

Exercise 3B: Data Management

**Problem 1.**

There are 592 cases of type 2 diabetes in our dataset.

**Problem 2.**

T2DM	Sex	Waist circumference mean (SD)	Total cholesterol mean (SD)
Yes	Male	111.04 (16.29)	186.20 (50.56)
Yes	Female	108.25 (16.46)	202.73 (44.37)
No	Male	99.98 (14.80)	196.49 (41.44)
No	Female	94.62 (15.13)	201.07 (41.27)

**CODE:**

```
> #Homework 3B | EPID 674.002 | Stephanie Mecham
>
> #Creating the T2DM variable
>
> diabetic <- ifelse(a1c >= 6.5, 1, 0)
> diabetic_meds <- ifelse (dmmed == 1,1,0)
> T2DM <- ifelse (!is.na(diabetic) & (diabetic==1) | (!is.na(diabetic_meds) & (diabetic_meds
==1)),1, 0)
>
> #Problem 1: Finding Frequency of Type 2 Diabetes
> sum(T2DM == 1)
[1] 592
>
>
> #Problem 2: Cross-tabulating
> aggregate(waist, by=list(Diabetes=T2DM, Sex=gender), mean, na.rm=T)
> Diabetes Sex      x
1      0 1 99.97987
```

```

2    1  1 111.03724
3    0  2 94.61666
4    1  2 108.25000
>
> aggregate(waist, by=list(Diabetes=T2DM, Sex=gender), sd, na.rm=T)
  Diabetes Sex    x
1     0  1 14.80492
2     1  1 16.28855
3     0  2 15.13392
4     1  2 16.46396
>
> aggregate(chol, by=list(Diabetes=T2DM, Sex=gender), mean, na.rm=T)
  Diabetes Sex    x
1     0  1 196.4940
2     1  1 186.2041
3     0  2 201.0700
4     1  2 202.7323
>
> aggregate(chol, by=list(Diabetes=T2DM, Sex=gender), sd, na.rm=T)
  Diabetes Sex    x
1     0  1 41.43644
2     1  1 50.55547
3     0  2 41.27048
4     1  2 44.36705
>
>
>#End of code

```