

Laboratorio 05

EmilySoto

6/9/2021

```
library(lubridate)
```

```
## Warning: package 'lubridate' was built under R version 4.1.1
```

```
##
```

```
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      date, intersect, setdiff, union
```

```
library(nycflights13)
```

```
## Warning: package 'nycflights13' was built under R version 4.1.1
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(readxl)
```

```
library(zoo)
```

```
## Warning: package 'zoo' was built under R version 4.1.1
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```



```
data$nuevas=as.Date(data$nuevas)
data$mes=format(data$nuevas, "%m")
data$dia=weekdays(data$nuevas)
```

Pregunta 1

1. ¿En qué meses existe una mayor cantidad de llamadas por código?

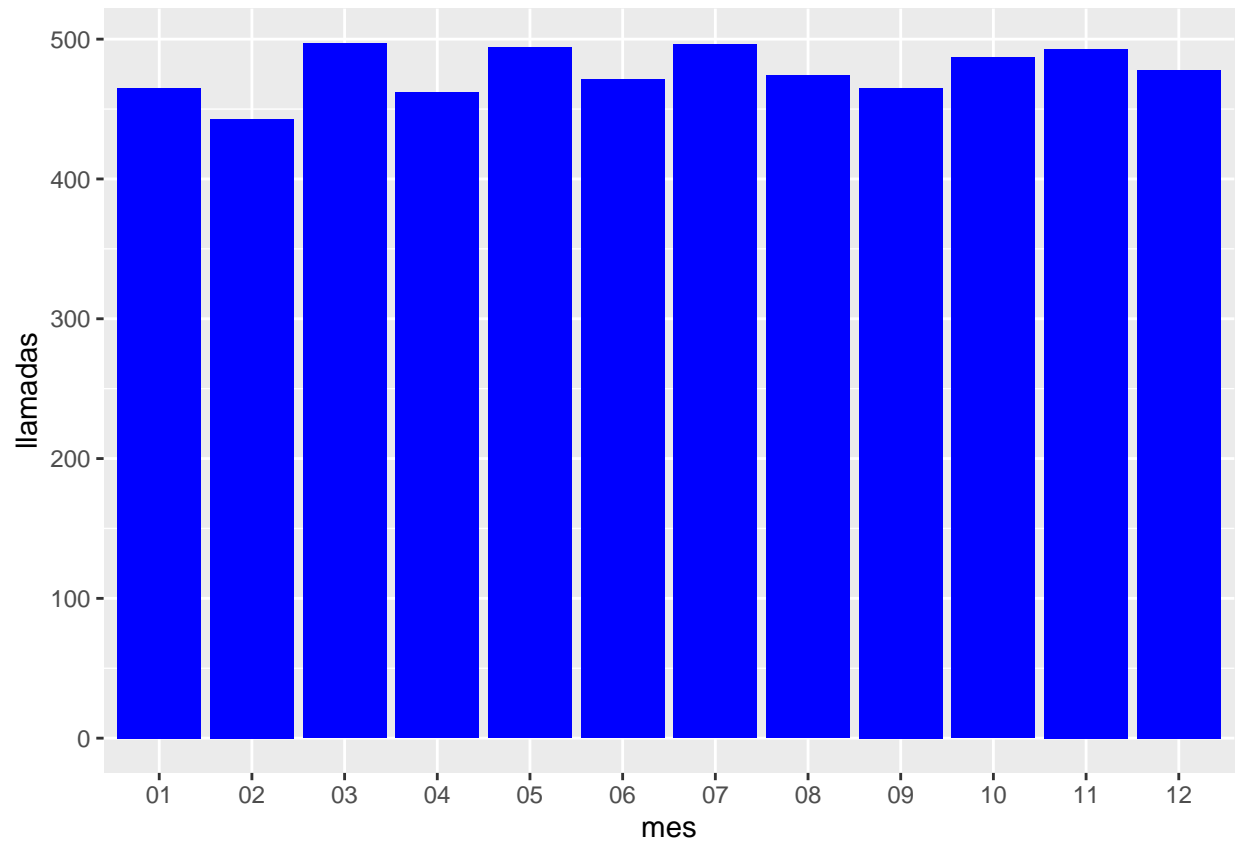
```
llamadasxmes=data %>%filter(Call==1) %>% group_by(mes, Cod) %>%
  summarize(llamadas = n())
```

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

```
llamadasxmes=llamadasxmes[order(llamadasxmes$llamadas, decreasing = T),]
llamadasxmes
```

```
## # A tibble: 12 x 3
## # Groups:   mes [12]
##   mes   Cod               llamadas
##   <chr> <chr>             <int>
## 1 03   Actualización de Información    497
## 2 07   Actualización de Información    496
## 3 05   Actualización de Información    494
## 4 11   Actualización de Información    493
## 5 10   Actualización de Información    487
## 6 12   Actualización de Información    478
## 7 08   Actualización de Información    474
## 8 06   Actualización de Información    471
## 9 01   Actualización de Información    465
## 10 09   Actualización de Información    465
## 11 04   Actualización de Información    462
## 12 02   Actualización de Información    443
```

```
ggplot(data=llamadasxmes)+geom_bar(aes(x=mes, y=llamadas),fill='blue', stat="identity")
```



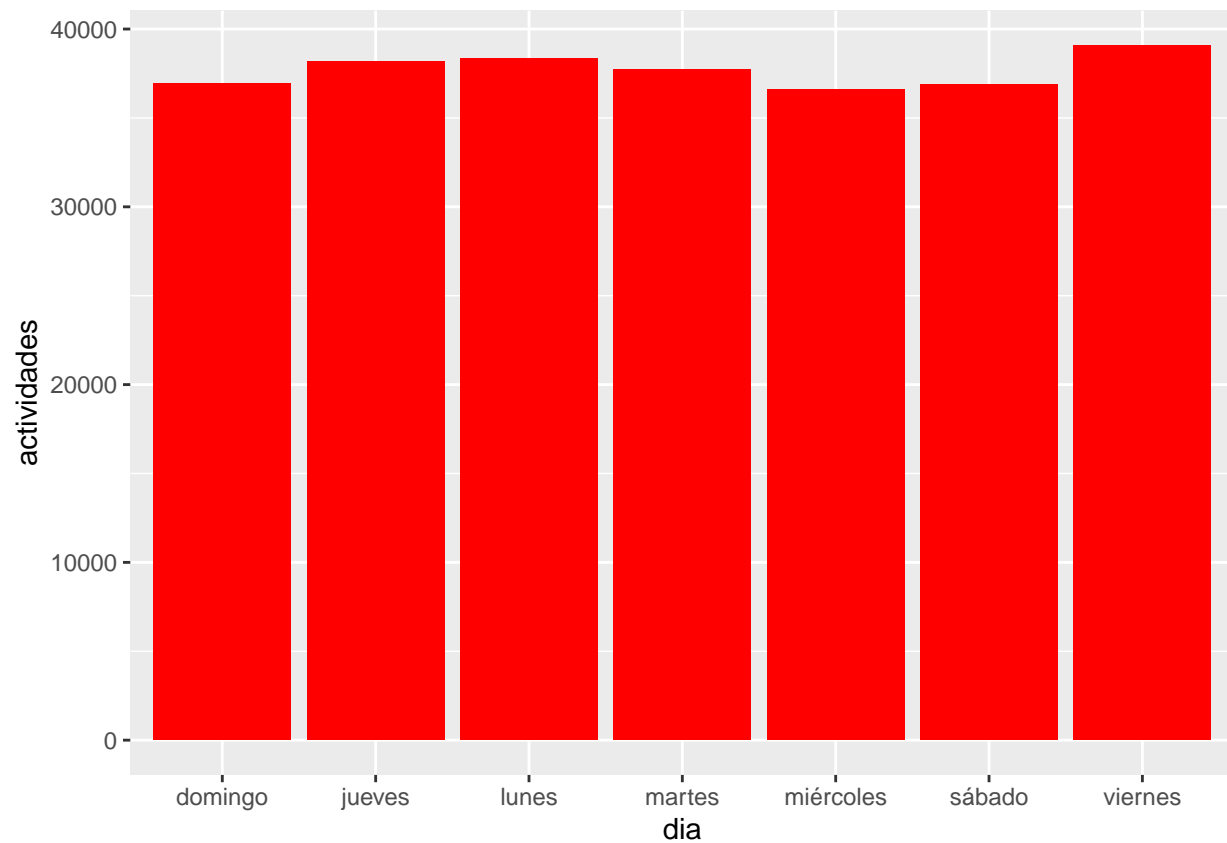
R/ Los meses con más llamadas son marzo, julio y mayo

Pregunta 2 ¿Qué día de la semana es el más ocupado? Assumption: Más ocupado=Más llamadas, sms o emails

```
dia_ocu=data %>%group_by(dia) %>% summarize(actividades= n())
dia_ocu=dia_ocu[order(dia_ocu$actividades, decreasing = T),]
dia_ocu
```

```
## # A tibble: 7 x 2
##   dia      actividades
##   <chr>         <int>
## 1 viernes      39083
## 2 lunes        38339
## 3 jueves       38179
## 4 martes       37726
## 5 domingo      36921
## 6 sábado       36873
## 7 miércoles    36604
```

```
ggplot(data=dia_ocu)+geom_bar(aes(x=dia, y=actividades),fill='red', stat="identity")
```



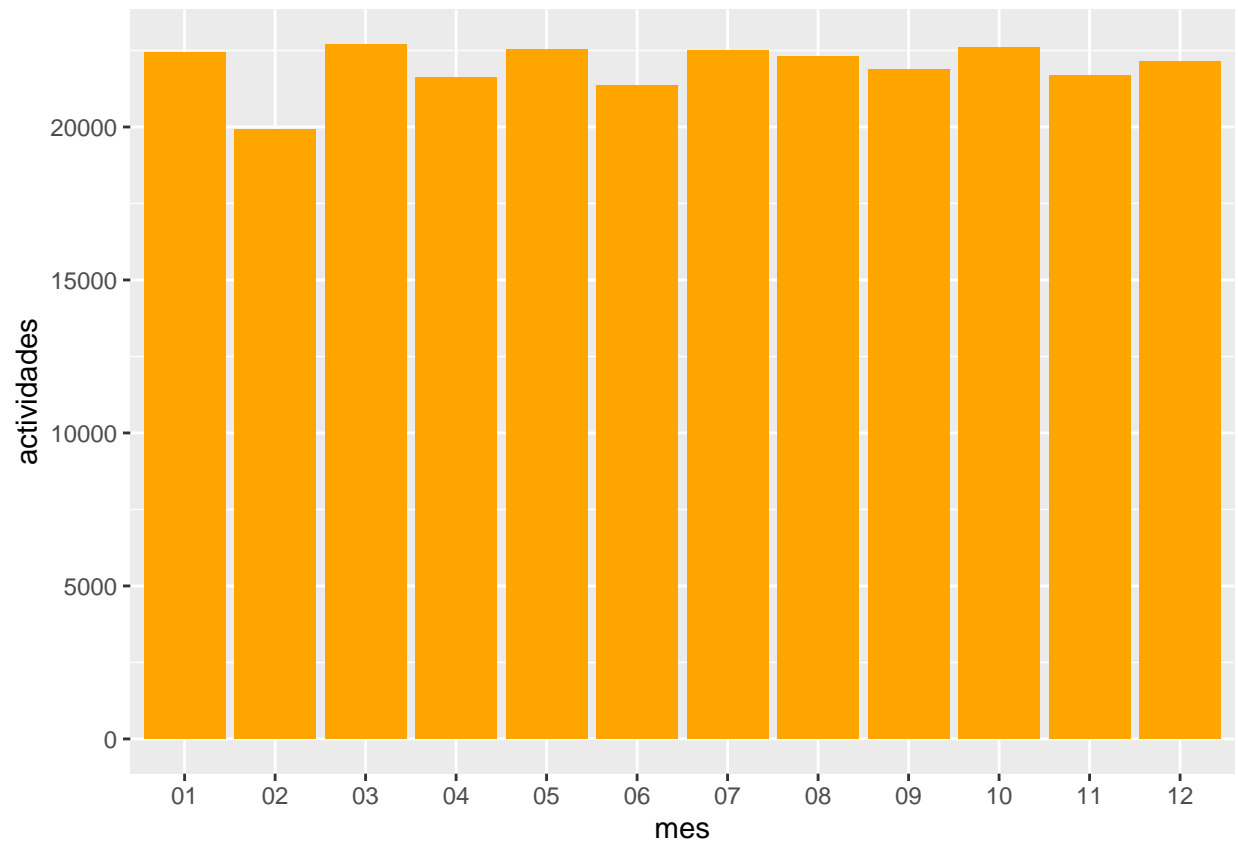
R/ Los días más ocupados son los viernes

Pregunta 3 ¿Qué mes es el más ocupado?

```
mes_ocu=data %>%group_by(mes) %>% summarize(actividades= n())
mes_ocu=mes_ocu[order(mes_ocu$actividades, decreasing = T),]
mes_ocu
```

```
## # A tibble: 12 x 2
##   mes   actividades
##   <chr>         <int>
## 1 03           22708
## 2 10           22601
## 3 05           22525
## 4 07           22514
## 5 01           22425
## 6 08           22316
## 7 12           22151
## 8 09           21891
## 9 11           21681
## 10 04          21611
## 11 06          21370
## 12 02          19932
```

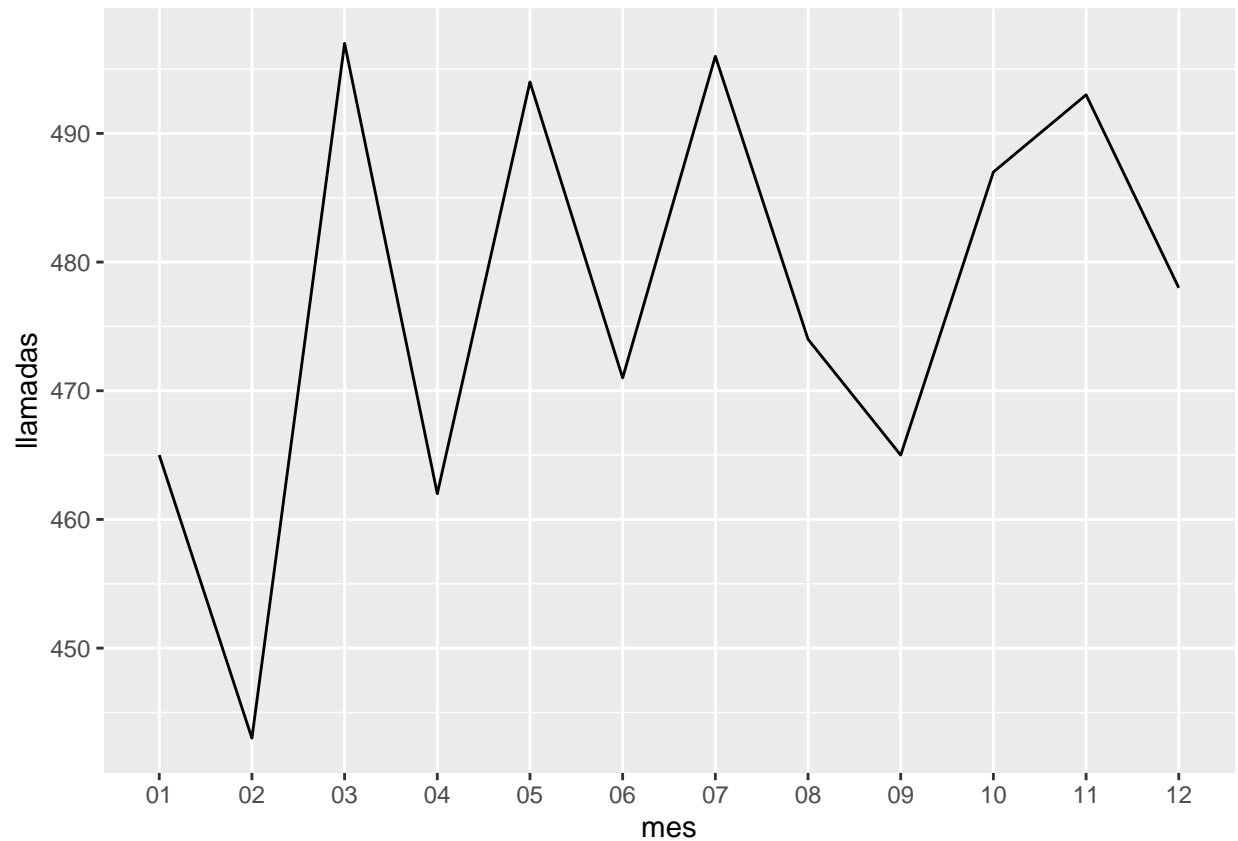
```
ggplot(data=mes_ocu)+geom_bar(aes(x=mes, y=actividades),fill='orange', stat="identity")
```



R/ El mes más ocupado es marzo

Pregunta 4 ¿Existe una concentración o estacionalidad en la cantidad de llamadas?

```
ggplot(llamadasxmes, aes(x=mes, y=llamadas, group = 1)) +  
  geom_line()
```



R/ Se puede notar en la gráfica que en el inicio del año existe una tendencia a la alza de llamadas cada dos meses, especialmente en marzo, mayo y julio. Sin embargo, el mes con un pico más alto de llamadas es marzo.

Pregunta 5 ¿Cuántos minutos dura la llamada promedio?

```
data$duracion=difftime((data$`Hora Final`),(data$`Hora Creación`),units = "mins")
llamadas=data %>%filter(Call==1)%>%select((duracion))
mean(llamadas$duracion)
```

```
## Time difference of 14.5579 mins
```

```
frec_minllamada=as.data.frame(table(llamadas))
names(frec_minllamada)[1]="Tiempo llamada en minutos"
names(frec_minllamada)[2]="Cantidad de llamadas"
frec_minllamada
```

Pregunta 6

```
##      Tiempo llamada en minutos Cantidad de llamadas
## 1              0              221
## 2              1              211
```

## 3	2	173
## 4	3	195
## 5	4	193
## 6	5	184
## 7	6	194
## 8	7	197
## 9	8	212
## 10	9	166
## 11	10	190
## 12	11	197
## 13	12	169
## 14	13	163
## 15	14	203
## 16	15	188
## 17	16	181
## 18	17	178
## 19	18	186
## 20	19	190
## 21	20	179
## 22	21	205
## 23	22	175
## 24	23	192
## 25	24	186
## 26	25	174
## 27	26	157
## 28	27	173
## 29	28	158
## 30	29	171
## 31	30	164

Parte 3

```

zodiac_sign=function(date){
  birth=ymd(date)
  month <- month(birth)
  day <- day(birth)
  if(month == 1){
    if(day < 20){
      zodic_sign= "capricorn"
    }else{
      zodic_sign= "aquarius"
    }
  }else{
    if(month == 2){
      if(dia < 19){
        zodic_sign= "aquarius"
      }else{
        zodic_sign= "piscis"
      }
    }else{
      if(month == 3){
        if(day < 21){
          zodic_sign= "piscis"
        }
      }
    }
  }
}

```



```

    }else{
        zodic_sign= "aries"
    }
}
}else{
    if(month == 4){
        if(day < 20){
            zodic_sign= "aries"
        }else{
            zodic_sign= "taurus"
        }
    }
    }else{
        if(month == 5){
            if(day < 21){
                zodic_sign= "taurus"
            }else{
                zodic_sign= "gemini"
            }
        }
        }else{
            if(month == 6){
                if(day < 21){
                    zodic_sign= "gemini"
                }else{
                    zodic_sign= "cancer"
                }
            }
            }else{
                if(month == 7){
                    if(day < 23){
                        zodic_sign= "cancer"
                    }else{
                        zodic_sign= "leo"
                    }
                }
                }else{
                    if(month == 8){
                        if(day < 23){
                            zodic_sign= "leo"
                        }else{
                            zodic_sign= "virgo"
                        }
                    }
                    }else{
                        if(month == 9){
                            if(day < 23){
                                zodic_sign= "virgo"
                            }else{
                                zodic_sign= "libra"
                            }
                        }
                        }else{
                            if(month == 10){
                                if(day < 23){
                                    zodic_sign= "libra"
                                }else{
                                    zodic_sign= "scorpio"
                                }
                            }
                            }else{

```

```

        if(month == 11){
          if(day < 22){
            zodic_sign= "scorpio"
          }else{
            zodic_sign= "Sagittarius"
          }
        }else{
          if(month == 12){
            if(day < 22){
              zodic_sign= "Sagittarius"
            }else{
              zodic_sign= "capricorn"
            }
          }
        }
      }
    }
  }
}
return(zodic_sign)
}
zodiac_sign( "2001-march-21")

```

```
## [1] "aries"
```

Parte 4

```

flights=flights
time_hm= function(year, month, day, time) {
  make_datetime(year, month, day, time %% 100, time %% 100)
}

new=flights %>%
  mutate(new_dep_time=time_hm(year, month, day, dep_time),
         new_arr_time=time_hm(year, month, day, arr_time),
         new_sch_dep=time_hm(year, month, day, sched_dep_time),
         new_sch_arr=time_hm(year, month, day, sched_arr_time))
new[,20:23]

```

Pregunta 1

```

## # A tibble: 336,776 x 4
##   new_dep_time      new_arr_time      new_sch_dep
##   <dtm>          <dtm>          <dtm>

```

```
## 1 2013-01-01 05:17:00 2013-01-01 08:30:00 2013-01-01 05:15:00
## 2 2013-01-01 05:33:00 2013-01-01 08:50:00 2013-01-01 05:29:00
## 3 2013-01-01 05:42:00 2013-01-01 09:23:00 2013-01-01 05:40:00
## 4 2013-01-01 05:44:00 2013-01-01 10:04:00 2013-01-01 05:45:00
## 5 2013-01-01 05:54:00 2013-01-01 08:12:00 2013-01-01 06:00:00
## 6 2013-01-01 05:54:00 2013-01-01 07:40:00 2013-01-01 05:58:00
## 7 2013-01-01 05:55:00 2013-01-01 09:13:00 2013-01-01 06:00:00
## 8 2013-01-01 05:57:00 2013-01-01 07:09:00 2013-01-01 06:00:00
## 9 2013-01-01 05:57:00 2013-01-01 08:38:00 2013-01-01 06:00:00
## 10 2013-01-01 05:58:00 2013-01-01 07:53:00 2013-01-01 06:00:00
## # ... with 336,766 more rows, and 1 more variable: new_sch_arr <dtm>
```

```
total_delay=new %>%
  select(flight,dep_delay, arr_delay) %>%
  mutate(delay=minutes(dep_delay+arr_delay))
total_delay
```

Pregunta 2

```
## # A tibble: 336,776 x 4
##   flight dep_delay arr_delay delay
##   <int>      <dbl>      <dbl> <Period>
## 1 1545         2        11 13M OS
## 2 1714         4        20 24M OS
## 3 1141         2        33 35M OS
## 4 725        -1       -18 -19M OS
## 5 461        -6       -25 -31M OS
## 6 1696        -4        12 8M OS
## 7 507         -5        19 14M OS
## 8 5708        -3       -14 -17M OS
## 9 79          -3        -8 -11M OS
## 10 301        -2         8 6M OS
## # ... with 336,766 more rows
```