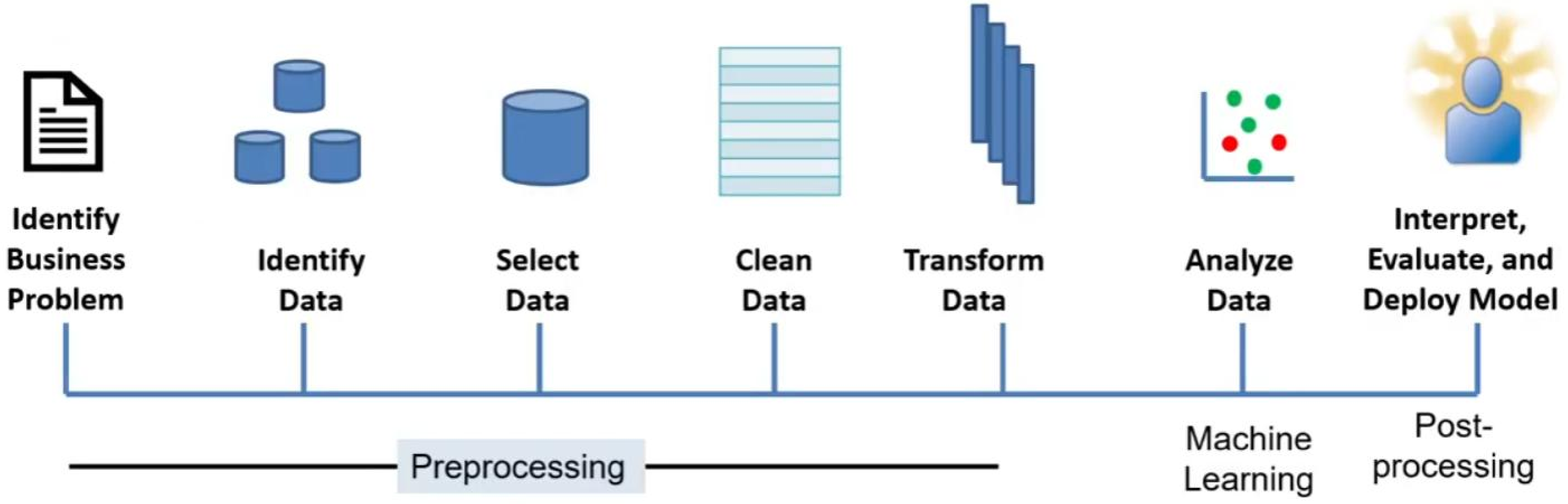
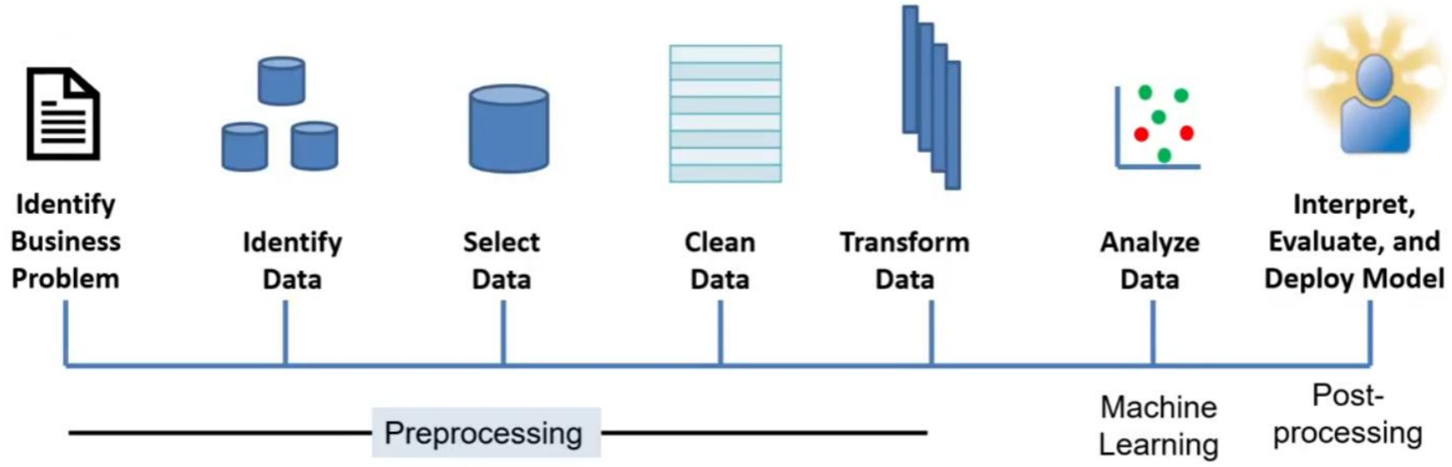
Fraud detection

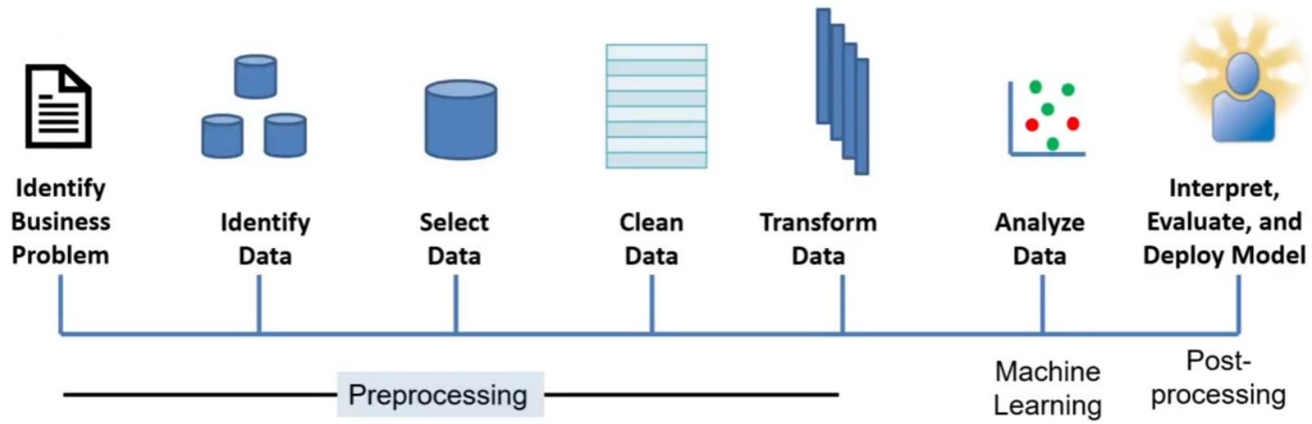
Roadmap



Identify Business Problem
Identify Data
Select Data
Clean Data
Transform Data
Analyze Data
Interpret, Evaluate, and Deploy Model
Preprocessing
Machine Learning
Post-processing



The business objective is to analyze the provided transaction data to **identify outliers** whose behavior significantly deviates from normal customer patterns. The goal is to **detect patterns** associated with fraudulent transactions and generate actionable insights to help **mitigate and prevent** fraudulent activities.



Data preprocessing:

(1). How does the data look like?
(2). Does our dataset contain any missing values?
(3). What is the distribution of our target variable?
(4). Feature engineering
(5). What is the distribution of our predictive variables?

How does the data look like?

column name	description
transaction_id	Unique identifier of a transaction
creation_date	Date in which a transaction was created
credit_card_id	Encrypted ID associated with a credit card
booker_country	Country from which a person is trying to make a reservation
card_issuing_country	Country from which the credit card was issued
hotel_country	Country where the hotel the person is trying to book is located
language	Language in which the person who is creating the reservation is viewing the website
price_euro	Price of the transaction
book_window	Time difference in days between creation date of the transaction and the check-in date
ip_id	Encrypted ID associated with the IP the booker is using
hotel_id	ID related to a property (hotel or other kind of accommodation) that's available on Booking.com
length_of_stay	Time difference in days between checkin and checkout date (the number of nights booked at the hotel)
auth_result	Result obtained from the Payment Service Provider after attempting to process the payment.
property_age	Number of months since the property was registered in Booking.com, at the time the transaction was created
payment_method	Method used for payment
email_id	Encrypted ID associated with the email address used to make a transaction
email_domain	Domain of the email address used to make the transaction

	transaction_id	creation_date	ip_id	credit_card_id	booker_country	card_issuing_country	hotel_country	hotel_id
0	922278674	2022-01-07	e1fcad6672daa92bd6b5020fd954ff5a	0bb5ba2d7dd11cd0cf12f6db8fba25b0	Germany	Germany	Germany	90084
1	819686771	2022-01-22	5a71be20c6d46f2393be71d7199b607e	bb01aa67b20f28b8ce29e3a60d8ab276	Iceland	Iceland	Bulgaria	56417
2	755157875	2022-01-09	c25608c42c31aeafa1d7ad8012906b3c	09f5288f5bdd02b54c38825d1c29f5c9	Norway	Norway	Norway	22278
3	601281336	2022-01-17	e782033ba3a8dd86edd4f798579a0f92	288c1c58f5255956d78e2e526f9a2910	Spain	Spain	Italy	44946
4	152684295	2022-01-25	3895fc0fecfd3d4f31d7d708fc592098	95d0d8ec61b03bf7fd4250086313ab22	Germany	Germany	Norway	91658

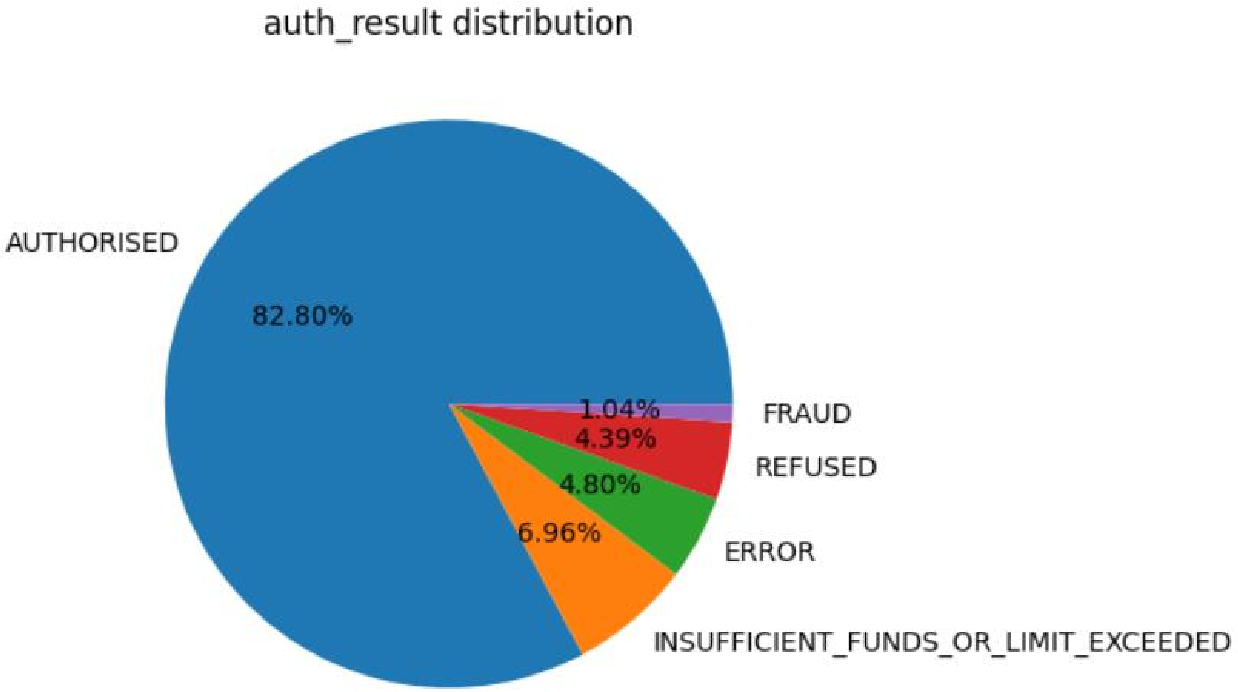
- structured data
- 79,557 records
- 16 predictive variables
- 1 target variable
- numerical, categorical and timestamp data types

Does our dataset contain any missing values?

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 79557 entries, 0 to 79556
Data columns (total 17 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   transaction_id       79557 non-null  int64
 1   creation_date        79557 non-null  object
 2   ip_id                79557 non-null  object
 3   credit_card_id       79557 non-null  object
 4   booker_country       79557 non-null  object
 5   card_issuing_country 79557 non-null  object
 6   hotel_country        79557 non-null  object
 7   hotel_id             79557 non-null  int64
 8   language             79557 non-null  object
 9   price_euro           79557 non-null  float64
 10  book_window          79557 non-null  int64
 11  length_of_stay       79557 non-null  int64
 12  email_id             79557 non-null  object
 13  email_domain         79557 non-null  object
 14  auth_result          79557 non-null  object
 15  property_age         79557 non-null  int64
 16  payment_method       79557 non-null  object
dtypes: float64(1), int64(5), object(11)
memory usage: 10.3+ MB
```

What is the distribution of our target variable?



Highly imbalanced data

Feature Engineering

Check each original feature one by one to decide which features we can keep, and which features can be used to derive new fearures.

column name	description	feature engineering
transaction_id	Unique identifier of a transaction	This works as an index for each transaction.
creation_date	Date in which a transaction was created	The time variable spans one month and can be converted into a relative time order.
credit_card_id	Encrypted ID associated with a credit card	Compute the number of times each credit card has been used.
booker_country	Country from which a person is trying to make a reservation	New variables can be derived from these four features. For each transaction, create indicators to check: Whether card_issuing_country = language. Whether card_issuing_country = booker_country. Whether booker_country = language.
card_issuing_country	Country from which the credit card was issued	
hotel_country	Country where the hotel the person is trying to book is located	
language	Language in which the person who is creating the reservation is viewing the website	
price_euro	Price of the transaction	Transform it to price per night
book_window	Time difference in days between creation date of the transaction and the check-in date	Keep it
ip_id	Encrypted ID associated with the IP the booker is using	Compute the number of times each ip has been used.
hotel_id	ID related to a property (hotel or other kind of accommodation) that's available on Booking.com	Compute the number of times each hotel has been used.
length_of_stay	Time difference in days between checkin and checkout date (the number of nights booked at the hotel)	Keep it
auth_result	Result obtained from the Payment Service Provider after attempting to process the payment.	Keep it
property_age	Number of months since the property was registered in Booking.com, at the time the transaction was created	Keep it
payment_method	Method used for payment	Keep it
email_id	Encrypted ID associated with the email address used to make a transaction	Compute the number of times each email has been used.
email_domain	Domain of the email address used to make the transaction	Keep it

Feature Engineering

column name	description	feature engineering
transaction_id	Unique identifier of a transaction	This works as an index for each transaction.
creation_date	Date in which a transaction was created	The time variable spans one month and can be converted into a relative time order.
credit_card_id	Encrypted ID associated with a credit card	Compute the number of times each credit card has been used.
booker_country	Country from which a person is trying to make a reservation	New variables can be derived from these four features. For each transaction, create indicators to check: Whether card_issuing_country = language. Whether card_issuing_country = booker_country. Whether booker_country = language.
card_issuing_country	Country from which the credit card was issued	
hotel_country	Country where the hotel the person is trying to book is located	
language	Language in which the person who is creating the reservation is viewing the website	
price_euro	Price of the transaction	Transform it to price per night
book_window	Time difference in days between creation date of the transaction and the check-in date	Keep it
ip_id	Encrypted ID associated with the IP the booker is using	Compute the number of times each ip has been used.
hotel_id	ID related to a property (hotel or other kind of accommodation) that's available on Booking.com	Compute the number of times each hotel has been used.
length_of_stay	Time difference in days between checkin and checkout date (the number of nights booked at the hotel)	Keep it
auth_result	Result obtained from the Payment Service Provider after attempting to process the payment.	Keep it
property_age	Number of months since the property was registered in Booking.com, at the time the transaction was created	Keep it
payment_method	Method used for payment	Keep it
email_id	Encrypted ID associated with the email address used to make a transaction	Compute the number of times each email has been used.
email_domain	Domain of the email address used to make the transaction	Keep it

	transaction_id	price_euro	length_of_stay	price_per_night
0	922278674	3478.16	5	579.693333
1	819686771	2476.98	13	176.927143
2	755157875	2855.06	17	158.614444
3	601281336	3780.31	17	210.017222
4	152684295	2115.12	13	151.080000
...	...	...	...	...
79552	807601027	4016.44	2	1338.813333
79553	294578496	4082.72	1	2041.360000
79554	762036796	4060.72	0	4060.720000
79555	221272161	3190.92	3	797.730000
79556	752996119	4479.39	2	1493.130000

Feature Engineering

column name	description	feature engineering
transaction_id	Unique identifier of a transaction	This works as an index for each transaction.
creation_date	Date in which a transaction was created	The time variable spans one month and can be converted into a relative time order.
credit_card_id	Encrypted ID associated with a credit card	Compute the number of times each credit card has been used.
booker_country	Country from which a person is trying to make a reservation	New variables can be derived from these four features. For each transaction, create indicators to check: Whether card_issuing_country = language. Whether card_issuing_country = booker_country. Whether booker_country = language.
card_issuing_country	Country from which the credit card was issued	
hotel_country	Country where the hotel the person is trying to book is located	
language	Language in which the person who is creating the reservation is viewing the website	
price_euro	Price of the transaction	Transform it to price per night
book_window	Time difference in days between creation date of the transaction and the check-in date	Keep it
ip_id	Encrypted ID associated with the IP the booker is using	Compute the number of times each ip has been used.
hotel_id	ID related to a property (hotel or other kind of accommodation) that's available on Booking.com	Compute the number of times each hotel has been used.
length_of_stay	Time difference in days between checkin and checkout date (the number of nights booked at the hotel)	Keep it
auth_result	Result obtained from the Payment Service Provider after attempting to process the payment.	Keep it
property_age	Number of months since the property was registered in Booking.com, at the time the transaction was created	Keep it
payment_method	Method used for payment	Keep it
email_id	Encrypted ID associated with the email address used to make a transaction	Compute the number of times each email has been used.
email_domain	Domain of the email address used to make the transaction	Keep it

	transaction_id	card_issuing_country	language	language_flag1
0	922278674	Germany	Germany	0
1	819686771	Iceland	Iceland	0
2	755157875	Norway	Slovakia	1
3	601281336	Spain	Spain	0
4	152684295	Germany	Germany	0
...	...	...	...	...
79552	807601027	Malta	Italy	1
79553	294578496	Malta	Italy	1
79554	762036796	Malta	Italy	1
79555	221272161	Malta	Italy	1
79556	752996119	Malta	Italy	1

Feature Engineering

column name	description	feature engineering
transaction_id	Unique identifier of a transaction	This works as an index for each transaction.
creation_date	Date in which a transaction was created	The time variable spans one month and can be converted into a relative time order.
credit_card_id	Encrypted ID associated with a credit card	Compute the number of times each credit card has been used.
booker_country	Country from which a person is trying to make a reservation	New variables can be derived from these four features. For each transaction, create indicators to check: Whether card_issuing_country = language. Whether card_issuing_country = booker_country. Whether booker_country = language.
card_issuing_country	Country from which the credit card was issued	
hotel_country	Country where the hotel the person is trying to book is located	
language	Language in which the person who is creating the reservation is viewing the website	
price_euro	Price of the transaction	Transform it to price per night
book_window	Time difference in days between creation date of the transaction and the check-in date	Keep it
ip_id	Encrypted ID associated with the IP the booker is using	Compute the number of times each ip has been used.
hotel_id	ID related to a property (hotel or other kind of accommodation) that's available on Booking.com	Compute the number of times each hotel has been used.
length_of_stay	Time difference in days between checkin and checkout date (the number of nights booked at the hotel)	Keep it
auth_result	Result obtained from the Payment Service Provider after attempting to process the payment.	Keep it
property_age	Number of months since the property was registered in Booking.com, at the time the transaction was created	Keep it
payment_method	Method used for payment	Keep it
email_id	Encrypted ID associated with the email address used to make a transaction	Compute the number of times each email has been used.
email_domain	Domain of the email address used to make the transaction	Keep it

	credit_card_id	transaction_id	days_since_first	auth_result	credit_card_id_count
77378	5e3a0db720f17bdf1a4ec149cdeab0e6	755215803	10	REFUSED	412
77961	5e3a0db720f17bdf1a4ec149cdeab0e6	606068615	10	REFUSED	412
78422	5e3a0db720f17bdf1a4ec149cdeab0e6	432254710	9	REFUSED	412
77180	5e3a0db720f17bdf1a4ec149cdeab0e6	602565752	9	AUTHORISED	412
78124	5e3a0db720f17bdf1a4ec149cdeab0e6	982820092	10	REFUSED	412
78048	5e3a0db720f17bdf1a4ec149cdeab0e6	140542727	9	REFUSED	412
77987	5e3a0db720f17bdf1a4ec149cdeab0e6	692372724	9	AUTHORISED	412
78418	5e3a0db720f17bdf1a4ec149cdeab0e6	745075293	10	REFUSED	412
77536	5e3a0db720f17bdf1a4ec149cdeab0e6	223752309	9	REFUSED	412
77186	5e3a0db720f17bdf1a4ec149cdeab0e6	778217924	9	REFUSED	412
77866	5e3a0db720f17bdf1a4ec149cdeab0e6	633261294	10	REFUSED	412
78416	5e3a0db720f17bdf1a4ec149cdeab0e6	262134043	10	REFUSED	412
77189	5e3a0db720f17bdf1a4ec149cdeab0e6	863910767	9	REFUSED	412
77190	5e3a0db720f17bdf1a4ec149cdeab0e6	228562351	9	AUTHORISED	412
78414	5e3a0db720f17bdf1a4ec149cdeab0e6	555542636	10	REFUSED	412
77192	5e3a0db720f17bdf1a4ec149cdeab0e6	984191971	10	AUTHORISED	412
78412	5e3a0db720f17bdf1a4ec149cdeab0e6	415895389	9	REFUSED	412
77714	5e3a0db720f17bdf1a4ec149cdeab0e6	405270490	10	REFUSED	412
78411	5e3a0db720f17bdf1a4ec149cdeab0e6	497911184	9	AUTHORISED	412
77196	5e3a0db720f17bdf1a4ec149cdeab0e6	172199164	9	REFUSED	412

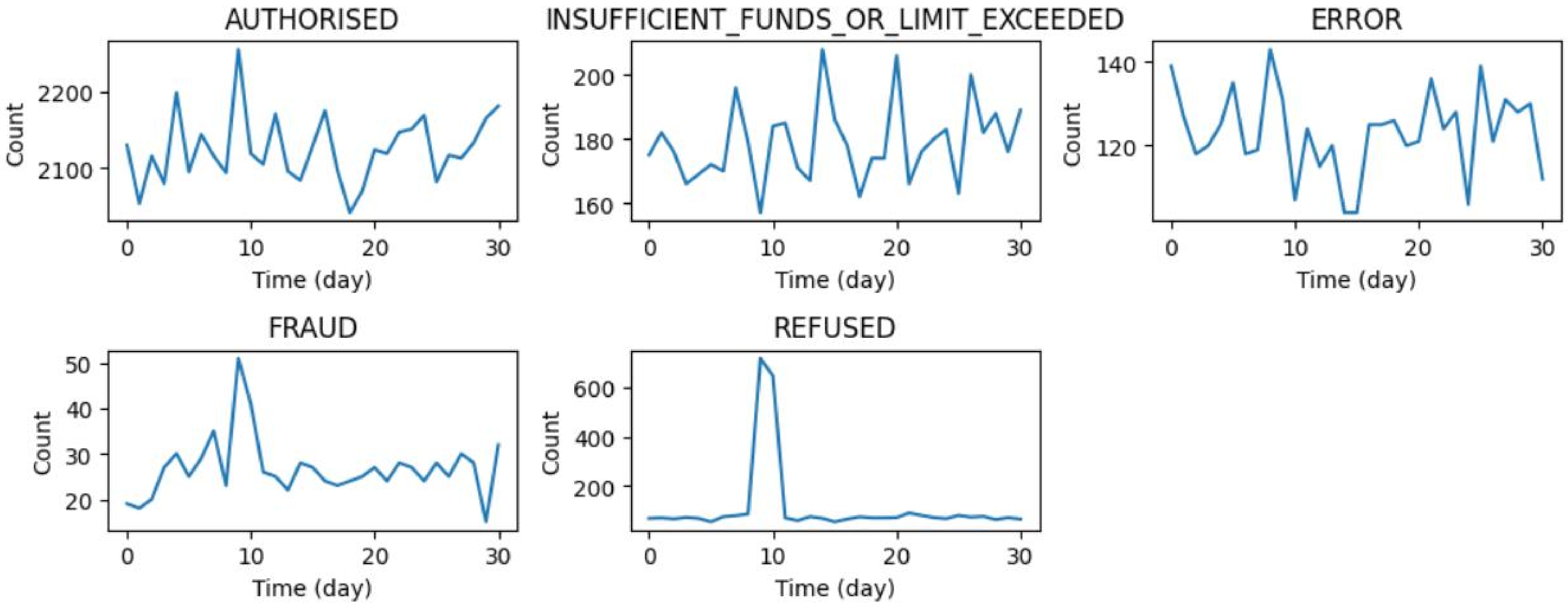
The outcome of feature engineering

	transaction_id	ip_id	credit_card_id	booker_country	card_issuing_country	hotel_country	hotel_id	language
0	922278674	e1fcad6672daa92bd6b5020fd954ff5a	0bb5ba2d7dd11cd0cf12f6db8fba25b0	Germany	Germany	Germany	90084	Germany
1	819686771	5a71be20c6d46f2393be71d7199b607e	bb01aa67b20f28b8ce29e3a60d8ab276	Iceland	Iceland	Bulgaria	56417	Iceland
2	755157875	c25608c42c31aeafa1d7ad8012906b3c	09f5288f5bdd02b54c38825d1c29f5c9	Norway	Norway	Norway	22278	Slovakia
3	601281336	e782033ba3a8dd86edd4f798579a0f92	288c1c58f5255956d78e2e526f9a2910	Spain	Spain	Italy	44946	Spain
4	152684295	3895fc0fecfd3d4f31d7d708fc592098	95d0d8ec61b03bf7fd4250086313ab22	Germany	Germany	Norway	91658	Germany

price_euro	book_window	...	payment_method	price_per_night	days_since_first	credit_card_id_count	ip_id_count	email_id_count	hotel_id_count	language_flag1	langu
3478.16	89	...	mastercard	579.693333	6	1	1	1	3	0	
2476.98	30	...	visa	176.927143	21	1	1	1	1	0	
2855.06	68	...	mastercard	158.614444	8	1	1	1	1	1	
3780.31	11	...	visa	210.017222	16	1	1	1	1	0	
2115.12	64	...	paypal	151.080000	24	1	1	1	1	0	

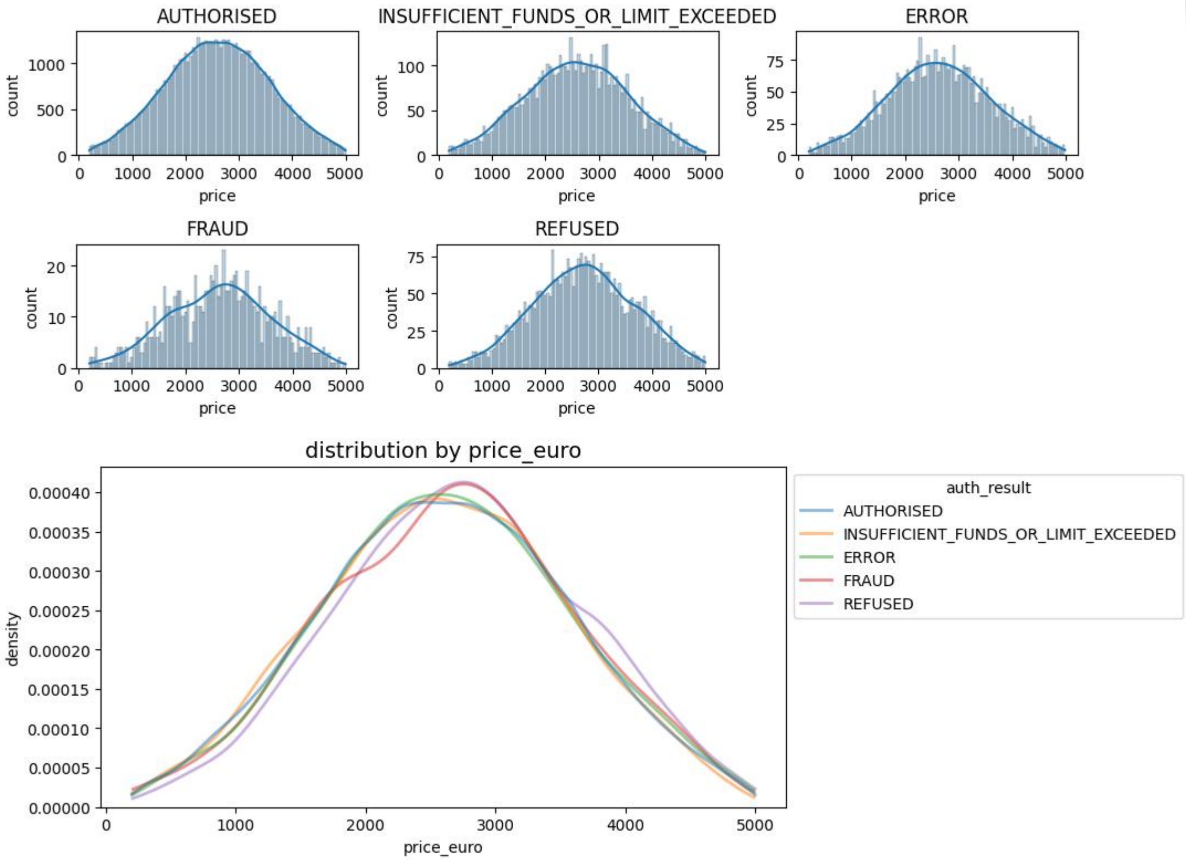
Feature engineering can enrich our data, such that we give the analytical or ML models more opportunities to find hidden fraudulent patterns.

What is the distribution or trend of our predictive variables?

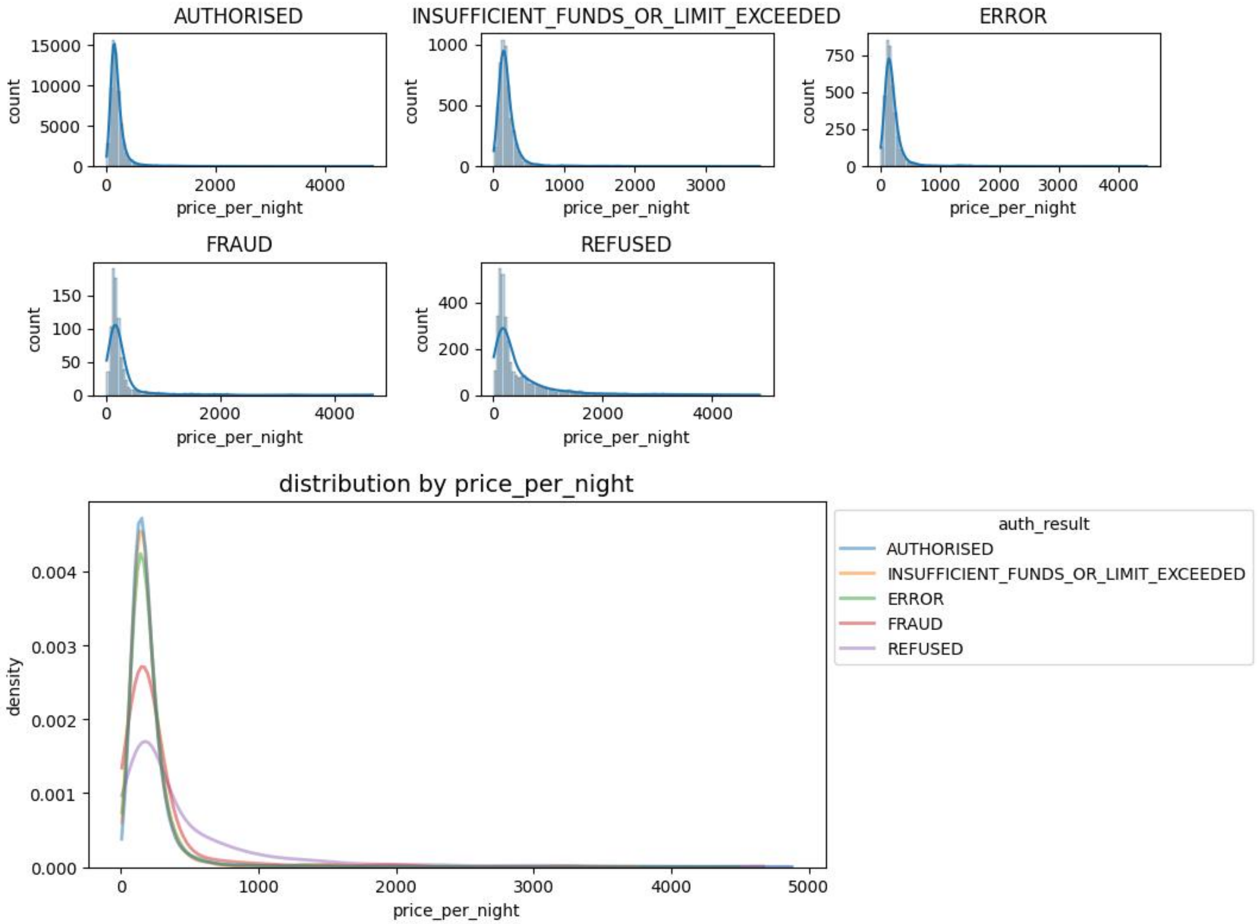


- Daily transaction number of each target group.
- There is a sudden spike in transaction number around day 9 for AUTHORISED, FRAUD and REFUSED groups. The other two goups do not exhibit any unusual patterns.

What is the distribution or trend of our predictive variables?



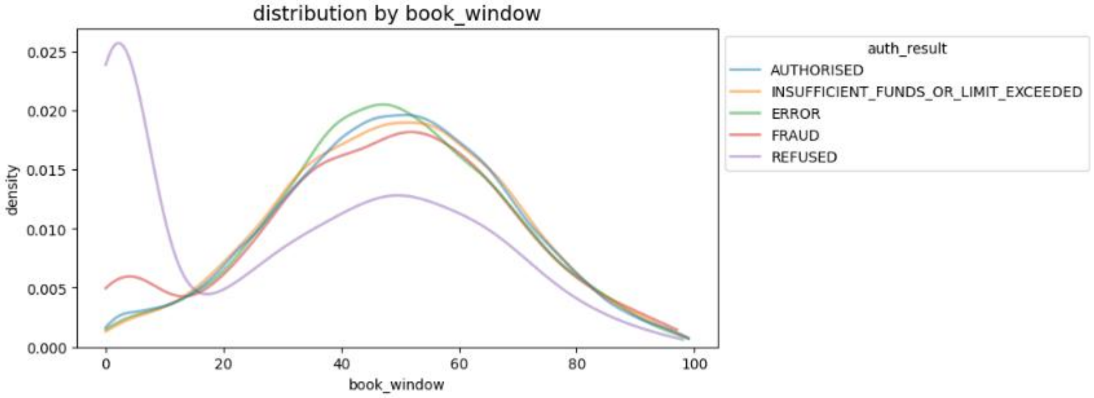
•price distribution for each group.

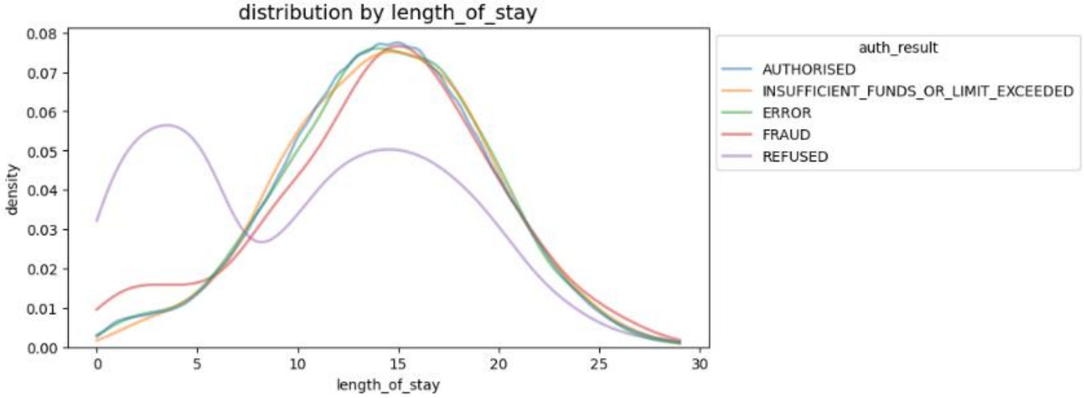


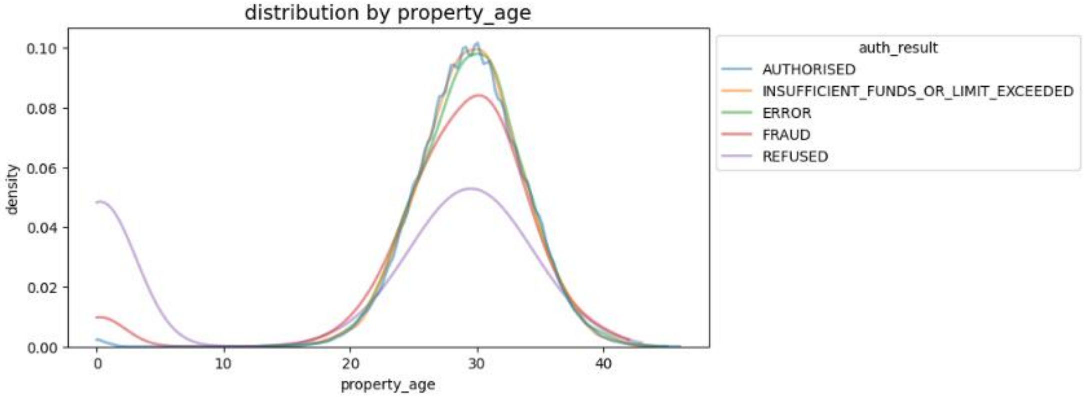
• price_per_night distribution for each group.

• There is almost no difference across different groups in terms of the total price.

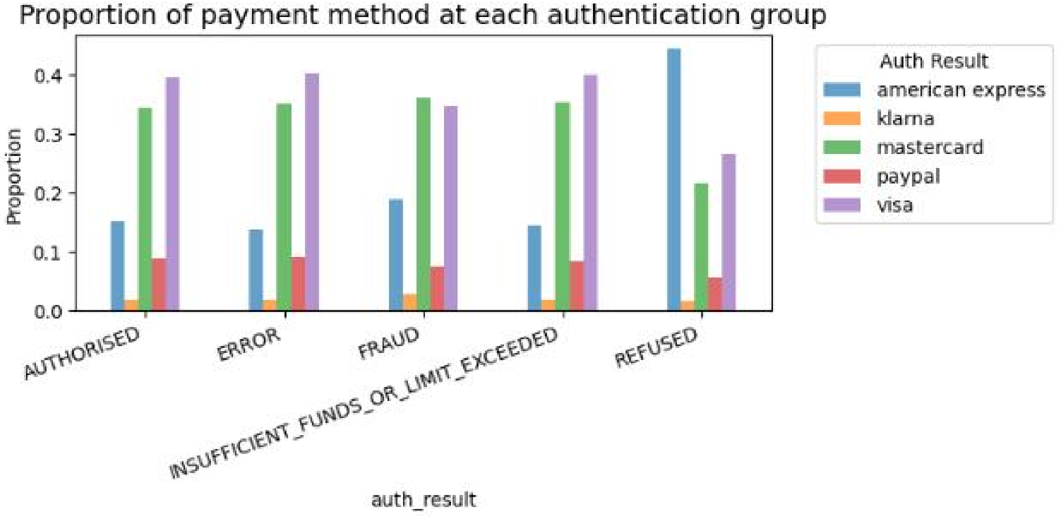
• For price_per_night, the density curve for Fraud and Refused group is slightly different with other groups, and expensive transactions are likely to be made in these two groups.

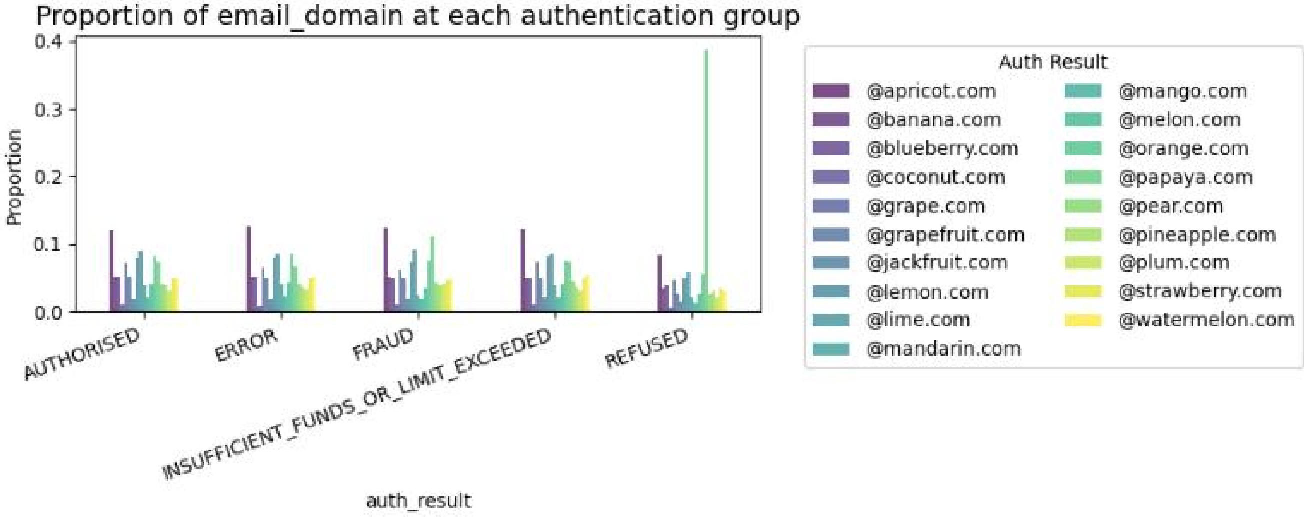






Transactions in the FRAUD and REFUSED groups tend to be made with very short book_window, short length_of_stay, and smaller property_age, indicating that these transactions are often made in a rush.

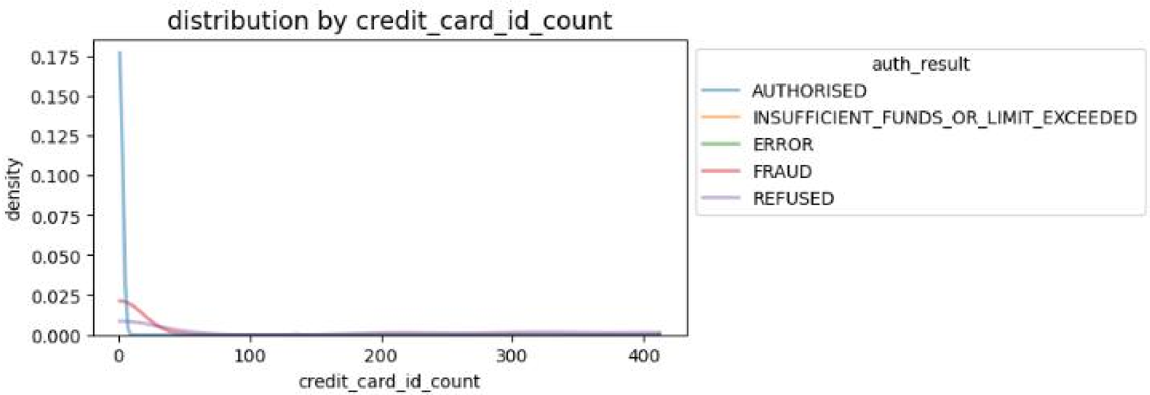


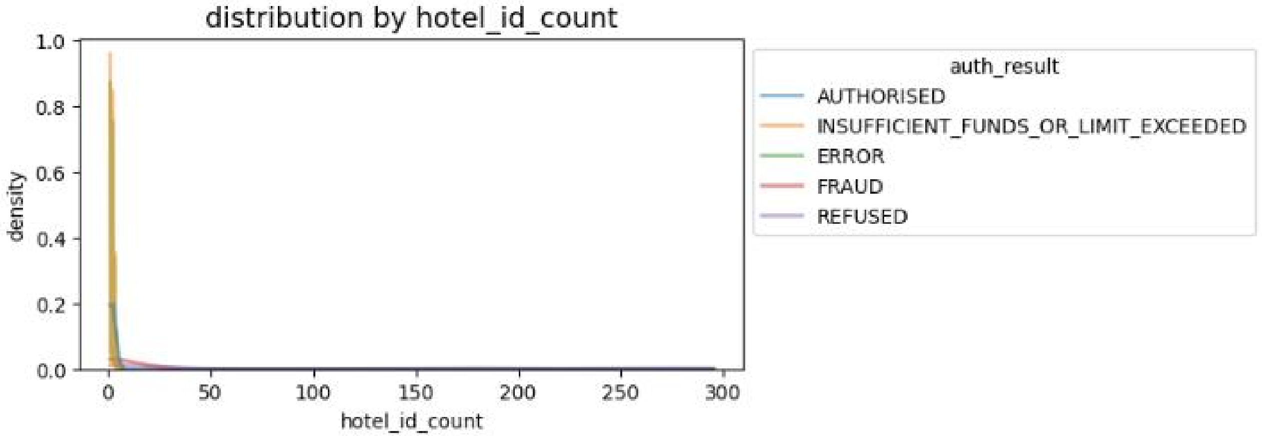


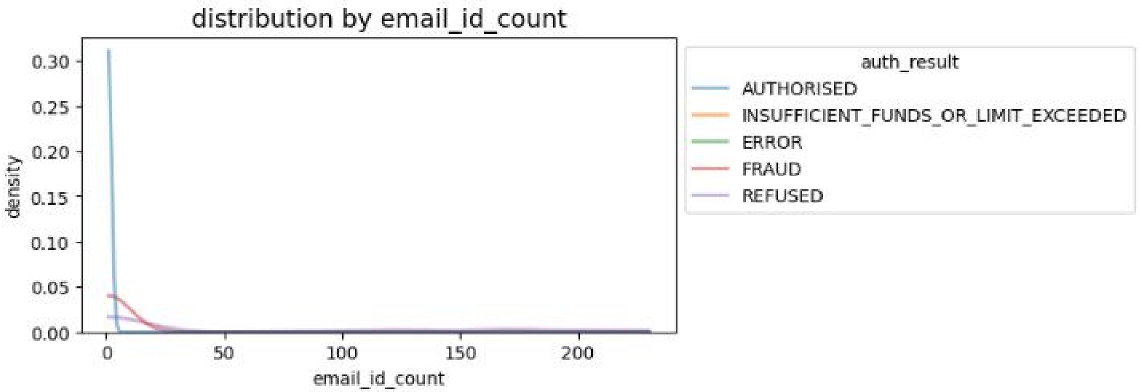
- When we look at payment method, Refused group like to use american express. However other 4 groups like to use visa and mastercard.

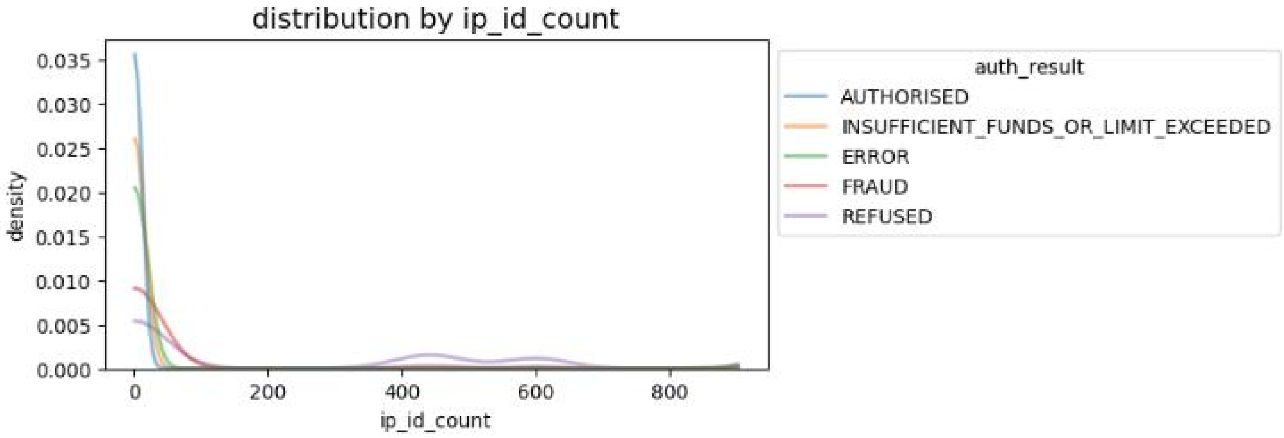
- When we look at email_domain, usage of @papaya.com by Refuse group is extremely high.

What is the distribution or trend of our predictive variables?





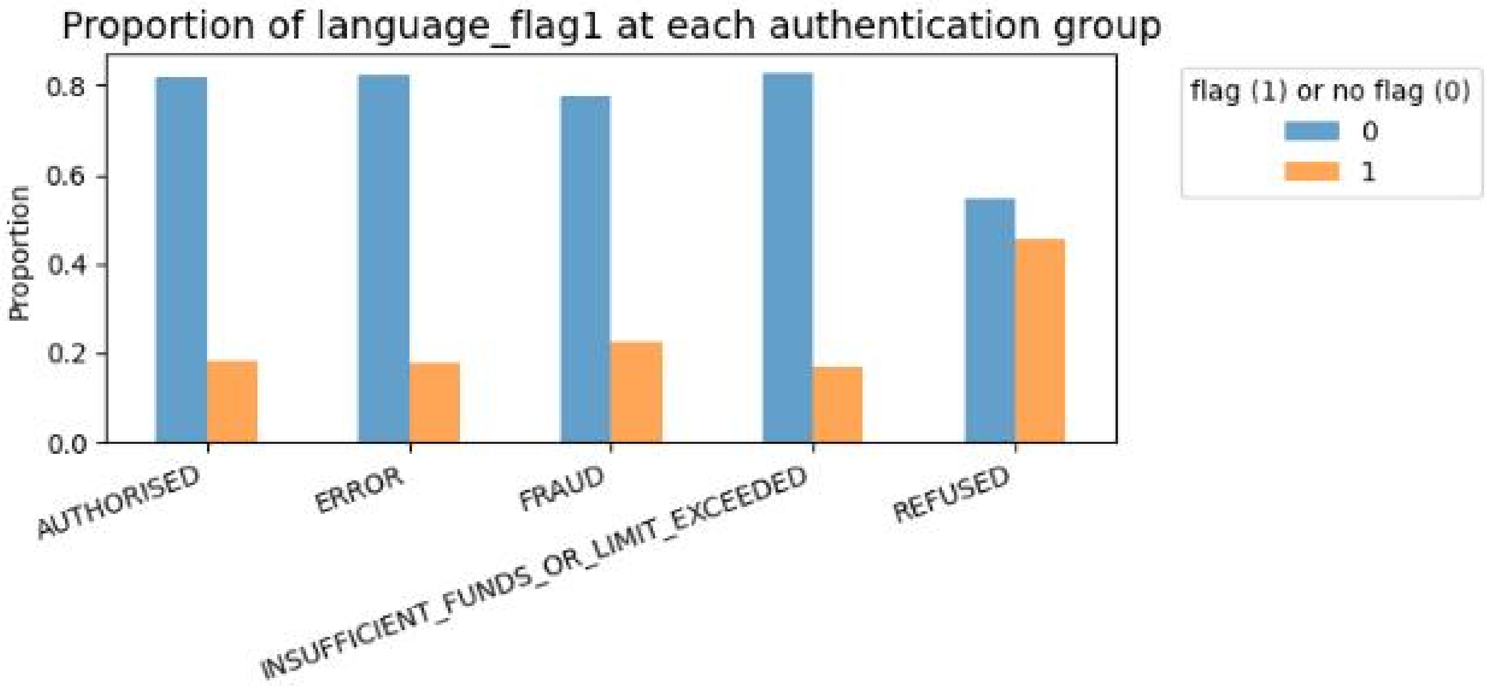




- Counted how many times each credit card, email address, IP address, and hotel ID was used and plotted their distributions for each target group.

- In the FRAUD and REFUSED groups, certain emails, hotels, IPs, and credit cards tend to be used quite often.

What is the distribution or trend of our predictive variables?



- Intuitively, when making a transaction, the language used should match the language associated with the card's issuing country; otherwise, it may trigger a flag (language_flag1).

- Our intuition holds true for most groups, but it does not apply to the REFUSED group.

What we can get from data processing:

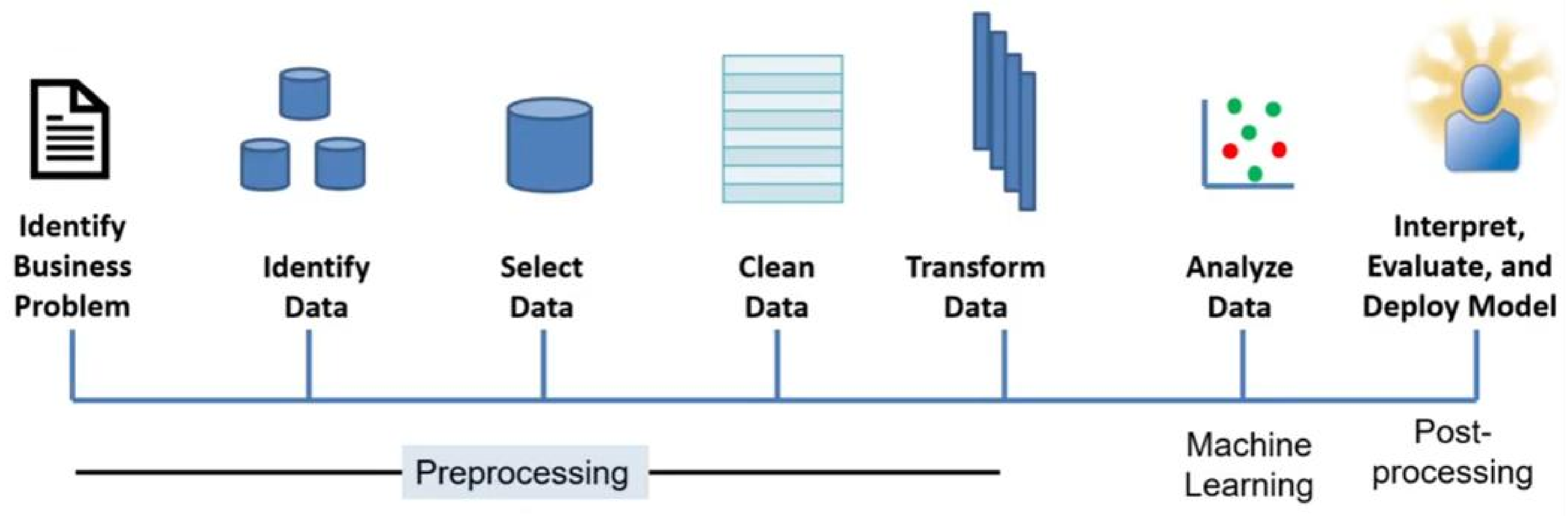
1. The derived new features appear to be effective in differentiating different groups.
(price_per_night, language_flag1, language_flag2, language_flag3, ip_id_count, hotel_id_count, credit_card_id_count, email_id_count)

2. The groups 'AUTHORISED', 'INSUFFICIENT_FUNDS_OR_LIMIT_EXCEEDED', and 'ERROR' have nearly identical patterns. This suggests that it may be reasonable to combine them into a single legitimate group.

3.The 'FRAUD' group shows some deviations in variable distributions compared to 'AUTHORISED', 'INSUFFICIENT_FUNDS_OR_LIMIT_EXCEEDED', and 'ERROR', but these differences are subtle. This highlights the difficulty and challenge of detecting 'FRAUD' transactions from the legitimate ones. We may not be able to achieve high detection accuracy with the current dataset.

4. The 'REFUSED' group displays distinct patterns compared with the other four groups. Transactions in this group are more likely to represent anomalies in the dataset than those in the 'FRAUD' group.

5. The outlier behaviour
(1). The payment tends to be expensive in terms of pirce per night
(2). The book_windown tends to be small
(4). The length of stay tends to be small
(5). The transaction tends to go the new hotel whose property_age is close to 0
(6). Payment methord “american express” likely to be used
(7). Email domain “papaya.com” tends to used.
(8). A card/email/ip/hotel tends to be used largely
(9). There is a mismatch between card_issing_country and language



Building detetion model:

Anomaly detection:

Hard-code method (Test each feature one by one)

Unsupervised learning (Isolation forest, local outlier factor)

Classification:

Supervised learning (logistic regression, randomforest, gradient boosting)

		Method	f1-score	precision	recall	roc-auc	auc
anomaly detection	hard-code	book_window	0.38	0.54	0.29	0.23	0.64
		length_of_stay	0.27	0.25	0.28	0.15	0.63
		property_age	0.43	0.84	0.29	0.31	0.65
		credit_card_id_count	0.43	0.84	0.29	0.28	0.64
		ip_id_count	0.38	0.53	0.29	0.17	0.64
		email_id_count	0.42	0.77	0.29	0.27	0.64
		hotel_id_count	0.42	0.77	0.29	0.3	0.65
		language_flag1	0.18	0.12	0.41	0.08	0.62
		language_flag2	0.23	0.17	0.35	0.09	0.63
		language_flag3	0.23	0.17	0.36	0.1	0.63
	unsupervised learning	isolation_forest	0.42	0.78	0.29	0.32	0.64
		local_outlier_factor	0.11	0.06	0.84	0.06	0.51
classification	supervised learning	logistic_regression	0.42	0.84	0.28	0.31	0.65
		random_forest_classifier	0.42	0.83	0.28	0.31	0.64
		gradient_boosting_classifier	0.42	0.83	0.28	0.31	0.65

Maximizing fraud detection:

Prioritize recall

Local_outlier_factor

Any mismatch between (card_issuing_country, language), (card_issuing_country, booker_country), or (booker_country, language) should trigger a flag.

Minimizing harm to legitimate customers:

Prioritize precision.

property_age and credit_card_id_count (a frequently used credit card or a booking for a newly listed hotel may trigger a flag)

Supervised learning

Balancing both false positives and false negatives:

F1-score or ROC-AUC.

In this scenario, ensemble learning methods such as Isolation Forest can help detect outliers, while features like property_age and credit_card_id_count remain important for risk assessment.

Mitigation

1. Correct the label of our data

	credit_card_id	transaction_id	days_since_first	auth_result	credit_card_id_count
77378	5e3a0db720f17bdf1a4ec149cdeab0e6	755215803	10	REFUSED	412
77961	5e3a0db720f17bdf1a4ec149cdeab0e6	606068615	10	REFUSED	412
78422	5e3a0db720f17bdf1a4ec149cdeab0e6	432254710	9	REFUSED	412
77180	5e3a0db720f17bdf1a4ec149cdeab0e6	602565752	9	AUTHORISED	412
78124	5e3a0db720f17bdf1a4ec149cdeab0e6	982820092	10	REFUSED	412
78048	5e3a0db720f17bdf1a4ec149cdeab0e6	140542727	9	REFUSED	412
77987	5e3a0db720f17bdf1a4ec149cdeab0e6	692372724	9	AUTHORISED	412
78418	5e3a0db720f17bdf1a4ec149cdeab0e6	745075293	10	REFUSED	412
77536	5e3a0db720f17bdf1a4ec149cdeab0e6	223752309	9	REFUSED	412
77186	5e3a0db720f17bdf1a4ec149cdeab0e6	778217924	9	REFUSED	412
77866	5e3a0db720f17bdf1a4ec149cdeab0e6	633261294	10	REFUSED	412
78416	5e3a0db720f17bdf1a4ec149cdeab0e6	262134043	10	REFUSED	412
77189	5e3a0db720f17bdf1a4ec149cdeab0e6	863910767	9	REFUSED	412
77190	5e3a0db720f17bdf1a4ec149cdeab0e6	228562351	9	AUTHORISED	412
78414	5e3a0db720f17bdf1a4ec149cdeab0e6	555542636	10	REFUSED	412
77192	5e3a0db720f17bdf1a4ec149cdeab0e6	984191971	10	AUTHORISED	412
78412	5e3a0db720f17bdf1a4ec149cdeab0e6	415895389	9	REFUSED	412
77714	5e3a0db720f17bdf1a4ec149cdeab0e6	405270490	10	REFUSED	412
78411	5e3a0db720f17bdf1a4ec149cdeab0e6	497911184	9	AUTHORISED	412
77196	5e3a0db720f17bdf1a4ec149cdeab0e6	172199164	9	REFUSED	412

Mitigation:

1. Correct the label of our data
2. Some suspicious card, hotel, ip, and email should be blocked.

ip_id	credit_card_id	booker_country	card_issuing_country	hotel_country	hotel_id	language	price_euro	book_window	...	payment_method
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	98294	Denmark	3488.54	3	...	american express
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Germany	Iceland	Norway	70304	Denmark	2505.04	4	...	american express
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	42442	Denmark	1513.53	3	...	american express
70d735	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	42442	Denmark	1952.21	3	...	american express
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	70304	Denmark	3694.75	3	...	american express
...	...	...	...	...	...	...	...	...	...	...
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	70304	Denmark	2434.48	4	...	american express
6cde52	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	98294	Denmark	2522.79	2	...	american express
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	70304	Spain	1831.31	4	...	american express
154d03	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	70304	Malta	1344.46	4	...	american express
70d735	5e3a0db720f17bdf1a4ec149cdeab0e6	Spain	Iceland	Norway	43013	Denmark	4057.96	3	...	american express

Mitigation:

1. Correct the label of our data

2. For mitigating fraud, some suspicious card, hotel, email should be blocked.

3. Special investigative emphasis should be placed on newly onboarded hotels, as they may present higher risk due to limited historical data and evolving transaction patterns

Thank you!

	book_window	length_of_stay	property_age	credit_card_id_count	ip_id_count	email_id_count	hotel_id_count	language_flag1	language_flag2	language_flag
f1-score	0.38	0.27	0.43	0.43	0.38	0.42	0.42	0.18	0.23	0.2
precision	0.54	0.25	0.84	0.84	0.53	0.77	0.77	0.12	0.17	0.1
recall	0.29	0.28	0.29	0.29	0.29	0.29	0.29	0.41	0.35	0.3
roc-auc	0.23	0.15	0.31	0.28	0.17	0.27	0.30	0.08	0.09	0.1
auc	0.64	0.63	0.65	0.64	0.64	0.64	0.65	0.62	0.63	0.6