
71380233	Existing Customer	50 F	F	1	College	Single	\$40K - \$60K	Blue	43	3	0	0	17034	2517	14787	1449	1708	33	0	0	
77394038	Existing Customer	49 M	M	3	High School	Married	\$60K - \$80K	Blue	37	3	2	1	3906	0	3906	1214	1756	32	0	0	
70910046	Existing Customer	38 F	F	3	Graduate	Single	\$60K - 1	Blue	3	0	0	0	8450	2855	1776	3397	1640	23	0	0	
70910046	Existing Customer	40 M	M	4	Uneducated	Single	\$60K - \$120K	Blue	30	3	0	0	34516	0	34516	1444	1444	28	1	0	
78917625	Existing Customer	50 M	M	2	Doctorate	Married	\$60K - \$80K	Blue	45	6	0	0	2283	1430	853	2336	1741	27	0	0	

Total Revolving Bal

2. CSV file imported into AWS S3 Bucket



3. IAM Role created to give permission to user for AWS Glue

AWSGlueServiceRole-ccdebt

Delete

Summary

Edit

Creation date

November 08, 2022, 10:39 (UTC-05:00)

ARN

arn:aws:iam::909304551513:role/service-role/AWSGlueServiceRole-ccdebt

Last activity

✓ 25 days ago

Maximum session duration

1 hour

Permissions

Trust relationships

Tags

Access Advisor

Revoke sessions

Permissions policies (2) Info

You can attach up to 10 managed policies.



Simulate

Remove

Add permissions

Filter policies by property or policy name and press enter.

<input type="checkbox"/>	Policy name	Type	Description
<input type="checkbox"/>	AWSGlueServiceRole-ccdebt	Customer managed	This policy will be used for Glue
<input type="checkbox"/>	AWSGlueServiceRole	AWS managed	Policy for AWS Glue service rol

4. AWS Glue Crawler: ETL job populates Glue Catalog Tables from CSV file in S3 (uses metadata to assess table format)

Crawlers > ccdebtdata

[Run crawler](#) [Edit](#)

Name	ccdebtdata
Description	
Create a single schema for each S3 path	false
Security configuration	
Tags	-
State	Running
Schedule	
Last updated	Tue Nov 08 10:40:27 GMT-500 2022
Date created	Tue Nov 08 10:40:27 GMT-500 2022
Database	default
Table level	
Table threshold	
Service role	service-role/AWSGlueServiceRole-ccdebt
Selected classifiers	
Data store	S3
Include path	s3://cccard
Connection	
Exclude patterns	

Configuration options

Schema updates in the data store	Update the table definition in the data catalog.
Object deletion in the data store	Mark the table as deprecated in the data catalog.

5. Athena Queries: SQL statements executed on data stored in Data catalog, extracted from the Glue crawler job. Results stored in a separate S3 bucket.

Amazon Athena > Query editor

Editor | Recent queries | Saved queries | Settings

Workgroup: primary

Data

Data source: AwsDataCatalog

Database: default

Tables and views: Create

Filter tables and views

Tables (1): cccard

Views (0)

SQL: `1 SELECT "Gender", "Customer_Age", "Income_Category", "Total_Revolving_Bal" FROM "cccard"`

Run again | Explain | Cancel | Clear | Create

Reuse query results (disabled)

Athena engine version 3 only

Query results | Query stats

Completed | Time in queue: 125 ms | Run time: 559 ms | Data scanned: 27.45 KB

6. Quicksight Visualizations executed from S3 bucket containing results of Athena Queries

