



**UNIVERSIDAD TÉCNICA DE AMBATO**  
**FACULTAD DE INGENIERÍA MECÁNICA**  
**TALLER MECÁNICO INDUSTRIAL II**



**Nombre:** Edison Quispe

**Curso:** Tercero

**Paralelo:** "A"

**Deber 1 Estadística**

**Federico Zertuche**

Importante: Escriban el código R para TODAS las preguntas de programación. Cuando terminen suban un pdf con las respuestas a la sección correspondiente del aula virtual.

**Ejercicio 1 Un poco de R.**

✓ **Pregunta 1**

**¿Cuáles de las siguientes expresiones valen 99 para  $x = 10$  en R?**

**Analicen la sintaxis como si estuvieran programando.**

1.  $10 * x - 1$   
R: 99
2.  $(x)(x) - 1$   
R: 99
3.  $\text{abs}(x * x) - \text{abs}(9 - x)$   
R: 99
4.  $11 * x - x + 1$   
R: 101

**Las tres primeras formas de respuestas están bien escritas la numero 4 nos da como resultado 101 y no concuerda con el resultado.**

✓ **Pregunta 2:**

**Un vector contiene una serie de ganancias ordenadas de manera creciente. Escriban el código que genera:**

**`x <- c(20, 80, 310, 655, 1962)`**

- La suma de todas las ganancias.  
`sum(g)`
- La segunda ganancia más grande.  
`x <- length(g)`  
`g[x-1]`
- La diferencia más grande entre las ganancias.  
`d <- diff(g)`  
`max(d)`
- Un booleano que responda a la pregunta:
- ¿La más grande diferencia ente dos ganancias es mayor a 10?  
`t <- max(d)`  
`nchar(t) < 10`  
**Argumento:** La respuesta es verdadera porque en el booleano se cuentan los dígitos de la más grande diferencia que en mi caso es 655 y serían 3 dígitos así que `3 < 10` es verdadero.
- La menor diferencia positiva entre dos ganancias.  
`d <- diff(g)`  
`min(d)`
- El máximo número de ganancias que pueden sumar sin pasar de 10000.  
`cumsum(g)<10000`

## Ejercicio 2 Dplyr en los Aeorpuertos

**Pregunta 3** Instalen la librería `nycflights13`. Escriban el código para encontrar todos los vuelos que:

- Fueron de SFO(San Francisco) hasta OAK(Oakland).  
`flights %>% filter(origin=='SFO', dest=='OAK') %>%`  
`View()`
- Salieron en Enero.  
`flights %>% filter(month=='1') %>%`  
`View()`
- Tienen demoras de mas de una hora (las demoras están en minutos).  
`flights %>%`  
`select(year, month, day,dep_delay )%>%`  
`filter(dep_delay>60)`  
`View()`
- Salieron entre medianoche y las 5 a.m.  
`flights %>% filter(hour==24, hour==5) %>%`

View()

- Tuvieron una demora de llegada 2 veces mas grande que la de salida.

```
flights %>% filter(arr_delay==2*dep_delay) %>%
```

View()

#### Pregunta 4

Lean la ayuda de select(). Escriban 2 formas de seleccionar las dos variables de retraso.

- `flights %>% select(dep_delay)`
- `flights %>% select(arr_delay, -starts_with("-"))`

#### Pregunta 5

Ordenen la tabla por fecha de salida y tiempo. ¿Cuáles fueron los vuelos que sufrieron las mayores demoras? ¿Cuáles recuperaron la mayor cantidad de tiempo durante el vuelo?

```
flights %>% arrange(dep_time, year, month, day) %>%
```

View()

#### Pregunta 6

Calculen la velocidad en mph usando el tiempo (que está en minutos) y la distancia (que está en millas). ¿Cuál fue el avión que voló más rápido?

```
velocidad <- flights %>% select(air_time, distance) %>%
```

```
  mutate(tiempo.horas = (air_time/60),
```

```
         distancia.metros = distance*1609.34)
```

```
velocidad %>% filter(tiempo.horas, distancia.metros) %>%
```

```
  select(tiempo.horas, distancia.metros) %>%
```

```
  mutate(velocidad.mph = (distancia.metros / tiempo.horas)) %>% View()
```

#### Pregunta 7

En dplyr el comando pipeline%>% se lee entonces. Significa:

$x \%>\% f(y) \rightarrow f(x,y)$ .

Es decir pasa x como primer argumento de f. Qué significan las siguientes líneas de código:

```
flights %>% filter(! is.na(dep_delay)) %>%
```

```
group_by(date, hour) %>%
```

```
summarise(delay = mean(dep_delay), n = n()) %>%
```

```
filter(n > 10)
```

- `flights %>% filter(! is.na(dep_delay))`  
Este filtra los valores de elementos en los retrasos de salida que se están perdiendo.
- `group_by(date, hour)`  
Indica los valores agrupados entre la fecha y la hora.
- `summarise(delay=mean(dep_delay, n=n()))`  
Hace un resumen del promedio de los retrasos.
- `filter(n > 10)`  
Se produjo Error in `n > 10` : comparison (6) is possible only for atomic and list types

### OTRA RESPUESTA

```
install.packages("magrittr")
library(magrittr)
flights %>%
  filter(!is.na(dep_delay))%>%
  group_by(year, month, day, hour)%>%
  summarise(delay = mean(dep_delay), n=n())%>%
  filter(n>10)
# filtra los retrasos diferentes de nulo agrupados por año,
# mes, día, hora y me devuelve el retraso de salida promedio cuyo
# valor sea mayor que 10.
```

### Pregunta 8

¿Cuál es la destinación que tiene las demoras promedio más grandes? ¿Cuántos vuelos diarios hay? ¿Cuál es la mejor hora para viajar sin retraso?

- `flights %>% group_by(dest, arr_delay) %>%`  
`summarize(mean(arr_delay, na.rm=TRUE)) %>%`  
`View()`