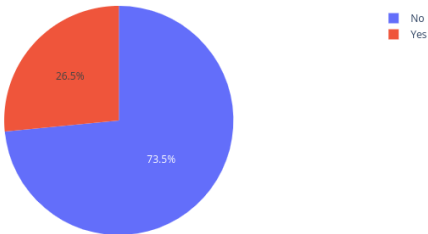# Customer churn

Michael O'Keefe

15 July 2021

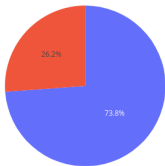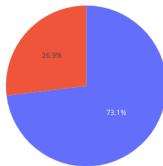# What fraction of customers churn?

How many customers churn?

- 73.5% of customers are loyal
- 26.5% have churned
- Can we predict when this is going to happen?

Churn demographics

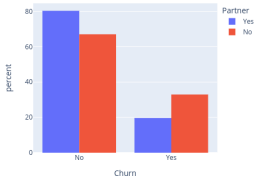Percentage of males that churn
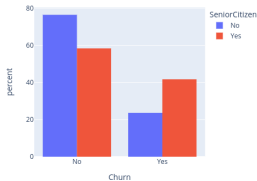

Percentage of females that churn

- Firstly, we will look at how demographic information plays a part in predicting customer churn
- Looks like gender does not provide any useful insight into whether or not a customer will churn
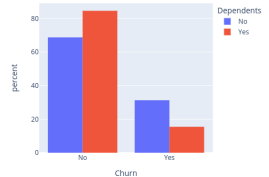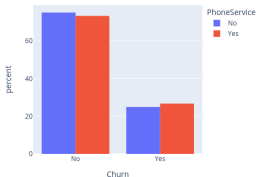
Churn demographics

- Customers who:
  - Do not have partners are 13% more likely to churn
  - Are senior citizens are 18% more likely to churn
  - Do not have dependents are 16% more likely to churn

Services used
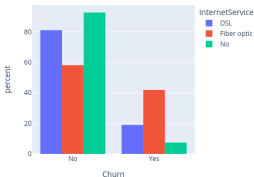
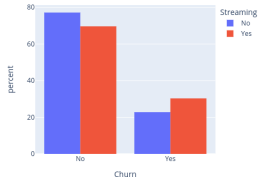- Whether or not a customer uses a phone service does not seem to effect whether or not they will churn
- Fibre optic internet users are much more likely to churn (42%) than DSL (19%) and non internet users (7%)
- Streaming service users are 8% more likely to churn

How do financial aspects effect churn?

Customer tenure churn
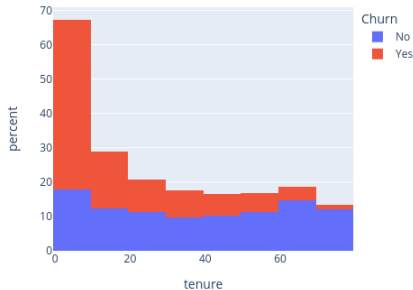
Monthly charges churn
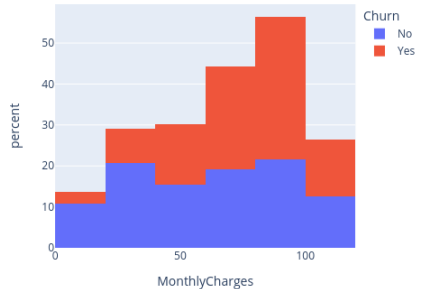
- About 50% of customers churn within the first 10 units of tenure (months? presumably)
- Monthly payments are also a significant factor in customer churn, higher prices higher churn!

How do financial aspects effect churn?

Contract type churn · Paperless billing churn · Payment method churn
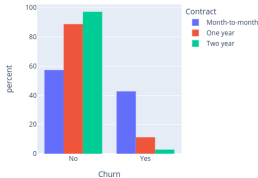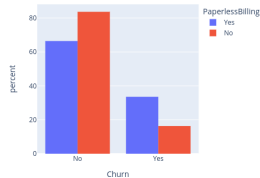
- Customers on monthly contracts are much more likely to churn (43%), compare with one (11%) and two (3%) year contracts
- Customers who use paperless billing are 17% more likely to churn
- Customer who pay by electronic cheque are much more likely to churn (45%) compared with other payment methods ($\approx 20\%$)

What can we do with this information?

- We have seen that demographic and financial information can provide insight on whether or not a customer is going to churn
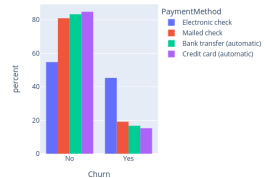- We can use machine learning techniques to build a model which will determine the probability of a customer churning based on their information
- This is known as a classification problem, and there are many models that can be used
- In this case I am going to use logistic regression
- We will use all of the information we just looked at to try and predict whether a customer is going to churn or not
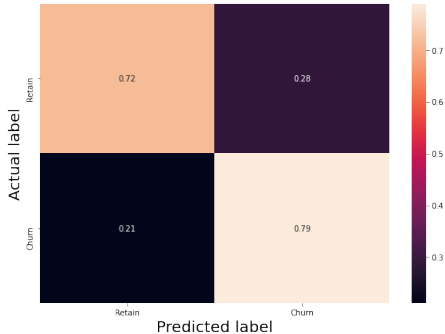
- Correctly predict:
  - Customer will not churn 72% of the time
  - Customers will churn 79% of the time

Evaluating our model

- We will use precision (P) and recall (R) to evaluate the performance of our model

$$P = \frac{T_p}{T_p + F_p}$$

$$R = \frac{T_p}{T_p + F_n}$$

- These metrics can then be combined into an F_1 score

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} = 0.76$$

- The best possible F_1 score is 1, and the worst is 0
- So our model is doing pretty well

Conclusion

- We found that:
    - Single customers are more likely to churn
    - Older customers are more likely to churn
    - Customers without children are more likely to churn
    - Fibre optic and streaming users are more likely to churn
    - New customers on month-to-month contracts paying a high premium are more likely to churn
    - Customers paying via electronic cheque are more likely to churn
- What can you do about this?
    - Start new customers out on lower premiums over a short fixed term contract (3 months)?
    - Move electronic cheque customers to a new payment model?
    - Send all customers paper bills
- Built a classification model which correctly predicts customer churn with reasonable accuracy
    - Use this to target prospective churners before they churn!