

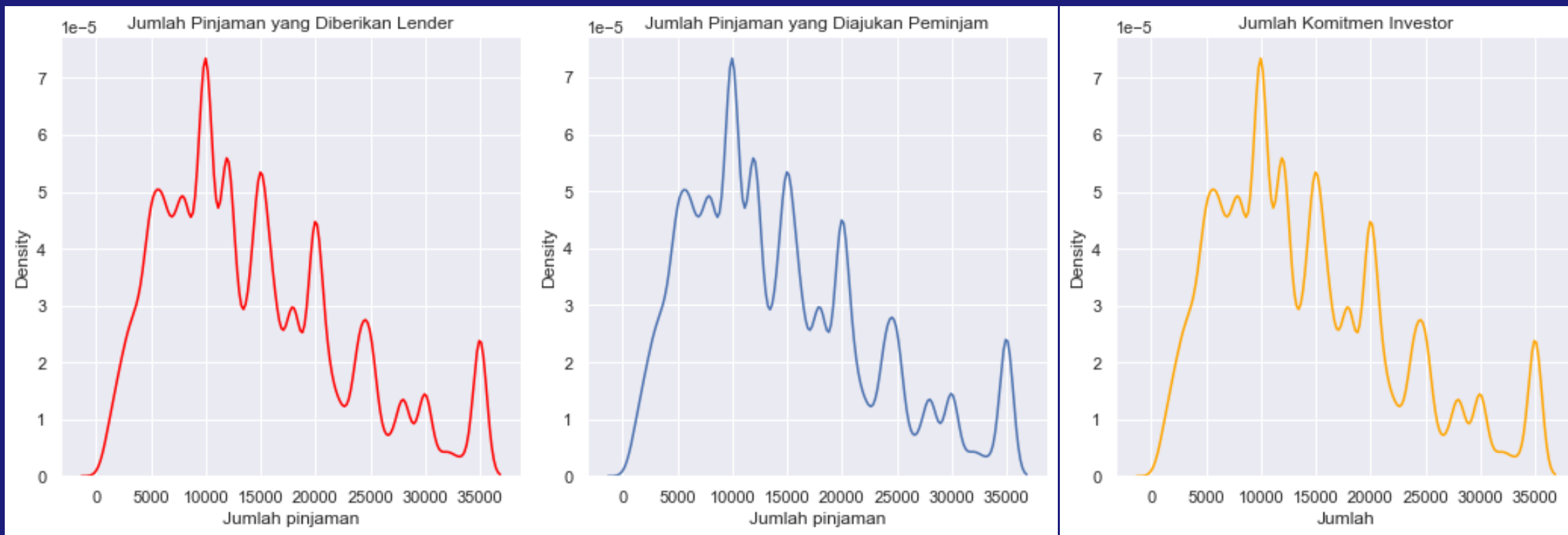
CREDIT RISK PREDICTION MODEL

Latar Belakang

Perusahaan loan merupakan sebuah perusahaan yang bergerak di bidang jasa peminjaman dana dengan pengembalian berupa dana pokok ditambah suku bunga yang ditetapkan perusahaan. Pihak lender memperoleh pendapatan dari suku bunga yang ditetapkan, sedangkan pihak peminjam memperoleh keuntungan dari dana yang dipinjamkan untuk keperluannya. Namun, perlu dilakukan pemilihan dalam penentuan apakah peminjam tersebut akan bermasalah atau tidak di kemudian hari. Dan apakah pihak lender perlu menerima atau menolak pihak peminjam dengan berbagai faktor yang dipertimbangkan. Oleh karena itu, perlu dibuat model untuk memprediksi apakah pihak peminjam akan default di kemudian hari.

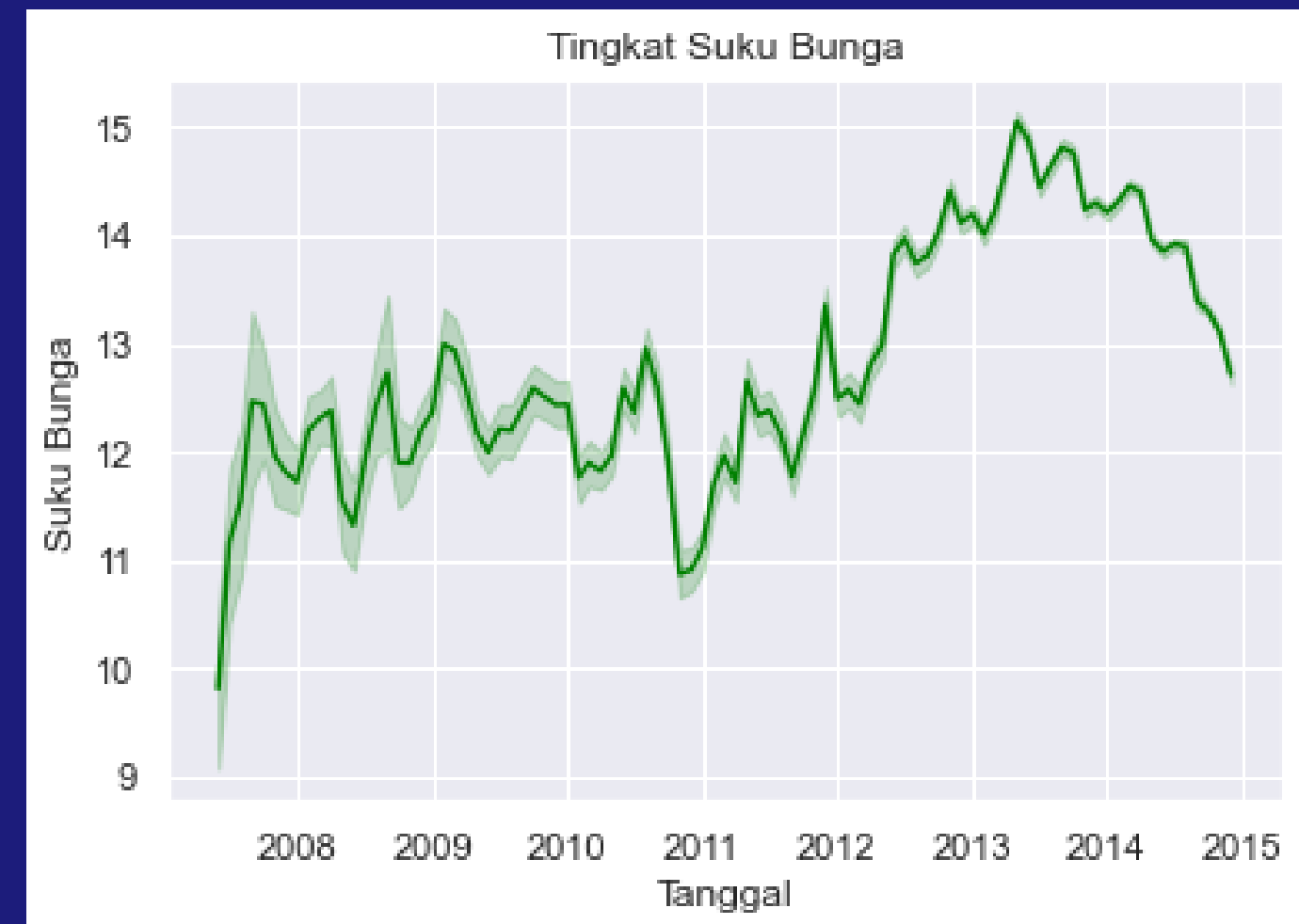
Pihak perusahaan meminta saya untuk membuat model yang dapat memprediksi apakah peminjam yang meminjam di perusahaan terkait, ditolak atau diterima. Dengan diberikannya data yang dibutuhkan, diharapkan model yang diperoleh bisa membantu dalam penentuan keputusan pihak perusahaan.

Analisis Deskriptif

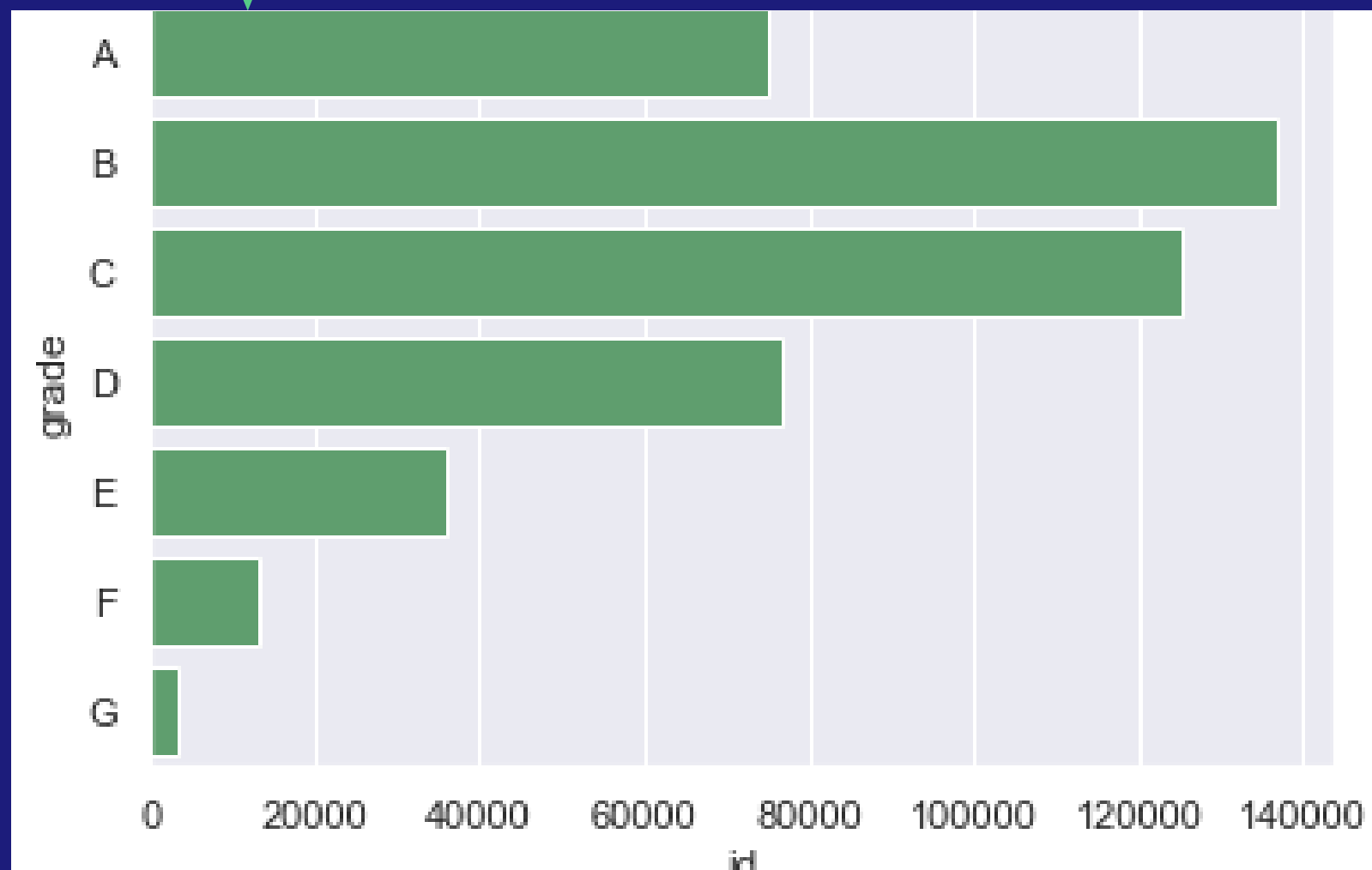
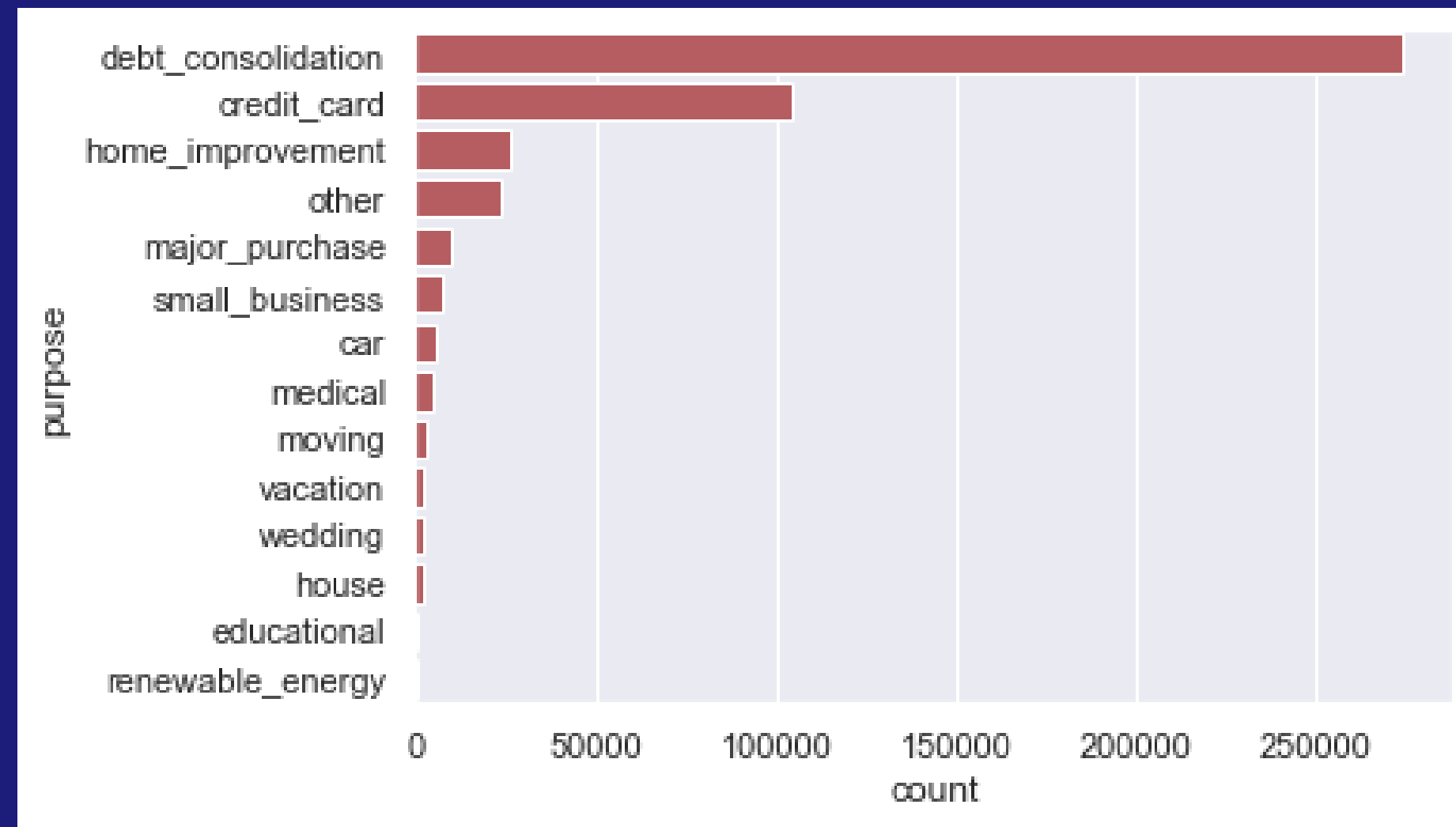


Sebelum melakukan modeling, maka perlu dilakukan analisis deskriptif pada data loan. Di samping merupakan grafik distribusi dari jumlah yang diajukan peminjam, jumlah pinjaman yang disetujui lender dan total komitmen investor. dapat dilihat bahwa grafik dari ketiga faktor tersebut identik. Artinya, jumlah pinjaman yang diajukan sama dengan jumlah pinjaman yang diberikan oleh lender.

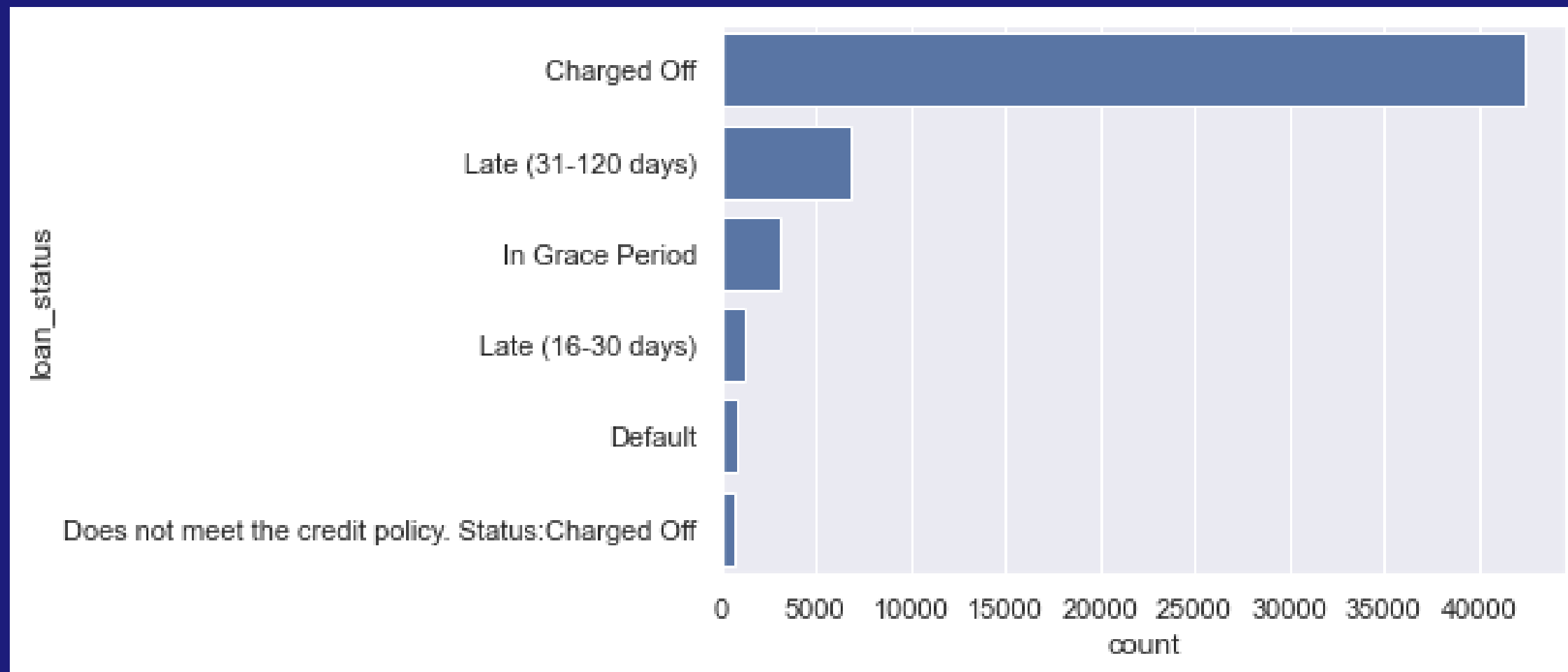
Suku bunga yang diberikan oleh lender kepada peminjam terlihat setiap tahunnya memiliki tren atau cenderung mengalami kenaikan. Namun disetiap tahunnya tentunya nilainya sangat berfluktuasi. Pada awal tahun 2013 hingga akhir tahun 2014, suku bunga mengalami penurunan yang cukup signifikan. Hal ini mungkin strategi dari pihak perusahaan lender untuk menurunkan suku bunga agar nantinya para calon peminjam bisa tertarik untuk melakukan peminjaman.



Selanjutnya akan dilihat sebenarnya tujuan peminjam dalam melakukan pinjaman itu apa. Dapat dilihat dari diagram batang disamping bahwa jumlah terbanyak yaitu untuk membayar hutang lain. Kemudian diikuti oleh pembayaran kartu kredit. Hal ini bisa menjadi pertimbangan lender untuk meminjamkan dana dan berfikir apakah nantinya si peminjam akan bermasalah dalam proses pembayarannya atau tidak.



Persebaran peminjam dalam setiap grade bisa dilihat pada grafik disamping. Grade dengan jumlah peminjam terbanyak ada di grade B dan disusul grade C. Sedangkan peminjam paling sedikit ada pada grade G. Hal ini juga bisa menjadi pertimbangan lender dalam penentuan pinjaman.



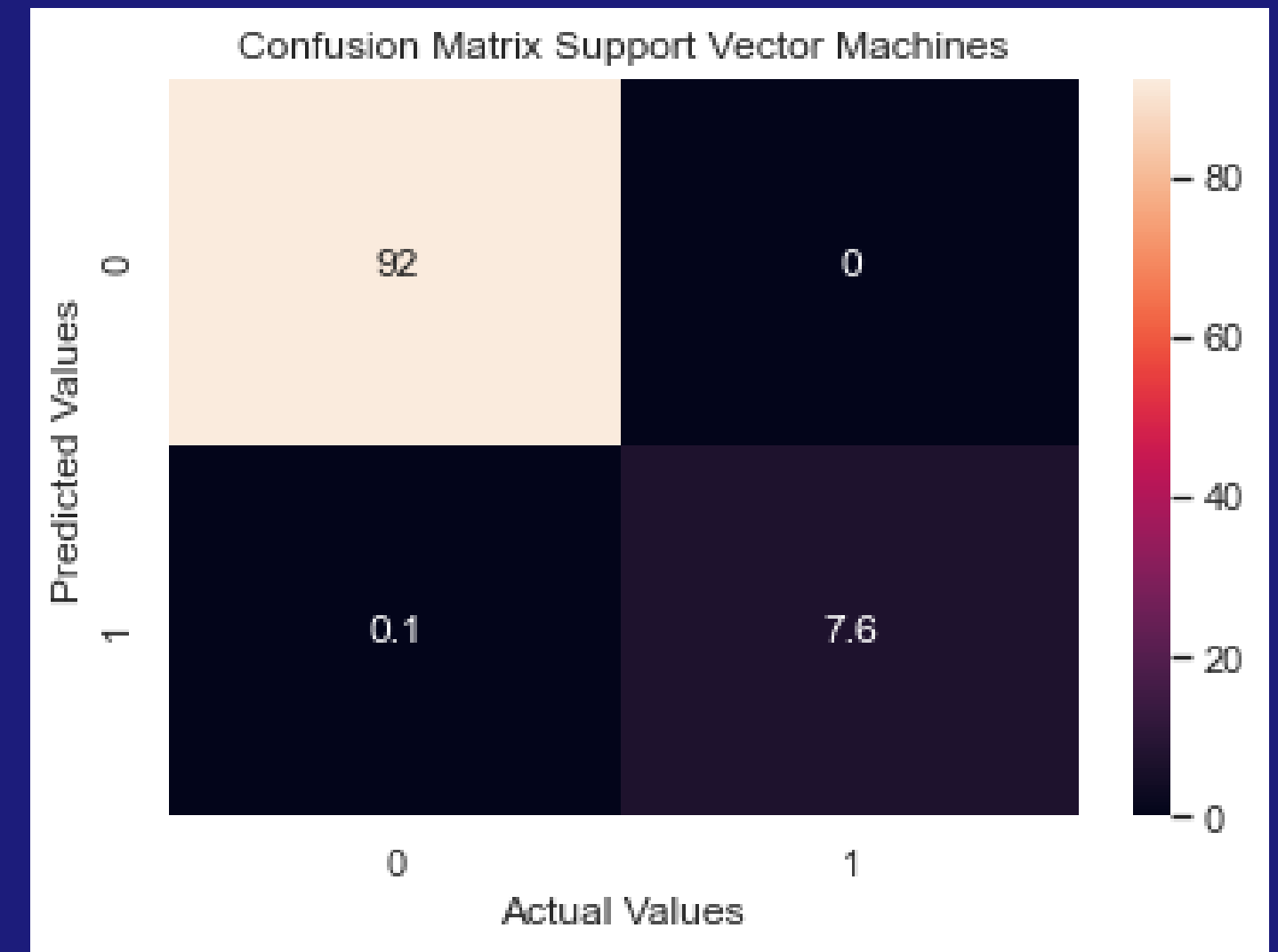
Lalu bagaimana dengan pinjaman yang bermasalah? Disini bisa dilihat bahwa grafik persebaran peminjam yang bermasalah dalam proses pembayaran. Terlihat bahwa peminjam yang sudah dipastikan bermasalah dan akan default berada di urutan paling tinggi. dilanjutkan dengan peminjam yang telat dalam proses pembayaran selama 1-4 bulan. Hal ini yang nantinya akan menjadi faktor penentu apakah peminjam default atau tidak

Hasil Pemodelan

Dari proses pemilihan model dengan beberapa model yang digunakan seperti Logistic Regression, Random Forest, K-Nearest Neighbor, Support Vector Machine dan Neural Network, diperoleh model terbaik yaitu dengan Model Support Vector Machine. Kriteria pemilihan model dapat dilihat dari nilai Akurasi, Presisi, Recall dan Skor F1. Disamping terlihat bahwa akurasi dari model sebesar 99,9 %, artinya model memiliki akurasi yang sangat baik dalam memprediksi default dan tidak default. Nilai presisi sebesar 100% menunjukkan bahwa model sangat baik dalam ketepatan prediksi.

Dari Confusion Matrix diperoleh bahwa:

1. Sekitar 92% yang diprediksi tidak akan default, ternyata tidak default.
2. Sekitar 0% yang diprediksi tidak akan default, ternyata default
3. Sekitar 0,1% yang diprediksi default, ternyata tidak default
4. Sekitar 7,6% yang diprediksi default, ternyata default



```
Accuracy predict: 99.90%  
Precision: 100.00%  
Recall: 98.64%  
F1_score: 99.32%
```

Kesimpulan

1. Analisis Deskriptif:

- a) Terjadi kesamaan antara jumlah pinjaman yang diajukan oleh peminjam dengan jumlah yang diberikan oleh pihak Lender
- b) Tingkat suku bunga cenderung memiliki tren naik. Tetapi tetap berfluktuasi setiap tahunnya. Pada awal 2013 sampai akhir 2014 mengalami penurunan yang mungkin disebabkan oleh kebijakan lender
- c) Peminjam terbanyak melakukan pinjaman dengan tujuan untuk membayar hutang
- d) Peminjam bermasalah terbanyak yaitu Charged Off
- g) Peminjam pada grade B memiliki jumlah terbanyak, dan grade G memiliki jumlah paling sedikit

2. Analisis Prediktif :

Didapatkan model yang terbaik adalah model Support Vector Machine dengan nilai akurasi 99,9% dan presisi 100%