### **F1 score**

Let's introduce something a bit more general. If we have precision and recall, then the F1 score is somewhere around the middle. That gives us much importance to both. Of course, if one of them is smaller, it raises a flag, but it treats them both the same. Now let's say our model cares a bit more about precision than recall, then we want something more skewed towards precision.

So, we'll say it's F0.5 score. So we call that beta. Beta is 0.5.

The smaller the beta, the more towards precision that we get.

Or if we want our model to care more about the recall, then we pick a larger beta.

The larger the beta, the more towards recall that we get.

### **Fraud detection example**

In the fraud detection example, which beta should we use? Since it needs to be a high recall model since we need to catch all the fraud cases, and it's okay if we accidentally detect and investigate some that are not.

So something like F10. But then maybe, we're sacrificing too much precision, and we're accidentally sending our customers too many notifications about their transactions without them being fraudulent, and they're starting to get annoyed. So, we can move a bit toward say F2. But then maybe we discover that we need to focus a bit more on recall because we really don't want to miss too many fraudulent transactions, so we go here to F5.

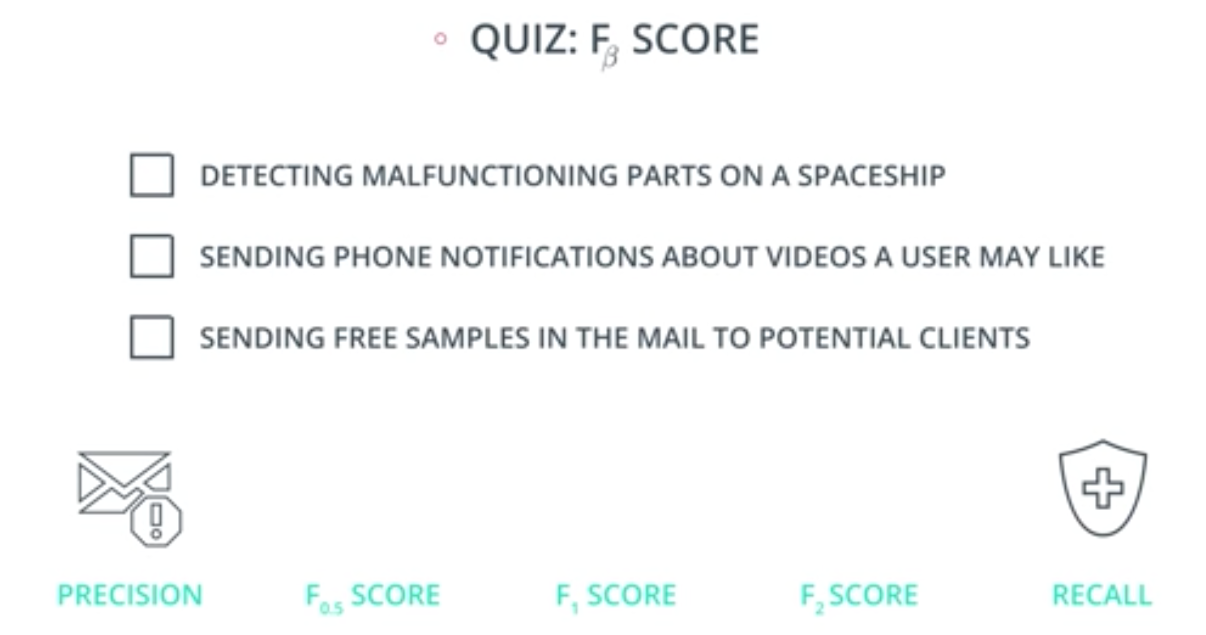
It's not an exact science. Finding a good value of beta requires a lot of intuition of your data and a lot of experimentation.

### Quiz

Now, let's test our knowledge. So, let's look at three possible models:

* In the first one, we are NASA, and we have a model for detecting malfunctioning parts on a spaceship.
* In the second, we have a video recommender system, and we have a model for sending users phone notifications about new videos they may like.
* And in the third one, we are a store, and we have a model for sending free samples in the mail to potential clients.

And let's say one of them has an F beta score of F1, the other one is F0.5, and the other one is F2. Which one is which?



### Quiz Question

Out of the following three models, which one should have an F-beta score of 2, 1, and 0.5?

Match each model with its corresponding score.

* Detecting malfunctioning parts in a spaceship
* Sending phone notifications about videos a user may like
* Sending promotional material in the mail to potential clients

#### **Model F-beta Score**

Spaceship \_\_\_\_\_\_\_\_\_\_\_

Notifications \_\_\_\_\_\_\_\_\_\_\_

Promotional Material \_\_\_\_\_\_\_\_\_\_\_

|  |  |  |
| --- | --- | --- |
| Model | F-beta Score | Explanation |
| Detecting malfunctioning parts in a spaceship | F₂ | Recall is critical - missing faulty parts could be catastrophic (prioritize recall) |
| Sending phone notifications about videos | F₁ | Balanced approach - both precision and recall matter for user experience |
| Sending promotional material | F₀.₅ | Precision is key - avoid wasting resources on uninterested customers |