Homework 1

Ama Konadu Appau & Akash Agrawal Bejarano

2023-01-26

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.3.6 ✔ purrr 0.3.4  
## ✔ tibble 3.1.8 ✔ dplyr 1.0.9  
## ✔ tidyr 1.2.0 ✔ stringr 1.4.1  
## ✔ readr 2.1.2 ✔ forcats 0.5.2  
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(nycflights13)  
library(lubridate)

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

library(dplyr)

Question 1

(("2020-06-06") %--% ("2021-07-14"))/days(1)

## [1] 403

Question 2

leap\_year(1988)

## [1] TRUE

Question 3

leap\_year(1989)

## [1] FALSE

Question 4

There is no standard value of months in terms of seconds since months have different numbers of days.

Question 5

(("1962-04-11") %--%today())/years(1)

## [1] 60.79452

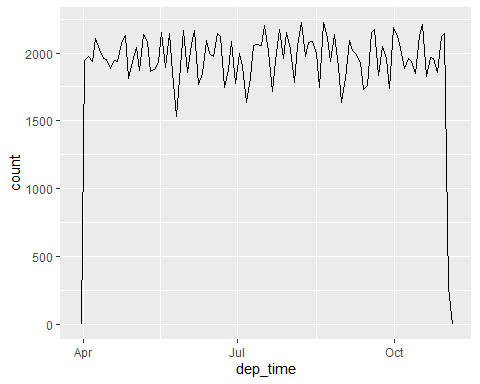
Question 6

make\_datetime\_100 <- function(year, month, day, time) {  
 make\_datetime(year, month, day, time %/% 100, time %% 100)  
}  
flights %>%   
 filter(!is.na(dep\_time), !is.na(arr\_time)) %>%   
 mutate(  
 dep\_time = make\_datetime\_100(year, month, day, dep\_time),  
 arr\_time = make\_datetime\_100(year, month, day, arr\_time),  
 ) %>%   
 select(origin, dest,carrier,arr\_time,dep\_time) ->flights\_dt  
flights\_dt

## # A tibble: 328,063 × 5  
## origin dest carrier arr\_time dep\_time   
## <chr> <chr> <chr> <dttm> <dttm>   
## 1 EWR IAH UA 2013-01-01 08:30:00 2013-01-01 05:17:00  
## 2 LGA IAH UA 2013-01-01 08:50:00 2013-01-01 05:33:00  
## 3 JFK MIA AA 2013-01-01 09:23:00 2013-01-01 05:42:00  
## 4 JFK BQN B6 2013-01-01 10:04:00 2013-01-01 05:44:00  
## 5 LGA ATL DL 2013-01-01 08:12:00 2013-01-01 05:54:00  
## 6 EWR ORD UA 2013-01-01 07:40:00 2013-01-01 05:54:00  
## 7 EWR FLL B6 2013-01-01 09:13:00 2013-01-01 05:55:00  
## 8 LGA IAD EV 2013-01-01 07:09:00 2013-01-01 05:57:00  
## 9 JFK MCO B6 2013-01-01 08:38:00 2013-01-01 05:57:00  
## 10 LGA ORD AA 2013-01-01 07:53:00 2013-01-01 05:58:00  
## # … with 328,053 more rows

Question 7

flights %>%   
 filter(!is.na(dep\_time), !is.na(arr\_time)) %>%   
 mutate(  
 dep\_time = make\_datetime\_100(year, month, day, dep\_time),  
 arr\_time = make\_datetime\_100(year, month, day, arr\_time),  
 sched\_dep\_time = make\_datetime\_100(year, month, day, sched\_dep\_time),  
 sched\_arr\_time = make\_datetime\_100(year, month, day, sched\_arr\_time)  
 ) %>%   
 select(origin, dest, ends\_with("delay"), ends\_with("time")) -> flights\_dt1  
  
flights\_dt1 %>%   
 filter(month(dep\_time)>=4, month(dep\_time)<=10) -> flights\_dt\_7   
   
 ggplot(flights\_dt\_7, aes(x = dep\_time)) + geom\_freqpoly(bins = 100)



Question 8

flights\_dt%>%   
 filter(origin=="LGA")%>%  
 filter(dest=="DFW")%>%  
 filter(carrier=="AA")%>%  
 slice(115:125)

## # A tibble: 11 × 5  
## origin dest carrier arr\_time dep\_time   
## <chr> <chr> <chr> <dttm> <dttm>   
## 1 LGA DFW AA 2013-01-09 16:16:00 2013-01-09 13:25:00  
## 2 LGA DFW AA 2013-01-09 19:17:00 2013-01-09 15:31:00  
## 3 LGA DFW AA 2013-01-09 19:36:00 2013-01-09 16:31:00  
## 4 LGA DFW AA 2013-01-09 20:53:00 2013-01-09 17:40:00  
## 5 LGA DFW AA 2013-01-09 22:24:00 2013-01-09 19:29:00  
## 6 LGA DFW AA 2013-01-10 08:37:00 2013-01-10 05:53:00  
## 7 LGA DFW AA 2013-01-10 10:20:00 2013-01-10 07:22:00  
## 8 LGA DFW AA 2013-01-10 11:22:00 2013-01-10 08:29:00  
## 9 LGA DFW AA 2013-01-10 12:16:00 2013-01-10 09:14:00  
## 10 LGA DFW AA 2013-01-10 13:19:00 2013-01-10 10:16:00  
## 11 LGA DFW AA 2013-01-10 13:23:00 2013-01-10 10:48:00

Question 9

d <-"2013-01-09 16:16:00 "  
w <-"2013-01-09 19:17:00"   
difftime(w,d,units = "mins")

## Time difference of 181 mins