

**FINAL TASK**

# **ID/X-PARTNER-RAKAMIN**



**IBRAHIM ERLANGGA**



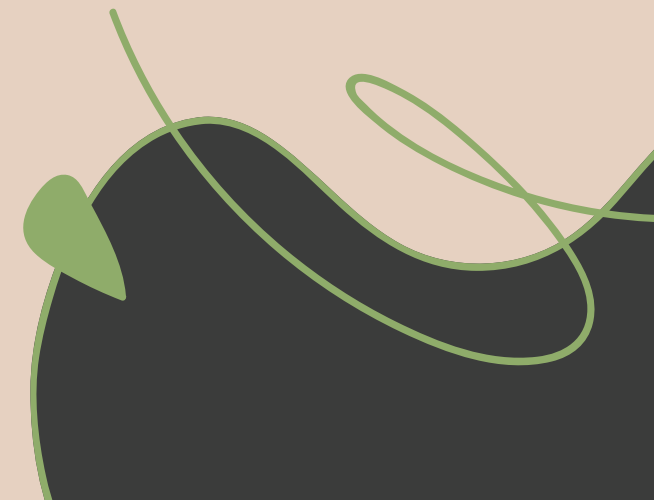
# BUSSINES UNDERSTANDING

Manajemen risiko kredit adalah mengelola risiko kredit. Risiko kredit adalah suatu risiko kerugian yang disebabkan oleh ketidakmampuan (gagal bayar) dari debitur atas kewajiban oembayaran utangnya, baik utang pokok maupun bunga, ataupun keduanya.

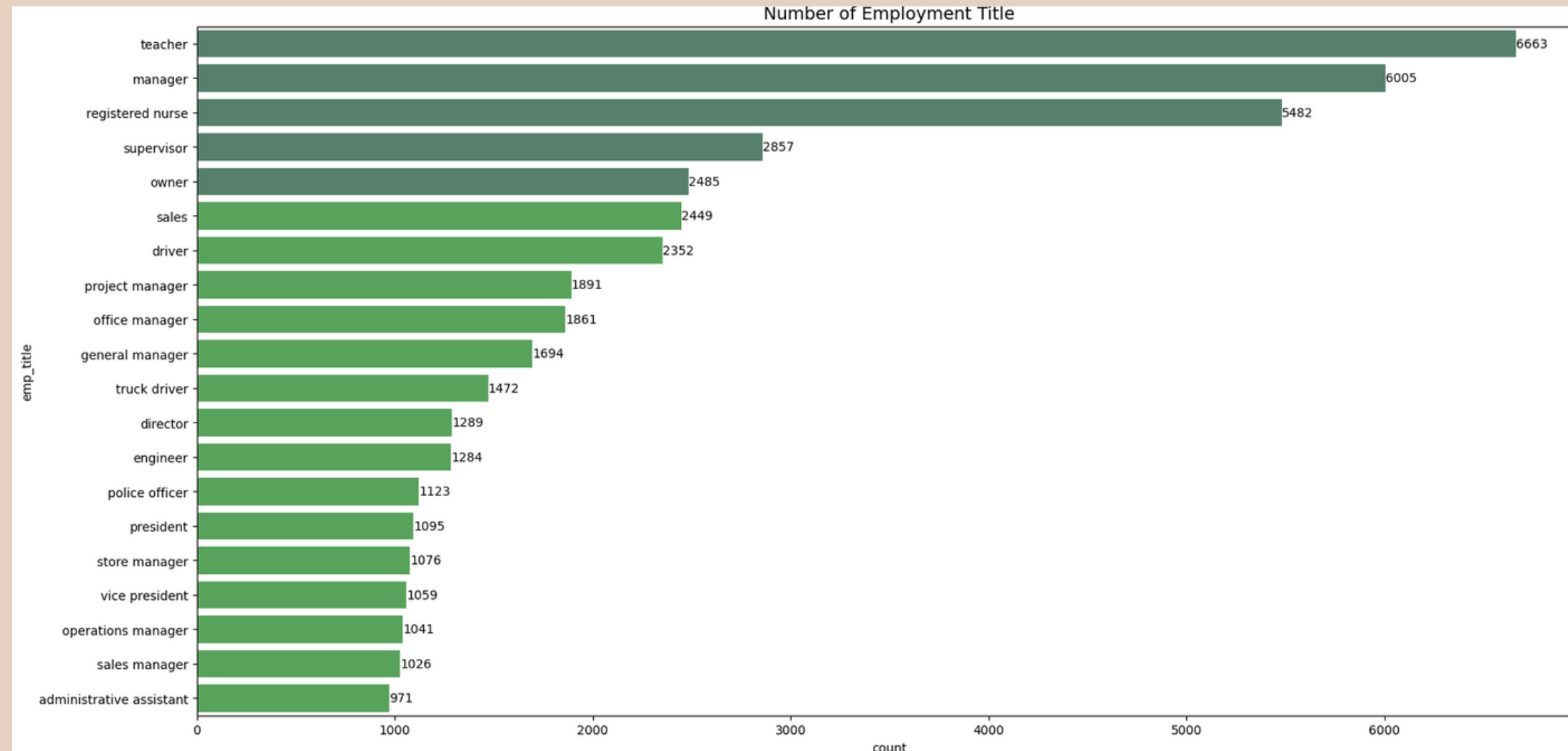
Menurut para ahli, risiko kredit didefinisikan sebagai risiko kerugian sehubungan dengan pihak peminjam (counterparty) tidak dapat dan tidak mau memenuhi kewajiban untuk membayar kembali dana yang dipinjamnya secara penuh pada saat jatuh tempo atau sesudahnya (Ferry dan Sugiarto)



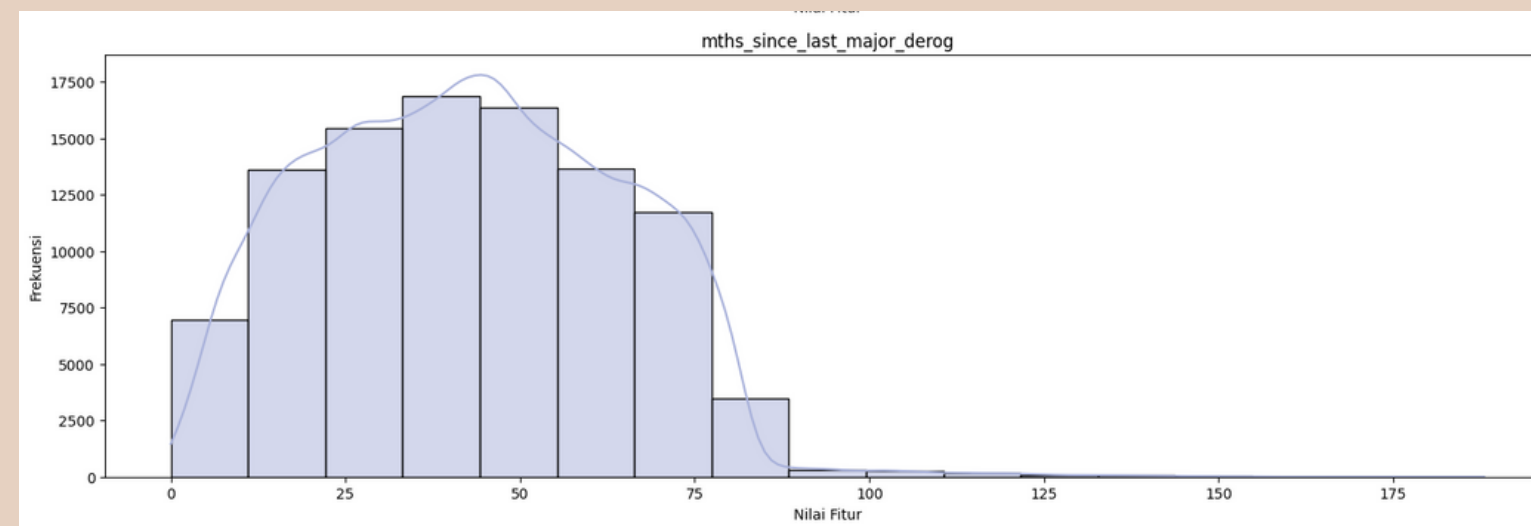
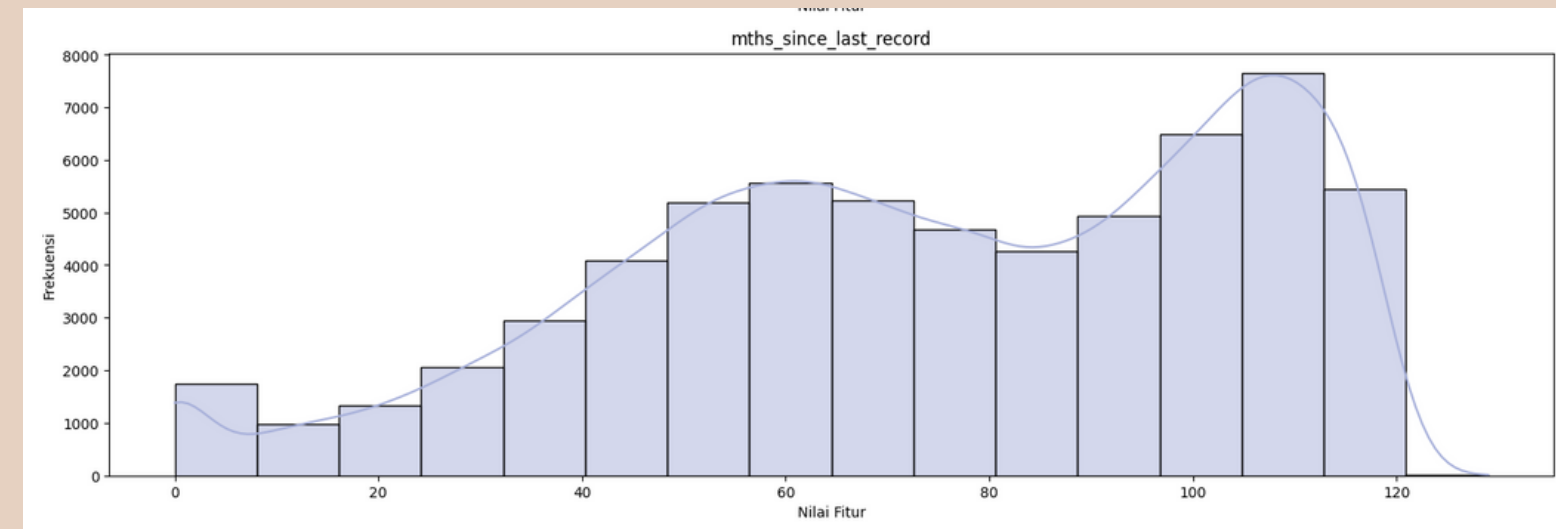
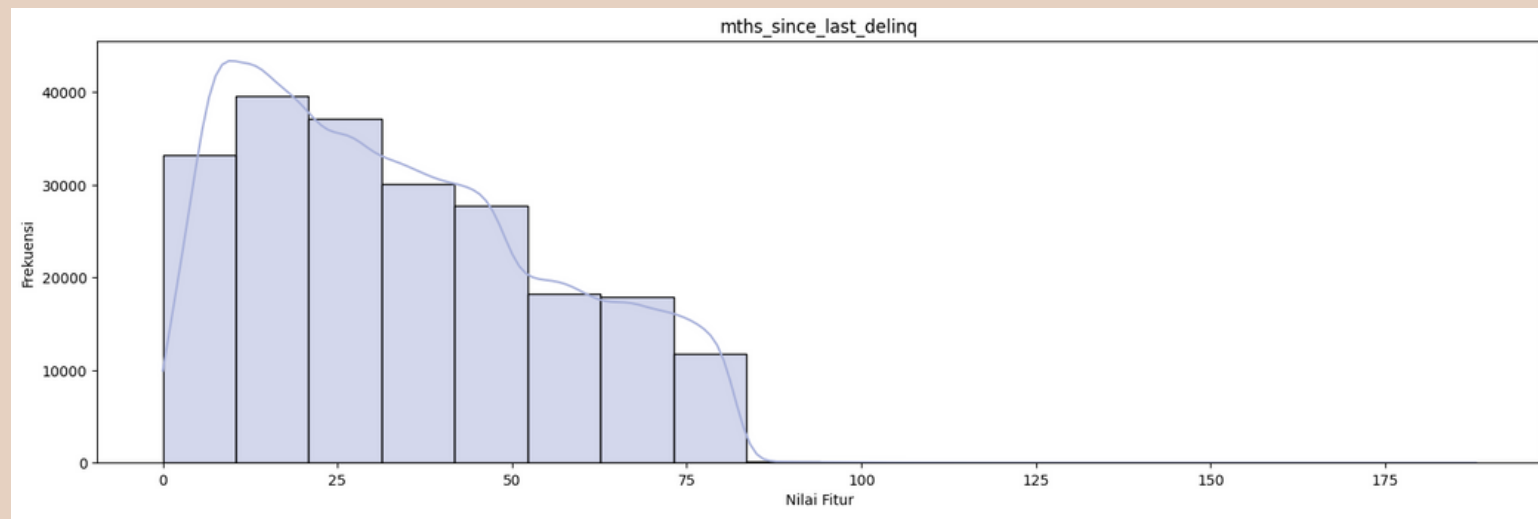
R I S K



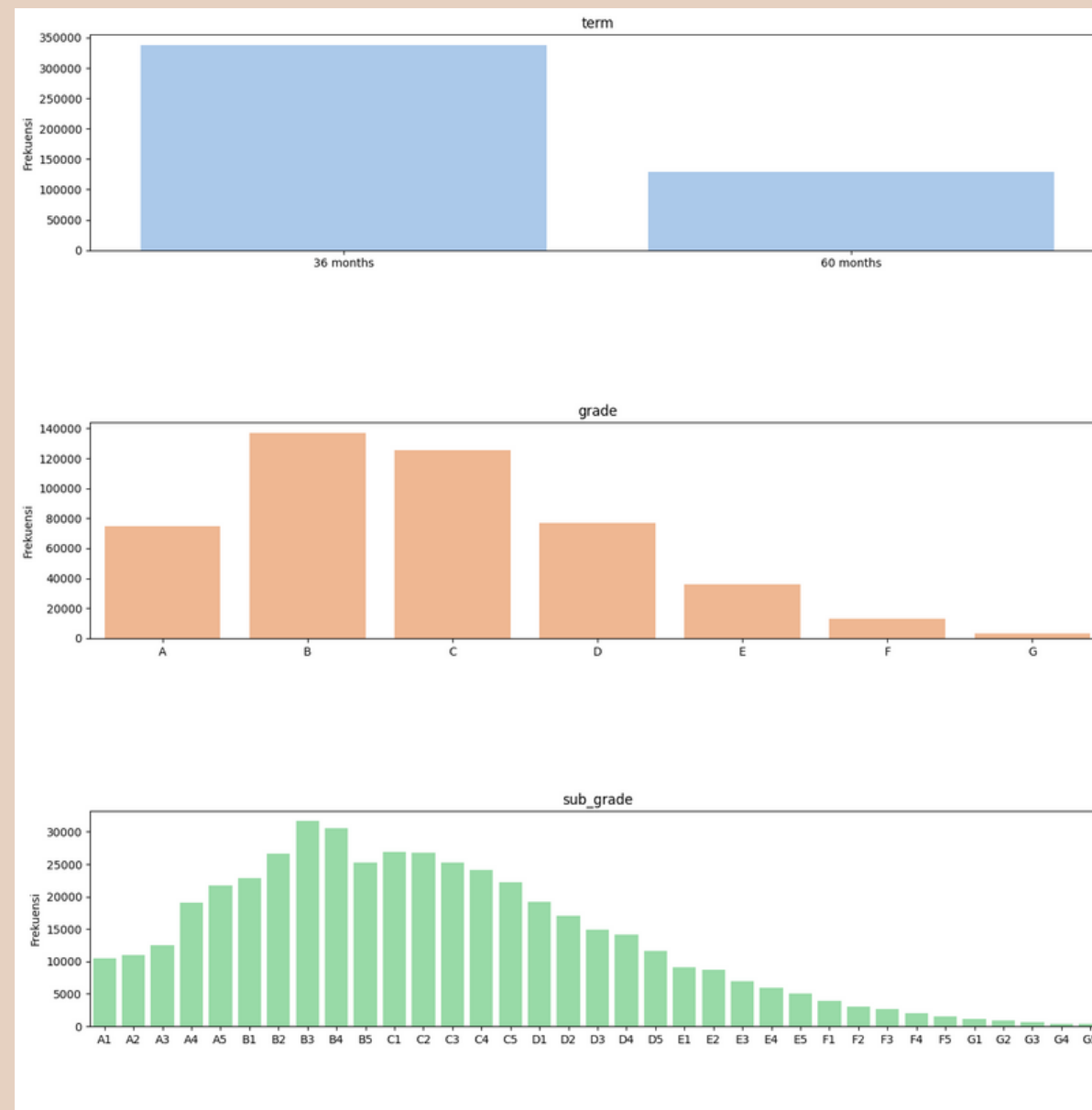
# EDA BEFORE DROPPING emp\_title



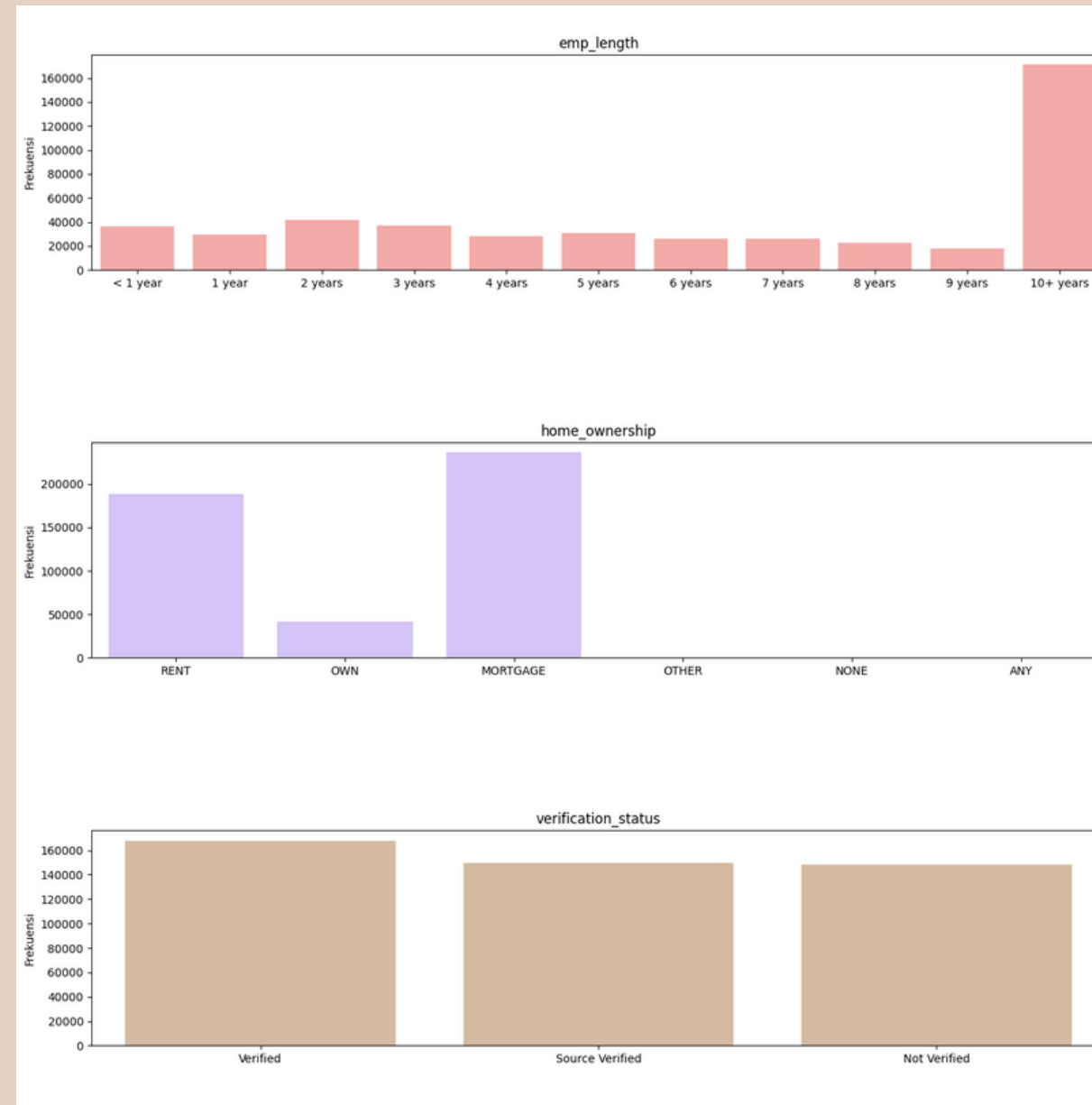
# EDA BEFORE DROPPING NUMERICAL DISTRIBUTION



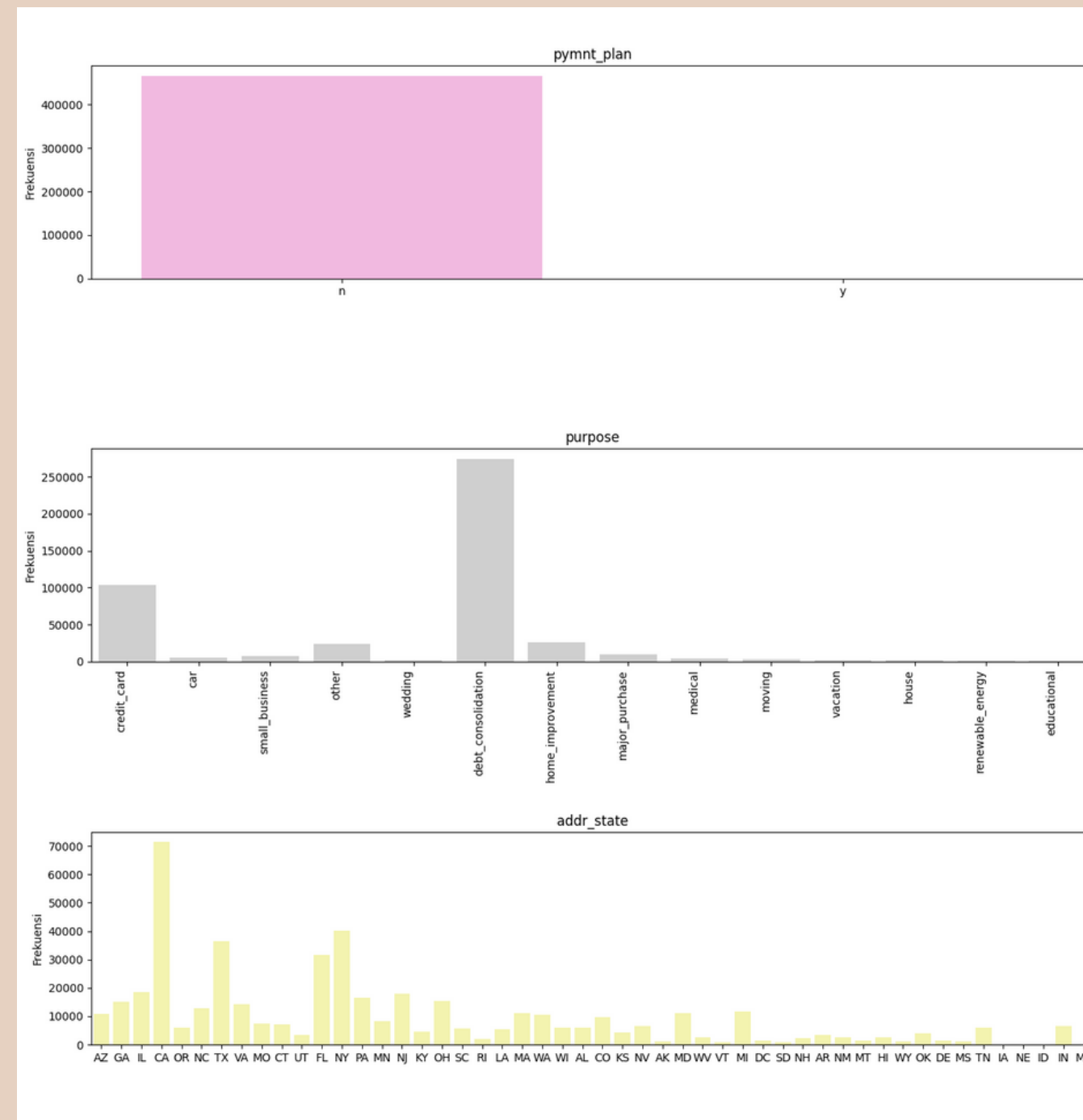
# EDA AFTER DROPPING ALL CATEGORICAL BAR CHART



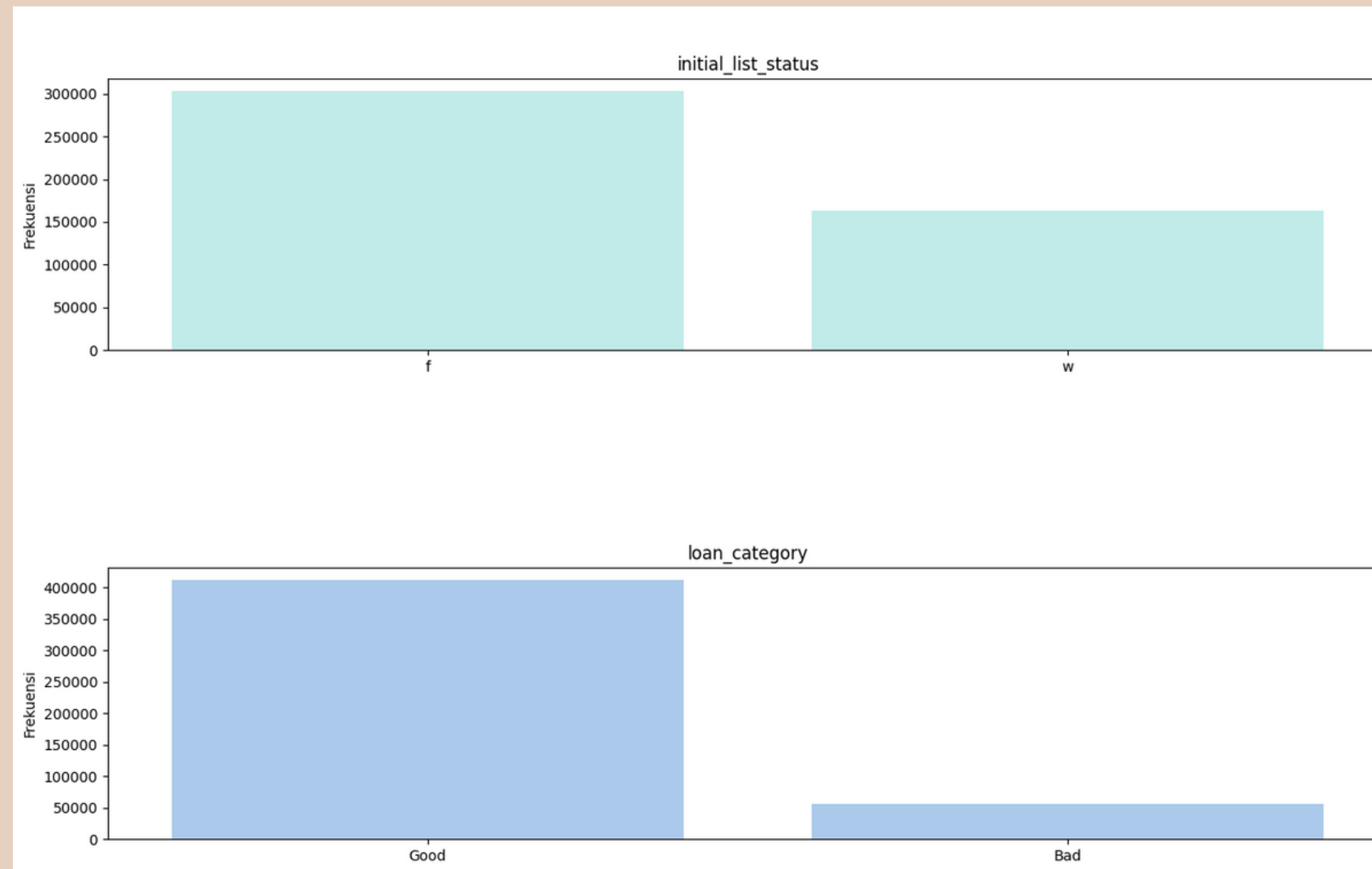
# EDA AFTER DROPPING ALL CATEGORICAL BAR CHART



# EDA AFTER DROPPING ALL CATEGORICAL BAR CHART



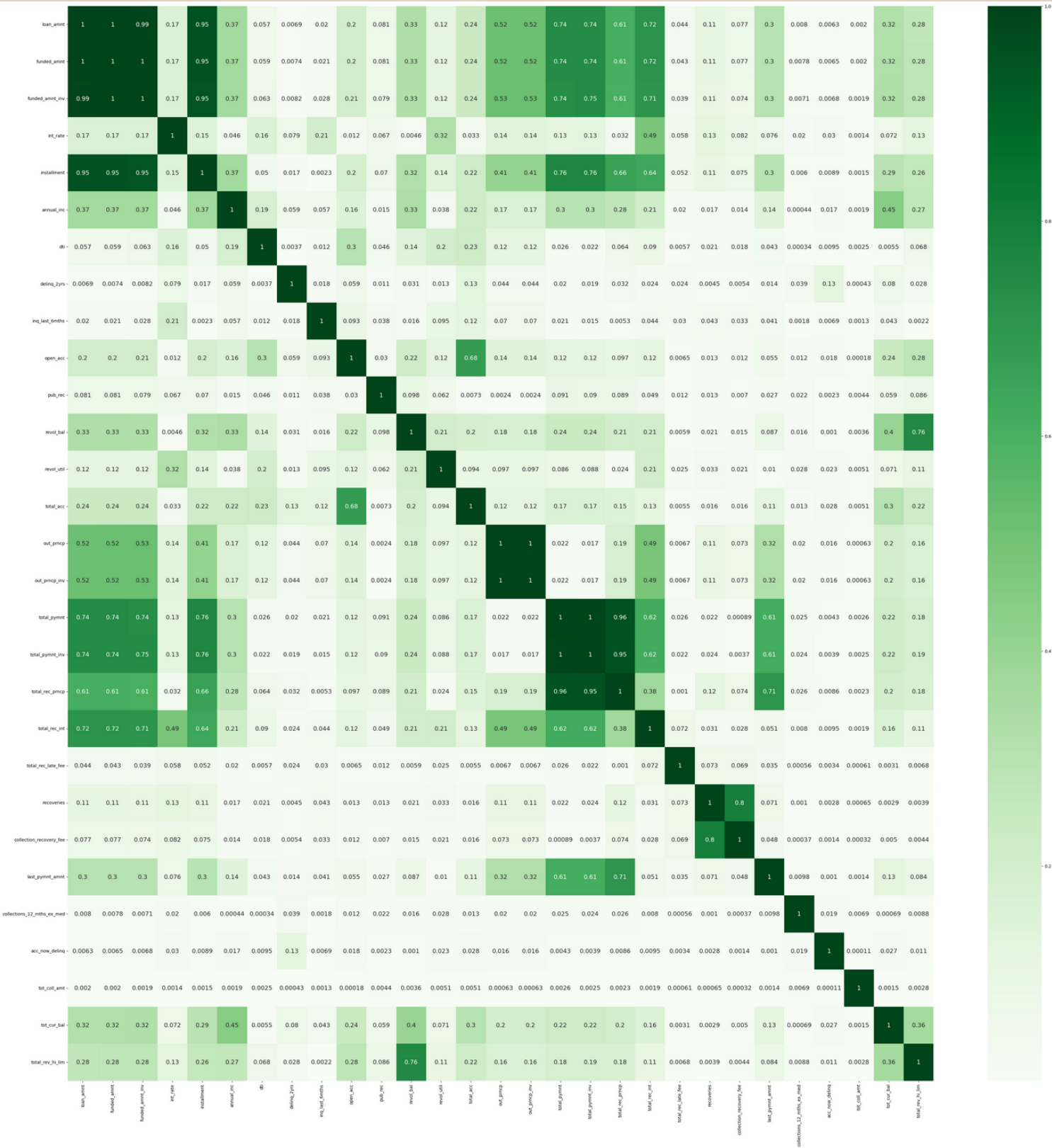
# EDA AFTER DROPPING ALL CATEGORICAL BAR CHART



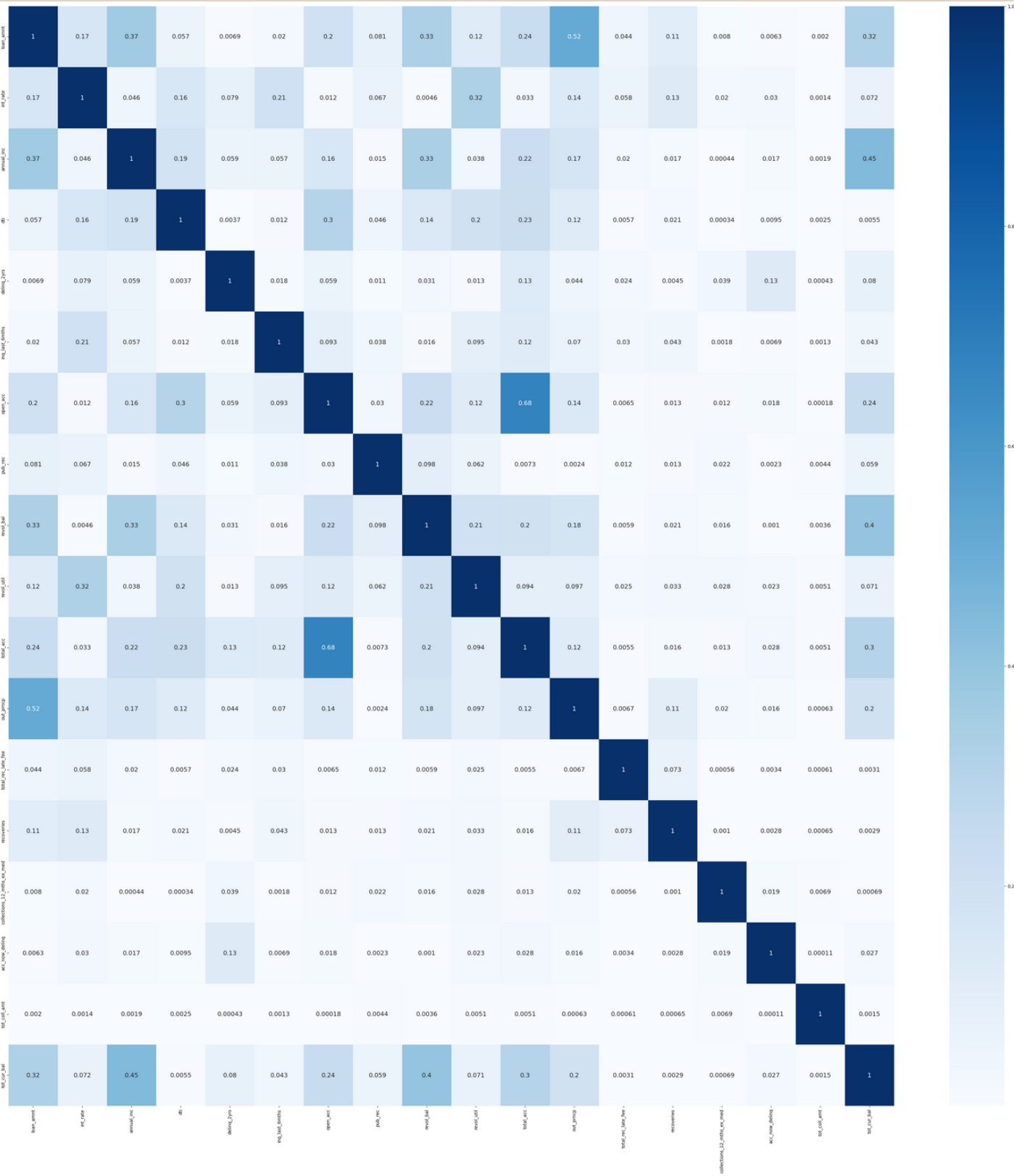


# CORRELATION MATRIX

## NUMERICAL FEATURE



# CORRELATION MATRIX AFTER CLEANING



# MACHINE LEARNING

```
Training Random Forest...
Evaluating Random Forest...
```

Classification Report:

	precision	recall	f1-score	support
Good	0.97	0.45	0.62	11080
Bad	0.93	1.00	0.96	82177
accuracy			0.93	93257
macro avg	0.95	0.73	0.79	93257
weighted avg	0.94	0.93	0.92	93257

Random Forest Accuracy: 0.9336

```
Training Logistic Regression...
Evaluating Logistic Regression...
```

Classification Report:

	precision	recall	f1-score	support
Good	0.98	0.42	0.59	11080
Bad	0.93	1.00	0.96	82177
accuracy			0.93	93257
macro avg	0.95	0.71	0.78	93257
weighted avg	0.93	0.93	0.92	93257

Logistic Regression Accuracy: 0.9302

# MACHINE LEARNING

```
Training Gradient Boosting...
Evaluating Gradient Boosting...
```

```
Classification Report:
```

	precision	recall	f1-score	support
Good	0.99	0.45	0.62	11080
Bad	0.93	1.00	0.96	82177
accuracy			0.93	93257
macro avg	0.96	0.72	0.79	93257
weighted avg	0.94	0.93	0.92	93257

```
Gradient Boosting Accuracy: 0.9338
```

```
Training K-Nearest Neighbors...
Evaluating K-Nearest Neighbors...
```

```
Classification Report:
```

	precision	recall	f1-score	support
Good	0.81	0.36	0.50	11080
Bad	0.92	0.99	0.95	82177
accuracy			0.91	93257
macro avg	0.87	0.67	0.73	93257
weighted avg	0.91	0.91	0.90	93257

```
K-Nearest Neighbors Accuracy: 0.9141
```

# MACHINE LEARNING OVERSAMPLING

Training Random Forest...  
Evaluating Random Forest...

Classification Report:

	precision	recall	f1-score	support
Good	0.99	0.96	0.98	81995
Bad	0.97	0.99	0.98	82387
accuracy			0.98	164382
macro avg	0.98	0.98	0.98	164382
weighted avg	0.98	0.98	0.98	164382

Random Forest Accuracy: 0.9769

Training Logistic Regression...  
Evaluating Logistic Regression...

Classification Report:

	precision	recall	f1-score	support
Good	0.87	0.60	0.71	81995
Bad	0.70	0.91	0.79	82387
accuracy			0.76	164382
macro avg	0.78	0.75	0.75	164382
weighted avg	0.78	0.76	0.75	164382

Logistic Regression Accuracy: 0.7552

# MACHINE LEARNING OVERSAMPLING

```
Training Gradient Boosting...
Evaluating Gradient Boosting...
```

Classification Report:

	precision	recall	f1-score	support
Good	0.69	0.80	0.74	81995
Bad	0.77	0.64	0.70	82387
accuracy			0.72	164382
macro avg	0.73	0.72	0.72	164382
weighted avg	0.73	0.72	0.72	164382

Gradient Boosting Accuracy: 0.7227

```
Training K-Nearest Neighbors...
Evaluating K-Nearest Neighbors...
```

Classification Report:

	precision	recall	f1-score	support
Good	0.81	0.98	0.89	81995
Bad	0.97	0.77	0.86	82387
accuracy			0.87	164382
macro avg	0.89	0.87	0.87	164382
weighted avg	0.89	0.87	0.87	164382

K-Nearest Neighbors Accuracy: 0.8741