Práctica 2: Booking Data Cleaning

Gerard Alcalde and Guillem Rochina

2022-12-24

## 1. DESCRIPCIÓN DEL DATASET

### ¿Por qué es importante y qué pregunta/problema pretende responder?

#### 1.1. Descripción general

El conjunto de datos ha tratar es el obtenido como resultado de la práctica anterior de Web Scraping. El dataset presenta los datos e indicadores más relevantes de cada uno de los hoteles encontrados en función de determinados criterios de búsqueda (ciudad, fecha de check-in, fecha de check-out, número de adultos, niños y habitaciones). En este dataset en concreto presentamos datos para distintas fechas (diciembre, marzo y junio), ciudades (barcelona, madrid, valencia), número de adultos, niños y habitaciones a fin de tener una muestra más amplia e informativa que la que nos daría una búsqueda con tan sólo unos parámetros fijos.

A continuación, se realiza una breve descripción de cada una de las variables disponibles antes de realizar la limpieza:

* **Name (str)**: Nombre del hotel.
* **City (str)**: Nombre de la ciudad donde se realiza la búsqueda.
* **Check-in (str)**: Fecha de entrada al hotel.
* **Check-out (str)**: Fecha de salida del hotel.
* **Adults (int)**: Nº de adultos para los que se realiza la reserva.
* **Num\_rooms (int)**: Nº de habitaciones reservadas.
* **Address (str)**: Dirección postal del hotel, nombre de calle, barrio de la ciudad, código postal, etc.
* **Hotel\_coordinates (str)**: Latitud y longitud de la ubicación del hotel.
* **Hotel\_score (int)**: Nota general que recibe el hotel por los usuarios.
* **Hotel\_scores (dict)**: Puntuaciones de cada una de las dimensiones que puede valorar un usuario.
* **Hotel description (list)**: Descripción aportada por el propietario del hotel
* **Features (list)**: Lista de servicios añadidos del hotel.
* **Room\_data (dict)**: Conjunto de diccionarios que recogen las características de cada una de las habitaciones disponibles en el hotel.
* **Page\_count (int)**: Posición en la que ha aparecido el hotel en el buscador.
* **Current\_page (int)**: Página en la que ha aparecido el hotel en el buscador.

#### 1.2. Importancia y problema a responder

Aparecer en los primeros resultados de una búsqueda de un usuario por internet se ha convertido en los últimos años en una de las preocupaciones principales de cualquier negocio que busca hacerse un hueco en el mundo digital. Este fenómeno explica la proliferación de herramientas como el SEO y el SEM, mediante las cuales se intenta innovar (más allá del pago de tarifas para aparecer en las primeras páginas de búsqueda) a fin de lograr un mejor “posicionamiento” en buscadores como Google o Bing. El negocio hotelero no es una excepción en este caso. No obstante, el uso de dichas herramientas no resulta útil si nos enfrentamos a un buscador dentro de una página web, como el caso que nos ocupa con Booking.com. ¿Qué hacer entonces? ¿Cómo mejorar el posicionamiento dentro de la web? Asimismo, la fijación de precios o “pricing” puede ser una tardea ardua en los últimos años, pues la aparición de apartamentos vacacionales a través de plataformas como AirBnB ha incrementado de manera significativa la competencia a la que se deben enfrentar los propietarios de los hoteles, lo que influye en última instancia los precios a establecer por sus servicios. Dicho esto, ¿Qué precios se deben fijar dadas las características de un hotel? ¿Añadir un servicio adicional permitiría a un hotel elevar los precios de sus habitaciones?. Ante las preguntas presentadas en los párrafos anteriores, el objetivo de este proyecto es dual. Por un lado se busca obtener las características que hacen que un hotel esté mejor posicionado que otro en el metabuscador de Booking.com y, por el otro, se persigue desarrollar un modelo que indique qué precios se deberían establecer dadas las características de una habitación y qué elementos pueden incrementar el precio de forma relevante. En conclusión, el presente proyecto pretende dar continuidad a la Práctica 1, explotando los datos que se obtuvieron mediante web scrapping para ayudar a superar dos de los desafíos más relevantes para el sector hotelero.

## 2. LIMPIEZA DE LOS DATOS

A fin de iniciar el proceso de limpieza de datos, en primer lugar, se debe realizar la lectura del fichero csv que obtuvimos en la práctica anterior. Como resultado de la función de R-base “read.csv” obtenemos un objeto data.frame, que manipularemos en la presente sección.

# Lectura y breve descripción de los datos  
  
booking <- read.csv("hotels\_data.csv", header = TRUE)  
summary(booking)

## name city check.in adults   
## Length:3912 Length:3912 Length:3912 Min. :2.000   
## Class :character Class :character Class :character 1st Qu.:2.000   
## Mode :character Mode :character Mode :character Median :4.000   
## Mean :3.329   
## 3rd Qu.:5.000   
## Max. :5.000   
##   
## children check.out num\_rooms address   
## Min. :1.000 Length:3912 Min. :2.000 Length:3912   
## 1st Qu.:1.000 Class :character 1st Qu.:2.000 Class :character   
## Median :3.000 Mode :character Median :2.000 Mode :character   
## Mean :2.475 Mean :2.268   
## 3rd Qu.:3.000 3rd Qu.:3.000   
## Max. :3.000 Max. :3.000   
##   
## hotel\_coordinates hotel\_score hotel\_scores hotel\_description   
## Length:3912 Min. :-1.000 Length:3912 Length:3912   
## Class :character 1st Qu.: 8.000 Class :character Class :character   
## Mode :character Median : 8.400 Mode :character Mode :character   
## Mean : 8.259   
## 3rd Qu.: 8.800   
## Max. :10.000   
## NA's :8   
## features room\_data page\_count current\_page   
## Length:3912 Length:3912 Min. : 0.00 Min. :1.000   
## Class :character Class :character 1st Qu.: 24.00 1st Qu.:1.000   
## Mode :character Mode :character Median : 53.00 Median :3.000   
## Mean : 57.19 Mean :2.741   
## 3rd Qu.: 88.00 3rd Qu.:4.000   
## Max. :133.00 Max. :5.000   
##   
## in\_page\_count search\_date   
## Min. : 0.00 Length:3912   
## 1st Qu.: 6.00 Class :character   
## Median :13.00 Mode :character   
## Mean :12.66   
## 3rd Qu.:19.00   
## Max. :27.00   
##

Tanto con la función “summary” como con la función “class” nos muestran que a cada variable se le ha asignado o bien la clase de entero o la de string. En concreto, muchas de las variables interpretadas como string se tratan de diccionarios y listas, factor que deberemos tener en cuenta a la hora de trabajar con la limpieza de estas variables.

# Tipo de dato asignado a cada variable  
  
res <- sapply(booking, function(x) class(x))  
kable(data.frame(var=names(res), clase=as.vector(res)))

| var | clase |
| --- | --- |
| name | character |
| city | character |
| check.in | character |
| adults | integer |
| children | integer |
| check.out | character |
| num\_rooms | integer |
| address | character |
| hotel\_coordinates | character |
| hotel\_score | numeric |
| hotel\_scores | character |
| hotel\_description | character |
| features | character |
| room\_data | character |
| page\_count | integer |
| current\_page | integer |
| in\_page\_count | integer |
| search\_date | character |

#### 2.1. Transformación de las variables en el formato adecuado

Como se ha comentado, existen algunas variables que presentan todavía una estructura no válida para tratarlos estadísticamente o introducirlos en un modelo de data mining. En esta sección nos encargamos de tratar dichas columnas a fin de obtener nuevas variables con un formato adecuado, eliminando a su vez cualquier información que no nos sea de utilidad.

La columna hotel\_coordinates incluye tanto los valores de latitud como los de longitud, separados por una coma. Dado que ambos dos valores nos indican información distinta, pues la latitud nos proporcionar información de la posición en dirección norte o sur del ecuador y la longitud información de la posición en dirección este u oeste, en este caso hemos optado por separarlos en dos columnas distintas.

# Separamos los valores de la variable hotel coordinates en dos nuevas columnas, latitud y longitud  
# La columna hotel\_coordinates se elimina en el proceso  
  
booking <- booking %>%  
 separate(hotel\_coordinates, c("latitude", "longitude"), sep = ",", remove = FALSE)

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 16 rows [202,  
## 203, 223, 224, 225, 226, 227, 664, 665, 765, 1598, 2455, 2672, 2789, 2993,  
## 2999].

booking$latitude <- as.numeric(booking$latitude)  
booking$longitude <- as.numeric(booking$longitude)

Por su parte, “hotel\_scores” alberga diccionarios en cada uno de los registros. Cada diccionario consta del nombre de cada dimensión (caracterísitca a valorar) como clave y la valoración que recibe como valor. Dado que R lo ha interpretado como un string, nuestro enfoque en su limpieza se ha basado en separar la información con comas y, una vez obtenidas todas las columnas resultantes, extraer los datos numéricos de cada registro para quedarnos solo con los scores de cada dimensión.

# Definimos un vector con los nombres de las columnas a crear  
  
columns <- c("staff\_score", "facilities\_score", "cleanliness\_score", "comfort\_score", "value\_for\_money\_score", "location\_score", "free\_wifi\_score")  
  
# Separamos los strings (diccionarios) en función de las comas y creamos las columnas correspondientes. La primera columna creada es una NA porque todos los primeros registros en cada diccionario están vacíos (por un fallo en la extracción de datos). Estableciendo la columna como NA, indicamos a la función separate que ignore dichos datos y no cree columna alguna para ellos.  
  
booking <- booking %>%  
 separate(hotel\_scores, c(NA, "staff\_score", "facilities\_score", "cleanliness\_score", "comfort\_score", "value\_for\_money\_score", "location\_score", "free\_wifi\_score"), sep = ",", remove = FALSE)

## Warning: Expected 8 pieces. Missing pieces filled with `NA` in 423 rows [9, 80,  
## 105, 125, 141, 202, 203, 218, 223, 224, 225, 226, 227, 247, 303, 311, 322, 327,  
## 336, 370, ...].

# Extraemos los valores numéricos de cada columna con la función del paquete readr, "parse\_number"  
  
booking[columns] <- apply(booking[columns], 2, readr::parse\_number)

Dado que de las fechas tan sólo nos interesa el mes en el que se realiza la reserva (un alojamiento con unas características determinadas podría estar mejor posicionado en un determinado mes que en otro o tener un precio distinto) procedemos a extraer dicha información de las variables check-in y check-out. De hecho, dado que los datos que hemos extraído corresponden al mismo mes en todos los registros, tan sólo nos quedaremos con una columna de mes. Asimismo, no crearemos una columna is\_june, pues la dejaremos como caso base, es decir, cuando no sea ninguna de las otras dos opciones (o Diciembre o Marzo).

booking <- booking %>%  
 # Separamos el nombre del mes del resto de elementos del registro para check\_in y check\_out. Eliminamos las columnas en el proceso.  
 separate(check.in, c(NA, "month", NA), sep = "-", remove = FALSE) %>%  
 mutate(is\_december = ifelse(month == "December", 1, 0),  
 is\_march = ifelse(month == "March", 1, 0))

De la columna address tan sólo nos interesa el código postal (el barrio no se incluye en todos los registros a diferencia del CP, así que hemos optado por quedarnos con estos valores), por lo que extraemos dicha información para crear una nueva columna llamada postal\_code, eliminando la información redundante en el proceso.

# Hemos tenido que cambiar esta función para que el extract no diera problemas  
Sys.setlocale('LC\_ALL', 'C')

## [1] "C"

# Extraemos los valores de codigo postal de la columna addres utilizando expressiones regulares  
booking <- booking %>%  
 extract(address, c("postal\_code"), regex = "( [0-9]{5} )", remove = FALSE)

Los features, entendidos como servicios adicionales más allá de la propia habitación, fueron guardados como una lista. No obstante, de nuevo nos encontramos como R lo ha identificado como string. Es por ello por lo que realizamos un tratamiento similar al de la columna de scores. En este caso creamos tantas variables dummy (columnas dicotómicas con valores de 1, en caso de que el feature exista en dicho registro, o 0, en caso de que no exista) como servicios adicionales de hotel consideramos interesantes.

# Definimos las features que consideramos que son interesantes. Utilizamos este formato para filtrar por expresión regular.  
feature\_list <- c("('Free WiFi')", "('Air conditioning')", "('24-hour front desk')", "('Safe')", "('Heating')", "('Elevator')", "('Private Bathroom')", "('Non-smoking rooms')", "('Aparments')", "('City view')","('Kitchen')", "('Pet friendly')", "('Swimming pool')", "('Balcony')")  
  
# Creamos un bucle en el que se recorre cada elemento de la lista de features  
for (feature in feature\_list){  
 # Se eliminan los parentésis y comillas de la variable local feature para crear el nombre de la columna.  
 col\_name <- str\_replace(str\_replace(str\_to\_lower(str\_extract(feature, "([A-Z][a-z]\*( |-)?[A-Z]?[a-z]\* ? ?[A-Z]?[a-z]\*)")), " ", "\_"), "-", "\_")  
 booking <- booking %>%  
 # Extraemos el nombre del feature de cada uno de los registros (strings). En caso de que no encuentre ningún valor devuelve un NA.  
 extract(features, c(col\_name), regex = feature, remove = FALSE) %>%  
 # Transformamos la columna que acaba de ser creada para que indique con un 1 si el registro tenia dicho servicio y 0 si el valor era NA (no tenida dicho servicio)  
 mutate\_(.dots = setNames(list(paste0("as.integer(!is.na(",col\_name,"))")), col\_name))  
}

## Warning: `mutate\_()` was deprecated in dplyr 0.7.0.  
## i Please use `mutate()` instead.  
## i See vignette('programming') for more help  
## i The deprecated feature was likely used in the dplyr package.  
## Please report the issue at <]8;;https://github.com/tidyverse/dplyr/issueshttps://github.com/tidyverse/dplyr/issues]8;;>.

Por otro lado, nos encontramos con la columna que incluye la descripción del hotel, esta descripción es muy amplia y un análisis profundo de ella requeriría de técnicas de NLP que no son el objetivo de esta práctica. No obstante, consideramos que la longitud de la descripción del hotel sí que puede tener una relación con la posición en el buscador, por lo que transformaremos esta variable a una nueva variable que contenga el número total de palabras que contiene la descripción, eliminando de nuevo dicha columna en el proceso.

# Evaluamos la longitud de la descripción del hotel  
booking$length\_description <- lengths(gregexpr("\\W+", booking$hotel\_description)) + 1

Finalmente, en la columna Room\_data, se incluye mucha información referente a los tipos de habitaciones disponibles, precios, características, etc. Esta se había extraído de esta forma con el objetivo de procesarla mediante un diccionario de Python. No obstante, al haber optado por un procesado con R este se complicará un poco más. En esta columna tenemos distintos datos que pueden ser muy relevantes, de los cuales extraeremos los siguientes: • min\_price: Precio de la habitación más económica. • max\_price: Precio de las habitaciones más caras del hotel. • suite\_available: Una de las habitaciones es de tipo suite. • is\_apartment: Una de las habitaciones es de tipo apartamento. • free\_cancelation\_available: Alguna de las habitaciones tiene cancelación gratuita.

# Función que encuentra el precio mínimo de la habitación dado un string con la información de las habitaciones  
find\_min\_price <- function(text) {  
 if (text=="{}"){  
 return(NA)  
 }  
 a <- gregexpr("([0-9]\*', 'room\_capacity)", text)  
   
 min\_price <- NA  
 for (value in a[[1]]){  
 substring <- substr(text, value, value+20)  
 price <- as.numeric(strsplit(substring, "'")[[1]][1])  
 if (is.na(min\_price)) {  
 min\_price <- price  
 }  
   
 if (price < min\_price) {  
 min\_price <- price  
 }   
 }  
 if (min\_price == 0){  
 return(NA)  
 }  
 else {  
 return(min\_price)  
 }  
}  
  
# Aplicamos la función a todos los valores  
min\_price\_vector <- c()  
for (i in seq(1, length(booking$room\_data))) {  
 min\_price\_vector <- append(min\_price\_vector, find\_min\_price(booking$room\_data[i]))  
}  
  
# Guardamos los resultados en una nueva columna.  
booking$min\_price <- min\_price\_vector  
  
# Función que encuentra el precio máximo de la habitación dado un string con la información de las habitaciones  
find\_max\_price <- function(text) {  
 if (text=="{}"){  
 return(NA)  
 }  
 a <- gregexpr("([0-9]\*', 'room\_capacity)", text)  
   
 max\_price <- NA  
 for (value in a[[1]]){  
 substring <- substr(text, value, value+20)  
 price <- as.numeric(strsplit(substring, "'")[[1]][1])  
 if (is.na(max\_price)) {  
 max\_price <- price  
 }  
   
 if (price > max\_price) {  
 max\_price <- price  
 }   
 }  
 if (max\_price == 0){  
 return(NA)  
 }  
 else {  
 return(max\_price)  
 }  
}  
  
# Aplicamos la función a todos los valores  
max\_price\_vector <- c()  
for (i in seq(1, length(booking$room\_data))) {  
 max\_price\_vector <- append(max\_price\_vector, find\_max\_price(booking$room\_data[i]))  
}  
  
# Guardamos los resultados en una nueva columna.  
booking$max\_price <- max\_price\_vector  
  
# Buscamos si hay una habitación en suite  
is\_suite\_vector <- c()  
for (i in seq(1, length(booking$room\_data))) {  
 is\_suite <- gregexpr("(suite)", booking$room\_data[i])[[1]][1]  
   
 if (is\_suite==-1) {  
 is\_suite\_vector <- append(is\_suite\_vector, 0)  
 } else {  
 is\_suite\_vector <- append(is\_suite\_vector, 1)  
 }  
}  
  
# Guardamos los resultados en una nueva columna.  
booking$is\_suite <- is\_suite\_vector  
  
# Buscamos si hay opción de apartamento  
is\_apartment\_vector <- c()  
for (i in seq(1, length(booking$room\_data))) {  
 is\_apartment <- gregexpr("(Apartment)", booking$room\_data[i])[[1]][1]  
   
 if (is\_apartment==-1) {  
 is\_apartment\_vector <- append(is\_apartment\_vector, 0)  
 } else {  
 is\_apartment\_vector <- append(is\_apartment\_vector, 1)  
 }  
}  
  
# Guardamos los resultados en una nueva columna.  
booking$is\_apartment <- is\_apartment\_vector  
  
# Buscamos si tiene cancelación gratuita  
free\_cancelation\_vector <- c()  
for (i in seq(1, length(booking$room\_data))) {  
 free\_cancelation <- gregexpr("(Free cancellation)", booking$room\_data[i])[[1]][1]  
   
 if (free\_cancelation==-1) {  
 free\_cancelation\_vector <- append(free\_cancelation\_vector, 0)  
 } else {  
 free\_cancelation\_vector <- append(free\_cancelation\_vector, 1)  
 }  
}  
  
# Guardamos los resultados en una nueva columna.  
booking$has\_free\_cancelation <- free\_cancelation\_vector

#### 2.2. Ceros y elementos vacíos

Dando paso a el tratamiento de ceros y elementos vacios, lo primero que debemos hacer es identificar aquellas variables en las que encontramos valores faltantes. Como observamos, “postal\_code” (19), “longitude” (16), “hotel\_score” (8), min (45) y max\_price (17) y el resto de variables relacionadas con los scores (60) presentan valores faltantes, destacando free\_wifi\_score con 423 registros faltantes, consecuencia con toda seguridad de la ausencia de este servicio en dichos hoteles.

sapply(booking, function(y) sum(length(which(is.na(y)))))

## name city check.in   
## 0 0 0   
## month adults children   
## 0 0 0   
## check.out num\_rooms address   
## 0 0 0   
## postal\_code hotel\_coordinates latitude   
## 19 0 16   
## longitude hotel\_score hotel\_scores   
## 16 8 0   
## staff\_score facilities\_score cleanliness\_score   
## 60 60 60   
## comfort\_score value\_for\_money\_score location\_score   
## 60 60 60   
## free\_wifi\_score hotel\_description features   
## 423 0 0   
## balcony swimming\_pool pet\_friendly   
## 0 0 0   
## kitchen city\_view aparments   
## 0 0 0   
## non\_smoking\_rooms private\_bathroom elevator   
## 0 0 0   
## heating safe air\_conditioning   
## 0 0 0   
## free\_wifi room\_data page\_count   
## 0 0 0   
## current\_page in\_page\_count search\_date   
## 0 0 0   
## is\_december is\_march as.integer(!is.na(NA))   
## 0 0 0   
## length\_description min\_price max\_price   
## 0 45 17   
## is\_suite is\_apartment has\_free\_cancelation   
## 0 0 0

Por lo que respecta a la variable “postal\_code” encontramos que en todos aquellos registros donde se encuentan NAs, también hallamos datos faltantes en la mayoría del resto de columnas. Por tanto, dado que dichos registros no nos aportan información de valor, procedemos a eliminarlos del dataset.

# Comprobamos algunos de los registros donde hay datos faltantes de postal\_code  
booking[is.na(booking$postal\_code),] %>%  
 dplyr::select(c("city", "month", "postal\_code", "longitude", "hotel\_score", "staff\_score")) %>%  
 head(3)

## city month postal\_code longitude hotel\_score staff\_score  
## 202 Barcelona December <NA> NA -1 NA  
## 203 Barcelona December <NA> NA -1 NA  
## 223 Barcelona December <NA> NA -1 NA

# Filtramos los datos para eliminar aquellos registros que tienen NAs en postal\_code  
booking <- booking[!is.na(booking$postal\_code),]

Si comprobamos de nuevo que columnas presentan todavía datos faltantes, observamos que en algunos casos (como “longitude”) hemos logrado eliminar todos los datos faltantes, mientras que en el resto se han reducido en 16 unidades el número de registros con NAs, pues como se había comentado todos aquellos registros con ausencia de postal\_code, también carecía del resto de datos.

sapply(booking, function(y) sum(is.na(y)))

## name city check.in   
## 0 0 0   
## month adults children   
## 0 0 0   
## check.out num\_rooms address   
## 0 0 0   
## postal\_code hotel\_coordinates latitude   
## 0 0 0   
## longitude hotel\_score hotel\_scores   
## 0 5 0   
## staff\_score facilities\_score cleanliness\_score   
## 44 44 44   
## comfort\_score value\_for\_money\_score location\_score   
## 44 44 44   
## free\_wifi\_score hotel\_description features   
## 407 0 0   
## balcony swimming\_pool pet\_friendly   
## 0 0 0   
## kitchen city\_view aparments   
## 0 0 0   
## non\_smoking\_rooms private\_bathroom elevator   
## 0 0 0   
## heating safe air\_conditioning   
## 0 0 0   
## free\_wifi room\_data page\_count   
## 0 0 0   
## current\_page in\_page\_count search\_date   
## 0 0 0   
## is\_december is\_march as.integer(!is.na(NA))   
## 0 0 0   
## length\_description min\_price max\_price   
## 0 29 1   
## is\_suite is\_apartment has\_free\_cancelation   
## 0 0 0

A continuación, analizamos los datos faltantes de hotel\_score. En este caso, encontramos que la ausencia de datos se registra de dos formas distintas, o bien con la introducción de NAs o bien con el valor -1. Por tanto, dado que el tratamiento en ambos casos va a ser el mismo, transformamos todos los valores -1 en NAs para poder llevar a cabo el siguiente proceso de inputación.

# Comprobamos algunos de los registros donde hay datos faltantes de hotel\_score  
booking[is.na(booking$hotel\_score),] %>%  
 dplyr::select(c("city", "month", "hotel\_score", "staff\_score", "location\_score", "longitude", "facilities\_score", "cleanliness\_score", "comfort\_score")) %>%  
 head(5)

## city month hotel\_score staff\_score location\_score longitude  
## 310 Barcelona December NA 9.4 9.4 2.1787906  
## 844 Valencia December NA 7.7 9.5 -0.3775288  
## 958 Valencia December NA 9.2 9.3 -0.3247243  
## 1210 Barcelona March NA 8.1 8.6 2.1429321  
## 2708 Barcelona June NA 8.1 8.6 2.1429321  
## facilities\_score cleanliness\_score comfort\_score  
## 310 8.6 8.9 8.9  
## 844 7.2 7.5 7.6  
## 958 9.1 9.3 9.3  
## 1210 7.3 7.7 7.7  
## 2708 7.3 7.7 7.7

# Comprobamos algunos de los registros donde hay datos faltantes de hotel\_score  
booking[booking$hotel\_score == "-1",] %>%  
 dplyr::select(c("city", "month", "hotel\_score", "staff\_score", "location\_score", "longitude", "facilities\_score", "cleanliness\_score", "comfort\_score")) %>%  
 tail(5)

## city month hotel\_score staff\_score location\_score longitude  
## 3690 Madrid June -1 NA NA -3.699663  
## 3736 Madrid March -1 NA NA -3.709937  
## 3855 Madrid March -1 NA NA -3.700559  
## 3877 Madrid March -1 NA NA -3.699663  
## 3909 Madrid March -1 NA NA -3.699676  
## facilities\_score cleanliness\_score comfort\_score  
## 3690 NA NA NA  
## 3736 NA NA NA  
## 3855 NA NA NA  
## 3877 NA NA NA  
## 3909 NA NA NA

booking[booking$hotel\_score == "-1",] %>%  
 dplyr::select(c("city", "month", "hotel\_score", "staff\_score", "free\_wifi\_score", "free\_wifi")) %>%  
 tail(10)

## city month hotel\_score staff\_score free\_wifi\_score free\_wifi  
## 3511 Madrid June -1 NA NA 1  
## 3544 Madrid June -1 NA NA 1  
## 3625 Madrid June -1 NA NA 1  
## 3667 Madrid June -1 NA NA 1  
## 3668 Madrid June -1 NA NA 1  
## 3690 Madrid June -1 NA NA 1  
## 3736 Madrid March -1 NA NA 0  
## 3855 Madrid March -1 NA NA 1  
## 3877 Madrid March -1 NA NA 1  
## 3909 Madrid March -1 NA NA 1

# Tranformamos los valores -1 en NAs para poder llevar a cabo el proceso de inputación correctamente.  
booking$hotel\_score <- ifelse(booking$hotel\_score == "-1", NA, booking$hotel\_score)

Dado que en este caso tan sólo encontramos datos faltantes en las columnas de score no consideramos eliminar estos registros, sino llevar a cabo una imputación. Aunque se podría haber utilizado una metología más simple como el uso de la media o la mediana, hemos optado por hacer uso de un método más complejo, basado en la similitud entre registros, en concreto llevar a cabo una imputación mediante el algoritmo de K vecinos más próximos. En este caso, la función se encarga de devolver los mismos valores en caso de que ya existan en el dataset y el valor de los registros más “similares” en el caso de un registro con NAs.

# Imputamos los datos con K Neirest Neighbours  
booking$hotel\_score <- kNN(booking)$hotel\_score  
booking$staff\_score <- kNN(booking)$staff\_score   
booking$facilities\_score <- kNN(booking)$facilities\_score  
booking$cleanliness\_score <- kNN(booking)$cleanliness\_score  
booking$comfort\_score <- kNN(booking)$comfort\_score  
booking$value\_for\_money\_score <- kNN(booking)$value\_for\_money\_score  
booking$location\_score <- kNN(booking)$location\_score  
booking$free\_wifi\_score <- kNN(booking)$free\_wifi\_score

En el caso de los precios, encontramos NAs en aquellos registros que presentaban valores iguales a 0 (no son valores lógicos). En este caso también encontramos que los únicos valores faltantes en estos registros corresponde a precio, por lo que procedemos de la misma forma que con los scores.

booking[is.na(booking$min\_price),] %>%  
 dplyr::select(c("city", "month", "hotel\_score", "max\_price", "min\_price")) %>%  
 tail(10)

## city month hotel\_score max\_price min\_price  
## 2767 Barcelona June 8.8 992 NA  
## 3219 Valencia March 9.6 900 NA  
## 3369 Valencia March 8.7 992 NA  
## 3467 Madrid June 8.9 988 NA  
## 3558 Madrid June 8.0 943 NA  
## 3620 Madrid June 9.5 903 NA  
## 3698 Madrid June 8.2 964 NA  
## 3846 Madrid March 8.8 981 NA  
## 3847 Madrid March 8.6 960 NA  
## 3849 Madrid March 8.9 988 NA

# Imputamos los datos con K Neirest Neighbours  
booking$max\_price <- kNN(booking)$max\_price  
booking$min\_price <- kNN(booking)$min\_price

Una última comprobación nos revela que el dataset ya no cuenta con datos faltantes, con independencia del formato que pudieran presenta (strings vacías, valores de -1 en integers, o simplemente NAs).

sapply(booking, function(y) sum(is.na(y)))

## name city check.in   
## 0 0 0   
## month adults children   
## 0 0 0   
## check.out num\_rooms address   
## 0 0 0   
## postal\_code hotel\_coordinates latitude   
## 0 0 0   
## longitude hotel\_score hotel\_scores   
## 0 0 0   
## staff\_score facilities\_score cleanliness\_score   
## 0 0 0   
## comfort\_score value\_for\_money\_score location\_score   
## 0 0 0   
## free\_wifi\_score hotel\_description features   
## 0 0 0   
## balcony swimming\_pool pet\_friendly   
## 0 0 0   
## kitchen city\_view aparments   
## 0 0 0   
## non\_smoking\_rooms private\_bathroom elevator   
## 0 0 0   
## heating safe air\_conditioning   
## 0 0 0   
## free\_wifi room\_data page\_count   
## 0 0 0   
## current\_page in\_page\_count search\_date   
## 0 0 0   
## is\_december is\_march as.integer(!is.na(NA))   
## 0 0 0   
## length\_description min\_price max\_price   
## 0 0 0   
## is\_suite is\_apartment has\_free\_cancelation   
## 0 0 0

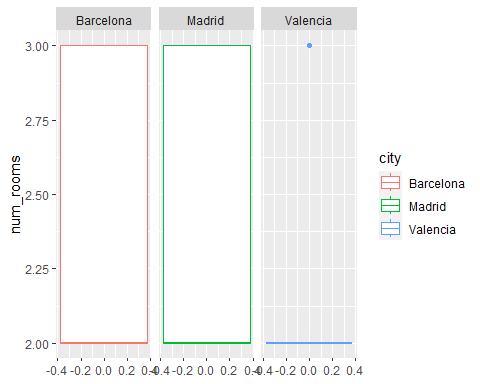
sapply(booking, function(y) sum(y == ""))

## name city check.in   
## 0 0 0   
## month adults children   
## 0 0 0   
## check.out num\_rooms address   
## 0 0 0   
## postal\_code hotel\_coordinates latitude   
## 0 0 0   
## longitude hotel\_score hotel\_scores   
## 0 0 0   
## staff\_score facilities\_score cleanliness\_score   
## 0 0 0   
## comfort\_score value\_for\_money\_score location\_score   
## 0 0 0   
## free\_wifi\_score hotel\_description features   
## 0 0 0   
## balcony swimming\_pool pet\_friendly   
## 0 0 0   
## kitchen city\_view aparments   
## 0 0 0   
## non\_smoking\_rooms private\_bathroom elevator   
## 0 0 0   
## heating safe air\_conditioning   
## 0 0 0   
## free\_wifi room\_data page\_count   
## 0 0 0   
## current\_page in\_page\_count search\_date   
## 0 0 0   
## is\_december is\_march as.integer(!is.na(NA))   
## 0 0 0   
## length\_description min\_price max\_price   
## 0 0 0   
## is\_suite is\_apartment has\_free\_cancelation   
## 0 0 0

#### 2.3. Valores extremos

Los outliers son aquellos datos extremos que dada su distancia con respecto al grueso de la distribución a priori resultan no ser congruentes si los comparamos con la población/muestra analizada. Existen varias vías para detectarlos, no obstante el método más rápdio y común se basa en el uso de boxplots, donde se detectan aquellos valores que se sitúan 1.5 veces más allá del rango intercuartílico con respecto a la mediana. Asimismo, en este trabajo también haremos uso de la función “boxplots.stats()” a fin de detectar, con un output más allá del visual, los outliers que encontramos usando boxplots. Dado que muchas de las variables que hemos generado en la fase de limpieza son dicotómicas, tan solo nos encargaremos de comprobar si existen valores extremos en las columnas de **num\_rooms**, **latitude**, **longitude**, **hotel\_score** y el resto de **scores del hotel**, **length\_description**, **min\_price** y **max\_price**.

ggplot(booking, aes(y = num\_rooms, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



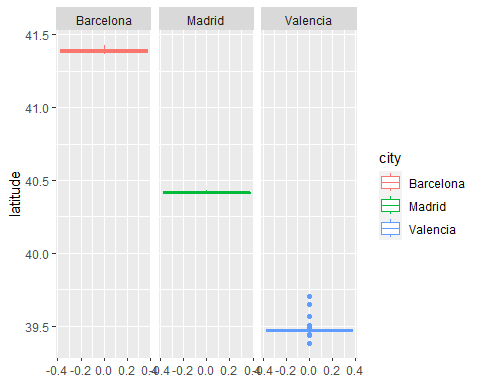
boxplot.stats(booking$num\_rooms)$out

## integer(0)

Empezamos analizando el número de habitaciones y observamos que en las 3 ciudades el número oscila entre 2 y 3 habitaciones para las búsquedas realizadas, lo cual son números razonables, aunque en Valencia esta aparezca como un outlier. Por lo tanto, aceptamos los valores outliers observados en la ciudad de Valencia como valores correctos.

A continuación se analiza la ubicación de los hoteles mediante su longitud y latitud.

ggplot(booking, aes(y = latitude, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



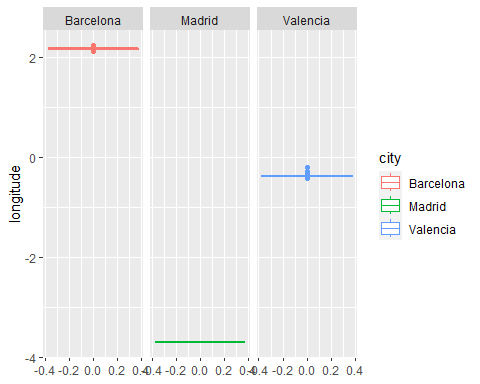
boxplot.stats(booking$latitude)$out

## numeric(0)

Observamos que la latitud contiene valores distintos para cada ciudad, como es de esperar al tener cada una una ubicación distinta. En el caso de Valencia vemos una mayor dispersión y tenemos la presencia de algunos outliers, esto es debido a que probablemente, al ser una ciudad más pequeña que Madrid o Barcelona tenga una menor disponibilidad hotelera y algunos de los hoteles hayados no pertenezcan a Valencia ciudad sino a los alrededores de esta.

A continuación evaluaremos si obtenemos un comportamiento similar con la longitud.

ggplot(booking, aes(y = longitude, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



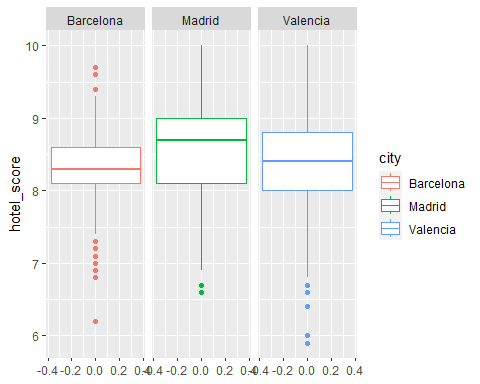
boxplot.stats(booking$longitude)$out

## numeric(0)

Para la longitud obtenemos unos resultados parecidos en los que hay una menor dispersión en los outliers de Valencia. Dado que estas variaciones tan pequeñas que se observan en los resultados son razonables aceptamos los resultados como validos.

A continuación se evalúan las distintas puntuaciones del hotel.

ggplot(booking, aes(y = hotel\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$hotel\_score)$out

## [1] 6.9 6.9 6.9 6.9 6.9 10.0 6.6 6.6 5.9 7.0 6.6 6.0 6.6 5.9 6.7  
## [16] 10.0 5.9 7.0 6.8 6.6 6.7 6.9 6.0 6.9 6.4 10.0 6.6 6.6 10.0 7.0  
## [31] 6.9 6.9 6.9 10.0 7.0 6.6 7.0 6.6 6.7 7.0 6.4 7.0 6.9 6.8 10.0  
## [46] 6.9 6.9 7.0 6.2 6.8 7.0 6.6 6.7 7.0 6.9 6.9 6.9 6.9 7.0 7.0  
## [61] 6.9 6.9 7.0 5.9 6.0 7.0 6.9 7.0 6.8 6.8 5.9 7.0 10.0 6.6 6.6  
## [76] 6.6 7.0 6.9 10.0 6.6 10.0 10.0 6.6

Por un lado observamos que la puntuación general del hotel oscila entre el 8 y el 9 siendo mayor en Madrid que en Barcelona o Valencia. También nos encontramos que hay outliers en las 3 ciudades no obstante todos se hayan en una puntuación entre el 5 y el 10, cuando el rango factible es entre 0 y 10, por lo que pueden ser valores anómalos de hoteles que han obtenido una puntuación más baja de lo normal, pero no por ello deben ser descartados.

A continuación, evaluamos aquellos hoteles que han obtenido una nota o muy baja o excepcionalmente buena para verificar que los resultados sean coherentes.

booking %>%  
 filter(hotel\_score <= 7.0) %>%  
 dplyr::select(c("hotel\_score", "facilities\_score", "cleanliness\_score", "location\_score")) %>%  
 sample\_n(5)

## hotel\_score facilities\_score cleanliness\_score location\_score  
## 1 6.8 6.7 7.1 8.5  
## 2 6.9 7.1 7.4 6.6  
## 3 6.9 6.8 7.5 8.2  
## 4 6.9 6.9 7.6 8.1  
## 5 6.9 7.0 7.4 8.0

Se observa que los hoteles con una mala puntuación lo hacen en la mayoría de sus características, lo que corrobora los resultados obtenidos y no es una sección que ha penalizado a los resultados obtenidos.

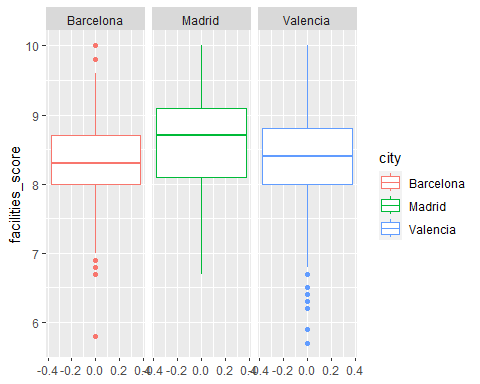
booking %>%  
 filter(hotel\_score == 10.0) %>%  
 sample\_n(3)

## name city check.in  
## 1 Alameda Velvet, loft en ciudad de las ciencias Valencia 12-June-2023  
## 2 CASA ESPECTACULAR DE DISE\357\277\275O CON JARDIN INTERIOR Valencia 12-December-2022  
## 3 Luxury Madrid Centro Madrid 12-December-2022  
## month adults children check.out num\_rooms  
## 1 June 4 1 16-June-2023 2  
## 2 December 2 3 16-December-2022 2  
## 3 December 5 3 16-December-2022 3  
## address  
## 1 38 Passeig de l'Albereda, Camins al Grau, 46023 Valencia, Spain  
## 2 Carrer Isabel la Catolica, 12 (Mislata) VALENCIA, 46920 Valencia, Spain  
## 3 44 Carrera de San Jer\357\277\275nimo, Madrid City Center, 28014 Madrid, Spain  
## postal\_code hotel\_coordinates latitude longitude hotel\_score  
## 1 46023 39.46341780,-0.35751330 39.46342 -0.3575133 10  
## 2 46920 39.47250590,-0.41545640 39.47251 -0.4154564 10  
## 3 28014 40.41611950,-3.69754890 40.41612 -3.6975489 10  
## hotel\_scores  
## 1 {'': '', 'Staff': '10', 'Facilities': '10', 'Cleanliness': '10', 'Comfort': '10', 'Value for money': '10', 'Location': '10'}  
## 2 {'': '', 'Staff': '10', 'Facilities': '10', 'Cleanliness': '10', 'Comfort': '10', 'Value for money': '10', 'Location': '7.5'}  
## 3 {'': '', 'Staff': '10', 'Facilities': '10', 'Cleanliness': '10', 'Comfort': '10', 'Value for money': '10', 'Location': '10'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 10 10 10 10  
## 2 10 10 10 10  
## 3 10 10 10 10  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 10 10.0 8.3  
## 2 10 7.5 10.0  
## 3 10 10.0 9.0  
## hotel\_description  
## 1 ['Located in Valencia, a 16-minute walk from Turia Gardens and 1.9 km from Jardines de Monforte, Alameda Velvet, loft en ciudad de las ciencias offers free WiFi and air conditioning. Offering private parking, the apartment is 2.6 km from Valencia North Station.', 'The apartment is equipped with 1 bedroom, a flat-screen TV and a fully equipped kitchen that provides guests with a dishwasher, a microwave, a washing machine, a fridge and an oven. Towels and bed linen are available.', 'At Alameda Velvet, loft en ciudad de las ciencias guests are welcome to use a hot tub.', 'Oceanografic is 1.9 km from the accommodation, while Port of Valencia is 3.1 km from the property. The nearest airport is Valencia Airport, 13 km from Alameda Velvet, loft en ciudad de las ciencias.']  
## 2 ['Located in Valencia, one kilometer from Bioparc Valencia and 3.5 km from Church of Saint Nicol\357\277\275s, CASA ESPECTACULAR DE DISE\357\277\275O CON JARDIN INTERIOR provides accommodations with free WiFi, air conditioning, a garden and a terrace. Providing private parking, the vacation home is 4.8 km from Valencia North Station.', 'The vacation home has 2 bedrooms, 2 bathrooms, bed linen, towels, a flat-screen TV, a fully equipped kitchen, and a balcony with garden views.', 'Basilica de la Virgen de los Desamparados is 5.1 km from CASA ESPECTACULAR DE DISE\357\277\275O CON JARDIN INTERIOR, while Jardines de Monforte is 5.1 km away. The nearest airport is Valencia Airport, 6.9 km from the accommodation.']  
## 3 ['In the Madrid City Center district of Madrid, close to Retiro Park, Luxury Madrid Centro has free WiFi and a washing machine. The property is a 4-minute walk from Thyssen-Bornemisza Museum and half a kilometer from Gran Via.', 'The air-conditioned apartment is composed of 2 separate bedrooms, a living room, a fully equipped kitchen with a dishwasher and microwave, and 2 bathrooms. Towels and bed linen are offered in this accommodation.', 'Popular points of interest near the apartment include Gran Via Station Metro Station, Reina Sofia Museum and Atocha Train Station. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 16.1 km from Luxury Madrid Centro.']  
## features  
## 1 ['Whole apartment', 'Size', 'Kitchen', 'Pet friendly', 'Washing machine', 'Free WiFi', 'Bathtub', 'Air conditioning']  
## 2 ['Whole house', 'Size', 'Kitchen', 'Garden', 'Free WiFi', 'Terrace', 'Balcony', 'Air conditioning', 'Private Bathroom', 'Heating']  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'View', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Elevator']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 0 0 1 1 0 0  
## 2 1 0 0 1 0 0  
## 3 0 0 0 1 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 0 0 0 0 0 1  
## 2 0 1 0 1 0 1  
## 3 0 1 1 0 0 1  
## free\_wifi  
## 1 1  
## 2 1  
## 3 1  
## room\_data  
## 1 {'909863001\_362167438\_4\_0\_0': {'room\_name': 'One-Bedroom Apartment', 'room\_price': '\357\277\275 593', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on June 10, 2023', 'Pay nothing until June 8, 2023'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 2/ '}}  
## 2 {'888411801\_358874169\_5\_0\_0': {'room\_name': 'Two-Bedroom House', 'room\_price': '\357\277\275 630', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on December 7, 2022', 'Pay nothing until December 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire vacation home', '120 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Garden view', 'Air conditioning', 'Flat-screen TV', 'Terrace'], 'room\_bed\_type': '1 full bed/ 1 bunk bed/ 1 sofa bed/ 2/ '}, '888411801\_358874169\_3\_0\_0': {'room\_name': 'Two-Bedroom House', 'room\_price': '\357\277\275 573', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on December 7, 2022', 'Pay nothing until December 5, 2022'], 'room\_facilities': ['Entire vacation home', '120 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Garden view', 'Air conditioning', 'Flat-screen TV', 'Terrace'], 'room\_bed\_type': '1 full bed/ 1 bunk bed/ 1 sofa bed/ 2/ '}, '888411801\_358874169\_1\_0\_0': {'room\_name': 'Two-Bedroom House', 'room\_price': '\357\277\275 516', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Free cancellation until 11:59 PM on December 7, 2022', 'Pay nothing until December 5, 2022'], 'room\_facilities': ['Entire vacation home', '120 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Garden view', 'Air conditioning', 'Flat-screen TV', 'Terrace'], 'room\_bed\_type': '1 full bed/ 1 bunk bed/ 1 sofa bed/ 2/ '}}  
## 3 {'907164701\_361787894\_5\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 947', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on December 8, 2022', 'Pay nothing until December 6, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_364062927\_5\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 861', 'room\_capacity': 'Max people: 5', 'room\_options': ['Partially refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_364062927\_6\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 964', 'room\_capacity': 'Max people: 6', 'room\_options': ['Partially refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_361787894\_6\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 1,062', 'room\_capacity': 'Max people: 6', 'room\_options': ['Free cancellation until 11:59 PM on December 8, 2022', 'Pay nothing until December 6, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_364062927\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 773', 'room\_capacity': 'Max people: 4', 'room\_options': ['Partially refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_361787894\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 850', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on December 8, 2022', 'Pay nothing until December 6, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_364062927\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 707', 'room\_capacity': 'Max people: 3', 'room\_options': ['Partially refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_361787894\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 776', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on December 8, 2022', 'Pay nothing until December 6, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_364062927\_2\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 641', 'room\_capacity': 'Max people: 2', 'room\_options': ['Partially refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}, '907164701\_361787894\_2\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 705', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on December 8, 2022', 'Pay nothing until December 6, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', '83 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'View', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 sofa bed/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 94 4 19 2022-11-19 0 0  
## 2 123 5 18 2022-11-18 1 0  
## 3 36 2 11 2022-11-18 1 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 138 593 593 0  
## 2 0 120 516 630 0  
## 3 0 112 62 964 0  
## is\_apartment has\_free\_cancelation  
## 1 1 1  
## 2 0 1  
## 3 1 1

Los hoteles con una puntuación de 10 destacan por obtener esta puntuación en todos los aspectos y tienen unas descripciones relativamente cortas, los precios mínimos y los máximos son muy distantes lo que sorprende. Dado que solo corresponde a 3 registros y que los resultados son razonables, puede corresponder a un hotel nuevo con muy pocas valoraciones, lo que permitiría una nota excepcionalmente alta otorgada por unos pocos clientes.

A continuación, se evalúan las instalaciones del hotel.

ggplot(booking, aes(y = facilities\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$facilities\_score)$out

## [1] 5.9 6.5 5.7 6.4 5.9 5.9 6.5 6.4 6.3 5.9 6.2 6.5 6.5 5.8 6.5 5.9 5.7 6.5 6.5  
## [20] 5.9

Vemos que los resultados obtenidos continuan en la línea anterior con algunos outliers en Barcelona o Valencia, pero con resultados lógicos.

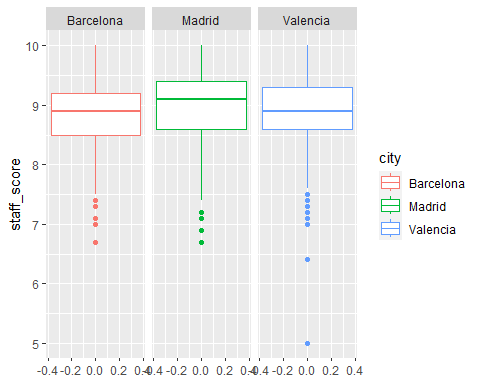
A continuación evaluamos aquellos hoteles que han obtenido una nota inferior al 6.5 para identificar si se trata de un error o no.

booking %>%  
 filter(facilities\_score <= 6.5) %>%  
 sample\_n(3)

## name city check.in month adults children  
## 1 B&B El Mercat Valencia 12-December-2022 December 5 3  
## 2 Pensi\357\277\275n Moratin Valencia 12-December-2022 December 5 3  
## 3 Total Valencia Terrace Valencia 12-December-2022 December 5 3  
## check.out num\_rooms  
## 1 16-December-2022 3  
## 2 16-December-2022 3  
## 3 16-December-2022 3  
## address postal\_code  
## 1 Carrer del Peu de la Creu 5, Ciutat Vella, 46001 Valencia, Spain 46001   
## 2 Moratin, 15, Ciutat Vella, 46002 Valencia, Spain 46002   
## 3 Reino de Valencia, 8, Eixample, 46005 Valencia, Spain 46005   
## hotel\_coordinates latitude longitude hotel\_score  
## 1 39.47309873,-0.38039260 39.47310 -0.3803926 6.4  
## 2 39.47079701,-0.37568763 39.47080 -0.3756876 6.6  
## 3 39.46465310,-0.37290300 39.46465 -0.3729030 6.7  
## hotel\_scores  
## 1 {'': '', 'Staff': '7.3', 'Facilities': '6.2', 'Cleanliness': '6.5', 'Comfort': '6.6', 'Value for money': '6.9', 'Location': '8.7', 'Free WiFi': '8.0'}  
## 2 {'': '', 'Staff': '8.1', 'Facilities': '6.4', 'Cleanliness': '7.1', 'Comfort': '6.7', 'Value for money': '7.4', 'Location': '9.3', 'Free WiFi': '7.3'}  
## 3 {'': '', 'Staff': '7.7', 'Facilities': '6.3', 'Cleanliness': '6.8', 'Comfort': '6.6', 'Value for money': '6.8', 'Location': '9.0', 'Free WiFi': '6.7'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 7.3 6.2 6.5 6.6  
## 2 8.1 6.4 7.1 6.7  
## 3 7.7 6.3 6.8 6.6  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 6.9 8.7 8.0  
## 2 7.4 9.3 7.3  
## 3 6.8 9.0 6.7  
## hotel\_description  
## 1 ['B&B El Mercat is located in the Ciutat Vella district in Valencia, 601 m from Barrio del Carmen and 1.7 mi from Bioparc Valencia.', 'Guests have access to 3 shared bathrooms. For your comfort, you will find free toiletries and a hairdryer. B&B El Mercat features free WiFi throughout the property.', 'There is luggage storage space at the property.', 'City of Arts & Sciences is 1.8 mi from B&B El Mercat, while Oceanografic is 2.2 mi from the property. Valencia Airport is 5 mi away.']  
## 2 ['Pensi\357\277\275n Moratin offers accommodations with free Wi-Fi and is 80 m from Valencia\357\277\275s main square, Plaza del Ayuntamiento.', 'All rooms feature a fan. Some rooms come with a shower and sink in the room and have access to shared toilet facilities, while others have a private bathroom. Towels and bed linen are provided.', 'The family-run Moratin is 5 minutes\357\277\275 walk from Valencia Nord Train Station and X\357\277\275tiva Metro Station, which offer direct connections to the airport. Thanks to its central location there are many caf\357\277\275s, bars and restaurants within a 5-minute walk.', 'Valencia Cathedral is 550 m away, while City of Arts & Sciences is 1.6 mi from the guest house. The nearest airport is Valencia Airport, 5.6 mi from Pensi\357\277\275n Moratin.']  
## 3 ['Total Valencia Terrace offers free WiFi, air conditioning and has access to a furnished terrace with barbecue facilities and city views. Located in Valencia, the accommodations is 1.2 mi from Valencia Cathedral.', 'Featuring plenty of natural light, this apartment comes with a seating area with a flat-screen TV. The kitchen is equipped with a microwave, oven and kitchenware. A washing machine is also available.', 'This spacious accommodations includes 4 twin bedrooms and 1 double bedroom. There is a bathroom with a shower and a second bathroom with a bath. Bed linen and towels are provided.', "A variety of restaurants, shops and bars are within 5 minutes' walk from the accommodations. The city of Arts and Sciences is 1.2 mi away. Valencia Nord Train Station is 400 m from the apartment.", 'Valencia Airport is 15 minutes away by car.']  
## features  
## 1 []  
## 2 []  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'City view', 'Washing machine', 'Free WiFi', 'Terrace', 'Air conditioning', 'Private Bathroom', 'Heating']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 0 0 1 1 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 1 0 1 0 1  
## free\_wifi  
## 1 0  
## 2 0  
## 3 1  
## room\_data  
## 1 {'220353003\_101557264\_0\_2\_0': {'room\_name': 'Quadruple Room', 'room\_price': '\357\277\275 430', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Breakfast \357\277\275 3 (optional)', 'Free cancellation until 11:59 PM on November 27, 2022', 'Pay nothing until November 25, 2022', 'Only 1 room left on our site'], 'room\_facilities': ['28 m\357\277\275', 'City view', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 full beds/ '}, '220353001\_265010488\_2\_2\_0': {'room\_name': 'Double or Twin Room', 'room\_price': '\357\277\275 221', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Breakfast \357\277\275 3 (optional)', 'Free cancellation until 11:59 PM on December 10, 2022', 'Pay nothing until December 8, 2022', 'Only 4 rooms left on our site'], 'room\_facilities': ['City view', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ '}, '220353002\_276866871\_0\_2\_0': {'room\_name': 'Single Room', 'room\_price': '\357\277\275 138', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Breakfast \357\277\275 3 (optional)', 'Free cancellation until 11:59 PM on December 7, 2022', 'Pay nothing until December 5, 2022', 'Only 1 room left on our site'], 'room\_facilities': ['City view', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 twin bed/ '}}  
## 2 {'32601505\_269319233\_0\_0\_0': {'room\_name': 'Triple Room with Shower', 'room\_price': '\357\277\275 166', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 1 room left on our site'], 'room\_facilities': ['15 m\357\277\275', 'Free WiFi'], 'room\_bed\_type': '3 twin beds/ '}, '32601506\_269319233\_0\_0\_0': {'room\_name': 'Triple Room', 'room\_price': '\357\277\275 181', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 1 room left on our site'], 'room\_facilities': ['21 m\357\277\275', 'Attached bathroom', 'Free WiFi'], 'room\_bed\_type': '3 twin beds/ '}, '32601511\_269319233\_0\_0\_0': {'room\_name': 'Double Room with Shower', 'room\_price': '\357\277\275 124', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 2 rooms left on our site'], 'room\_facilities': ['Free WiFi'], 'room\_bed\_type': '1 full bed/ '}, '32601504\_269319233\_0\_0\_0': {'room\_name': 'Twin Room with Shower', 'room\_price': '\357\277\275 124', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 3 rooms left on our site'], 'room\_facilities': ['Free WiFi'], 'room\_bed\_type': '2 twin beds/ '}, '32601507\_269319233\_0\_0\_0': {'room\_name': 'Double Room', 'room\_price': '\357\277\275 143', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 1 room left on our site'], 'room\_facilities': ['Private Bathroom', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ '}, '32601512\_269319233\_0\_0\_0': {'room\_name': 'Single Room with Shared Bathroom', 'room\_price': '\357\277\275 79', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'Pay in advance', 'Only 1 room left on our site'], 'room\_facilities': ['Free WiFi'], 'room\_bed\_type': '1 twin bed/ '}}  
## 3 {'107579801\_351582972\_8\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 854', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_8\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 757', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_9\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 764', 'room\_capacity': 'Max adults: 9', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_9\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 862', 'room\_capacity': 'Max adults: 9', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_10\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 772', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_10\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 870', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_7\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 750', 'room\_capacity': 'Max adults: 7', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_7\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 845', 'room\_capacity': 'Max adults: 7', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_6\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 743', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_6\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 837', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_5\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 736', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_5\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 829', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_4\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 728', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_4\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 821', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_3\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 721', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_3\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 813', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_2\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 714', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_2\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 804', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_1\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 707', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_1\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 796', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 118 5 13 2022-11-18 1 0  
## 2 82 4 5 2022-11-18 1 0  
## 3 86 4 9 2022-11-18 1 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 92 138 430 0  
## 2 0 133 79 181 0  
## 3 0 143 707 870 0  
## is\_apartment has\_free\_cancelation  
## 1 0 1  
## 2 0 1  
## 3 1 1

Observamos que los hoteles parecen tener unas características consistentes y que la extracción de características ha sido correcta por lo que no podemos atribuir los resultados a un error sino meramente a una baja calidad del hotel.

ggplot(booking, aes(y = staff\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$staff\_score)$out

## [1] 7.0 7.5 7.4 7.4 7.5 7.4 7.5 7.5 7.5 7.5 6.7 7.5 7.1 7.3 7.3 6.4 7.1 7.2  
## [19] 7.1 7.4 7.4 7.5 7.0 7.4 7.1 7.3 7.2 7.4 7.5 7.5 7.5 6.7 7.4 7.5 7.2 7.1  
## [37] 6.9 7.5 7.4 7.5 7.5 7.5 7.4 7.4 5.0 5.0 7.3 7.3 7.2 7.5 7.1 7.4 7.2 7.3  
## [55] 7.3 7.4 7.3 7.1 5.0 7.3 7.4 7.5 7.0 7.5 7.3 7.4 7.4 7.1 6.7 7.5 7.4 7.5  
## [73] 7.4 7.5 6.7 7.5 7.5 7.4 7.4 7.1 7.5 7.5 7.4 7.5 7.5 7.5 7.4 7.2 7.3 7.5  
## [91] 7.5 7.4 6.7 7.5 7.4 7.4 7.5 7.3 7.3 7.1 6.4 7.3 7.3 7.3 7.3 7.1 7.3 7.5  
## [109] 7.5 6.7 7.5 7.5 7.4 7.5 7.5 7.1 7.4 6.9 7.4 7.5 7.4 6.7 7.4 7.5 6.7 7.4  
## [127] 7.1 7.5 7.5 7.5 7.2

Para la puntuación otorgada a los trabajadores del hotel obtenemos unos resultados similares a los previos, con la diferencia de que en este apartado la puntuación promedio es más elevada rondando el 9 en las 3 ciudades y todas presentan outliers con valores inferiores.

Evaluamos aquellos resultados con puntuaciones muy bajas.

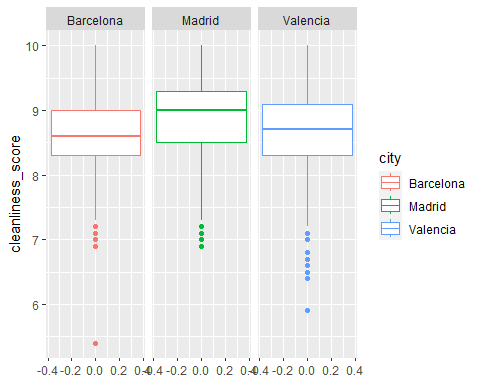
booking %>%  
 filter(staff\_score <= 6.7) %>%  
 sample\_n(3)

## name city check.in month  
## 1 Apartamento 3 habitaciones en el barrio Justicia Madrid 12-March-2023 March  
## 2 Apartamento 3 habitaciones en el barrio Justicia Madrid 12-June-2023 June  
## 3 \357\277\275tico en Paiporta con 2 grandes terrazas Valencia 12-June-2023 June  
## adults children check.out num\_rooms  
## 1 5 3 16-March-2023 3  
## 2 2 3 16-June-2023 2  
## 3 4 1 16-June-2023 2  
## address  
## 1 6 Calle de Hern\357\277\275n Cort\357\277\275s, Madrid City Center, 28004 Madrid, Spain  
## 2 6 Calle de Hern\357\277\275n Cort\357\277\275s, Madrid City Center, 28004 Madrid, Spain  
## 3 c/Bonrempos i Mirambell n\357\277\27517, Bloque2 pta9, 46200 Valencia, Spain  
## postal\_code hotel\_coordinates latitude longitude hotel\_score  
## 1 28004 40.42346550,-3.70014820 40.42347 -3.700148 7.3  
## 2 28004 40.42346550,-3.70014820 40.42347 -3.700148 7.3  
## 3 46200 39.43310010,-0.40836704 39.43310 -0.408367 9.0  
## hotel\_scores  
## 1 {'': '', 'Staff': '6.7', 'Facilities': '7.0', 'Cleanliness': '7.5', 'Comfort': '7.4', 'Value for money': '7.7', 'Location': '9.4'}  
## 2 {'': '', 'Staff': '6.7', 'Facilities': '7.0', 'Cleanliness': '7.5', 'Comfort': '7.4', 'Value for money': '7.7', 'Location': '9.4'}  
## 3 {'': '', 'Staff': '5.0', 'Facilities': '10', 'Cleanliness': '10', 'Comfort': '10', 'Value for money': '10', 'Location': '7.5'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 6.7 7 7.5 7.4  
## 2 6.7 7 7.5 7.4  
## 3 5.0 10 10.0 10.0  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 7.7 9.4 8.1  
## 2 7.7 9.4 8.2  
## 3 10.0 7.5 8.3  
## hotel\_description  
## 1 ['Apartamento 3 habitaciones en el barrio Justicia is located in the Madrid City Center district of Madrid, a 6-minute walk from Gran Via Station Metro Station, 641 m from Gran Via and a 14-minute walk from Thyssen-Bornemisza Museum. The property is 1.9 km from Retiro Park and a 17-minute walk from Mercado San Miguel.', 'The apartment is fitted with 3 bedrooms, a flat-screen TV and a fully equipped kitchen that provides guests with a dishwasher, a microwave, a washing machine, a fridge and an oven. Towels and bed linen are offered in the apartment.', 'Popular points of interest near the apartment include Plaza de Espa\357\277\275a Metro Station, Temple of Debod and Royal Palace of Madrid. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 15 km from Apartamento 3 habitaciones en el barrio Justicia.']  
## 2 ['Apartamento 3 habitaciones en el barrio Justicia is located in the Madrid City Center district of Madrid, a 6-minute walk from Gran Via Station Metro Station, 641 m from Gran Via and a 14-minute walk from Thyssen-Bornemisza Museum. The property is 1.9 km from Retiro Park and a 17-minute walk from Mercado San Miguel.', 'The apartment is fitted with 3 bedrooms, a flat-screen TV and a fully equipped kitchen that provides guests with a dishwasher, a microwave, a washing machine, a fridge and an oven. Towels and bed linen are offered in the apartment.', 'Popular points of interest near the apartment include Plaza de Espa\357\277\275a Metro Station, Temple of Debod and Royal Palace of Madrid. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 15 km from Apartamento 3 habitaciones en el barrio Justicia.']  
## 3 ['\357\277\275tico en Paiporta con 2 grandes terrazas, a property with a garden and a terrace, is located in Valencia, 7.4 km from Church of Saint Nicol\357\277\275s, 7.7 km from Oceanografic, as well as 8 km from Turia Gardens. The air-conditioned property is 6.9 km from Valencia North Station, and guests benefit from complimentary WiFi and private parking available on site.', 'The apartment has 3 bedrooms, a flat-screen TV and a fully equipped kitchen that provides guests with a dishwasher, a microwave, a washing machine, a fridge and an oven. Towels and bed linen are provided.', 'Gonz\357\277\275lez Mart\357\277\275 National Museum of Ceramics and Decorative Arts is 8.2 km from the apartment, while Basilica de la Virgen de los Desamparados is 8.2 km away. The nearest airport is Valencia Airport, 13 km from \357\277\275tico en Paiporta con 2 grandes terrazas.']  
## features  
## 1 ['Whole apartment', 'Kitchen', 'Washing machine']  
## 2 ['Whole apartment', 'Kitchen', 'Washing machine']  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'Garden', 'Pet friendly', 'Washing machine', 'Free WiFi', 'Terrace', 'Air conditioning', 'Baggage storage']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 0 0 0 1 0 0  
## 2 0 0 0 1 0 0  
## 3 0 0 1 1 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 0 0 0 0 1  
## free\_wifi  
## 1 0  
## 2 0  
## 3 1  
## room\_data  
## 1 {'794449001\_339672653\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 720', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on March 7, 2022', 'Pay nothing until March 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 653', 'room\_capacity': 'Max people: 5', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 685', 'room\_capacity': 'Max people: 6', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 756', 'room\_capacity': 'Max people: 6', 'room\_options': ['Free cancellation until 11:59 PM on March 7, 2022', 'Pay nothing until March 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 621', 'room\_capacity': 'Max people: 4', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 685', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on March 7, 2022', 'Pay nothing until March 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 633', 'room\_capacity': 'Max people: 3', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 699', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on March 7, 2022', 'Pay nothing until March 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 608', 'room\_capacity': 'Max people: 2', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 670', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on March 7, 2022', 'Pay nothing until March 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}}  
## 2 {'794449001\_339522317\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 653', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 720', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 685', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 756', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 621', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 685', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 633', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 699', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339522317\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 608', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Non-refundable', 'Pay in advance', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}, '794449001\_339672653\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 670', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022', 'FREE taxi from the airport to this property '], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 1 full bed/ '}}  
## 3 {'922256801\_363993621\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 283', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on June 9, 2023', 'Pay nothing until June 7, 2023'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 2 twin beds/ '}, '922256801\_363927374\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 268', 'room\_capacity': 'Max people: 4', 'room\_options': ['Flexible to reschedule if plans change', 'Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 2 twin beds/ '}, '922256801\_363927374\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 270', 'room\_capacity': 'Max people: 5', 'room\_options': ['Flexible to reschedule if plans change', 'Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 2 twin beds/ '}, '922256801\_363927374\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 263', 'room\_capacity': 'Max people: 3', 'room\_options': ['Flexible to reschedule if plans change', 'Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 2 twin beds/ '}, '922256801\_363927374\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 259', 'room\_capacity': 'Max people: 2', 'room\_options': ['Flexible to reschedule if plans change', 'Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 full bed/ 2 twin beds/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 46 2 21 2022-11-18 0 1  
## 2 68 3 17 2022-11-19 0 0  
## 3 82 4 7 2022-11-19 0 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 143 608 756 0  
## 2 0 143 608 756 0  
## 3 0 151 259 283 0  
## is\_apartment has\_free\_cancelation  
## 1 1 1  
## 2 1 1  
## 3 1 1

De los resultados obtenidos sorprende uno en Valencia que tiene unas puntuaciones muy elevadas excepto en **staff\_score** la cual es bastante baja. Si nos fijamos estas puntuaciones son muy exactas (10.0, 9.0, 7.5 o 5.0) lo cual puede venir dado por un número de puntuaciones muy reducidas en las que un cliente ha puntuado de forma muy negativa este aspecto. No por ello eliminaremos el registro o lo modificaremos ya que es razonable el resultado, y afectará en la posición que presenta el hotel en el buscador.

A continuación se evalúa la limpieza del hotel.

ggplot(booking, aes(y = cleanliness\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$cleanliness\_score)$out

## [1] 7.2 6.9 7.2 7.0 7.0 6.9 7.1 5.9 6.8 7.0 7.1 6.4 7.1 5.9 7.2 6.7 7.1 7.1 7.2  
## [20] 5.9 6.4 6.7 7.1 6.8 6.6 6.5 7.0 6.9 7.1 7.0 7.1 7.1 7.0 6.9 6.8 7.0 6.7 7.1  
## [39] 7.0 6.8 7.0 6.8 7.1 7.1 7.2 7.2 6.9 6.9 5.4 7.1 7.2 7.0 6.9 6.8 7.2 6.5 7.2  
## [58] 6.9 6.8 5.9 6.4 6.8 6.5 7.1 6.8 6.7 7.1 6.4 5.9 6.8 7.2 6.9 6.9 7.0 7.1 7.1  
## [77] 7.0 7.0 7.0 7.2 7.1 6.9 6.9 7.0

Observamos unos resultados muy similares a los anteriores. Para descartar algún error observamos aquellos registros con puntuaciones excepcionalmente bajas.

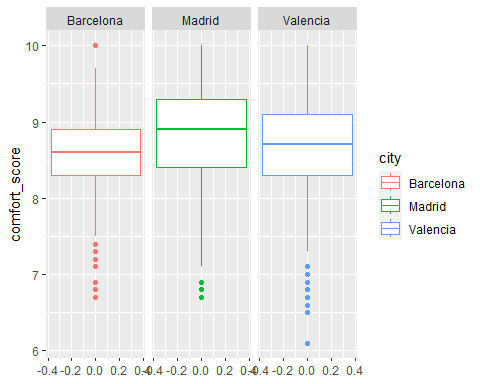
booking %>%  
 filter(cleanliness\_score < 7.0) %>%  
 sample\_n(3)

## name city check.in month adults children  
## 1 Coroa del Mar Valencia 12-June-2023 June 2 3  
## 2 Benlloch Silence Valencia 12-March-2023 March 2 3  
## 3 Total Valencia Terrace Valencia 12-December-2022 December 5 3  
## check.out num\_rooms  
## 1 16-June-2023 2  
## 2 16-March-2023 2  
## 3 16-December-2022 3  
## address postal\_code  
## 1 Calle Navardera, 2 , Poblats Maritims, 46024 Valencia, Spain 46024   
## 2 78 Avinguda del Cardenal Benlloch, Algiros, 46021 Valencia, Spain 46021   
## 3 Reino de Valencia, 8, Eixample, 46005 Valencia, Spain 46005   
## hotel\_coordinates latitude longitude hotel\_score  
## 1 39.45995123,-0.33388876 39.45995 -0.3338888 5.9  
## 2 39.47227930,-0.35369160 39.47228 -0.3536916 7.0  
## 3 39.46465310,-0.37290300 39.46465 -0.3729030 6.7  
## hotel\_scores  
## 1 {'': '', 'Staff': '7.1', 'Facilities': '5.9', 'Cleanliness': '5.9', 'Comfort': '6.1', 'Value for money': '6.7', 'Location': '7.6', 'Free WiFi': '7.7'}  
## 2 {'': '', 'Staff': '7.3', 'Facilities': '7.4', 'Cleanliness': '6.8', 'Comfort': '7.1', 'Value for money': '7.5', 'Location': '6.8', 'Free WiFi': '5.0'}  
## 3 {'': '', 'Staff': '7.7', 'Facilities': '6.3', 'Cleanliness': '6.8', 'Comfort': '6.6', 'Value for money': '6.8', 'Location': '9.0', 'Free WiFi': '6.7'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 7.1 5.9 5.9 6.1  
## 2 7.3 7.4 6.8 7.1  
## 3 7.7 6.3 6.8 6.6  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 6.7 7.6 7.7  
## 2 7.5 6.8 5.0  
## 3 6.8 9.0 6.7  
## hotel\_description  
## 1 ["You're eligible for a Genius discount at Coroa del Mar! To save at this property, all you have to do is sign in.", 'Coroa del Mar is located in Valencia, 200 m from Puerto de Valencia. Oceanografic is 0.9 mi away. Free WiFi is available throughout the property.', 'Featuring a balcony, all units are air conditioned and feature a seating and dining area. All units include a kitchen equipped with a microwave. A fridge and stovetop are also offered, as well as a coffee machine and a kettle. Every unit features a private bathroom with a bidet and a hairdryer. Towels are provided.', 'City of Arts & Sciences is 1 mi from Coroa del Mar. The nearest airport is Valencia Airport, 8.1 mi from Coroa del Mar.']  
## 2 ['In the Algiros district of Valencia, close to Turia Gardens, Benlloch Silence features free WiFi and a washing machine. It is located 1.8 km from Jardines de Monforte and provides a tour desk.', "The air-conditioned apartment consists of 2 separate bedrooms, 2 bathrooms with free toiletries, a seating area area, and a living room. There's a dining area and a kitchen complete with a microwave, a fridge and a stovetop.", 'Gonz\357\277\275lez Mart\357\277\275 National Museum of Ceramics and Decorative Arts is 3.1 km from the apartment, while Valencia North Station is 3.1 km away. The nearest airport is Valencia Airport, 13 km from Benlloch Silence.']  
## 3 ['Total Valencia Terrace offers free WiFi, air conditioning and has access to a furnished terrace with barbecue facilities and city views. Located in Valencia, the accommodations is 1.2 mi from Valencia Cathedral.', 'Featuring plenty of natural light, this apartment comes with a seating area with a flat-screen TV. The kitchen is equipped with a microwave, oven and kitchenware. A washing machine is also available.', 'This spacious accommodations includes 4 twin bedrooms and 1 double bedroom. There is a bathroom with a shower and a second bathroom with a bath. Bed linen and towels are provided.', "A variety of restaurants, shops and bars are within 5 minutes' walk from the accommodations. The city of Arts and Sciences is 1.2 mi away. Valencia Nord Train Station is 400 m from the apartment.", 'Valencia Airport is 15 minutes away by car.']  
## features  
## 1 ['Apartments', 'Kitchen', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms', 'Heating', 'Elevator']  
## 2 ['Whole apartment', 'Size', 'Kitchen', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms']  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'City view', 'Washing machine', 'Free WiFi', 'Terrace', 'Air conditioning', 'Private Bathroom', 'Heating']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 0 0 0 1 0 0  
## 2 0 0 0 1 0 0  
## 3 0 0 0 1 1 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 1 1 1 1 0 1  
## 2 1 1 0 0 0 1  
## 3 0 1 0 1 0 1  
## free\_wifi  
## 1 1  
## 2 1  
## 3 1  
## room\_data  
## 1 {'154939904\_364159804\_0\_0\_0': {'room\_name': 'Duplex Apartment', 'room\_price': '\357\277\275 196', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on June 10, 2022', 'Pay nothing until June 8, 2022', 'Only 1 left on our site'], 'room\_facilities': ['Entire apartment', '60 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 twin bed/ 1 full bed/ 1 sofa bed/ '}, '154939903\_364159804\_0\_0\_0': {'room\_name': 'Duplex Apartment', 'room\_price': '\357\277\275 168', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Free cancellation until 11:59 PM on June 10, 2022', 'Pay nothing until June 8, 2022', 'Only 1 left on our site'], 'room\_facilities': ['Entire apartment', '55 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ '}, '154939901\_364159804\_0\_0\_0': {'room\_name': 'Apartment Studio 3 Pax', 'room\_price': '\357\277\275 156', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on June 10, 2022', 'Pay nothing until June 8, 2022', 'Only 1 left on our site'], 'room\_facilities': ['Entire studio', '45 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '3 twin beds/ '}}  
## 2 {'794422401\_364182307\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 195', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on March 10, 2022', 'Pay nothing until March 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}, '794422401\_364182307\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 210', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on March 10, 2022', 'Pay nothing until March 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}, '794422401\_364182307\_5\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 217', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on March 10, 2022', 'Pay nothing until March 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}}  
## 3 {'107579801\_351582972\_8\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 854', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_8\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 757', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_9\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 764', 'room\_capacity': 'Max adults: 9', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_9\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 862', 'room\_capacity': 'Max adults: 9', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_10\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 772', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_10\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 870', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_7\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 750', 'room\_capacity': 'Max adults: 7', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_7\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 845', 'room\_capacity': 'Max adults: 7', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_6\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 743', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_6\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 837', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_5\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 736', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_5\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 829', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_4\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 728', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_4\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 821', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_3\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 721', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_3\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 813', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_2\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 714', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_2\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 804', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_91345125\_1\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 707', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}, '107579801\_351582972\_1\_0\_0': {'room\_name': 'Five-Bedroom Apartment with Terrace', 'room\_price': '\357\277\275 796', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022'], 'room\_facilities': ['Entire apartment', '167 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 11 1 11 2022-11-18 0 0  
## 2 38 2 13 2022-11-19 0 1  
## 3 86 4 9 2022-11-18 1 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 131 156 196 0  
## 2 0 112 195 217 0  
## 3 0 143 707 870 0  
## is\_apartment has\_free\_cancelation  
## 1 1 1  
## 2 1 1  
## 3 1 1

Vemos un comportamiento razonable en estos hoteles, por lo que, siendo resultados razonables, aceptamos estos valores outliers.

A continuación, se valora la puntuación de la comodidad.

ggplot(booking, aes(y = comfort\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



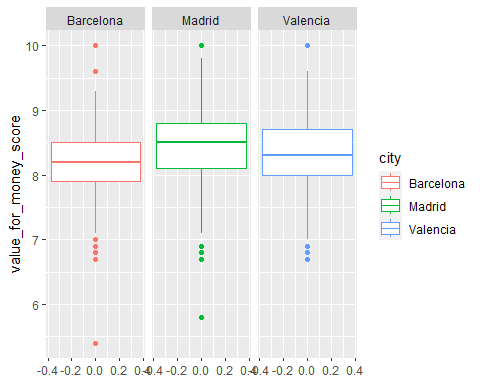
boxplot.stats(booking$comfort\_score)$out

## [1] 7.1 7.1 7.1 6.8 6.9 7.1 6.7 6.1 7.1 6.8 6.1 6.7 6.1 7.0 7.1 7.1 6.1 7.0 6.7  
## [20] 6.7 6.6 6.9 6.5 6.6 6.9 7.1 6.7 7.1 7.1 6.9 6.8 7.1 6.9 6.8 7.0 7.1 7.1 6.8  
## [39] 6.7 7.1 6.7 7.1 7.1 6.9 7.1 7.1 7.0 7.1 7.1 7.1 7.1 6.7 7.1 6.1 6.1 7.1 7.0  
## [58] 7.1 7.1 6.7 6.7 6.1 6.9 7.1 6.9 6.7 6.8 7.1 6.9 6.9

booking %>%  
 filter(comfort\_score < 7.0) %>%  
 sample\_n(3)

## name city check.in  
## 1 Charming Gran V\357\277\275a Luna Madrid 12-December-2022  
## 2 Design Flats Botanico Valencia 12-March-2023  
## 3 Modern 3 Bedroom in La Latina 10 minutes to Metro Madrid 12-March-2023  
## month adults children check.out num\_rooms  
## 1 December 5 3 16-December-2022 3  
## 2 March 2 3 16-March-2023 2  
## 3 March 2 3 16-March-2023 2  
## address postal\_code  
## 1 11 Calle de la Luna, Madrid City Center, 28004 Madrid, Spain 28004   
## 2 10 Carrer del T\357\277\275ria, Extramurs, 46008 Valencia, Spain 46008   
## 3 Madrid City Center, 28005 Madrid, Spain 28005   
## hotel\_coordinates latitude longitude hotel\_score  
## 1 40.42239400,-3.70621100 40.42239 -3.7062110 6.6  
## 2 39.47455450,-0.38517390 39.47455 -0.3851739 6.8  
## 3 40.40919100,-3.71138920 40.40919 -3.7113892 6.6  
## hotel\_scores  
## 1 {'': '', 'Staff': '7.9', 'Facilities': '6.7', 'Cleanliness': '7.1', 'Comfort': '6.7', 'Value for money': '5.8', 'Location': '8.8'}  
## 2 {'': '', 'Staff': '8.2', 'Facilities': '6.7', 'Cleanliness': '7.1', 'Comfort': '6.7', 'Value for money': '6.9', 'Location': '8.5', 'Free WiFi': '6.9'}  
## 3 {'': '', 'Staff': '7.5', 'Facilities': '7.2', 'Cleanliness': '6.9', 'Comfort': '6.9', 'Value for money': '7.5', 'Location': '9.1'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 7.9 6.7 7.1 6.7  
## 2 8.2 6.7 7.1 6.7  
## 3 7.5 7.2 6.9 6.9  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 5.8 8.8 8.9  
## 2 6.9 8.5 6.9  
## 3 7.5 9.1 8.3  
## hotel\_description  
## 1 ['Located within a 4-minute walk of Gran Via and 549 m of Gran Via Station Metro Station in the center of Madrid, Charming Gran V\357\277\275a Luna has accommodations with free WiFi and seating area.', 'All units here are air-conditioned and feature a flat-screen TV, a living room with a sofa, a well-equipped kitchen with a dining area, and a private bathroom with shower, a hairdryer and free toiletries. A microwave, a fridge and stovetop are also available, as well as an electric tea pot and a coffee machine.', 'Popular points of interest near the apartment include Thyssen-Bornemisza Museum, Plaza de Espa\357\277\275a Metro Station and Temple of Debod. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 17.7 km from Charming Gran V\357\277\275a Luna.']  
## 2 ['Located in Valencia, a 8-minute walk from Church of Saint Nicol\357\277\275s and 1.1 km from Basilica de la Virgen de los Desamparados, Design Flats Botanico provides accommodations with amenities such as free WiFi and a flat-screen TV. This apartment is 2.7 km from Jardines de Monforte and 4.8 km from Bioparc Valencia.', 'The air-conditioned apartment is composed of 2 separate bedrooms, a living room, a fully equipped kitchen with a microwave and fridge, and 1 bathroom. Towels and bed linen are available.', 'Popular points of interest near the apartment include Turia Gardens, Gonz\357\277\275lez Mart\357\277\275 National Museum of Ceramics and Decorative Arts and Valencia North Station. The nearest airport is Valencia, 10 km from Design Flats Botanico, and the property offers a paid airport shuttle service.']  
## 3 ['Well set in the center of Madrid, Modern 3 Bedroom in La Latina 10 minutes to Metro offers a patio. It is located a 9-minute walk from Puerta de Toledo and provides free WiFi plus a 24-hour front desk.', 'The air-conditioned apartment is composed of 3 separate bedrooms, a fully equipped kitchen with a microwave and a fridge, and 2 bathrooms. Towels and bed linen are featured in the apartment.', 'Popular points of interest near the apartment include Plaza Mayor, Mercado San Miguel and Puerta del Sol. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 19.3 km from Modern 3 Bedroom in La Latina 10 minutes to Metro.']  
## features  
## 1 ['Apartments', 'Kitchen', 'City view', 'Washing machine', 'Free WiFi', 'Balcony', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms', 'Safe']  
## 2 ['Whole apartment', 'Size', 'Kitchen', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms']  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'Washing machine', 'Free WiFi', 'Balcony', 'Air conditioning', 'Private Bathroom', '24-hour front desk', 'Non-smoking rooms']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 1 0 0 1 1 0  
## 2 0 0 0 1 0 0  
## 3 1 0 0 1 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 1 1 0 0 1 1  
## 2 1 1 0 0 0 1  
## 3 1 1 0 0 0 1  
## free\_wifi  
## 1 1  
## 2 1  
## 3 1  
## room\_data  
## 1 {'786984801\_338289866\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 908', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_5\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 826', 'room\_capacity': 'Max people: 5', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 830', 'room\_capacity': 'Max people: 6', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_6\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 913', 'room\_capacity': 'Max people: 6', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_7\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 894', 'room\_capacity': 'Max people: 7', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_7\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 983', 'room\_capacity': 'Max people: 7', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_8\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 973', 'room\_capacity': 'Max people: 8', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_8\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 1,071', 'room\_capacity': 'Max people: 8', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 822', 'room\_capacity': 'Max people: 4', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 903', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 818', 'room\_capacity': 'Max people: 3', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 899', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289867\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 816', 'room\_capacity': 'Max people: 2', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984801\_338289866\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 895', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Attached bathroom', 'Balcony', 'City view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 sofa bed/ 1 full bed/ 1 sofa bed/ 3/ '}, '786984802\_338289867\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 891', 'room\_capacity': 'Max people: 4', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '786984802\_338289866\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 979', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '786984802\_338289867\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 813', 'room\_capacity': 'Max people: 3', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '786984802\_338289866\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 892', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on December 6, 2022', 'Pay nothing until December 4, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '786984802\_338289867\_2\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 733', 'room\_capacity': 'Max people: 2', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '786984802\_338522788\_2\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 804', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on December 7, 2022', 'Pay nothing until December 5, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'City view', 'Air conditioning', 'Spa tub', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}}  
## 2 {'277969801\_351582940\_2\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 500', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on May 28, 2023'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_106881216\_2\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 451', 'room\_capacity': 'Max people: 2', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_106881216\_3\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 456', 'room\_capacity': 'Max people: 3', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_351582940\_3\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 510', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on May 28, 2023'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_106881216\_4\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 460', 'room\_capacity': 'Max people: 4', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_351582940\_4\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 520', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on May 28, 2023'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_106881216\_1\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 447', 'room\_capacity': 'Max people: 1', 'room\_options': ['Non-refundable'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}, '277969801\_351582940\_1\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 490', 'room\_capacity': 'Max people: 1', 'room\_options': ['Free cancellation until 11:59 PM on May 28, 2023'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Flat-screen TV'], 'room\_bed\_type': '1 queen bed/ 1 full bed/ '}}  
## 3 {'685428002\_310811050\_6\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 558', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on March 6, 2022', 'Pay nothing until March 4, 2022'], 'room\_facilities': ['Entire apartment', '95 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Air conditioning', 'Patio', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 full bed/ 2/ '}, '685428002\_331007874\_6\_0\_0': {'room\_name': 'Apartment', 'room\_price': '\357\277\275 536', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '95 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Air conditioning', 'Patio', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 full bed/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 76 3 25 2022-11-18 1 0  
## 2 123 5 22 2022-11-19 0 1  
## 3 78 4 1 2022-11-19 0 1  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 134 71 983 0  
## 2 0 135 447 520 0  
## 3 0 115 536 558 0  
## is\_apartment has\_free\_cancelation  
## 1 1 1  
## 2 1 1  
## 3 1 1

ggplot(booking, aes(y = value\_for\_money\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



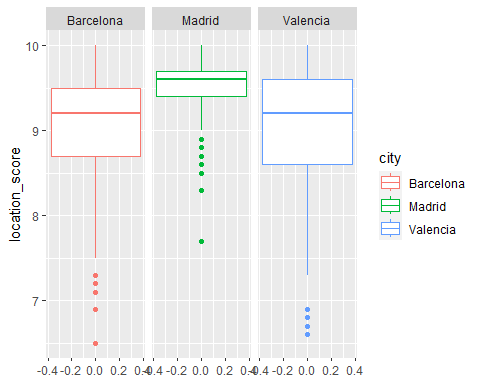
boxplot.stats(booking$value\_for\_money\_score)$out

## [1] 6.9 6.9 10.0 10.0 10.0 5.8 6.7 10.0 6.7 10.0 10.0 6.7 6.9 10.0 6.7  
## [16] 6.8 6.9 6.7 6.7 6.9 10.0 10.0 10.0 10.0 5.8 10.0 10.0 9.8 10.0 10.0  
## [31] 10.0 6.8 9.8 10.0 10.0 10.0 10.0 10.0 6.9 10.0 9.8 10.0 10.0 6.7 6.9  
## [46] 10.0 10.0 6.9 6.9 10.0 6.7 6.7 5.4 6.9 10.0 10.0 6.8 10.0 10.0 10.0  
## [61] 6.9 6.8 10.0 10.0 6.7 6.7 6.7 10.0 6.9 6.7 10.0 10.0 10.0 5.8 6.9  
## [76] 10.0 10.0 9.8 10.0 10.0 10.0

booking %>%  
 filter(value\_for\_money\_score <= 7.0) %>%  
 sample\_n(3)

## name city check.in month  
## 1 ART & FLATS Pleno Centro Valencia Valencia 12-June-2023 June  
## 2 Apartamento Tirso de Molina Madrid 12-June-2023 June  
## 3 Rent Top Apartments Beach-Diagonal Mar Barcelona 12-December-2022 December  
## adults children check.out num\_rooms  
## 1 2 3 16-June-2023 2  
## 2 4 1 16-June-2023 2  
## 3 5 3 16-December-2022 3  
## address  
## 1 Av. del Bar\357\277\275 de C\357\277\275rcer, 37., Ciutat Vella, 46001 Valencia, Spain  
## 2 Calle Jes\357\277\275s y Mar\357\277\275a, 17, Madrid City Center, 28012 Madrid, Spain  
## 3 Passeig de taulat 279 C , Sant Mart\357\277\275, 08019 Barcelona, Spain  
## postal\_code hotel\_coordinates latitude longitude hotel\_score  
## 1 46001 39.47047180,-0.37957606 39.47047 -0.3795761 6.0  
## 2 28012 40.41116930,-3.70356840 40.41117 -3.7035684 6.9  
## 3 08019 41.40493361,2.21523240 41.40493 2.2152324 7.5  
## hotel\_scores  
## 1 {'': '', 'Staff': '6.4', 'Facilities': '5.7', 'Cleanliness': '6.4', 'Comfort': '6.1', 'Value for money': '6.7', 'Location': '8.9', 'Free WiFi': '10'}  
## 2 {'': '', 'Staff': '9.3', 'Facilities': '7.1', 'Cleanliness': '8.3', 'Comfort': '6.8', 'Value for money': '6.9', 'Location': '9.4', 'Free WiFi': '7.5'}  
## 3 {'': '', 'Staff': '7.4', 'Facilities': '7.6', 'Cleanliness': '7.6', 'Comfort': '7.9', 'Value for money': '7.0', 'Location': '8.1', 'Free WiFi': '7.5'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 6.4 5.7 6.4 6.1  
## 2 9.3 7.1 8.3 6.8  
## 3 7.4 7.6 7.6 7.9  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 6.7 8.9 10.0  
## 2 6.9 9.4 7.5  
## 3 7.0 8.1 7.5  
## hotel\_description  
## 1 ['Located in Valencia, a 7-minute walk from Valencia North Station and half a kilometer from Church of Saint Nicol\357\277\275s, ART & FLATS Pleno Centro Valencia provides city views and free WiFi.', 'Each unit has a fully equipped kitchen with a dishwasher, a seating area, a flat-screen TV, a washing machine, and a private bathroom with bidet and a hairdryer. A microwave, a fridge and oven are also provided, as well as an electric tea pot and a coffee machine.', 'Popular points of interest near the apartment include Turia Gardens, Gonz\357\277\275lez Mart\357\277\275 National Museum of Ceramics and Decorative Arts and Basilica de la Virgen de los Desamparados. The nearest airport is Valencia Airport, 10 km from ART & FLATS Pleno Centro Valencia.']  
## 2 ['In the Madrid City Center district of Madrid, close to Plaza Mayor, Apartamento Tirso de Molina features free WiFi and a washing machine. This apartment is a 11-minute walk from Reina Sofia Museum and 1.3 km from Atocha Train Station.', "The air-conditioned apartment consists of 2 separate bedrooms, 1 bathroom with a hairdryer and free toiletries, and a seating area. There's a dining area and a kitchen complete with a dishwasher, a microwave and a fridge.", 'Popular points of interest near the apartment include Mercado San Miguel, Puerta del Sol and Puerta de Toledo. The nearest airport is Adolfo Suarez Madrid-Barajas Airport, 17.7 km from Apartamento Tirso de Molina.']  
## 3 ["You're eligible for a Genius discount at Rent Top Apartments Beach-Diagonal Mar! To save at this property, all you have to do is sign in.", "Offering different apartments in Barcelona\357\277\275s modern district of Diagonal Mar, Rent Top Apartments is within 10 minutes' walk from the beach. It offers stylish apartments with private terraces with city, garden or beach views.", 'Rent Top Apartments is less than 5 minutes\357\277\275 walk from Diagonal Mar Shopping Center and from the Forum Building. There is a Tourist Bus stop and a taxi rank just in front of the property, which is also accessible by metro.', 'Each spacious and bright air-conditioned apartment features modern d\357\277\275cor in light colors, and has free Wi-Fi and a 24-hour security service. All come with a lounge with flat-screen TV and DVD player, a well-equipped kitchen and private bathrooms. Some apartments have access to a shared seasonal swimming pool.', 'Featuring city and sea views, apartments at Rent Top include bed linen, towels, heating and a washing machine.']  
## features  
## 1 ['Apartments', 'Kitchen', 'City view', 'Pet friendly', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms', 'Heating']  
## 2 ['Whole apartment', 'Size', 'Kitchen', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms', 'Heating']  
## 3 ['Apartments', 'Kitchen', 'Garden', 'Pet friendly', 'Swimming pool', 'Washing machine', 'Free WiFi', 'Terrace', 'Air conditioning', 'Private Bathroom']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 0 0 1 1 1 0  
## 2 0 0 0 1 0 0  
## 3 0 1 1 1 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 1 1 0 1 0 1  
## 2 1 1 0 1 0 1  
## 3 0 1 0 0 0 1  
## free\_wifi  
## 1 1  
## 2 1  
## 3 1  
## room\_data  
## 1 {'155438801\_329866080\_5\_0\_0': {'room\_name': 'Two-Bedroom Apartment-Door 4', 'room\_price': '\357\277\275 354', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 1 sofa bed/ 2/ '}, '155438801\_329866080\_6\_0\_0': {'room\_name': 'Two-Bedroom Apartment-Door 4', 'room\_price': '\357\277\275 373', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 1 sofa bed/ 2/ '}, '155438801\_329866080\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment-Door 4', 'room\_price': '\357\277\275 335', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 1 sofa bed/ 2/ '}, '155438801\_329866080\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment-Door 4', 'room\_price': '\357\277\275 316', 'room\_capacity': 'Max adults: 3', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 1 sofa bed/ 2/ '}, '155438801\_145852844\_2\_0\_0': {'room\_name': 'Two-Bedroom Apartment-Door 4', 'room\_price': '\357\277\275 272', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', 'Private kitchen', 'Private Bathroom', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Soundproof', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 1 sofa bed/ 2/ '}}  
## 2 {'281002801\_334200198\_4\_0\_0': {'room\_name': 'Apartment - Ground Floor', 'room\_price': '\357\277\275 439', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on June 7, 2022', 'Pay nothing until June 5, 2022'], 'room\_facilities': ['Entire apartment', '59 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 1 twin bed/ 1 sofa bed/ 1 sofa bed/ '}}  
## 3 {'38484103\_329102250\_0\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 1,308', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Free cancellation until 11:59 PM on December 9, 2022', 'Pay nothing until December 7, 2022'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484103\_89623448\_0\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 1,252', 'room\_capacity': 'Max adults: 8', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '90 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484108\_89623448\_0\_0\_0': {'room\_name': 'Superior Four-Bedroom Apartment', 'room\_price': '\357\277\275 1,705', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Non-refundable', 'Pay in advance'], 'room\_facilities': ['Entire apartment', '130 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Ocean view', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484108\_329102250\_0\_0\_0': {'room\_name': 'Superior Four-Bedroom Apartment', 'room\_price': '\357\277\275 1,730', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Free cancellation until 11:59 PM on December 9, 2022', 'Pay nothing until December 7, 2022'], 'room\_facilities': ['Entire apartment', '130 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Ocean view', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484109\_89623448\_0\_0\_0': {'room\_name': 'Penthouse Four-Bedroom Apartment', 'room\_price': '\357\277\275 1,811', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Non-refundable', 'Pay in advance', 'discount available'], 'room\_facilities': ['Entire apartment', '175 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Ocean view', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484109\_329102250\_0\_0\_0': {'room\_name': 'Penthouse Four-Bedroom Apartment', 'room\_price': '\357\277\275 1,895', 'room\_capacity': 'Max adults: 10', 'room\_options': ['Free cancellation until 11:59 PM on December 9, 2022', 'Pay nothing until December 7, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '175 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Ocean view', 'City view', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 1 queen bed/ 2 twin beds/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484111\_89623448\_0\_0\_0': {'room\_name': 'Standard Two-Bedroom Apartment', 'room\_price': '\357\277\275 897', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable', 'Pay in advance', 'discount available'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 1 sofa bed/ '}, '38484111\_329102250\_0\_0\_0': {'room\_name': 'Standard Two-Bedroom Apartment', 'room\_price': '\357\277\275 937', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on December 9, 2022', 'Pay nothing until December 7, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 1 sofa bed/ '}, '38484102\_89623448\_0\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 965', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Non-refundable', 'Pay in advance', 'discount available'], 'room\_facilities': ['Entire apartment', '75 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 1 sofa bed/ 2/ '}, '38484102\_329102250\_0\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 1,009', 'room\_capacity': 'Max adults: 6', 'room\_options': ['Free cancellation until 11:59 PM on December 9, 2022', 'Pay nothing until December 7, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '75 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Air conditioning', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 queen bed/ 2 twin beds/ 1 sofa bed/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 99 4 22 2022-11-18 0 0  
## 2 100 4 23 2022-11-18 0 0  
## 3 61 3 10 2022-11-18 1 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 125 272 373 0  
## 2 0 117 439 439 0  
## 3 0 177 9 965 0  
## is\_apartment has\_free\_cancelation  
## 1 1 0  
## 2 1 1  
## 3 1 1

ggplot(booking, aes(y = location\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



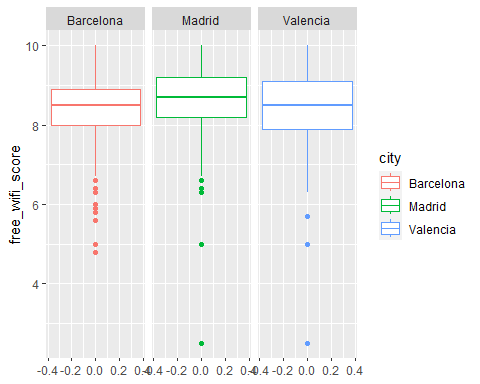
boxplot.stats(booking$location\_score)$out

## [1] 7.6 7.3 7.6 7.6 7.2 7.6 6.8 6.7 7.6 7.5 7.7 7.3 7.8 7.6 7.8 7.5 7.8 7.8  
## [19] 7.6 6.8 7.8 6.8 7.8 7.5 7.7 6.6 7.3 7.8 7.5 7.8 7.5 7.8 7.5 7.3 7.6 7.6  
## [37] 7.6 7.7 6.8 7.5 7.5 6.7 7.6 7.6 7.3 7.3 7.5 6.8 6.8 6.8 7.5 7.7 7.7 6.9  
## [55] 7.7 7.8 6.8 7.5 7.8 7.3 7.6 7.5 7.2 7.3 7.8 7.1 7.8 7.2 6.5 6.9 7.6 6.8  
## [73] 7.7 6.8 6.7 7.6 7.6 7.3 7.6 7.6 7.6 7.8 7.5 6.9 7.2 7.2 6.8 7.6 7.8 7.8  
## [91] 6.8 7.7 7.5 7.5 7.8 7.7 7.5 6.8 6.8 6.8 7.8 7.8 7.8 7.7 7.6 7.8 7.7

booking %>%  
 filter(location\_score < 7.0) %>%  
 sample\_n(3)

## name city check.in  
## 1 Park Guell Barcelona 12-March-2023  
## 2 PISO DE DISE\357\277\275O A 5 MIN ANDANDO DE LA PLAYA Valencia 12-March-2023  
## 3 Benlloch Silence Valencia 12-December-2022  
## month adults children check.out num\_rooms  
## 1 March 5 3 16-March-2023 3  
## 2 March 5 3 16-March-2023 3  
## 3 December 2 3 16-December-2022 2  
## address  
## 1 carrer gran vista 127, Horta-Guinard\357\277\275, 08032 Barcelona, Spain  
## 2 35 Carrer de Sant Joan de D\357\277\275u, Poblats Maritims, 46011 Valencia, Spain  
## 3 78 Avinguda del Cardenal Benlloch, Algiros, 46021 Valencia, Spain  
## postal\_code hotel\_coordinates latitude longitude hotel\_score  
## 1 08032 41.41887120,2.15781600 41.41887 2.1578160 7.2  
## 2 46011 39.47732309,-0.32829236 39.47732 -0.3282924 8.6  
## 3 46021 39.47227930,-0.35369160 39.47228 -0.3536916 7.0  
## hotel\_scores  
## 1 {'': '', 'Staff': '8.1', 'Facilities': '7.3', 'Cleanliness': '7.7', 'Comfort': '7.6', 'Value for money': '7.4', 'Location': '6.9', 'Free WiFi': '5.8'}  
## 2 {'': '', 'Staff': '9.2', 'Facilities': '9.2', 'Cleanliness': '9.1', 'Comfort': '9.4', 'Value for money': '9.1', 'Location': '6.7', 'Free WiFi': '10'}  
## 3 {'': '', 'Staff': '7.3', 'Facilities': '7.4', 'Cleanliness': '6.8', 'Comfort': '7.1', 'Value for money': '7.5', 'Location': '6.8', 'Free WiFi': '5.0'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 8.1 7.3 7.7 7.6  
## 2 9.2 9.2 9.1 9.4  
## 3 7.3 7.4 6.8 7.1  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 7.4 6.9 5.8  
## 2 9.1 6.7 10.0  
## 3 7.5 6.8 5.0  
## hotel\_description  
## 1 ["You're eligible for a Genius discount at Park Guell! To save at this property, all you have to do is sign in.", 'In the Horta-Guinard\357\277\275 district of Barcelona, close to Park G\357\277\275ell, Park Guell has a terrace, free WiFi and a washing machine. The property features mountain views and is 4.3 km from Sagrada Familia and 4.3 km from La Pedrera.', 'The apartment is equipped with 3 bedrooms, 1 bathroom, bed linen, towels, a flat-screen TV, a dining area, a fully equipped kitchen, and a balcony with city views.', 'Casa Batllo is 4.8 km from the apartment, while Passeig de Gracia is 5.1 km from the property. The nearest airport is Barcelona-El Prat Airport, 19.3 km from Park Guell.']  
## 2 ['Located in Valencia, a 6-minute walk from Malvarrosa Beach and half a kilometer from Playa de las Arenas, PISO DE DISE\357\277\275O A 5 MIN ANDANDO DE LA PLAYA offers free WiFi and air conditioning. Housed in a building dating from 1978, this apartment is 4.8 km from Jardines de Monforte and 5 km from Oceanografic.', 'The apartment has 2 bedrooms, 1 bathroom, bed linen, towels, a flat-screen TV with satellite channels, a dining area, a fully equipped kitchen, and a patio with city views. For added convenience, the property can provide towels and bed linen for an extra charge.', 'Patacona Beach is a 13-minute walk from the apartment, while Port of Valencia is 2.4 km away. The nearest airport is Valencia Airport, 16.1 km from PISO DE DISE\357\277\275O A 5 MIN ANDANDO DE LA PLAYA.']  
## 3 ['In the Algiros district of Valencia, close to Turia Gardens, Benlloch Silence features free WiFi and a washing machine. It is located 1.8 km from Jardines de Monforte and provides a tour desk.', "The air-conditioned apartment consists of 2 separate bedrooms, 2 bathrooms with free toiletries, a seating area area, and a living room. There's a dining area and a kitchen complete with a microwave, a fridge and a stovetop.", 'Gonz\357\277\275lez Mart\357\277\275 National Museum of Ceramics and Decorative Arts is 3.1 km from the apartment, while Valencia North Station is 3.1 km away. The nearest airport is Valencia Airport, 13 km from Benlloch Silence.']  
## features  
## 1 ['Whole apartment', 'Size', 'Kitchen', 'City view', 'Washing machine', 'Free WiFi', 'Terrace', 'Balcony', 'Private Bathroom', 'Non-smoking rooms']  
## 2 ['Whole apartment', 'Size', 'Kitchen', 'City view', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Heating']  
## 3 ['Whole apartment', 'Size', 'Kitchen', 'Washing machine', 'Free WiFi', 'Air conditioning', 'Private Bathroom', 'Non-smoking rooms']  
## balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 1 0 0 1 1 0  
## 2 0 0 0 1 1 0  
## 3 0 0 0 1 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 1 1 0 0 0 0  
## 2 0 1 0 1 0 1  
## 3 1 1 0 0 0 1  
## free\_wifi  
## 1 1  
## 2 1  
## 3 1  
## room\_data  
## 1 {'143465501\_337923975\_0\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 459', 'room\_capacity': 'Max people: 5', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '65 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Mountain view', 'City view', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 1 twin bed/ 1 full bed/ '}, '143465501\_337923975\_4\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 408', 'room\_capacity': 'Max people: 4', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '65 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Mountain view', 'City view', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 1 twin bed/ 1 full bed/ '}, '143465501\_337923975\_3\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 356', 'room\_capacity': 'Max people: 3', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '65 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Mountain view', 'City view', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 1 twin bed/ 1 full bed/ '}, '143465501\_337923975\_2\_0\_0': {'room\_name': 'Three-Bedroom Apartment', 'room\_price': '\357\277\275 303', 'room\_capacity': 'Max people: 2', 'room\_options': ['Free cancellation until 11:59 PM on November 27, 2022', 'discount available'], 'room\_facilities': ['Entire apartment', '65 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Balcony', 'Mountain view', 'City view', 'Dishwasher', 'Flat-screen TV', 'Terrace', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '2 twin beds/ 1 twin bed/ 1 full bed/ '}}  
## 2 {'554626501\_339203290\_6\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 524', 'room\_capacity': 'Max people: 6', 'room\_options': ['Free cancellation until 11:59 PM on March 6, 2023', 'Pay nothing until March 4, 2023'], 'room\_facilities': ['Entire apartment', '65 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Landmark view', 'City view', 'Inner courtyard view', 'Air conditioning', 'Patio', 'Dishwasher', 'Flat-screen TV'], 'room\_bed\_type': '1 full bed/ 1 bunk bed/ 1 sofa bed/ '}}  
## 3 {'794422401\_364182307\_3\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 201', 'room\_capacity': 'Max adults: 3 <br> Max children: 2', 'room\_options': ['Free cancellation until 11:59 PM on December 10, 2022', 'Pay nothing until December 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}, '794422401\_364182307\_4\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 217', 'room\_capacity': 'Max adults: 4 <br> Max children: 1', 'room\_options': ['Free cancellation until 11:59 PM on December 10, 2022', 'Pay nothing until December 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}, '794422401\_364182307\_5\_0\_0': {'room\_name': 'Two-Bedroom Apartment', 'room\_price': '\357\277\275 224', 'room\_capacity': 'Max adults: 5', 'room\_options': ['Free cancellation until 11:59 PM on December 10, 2022', 'Pay nothing until December 8, 2022'], 'room\_facilities': ['Entire apartment', '70 m\357\277\275', 'Private kitchen', 'Private Bathroom', 'Inner courtyard view', 'Air conditioning', 'Flat-screen TV', 'Coffee machine', 'Free WiFi'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ 1 sofa bed/ 2/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 127 5 22 2022-11-19 0 1  
## 2 5 1 5 2022-11-18 0 1  
## 3 13 1 13 2022-11-18 1 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 131 303 459 0  
## 2 0 145 524 524 0  
## 3 0 112 201 224 0  
## is\_apartment has\_free\_cancelation  
## 1 1 1  
## 2 1 1  
## 3 1 1

ggplot(booking, aes(y = free\_wifi\_score, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$free\_wifi\_score)$out

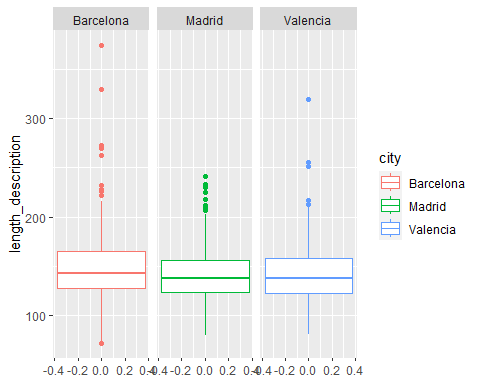
## [1] 5.9 5.6 6.6 5.9 5.6 6.4 6.3 6.3 6.6 5.0 6.6 6.5 5.0 5.0 6.5 6.3 5.0 6.5  
## [19] 6.3 2.5 6.7 5.0 6.3 6.6 4.8 5.9 5.6 6.6 5.9 5.6 6.6 4.8 5.0 2.5 6.4 5.9  
## [37] 5.6 6.6 4.8 6.0 5.6 6.4 5.0 6.7 5.0 5.0 5.7 6.4 6.5 5.0 5.0 5.0 5.6 6.4  
## [55] 6.4 6.6 5.0 5.9 6.3 5.6 6.7 5.8 6.6 5.0 6.3 5.0 6.3 5.0 5.9 5.6 6.6 4.8  
## [73] 5.6 5.6 5.9 5.6 4.8 5.9 5.9 6.3 5.9 6.6 6.4 5.9 5.6 5.6 5.8 6.7 5.9 4.8  
## [91] 5.9 5.0 6.6 5.6 5.0 6.3 5.0 5.0 6.5 6.5 5.0 6.5 5.0 5.7 6.5 5.0 6.6 6.6  
## [109] 6.3 5.0 6.6

booking %>%  
 filter(free\_wifi\_score < 5.0 & free\_wifi\_score > 0) %>%  
 sample\_n(3)

## name city check.in month adults children check.out  
## 1 Hotel Peninsular Barcelona 12-March-2023 March 2 3 16-March-2023  
## 2 Hotel Peninsular Barcelona 12-March-2023 March 2 3 16-March-2023  
## 3 Hotel Peninsular Barcelona 12-June-2023 June 5 3 16-June-2023  
## num\_rooms address  
## 1 2 Carrer de Sant Pau, 34, Ciutat Vella, 08001 Barcelona, Spain  
## 2 2 Carrer de Sant Pau, 34, Ciutat Vella, 08001 Barcelona, Spain  
## 3 3 Carrer de Sant Pau, 34, Ciutat Vella, 08001 Barcelona, Spain  
## postal\_code hotel\_coordinates latitude longitude hotel\_score  
## 1 08001 41.37962423,2.17249945 41.37962 2.172499 7.1  
## 2 08001 41.37962423,2.17249945 41.37962 2.172499 7.1  
## 3 08001 41.37962423,2.17249945 41.37962 2.172499 7.1  
## hotel\_scores  
## 1 {'': '', 'Staff': '8.6', 'Facilities': '6.7', 'Cleanliness': '7.5', 'Comfort': '7.1', 'Value for money': '7.7', 'Location': '9.2', 'Free WiFi': '4.8'}  
## 2 {'': '', 'Staff': '8.6', 'Facilities': '6.7', 'Cleanliness': '7.5', 'Comfort': '7.1', 'Value for money': '7.7', 'Location': '9.1', 'Free WiFi': '4.8'}  
## 3 {'': '', 'Staff': '8.6', 'Facilities': '6.7', 'Cleanliness': '7.5', 'Comfort': '7.1', 'Value for money': '7.7', 'Location': '9.1', 'Free WiFi': '4.8'}  
## staff\_score facilities\_score cleanliness\_score comfort\_score  
## 1 8.6 6.7 7.5 7.1  
## 2 8.6 6.7 7.5 7.1  
## 3 8.6 6.7 7.5 7.1  
## value\_for\_money\_score location\_score free\_wifi\_score  
## 1 7.7 9.2 4.8  
## 2 7.7 9.1 4.8  
## 3 7.7 9.1 4.8  
## hotel\_description  
## 1 ["You're eligible for a Genius discount at Hotel Peninsular! To save at this property, all you have to do is sign in.", 'Just a one minute walk from Las Ramblas, this charming hotel features a free Wi-Fi zone and a bright central courtyard filled with plants. Liceu Metro Station is only 202 m away.', 'Hotel Peninsular features functional, air-conditioned rooms with private bathrooms. The hotel\357\277\275s 24-hour front desk offers luggage storage and information about Barcelona.', 'The Peninsular is on the edge of the trendy Raval district, which is full of lively bars. The colorful street performers and sidewalk caf\357\277\275s of Las Ramblas are just around the corner.', 'The famous Boqueria Market is 247 m from the hotel, while Barcelona Cathedral is 0.5 km away. Airport buses and local trains stop in Pla\357\277\275a Catalunya, a 10-minute walk away.']  
## 2 ["You're eligible for a Genius discount at Hotel Peninsular! To save at this property, all you have to do is sign in.", 'Just a one minute walk from Las Ramblas, this charming hotel features a free Wi-Fi zone and a bright central courtyard filled with plants. Liceu Metro Station is only 202 m away.', 'Hotel Peninsular features functional, air-conditioned rooms with private bathrooms. The hotel\357\277\275s 24-hour front desk offers luggage storage and information about Barcelona.', 'The Peninsular is on the edge of the trendy Raval district, which is full of lively bars. The colorful street performers and sidewalk caf\357\277\275s of Las Ramblas are just around the corner.', 'The famous Boqueria Market is 247 m from the hotel, while Barcelona Cathedral is 0.5 km away. Airport buses and local trains stop in Pla\357\277\275a Catalunya, a 10-minute walk away.']  
## 3 ["You're eligible for a Genius discount at Hotel Peninsular! To save at this property, all you have to do is sign in.", 'Just a one minute walk from Las Ramblas, this charming hotel features a free Wi-Fi zone and a bright central courtyard filled with plants. Liceu Metro Station is only 202 m away.', 'Hotel Peninsular features functional, air-conditioned rooms with private bathrooms. The hotel\357\277\275s 24-hour front desk offers luggage storage and information about Barcelona.', 'The Peninsular is on the edge of the trendy Raval district, which is full of lively bars. The colorful street performers and sidewalk caf\357\277\275s of Las Ramblas are just around the corner.', 'The famous Boqueria Market is 247 m from the hotel, while Barcelona Cathedral is 0.5 km away. Airport buses and local trains stop in Pla\357\277\275a Catalunya, a 10-minute walk away.']  
## features balcony swimming\_pool pet\_friendly kitchen city\_view aparments  
## 1 [] 0 0 0 0 0 0  
## 2 [] 0 0 0 0 0 0  
## 3 [] 0 0 0 0 0 0  
## non\_smoking\_rooms private\_bathroom elevator heating safe air\_conditioning  
## 1 0 0 0 0 0 0  
## 2 0 0 0 0 0 0  
## 3 0 0 0 0 0 0  
## free\_wifi  
## 1 0  
## 2 0  
## 3 0  
## room\_data  
## 1 {'38860904\_84885129\_0\_1\_0': {'room\_name': 'Quadruple Room', 'room\_price': '\357\277\275 568', 'room\_capacity': 'Max adults: 4', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 4 rooms left on our site'], 'room\_facilities': ['15 m\357\277\275', 'Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '/multiple choice/'}, '38860901\_84885129\_0\_1\_0': {'room\_name': 'Single Room', 'room\_price': '\357\277\275 240', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 2 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '4 twin beds/ '}, '38860902\_84885129\_0\_1\_0': {'room\_name': 'Double or Twin Room', 'room\_price': '\357\277\275 324', 'room\_capacity': 'Max adults: 2', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'discount available', 'Only 3 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '1 twin bed/ '}, '38860903\_84885129\_1\_1\_0': {'room\_name': 'Triple Room', 'room\_price': '\357\277\275 432', 'room\_capacity': 'Max adults: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 6 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '2 twin beds/ 1 full bed/ /multiple choice/'}}  
## 2 {'38860901\_84885129\_0\_1\_0': {'room\_name': 'Single Room', 'room\_price': '\357\277\275 240', 'room\_capacity': 'Max people: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 2 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '/multiple choice/'}, '38860902\_84885129\_0\_1\_0': {'room\_name': 'Double or Twin Room', 'room\_price': '\357\277\275 324', 'room\_capacity': 'Max people: 2', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'discount available', 'Only 3 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '1 twin bed/ '}, '38860904\_84885129\_0\_1\_0': {'room\_name': 'Quadruple Room', 'room\_price': '\357\277\275 544', 'room\_capacity': 'Max people: 4', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 4 rooms left on our site'], 'room\_facilities': ['15 m\357\277\275', 'Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ /multiple choice/'}, '38860903\_84885129\_1\_1\_0': {'room\_name': 'Triple Room', 'room\_price': '\357\277\275 432', 'room\_capacity': 'Max people: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on March 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 6 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '4 twin beds/ '}}  
## 3 {'38860902\_84885129\_0\_1\_0': {'room\_name': 'Double or Twin Room', 'room\_price': '\357\277\275 464', 'room\_capacity': 'Max people: 2', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on June 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'discount available', 'Only 7 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '/multiple choice/'}, '38860901\_84885129\_0\_1\_0': {'room\_name': 'Single Room', 'room\_price': '\357\277\275 284', 'room\_capacity': 'Max people: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on June 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 3 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '1 full bed/ 2 twin beds/ /multiple choice/'}, '38860904\_84885129\_0\_1\_0': {'room\_name': 'Quadruple Room', 'room\_price': '\357\277\275 688', 'room\_capacity': 'Max people: 4', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on June 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 4 rooms left on our site'], 'room\_facilities': ['15 m\357\277\275', 'Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '1 twin bed/ '}, '38860903\_84885129\_1\_1\_0': {'room\_name': 'Triple Room', 'room\_price': '\357\277\275 532', 'room\_capacity': 'Max people: 1', 'room\_options': ['Continental breakfast included', 'Free cancellation until 11:59 PM on June 8, 2023', 'NO PREPAYMENT NEEDED \357\277\275 pay at the property', 'Only 6 rooms left on our site'], 'room\_facilities': ['Air conditioning', 'Private Bathroom'], 'room\_bed\_type': '4 twin beds/ '}}  
## page\_count current\_page in\_page\_count search\_date is\_december is\_march  
## 1 45 2 20 2022-11-19 0 1  
## 2 6 1 6 2022-11-18 0 1  
## 3 16 1 16 2022-11-18 0 0  
## as.integer(!is.na(NA)) length\_description min\_price max\_price is\_suite  
## 1 0 148 240 568 0  
## 2 0 148 240 544 0  
## 3 0 148 284 688 0  
## is\_apartment has\_free\_cancelation  
## 1 0 1  
## 2 0 1  
## 3 0 1

A continuación se evalúa la longitud de las descripciones.

ggplot(booking, aes(y = length\_description, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



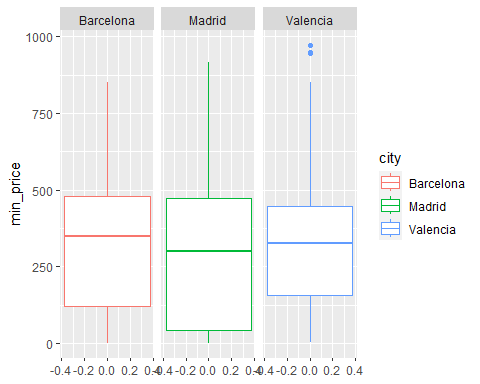
boxplot.stats(booking$length\_description)$out

## [1] 226 232 215 270 226 232 270 374 215 226 232 270 330 215 273 228 222 217 213  
## [20] 320 251 256 222 226 232 72 216 215 226 218 212 232 225 263 222 233 230 212  
## [39] 232 270 330 222 263 73 212 212 218 232 212 241 217 215 226 232 232 228 215  
## [58] 330 216 222 226 72 270 232 222 222 222 222 270 222 226 228 330 263 73 251  
## [77] 213 320 212 218 230 241 212

Observamos que las descripciones tienen una longitud muy similar en las 3 ciudades en torno a las 120-170 palabras, presentando algún outlier por la parte superior, pero siempre con descripciones inferiores a las 400 palabras. Aunque corresponda a una descripción un poco extensa, no parece que corresponda a un outlier sino simplemente a una descripción más extensa de lo habitual.

A continuación evaluaremos el precio mínimo de las habitaciones del hotel.

ggplot(booking, aes(y = min\_price, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



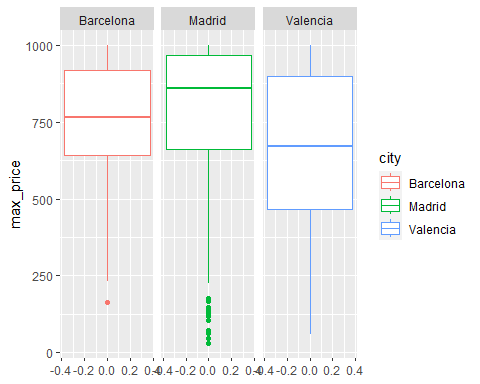
boxplot.stats(booking$min\_price)$out

## numeric(0)

Observamos que el precio mínimo tiene una gran disparidad oscilando de promedio entre los 100€ y los 500€, siendo Valencia la ciudad con el menor rango intercuartil con presencia de outliers en el rango superior de precios, con alguna habitación con precios cercanos a los 1000€, que aunque puedan parecer caros, son razonables en algún hotel de lujo de la ciudad.

A continuación se evalúan los precios máximos.

ggplot(booking, aes(y = max\_price, color = city))+  
 geom\_boxplot() +  
 facet\_grid(~city)



boxplot.stats(booking$max\_price)$out

## [1] 30 47 30 47

Observamos unos precios promedio mucho más elevados con un mayor rango intercuartil en la ciudad de Valencia, lo que sorprende al ser lo contrario que con los precios mínimos. También se observa que los outliers en este caso aparecen por el lado inferior especialmente en la ciudad de Madrid, aunque con precios positivos, que podrían corresponder a hoteles muy económicos o albergues.

Por otro lado sorprende ver que los precios máximos no superan el valor máximo observado en el precio mínimo, lo que se puede atribuir a que aquellos hoteles con precios muy elevados no ofrecían una variedad de habitaciones y precios.

#### 2.4. Exportación de los datos postprocesados

Una vez se han procesado los datos, se exportaran a un fichero CSV.

write.csv(booking, "hotel\_data\_processed.csv")

## 3. INTEGRACIÓN Y SELECCIÓN

La sección de integración y selección la encontramos tras la de limpieza, pues muchas de las columnas que queriamos descartar han sido tratadas en la fase de limpieza para extraer información de utilidad. Hecho esto, en la presente sección procedemos a eliminar todas las columnas del dataset origial que ya no vamos a necesitar para los siguientes apartados.

# Seleccionamos con la función "select" de dplyr todas las columnas excepto las que ya hemos limpiado para crear otras, además de aquellasque nos van a ser de poca utilidad.  
  
booking <- booking %>%  
 dplyr::select(-c("name", "search\_date", "hotel\_coordinates", "hotel\_scores", "check.in", "check.out", "address", "features", "hotel\_description", "room\_data"))

En primer lugar, hemos obtado por añadir datos de vuelos por varias razones:

* Un mayor número de vuelos en un mes podría aumentar las búsquedas de hoteles cerca del aeropuerto o los precios de dichos alojamientos.
* Un mayor número de vuelos en un mes indica más turismo internacional (en lugar de sólo turismo nacional). Un tipo de turismo que podría estar buscando hoteles con características diferentes a los turistas nacionales.
* Un mayor número de vuelos en un mes indica más turismo general, lo que puede influir en las estrategias de promoción y ofertas de determinados hoteles, influyendo en última instancia en su posición en el buscador Booking y, a su vez, directamente en los precios de los hoteles.

# Leemos los datos de vuelos  
flights <- read.csv("avia\_tf\_apal\_linear.csv.gz")  
summary(flights)

## DATAFLOW LAST.UPDATE freq unit   
## Length:1502955 Length:1502955 Length:1502955 Length:1502955   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## tra\_meas airline rep\_airp TIME\_PERIOD   
## Length:1502955 Length:1502955 Length:1502955 Length:1502955   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## OBS\_VALUE OBS\_FLAG   
## Min. : 0 Mode:logical   
## 1st Qu.: 0 NA's:1502955   
## Median : 264   
## Mean : 75333   
## 3rd Qu.: 3166   
## Max. :289204357

Para añadir dicha información hacemos un tratamiento de los datos obtenidos en Eurostat a fin de obtener una media de vuelos por mes para los últimos 5 años. Una vez obtenidos, tan solo debemos hacer un merge con nuestro dataset original.

proc\_flights <- flights %>%   
 # Filtramos para quedarnos solo con los aeropuertos de interés  
 # Código OACI/ICAO: Barcelona --> ES\_LEBL; Valencia --> ES\_LEVC; Madrid: ES\_LEMD  
 filter(rep\_airp %in% c("ES\_LEBL", "ES\_LEVC", "ES\_LEMD")) %>%   
 # Filtramos para quedarnos solo con los datos de carga de pasajeros  
 filter(tra\_meas == "PAS\_CRD") %>%  
 # Nos interesa solo los datos mensuales, así que filtramos por ellos  
 filter(freq == "M") %>%  
 separate(TIME\_PERIOD, c("YEAR", "MONTH"), sep = "-") %>%  
 # Nos quedamos solo con los últimos 5 años y los meses de Marzo, Junio y Diciembre  
 filter(YEAR %in% c("2022", "2021", "2020", "2019", "2018") & MONTH %in% c("03", "06", "12")) %>%  
 # No nos interesa los datos por aerolínea, solo los generales  
 filter(airline == "TOTAL") %>%  
 # Modificamos la columna de MONTH para tener los nombres del mes y cambiamos el código del aeropuerto por el nombre de la ciudad  
 mutate(month\_name = ifelse(MONTH == "03", "March", ifelse(MONTH == "06", "June", "December")),  
 city\_airp = ifelse(rep\_airp == "ES\_LEBL", "Barcelona", ifelse(rep\_airp == "ES\_LEVC", "Valencia", "Madrid"))) %>%  
 # Agrupamos por ciudad y mes y calculamos la media de vuelos.  
 group\_by(rep\_airp, month\_name) %>%  
 mutate(mean\_flights = mean(OBS\_VALUE)) %>%  
 ungroup() %>%  
 dplyr::select(c(city\_airp, month\_name, mean\_flights))  
  
# Nos quedamos con los registros únicos  
final\_flights <- unique(proc\_flights)  
  
# Hacemos un merge con bookings  
booking <- booking %>%  
 merge(final\_flights, by.x = c("city", "month"), by.y = c("city\_airp", "month\_name"))

## 4. ANÁLISIS DE LOS DATOS

A continuación, se evaluará la influencia de las distintas variables sobre la posición de un hotel en el buscador y sobre el precio que este fija para sus habitaciones. Para ello se empleará una matriz de correlaciones (de Spearman, dado que no asumimos normalidad). Una vez obtenidas las variables con mayor correlación con la variable objetivo generaremos una selección de los atributos a analizar, un grupo de variables distinto por cada problema a tratar.

Tras generar los distintos grupos procederemos a evaluar la normalidad y la homogeneidad de la varianza para así poder determinar en el siguiente punto que pruebas estadísticas se pueden aplicar.

Finalmente aplicaremos distintos tests estadísticos adicionales a la correlación hallada inicialmente. Estas pruebas estadísticas incluirán contrastes de hipótesis y regresiones lineales (enfocados dese un punto de vista inferencial, no predictivo).

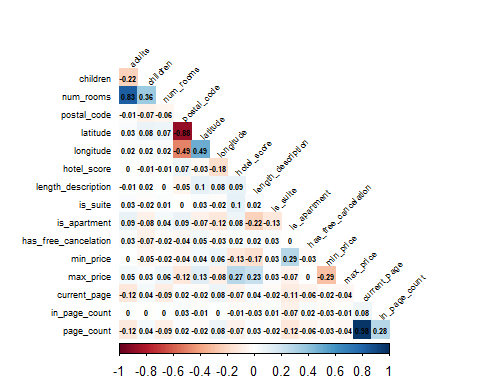
#### 4.1 Correlación con la variable objetivo y selección de los grupos de datos que se quieren analizar

Comenzamos evaluando la correlación (de spearman, dado que no podemos asumir normalidad y utilizamos variables categóricas también) entre las distintas variables con nuestra variable objetivo, para ello creamos una matriz de correlación y evaluamos los resultados obtenidos.

# Transformamos las variables a numérico  
booking$postal\_code <- as.numeric(booking$postal\_code)  
booking$latitude <- as.numeric(booking$latitude)  
booking$longitude <- as.numeric(booking$longitude)  
booking$free\_wifi <- as.numeric(booking$free\_wifi)  
booking$is\_apartment <- as.numeric(booking$is\_apartment)  
booking$pet\_friendly <- as.numeric(booking$pet\_friendly)  
booking$page\_count <- booking$page\_count + 1  
  
# Seleccionamos solo las variables numéricas  
booking\_numeric <- booking %>%  
 dplyr::select(-c("city", "month", "as.integer(!is.na(NA))"))  
  
mattmp <- booking\_numeric %>%  
 dplyr::select(c("adults", "children", "num\_rooms", "postal\_code", "latitude", "longitude", "hotel\_score", "length\_description",   
 "is\_suite", "is\_apartment", "has\_free\_cancelation", "min\_price","max\_price", "current\_page", "in\_page\_count", "page\_count"))  
  
sapply(mattmp, function(y) sum(is.na(y)))

## adults children num\_rooms   
## 0 0 0   
## postal\_code latitude longitude   
## 0 0 0   
## hotel\_score length\_description is\_suite   
## 0 0 0   
## is\_apartment has\_free\_cancelation min\_price   
## 0 0 0   
## max\_price current\_page in\_page\_count   
## 0 0 0   
## page\_count   
## 0

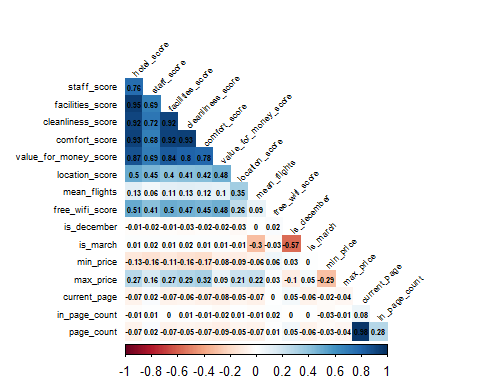
# Creamos una matriz de correlación del dataset  
cor\_mat <- cor(mattmp, method = "spearman")  
plot.new()  
plot.window(xlim=c(-2,2), ylim=c(5,10))  
corrplot(cor\_mat, method="color", type = "lower", tl.col = "black", tl.srt = 45, diag = FALSE, tl.cex = 0.55, addCoef.col = 'black', number.cex = 0.5)



mattmp <- booking\_numeric %>%  
 dplyr::select(c("hotel\_score", "staff\_score", "facilities\_score", "cleanliness\_score", "comfort\_score", "value\_for\_money\_score", "location\_score",  
 "mean\_flights", "free\_wifi\_score", "is\_december", "is\_march", "min\_price", "max\_price", "current\_page", "in\_page\_count", "page\_count"))  
  
sapply(mattmp, function(y) sum(is.na(y)))

## hotel\_score staff\_score facilities\_score   
## 0 0 0   
## cleanliness\_score comfort\_score value\_for\_money\_score   
## 0 0 0   
## location\_score mean\_flights free\_wifi\_score   
## 0 0 0   
## is\_december is\_march min\_price   
## 0 0 0   
## max\_price current\_page in\_page\_count   
## 0 0 0   
## page\_count   
## 0

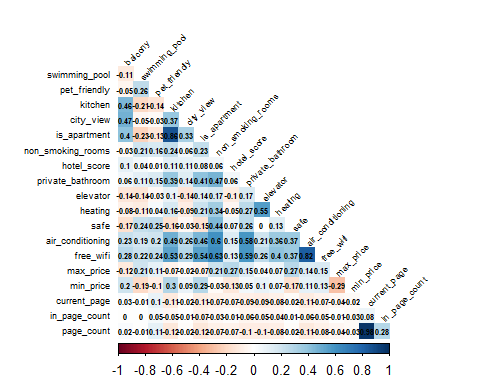
# Creamos una matriz de correlación del dataset  
cor\_mat <- cor(mattmp, method = "spearman")  
plot.new()  
plot.window(xlim=c(-2,2), ylim=c(5,10))  
corrplot(cor\_mat, method="color", type = "lower", tl.col = "black", tl.srt = 45, diag = FALSE, tl.cex = 0.55, addCoef.col = 'black', number.cex = 0.5)



mattmp <- booking\_numeric %>%  
 dplyr::select(c("balcony", "swimming\_pool", "pet\_friendly", "kitchen", "city\_view", "is\_apartment", "non\_smoking\_rooms", "hotel\_score",   
 "private\_bathroom", "elevator", "heating", "safe", "air\_conditioning", "free\_wifi", "max\_price", "min\_price", "current\_page", "in\_page\_count", "page\_count"))  
  
sapply(mattmp, function(y) sum(is.na(y)))

## balcony swimming\_pool pet\_friendly kitchen   
## 0 0 0 0   
## city\_view is\_apartment non\_smoking\_rooms hotel\_score   
## 0 0 0 0   
## private\_bathroom elevator heating safe   
## 0 0 0 0   
## air\_conditioning free\_wifi max\_price min\_price   
## 0 0 0 0   
## current\_page in\_page\_count page\_count   
## 0 0 0

# Creamos una matriz de correlación del dataset  
cor\_mat <- cor(mattmp, method = "spearman")  
plot.new()  
plot.window(xlim=c(-2,2), ylim=c(5,10))  
corrplot(cor\_mat, method="color", type = "lower", tl.col = "black", tl.srt = 45, diag = FALSE, tl.cex = 0.55, addCoef.col = 'black', number.cex = 0.5)



Como bien se puede observar en las matrices, las variables que recogen las distintas puntuaciones están correlacionadas entre ellas, por lo que un hotel que haya obtenido una puntuación elevada en un aspecto generalmente también lo obtendrá en el resto. Este factor es de especial interés si se seleccionan varias columnas de este tipo, pues corremos el riesgo de sufrir un problema de multicolinealidad. También encontramos que los hoteles con distintos atributos como aire acondicionado o baño privado suelen tener el resto de servicios adicionales, aunque este factor puede deberse a que es más habitual encontrar este tipo de elementos en una habitación de hotel, por lo que es común que cuando exista una habitación con calefacción, por ejemplo, también tenga aire acondicionado o baños privado. Por último, aunque es evidente, cabe destacar que las mayores colinealidades las tenemos con variables directamente relacionadas o que presentan información de la misma fuente (current\_page con page\_count o max\_price con min\_price). No obstante, esto no es un problema, pues eliminaremos estas variables del conjunto de entrenamiento para no incurrir en “data leakage”.

# Filtramos para obtener cuales son las variables más correlacionadas con nuestra variable objetivo, current\_page  
cor\_mat <- cor(booking\_numeric, method = "spearman")

## Warning in cor(booking\_numeric, method = "spearman"): the standard deviation is  
## zero

cor\_mat\_df <- as.data.frame(cor\_mat)  
cor\_mat\_df %>%  
 filter(abs(page\_count) > 0.065) %>%  
 filter(page\_count < 0.95) %>%  
 arrange(page\_count) %>%  
 dplyr::select(page\_count)

## page\_count  
## kitchen -0.12023093  
## is\_apartment -0.12015885  
## adults -0.11580745  
## air\_conditioning -0.11400432  
## elevator -0.09860877  
## private\_bathroom -0.09661382  
## num\_rooms -0.08662662  
## value\_for\_money\_score -0.08574257  
## heating -0.08399204  
## free\_wifi -0.07529422  
## non\_smoking\_rooms -0.07424905  
## mean\_flights -0.06952514  
## comfort\_score -0.06936356  
## hotel\_score -0.06765554  
## facilities\_score -0.06508687  
## longitude 0.07601044  
## pet\_friendly 0.10535168  
## in\_page\_count 0.28025901

Cuando evaluamos la posición en la página web observamos que no hay una correlación elevada con ninguna variable, encontrando valores por debajo de 0.12 (en valor absoluto) para todos los casos. No obstante, para el caso de page\_count, hemos optado por someter a estudio aquellas variables que presentan mayor correlación de spearman con la variable dependiente (pues evaluando con distintas métricas que buscan determinar el nivel de dependencia no-lineal hemos obtenido resultados similares). Las variables seleccionadas en cuestión son: kitchen, is\_apartment, adults, air\_conditioning, elevator, private\_bathroom, num\_rooms, value\_for\_money\_score, heating, free\_wifi, non\_smoking\_rooms, comfort\_score, hotel\_score, longitude, facilities\_score pet\_friendly, mientras que la variable dependiente será page\_count.

booking\_pcount <- booking\_numeric %>%  
 dplyr::select(c("kitchen", "is\_apartment", "adults", "air\_conditioning", "elevator", "private\_bathroom",  
 "num\_rooms", "value\_for\_money\_score", "heating", "non\_smoking\_rooms", "comfort\_score", "hotel\_score",  
 "longitude", "has\_free\_cancelation", "page\_count"))

En el caso del problema a resolver relacionado con el precio, al analizar las correlaciones sí que encontramos valores entre 0.15 y 0.20. Aunque estas correlaciones no sean muy elevadas nos ofrecen un escenario más prometedor que el anteriormente presentado. A continuación, procedemos a seguir con el mismo procedimiento que con page\_count, esta vez seleccionando las siguientes variables:

* **postal\_code**
* **balcony**
* **pet\_friendly**
* **elevator**
* **heating**
* **cleanliness\_score**
* **latitude**
* **mean\_flights**
* **location\_score**
* **hotel\_score**
* **facilities\_score**
* **non\_smoking\_rooms**
* **comfort\_score**
* **is\_apartment**
* **private\_bathroom**
* **kitchen**
* **air\_conditioning**

# Filtramos para obtener cuales son las variables más correlacionadas con nuestra variable objetivo, page\_count  
cor\_mat <- cor(booking\_numeric, method = "spearman")

## Warning in cor(booking\_numeric, method = "spearman"): the standard deviation is  
## zero

cor\_mat\_df <- as.data.frame(cor\_mat)  
cor\_mat\_df %>%  
 filter(abs(max\_price) > 0.1) %>%  
 filter(max\_price < 0.95) %>%  
 arrange(max\_price) %>%  
 dplyr::select(max\_price)

## max\_price  
## min\_price -0.2933215  
## postal\_code -0.1244939  
## balcony -0.1185999  
## pet\_friendly 0.1073362  
## latitude 0.1271236  
## air\_conditioning 0.1408840  
## private\_bathroom 0.1452074  
## free\_wifi 0.1523803  
## staff\_score 0.1640168  
## swimming\_pool 0.2055779  
## non\_smoking\_rooms 0.2110602  
## location\_score 0.2116932  
## mean\_flights 0.2224695  
## length\_description 0.2287604  
## facilities\_score 0.2667483  
## hotel\_score 0.2700575  
## safe 0.2719657  
## cleanliness\_score 0.2919506  
## comfort\_score 0.3206842

booking\_price <- booking\_numeric %>%  
 dplyr::select(c("postal\_code", "balcony", "pet\_friendly", "latitude", "air\_conditioning", "private\_bathroom", "free\_wifi", "staff\_score",  
 "swimming\_pool", "location\_score", "non\_smoking\_rooms", "mean\_flights", "length\_description", "facilities\_score", "hotel\_score",  
 "safe", "cleanliness\_score", "comfort\_score", "max\_price"))

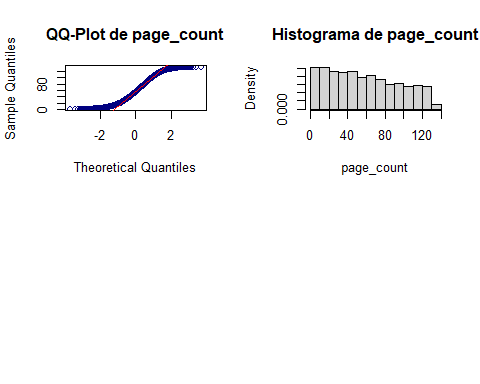
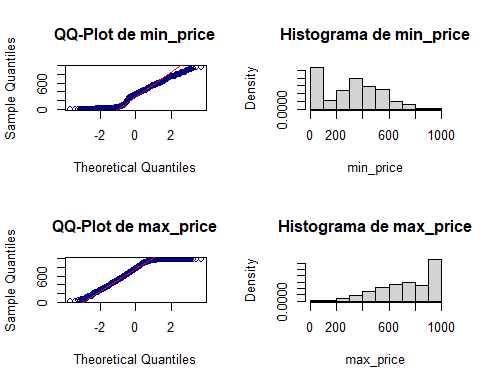
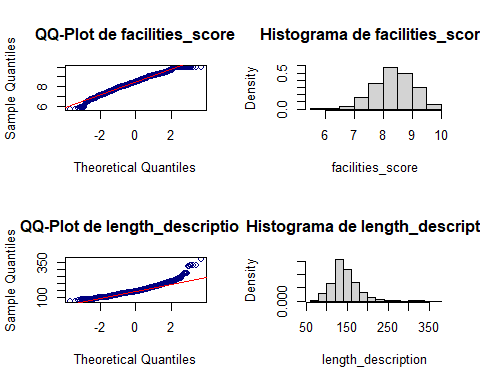
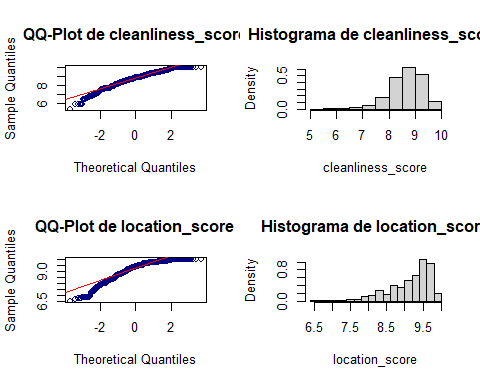
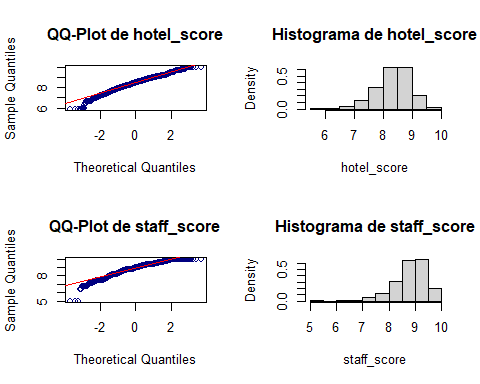
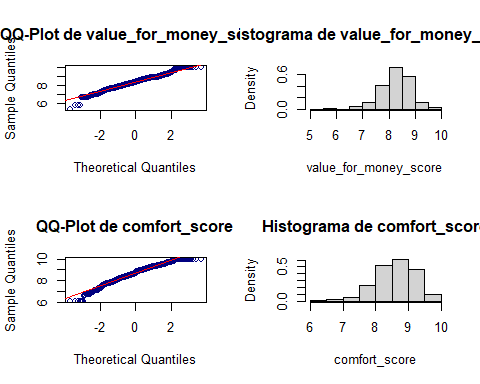
#### 4.2 Comprobación de la normalidad y homogeneidad de la varianza

Para comprobar si los valores de las variables de nuestro dataset se distribuyen o se aproximan a una población distribuida normalmente haremos un análisis visual, con QQ-Plots y Histogramas, y un análisis mediante el test Shapiro-Wilk. Concretamente solo comprobaremos la normalidad para las variables seleccionadas en el apartado anterior.

norm\_cols <- c("value\_for\_money\_score", "comfort\_score", "hotel\_score", "staff\_score", "cleanliness\_score", "location\_score", "facilities\_score", "length\_description", "min\_price", "max\_price", "page\_count")  
  
booking\_norm <- booking\_numeric %>%  
 dplyr::select(norm\_cols)

## Warning: Using an external vector in selections was deprecated in tidyselect 1.1.0.  
## i Please use `all\_of()` or `any\_of()` instead.  
## # Was:  
## data %>% select(norm\_cols)  
##   
## # Now:  
## data %>% select(all\_of(norm\_cols))  
##   
## See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.

par(mfrow=c(2,2))  
for(i in 1:ncol(booking\_norm)){  
 qqnorm(booking\_norm[,i], main = paste("QQ-Plot de",colnames(booking\_norm)[i]), col = "navy")  
 qqline(booking\_norm[,i], col = "red")  
 hist(booking\_norm[,i], main = paste("Histograma de", colnames(booking\_norm)[i]), xlab=colnames(booking\_norm)[i], freq = FALSE)  
}



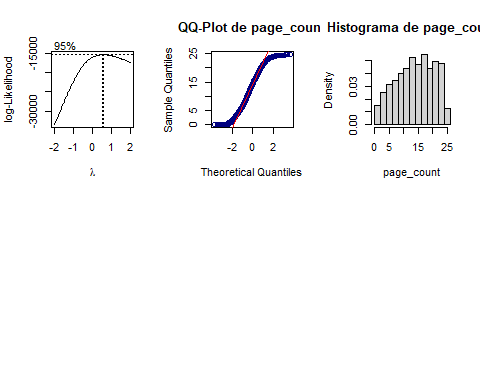
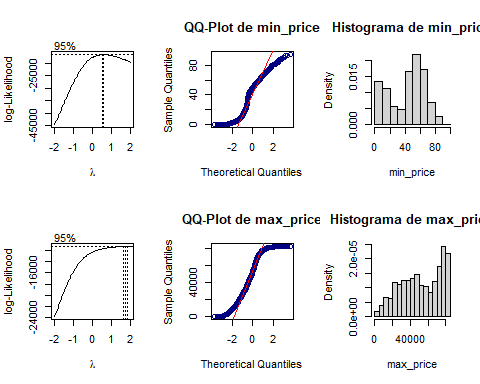
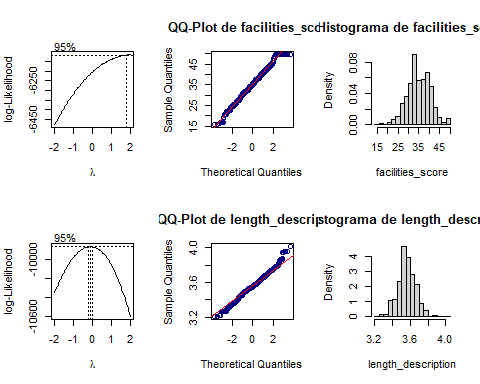
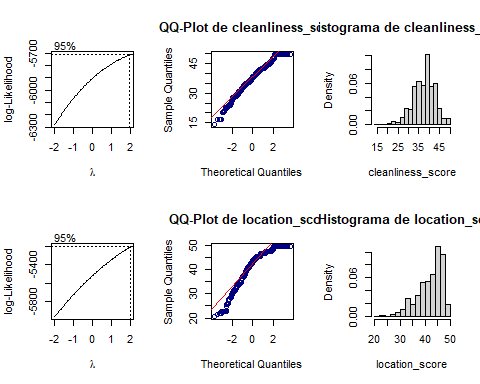
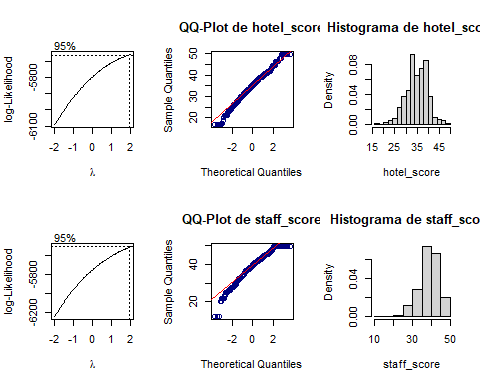
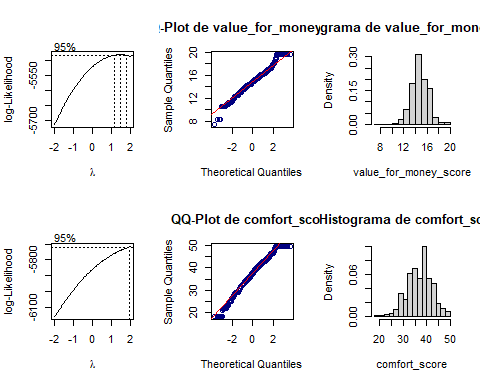
Los QQ-plots y los histogramas nos muestran una distribuciones con una evidente asimetría negativa en la mayoría de los casos (excepto variables como length\_description que muestran una asimetría más bien positiva, aunque podría ser influencia de un dato extremo que no ha sido considerado outlier), así como un evidente problema con las colas, pues en la mayoría de los casos los puntos se alejan de la línea roja, la cual corresponde a la hipotética distribución normal. Por tanto, a priori podríamos rechazar la hipótesis de normalidad de las variables. No obstante, tras ejecutar el test Shapiro-Wilk confirmamos de forma más fiable nuestras primeras conclusiones, pues todos los p-valores muestran valores muy inferiores a 0.05, rechazando en todos los casos la hipótesis nula sobre la normamlidad de los datos.

lapply(booking\_norm, shapiro.test)

## $value\_for\_money\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.98677, p-value < 2.2e-16  
##   
##   
## $comfort\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.98219, p-value < 2.2e-16  
##   
##   
## $hotel\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.9837, p-value < 2.2e-16  
##   
##   
## $staff\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.95898, p-value < 2.2e-16  
##   
##   
## $cleanliness\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.97328, p-value < 2.2e-16  
##   
##   
## $location\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.89384, p-value < 2.2e-16  
##   
##   
## $facilities\_score  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.99044, p-value = 1.539e-15  
##   
##   
## $length\_description  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.9554, p-value < 2.2e-16  
##   
##   
## $min\_price  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.94862, p-value < 2.2e-16  
##   
##   
## $max\_price  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.92609, p-value < 2.2e-16  
##   
##   
## $page\_count  
##   
## Shapiro-Wilk normality test  
##   
## data: X[[i]]  
## W = 0.95078, p-value < 2.2e-16

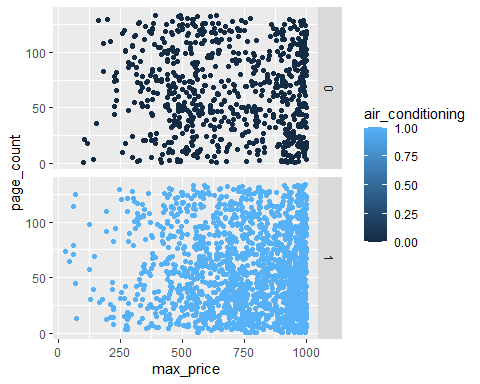
A fin de aproximar más los datos a una distribución normal, se propone hacer una transformación de box-cox y comprobar si los datos efectivamente se aproximan más a una normal. Tras llevar a cabo este proceso, si bien los histogramas muestran en algunos casos una distribución “visualmente más normales” y los QQ-plots parecen mostrar que la transformación ha aliviado ligeramente el problema con las colas, los resultados de la transformación no han sido suficientes como para poder considerar que los datos se distribuyen como una normal. No obstante, no debemos preocuparnos demasiado por este problema, pues tenemos suficientes registros como para poder apoyarnos en el Teorema del Límite Central. Por ejemplo, el t-test asume que las medias de las diferentes muestras se distribuyen normalmente (siendo una muestra de más de 30 registros suficiente, como rule of thumb, para que el t-test sea válido aunque los datos no se distribuyan normalmente).

par(mfrow=c(2,3))  
for(i in 1:ncol(booking\_norm)){  
 b <- boxcox(lm(booking\_norm[,i] ~ 1))  
 lambda <- b$x[which.max(b$y)]  
 qqnorm((booking\_norm[,i]^lambda - 1)/lambda, main = paste("QQ-Plot de",colnames(booking\_norm)[i]), col = "navy")  
 qqline((booking\_norm[,i]^lambda - 1)/lambda, col = "red")  
 hist((booking\_norm[,i]^lambda - 1)/lambda, main = paste("Histograma de", colnames(booking\_norm)[i]), xlab=colnames(booking\_norm)[i], freq = FALSE)  
}



Seguidamente, analizaremos la homocedasticidad mediante scatterplots y varios test de homocedasticidad, o de homogeneidad de varianzas. En este caso, no podemos utilizar el test de Levene ni el de Barlett ya que son muy sensibles ante datos no normales y, como acabamos de presentar, los datos analizados no se aproximan en ningún caso a una normal.

ggplot(booking, aes(x = max\_price, y = page\_count, color = air\_conditioning)) +   
 geom\_point() +  
 facet\_grid(air\_conditioning~.)



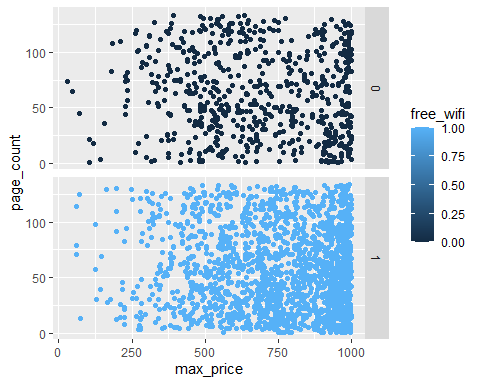
Como todos los test de homogeneidad de varianzas, nos basamos en las siguientes hipótesis:

Donde rechazaremos la hipótesis nula cuando el pvalor del test sea inferior a un nivel de significación igual a 0.05.

fligner.test(page\_count ~ air\_conditioning, data = booking)

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: page\_count by air\_conditioning  
## Fligner-Killeen:med chi-squared = 0.20066, df = 1, p-value = 0.6542

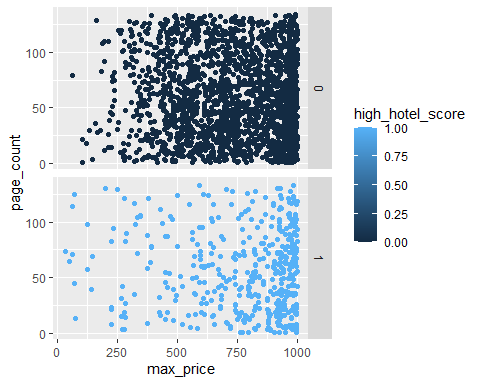
ggplot(booking, aes(x = max\_price, y = page\_count, color = free\_wifi)) +   
 geom\_point()+  
 facet\_grid(free\_wifi~.)



fligner.test(page\_count ~ free\_wifi, data = booking)

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: page\_count by free\_wifi  
## Fligner-Killeen:med chi-squared = 0.40063, df = 1, p-value = 0.5268

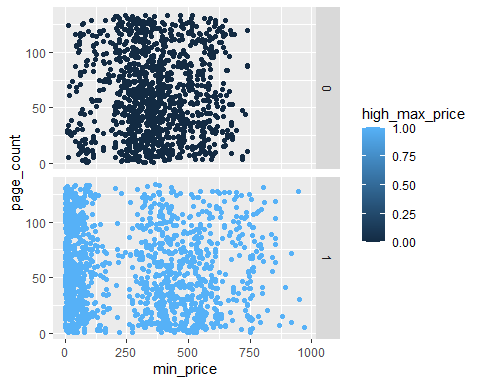
booking <- booking %>%  
 mutate(high\_hotel\_score = ifelse(hotel\_score >= 9, 1, 0))  
  
ggplot(booking, aes(x = max\_price, y = page\_count, color = high\_hotel\_score)) +   
 geom\_point()+  
 facet\_grid(high\_hotel\_score~.)



fligner.test(page\_count ~ high\_hotel\_score, data = booking)

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: page\_count by high\_hotel\_score  
## Fligner-Killeen:med chi-squared = 2.8946, df = 1, p-value = 0.08888

booking <- booking %>%  
 mutate(high\_max\_price = ifelse(max\_price >= mean(max\_price), 1, 0))  
  
ggplot(booking, aes(x = min\_price, y = page\_count, color = high\_max\_price)) +   
 geom\_point()+  
 facet\_grid(high\_max\_price~.)



fligner.test(page\_count ~ high\_max\_price, data = booking)

##   
## Fligner-Killeen test of homogeneity of variances  
##   
## data: page\_count by high\_max\_price  
## Fligner-Killeen:med chi-squared = 0.22846, df = 1, p-value = 0.6327

Tras realizar test a un gran número de variables de interés (principalmente las que van a ser sujeto de estudio para los tests estadísticos), se ha determinado que existe homogeneidad de varianzas, por lo que no nos encontraremos con problemas de heterogeneidad a la hora de realizar pruebas de hipótesis.

#### 4.3 Aplicación de pruebas estadísticas para comparar los grupos de datos

En este apartado se aplicarán distintas pruebas estadísticas para comparar los distintos grupos de datos de los que disponemos y ver cual es su relación con la variable objetivo. Esto nos permitirá evaluar que variables tienen un impacto significativo sobre la variable objetivo y cuales no. Nuestra variable objetivo es la posición en la página (page\_count) y relacionada directamente la página en la que aparece (current\_page), por lo que las distintas hipótesis que se plantean están asociadas con la influencia de distintas variables sobre el posicinamiento de un hotel en Booking. A continuación se recogen unas hipótesis, aunque sería posible evaluar otras, que también pareciesen interesantes, dado el carácter académico de este trabajo nos centraremos exclusivamente en estudiar la influencia de las variables:

* **free\_wifi**
* **has\_free\_cancelation**
* **hotel\_score**
* **min\_price**

Las hipótesis que trataremos de verificar serán las siguientes:

* Aquellos hoteles con wifi tendrán una mejor posición (page\_count inferior) a los que no dispongan de wifi
* Los hoteles con cancelación gratuita tendrán una posición diferente a los que no dispongan de esta.
* Los hoteles con una puntuación global superior a un 9 tendrán un mejor posicionamiento.
* Los hoteles con un precio mínimo inferior a la media tendrán un mejor posicionamiento.

A continuación se estudiará cada una de estas hipótesis y el resultado obtenido.

**Hipótesis 1** La primera hipótesis evalúa la influencia del wifi en el posicionamiento del hotel, suponiendo que un hotel con wifi gratuito tendrá un posicionamiento mejor, es decir menor **page\_count** que un hotel que no disponga de wifi gratuito. A continuación se plantea la hipótesis nula () y la alternativa ().

* La posición global de un hotel con wifi gratuito y uno sin serán iguales.
* La posición global de un hotel con wifi gratuito será mejor (inferior) a la de uno sin.

TEST NO PARAMETRICO AL NO PODER ASUMIR LA NORMALIDAD DE LOS DATOS

**Hipótesis 2** Esta segunda hipótesis trata de evaluar la influencia de la cancelación gratuita, la cual para aportar una mayor diversidad en los análisis realizados, puede estar asociada a una posición distinta (inferior o superior) que la de un hotel sin cancelación gratuita. Definimos esta hipótesis mediante las hipótesis nula () y alternativa () siguientes:

* La posición global de un hotel con cancelación gratuita y uno sin serán iguales.
* La posición global de un hotel con cancelación gratuita será distinta a la de uno sin.

TEST NO PARAMETRICO AL NO PODER ASUMIR LA NORMALIDAD DE LOS DATOS

**Hipótesis 3** A continuación evaluamos la influencia de la puntuación global del hotel. En este caso evaluaremos si un hotel con una puntuación global (**hotel\_score**) superior o igual a un 9.0 obtiene un mejor posicionamiento que un hotel con una peor puntuación. A continuación se plantea la hipótesis nula () y la alternativa ()

Los hoteles con una puntuación global superior a un 9 tendrán un mejor posicionamiento.

* La posición global de un hotel con una puntuación global mayor o igual a 9.0 y uno con una puntuación inferior serán iguales.
* La posición global de un hotel con una puntuación global mayor o igual a 9.0 será mejor (menor) que uno con una puntuación inferior.

TEST NO PARAMETRICO AL NO PODER ASUMIR LA NORMALIDAD DE LOS DATOS

**Hipótesis 4** Finalmente evaluamos la influencia del precio mínimo del hotel considerando que un hotel con un menor precio mínimo tendrá un mejor posicionamiento. Para ello consideraremos las siguientes hipótesis nula () y alternativa ()

Los hoteles con un precio mínimo inferior a la media tendrán un mejor posicionamiento.

* La posición global de un hotel con precio mínimo inferior a la media y uno superior serán iguales.
* La posición global de un hotel con precio mínimo inferior a la media será mejor (inferior) a la de uno con precio superior.

TEST NO PARAMETRICO AL NO PODER ASUMIR LA NORMALIDAD DE LOS DATOS

#### 4.4 Aplicación de funciones lineales para resolución del problema

Una vez analizados los distintos contrastes de hipótesis y hayada la correlación entre las distintas variables, tenemos un conocimiento del conjunto de datos lo suficientemente extenso como para analizar mediante una función lineal el posicionamiento de un hotel en el buscador de Booking.

## 5. REPRESENTACIÓN DE LOS RESULTADOS

A lo largo del documento se han presentado representaciones gráficas de los análisis realizados y de los resultados obtenidos.

## 6. RESOLUCIÓN DEL PROBLEMA