# Explanatory Notes for 6.390

Shaunticlair Ruiz (Current TA)

Fall 2022

## Applications: Downstream Tasks

Finally, there's one more way to think about clustering that is more **practical**, and closer to **objective**.

We use clustering to **sort** different data points that need **different** processing: this can make our model more **effective**, since different parts of the dataset may work **better** with different **treatment**.

**Example:** We could train a different regression model on each cluster: this can create a more accurate model.

We call this next problem a **downstream application**.

**Definition 1**

A **downstream application** is a **problem** that relies on a **different** process to make its work better or easier.

In this case, **clustering** has **downstream applications** that can **take advantage** of the **structure** it reveals.

If our clustering is **good**, we would expect it to **improve** the performance of downstream tasks.

**Concept 2**

We can indirectly **evaluate** a **clustering algorithm** based on how **successful** the **downstream application** is.

If it **improves** the performance of a downstream application, we could say it works **well**.