

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA
CARRERA DE INGENIERÍA MECÁNICA

ESTADISTICA TERCERO A
DEBER I
LUIS GUANANGA

EJERCICIO 1

Pregunta 1 Cuáles de las siguientes expresiones valen 99 para $x = 10$ en R?

Analicen la sintaxis como si estuvieran programando.

- `10*x-1`
[1] 99
- `x*x-1`
[1] 99
- `abs(x*x)-abs(9-x)`
[1] 99
- `11*x-x+1`
[1] 101

Pregunta 2 Un vector contiene una serie de ganancias ordenadas de manera creciente. Escriban el código que genera:

`x=c(10,20,30,40,50,60,70,80,90,110)`

- `sum([1seq(1,length(x),2)])`

```
x[c(9)]  
x[c(10)- c(1)]  
cummin(x)>10  
x[c(2)- c(1)]  
cummin(x) <100
```

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA
CARRERA DE INGENIERÍA MECÁNICA

EJERCICIO 2 Vamos a estudiar los datos de los vuelos locales en Estados Unidos durante el 2011. Usen los verbos:

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

- `flights %>% filter(origin=='SFO', dest=='OAK') %>%
View()`
- `flights %>% filter(month=='1') %>%
View()`
- `flights %>% filter(dep_delay>=1) %>%
View()`
- `flights %>% filter(hour==24, hour==5) %>%
View()`
- `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()`

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA
CARRERA DE INGENIERÍA MECÁNICA

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?

- `flights %>% filter(air_time, distance) %>%
mutate(Velocidad=air_time*distance, air_time=air_time/60) %>%
View()`

Pregunta 7 En dplyr el comando `pipeline()` se lee entonces.

Significa:

$x \%>\% f(y) \rightarrow f(x; y):$

- `flights %>% filter(! is.na(dep_delay))` 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)` En esta función apareció error. **Error in n > 10 : comparison (6) is possible only for atomic and list types**

**Pregunta 8Cuál 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()`

UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA CIVIL Y MECÁNICA
CARRERA DE INGENIERÍA MECÁNICA

EJERCICIO 3

Pregunta 9 Cuántas personas han habitado Sur América? ¿Hay alguna forma de usar los datos para saber cuál es la región que tiene las poblaciones más antiguas? Pueden usar los datos para tratar de entender si hay un patrón migratorio a lo largo de la historia