

**TUGAS KELOMPOK MODEL SURVIVAL**  
**ANALISIS PENYEBAB *CUSTOMER CHURN* PERUSAHAAN**  
**TELEKOMUNIKASI DENGAN MODEL SURVIVAL**



**Anggota Kelompok :**

1. Zhafirah Tsany - 2006572951
2. Rafly Witjaksana Hartantyo - 2106651572
3. Rachel Thyffani Margaretha S - 2106726900

**DEPARTEMEN MATEMATIKA**  
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN**  
**UNIVERSITAS INDONESIA**  
**TAHUN AJARAN 2022/2023**

## DAFTAR ISI

<b>DAFTAR ISI</b>	<b>1</b>
<b>BAB I</b>	<b>2</b>
<b>PENDAHULUAN</b>	<b>2</b>
A. LATAR BELAKANG	2
B. RUMUSAN MASALAH	2
C. TUJUAN PENGOLAHAN DATA	2
<b>BAB II</b>	<b>3</b>
<b>METODE PENGOLAHAN DATA</b>	<b>3</b>
A. DATA	3
B. PENGGUNAAN METODE SURVIVAL	3
<b>BAB III</b>	<b>9</b>
<b>INTERPRETASI DAN ANALISIS HASIL</b>	<b>9</b>
A. INTERPRETASI	9
B. ANALISIS	10
<b>BAB IV</b>	<b>11</b>
<b>KESIMPULAN</b>	<b>11</b>
<b>DAFTAR PUSTAKA</b>	<b>12</b>
<b>LAMPIRAN</b>	<b>13</b>
<b>PERSENTASE DAN JENIS KONTRIBUSI</b>	<b>15</b>

## **BAB I**

### **PENDAHULUAN**

#### **A. LATAR BELAKANG**

Analisis Survival adalah salah satu metode analisis untuk mempelajari berapa lama waktu yang dibutuhkan suatu objek untuk bertahan hidup. Analisis ini sering digunakan untuk mengetahui faktor-faktor yang mempengaruhi waktu bertahan hidup dan kemungkinan lama bertahan hidup suatu objek, serta untuk mengurangi kerugian sebuah perusahaan dengan menganalisis faktor risiko yang mempengaruhi keberlangsungan perusahaan.

Salah satu cara untuk menggunakan analisis survival akan ditunjukkan pada data *customer churn* berikut ini. *customer churn*, juga dikenal sebagai pengurangan pelanggan, adalah metrik bisnis yang menghitung jumlah pelanggan yang tidak lagi menggunakan produk atau layanan. Peningkatan *customer churn rate* menandakan adanya penurunan *customer lifetime value* (CLV), yang merupakan total pendapatan yang diharapkan dari rata-rata pelanggan, serta penurunan tingkat pertumbuhan perusahaan. Oleh karena itu, perlu ditelusuri faktor penyebabnya dan diperlukan strategi mumpuni untuk menghindari masalah tersebut.

Dalam laporan ini, akan dipelajari bagaimana analisis survival dapat digunakan untuk mengurangi potensi terjadinya *customer churn* dengan mengidentifikasi faktor-faktor penyebab pelanggan berhenti menggunakan layanan perusahaan dan menemukan solusi untuk menghindari atau mengurangi terjadinya *customer churn*.

#### **B. RUMUSAN MASALAH**

1. Bagaimana penerapan model survival dalam memahami data *customer churn* berikut?
2. Apa langkah yang perlu diambil perusahaan telekom untuk mengurangi *customer churn*?

#### **C. TUJUAN PENGOLAHAN DATA**

1. Mengetahui peranan model survival di dunia nyata, dimana pada laporan ini penggunaannya berfokus pada *customer churn*.
2. Mengetahui langkah selanjutnya yang perlu diambil perusahaan untuk mengurangi terjadinya *customer churn*.

## BAB II

### METODE PENGOLAHAN DATA

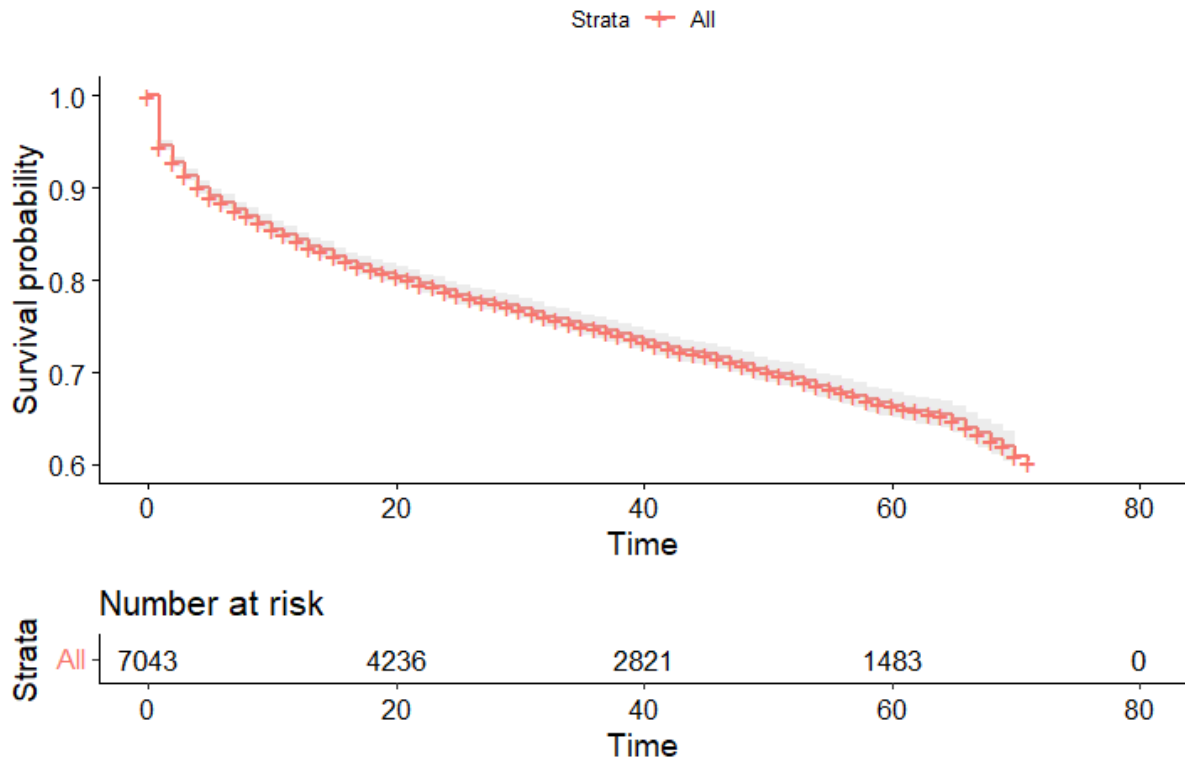
#### A. DATA

Data yang digunakan di penelitian ini adalah data *customer churn* suatu perusahaan dalam 72 bulan. Sumber data berasal dari [https://github.com/treselle-systems/customer\\_churn\\_analysis/blob/master/WA\\_Fn-UseC\\_-Telco-Customer-Churn.csv](https://github.com/treselle-systems/customer_churn_analysis/blob/master/WA_Fn-UseC_-Telco-Customer-Churn.csv). Berikut ini adalah cuplikan dari data:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity
0	7590-VHVEG	Female	0	Yes	No	1.0	No	No phone service	DSL	No
1	5575-GNVDE	Male	0	No	No	34.0	Yes	No	DSL	Yes
2	3668-QPYBK	Male	0	No	No	2.0	Yes	No	DSL	Yes
3	7795-CFOCW	Male	0	No	No	45.0	No	No phone service	DSL	Yes
4	9237-HQITU	Female	0	No	No	2.0	Yes	No	Fiber optic	No

#### B. PENGGUNAAN METODE SURVIVAL

Pertama, karena kita ingin menentukan langkah yang perlu diambil untuk mengurangi *customer churn* pada perusahaan akan dihapus seluruh baris dimana pengguna tidak mempunyai *phone service* dan *internet service* karena hasil yang diperoleh tidak akan sesuai dengan yang diinginkan. Sebelum melakukan analisis lebih lanjut, akan dilihat *Kaplan-Meier Survival Curve*.



**Gambar 2.1.** Grafik fungsi survival *customer churn* perusahaan teknologi dengan metode Kaplan-Meier

Setelah itu akan dibuat model Cox-PH, berikut adalah hasil dari Cox-Ph tersebut:

```
> summary(cox)
Call:
coxph(formula = Surv(tenure, churn) ~ ., data = df)

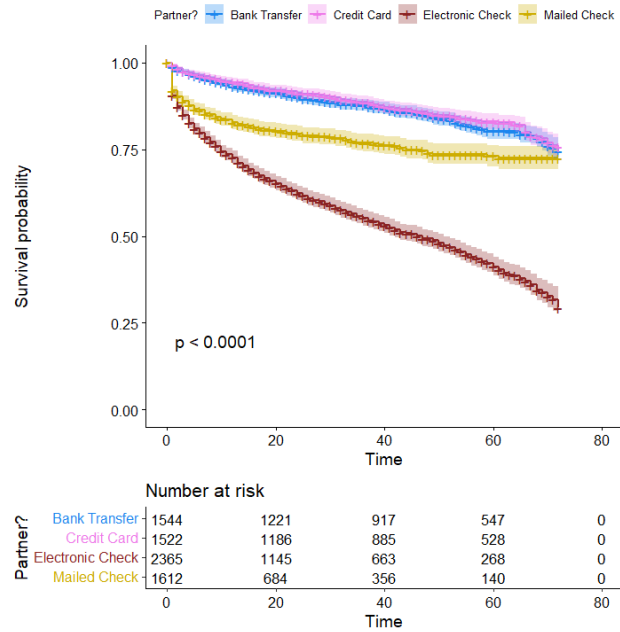
n= 4832, number of events= 1586
(3 observations deleted due to missingness)
```

	coef	exp(coef)	se(coef)	z	Pr(> z )
genderMale	-6.598e-03	9.934e-01	5.063e-02	-0.130	0.896323
SeniorCitizen	2.532e-02	1.026e+00	6.018e-02	0.421	0.673911
PartnerYes	-1.382e-01	8.709e-01	5.969e-02	-2.315	0.020612 *
DependentsYes	-6.324e-02	9.387e-01	7.618e-02	-0.830	0.406403
MultipleLinesYes	-1.999e-01	8.188e-01	5.697e-02	-3.508	0.000451 ***
OnlineSecurityYes	-6.045e-01	5.464e-01	7.358e-02	-8.215	< 2e-16 ***
OnlineBackupYes	-3.628e-01	6.957e-01	5.893e-02	-6.157	7.42e-10 ***
DeviceProtectionYes	-2.298e-01	7.947e-01	5.921e-02	-3.881	0.000104 ***
TechSupportYes	-3.731e-01	6.886e-01	7.031e-02	-5.308	1.11e-07 ***
StreamingTVYes	-3.050e-01	7.371e-01	6.845e-02	-4.455	8.37e-06 ***
StreamingMoviesYes	-3.077e-01	7.352e-01	6.795e-02	-4.528	5.95e-06 ***
ContractOne year	-8.639e-01	4.215e-01	1.090e-01	-7.927	2.24e-15 ***
ContractTwo year	-2.050e+00	1.287e-01	2.088e-01	-9.818	< 2e-16 ***
PaperlessBillingYes	2.488e-01	1.283e+00	6.349e-02	3.919	8.90e-05 ***
PaymentMethodCredit card (automatic)	2.440e-03	1.002e+00	9.778e-02	0.025	0.980091
PaymentMethodElectronic check	3.645e-01	1.440e+00	7.933e-02	4.595	4.33e-06 ***
PaymentMethodMailed check	3.170e-01	1.373e+00	9.996e-02	3.172	0.001516 **
Monthlycharges	9.386e-02	1.098e+00	3.932e-03	23.873	< 2e-16 ***
Totalcharges	-2.599e-03	9.974e-01	6.123e-05	-42.439	< 2e-16 ***

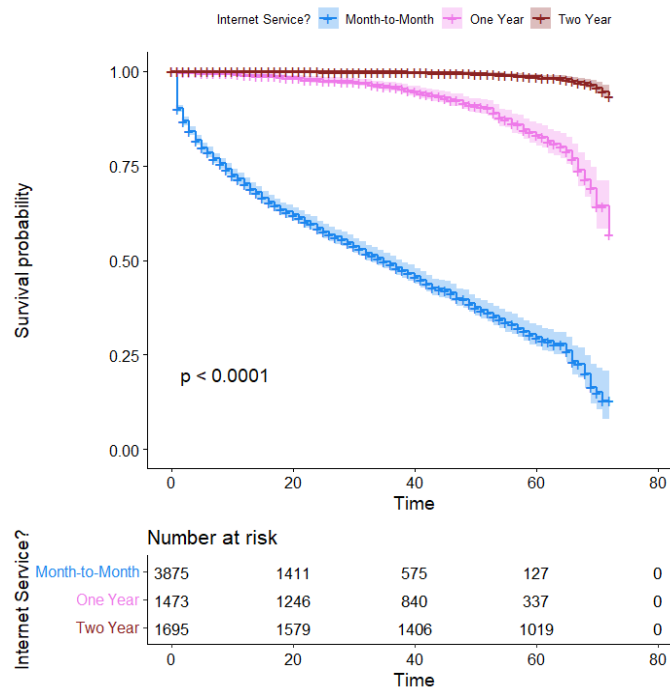
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari hasil diatas, output yang akan dilihat adalah exp(coef), kolom ini mewakili kenaikan hazard untuk setiap poin kenaikan variabel, dengan 1 adalah netral.

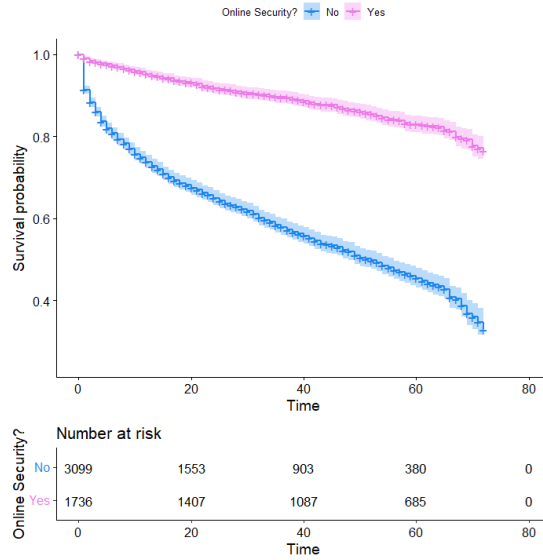
Untuk mempermudah interpretasi maka akan diberikan beberapa plot untuk variabel yang signifikan.



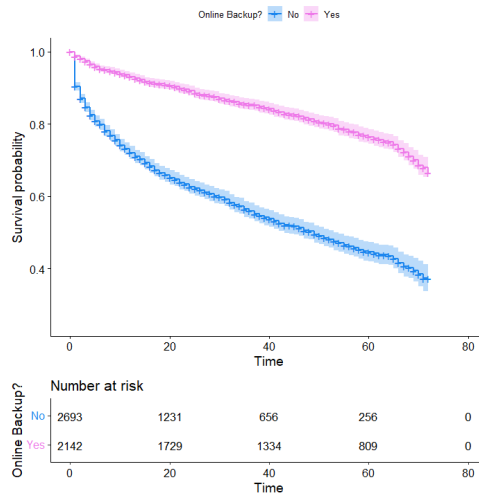
**Gambar 2.2.** Grafik fungsi survival *customer churn* berdasarkan metode pembayaran



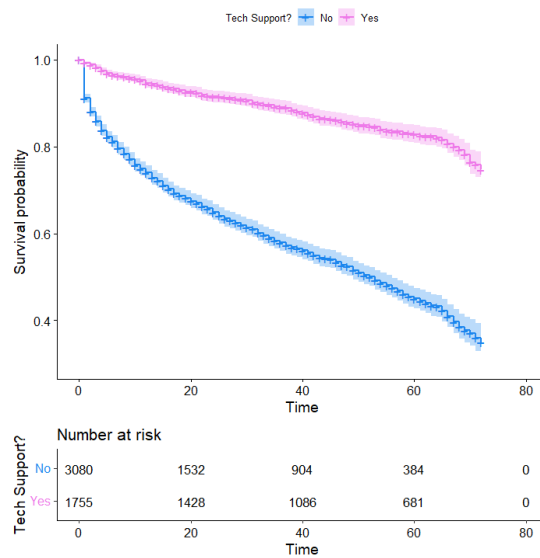
**Gambar 2.3.** Grafik fungsi survival *customer churn* berdasarkan tipe paket langganan



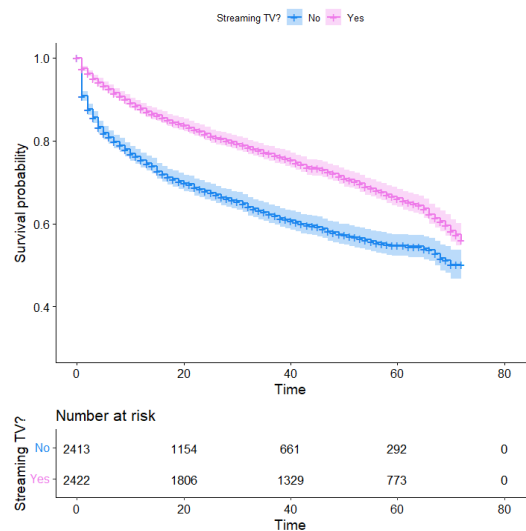
**Gambar 2.4.** Grafik fungsi survival *customer churn* berdasarkan keamanan *online*



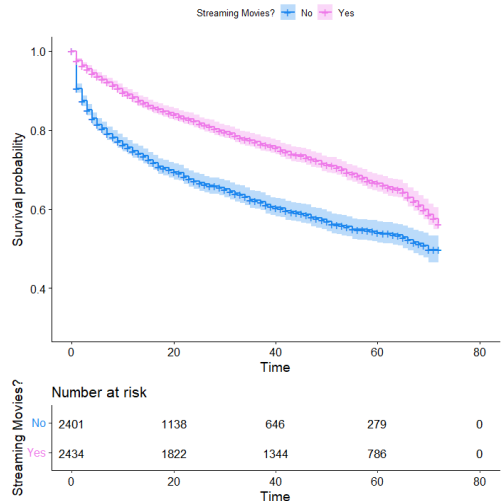
**Gambar 2.5.** Grafik fungsi survival *customer churn* berdasarkan *online backup*



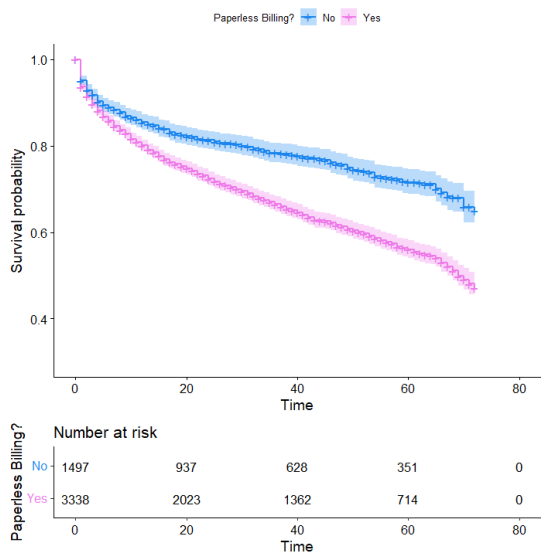
Gambar 2.6. Grafik fungsi survival *customer churn* berdasarkan *technical support*



Gambar 2.7. Grafik fungsi survival *customer churn* berdasarkan *streaming TV*



Gambar 2.8. Grafik fungsi survival *customer churn* berdasarkan *streaming film*





**Gambar 2.9.** Grafik fungsi survival *customer churn* berdasarkan pembayaran *paperless*

### BAB III

#### INTERPRETASI DAN ANALISIS HASIL

##### A. INTERPRETASI

Data yang dianalisis pada penelitian ini adalah *customer churn* dari suatu perusahaan telekomunikasi. Selama 72 bulan dilakukan penelitian dengan mengukur hari pertama pelanggan sampai berhenti berlangganan. Pada gambar 2.1 dengan metode Kaplan-Meier diperoleh bahwa tingkat *customer churn* masih termasuk rendah karena dalam 72 bulan, perusahaan masih dapat mempertahankan sekitar 60% pelanggannya. Selanjutnya untuk melihat survival berdasarkan variabel-variabel tertentu juga bisa dicari tahu dan pada penelitian ini memakai metode Cox-PH dan berikut ini interpretasi dari berbagai variabel yang dipilih.

Pada gambar 2.2 menunjukkan fungsi survival *customer churn* berdasarkan metode pembayaran. Terdapat empat pilihan pembayaran, yaitu transfer bank, kartu kredit, cek elektronik, dan cek email. Pembayaran menggunakan transfer bank dan kartu kredit memiliki nilai survival yang lebih besar dan penurunan survival yang cukup lambat. Keduanya juga memiliki nilai yang hampir sama. Sedangkan cek elektronik memiliki nilai survival. Dapat diinterpretasikan bahwa opsi pembayaran transfer bank dan kartu kredit memudahkan pembayaran langganan sehingga mendorong pelanggan untuk meneruskan langganannya.

Pada gambar 2.3 menunjukkan fungsi survival *customer churn* berdasarkan tipe paket langganan. Terdapat 3 paket langganan, yaitu per bulan, per tahun, dan per 2 tahun. Paket per 2 tahun memiliki survival tertinggi dan selama penelitian hampir tidak turun nilainya. Sedangkan paket per bulan memiliki survival terendah dan penurunan yang lebih cepat. Dapat diinterpretasikan paket per 2 tahun memang dibeli bagi pelanggan yang sudah puas dan menyukai paket langganan yang disediakan perusahaan telekomunikasi ini sehingga tidak ragu untuk berlangganan dalam jangka panjang. Sementara untuk paket per bulan mungkin paket yang dibeli bagi pelanggan yang masih ragu terhadap paket langganannya sehingga membeli paket jangka pendek agar jika tidak puas bisa lebih cepat diberhentikan.

Pada gambar 2.4 menunjukkan fungsi survival *customer churn* berdasarkan keamanan *online*. Pelanggan yang merasakan keamanan *online* yang diberikan sudah cukup memiliki survival yang lebih besar dan penurunan lebih lambat dibandingkan dengan pelanggan yang merasa tidak aman. Pada masa digital seperti sekarang keamanan digital sangatlah penting karena sangat mudah data diambil kemudian disalahgunakan jika tidak memiliki sistem keamanan yang baik. Jadi sangat masuk akal bagi pelanggan yang tidak merasa cukup terhadap keamanan *online* tidak melanjutkan langganan. Informasi yang diberikan dari gambar 2.5 mengenai survival *customer churn* berdasarkan *online backup* juga memiliki interpretasi dan alasan yang serupa.

Pada gambar 2.6 menunjukkan fungsi survival *customer churn* berdasarkan *technical support*. Bagi pelanggan yang merasa puas dengan layanan *technical support* yang diberikan perusahaan memiliki survival yang lebih besar dan penurunan

lebih lambat daripada pelanggan yang tidak puas. Hasil ini masuk akal karena *technical support* sangat penting bagi pelanggan agar kesulitan dan masalah yang dihadapi dapat terselesaikan sehingga menjadi alasan untuk terus berlangganan.

Gambar 2.7 dan 2.8 sama-sama menunjukkan menunjukkan fungsi survival *customer churn* berdasarkan penggunaan data untuk *streaming*, yang pertama untuk TV dan yang kedua untuk film. Pada kedua grafik pelanggan yang menggunakan paket data untuk *streaming* memiliki survival yang lebih besar daripada yang tidak. Dengan menonton acara atau film yang lebih leluasa saat ini banyak sekali situs *streaming* yang bisa dipilih. Bagi pelanggan yang memiliki hobi untuk menonton berlangganan data akan mendukung hobi tersebut juga memiliki alasan lebih untuk berlangganan paket data.

Pada gambar 2.9 menunjukkan fungsi survival *customer churn* berdasarkan pembayaran *paperless*. Pelanggan yang melakukan pembayaran *paperless* memiliki survival yang lebih besar dibandingkan dengan yang tidak. Hal ini dapat disebabkan dengan adanya digitalisasi yang memudahkan melakukan banyak aktivitas, termasuk pembayaran.

## B. ANALISIS

Salah satu tujuan dari penelitian ini adalah bagaimana perusahaan mengurangi *customer churn*. Walaupun sebelumnya melalui gambar 2.1 diketahui bahwa tingkat *customer churn* perusahaan termasuk rendah, tetap tidak ada salahnya untuk mengurangi *customer churn* sebisa mungkin. Berdasarkan penelitian survival setiap variabel, untuk variabel tipe paket langganan, telah diketahui bahwa pelanggan yang membeli paket per bulan memiliki survival terendah dan memiliki penurunan yang lumayan cepat pula. Salah satu rekomendasi yang bisa diberikan adalah membuat promo untuk langganan per bulan lebih sering. Misalkan potongan harga untuk beberapa bulan pertama atau harga spesial di bulan perayaan tertentu untuk menarik pelanggan baru dan memperoleh loyalitas dari pelanggan lama.

Dari model Cox-PH kedua variabel *streaming* memiliki koefisien yang bernilai negatif yang menyatakan untuk pelanggan yang menggunakan pakatnya untuk *streaming* dapat mengurangi kemungkinan untuk berhenti berlangganan. Sehingga untuk mempertahankan loyalitas dan juga menarik pelanggan lain yang sebelumnya jarang *streaming*, perusahaan dapat berkolaborasi dengan situs *streaming* untuk membuat paket khusus untuk *streaming*. Harga paket bisa dibuat lebih murah dibandingkan metode pembayaran lainnya.

Kemudian sepertinya bidang keamanan dan pelayanan *support* untuk pelanggan harus ditingkatkan karena selisih survival yang merasa puas terhadap keamanan dan *support* yang diberikan dengan yang tidak puas termasuk besar. Selain untuk membuat pelanggan tetap ingin melanjutkan berlangganan, sudah menjadi tanggung jawab perusahaan untuk menjaga data yang sudah dipercayakan pelanggan setelah mereka setuju untuk membeli paket langganan.

## BAB IV KESIMPULAN

Setelah melakukan analisis survival pada data *customer churn* perusahaan telekomunikasi diperoleh kesimpulan sebagai berikut:

1. Model survival bisa digunakan untuk melihat porsi *customer churn* yang terjadi pada suatu periode tertentu. Juga dapat mengetahui variabel-variabel yang dapat mempengaruhi pelanggan untuk lanjut berlangganan atau berhenti sehingga melalui informasi tersebut perusahaan dapat mengetahui bidang apa saja yang bisa ditingkatkan untuk mengurangi *customer churn*.
2. Beberapa rekomendasi yang dapat diberikan perusahaan telekomunikasi ini untuk mengurangi *customer churn* adalah mengadakan promo bulanan, berkolaborasi dengan situs *streaming* untuk membentuk paket khusus, dan meningkatkan keamanan *online* serta pelayanan yang diberikan kepada pelanggan.

## DAFTAR PUSTAKA

- Abdullah, S. (2022). *Analisis Survival Konsep dan Aplikasi Dengan R*. PT Bumi Aksara.
- Cinthy. (2022). *Customer churn: Pengertian, Cara Hitung, Dan Cara Menanggulangnya*. Accurate Online.  
[https://accurate.id/marketing-manajemen/customer-churn/#Apa\\_itu\\_customer\\_churn](https://accurate.id/marketing-manajemen/customer-churn/#Apa_itu_customer_churn)
- Treselle-Systems, uvarajseerangan. (2017). Treselle-systems/customer\_churn\_analysis. GitHub. [https://github.com/treselle-systems/customer\\_churn\\_analysis](https://github.com/treselle-systems/customer_churn_analysis)

## LAMPIRAN

### Kode R

```
library(survival)
library(ranger)
library(ggplot2)
library(dplyr)
library(survminer)
library(ggfortify)
library(naniar)
library(Publish)
library(ggtext)
library(fastDummies)
df=Churn.dataset
df$Churn[df$Churn == "Yes"] <- 1      # Replace "Yes" by 1
df$Churn[df$Churn == "No"] <- 0      # Replace "No" by 0
df$Churn=as.numeric(df$Churn)

km_fit <- survfit(Surv(tenure, Churn) ~ 1, data=df)
ggsurvplot(km_fit, conf.int=TRUE, risk.table=TRUE,, ylim = c(0.6,
1))

df=subset(df, select =
-c(customerID,PhoneService,InternetService) )

df=subset(df,MultipleLines!='No phone service' )
df=subset(df,OnlineSecurity!='No internet service' )
cox = coxph(Surv(tenure, Churn) ~., data =df)

summary(cox)
ggsurvplot(survfit(cox), data = df,break.time.by = 150,
risk.table=TRUE)

publish(cox)
plot(publish(cox), xlim=c(0,10), cex = 0.6)

km_fit <- survfit(Surv(tenure, Churn) ~ df$PaymentMethod,
data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
legend.labs=c("Bank Transfer","Credit Card",
"Electronic Check",'Mailed Check'), legend.title="Partner?",
palette=c("dodgerblue2", "orchid2","brown4",'gold3'),
main="Customer Churn by Partner")

km_fit <- survfit(Surv(tenure, Churn) ~ df$Contract, data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
```

```

        legend.labs=c("Month-to-Month", "One Year", 'Two
Year'), legend.title="Internet Service?",
        palette=c("dodgerblue2", "orchid2", 'brown4'),
        main="Customer Churn by Partner",ylim =c(0,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$OnlineSecurity,
data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Online
Security?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.25,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$OnlineBackup, data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Online
Backup?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.25,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$StreamingMovies, data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Streaming
Movies?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.3,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$TechSupport, data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Tech
Support?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.3,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$StreamingTV, data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Streaming
TV?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.3,1))

km_fit <- survfit(Surv(tenure, Churn) ~ df$PaperlessBilling,
data=df)
ggsurvplot(km_fit, conf.int=TRUE, pval=TRUE, risk.table=TRUE,
        legend.labs=c("No", 'Yes'), legend.title="Paperless
Billing?",
        palette=c("dodgerblue2", "orchid2"),
        main="Customer Churn by Partner",ylim =c(0.3,1))

```

### **PERSENTASE DAN JENIS KONTRIBUSI**

NAMA	NPM	KONTRIBUSI	PERSENTASE
Zhafirah Tsany	2006572951	Mencari data, membuat interpretasi dan analisis dari hasil pengolahan data dan menyusun laporan	100%
Rafly Witjaksana Hartantyo	2106651572	Mencari data, melakukan pengolahan data dan menyusun laporan data	100%
Rachel Thyffani Margaretha S	2106726900	Mencari data, memberikan latar belakang, merumuskan masalah, membentuk tujuan penelitian, membuat daftar pustaka dan menyusun laporan	100%