stat hw 1

2022-09-26

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
pain<- read.csv("pain.csv")
score<- pain$Score
haircolor<- pain$HairColor
```

```
mean(score)
```

```
## [1] 47.84211
```

```
sd(score)
```

```
## [1] 11.4565
```

```
db<- filter(pain, haircolor == "DarkBrunette")
db
```

```
##      HairColor Score
## 1 DarkBrunette    32
## 2 DarkBrunette    39
## 3 DarkBrunette    51
## 4 DarkBrunette    30
## 5 DarkBrunette    35
```

```
mean(db$Score)
```

```
## [1] 37.4
```

```
lb<- filter(pain, haircolor == "LightBrunette")
lb
```

```
##       HairColor Score
## 1 LightBrunette    42
## 2 LightBrunette    50
## 3 LightBrunette    41
## 4 LightBrunette    37
```

```
sd(lb$Score)
```

```
## [1] 5.446712
```

```
count(lb)
```

```
##   n
## 1 4
```

```
pain %>%
  group_by(HairColor) %>%
  summarise(HairColor, mean=mean(Score)) %>%
  arrange(desc(mean))
```

```
## `summarise()` has grouped output by 'HairColor'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 19 x 2
## # Groups:   HairColor [4]
##    HairColor      mean
##    <chr>         <dbl>
##  1 LightBlond     59.2
##  2 LightBlond     59.2
##  3 LightBlond     59.2
##  4 LightBlond     59.2
##  5 LightBlond     59.2
##  6 DarkBlond      51.2
##  7 DarkBlond      51.2
##  8 DarkBlond      51.2
##  9 DarkBlond      51.2
## 10 DarkBlond      51.2
## 11 LightBrunette  42.5
## 12 LightBrunette  42.5
## 13 LightBrunette  42.5
## 14 LightBrunette  42.5
## 15 DarkBrunette   37.4
## 16 DarkBrunette   37.4
## 17 DarkBrunette   37.4
## 18 DarkBrunette   37.4
## 19 DarkBrunette   37.4
```

```
pain %>%
  group_by(HairColor) %>%
  summarise(HairColor, standard_dev=sd(Score)) %>%
  arrange(standard_dev)
```

```
## `summarise()` has grouped output by 'HairColor'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 19 x 2
## # Groups:   HairColor [4]
##    HairColor     standard_dev
##    <chr>                <dbl>
##  1 LightBrunette         5.45
##  2 LightBrunette         5.45
##  3 LightBrunette         5.45
##  4 LightBrunette         5.45
##  5 DarkBrunette          8.32
##  6 DarkBrunette          8.32
##  7 DarkBrunette          8.32
##  8 DarkBrunette          8.32
##  9 DarkBrunette          8.32
## 10 LightBlond            8.53
## 11 LightBlond            8.53
## 12 LightBlond            8.53
## 13 LightBlond            8.53
## 14 LightBlond            8.53
## 15 DarkBlond             9.28
## 16 DarkBlond             9.28
## 17 DarkBlond             9.28
## 18 DarkBlond             9.28
## 19 DarkBlond             9.28
```

```
pain %>%
  group_by(HairColor) %>%
  count()
```

```
## # A tibble: 4 x 2
## # Groups:   HairColor [4]
##   HairColor         n
##   <chr>         <int>
## 1 DarkBlond         5
## 2 DarkBrunette      5
## 3 LightBlond        5
## 4 LightBrunette     4
```