Week 1 Assignment

MATH 511 Christina Morgenstern

---

1. Reading data file “ages-salaries.csv” into R and displaying data. The data is assigned to the data frame `df1`.

```
> df1<-read.csv("ages-salaries.csv")
> df1
```

```
name age salary
1     Alice  25  35000
2       Bob  30  65000
3     Carol  28  70000
4    Daniel  32  15000
5       Eve  42  43013
6      Fred  51  51777
7      Gail  31 100000
8     Harry  57  99999
9    Ileana  42  30001
10     John  40  45710
11    Kayla  27  54069
12     Lyle  57 303303
13    Maria  37  57634
14      Ned  45  45455
15     Olga  33  44449
16    Pedro  65 203315
17     Qing  31  87312
18   Robert  51  98765
19 Samantha  30  32567
20      Tom  70  42153
21   Ursula  55  77845
22   Victor  25  30000
23    Wanda  24  31000
24   Xavier  39  39000
25      Yun  35  50000
26     Zack  47  70000
```

2. Determine the number of observations in the dataset.

```
> str(df1)
```

```
'data.frame':    26 obs. of  3 variables:
 $ name  : Factor w/ 26 levels "Alice","Bob",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ age   : int  25 30 28 32 42 51 31 57 42 40 ...
 $ salary: int  35000 65000 70000 15000 43013 51777 100000 99999 30001 45710 ...
```

There are 26 observations in the dataset. An observation contains data of 3 variable types: name, age and salary. The variables age and salary are of `int` type.

3. Compute average age and average salary.

```
> mean(df1$age)
[1] 40.34615
```

The average age is about 40 years.

```
> mean(df1$salary)
[1] 70091.04
```

The average salary is about 70091$.

3. Compute standard deviation of age and salary.

```
> sd(df1$age)
[1] 12.92731
```

The sample standard deviation is about 12.93 years.

```
> sd(df1$salary)
[1] 60344.87
```

The sample standard deviation is about 60344.87 $.