Homework #14

Dates and Times

## Concepts

1. Software packages often use a reference date for their date-time functions. Excel uses January 1, 1990. One alternative for R is “Unix Epoch”, which is 1970-01-01. Explain how these systems represent date and date-time values as numbers.
2. Time spans can be measured as durations, periods, or intervals. Compare and contrast these three ways of considering time.

## Interpretation of code

Interpret the following two R code chunks:

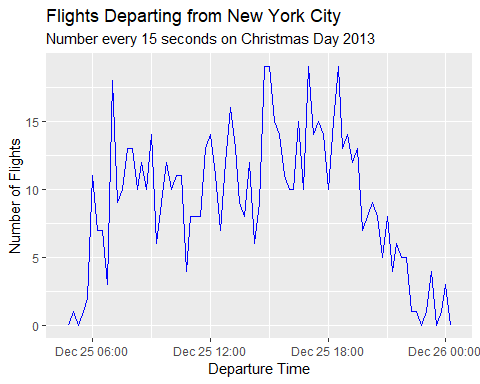
### Code chunk #1

# Create a Function  
make\_datetime\_100 <- function(year, month, day, time){  
 make\_datetime(year, month, day, time %/% 100, time %% 100)  
}  
# Convert departure and arrival times into date-time variables  
flights\_dt <- 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\_dt %>% print()

## # A tibble: 328,063 x 9  
## origin dest dep\_delay arr\_delay dep\_time sched\_dep\_time   
## <chr> <chr> <dbl> <dbl> <dttm> <dttm>   
## 1 EWR IAH 2 11 2013-01-01 05:17:00 2013-01-01 05:15:00  
## 2 LGA IAH 4 20 2013-01-01 05:33:00 2013-01-01 05:29:00  
## 3 JFK MIA 2 33 2013-01-01 05:42:00 2013-01-01 05:40:00  
## 4 JFK BQN -1 -18 2013-01-01 05:44:00 2013-01-01 05:45:00  
## 5 LGA ATL -6 -25 2013-01-01 05:54:00 2013-01-01 06:00:00  
## 6 EWR ORD -4 12 2013-01-01 05:54:00 2013-01-01 05:58:00  
## 7 EWR FLL -5 19 2013-01-01 05:55:00 2013-01-01 06:00:00  
## 8 LGA IAD -3 -14 2013-01-01 05:57:00 2013-01-01 06:00:00  
## 9 JFK MCO -3 -8 2013-01-01 05:57:00 2013-01-01 06:00:00  
## 10 LGA ORD -2 8 2013-01-01 05:58:00 2013-01-01 06:00:00  
## # ... with 328,053 more rows, and 3 more variables: arr\_time <dttm>,  
## # sched\_arr\_time <dttm>, air\_time <dbl>

### Code chunk #2

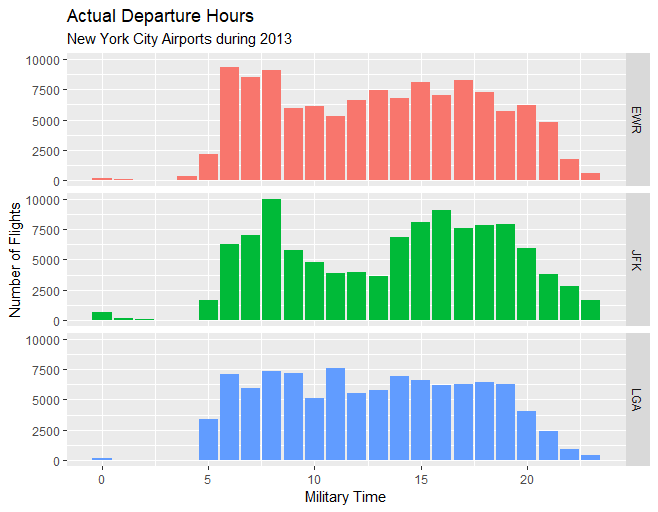
# Uses the flights\_dt from the first code chunk  
flights\_dt %>%   
 filter(dep\_time >= ymd(20131225)) %>%  
 filter(dep\_time < ymd(20131226)) %>%   
 ggplot(aes(dep\_time)) +  
 geom\_freqpoly(binwidth = 900, color = "blue") +  
 labs(  
 title = "Flights Departing from New York City",  
 subtitle = "Number every 15 seconds on Christmas Day 2013",  
 x = "Departure Time",  
 y = "Number of Flights"  
 )



## Write code

Problem 1

1. Remove observations with missing values for dep\_time
2. Select origin, year, month, day, and dep\_time from the flights data set
3. Convert dep\_time into a time-date variable using the make\_datetime function
4. Use the hour function to create a new hour variable from the dep\_time
5. Draw a bar chart for the new hour variable
6. Use faceting to create a chart for each airport



Problem 2

Use durations, periods, and spans to calculate your age today. The functions as.duration, as.period, years, and the %–% operator will be helpful.