

PREDICT DELIVERY TIME

June 27, 2023

Huy Hoang Vuong
BrainStation-Data Science



Problem Space

17%



Not return after first
late delivery

50%



Abandon company for
second time

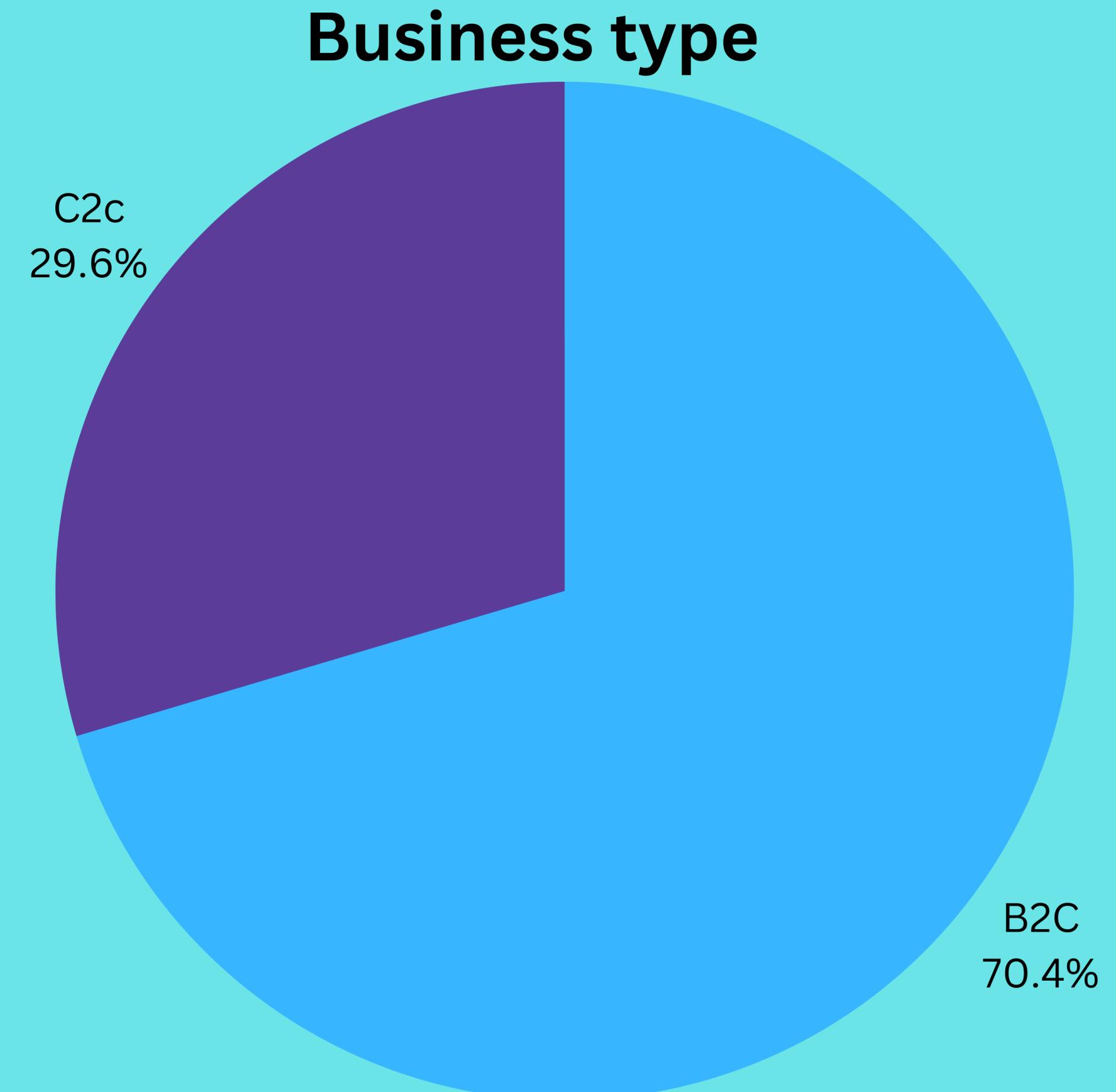
How might we use machine learning to let customers know the day
they would receive their package when they shop online?

Data

The dataset has 15 million rows

Record the transaction from
2017 to 2019

Target: calculate delivery time
of the order

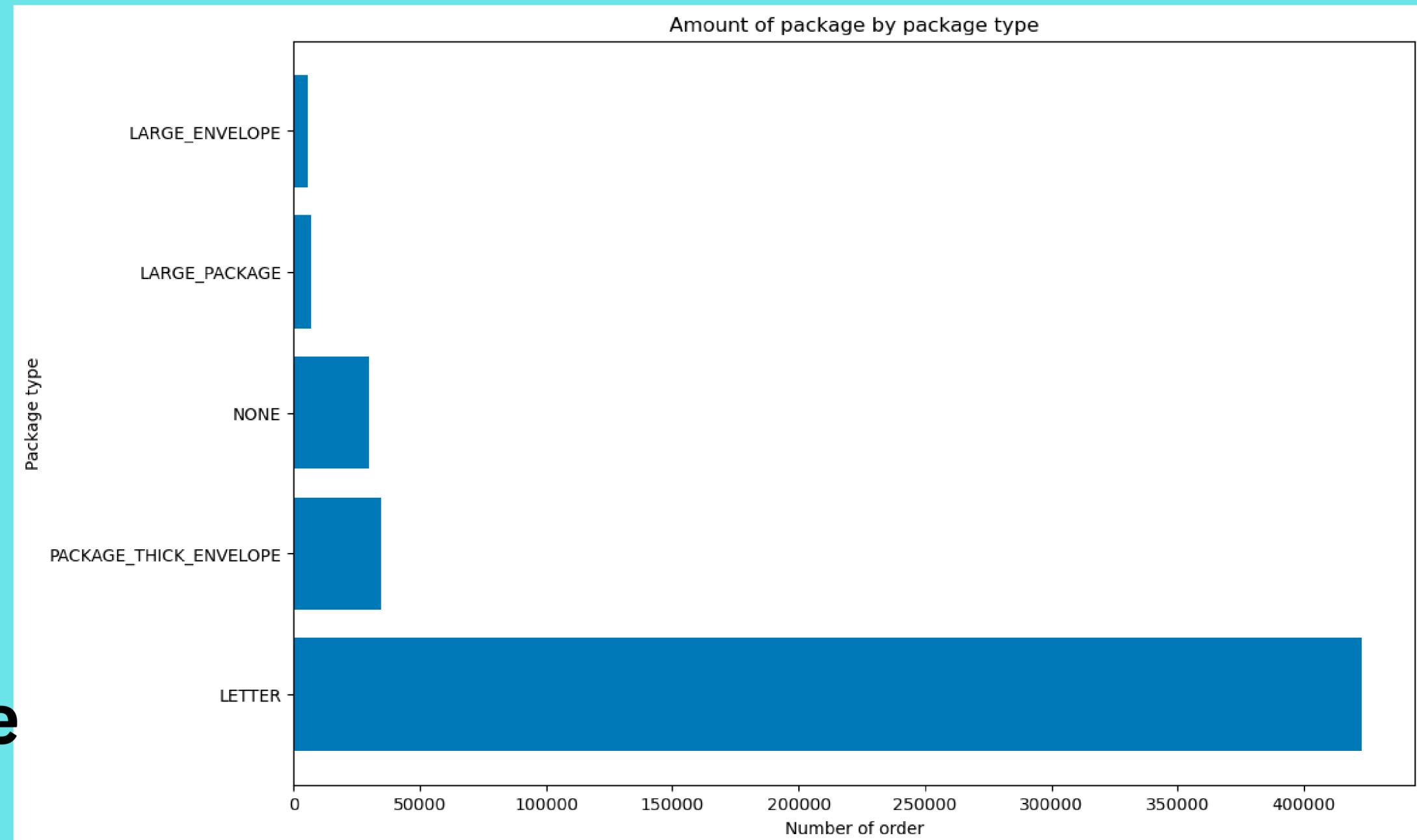


EDA

Most of the transaction contribute by Business to Consumer

Most of the order was sent in the small pack: letter size

Business type



Business type

Feature Selection

Feature convert to number

Add new features from existed
features which help model
understand

Feature importance:

- b2c_c2c
- distance
- weight
- package_size



Model

	<i>Percent of ontime or Early</i>	<i>Loss</i>
<i>Linear Regression</i>	70.6137	0.8007
<i>Ridge Regression</i>	70.6187	0.8007
<i>XGBoost</i>	72.0835	0.7612
<i>Neural Network</i>	72.0035	0.7589
<i>Recurrent NN</i>	70.6137	0.8007

Next Steps

- Consider having more features
- Future implementation: giving an early warning for business



THANK YOU!

