

Raport 1

SD

23 maja 2023

Spis treści

1	Opis eksperymentów	1
2	Etap.1	1
3	Etap.2a	2
4	Etap.2b Zmienna Ilościowe	3
5	Etap 2c Wnioski	9
6	Etap 2d Wnioski	9
7	Etap 3a	10
8	Etap 3b	14
9	Podsumowanie	14
9.1	Etap 4a	14
9.2	Etap 4b	15
9.3	Etap 4c	15

1 Opis eksperymentów

Celem sprawozdania jest odnalezienie przyczyny odejścia klientów. Do zbadania tego zjawiska zostaną użyte następujące narzędzia:

- Podstawowe wskaźniki sumaryczne.
- Metody graficzne: Wykresy słupkowe, rozrzutu oraz pudełkowe.

2 Etap.1

```
dane <- read.csv(file="WA_Fn-UseC_-Telco-Customer-Churn.csv", stringsAsFactors = TRUE)

attach(dane)
dane$SeniorCitizen <- as.factor(dane$SeniorCitizen)
```

```
#Powinna być to zmienna jakościowa więc musimy ją zamienić.
dim(dane) #liczba przypadków/cech
```

```
[1] 7043 21
```

```
sum(sapply(dane, is.factor)) #ile cech jakościowych
```

```
[1] 18
```

```
sum(sapply(dane, is.numeric)) #ile cech ilościowych
```

```
[1] 3
```

```
sum(sapply(dane, is.nan)) #Sprawdzamy czy jakaś kolumna jest pusta
```

```
[1] 0
```

```
which(sapply(dane, is.numeric) == TRUE)
```

```
tenure MonthlyCharges TotalCharges 6 19 20
```

```
#Poszczególne cechy ilościowe.
#Pozostałe dane mają typ jakościowy.
```

```
dane <- dane[-1] #usuwamy kolumnę z identyfikatorem klienta.
```

```
# 11 brakujących obserwacji znajdują się w kolumnie TotalCharges
which(is.na(dane), arr.ind=TRUE)
```

```
row col [1,] 489 19 [2,] 754 19 [3,] 937 19 [4,] 1083 19 [5,] 1341 19 [6,] 3332 19 [7,] 3827 19
[8,] 4381 19 [9,] 5219 19 [10,] 6671 19 [11,] 6755 19
```

3 Etap.2a

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	0.00	72.00	32.37	29.00	603.17	24.56	9.00	55.00	72.00

Tabela 1: Podstawowe wskaźniki sumaryczne względem kolumny tenure.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.25	118.75	64.76	70.35	905.41	30.09	35.50	89.85	100.50

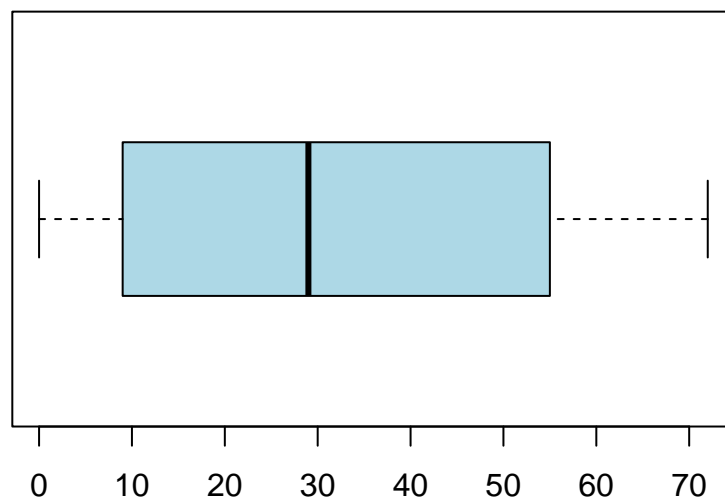
Tabela 2: Podstawowe wskaźniki sumaryczne względem kolumny MonthlyCharges

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.80	8684.80	2283.30	1397.47	5138252.41	2266.77	401.45	3794.74	8666.00

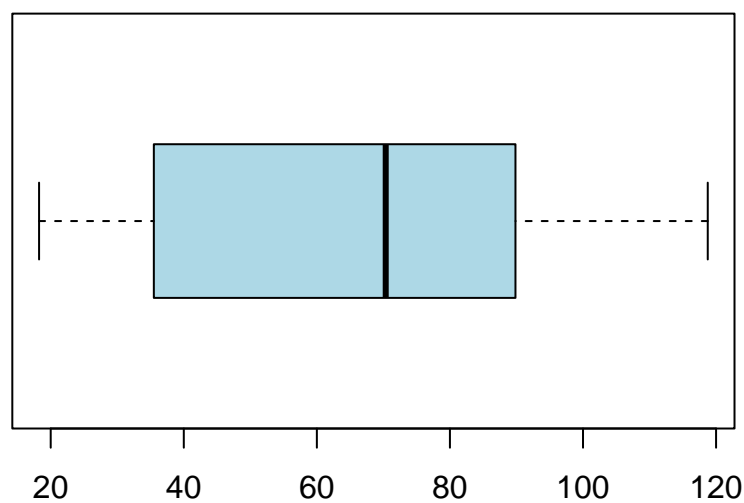
Tabela 3: Podstawowe wskaźniki sumaryczne względem kolumny TotalCharges.

4 Etap.2b Zmienna Ilościowe

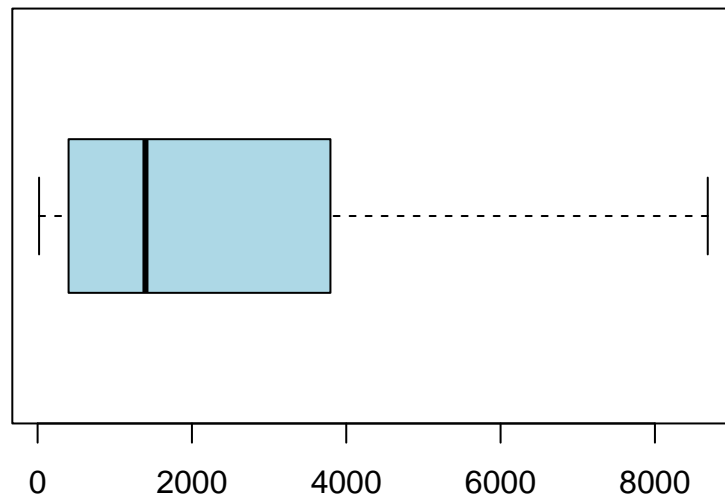
Wykres pudełkowy względem kolumny tenure.



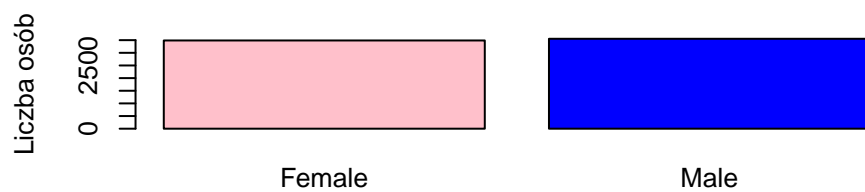
Wykres pudełkowy względem kolumny MonthlyCharg



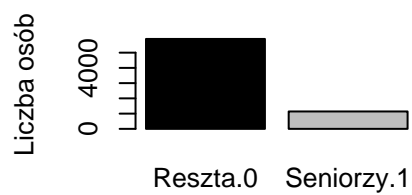
Wykres pudełkowy względem kolumny TotalCharge



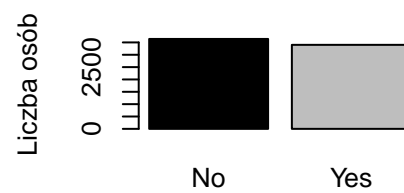
Klienci względem płci

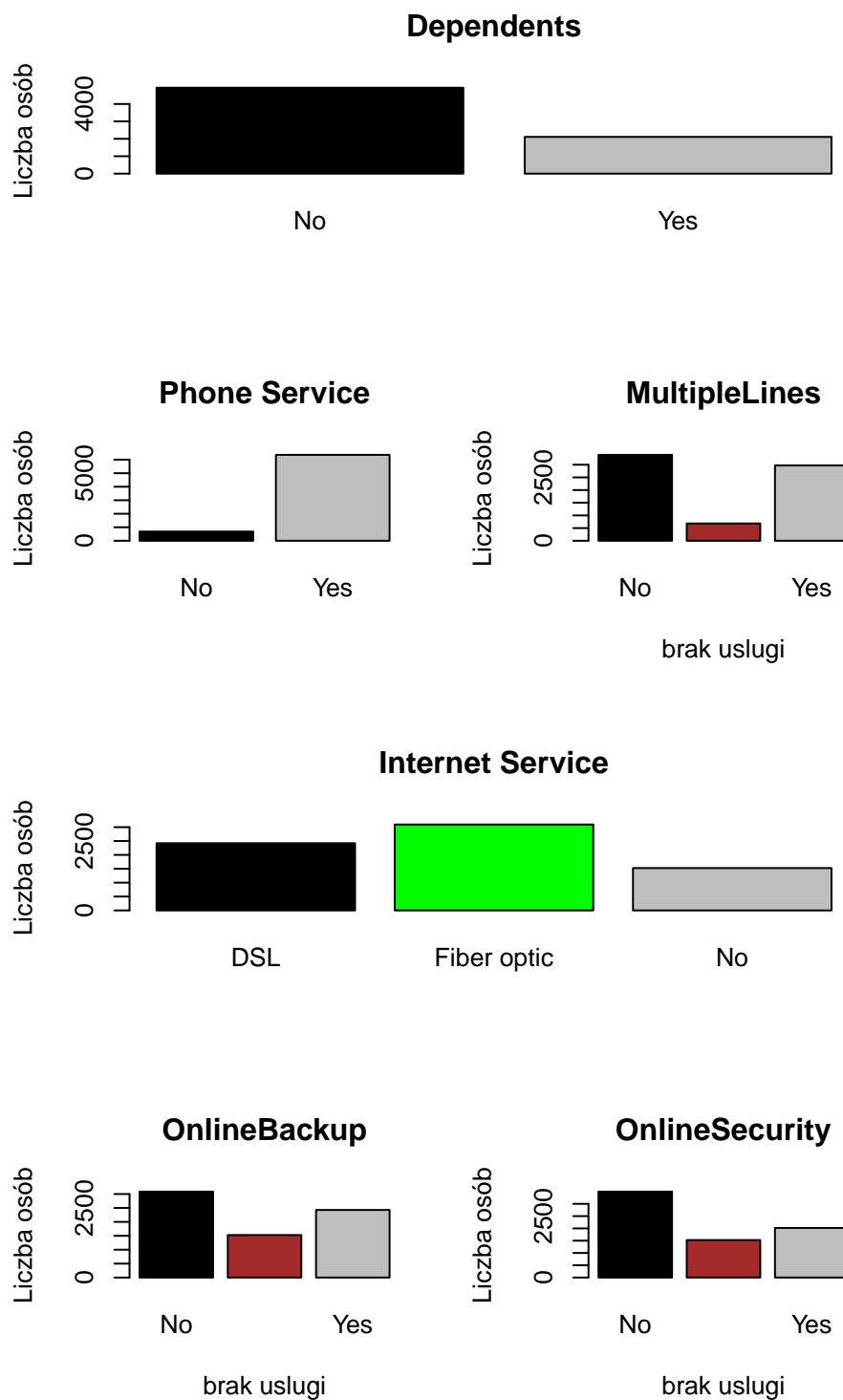


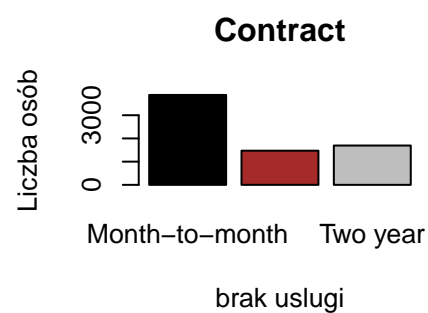
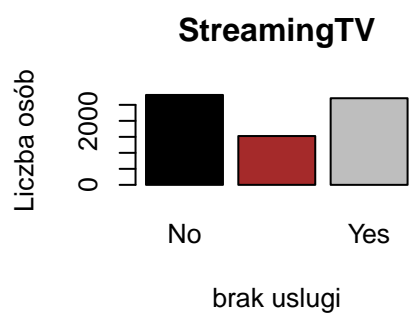
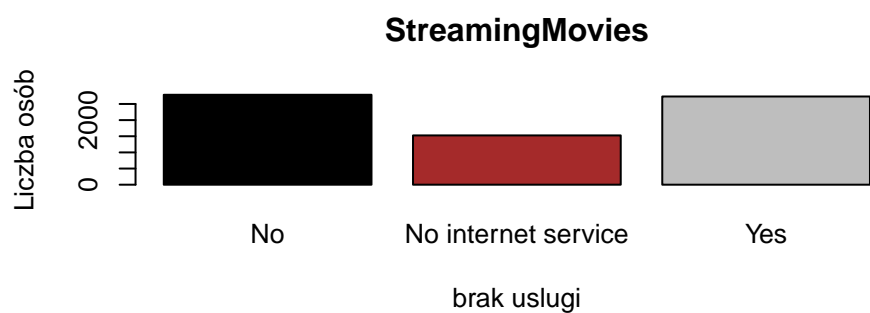
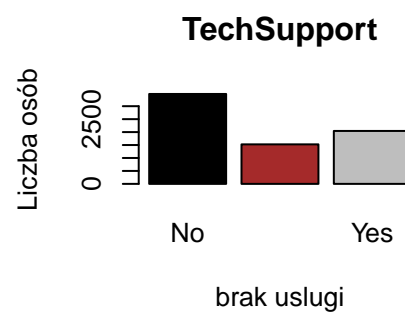
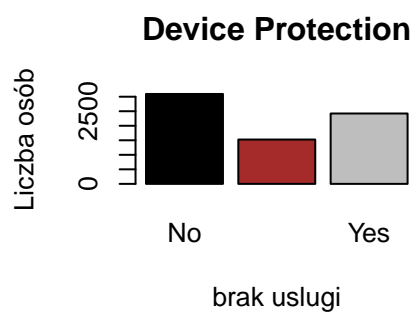
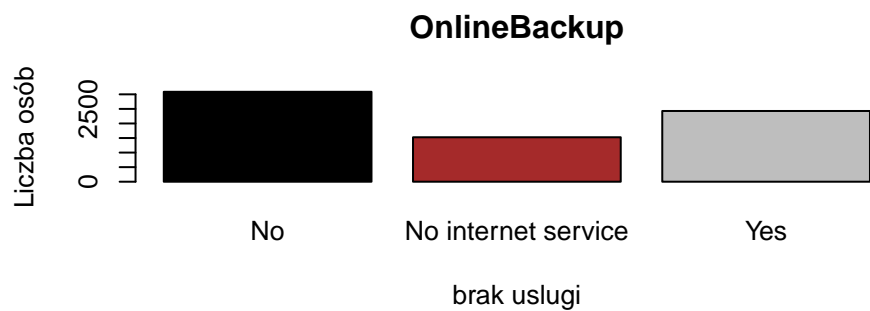
Liczba Seniorów

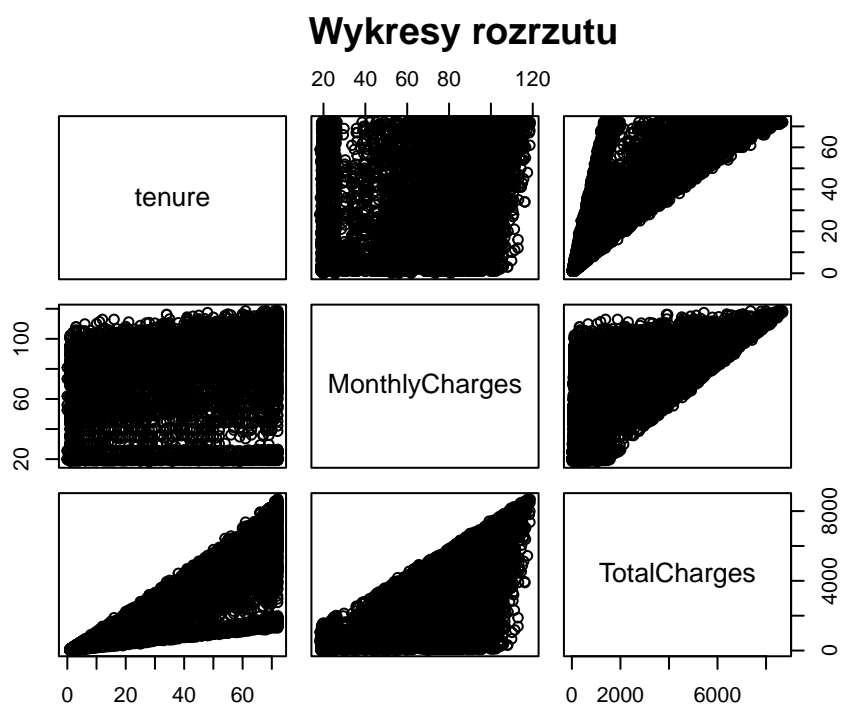
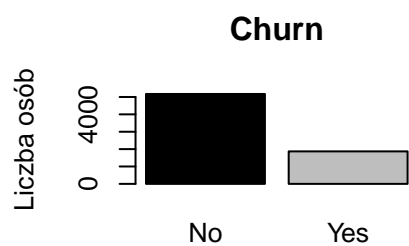
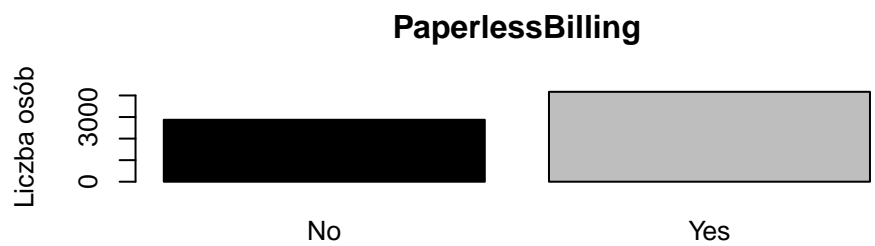


Partner

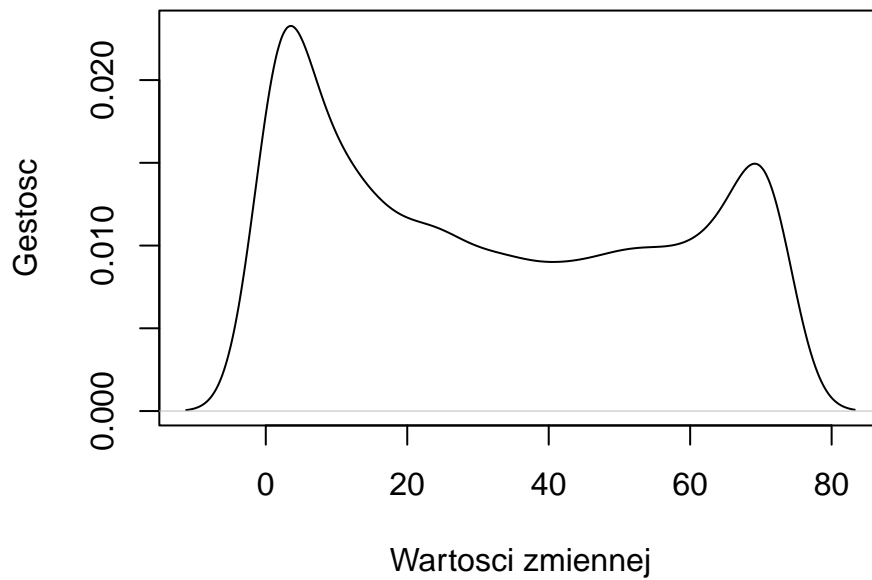




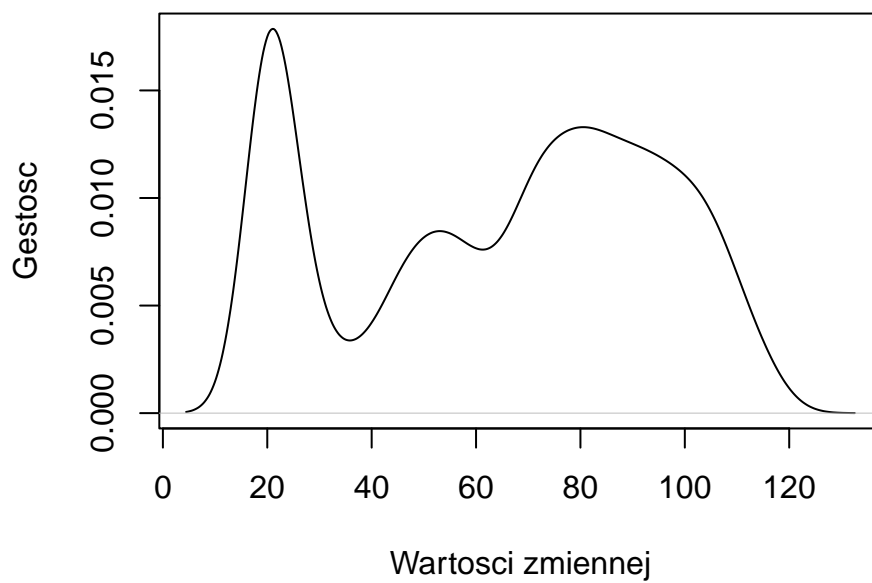




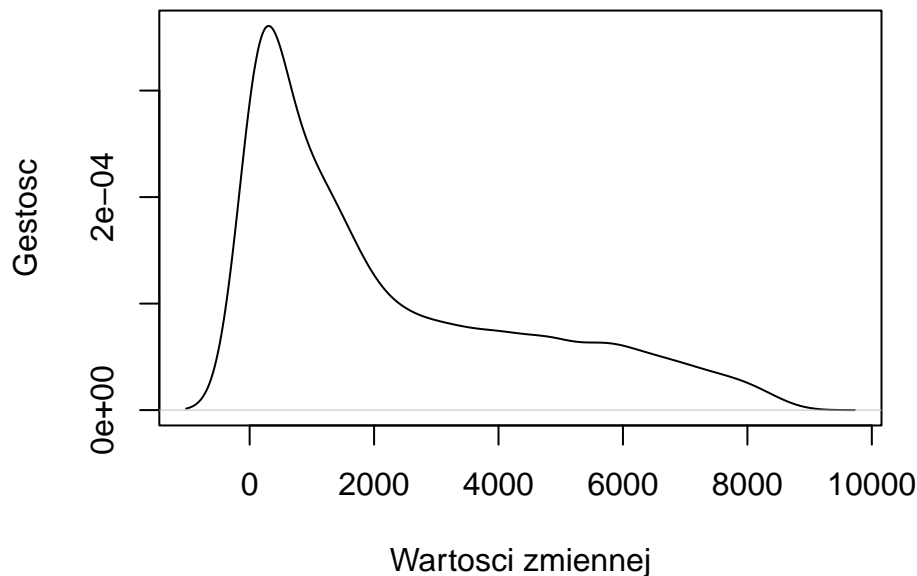
Wykres gestosci tenure



Wykres gestosci MonthlyCharges



Wykres gestosci TotalCharges



5 Etap 2c Wnioski

Wiadać korelacje miedzy tenure a TotalCharges. Wnioskując z wykresu- im dłużej ktoś jest klientem tym więcej ma opłat ogółem.

6 Etap 2d Wnioski

- Zakres dla zmiennych jakościowych to min-max, a dla zmiennych ilościowych został przedstawiony na wykresach kolumnowych.
- Wykres gestosci MonthlyCharges oraz TotalCharges nie jest symetryczny więc nie ma rozkładu symetrycznego.
- Największą zmiennością charakteryzuje się zmienna TotalCharges, ponieważ odchylenie standardowe jest największe.
- Częstość:
 - Liczba kobiet jest podobna do liczby mężczyzn
 - Seniorzy są w mniejszości
 - Liczba klientów którzy mają partnera jest podobna do tych bez.
 - Większość klientów ma usługę telefoniczną.
 - Większość klientów ma usługę Fiber optic.
 - Większość klientów nie posiada usługi OnlineBackup, OnlineSecurity, Device Protection, techSupport, StreamingMovies, StreamingTv.

- Większość klientów posiada kontrakt na miesiąc.
- Większość klientów otrzymuje elektroniczny rachunek.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	1.00	72.00	17.98	10.00	381.46	19.53	2.00	29.00	71.00

Tabela 4: Klienci którzy zrezygnowali względem kolumny tenure.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	0.00	72.00	37.57	38.00	581.47	24.11	15.00	61.00	72.00

Tabela 5: Aktywni Klienci względem kolumny tenure.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.85	118.35	74.44	79.65	608.41	24.67	56.15	94.20	99.50

Tabela 6: Klienci którzy zrezygnowali względem kolumny MonthlyCharges.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.25	118.75	61.27	64.43	966.75	31.09	25.10	88.40	100.50

Tabela 7: Aktywni klienci względem kolumny MonthlyCharges.

	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.85	8684.80	1531.80	703.55	3575211.60	1890.82	134.50	2331.30	8665.95

Tabela 8: Klienci którzy zrezygnowali względem kolumny TotalCharges.

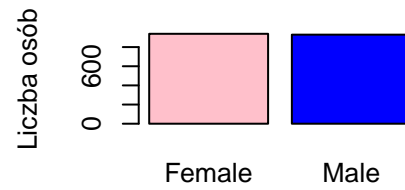
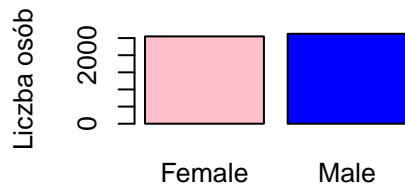
	Min	Max	Średnia	Mediana	War.	Odch.St	Q1..25%	Q3..75%	Rozstęp
1	18.80	8672.45	2555.34	1683.60	5426369.84	2329.46	577.83	4264.12	8653.65

Tabela 9: Aktywni klienci względem kolumny TotalCharges.

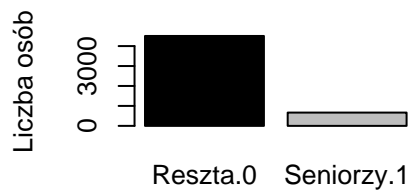
7 Etap 3a

0 1 4508 666 Female Male 2549 2625 Female Male 939 930

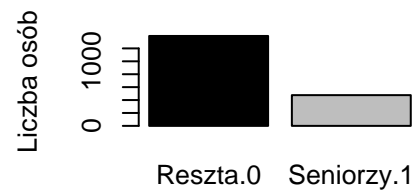
Aktywni Klienci względem płci którzy zrezygnowali względ



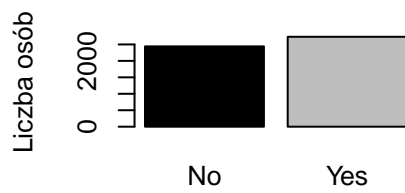
Aktywni Seniorów



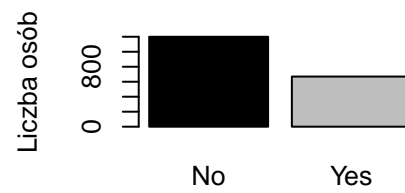
Seniorzy którzy zrezygnowal



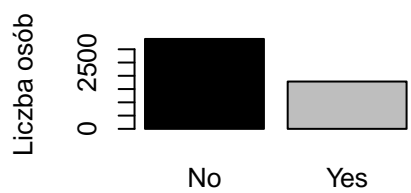
Partner Churn_n



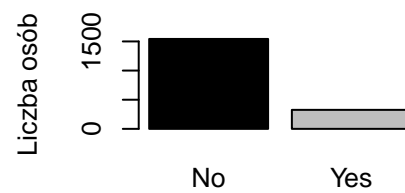
Partner Churn_y

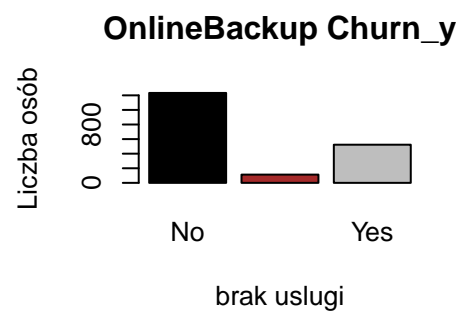
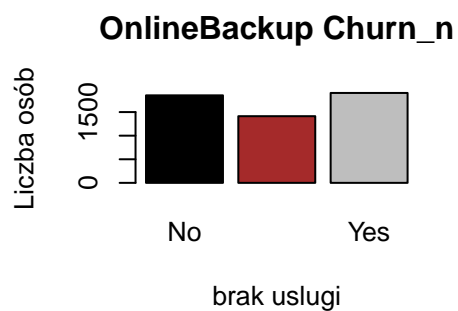
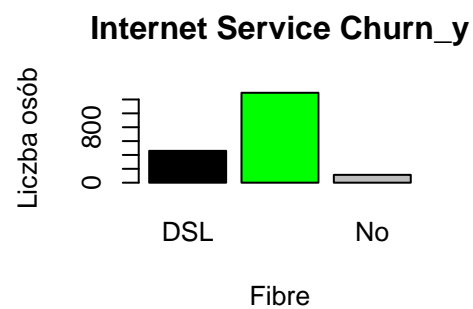
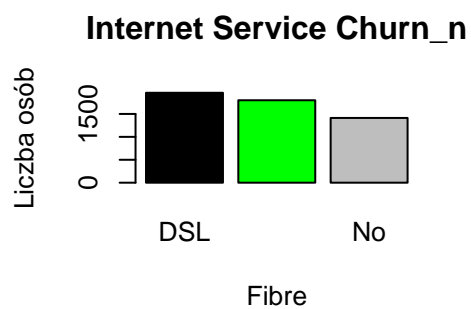
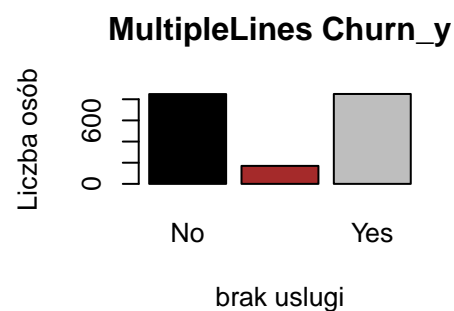
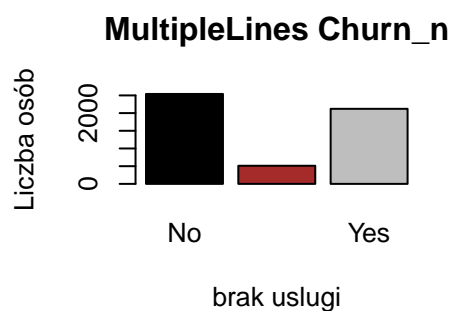
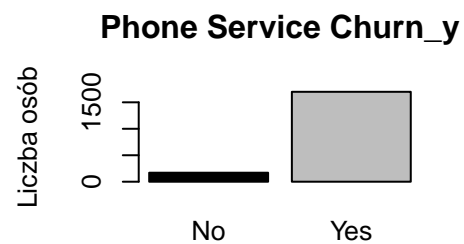
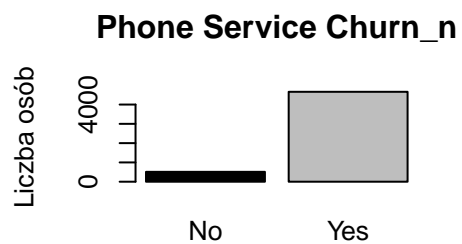


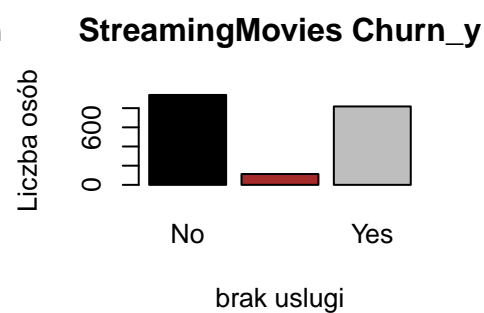
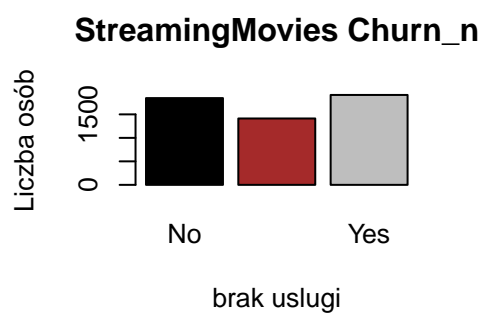
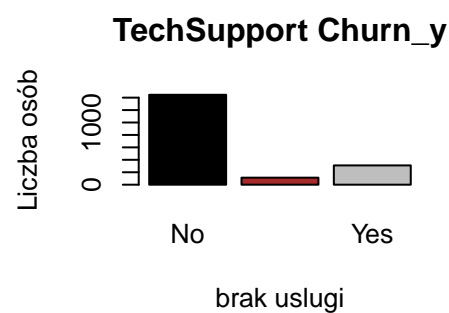
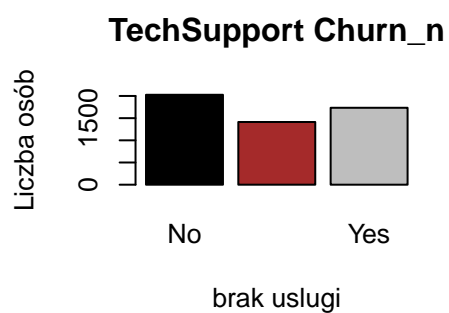
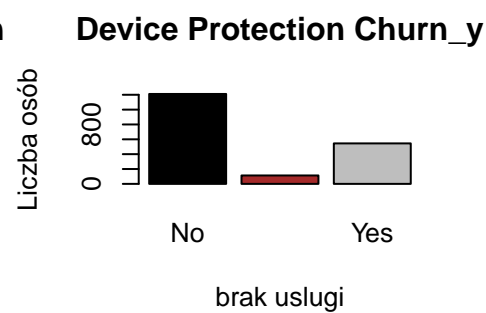
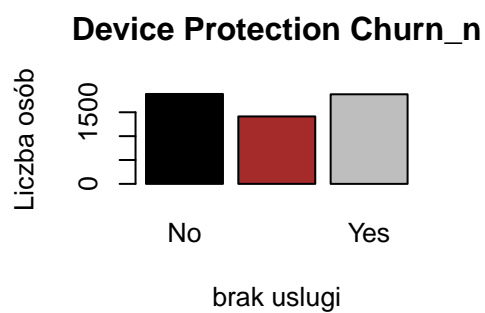
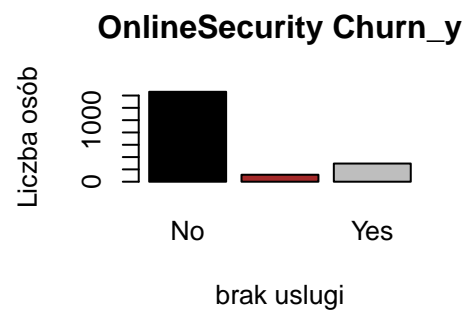
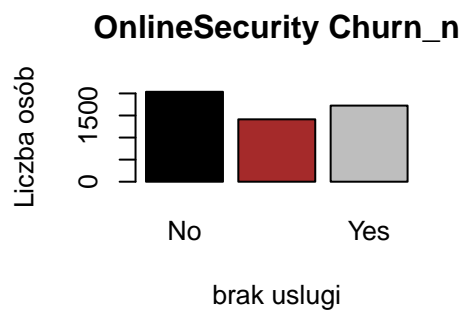
Dependents Churn_n

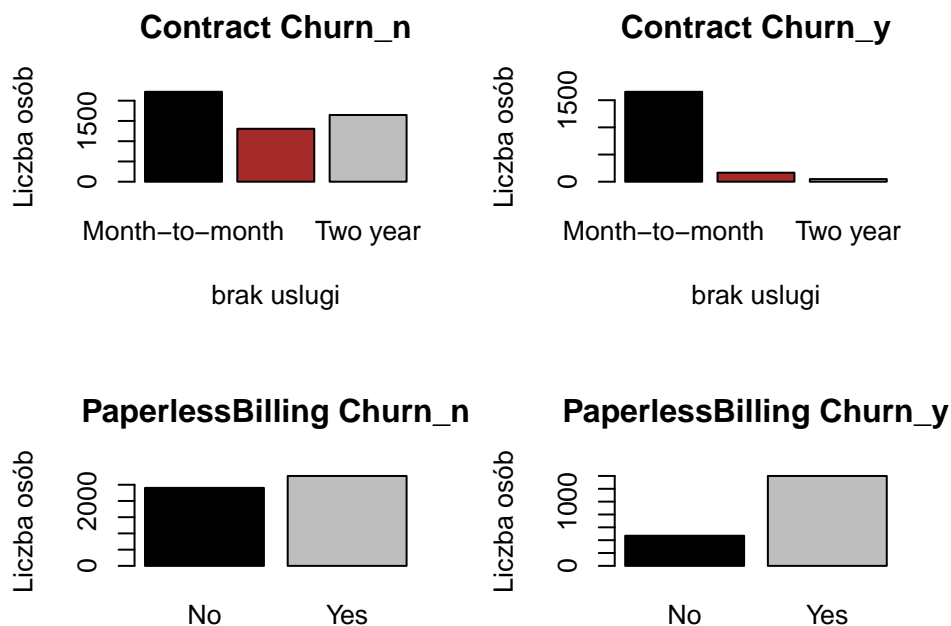


Dependents Churn_y









8 Etap 3b

Klienci którzy zrezygnowali:

- Częściej nie mają partnera.
- W większości korzystali z usługi Fibre.
- Rzadko korzystali z pomocy technicznej.
- W większości mieli umowę na miesiąc.

9 Podsumowanie

9.1 Etap 4a

Najważniejsze wnioski, jakie udało się wysnuć na podstawie przeprowadzonych analiz:

- Wysoka wariancja względem TotalCharges
- Ponad połowa klientów korzysta z usług firmy od więcej niż 30 miesięcy.
- Zakres miesięcznych wydatków klientów mieści się w przedziale 20-120.
- Ponad połowa klientów wydała więcej niż 2000 na usługi firmy, a 25% wydało więcej niż 4000.

9.2 Etap 4b

Z danych dowiadujemy się, że liczba klientek jest praktycznie równa liczbie klientów. Główną produktem u klientów są usługi telefoniczne, oraz internet, który jest w głównej mierze zapewniany poprzez technologie Fiber optic. Dużą popularność ma również usługa StreamingTV. Klienci w głównej mierze rezygnują z papierowych rachunków.

9.3 Etap 4c

Powody rezygnacji:

- Brak korzystnych ofert dla osób bez partera.
- Problem z usługą Fibre optic.
- Brak kontaktu z klientem w trakcie awarii.
- Umowa na krótki okres.

Aby klienci przestali odchodzić należy:

- Stworzyć więcej ofert dla singli.
- Przeciwdziałać awarii Fibre Optic lub poprawić działanie tej usługi.
- W trakcie awarii kontaktować się z klientami.
- Nakłaniać klientów na dłuższy okres trwania umowy.