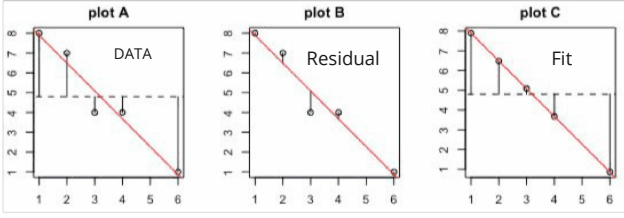
spencer Kittredge

Session 04 Group Quiz

A pair of students interested in studying how US students learn Spanish give five friends, who have taken 1 or more Spanish language courses, a challenging vocabulary test and note the number of mistakes.

```
semesters <- c(1,2,3,4,6) # number of semesters of Spanish language courses taken
mistakes  <- c(8,7,4,4,1) # number of mistakes on the vocabulary test
```

1. Using R they produce three plots showing different measurements of variance for their problem (sum of lengths of vertical lines squared). Identify which plot corresponds to each of the following: $s^2_{DATA}, s^2_{FIT}, s^2_{RESID}$



2. Find the values for $s^2_{DATA}, s^2_{FIT}, s^2_{RESID}$ in terms of the following statistics:

```
(s_sem <- sd(semesters))
```

```
## [1] 1.923538
```

S^2 FIT = (.9742202^2)(7.7)= 7.308

```
(s_mis <- sd(mistakes))
```

```
## [1] 2.774887
```

s^2 DATA = (2.774887)^2 = 7.7

```
(corr_sm <- cor(semesters,mistakes))
```

```
## [1] -0.9742202
```

s^2 RESID = (1-(-0.9742202^2))s^2 DATA = 0.3919

3. Confirm that $s^2_{DATA} = s^2_{FIT} + s^2_{RESID}$ for this case.

7.7 = 7.308 + 0.3919