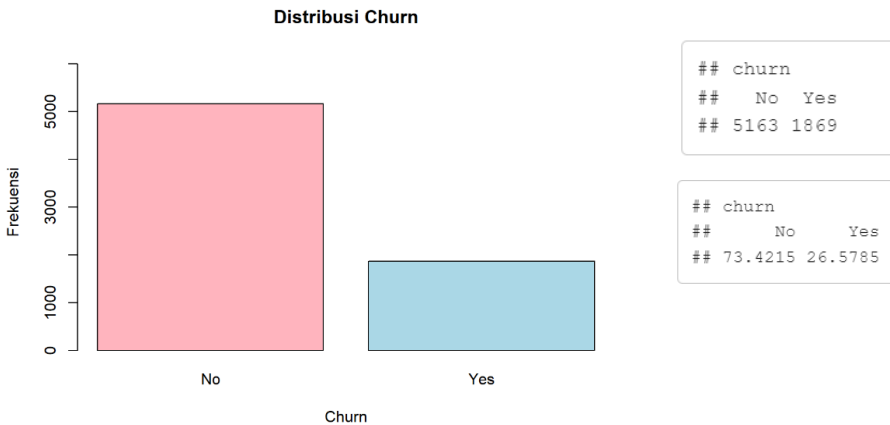


1. Churn Analysis using R and PowerBI

Dari dataset Telco, yaitu data mengenai pelanggan perusahaan telekomunikasi mengenai apa saja layanan yang digunakan oleh pelanggan, harga yang dibayar, serta apakah mereka berhenti menggunakan layanan perusahaan tersebut atau tidak. Pada laporan ini, dari dataset tersebut akan dianalisa faktor-faktor yang mempengaruhi berhentinya pelanggan dari menggunakan layanan perusahaan Telco menggunakan R.

Berikut adalah distribusi churn dari pelanggan perusahaan telco.

```
barplott <- barplot(Freq.Churn, col = c("lightpink", "lightblue"), ylim = c(0, max(Freq.C  
hurn) * 1.2), main = "Distribusi Churn", xlab = "Churn", ylab = "Frekuensi")
```



Dapat dilihat bahwa jumlah pelanggan yang berhenti berlangganan (**churn**) di bulan terakhir adalah sebanyak **1869** orang atau sekitar (**26.58%**).

Selanjutnya akan dilakukan analisis menggunakan chi-square untuk melihat hubungan antara churn dengan beberapa variabel-variabel lainnya yang jika dilihat dari hasil powerbi memiliki hubungan terhadap churn.

1. Analisa hubungan tipe kontrak dengan churn menggunakan chi-sq

```
##  
## Pearson's Chi-squared test  
##  
## data: contingency_table  
## X-squared = 1179.5, df = 2, p-value < 2.2e-16
```

Hasilnya $p\text{-value} < 0.05$, berarti terdapat hubungan yang sangat signifikan antara churn (pelanggan berhenti berlangganan) dan jenis kontrak yang digunakan oleh pelanggan. Ini juga berarti bahwa jenis kontrak yang dipilih oleh pelanggan mempengaruhi

kemungkinan mereka untuk berhenti berlangganan dari layanan perusahaan telekomunikasi.

Dalam hal ini, dari hasil yang di dapat, kebanyakan pelanggan yang melakukan churn adalah pelanggan yang berlangganan dengan tipe kontrak month-to-month atau bulanan, dapat disimpulkan bahwa tipe kontrak month-to-month juga sangat berpengaruh terhadap churn atau tidaknya pelanggan.

2. Analisa hubungan senior citizen dengan churn pelanggan menggunakan chi-sq

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: contingency_table  
## X-squared = 158.44, df = 1, p-value < 2.2e-16
```

Hasilnya $pvalue < 0.05$, menunjukkan bahwa terdapat hubungan yang signifikan antara churn dengan pelanggan yang adalah senior citizen. Berarti bahwa jenis pelanggan (senior atau bukan) mempengaruhi faktor churnnya pelanggan atau tidak.

Hal ini berarti bahwa kita perlu melakukan analisis lebih lanjut untuk memahami faktor-faktor lain yang mungkin mempengaruhi churn di antara senior citizen. Misalnya, apakah ada layanan atau produk tertentu yang lebih atau kurang digunakan oleh senior citizen? Serta perlu mengembangkan strategi khusus menargetkan pelanggan yang merupakan senior citizen.

3. Analisa hubungan dependents dengan churn menggunakan chi-sq

```
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data: contingency_table  
## X-squared = 186.32, df = 1, p-value < 2.2e-16
```

Hasilnya $pvalue < 0.05$, menunjukkan bahwa terdapat hubungan yang signifikan antara churn dengan pelanggan yang memiliki dependents. Berarti bahwa pelanggan memiliki dependents atau tidak mempengaruhi faktor churnnya pelanggan atau tidak.

4. Analisa hubungan tenure dengan churn menggunakan t-test

```
##
## Welch Two Sample t-test
##
## data: data$tenure by data$Churn
## t = 34.972, df = 4045.5, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group No and group Yes is not
equal to 0
## 95 percent confidence interval:
## 18.56811 20.77364
## sample estimates:
## mean in group No mean in group Yes
## 37.65001 17.97913
```

Hasilnya $p\text{-value} < 0.05$, menunjukkan bahwa terdapat hubungan yang signifikan antara churn dengan pelanggan tenure.

5. Analisa hubungan monthly charge dengan churn menggunakan t-test

```
##
## Welch Two Sample t-test
##
## data: data$MonthlyCharge by data$Churn
## t = -18.341, df = 4139.7, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group No and group Yes is not
equal to 0
## 95 percent confidence interval:
## -14.53786 -11.72998
## sample estimates:
## mean in group No mean in group Yes
## 61.30741 74.44133
```

Hasilnya $p\text{-value} < 0.05$, menunjukkan bahwa terdapat hubungan yang signifikan antara churn dengan pelanggan monthlycharge.

Saya melakukan Logistic Regression untuk melihat model yang cocok antara churn dengan variabel-variabel lainnya.

```
# Menghilangkan variabel non-signifikan dan membangun ulang model
m2 <- glm(churnNumerical ~ seniorCitizen + techSupportNumerical + multipleLinesNumerical +
paperlessBillingNumerical + oneYear + twoYear + onlineSecurityNumerical + fiberOptic + electronicCheck + tenure + TotalCharges,
          family = binomial, data = telco_final)
summary(m2)
```

```
##
## Call:
## glm(formula = churnNumerical ~ seniorCitizen + techSupportNumerical +
##     multipleLinesNumerical + paperlessBillingNumerical + oneYear +
##     twoYear + onlineSecurityNumerical + fiberOptic + electronicCheck +
##     tenure + TotalCharges, family = binomial, data = telco_final)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -7.830e-01  8.107e-02  -9.658  < 2e-16 ***
## seniorCitizen    3.046e-01  8.233e-02   3.700  0.000216 ***
## techSupportNumerical -2.368e-01  8.462e-02  -2.798  0.005138 **
## multipleLinesNumerical  1.951e-01  7.752e-02   2.517  0.011824 *
## paperlessBillingNumerical  4.566e-01  7.261e-02   6.288  3.22e-10 ***
## oneYear         -7.333e-01  1.048e-01  -6.997  2.61e-12 ***
## twoYear         -1.522e+00  1.724e-01  -8.830  < 2e-16 ***
## onlineSecurityNumerical -3.119e-01  8.370e-02  -3.727  0.000194 ***
## fiberOptic       8.095e-01  8.739e-02   9.262  < 2e-16 ***
## electronicCheck  4.937e-01  6.728e-02   7.337  2.18e-13 ***
## tenure          -6.183e-02  5.661e-03  -10.921  < 2e-16 ***
## TotalCharges     3.682e-04  6.104e-05   6.032  1.62e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 8143.4  on 7031  degrees of freedom
## Residual deviance: 5947.8  on 7020  degrees of freedom
## AIC: 5971.8
##
```

Melalui Logistic Regression yang telah dilakukan, model 3 memiliki variabel-variabel yang signifikan, sehingga kita akan menggunakan model 3 untuk prediksi modelnya. Dari hasil tersebut didapat hasil yang menarik bahwa:

- Senior Citizen dengan nilai ratio odds $\exp(0.304)$ menandakan bahwa pelanggan senior (Senior Citizen) memiliki kemungkinan untuk churn yang lebih tinggi daripada pelanggan yang lebih muda (non-senior). Dapat dikatakan juga bahwa pelanggan yang berusia lebih muda, memiliki kemungkinan yang lebih kecil untuk churn. Menariknya, pelanggan yang churn dari data tersebut adalah lebih banyak pelanggan yang non-senior. Hal ini karena pelanggan telco kebanyakan adalah non-senior (lebih muda), sedangkan pelanggan seniornya yang hanya sedikit juga, banyak yang keluar (churn).
- Koefisien positif dari TechSupport menunjukkan bahwa pelanggan tanpa dukungan teknis (TechSupport) lebih cenderung untuk churn dibanding pelanggan yang menggunakan.
- Koefisien untuk multipleLinesNumerical adalah 0.1951, menunjukkan bahwa pelanggan yang memiliki multiple lines memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan yang tidak memiliki multiple lines.

- Koefisien untuk `paperlessBillingNumerical` adalah 0.4566. Ini menunjukkan bahwa pelanggan yang menggunakan `paperless billing` memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan yang tidak menggunakan `paperless billing`.
- Koefisien untuk `oneYear` adalah -0.7333 dan untuk `twoYear` adalah -1.522. Ini menunjukkan bahwa pelanggan dengan kontrak selain `oneYear` dan `twoYear` yaitu `month to month` cenderung memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan dengan kontrak jangka panjang `one year` atau `two year`.
- Koefisien untuk `onlineSecurityNumerical` adalah -0.3119. Ini menunjukkan bahwa pelanggan yang tidak menggunakan layanan `online security` memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan yang menggunakan layanan tersebut.
- Koefisien untuk `fiberOptic` adalah 0.8095, menunjukkan bahwa pelanggan yang menggunakan layanan `Fiber Optic` memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan yang menggunakan `DSL` dan yang tidak menggunakan `internet service`.
- Koefisien untuk `electronicCheck` adalah 0.4937. Ini menunjukkan bahwa pelanggan yang menggunakan metode pembayaran elektronik (`electronicCheck`) memiliki peluang churn yang lebih tinggi dibandingkan dengan pelanggan yang menggunakan metode pembayaran lainnya.

Model regresi ini masih memiliki error rate hampir 20%, sehingga dibutuhkan analisa lebih lanjut untuk mendapatkan model yang lebih akurat.

```
gg2 = floor(pptest2+0.5)
ttt2 = table(ynew,gg2)
ttt2
```

```
##      gg2
## ynew  0   1
##    0 1857 188
##    1  382 386
```

```
error2 = (ttt2[1,2]+ttt2[2,1])/n2
error2
```

```
## [1] 0.2026306
```

Dari hasil yang di dapat, secara keseluruhan ada beberapa rekomendasi yang dapat diberikan:

1. **Tingkatkan Dukungan Teknis (TechSupport) dan layanan online security:**
 - Pelanggan yang tidak memiliki dukungan teknis dan tidak memiliki online security cenderung lebih tinggi untuk churn. Oleh karena itu, perusahaan dapat meningkatkan marketing, meningkatkan kualitas Tech Support, dapat mengurangi jumlah churn pelanggan. Namun perusahaan juga perlu memperhatikan apakah ada masalah-masalah lain dalam menggunakan layanan-layanan perusahaan sehingga pelanggan yang tidak menggunakan layanan TechSupport dan online security memilih untuk keluar (churn).
2. **Promosikan Kontrak Jangka Panjang :**
 - Pelanggan dengan kontrak satu tahun atau dua tahun lebih kecil kemungkinannya untuk churn dibandingkan dengan pelanggan dengan kontrak bulan-ke-bulan. Perusahaan mungkin bisa menawarkan promo atau diskon untuk mendorong pelanggan agar memilih kontrak jangka panjang, atau juga dapat memberikan tawaran-tawaran menarik lainnya. Namun perusahaan juga perlu melihat alasan-alasan lain mengapa pelanggan yang kontrak bulanan lebih cenderung untuk churn, apakah karena layanan perusahaan, atau alasan-alasan lainnya.
3. **Evaluasi Metode Pembayaran:**
 - Pelanggan yang membayar menggunakan electronic check cenderung lebih tinggi untuk churn. Perusahaan dapat mengedukasi dan menyarankan pelanggan tentang metode pembayaran yang lebih stabil dan mudah, seperti pembayaran otomatis melalui kartu kredit atau debit. Namun perusahaan juga perlu mengevaluasi kembali apakah pembayaran menggunakan electronic check masih efektif atau tidak, dan alternatif atau cara lain agar pembayaran menggunakan metode ini lebih efektif.
4. **Tingkatkan layanan untuk Pelanggan Senior:**
 - Pelanggan senior cenderung churn daripada pelanggan non-senior. Perusahaan dapat menawarkan layanan yang lebih disesuaikan untuk pelanggan yang lebih senior, seperti paket khusus atau layanan pelanggan yang lebih personal, dapat meningkatkan kualitas layanan bagi pelanggan yang senior. Perusahaan juga dapat meningkatkan strategi pemasaran yang lain yang lebih menarik dan sesuai kebutuhan bagi pelanggan yang senior.
5. **Tingkatkan Pengalaman Pelanggan dengan Layanan Internet:**
 - Pelanggan dengan layanan fiber optic cenderung lebih tinggi untuk churn. Hal ini dapat dievaluasi oleh perusahaan, apakah layanan fiber optic sering memiliki gangguan, dan lain-lain, serta hal apa yang membuat pelanggan tidak berlangganan lagi. Perusahaan perlu meningkatkan kualitas dari layanan FiberOptic
6. **Strategi Penetapan Harga dan Penawaran Khusus yang menarik:**
 - Pelanggan dengan total biaya yang lebih tinggi cenderung lebih tinggi untuk churn. Perusahaan dapat mempertimbangkan untuk memberikan penawaran

husus atau diskon untuk pelanggan dengan tagihan bulanan yang tinggi untuk meningkatkan loyalitas mereka.

7. Program Loyalitas:

- Perusahaan juga dapat mengembangkan program-program menarik seperti program yang memberikan penghargaan kepada pelanggan setia atau mereka yang telah menggunakan layanan dalam jangka waktu lama. Misalnya, memberikan bonus atau diskon berdasarkan masa layanan, agar pelanggan yang memiliki tenure yang lama juga dapat bertahan.

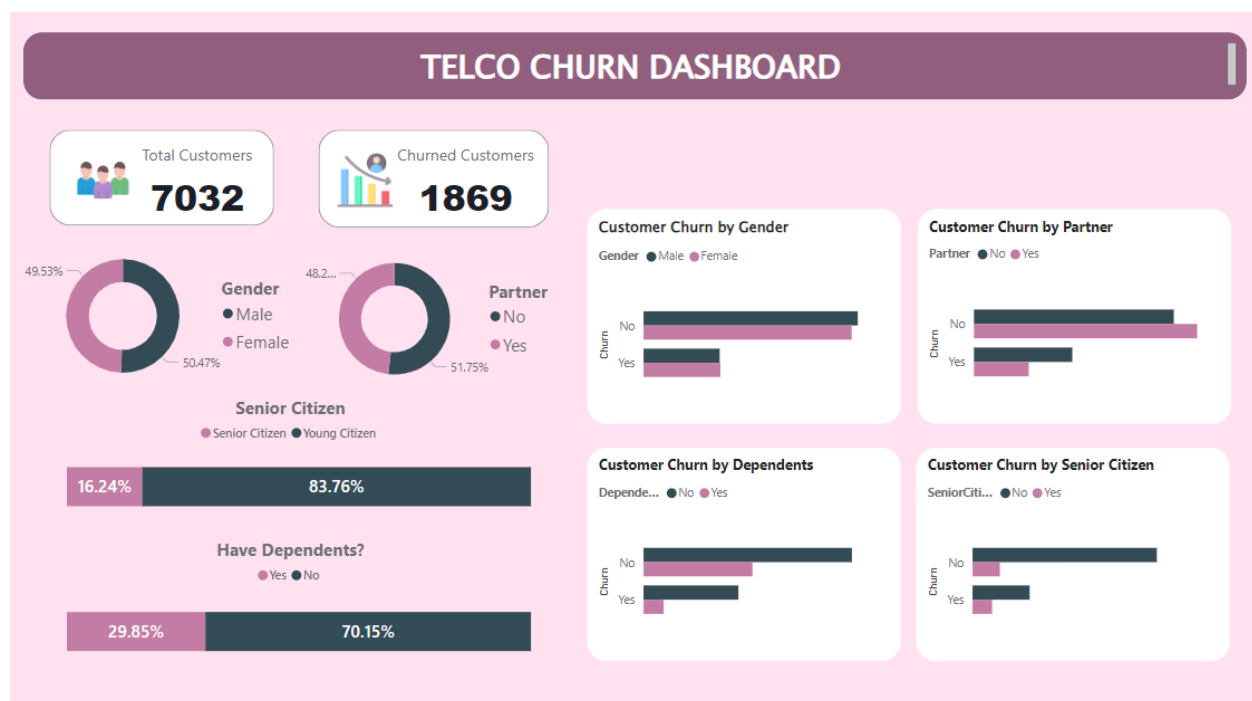
8. Penggunaan Teknologi dan Data:

- Menggunakan analitik prediktif untuk mengidentifikasi pelanggan yang berisiko churn dan mengembangkan intervensi yang disesuaikan. Perusahaan dapat menggunakan data dari model regresi untuk memprioritaskan upaya retensi pada kelompok pelanggan tertentu.

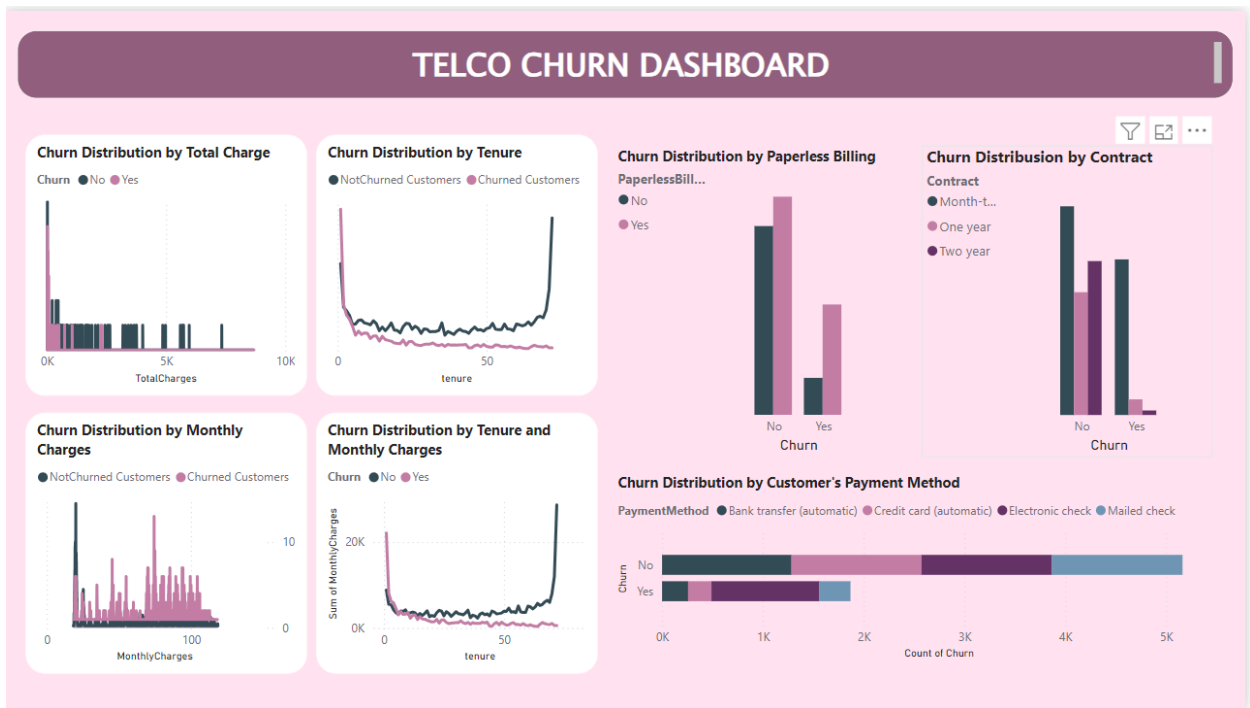
Perusahaan perlu mengecek dan memperbaiki kualitas dari layanan-layanan yang ditawarkan dan perlu mengembangkannya secara terus menerus.

Secara keseluruhan, perusahaan perlu secara rutin untuk mengumpulkan feedback dari pelanggan untuk mengidentifikasi area yang perlu diperbaiki. Dengan ini, perusahaan dapat meningkatkan retensi pelanggan dan mengurangi churn, yang pada akhirnya akan meningkatkan profit dan pertumbuhan jangka panjang.

Customer Churned:



Customer Account:



Service and Churn

