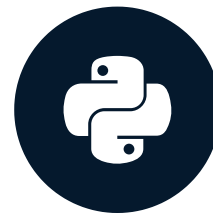


Vítejte v kurzu

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON



Mark Peterson

Director of Data Science, Infoblox

Analytika odchodů zákazníků





Odchod zákazníků

Stávající zákazník přestane obchodovat s danou společností

Smluvní odchod



Dobrovolný odchod



Nesmluvní odchod



Nedobrovolný odchod: Vypršení platební karty



Nedobrovolný odchod: Odpojení služeb



Využití vlastních zkušeností

- Zákazník
 - Nízké využívání
 - Nízká kvalita služeb
 - Lepší cena jinde
- Znalost oblasti/odvětví

Datová sada Telco Churn

Popis	Hodnota
Záznamy	3333
Příznaky	21
Spojité	15
Kategorické	6

Klíčové příznaky

- Hlasová pošta
- Mezinárodní volání
- Cena za službu
- Využívání zákazníkem
- Odchod zákazníka

Definice odchodu v tomto kurzu

Zákazník zruší svůj mobilní tarif v daném okamžiku

- "no"
- "yes"

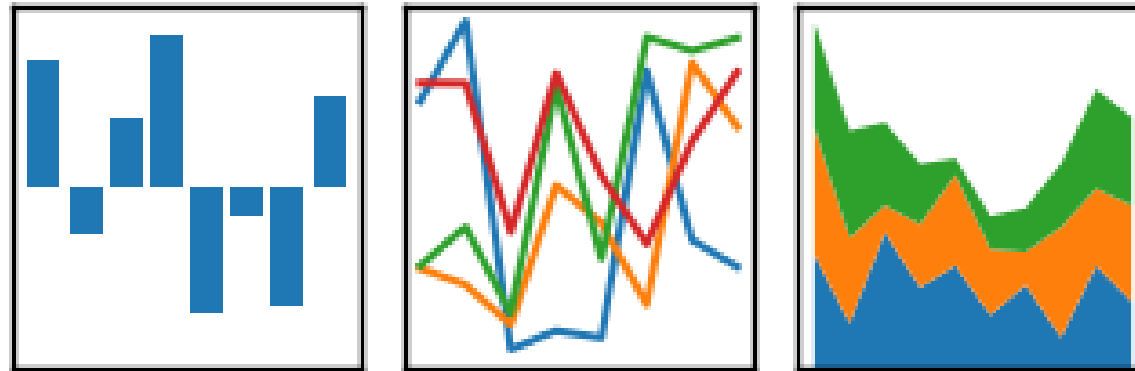
Exploratorní analýza dat pomocí pandas

- Pochopení příznaků datové sady
- Výpočet souhrnných statistik

Exploratorní analýza dat pomocí pandas

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



pandas Foundations

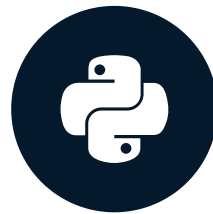
- `df.head()`
- `df.describe()`
- `df.mean()`

Prozkoumejme data!

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON

Seskupování a sumarizace dat

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON



Mark Peterson

Director of Data Science, Infoblox

Zákazníci s odchodem a bez odchodu

```
print(telco['Churn'].value_counts())
```

```
no      2850  
yes      483  
Name: Churn, dtype: int64
```

Výstupy modelu

Dvě třídy:

- `'yes'` : Zákazník odejde
- `'no'` : Zákazník neodejde

Rozdíly mezi zákazníky s odchodem a bez odchodu

- Volají zákazníci s odchodem zákaznickou podporu častěji?
- Má některý stát více odcházejících zákazníků než jiný?

Seskupování a sumarizace dat

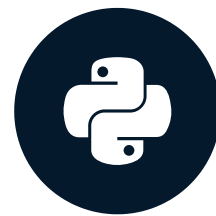
- `.groupby()`

Pojďme seskupovat a sumarizovat!

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON

Zkoumání dat pomocí vizualizací

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON



Mark Peterson

Director of Data Science, Infoblox

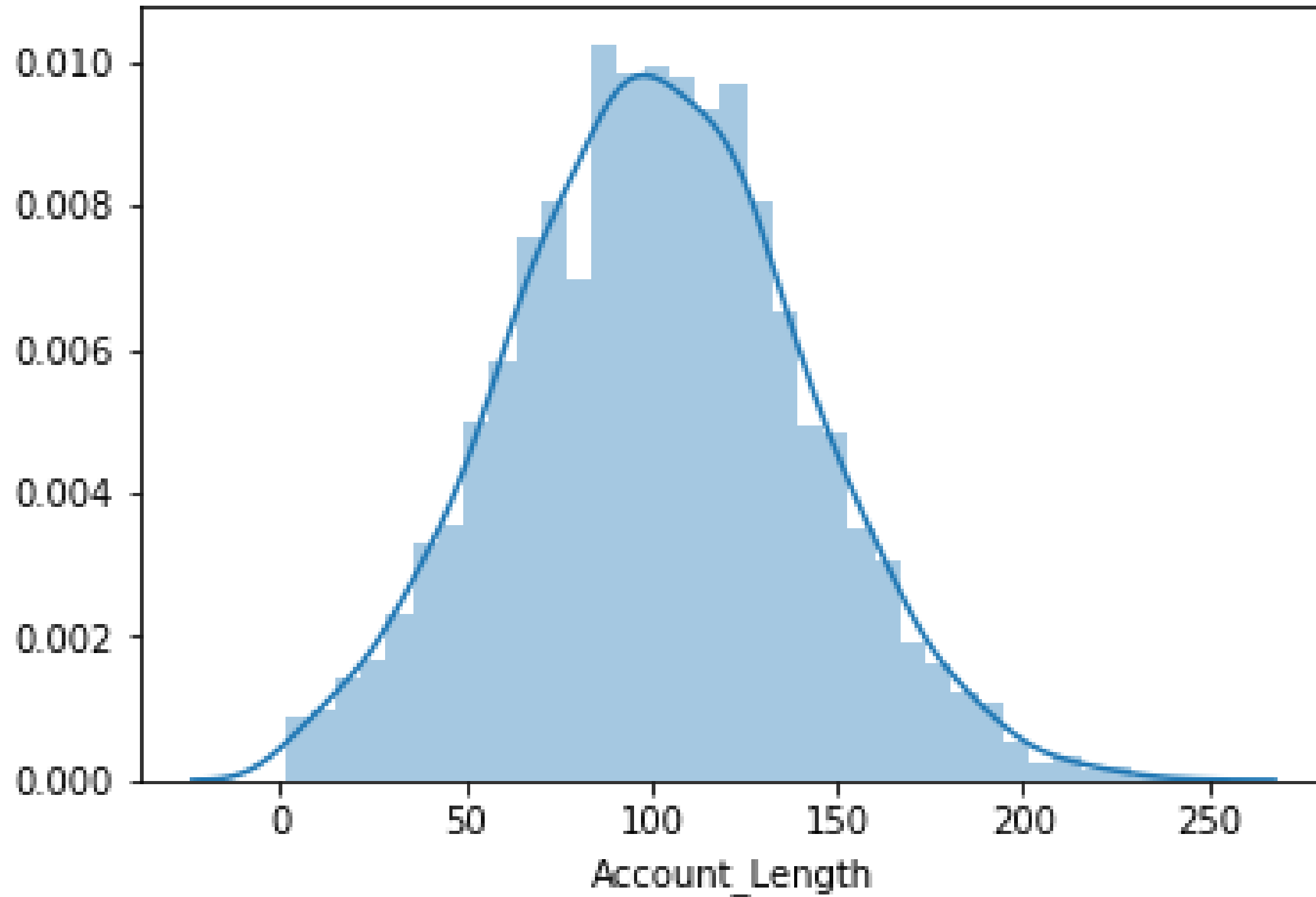
Vizualizace dat v Pythonu

- Knihovna `seaborn` umožňuje snadno vytvářet přehledné a atraktivní grafy
- Staví na `matplotlib`

Vizualizace rozdělení délky účtu

- Důležité je porozumět rozdělení proměnných

```
import matplotlib.pyplot as plt
import seaborn as sns
sns.distplot(telco['Account_Length'])
plt.show()
```

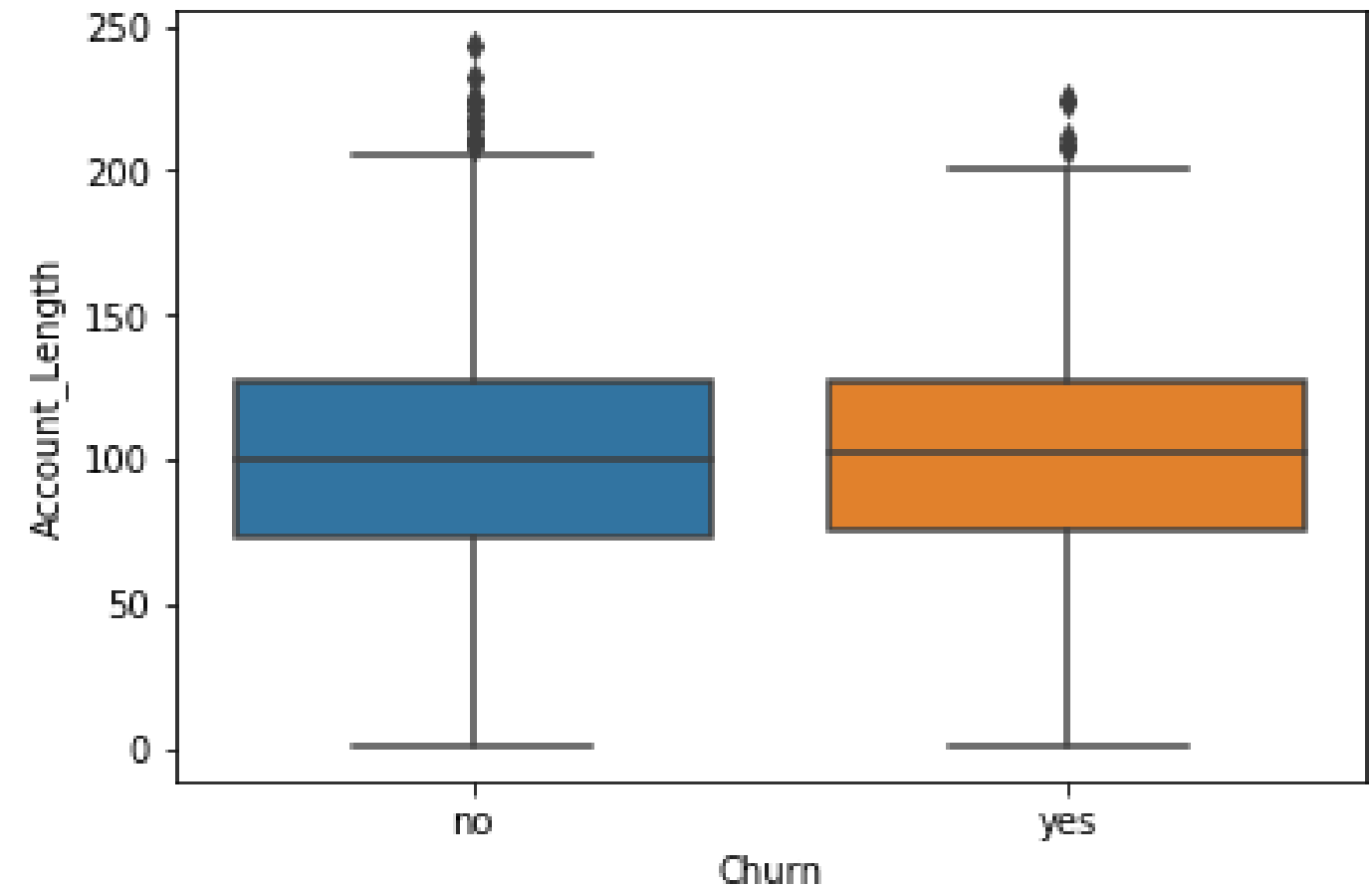


Rozdíly v délce účtu

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

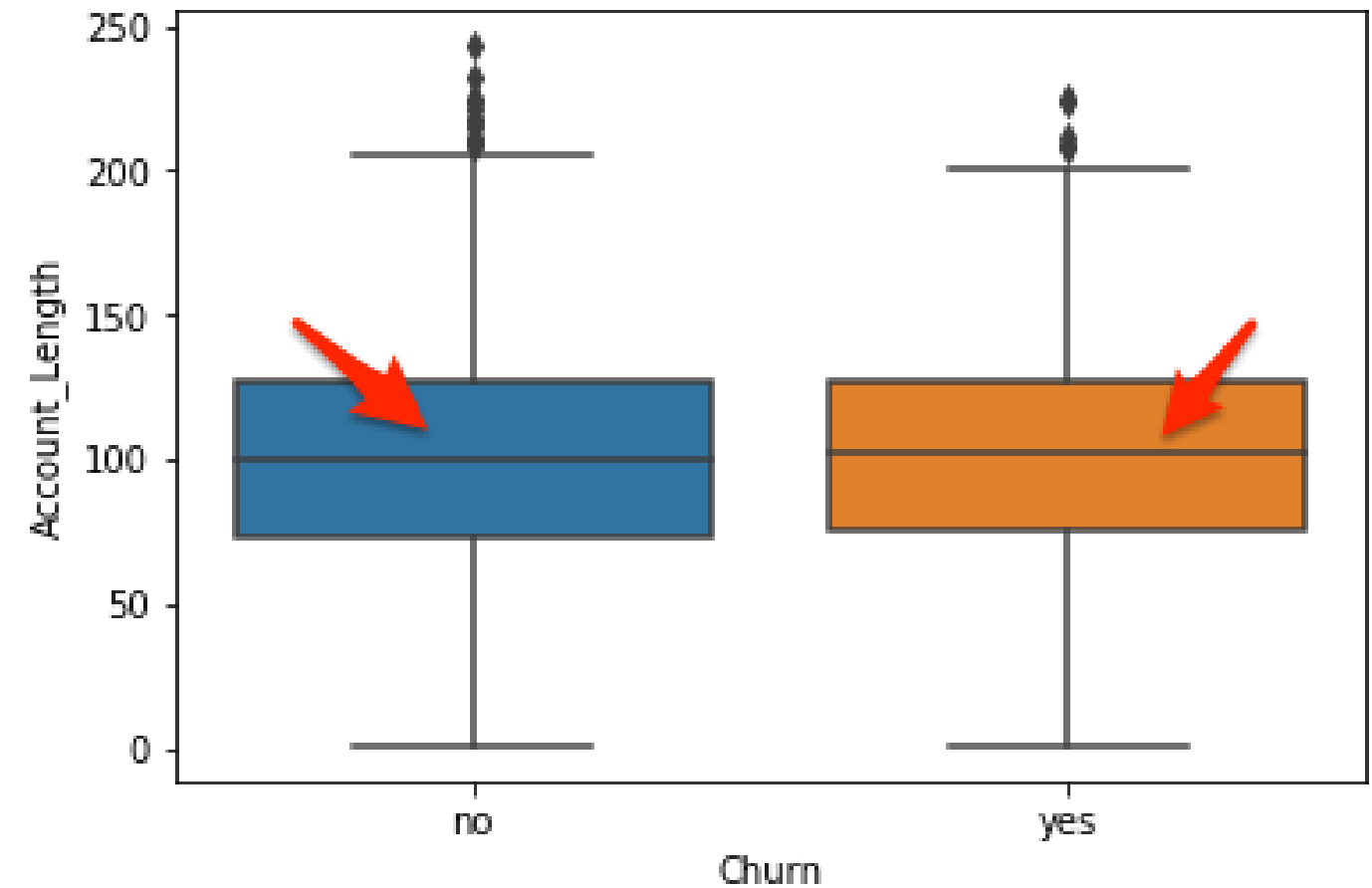


Rozdíly v délkách účtů

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

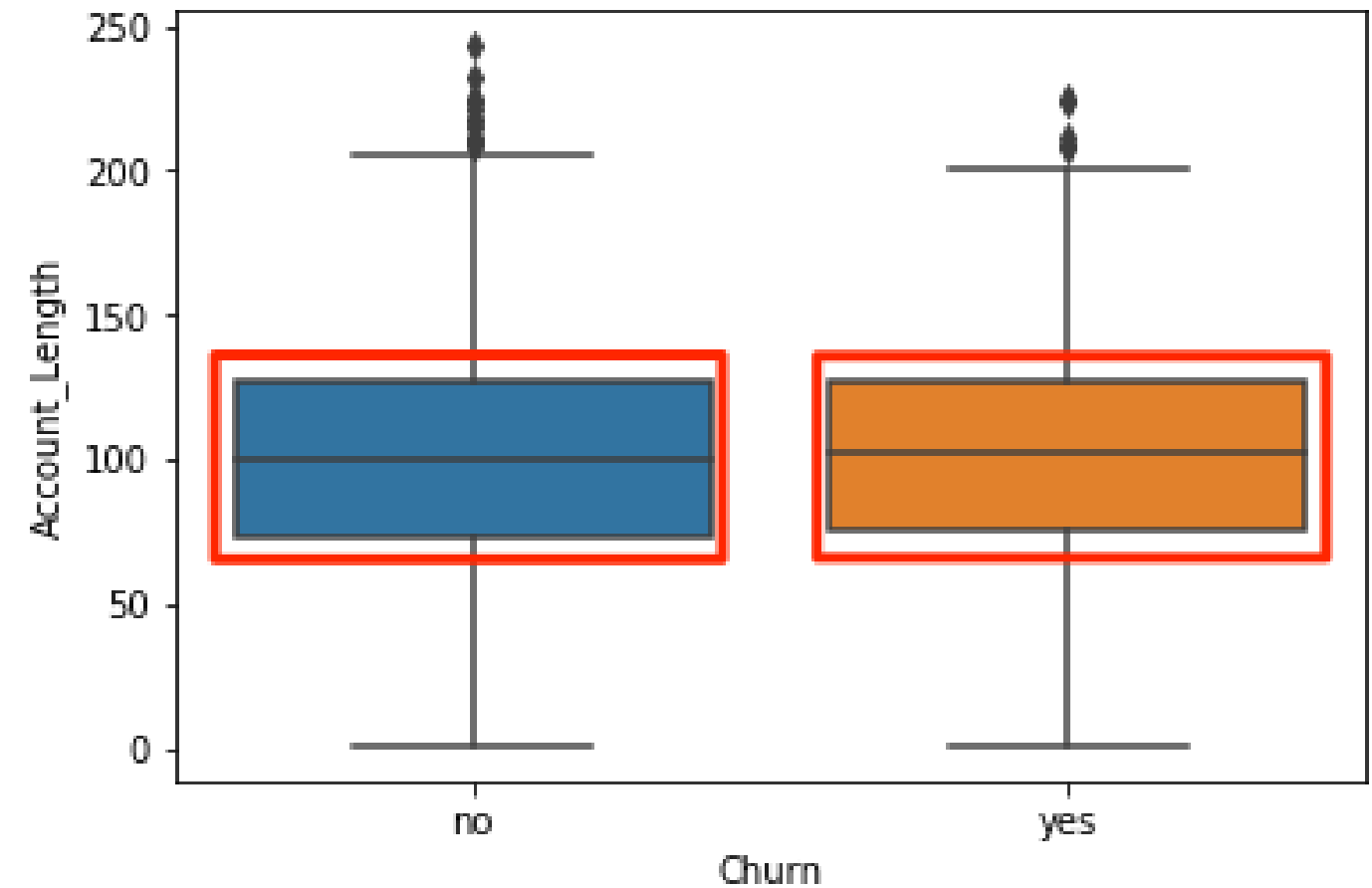


Rozdíly v délkách účtů

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

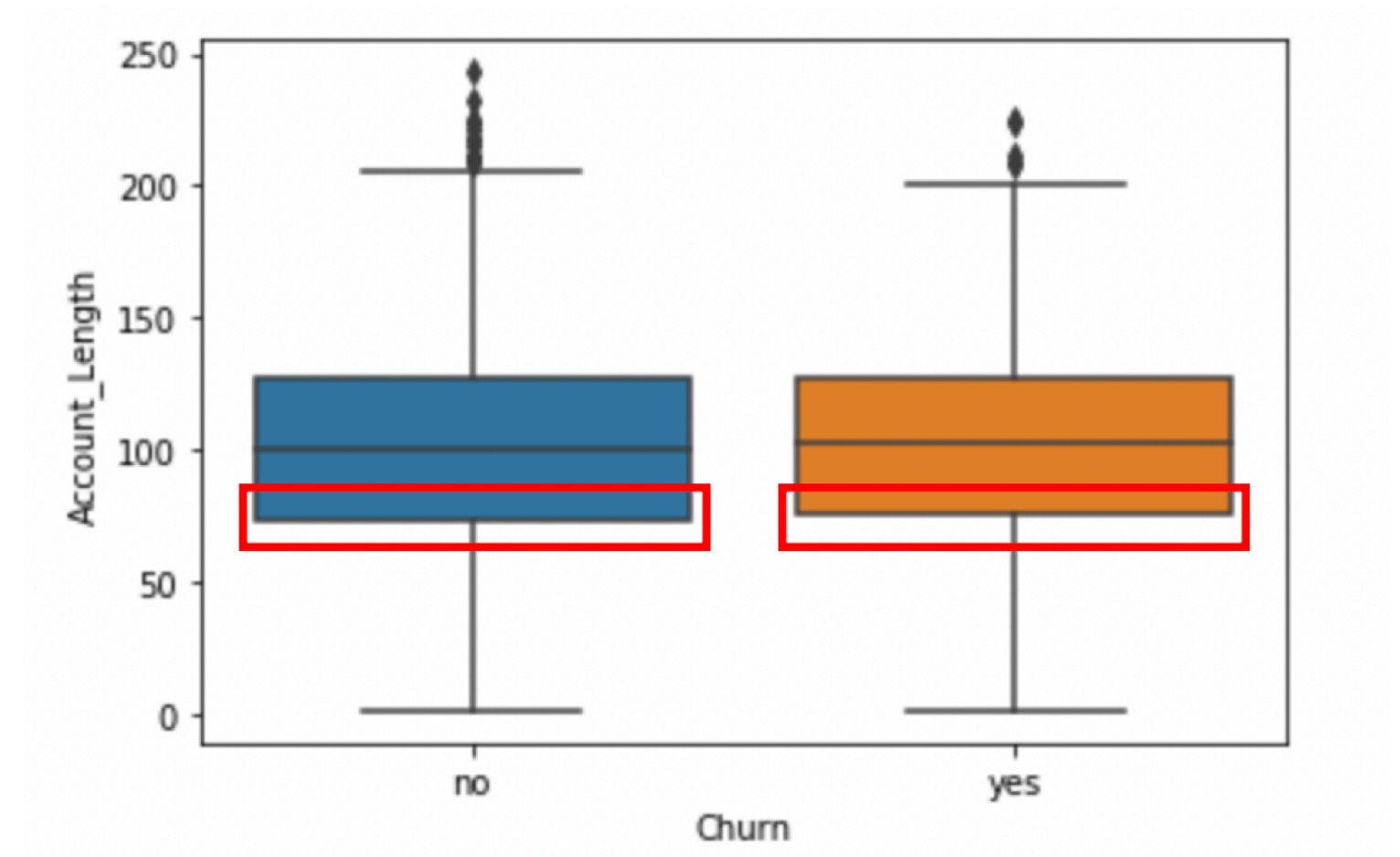


Rozdíly v délkách účtů

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

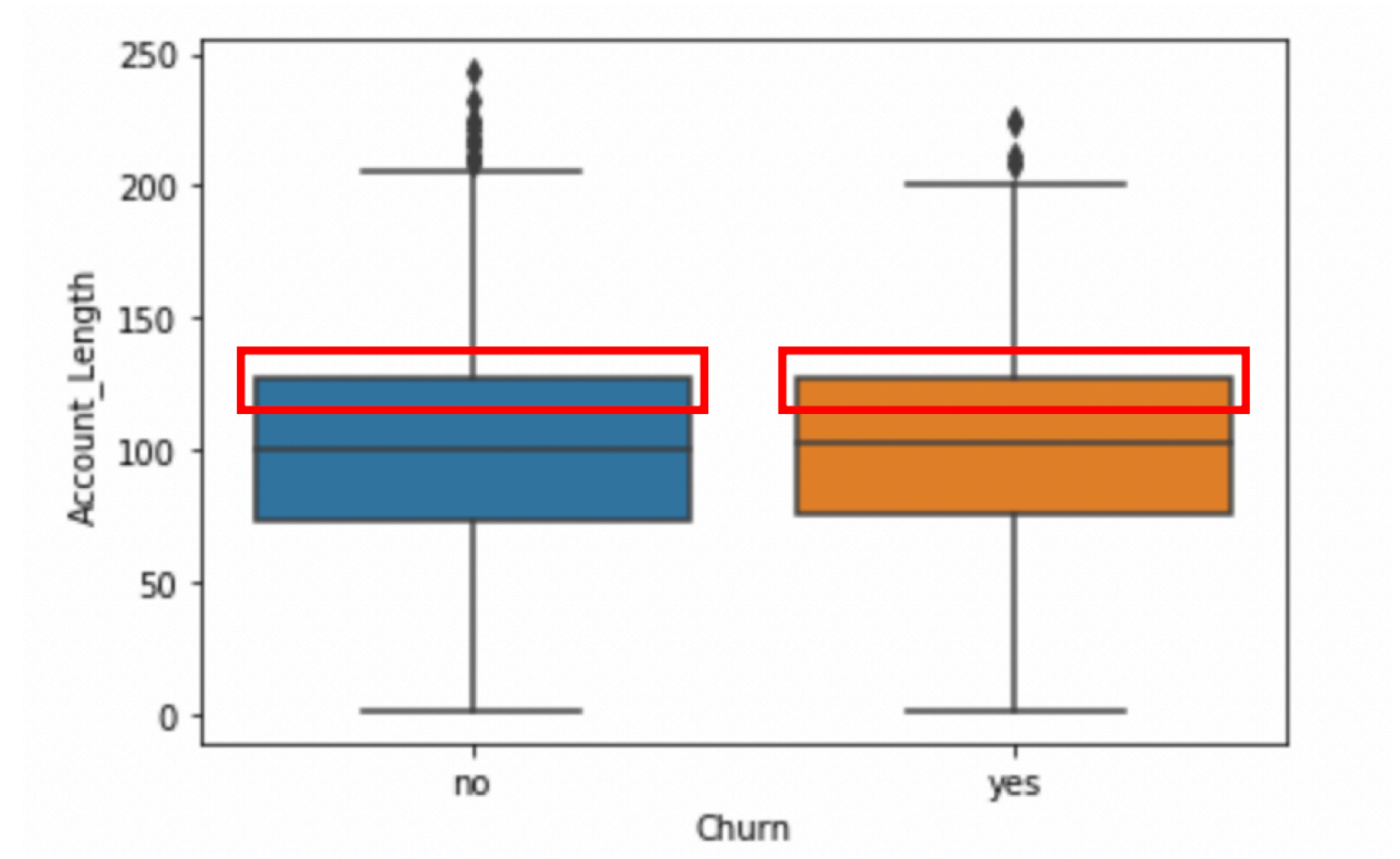


Rozdíly v délkách účtů

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

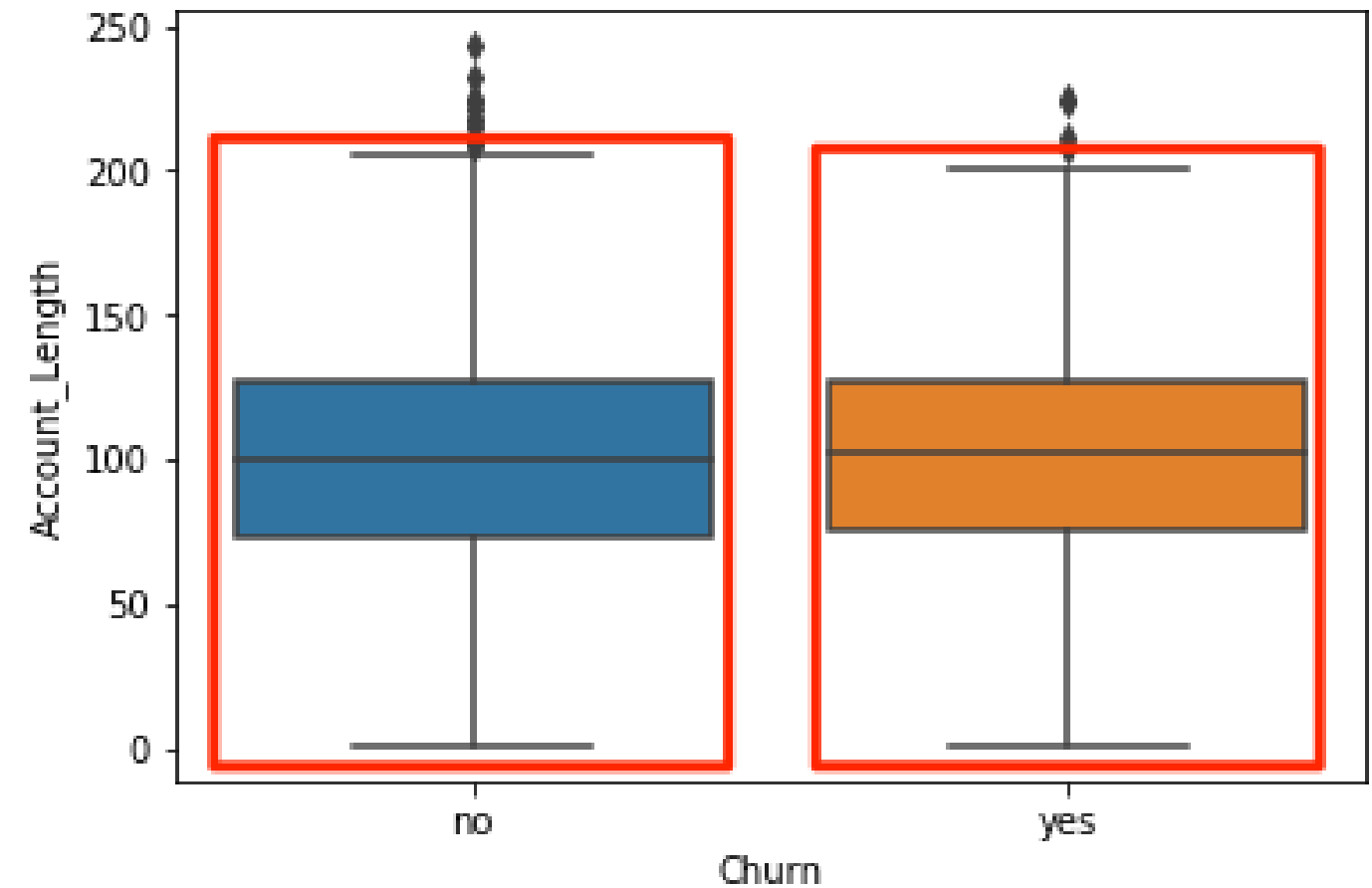


Rozdíly v délkách účtů

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

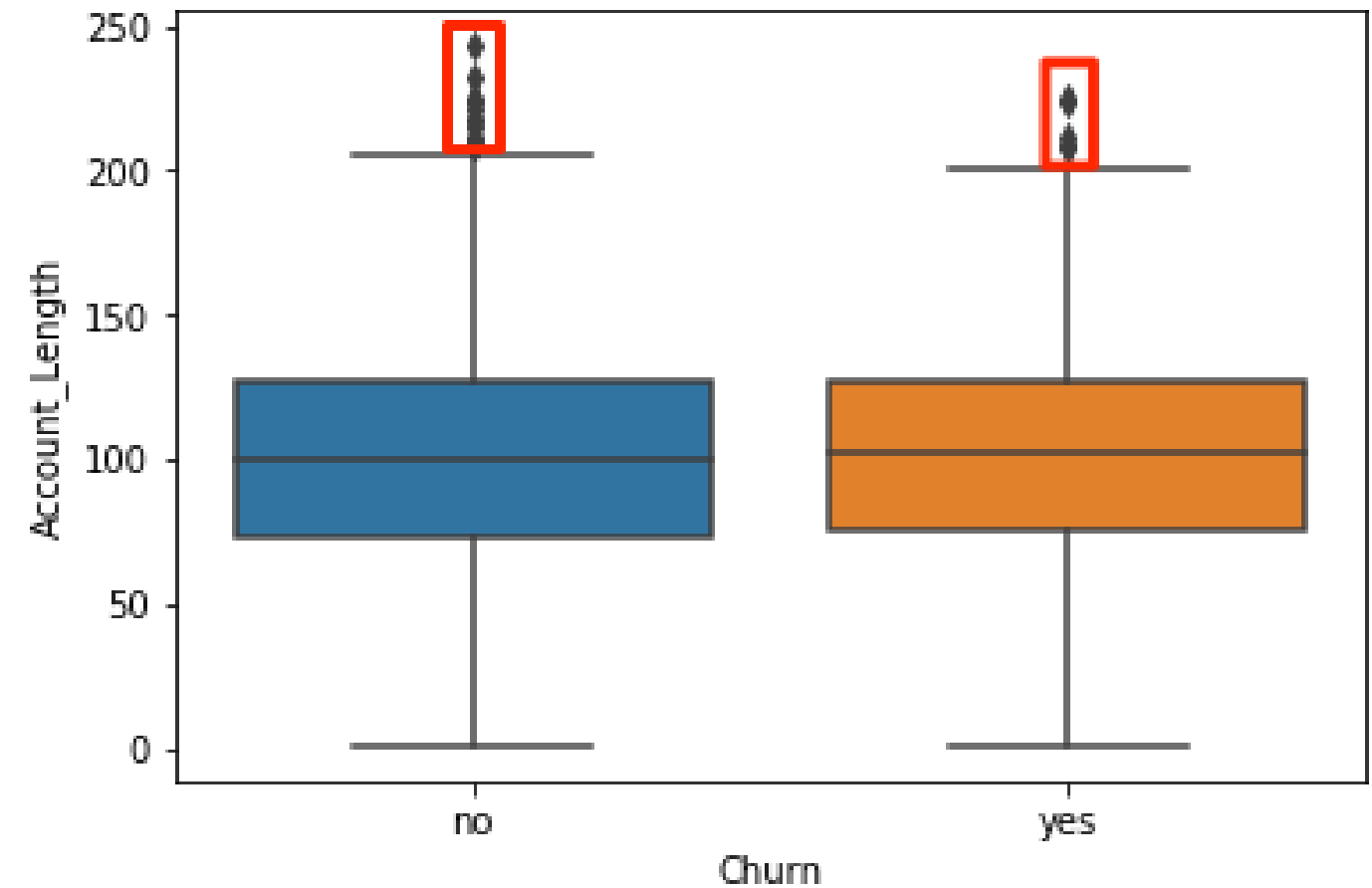


Rozdíly v délce účtu

- Box plot

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco)
```

```
plt.show()
```

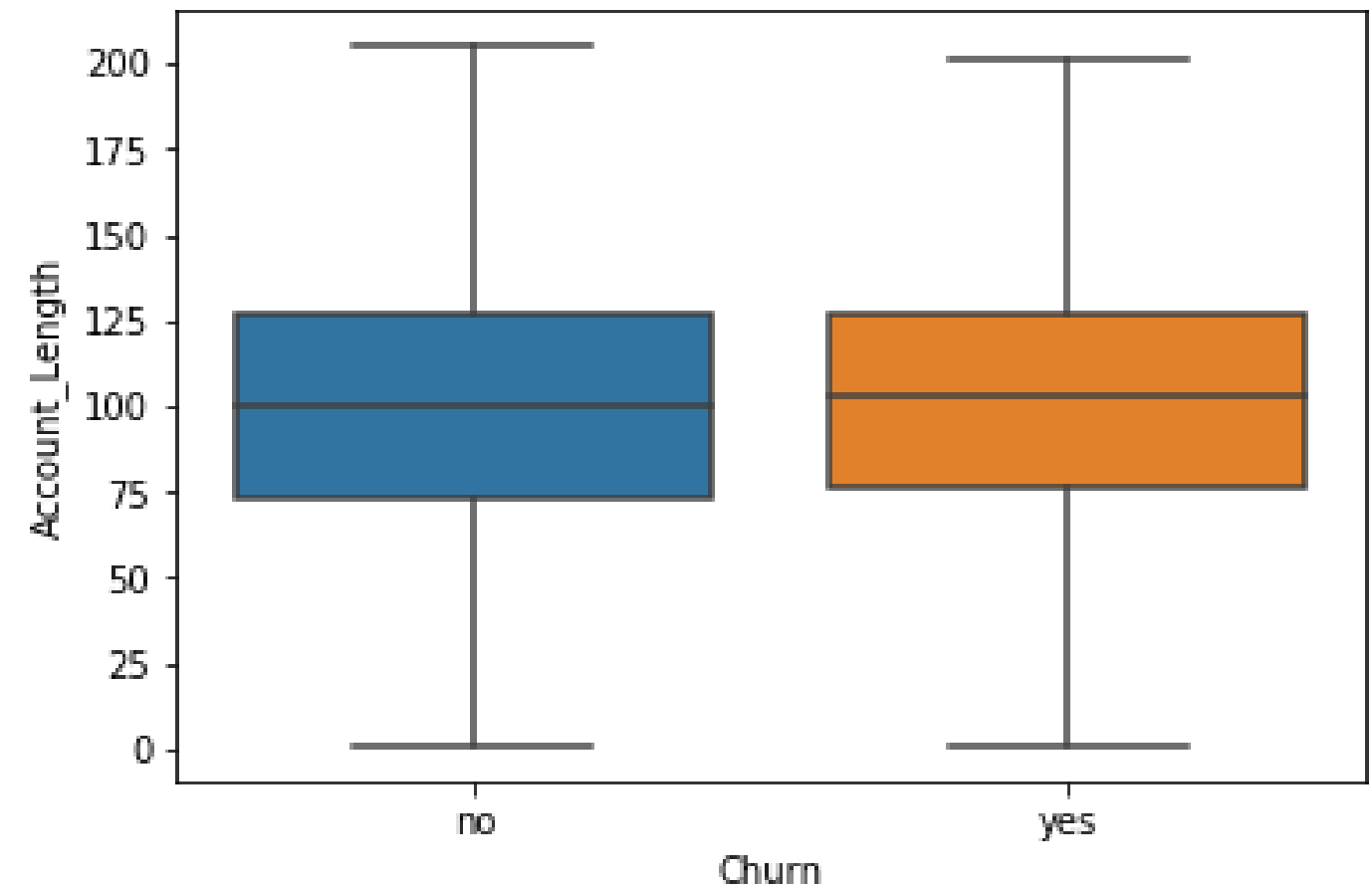


Rozdíly v délce účtu

- Box plot

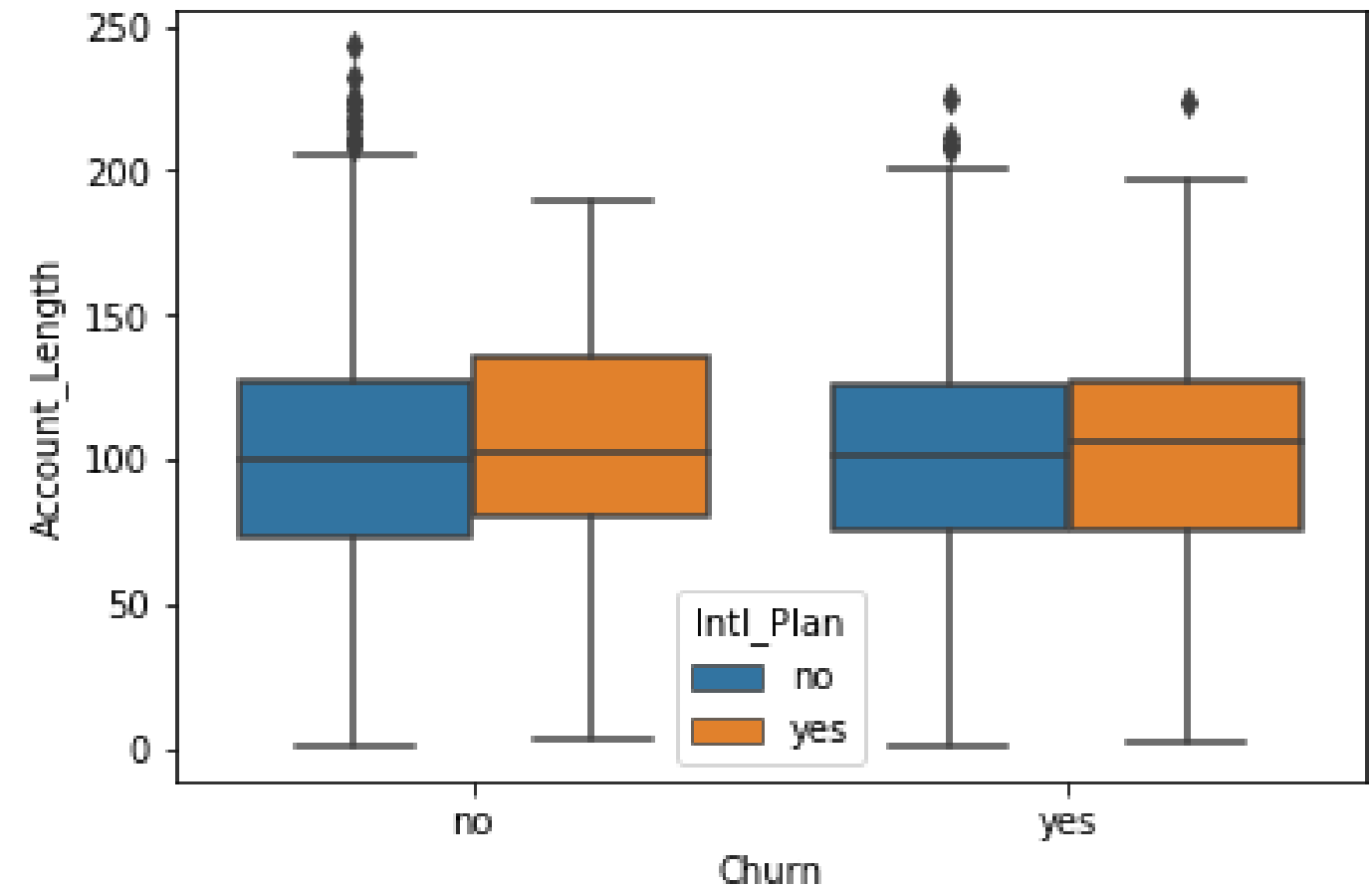
```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco,  
            sym="")
```

```
plt.show()
```



Přidání třetí proměnné

```
sns.boxplot(x = 'Churn',  
            y = 'Account_Length',  
            data = telco,  
            hue = 'Intl_Plan')  
  
plt.show()
```



Pojďme vytvořit grafy!

MARKETING ANALYTICS: PREDICTING CUSTOMER CHURN IN PYTHON

