Now imagine we find the distributions of reading times for our two conditions are positively skewed (and we discover the residuals are non-normal). We could log transform these two columns and have two new columns in our data frame - let's call them `log_simple` and `log_complex`. We can use the `mutate` function in the `dplyr` package to create two new columns.

```
> data_transformed <- mutate(dataRT_all, log_simple = log(Simple_Sentence),
log_complex = log(Complex_Sentence))
> data_transformed
     ID WM  IQ Comp Simple_Sentence Complex_Sentence log_Simple log_Complex
1    95 47  94   19            2154              2441   7.675082    7.758333
2   400 45 118   18            1824              2456   7.508787    7.825245
3   457 42 100   22            1857              2324   7.526718    7.912423
4  1138 41  77   18            1902              2341   7.550661    7.772753
5  1587 54  67   21            1844              2320   7.519692    7.685703
6  1805 52 109   19            2224              2256   7.707063    7.733684
7  1864 57 111   19            1880              2391   7.539027    7.800163
8  2006 44 110   19            2091              2456   7.645398    7.761745
9  2183 55 125   23            1926              2218   7.563201    7.771067
10 2318 51  91   21            1960              2440   7.580700    7.771489
```

Perhaps we have a reason to exclude a particular participant - number 2006 for example. We can use the filter function in `dplyr` to keep those participants where the ID number does not equal 2006.

```
filtered_data <- filter(data_transformed, ID != 2006)
```

!= stands for "not equal to"- here are other useful logical operators in R:

<	less than
<=	less than or equal to
>	greater than
>=	greater than or equal to
==	exactly equal to
!=	not equal to