Homework 3

Task 1

a.

We are unable to use read.csv because it deals with data that is delimited by commas, not semicolons. Instead, read.csv2 is able to be used.

```
read.csv2("C:/Users/ESLil/OneDrive/Desktop/ST 558/Repos/Homework3/Homework-3/Data/data.txt")
```

```
Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
incomplete final line found by readTableHeader on
'C:/Users/ESLil/OneDrive/Desktop/ST
558/Repos/Homework3/Homework-3/Data/data.txt'
```

```
  x y z
1 1 2 3
2 5 3 8
```

b.

```
read.table("C:/Users/ESLil/OneDrive/Desktop/ST 558/Repos/Homework3/Homework-3/Data/data2.txt
           skip = 1,
           sep = "6",
           col.names = c("x", "y", "z"),
           colClasses = c("factor", "numeric", "character")
           )
```

```
  x y z
1 1 2 3
2 5 3 8
3 7 4 2
```

Task 2

a.

```
trailblazer <- read.csv("C:/Users/ESLil/OneDrive/Desktop/ST 558/Repos/Homework3/Homework-3/Da

glimpse(trailblazer)
```

```
Rows: 9
Columns: 11
$ Player      <chr> "Damian Lillard", "CJ McCollum", "Norman Powell", "Robert ~
$ Game1_Home  <int> 20, 24, 14, 8, 20, 5, 11, 2, 7
$ Game2_Home  <int> 19, 28, 16, 6, 9, 5, 18, 8, 11
$ Game3_Away  <int> 12, 20, NA, 0, 4, 8, 12, 5, 5
$ Game4_Home  <int> 20, 25, NA, 3, 17, 10, 17, 8, 9
$ Game5_Home  <int> 25, 14, 12, 9, 14, 9, 5, 3, 8
$ Game6_Away  <int> 14, 25, 14, 6, 13, 6, 19, 8, 8
$ Game7_Away  <int> 20, 20, 22, 0, 7, 0, 17, 7, 4
$ Game8_Away  <int> 26, 21, 23, 6, 6, 7, 15, 0, 0
$ Game9_Home  <int> 4, 27, 25, 19, 10, 0, 16, 2, 7
$ Game10_Home <int> 25, 7, 13, 12, 15, 6, 10, 4, 8
```

b.

```
trailblazer_longer <- trailblazer |>
  pivot_longer(
    cols = -Player,
    names_to = c("Game", "Location"),
    names_sep = "_",
    values_to = "Points"
  )

head(trailblazer_longer)
```

```
# A tibble: 6 x 4
  Player         Game  Location Points
  <chr>          <chr> <chr>     <int>
1 Damian Lillard Game1 Home         20
2 Damian Lillard Game2 Home         19
```

```
3 Damian Lillard Game3 Away         12
4 Damian Lillard Game4 Home         20
5 Damian Lillard Game5 Home         25
6 Damian Lillard Game6 Away         14
```

c.

```
trailblazer_longer |>
  pivot_wider(names_from = Location,
              values_from = Points) |>
  group_by(Player) |>
  summarize(mean_home = mean(Home, na.rm = T),
            mean_away = mean (Away, na.rm = T)) |>
  mutate(Diff = mean_home - mean_away) |>
  arrange(Diff)
```

```
# A tibble: 9 x 4
  Player           mean_home mean_away   Diff
  <chr>                <dbl>     <dbl>  <dbl>
1 Norman Powell        16        19.7  -3.67
2 Anfernee Simons      12.8      15.8  -2.92
3 CJ McCollum          20.8      21.5  -0.667
4 Larry Nance Jr        4.5       5    -0.5
5 Cody Zeller           5.83      5.25  0.583
6 Damian Lillard       18.8      18     0.833
7 Nassir Little         8.33      4.25  4.08
8 Robert Covington      9.5       3     6.5
9 Jusuf Nurkic         14.2       7.5   6.67
```

The players that scored more points at home are the last 5 players in the tibble.

Task 3

a.

means that there are no bill length values for that observation.

<dbl [52]> means that it is a numeric vector with 52 observations in the vector.

means that each column is a list.

b.

```
penguins |>
  select(species, island) |>
  count(species, island) |>
  pivot_wider(names_from = island,
              values_fill = 0,
              values_from = n)
```

```
# A tibble: 3 x 4
  species   Biscoe Dream Torgersen
  <fct>      <int> <int>     <int>
1 Adelie        44    56        52
2 Chinstrap      0    68         0
3 Gentoo       124     0         0
```

Task 4

```
penguins |>
  select(species, bill_length_mm) |>
  mutate(bill_length_mm = case_when(
    is.na(bill_length_mm) & species == "Adelie" ~ 26,
    is.na(bill_length_mm) & species == "Gentoo" ~ 30,
    TRUE ~ bill_length_mm
  )) |>
  arrange(bill_length_mm)
```

```
# A tibble: 344 x 2
   species bill_length_mm
   <fct>            <dbl>
 1 Adelie            26
 2 Gentoo            30
 3 Adelie            32.1
 4 Adelie            33.1
 5 Adelie            33.5
 6 Adelie            34
 7 Adelie            34.1
 8 Adelie            34.4
 9 Adelie            34.5
```

```
10 Adelie            34.6
# i 334 more rows
```