

The suggestion is based on the classic hold-out validation, where we have:

**70% training** : 61 subjects

**30% validation:** 26 subjects

~~~~~

From the **26** (13 men, 13 women) Validation Subjects we keep:

--> **5** with low expression--> 100914\_m\_39,  
101114\_w\_37,  
082315\_w\_60,  
083114\_w\_55,  
083109\_m\_60

--> **21** with normal expression \*(age group)

--> **7**(20-35)--> 4 men, 3 women

--> 072514\_m\_27,  
080309\_m\_29,  
112016\_m\_25,  
112310\_m\_20,  
092813\_w\_24,  
112809\_w\_23,  
112909\_w\_20

--> **7**(36-50)--> 3 men, 4 women

--> 071313\_m\_41,  
101309\_m\_48,  
101609\_m\_36,  
091809\_w\_43,  
102214\_w\_36,  
102316\_w\_50,  
112009\_w\_43

--> **7**(51-65)--> 4 men, 3 women

--> 101814\_m\_58,  
101908\_m\_61,  
102309\_m\_61,  
112209\_m\_51,  
112610\_w\_60,  
112914\_w\_51,  
120514\_w\_56

The remaining subjects we keep them for the training part.

In this way we retain almost the same percentage of the gender variation, the age and the "low-expressive" subjects.