1. Imagine the business context is the telecommunication customer churn problem. Churn is defined as a customer’s switch over to a competing provider. A predictive model tries to predict the value of churn. We refer churn to be the actual realized value and prediction to be the predicted value of churn. In a predictive model of churn:

* The primary outcome churn = 1 when the customer is a churner;
* TP (true positive) is when prediction = 1 and churn = 1;
* FP (false positive) is when prediction = 1 but churn = 0;
* TN (true negative) is when prediction = 0 and churn = 0;
* FN (false negative) is when prediction = 0 but churn = 1.

If interceding with (i.e., trying to change a customer’s mind by using some intervention) a potential churner is relatively cheap but losing a customer is expensive:

**Question: Which error is costlier** (i.e., more expensive)**: a FP or a FN? Explain why.**

**As defined before, False-Positive means our model predicts that the customer will churn but in reality, he/she won’t. so, they have to do some intervention to change the customer’s mind which may have some cost for them. They might offer the customer some discounts but still, it’s cheaper than losing the customer.**

**False-negative means our model predicts that the customer won’t churn but in reality, she/he churns so we end up losing the customer and it costs a lot for the company.**

**Overall, the tradeoff between FPs and FNs will vary by dataset and variant-calling pipeline.**

**One other thing that we have to consider in our model is the Proportion of False-Positives and Proportion of False-negative which are defined as below:**

**Proportion of False Positives =**

**Proportion of False Negative =**

**lower is better for PFP, and PFN.**

**By comparing PFP and PFN, we can decide to go with FP or FN. in the other word, if the FP is way larger than FN, on the special case we might decide to deal with FN rather than FP.**

**But overall in this case scenario, dealing with FP is much cheaper than FN.**