# Module 6 - Assignment 1

## Tate, Levi

### Lubridate

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.3.2 ✓ purrr 0.3.4  
## ✓ tibble 3.0.4 ✓ dplyr 1.0.2  
## ✓ tidyr 1.1.2 ✓ stringr 1.4.0  
## ✓ readr 1.4.0 ✓ forcats 0.5.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(lubridate)

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':  
##   
## date, intersect, setdiff, union

Appointments <- read\_csv("Appointments.csv")

##   
## ── Column specification ────────────────────────────────────────────────────────  
## cols(  
## Age = col\_double(),  
## Sex = col\_double(),  
## ReservationMonth = col\_double(),  
## ReservationDay = col\_double(),  
## ReservationHour = col\_double(),  
## ReservationMinute = col\_double(),  
## CreationMonth = col\_double(),  
## CreationDay = col\_double(),  
## CreationYear = col\_double(),  
## CreationHour = col\_double(),  
## `Creation Minute` = col\_double(),  
## Show = col\_double()  
## )

### Working with Dates

Appointments <- Appointments %>%  
 mutate(ReservationYear = 2019,   
 ReservationDate = make\_date (month = ReservationMonth, day = ReservationDay, year = ReservationYear),  
 CreationDate = make\_date (month = CreationMonth, day = CreationDay, year = CreationYear),  
 ReservationSpan = ReservationDate - CreationDate  
 )

### Exploratory Analysis

summary(Appointments)

## Age Sex ReservationMonth ReservationDay   
## Min. : 0.00 Min. :1.000 Min. :1.000 Min. : 1.00   
## 1st Qu.: 16.00 1st Qu.:1.000 1st Qu.:2.000 1st Qu.: 7.00   
## Median : 34.00 Median :2.000 Median :3.000 Median :15.00   
## Mean : 34.64 Mean :1.597 Mean :2.527 Mean :14.51   
## 3rd Qu.: 51.00 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:21.00   
## Max. :101.00 Max. :2.000 Max. :4.000 Max. :28.00   
## ReservationHour ReservationMinute CreationMonth CreationDay   
## Min. : 0.00 Min. :0 Min. : 1.000 Min. : 1.00   
## 1st Qu.:11.00 1st Qu.:0 1st Qu.: 2.000 1st Qu.: 7.00   
## Median :14.00 Median :0 Median : 2.000 Median :14.00   
## Mean :13.85 Mean :0 Mean : 5.265 Mean :14.51   
## 3rd Qu.:17.00 3rd Qu.:0 3rd Qu.:12.000 3rd Qu.:22.00   
## Max. :21.00 Max. :0 Max. :12.000 Max. :28.00   
## CreationYear CreationHour Creation Minute Show   
## Min. :2018 Min. : 0.00 Min. :0 Min. :0.0000   
## 1st Qu.:2018 1st Qu.:10.00 1st Qu.:0 1st Qu.:1.0000   
## Median :2019 Median :13.00 Median :0 Median :1.0000   
## Mean :2019 Mean :13.42 Mean :0 Mean :0.7898   
## 3rd Qu.:2019 3rd Qu.:16.00 3rd Qu.:0 3rd Qu.:1.0000   
## Max. :2019 Max. :23.00 Max. :0 Max. :1.0000   
## ReservationYear ReservationDate CreationDate ReservationSpan   
## Min. :2019 Min. :2019-01-01 Min. :2018-10-03 Length:61214   
## 1st Qu.:2019 1st Qu.:2019-02-01 1st Qu.:2018-12-21 Class :difftime   
## Median :2019 Median :2019-03-04 Median :2019-01-20 Mode :numeric   
## Mean :2019 Mean :2019-03-01 Mean :2019-01-21   
## 3rd Qu.:2019 3rd Qu.:2019-03-28 3rd Qu.:2019-02-19   
## Max. :2019 Max. :2019-04-28 Max. :2019-03-28

Appointments$ReservationSpan <- as.numeric(Appointments$ReservationSpan)  
summary(Appointments)

## Age Sex ReservationMonth ReservationDay   
## Min. : 0.00 Min. :1.000 Min. :1.000 Min. : 1.00   
## 1st Qu.: 16.00 1st Qu.:1.000 1st Qu.:2.000 1st Qu.: 7.00   
## Median : 34.00 Median :2.000 Median :3.000 Median :15.00   
## Mean : 34.64 Mean :1.597 Mean :2.527 Mean :14.51   
## 3rd Qu.: 51.00 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:21.00   
## Max. :101.00 Max. :2.000 Max. :4.000 Max. :28.00   
## ReservationHour ReservationMinute CreationMonth CreationDay   
## Min. : 0.00 Min. :0 Min. : 1.000 Min. : 1.00   
## 1st Qu.:11.00 1st Qu.:0 1st Qu.: 2.000 1st Qu.: 7.00   
## Median :14.00 Median :0 Median : 2.000 Median :14.00   
## Mean :13.85 Mean :0 Mean : 5.265 Mean :14.51   
## 3rd Qu.:17.00 3rd Qu.:0 3rd Qu.:12.000 3rd Qu.:22.00   
## Max. :21.00 Max. :0 Max. :12.000 Max. :28.00   
## CreationYear CreationHour Creation Minute Show   
## Min. :2018 Min. : 0.00 Min. :0 Min. :0.0000   
## 1st Qu.:2018 1st Qu.:10.00 1st Qu.:0 1st Qu.:1.0000   
## Median :2019 Median :13.00 Median :0 Median :1.0000   
## Mean :2019 Mean :13.42 Mean :0 Mean :0.7898   
## 3rd Qu.:2019 3rd Qu.:16.00 3rd Qu.:0 3rd Qu.:1.0000   
## Max. :2019 Max. :23.00 Max. :0 Max. :1.0000   
## ReservationYear ReservationDate CreationDate ReservationSpan   
## Min. :2019 Min. :2019-01-01 Min. :2018-10-03 Min. : 1.00   
## 1st Qu.:2019 1st Qu.:2019-02-01 1st Qu.:2018-12-21 1st Qu.: 25.00   
## Median :2019 Median :2019-03-04 Median :2019-01-20 Median : 35.00   
## Mean :2019 Mean :2019-03-01 Mean :2019-01-21 Mean : 38.53   
## 3rd Qu.:2019 3rd Qu.:2019-03-28 3rd Qu.:2019-02-19 3rd Qu.: 49.00   
## Max. :2019 Max. :2019-04-28 Max. :2019-03-28 Max. :207.00

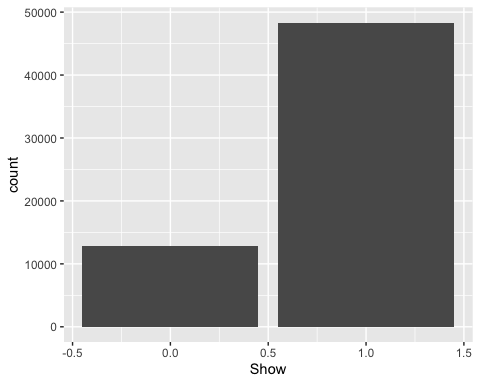
1. What is the average time between appointment creation and the actual reservation date?  
   Average time between appointment creation and actual reservation: 38.53
2. What was the max time?  
   Max time: 207
3. What was the min time?  
   Minimum time: 1

cor(Appointments$ReservationSpan, Appointments$Show, method = "pearson")

## [1] 0.002693853

1. Are ReservationSpan and Show highly correlated?  
   These variables have a correlation coefficient of .0027, therefore no correlation.

ggplot(data = Appointments, mapping = aes(x = Show)) +  
geom\_bar()



1. Based on 0 being “no show” and 1 being “show”, did most people make their appointments?  
   Most people showed up to their appointments(~47,000 > ~13,000).