

▼ 1 INIT

▼ 2 build\_data\_sets

▼ 2.1 read protein data

LFQ: rel. quantification – for same protein across different samples

IBAQ: orthogonal comparison – for different proteins in the same sample

Intensity: supposed to be both

```
C:\ProgramData\Anaconda3\lib\site-packages\ipykernel_launcher.py:2: DtypeWarning: Columns (5,6,2764,2765,2773,2774,2775,2776) have mixed types. Specify dtype option on import or set low_memory=False.
```

▼ 2.2 read mrna data

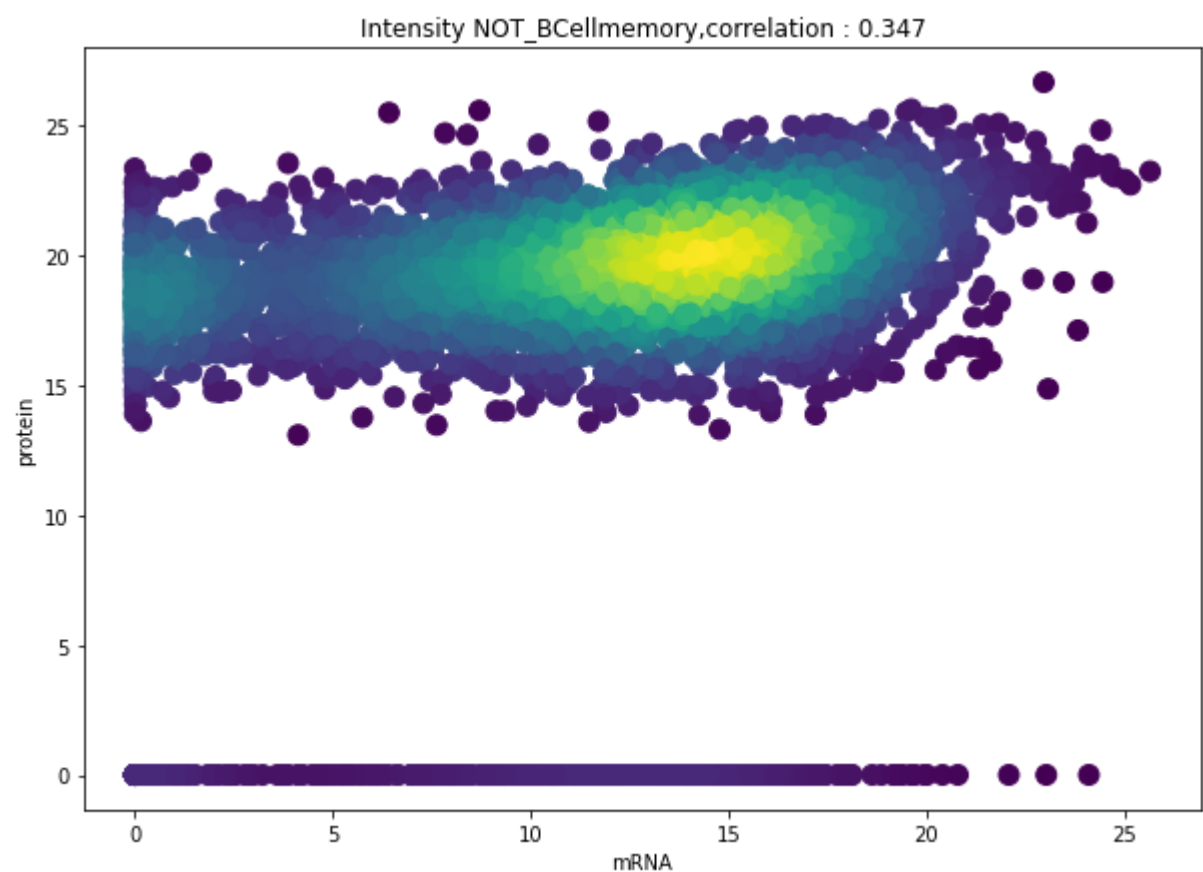
we start with mapping the cells annotations from the MassSpec to mRNA :

▼ 3 compare sets

▼ 3.1 compare all proteins to all genes

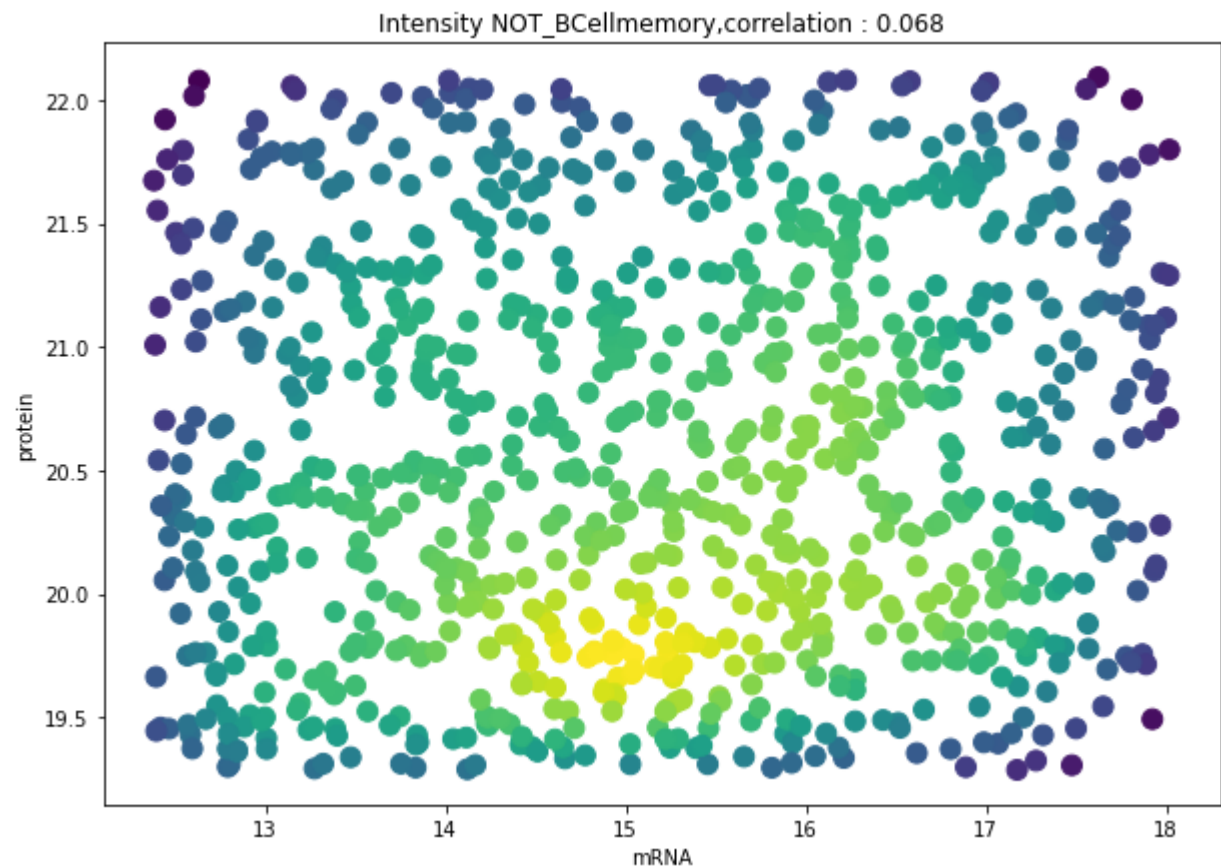
▼ 3.1.1 med value of genes

Here we took the median value of each protein (between Jan versions) and compare them to the adjusted mRNA values :

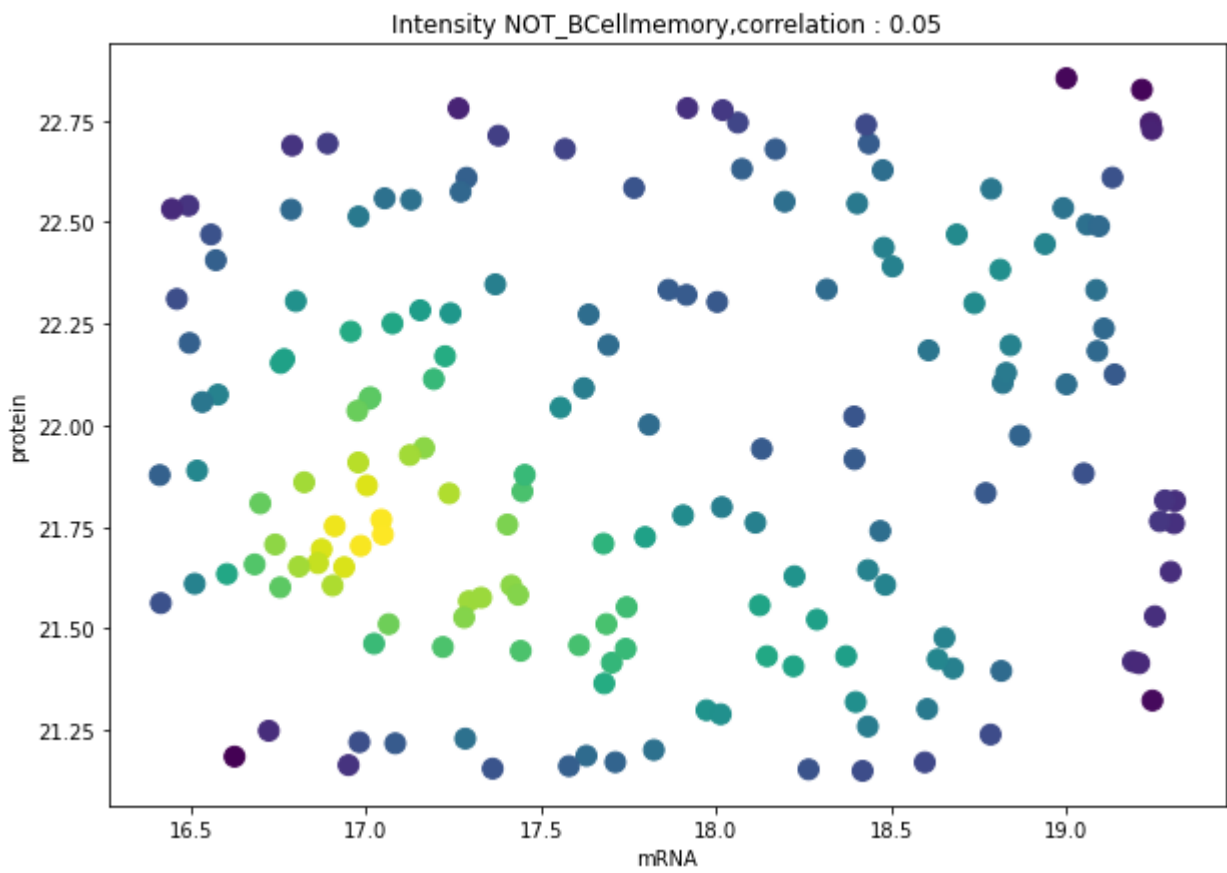


▼ 3.1.2 med values + taking only percantile 0.5 to 0.9

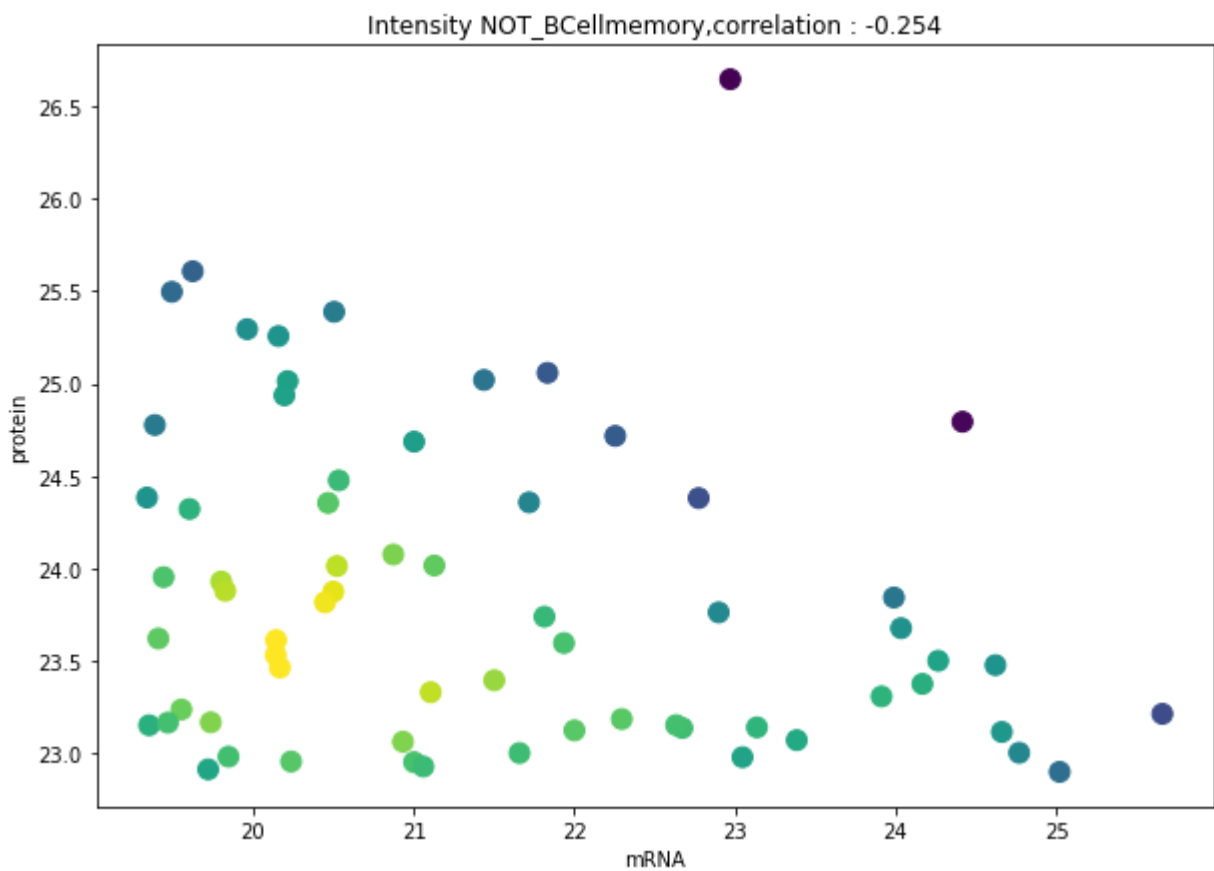
Here we kept only high values (mostly remove zeros from proteins)



▼ 3.1.3 med values + taking only percantile 0.95 to 0.8



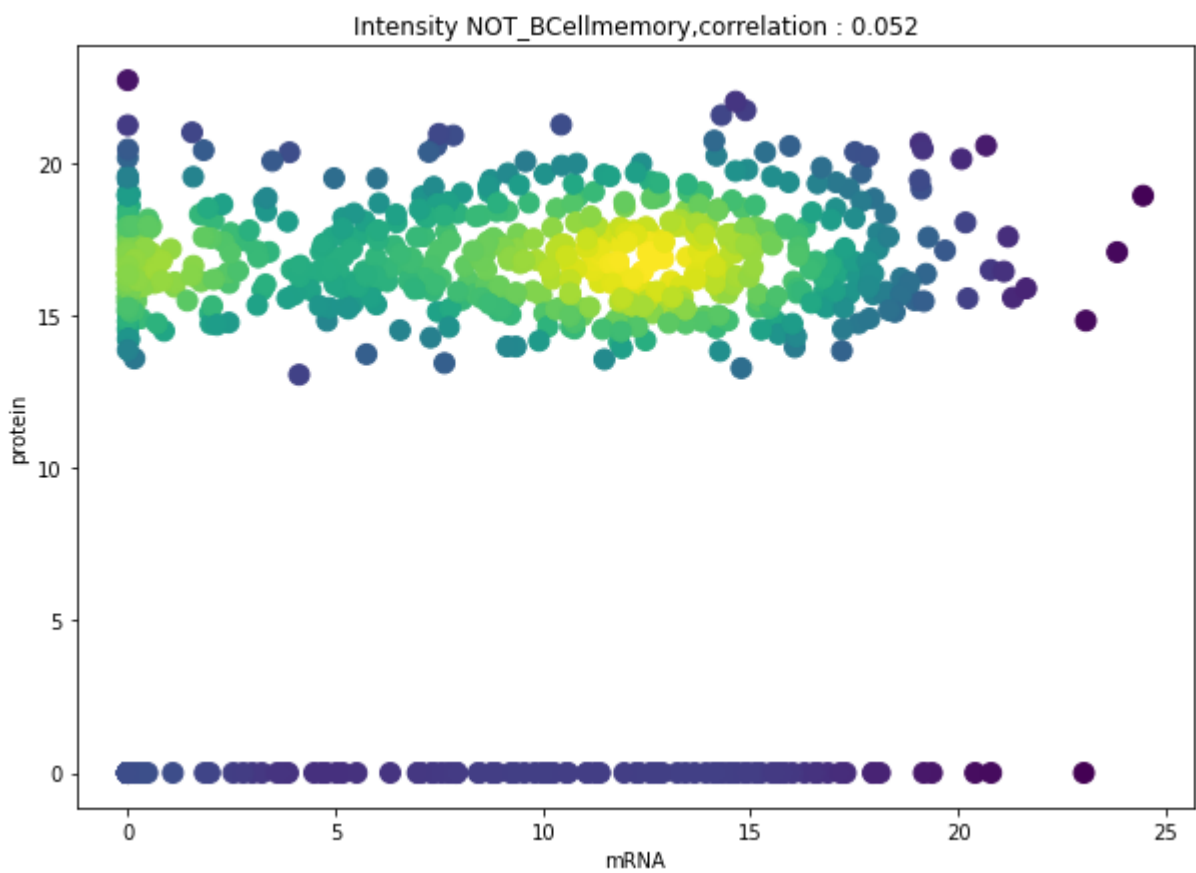
▼ 3.1.4 med values + taking only percantile 0.95 to 1



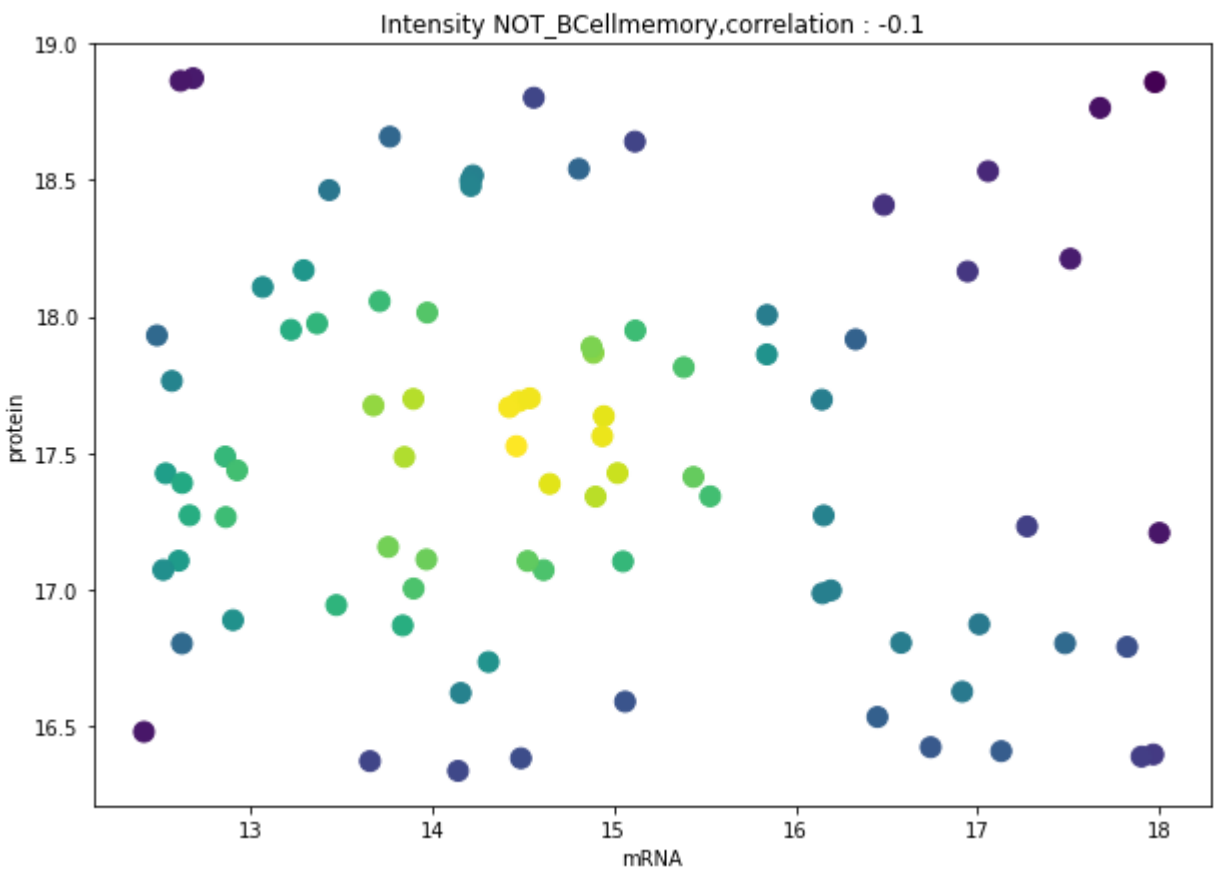
▼ 3.2 remove high STD proteins

Here we removed proteins we are "not sure about". aka - proteins where the std between versions is bigger then 0.8

▼ 3.2.1 med value of genes



▼ 3.2.2 med values + taking only percentile 0.5 to 0.9



▼ 3.2.3 med values + taking only percentile 0.99 to 0.7

