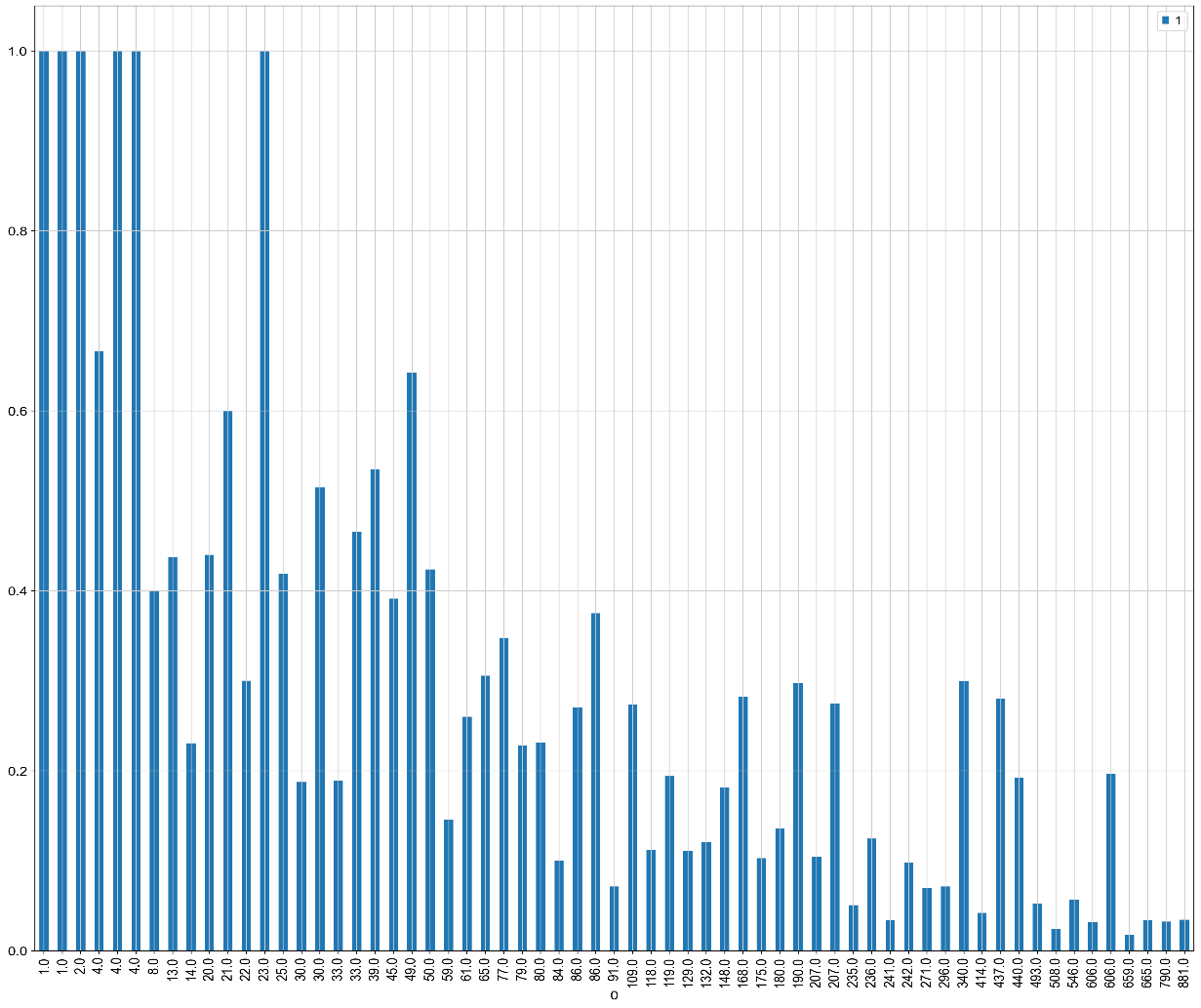
**Total approach:**

There are relationship and correlation between cell proportions and their knockout rows numbers in datasets like clon\_information or sc\_training dataset. We should model relationships and correlation and predict cell proportions for each knockout.

**Details:**

1. Inner join datasets clon\_information and sc\_training using common column barcode.
2. For each rows now we has cell condition, cell state.
3. Create five KNN regression algorithm for each of five states. Input knockout rows number and output is cell state proportion.

* For example below show figure between progenitor proportion and knockout rows numbers in clon\_information(x-axis show knockouts rows numbers and y-axis show their progenitor proportion):



1. For knockout input, we find its row numbers in clon\_information as input, then predict cell proportions for five states using five KNN regression models.