

Database Name	Field Name	Description	Type	Format	Data Codes	Missing values
Base.CSV	Income	Annual income of the applicant (in decile form)	Numeric	float64	Ranges between [0.1, 0.9]	N/A
Base.CSV	Name_Email_Similarity	Metric of similarity between email and applicant's name. Higher values represent higher similarity	Numeric	float64	Ranges between [0, 1]	N/A
Base.CSV	Prev_Address_Months_Count	Number of months in previous registered address of the applicant, i.e., the applicant's previous residence, if applicable	Numeric	int64	Ranges between [-1, 380] months	(-1 is a missing value)
Base.CSV	Current_Address_Months_Count	Months in currently registered address of the applicant	Numeric	int64	Ranges between [-1, 429] months (-1 is a missing value)	(-1 is a missing value)
Base.CSV	Customer_Age	Applicant's age in years, rounded to the decade	Numeric	int64	Ranges between [10, 90] years	N/A
Base.CSV	Days_Since_Request	Number of days passed since application was done	Numeric	float64	Ranges between [0, 79] days	N/A
Base.CSV	Intended_Balcon_Amount	Initial transferred amount for application	Numeric	float64	Ranges between [-16, 114] (negatives are missing values)	N/A
Base.CSV	Payment_Type	Credit payment plan type	Categorical	object	5 possible (anonymized) values: - AA - AB - AC - AD - AE	N/A
Base.CSV	Zip_Count_4W	Number of applications within same zip code in last 4 weeks	Numeric	int64	Ranges between [1, 6830]	N/A
Base.CSV	Velocity_6H	Velocity of total applications made in last 6 hours i.e., average number of applications per hour in the last 6 hours	Numeric	float64	Ranges between [-175, 16818]	N/A
Base.CSV	Velocity_24H	Velocity of total applications made in last 24 hours i.e., , average number of applications per hour in the last 24 hours	Numeric	float64	Ranges between [1297, 9586]	N/A
Base.CSV	Velocity_4W	Velocity of total applications made in last 4 weeks, i.e., average number of applications per hour in the last 4 weeks	Numeric	float64	Ranges between [2825, 7020]	N/A
Base.CSV	Bank_Branch_Count_8W	Number of total applications in the selected bank branch in last 8 weeks	Numeric	int64	Ranges between [0, 2404]	N/A
Base.CSV	Date_Of_Birth_Distinct_Emails_4W	Number of emails for applicants with same date of birth in last 4 weeks	Numeric	int64	Ranges between [0, 39]	N/A
Base.CSV	Employment_Status	Employment status of the applicant	Categorical	object	7 possible (anonymized) values: - CA - CB - CC - CD - CE - CF - CG	N/A
Base.CSV	Credit_Risk_Score	Internal score of application risk	Numeric	int64	Ranges between [-191, 389]	N/A
Base.CSV	Email_Is_Free	Domain of application email (either free or paid)	Binary	int64	Either 1 or 0.	N/A

Base.CSV	Housing_Status	Current residential status for applicant	Categorical	object	7 possible (anonymized) values: - BA - BB - BC - BD - BE - BF - BG	N/A
Base.CSV	Phone_Home_Valid	Validity of provided home phone	Binary	int64	Either 1 or 0.	N/A
Base.CSV	Phone_Mobile_Valid	Validity of provided mobile phone	Binary	int64	Either 1 or 0.	N/A
Base.CSV	Bank_Months_Count	How old is previous account (if held) in months	Numeric	int64	Ranges between [-1, 32] months	(-1 is a missing value)
Base.CSV	Has_Other_Cards	If applicant has other cards from the same banking company	Binary	int64	Either 1 or 0.	N/A
Base.CSV	Proposed_Credit_Limit	Applicant's proposed credit limit	Numeric	float64	Ranges between [200, 2000]	N/A
Base.CSV	Foreign_Request	If origin country of request is different from bank's country	Binary	int64	Either 1 or 0.	N/A
Base.CSV	Source	Online source of application	Categorical	object	Either browser (INTERNET) or app (TELEAPP)	N/A
Base.CSV	Session_Length_In_Minutes	Length of user session in banking website in minutes	Numeric	float64	Ranges between [-1, 107] minutes	(-1 is a missing value)
Base.CSV	Device_Os	Operative system of device that made request	Categorical	object	Possible values are: - Windows - MacOS - Linux - X11- - Other	N/A
Base.CSV	Keep_Alive_Session	User option on session logout	Binary	int64	Either 1 or 0.	N/A
Base.CSV	Device_Distinct_Emails	Number of distinct emails in banking website from the used device in last 8 weeks	Numeric	int64	Ranges between [-1, 2] emails	(-1 is a missing value)
Base.CSV	Device_Fraud_Count	Number of fraudulent applications with used device	Numeric	int64	Ranges between [0, 1]	N/A
Base.CSV	Month	Month where the application was made	Numeric	int64	Ranges between [0, 7]	N/A
Base.CSV	Fraud_Bool	If the application is fraudulent or not	Binary	int64	Either 1 or 0.	N/A