

# Intellegend Ecommerce Churn Prediction

Dokumen  
Laporan Final  
Project

(dipresentasikan setiap sesi mentoring)



# Latar Belakang Masalah

Belanja online telah menjadi kebiasaan baru di masa kini, lebih dari 80% pengguna internet di dunia telah berbelanja produk online. Adanya perkembangan tersebut menyebabkan persaingan yang ketat diantara E-commerce. Persaingan tersebut diantaranya untuk menarik pembeli dan membuat mereka nyaman sehingga akan menjadi pelanggan.

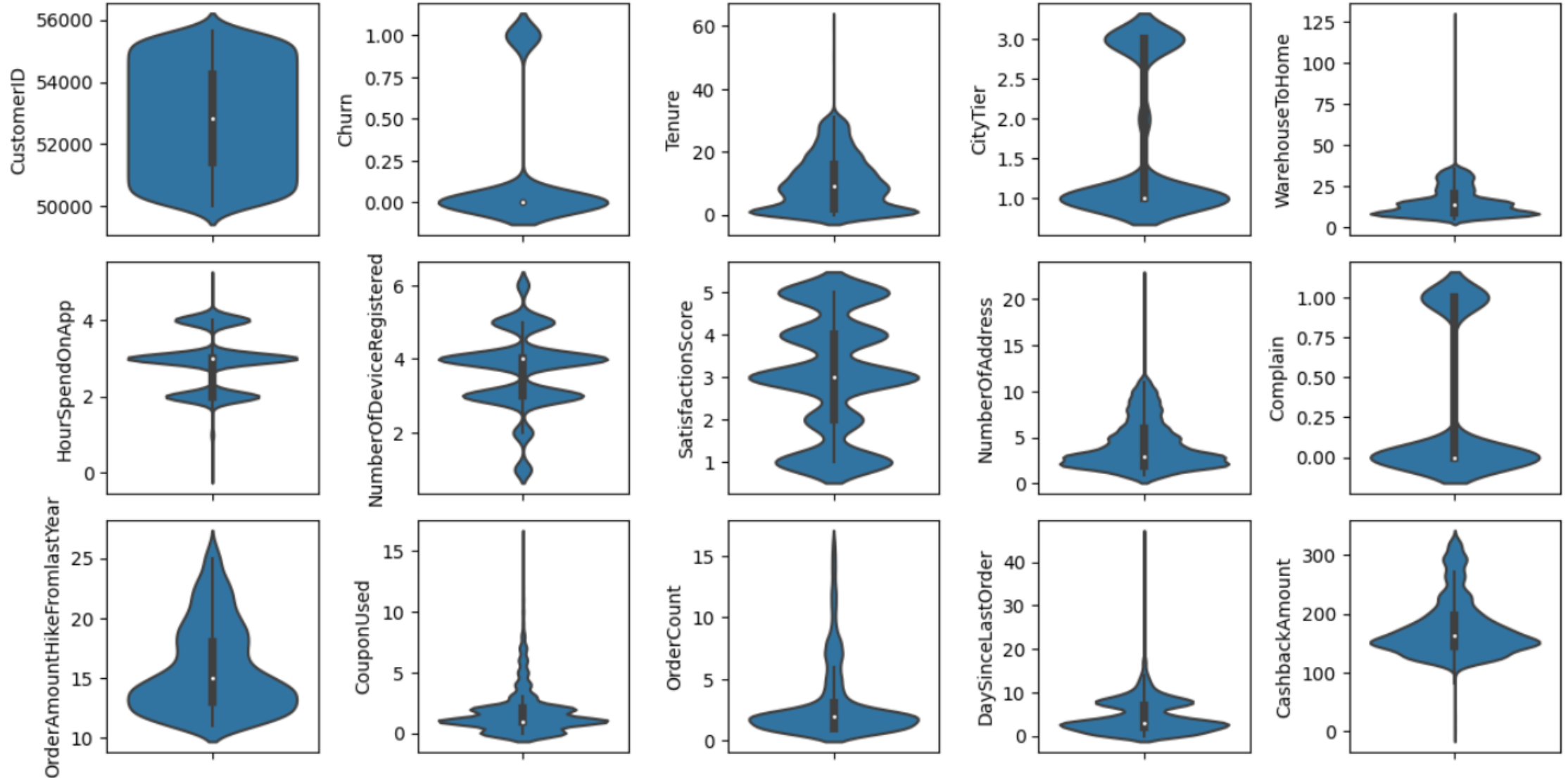
# Problem Statement

- Sebuah Ecommerce mempunyai customer yang churn sebanyak **16.83%**. yang akan mengakibatkan kerugian. Hal ini dikarenakan biaya untuk mempertahankan customer lebih kecil 5 kali lipat dibanding biaya untuk mencari customer baru.
- Cara yang dapat digunakan untuk mengurangi jumlah customer Ecommerce yang churn.
- Membuat sebuah machine learning model yang dapat memprediksi apakah customer akan churn serta faktor-faktor yang mempengaruhinya untuk churn.



# Rekomendasi Pre-processing

## DATA VISUALISASI



# Rekomendasi Pre-processing

## DATA VISUALISASI

Pada visualisasi sebelumnya kita dapat melihat distribusi dari beberapa data numerical yang membutuhkan preprocess disaat Stage 2 (Data Preprocessing):

- Uniform distribution: CustomerID --> karena ini adalah ID dan sangat banyak value unique, maka kita akan drop pada saat preprocessing
- Right-skewed: Tenure, WarehouseToHome, NumberOfAddress, OrderAmountHikeFromlastYear, CouponUsed, OrderCount, DaySinceLastOrder memiliki distribusi right skewed atau positively skewed, kita akan mentransformasi kolom ini dengan menggunakan log transformation
- Left-skewed: CashbackAmount

karena ada beberapa column yang memiliki left-skewed distribution maka kita akan melakukan square transformation agar distribusi kolom-kolom tersebut akan mendekati normal.

# Rekomendasi Pre-processing

## DATA VISUALISASI

Ada kolom yang memiliki 2 value yang sama namun dengan nama yang berbeda, yaitu:

- PreferredLoginDevice, Phone dan Mobile Phone memiliki arti yang sama, karena telepon yang tidak mobile (bukan smart phone) tidak bisa digunakan untuk login.
- PreferredPaymentMode juga terdapat beberapa category yang sama dengan nama berbeda, seperti CC dan Credit Card serta COD dan Cash on Delivery.

Nantinya category ini akan dijadikan satu pada saat preprocessing.

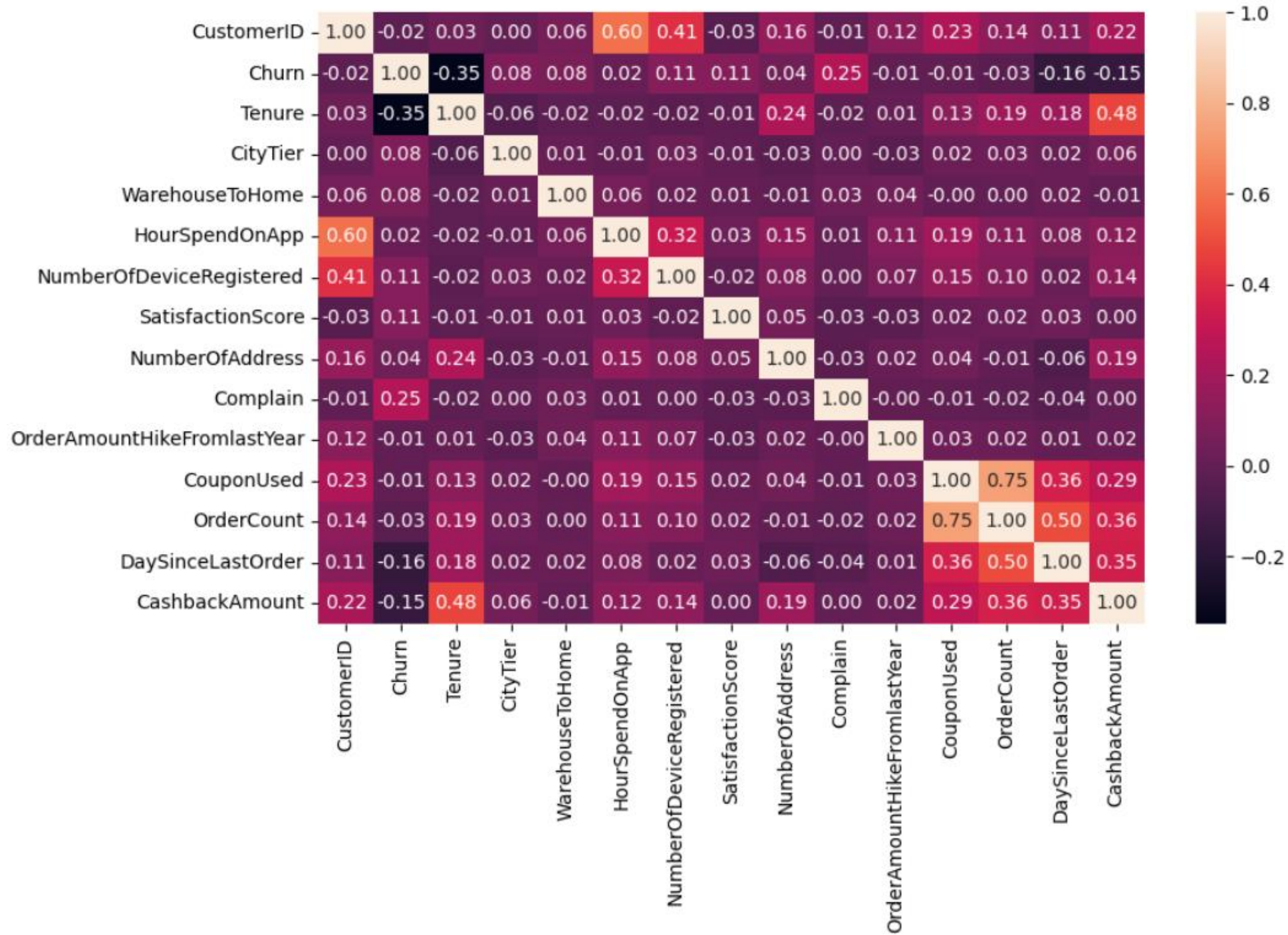
# Rekomendasi Pre-processing

Ada pula beberapa kolom yang memiliki missing value:

- Tenure
- WarehouseToHome
- HourSpendOnApp
- OrderAmountHikeFromlastYear
- CouponUsed
- OrderCount
- DaySinceLastOrder

Untuk kolom dengan tipe data numerical akan diisi dengan nilai Median, dan untuk column dengan tipe data Categorical, akan diisi dengan nilai Modus







# Rekomendasi Pre-processing

## DATA VISUALISASI

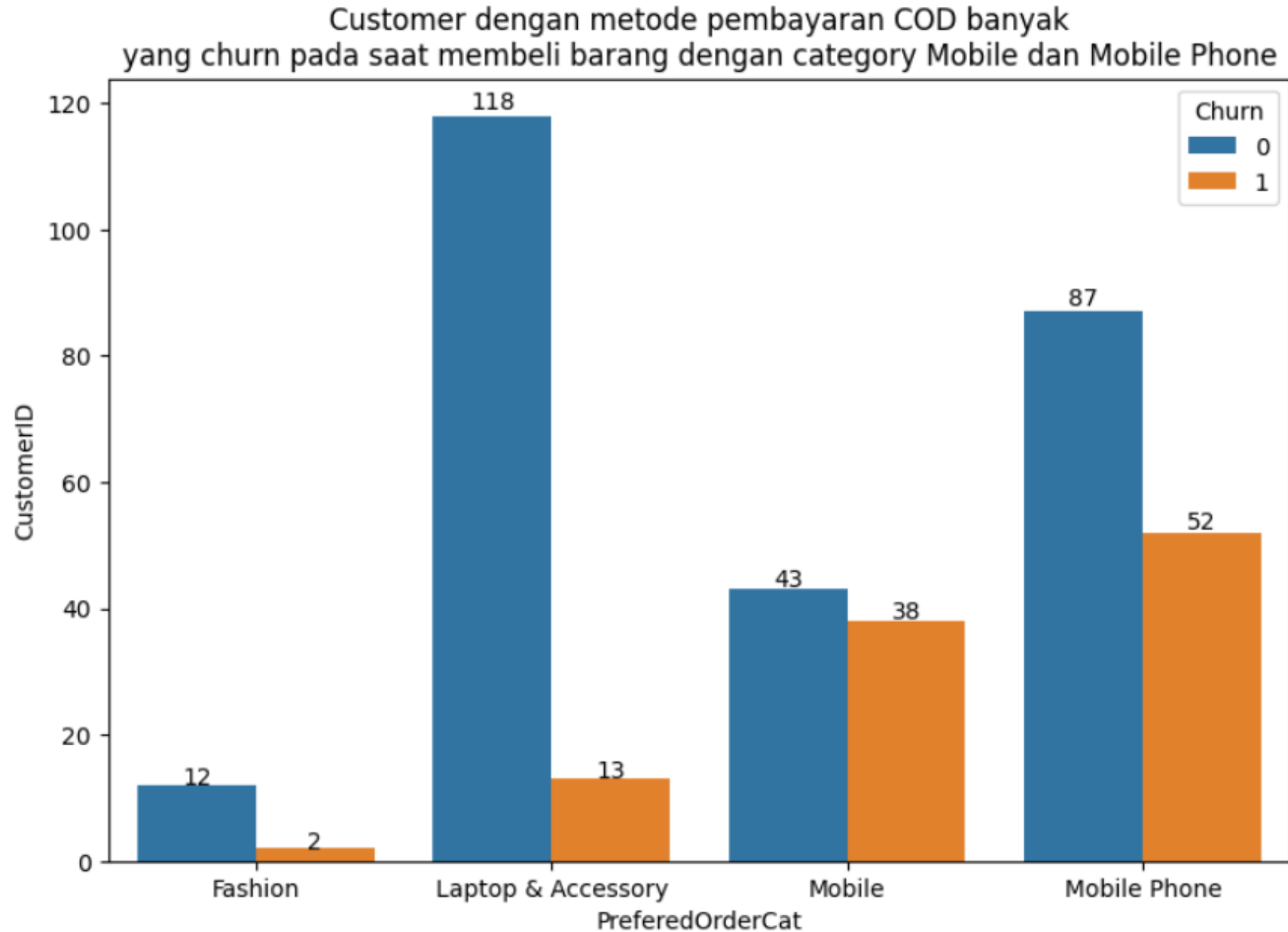
Dari Heatmap pada slide sebelumnya, kita dapat melihat Fitur yang memiliki korelasi tinggi yang relevan untuk dipertahankan:

- Tenure
- CityTier
- WarehouseToHome
- Complain
- DaySinceLastOrder
- CashbackAmount

Ada pula beberapa feature yang berkorelasi tinggi, sehingga dapat menimbulkan masalah multicollinearity. Feature-feature tersebut antara lain:

- CouponUsed dengan OrderCount --> tidak bisa dipakai bersama-sama karena korelasi tinggi. Keduanya tidak dipakai karena korelasi mereka rendah terhadap target.

# Business Insight (1)



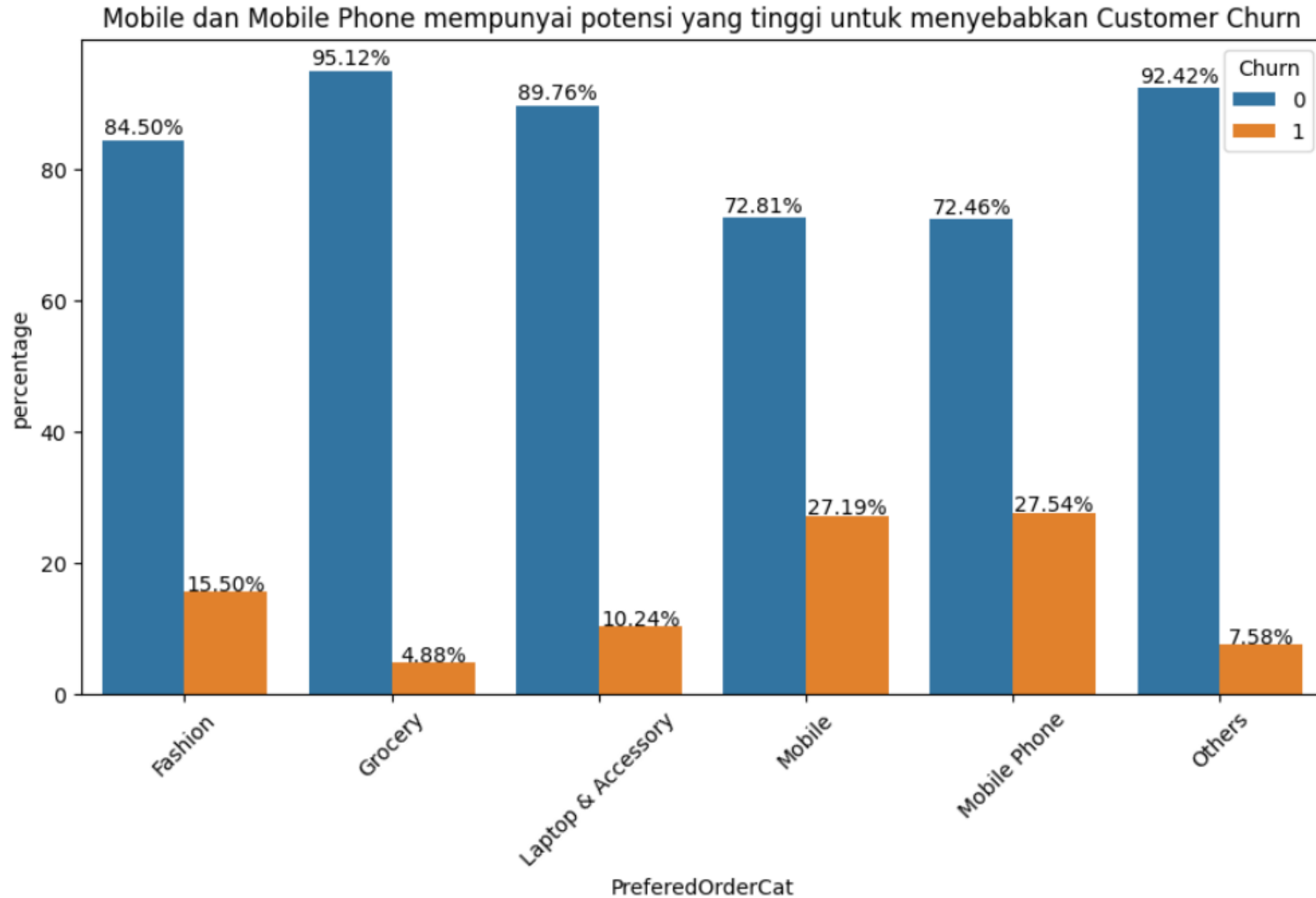
## Business Insight (1)

Customer dengan metode pembayaran COD banyak yang churn pada saat membeli barang dengan category Mobile dan Mobile Phone, Hal ini mungkin disebabkan oleh kurir COD yang kurang ramah dan berpengalaman sehingga menyebabkan barang menjadi cacat atau customer menjadi tidak puas terhadap pelayanannya.

Oleh karena itu untuk pembelian barang dengan category Mobile dan Mobile Phone disarankan untuk menghilangkan metode pembayaran COD.



# Business Insight (2)



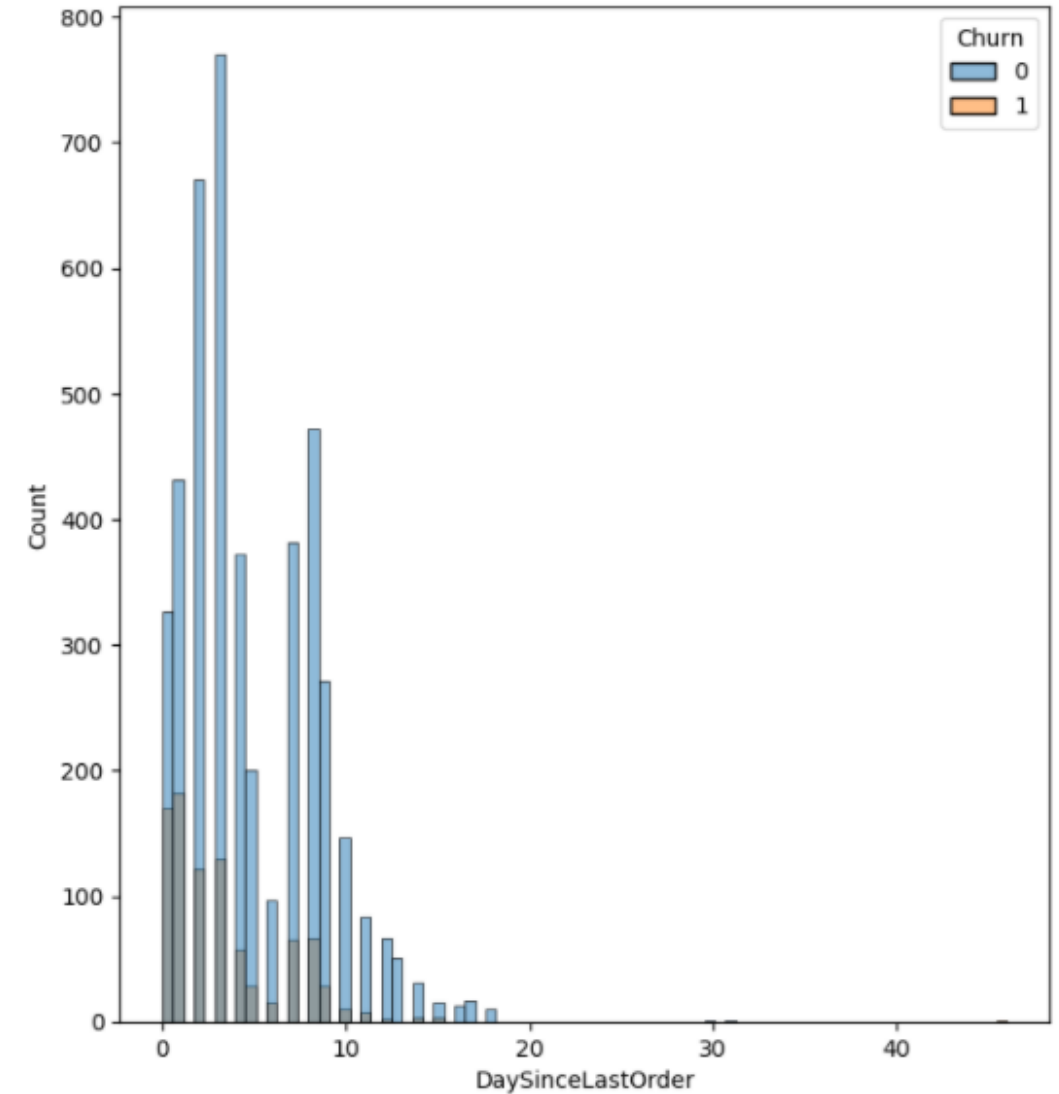
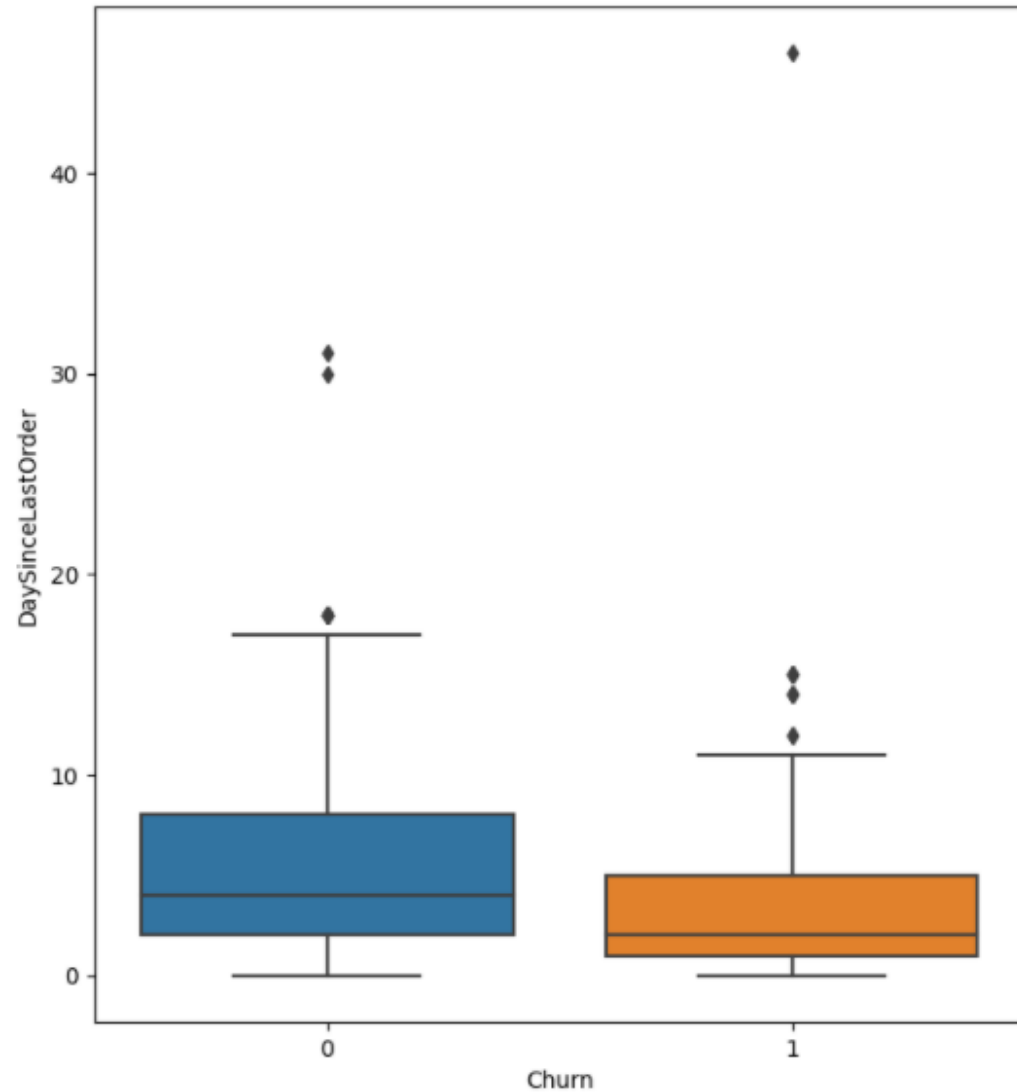
## Business Insight (2)

Mobile dan mobile phone memiliki potensi yang tinggi untuk menyebabkan customer churn. hal ini mungkin disebabkan karena kualitas barang dan juga mahalnya harga di Ecommerce tersebut, sehingga pelanggan pergi ke Ecommerce lain.

Hal yang dapat dilakukan untuk mengatasi ini adalah memberikan pilihan untuk menggunakan asuransi/garansi serta memberikan kupon diskon untuk category Mobile dan Mobile Phone.

# Business Insight (3)

DaySinceLastOrder yang rendah cenderung menyebabkan Customer Churn





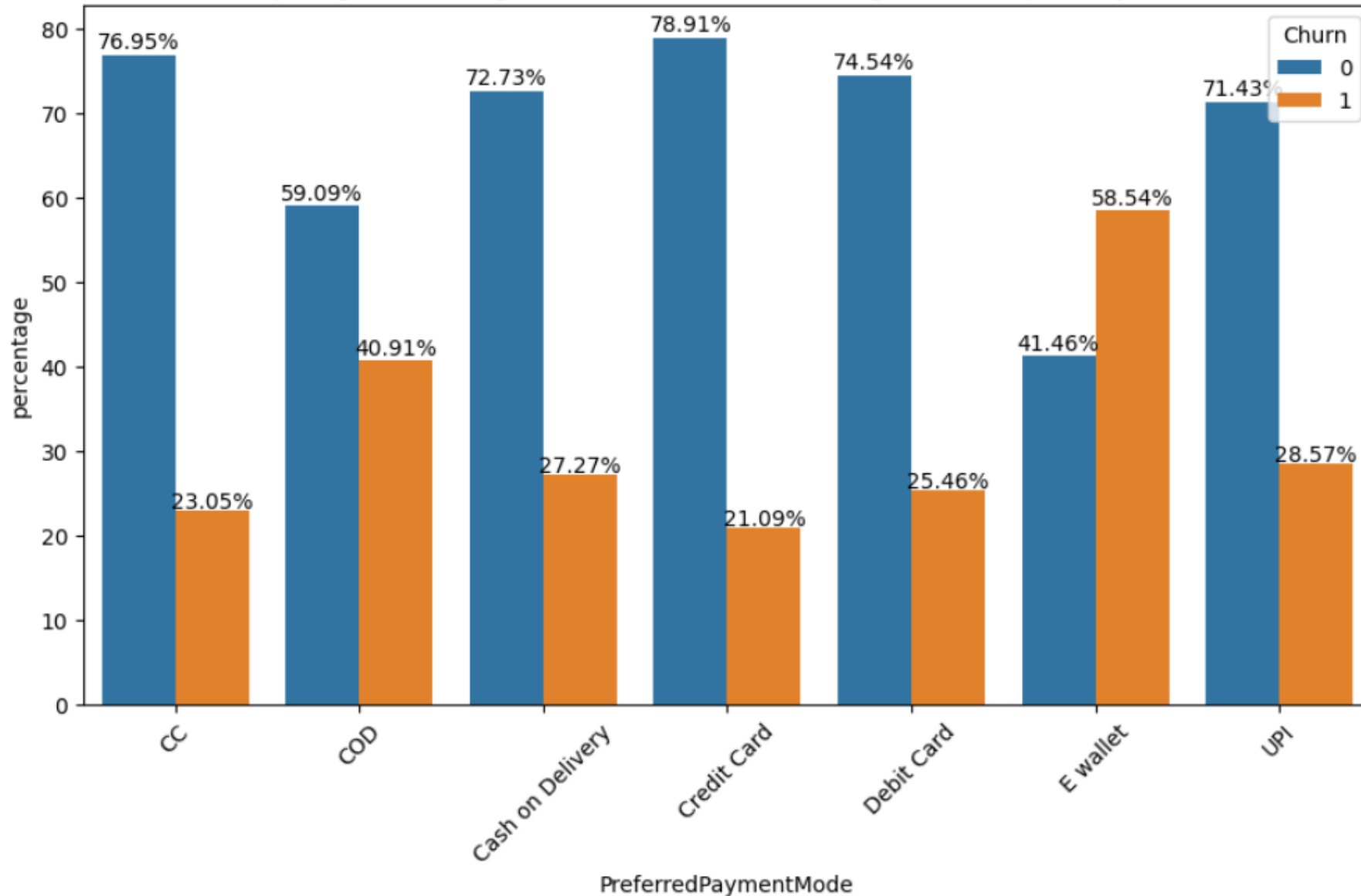
## Business Insight (3)

DaySinceLastOrder yang rendah cenderung menyebabkan customer churn. Hal ini mungkin dapat terjadi karena customer merasa terlalu sering berbelanja online, sehingga ia memutuskan untuk merubah kebiasaan itu dan akhirnya berhenti berbelanja sama sekali.

Hal ini dapat kita atasi dengan memberikan voucher setiap beberapa waktu, dalam kasus ini setiap minggu. Dengan melakukan ini, customer akan cenderung menunggu setiap minggu untuk berbelanja, sehingga customer tidak terlalu sering berbelanja.

# Business Insight (4)

Untuk pembelian dengan category Mobile dan Mobile Phone paling cocok dengan CC dan Credit Card sebagai metode pembayaran



## Business Insight (4)

Untuk Pembelian dengan kategori Mobile dan MobilePhone kebanyakan customer yang membayar dengan CC dan Credit Card cenderung lebih jarang churn dibanding dengan metode pembayaran lainnya.

Oleh karena itu, kita dapat memberikan promo untuk customer yang mau membeli produk dengan kategori Mobile dan MobilePhone dan membayar dengan CC dan Credit Card sehingga dapat mengurangi jumlah customer yang churn.