

Anatomy of Dplyr

Meredith Dodd <meredith.dodd@woz-u.com>

Wed 5/13/2020 4:08 PM

To: Meredith Dodd <meredith.dodd@woz-u.com>

Filter to only show those who have paid for their course

```
library("dplyr")
```

```
Paid <- student.mat %>% filter(paid == 'yes')
```

Filter to show only those who have either a mother or a father as a guardian

```
MotherFather <- student.mat %>% filter(guardian %in% c("mother", "father"))
```

Order that data frame so that we see all the mothers first - since it was just Mother's day

```
SortedMomFirst <- MotherFather %>% arrange(desc(guardian))
```

Include only the columns on their age, guardian, and whether they have paid

```
SortedMomFirstColumns = SortedMomFirst %>% select(age, guardian, paid)
```

Find the mean age of this dataset

```
MeanAge <- SortedMomFirstColumns %>% summarise(MeanAge=mean(age))
```

Find the mean age per guardian

```
MeanAgeByGuardian <- SortedMomFirstColumns %>% group_by(guardian) %>%  
summarise(MeanAge=mean(age))
```

Do this all at once - link them all!

```
Final <- student.mat %>%  
  filter(guardian %in% c("mother", "father")) %>%  
  arrange(desc(guardian)) %>%  
  select(age, guardian, paid) %>%  
  group_by(guardian) %>%  
  summarise(MeanAge=mean(age))
```

Meredith Dodd, Ph.D. | Data Science Program Chair and Instructor

meredith.dodd@woz-u.com

o: 480-291-8068

powered by



<https://woz-u.com>